DATA-DRIVEN APPROACHES FOR SUSTAINABLE OPERATION AND DEFENSIBLE RESULTS IN A LONG-TERM, MULTI-SITE ECOSYSTEM FLUX MEASUREMENT PROGRAM

DATA-DRIVEN APPROACHES FOR SUSTAINABLE OPERATION AND DEFENSIBLE RESULTS IN A LONG-TERM, MULTI-SITE ECOSYSTEM FLUX MEASUREMENT PROGRAM

By

JASON JOSEPH BRODEUR, B.SC. (HONS.), M.SC.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University
Doctor of Philosophy (2014)
Hamilton, Ontario, Canada
(Geography and Earth Sciences)

TITLE:

Data-driven approaches for sustainable operation and defensible results in a long-term, multi-site ecosystem flux measurement program

AUTHOR: Jason Joseph Brodeur   B.Sc.(Hons)

                                                   (McMaster University, Hamilton, 2003)

                                                   M.Sc.

                                                   (University of Guelph, Guelph, 2005)

SUPERVISOR: Professor M. Altaf Arain

NUMBER OF PAGES   xvi, 232

ABSTRACT

Modern advances in biometeorological monitoring technology have improved the capacity for measuring ecosystem exchanges of mass, energy and scalars (such as CO_2). Translating these measurements into robust and accurate scientific information (and ultimately, understanding) requires careful assessment of operations throughout the biometeorological data life cycle. In response, this research analyzed and optimized aspects of data collection, management and filtering for an ecosystem exchange measurement program over an age-sequence of temperate white pine forests.

A comprehensive data workflow and management system (DWMS) was developed and implemented to support the entire data life cycle for all past, present and future measurement operations in our research group, and meet the needs of a collaborative, student-led data management environment. Best practices for biometeorological data management were introduced and used as standards to assess system performance.

Roving eddy covariance (rEC) systems were examined as a means of producing reliable time-integrated carbon exchange estimates at multiple sites, by rotating an EC system in a resource-mindful approach. When used with an optimal gap-filling model and rEC rotation schedule (2 sites with 15-day rotations), the results suggested its viability, as annual NEE estimate uncertainties ranged between 35 and 63% of the annual NEE flux magnitude at our study sites – even though approximately 70% of half-hours were filled.

Lastly, a data-driven approach was used to investigate the effects of different friction velocity and footprint filtering methods on time-integrated carbon exchange estimates at our fetch-limited forests. Though predicted flux source areas varied considerably between footprint models, our objective analyses identified the model (Kljun et al., 2004) and within-fetch requirement (80%) that optimized reliability and representativeness of carbon exchange estimates. Applying this footprint model decreased annual NEE by 31 to 129% (59 to 207 g C m^{-2} y^{-1}) relative to no footprint application, and highlighted the importance of objective analyses of EC flux filtering methods.

ACKNOWLEDGEMENTS

At the risk of extending this document by even an additional page, it is very important that I take a few paragraphs to thank the long list of people who have made this work possible, and/or have shown an incredible amount of patience with me over the years (*and years, and years*) it has taken me to complete this work.

First of all, I would like to thank my supervisor, Dr. Altaf Arain for being a supportive and steadying voice during my time as a PhD student. The amount of patience that you've shown me, and all other members of the research group – even when we've crashed your truck – is both exceptional and enviable. I would also like to thank the members of my supervisory committee: Dr. Mike Waddington, Dr. Harry McCaughey, and Dr. Greg Slater, for the encouragement, guidance, and repeated appearances at annual committee meetings.

Second, I would like to extend a heartfelt thanks to all members of the "Climate Lab", who have supported me scientifically, morally, emotionally, (and probably financially?), throughout my time here. Doing anything of importance is always more enjoyable and rewarding when it can be shared with colleagues and friends, and I consider you all to be the utmost of both. As such, I'd like to recognize you each individually and say that it's been my pleasure: Matthias Peichl, Samantha MacKay, Emily Nicholas, Myroslava Khomik, Suo Huang, Michelle Molon, Rob Thorne, Janelle Trant, Michelle Kula, and Ananta Parsaud.

I could not end this section without acknowledging the incredible patience and support that I've received from my family. To both the Brodeur and Van De Walle families, thank you for putting up with me locking myself in your offices, or doing work on a laptop during holiday get-togethers. To Lisa, Lucy and Cameron: I am eternally grateful to you all for the patience you've shown, and the sacrifices you've made to help me finish this. Thank you for being the daily inspiration for everything I do, and for giving me perspective and making me smile every day.

Finally, I'd like to dedicate this work to my Dad. Though it's a shame I was not able to share this accomplishment with him, I am happy realizing that he would be proud of this work, as he always encouraged me to do things well, and do them to the highest moral and ethical standards. Thanks, Dad.

List of Abbreviations and Symbols

Table 1*: List of Abbreviations and Symbols*

Abbreviation or Symbol	Description
TPFS	Turkey Point Flux Station
TP39	TPFS flux study site, planted in 1939
TP74	TPFS flux study site, planted in 1974
TP89	TPFS flux study site, planted in 1989
TP02	TPFS flux study site, planted in 2002
TP_PPT	TPFS precipitation measurement site
EC	Eddy covariance
rEC	Roving eddy covariance
CPEC	Closed-path eddy covariance
OPEC	Open-path eddy covariance
ACS	Automated Chamber System
BACON	Biometeorological Analysis, Collection, and Organizational Node
DWMS	Data workflow and management system
CFS	Central file system
CO_2	Carbon Dioxide
F_c	Turbulent carbon flux
ΔS_c	Air column CO_2 storage change flux
NEE	Net Ecosystem Exchange
GEP	Gross Ecosystem Productivity
RE	Ecosystem Respiration
T_a	Air temperature
T_s	Soil temperature
PAR	Photosynthetically active radiation
PPFD	Photosynthetic photon flux density
GDD	Growing degree days
VWC_{30}	30 cm depth-weighted volumetric water content
WS, (u)	Horizontal wind speed, (velocity)
u_*	Friction velocity
u_*^{Th}	Friction velocity threshold
σ_w	Standard deviation of vertical wind velocity
L	Obukhov length
z_m	Measurement height
z_0	Roughness length
d	Displacement height
ζ	Atmospheric stability parameter

Table 1*: List of Abbreviations and Symbols (continued)*

Abbreviation or Symbol	Description
F_{in}	Within-fetch integrated footprint function
fp_m	Footprint model
fp^{Th}	Footprint threshold
LES	Large eddy simulation
ϕ	Available data fraction
$\phi_{RE, GEP}$	RE- and GEP-parameterizable available data fraction
IQR	Inter-quartile range
r_Φ	Pearson's correlation coefficient
CI_{95}	95% confidence interval
MPT-G	Moving point test method described by Gu et al. (2005)
MPT-P	Moving point test method described by Papale et al. (2006)
CPD	Change points detection method described by Barr et al. (2013)
SP90	Footprint model of Schuepp et al. (1990)
HS00	Footprint model of Hsieh et al. (2000)
KM01	Footprint model of Kormann and Meixner (2001)
KL04	Footprint model of Kjun et al. (2004)
ANN	Artificial neural network
MDS	Marginal distribution sampling gap-filling model of Reichstein et al. (2005)
NLR-FC	Fluxnet-Canada non-linear regression gap-filling model
NLR-HL	Howland forest non-linear regression gap-filling model of Richardson et al. (2007)

Table of Contents

Index of Tables

Index of Figures

1   Introduction

Modern advances in environmental monitoring technologies have revolutionized the way in which researchers observe and understand natural phenomena (Benson et al., 2010; Campbell et al., 2013; Porter et al., 2009). In the field of biometeorology, these technologies have improved both the temporal and spatial resolution at which ecosystems may be monitored, providing estimates of the distribution and exchanges of mass, energy and scalars (such as CO_2) (Baldocchi et al., 2001; Hamilton et al., 2007; Kao et al., 2012).   Among these advances is the development and wide implementation of the eddy covariance (EC) method to study the terrestrial carbon cycle (Baldocchi, 2003). Together with micrometeorological towers and chamber measurements, EC flux installations have served a wide range of interests in ecological studies, including: the quantification of global carbon sources and sinks (Coursolle et al., 2012b; Janssens et al., 2001; Valentini et al., 2000); investigations of environmental and biological controls on terrestrial carbon exchanges (Barr et al., 2007; Falge et al., 2002; Richardson et al., 2010, 2007); monitoring of ecosystem growth and health (McLaren et al., 2008; Richardson et al., 2013); evaluation and information on management practices (Olajuyigbe et al., 2012; Saunders et al., 2012; Son et al., 2004); and, parameterization and validation of simple empirical to complex process-based models of ecosystem exchange dynamics (Schwalm et al., 2010).

Though these technologies exhibit unmatched potential for understanding ecosystem-atmosphere interactions, it is important to note that their 'real-world' operation is accompanied by practical challenges and limitations. For example, when operated above forested sites, the simplifying assumptions of the EC method are not always met, leading to measurements that may not be representative of the true ecosystem-atmosphere carbon exchange (Aubinet, 2008; Göckede et al., 2004; Gu et al., 2005; Massman and Lee, 2002; Paw U et al., 2000). These ideal requirements include stationarity of the data, homogeneity of the underlying surface, fully-developed atmospheric turbulence, and the absence of horizontal and vertical advection (Aubinet et al., 2001; Baldocchi, 2003; Foken and Wichura, 1996). Since the derived knowledge of EC flux studies is dependent on the nature of the measured data, it is crucial that EC data be properly quality-controlled to identify and subsequently remove or correct non-representative measurements (Aubinet, 2008; Barr et al., 2013; Foken and Wichura, 1996; Xuhui Lee et al., 2004). In particular, when EC study sites are fetch-limited, it is often necessary to consider the potential contamination of EC flux measurements by non-target surfaces (Göckede et al., 2004; Neftel et al., 2008; van de Boer et al., 2013).

At a larger scale, the intensification of high-frequency measurement systems such as EC – and their continuous operation at annual to decadal time scales – has led to a substantial increase in the diversity and quantity of data that must be collected, managed, processed, transferred, stored and shared properly throughout its entire data life cycle (Benson et al., 2010; Campbell et al., 2013; Porter et al., 2012). Those responsible for this data must be able to manage large quantities of data in various formats, and ensure that it is downloaded, processed and stored in an efficient and reliable manner (Campbell et al., 2013; Strasser et al., 2012) .

When the aforementioned challenges are considered in the context of single research groups – often the organizational units responsible for implementing and maintaining these systems – it is also necessary to acknowledge the presence of resource constraints on their operations. Considering both a general lack of ideal measurement sites, and an operational environment of funding uncertainty, short grant cycles, restrictions on technical staff, and high graduate student turnover rates, compromises must often be made between approaches that are ideal, and those that are feasible. In such cases, it is prudent that researchers take efforts to quantify and subsequently minimize the effect of such compromises on research outcomes.

Following these recommendations, the research presented in this thesis aims to quantify and minimize the effect of operational compromises on the results and derived knowledge of an ecosystem exchange measurement program over an age-sequence of temperate white pine forests.

1.1 Turkey Point Flux Station

The foci of this study are the Turkey Point Flux Station (TPFS) research sites, located in proximity to the north shore of Lake Erie in Norfolk County (42.71 °N, 80.36 °W), southern Ontario, Canada. TPFS is comprised of an age-sequence of planted and managed eastern white pine (*Pinus strobus* L.) forests, planted in 1939, 1974, 1989 and 2002, and herein referred to as TP39, TP74, TP89 and TP02. EC and continuous meteorological measurements commenced at these sites in 2002. A closed-path eddy-covariance (CPEC) system has been operated at the primary site (TP39) from 2002 until the present. Between 2002 and 2007, a single roving open-path eddy covariance (OPEC) system was moved between the younger sites (TP74, TP89, TP02) at semi-regular intervals of 2 weeks to 1 month. In 2008, permanent CPEC systems were installed at TP74 and TP02, while the TP89 site was retired from all measurements. In the context of forest micrometeorology studies, these sites may all be classified as area-limited, ranging between 0.40 km^2 at TP39 and 0.07 km^2 at TP02 ; these fetch

limitations are a consequence of the intensively fragmented and managed nature of the greater region. Detailed description of site characteristics may be found in Arain and Restrepo-Coupe (2005), Peichl and Arain (2006), and Peichl et al., (2010a, 2010b).

1.2 Objectives of this Research

This set of works aims to address the most important considerations for obtaining accurate, well-constrained characterizations of ecosystem carbon exchanges in our age-sequence of temperate white pine forests. Three objectives have been developed to address these challenges and achieve the research goal.

The quantity and diversity of data collected through biometeorological and ecological measurements at our site presents substantial data management and processing challenges. In order to ensure integrity and accuracy of analyses and disseminated results, it is critical that data be collected, managed, processed, transferred, stored and shared properly throughout its entire data life cycle. In research groups such as ours, where there does not exist a dedicated data manager, an added challenge is to develop a data system that operates effectively in a collaborative setting. As such, the first objective of this research is to develop and implement an effective and sustainable data workflow and management system to enable collaborative data management in biometeorological research.

The application of a roving eddy covariance system at our youngest study sites between 2002 and 2007 extended the number of sites that could be covered by a single EC system. However, this was done at the cost of significant challenges to extracting reliable carbon exchange information for any given site from such temporally-fragmented EC measurement time series. Quantifying the uncertainties in time-integrated carbon exchanges for these ecosystems provides important context and insights into these data and their utility. Additionally, if reliable estimates are possible, then such an approach may be applicable to other sites, in order to increase EC measurement coverage while requiring limited resources. Therefore, the second objective of this research is to estimate the uncertainty in time-integrated carbon estimates associated with roving EC operation at our sites, and assess its potential for wider use.

Given the fetch limitations at all of our measurement sites, it is necessary to consider the extent of EC measurement source areas when conducting analyses or producing ecosystem exchange estimates. Though numerous footprint models have been developed to estimate the EC measurement source area, comparisons have shown considerable variation among predictions (e.g. Kljun et al., 2003; van de Boer

et al., 2013). Without tracer experiments to assess footprint model predictions, it is important to characterize differences between each model, understand their implications on analyses, and seek secondary approaches to select the most appropriate footprint filtering application. In response to this need, the third research objective is to compare a number of analytical footprint models for predicting EC measurement source area at our sites, characterize their influence on intermediate and final analyses results, and develop a data-driven approach to evaluating footprint model performance.

1.3 Thesis Structure

The aforementioned research objectives are addressed in Chapters 2 through 4, which have been written as standalone research papers. In the interest of providing a cohesive and fully-referenced document, a small number of references are made between chapters. This has been done in respect to the chronological order in which these papers will be submitted. The following paragraphs provide a brief outline of each chapter.

Chapter 2: A system for collaborative data workflow and management in a long-term, multi-site biometeorology measurement program.

This paper documents the background, design and implementation of a comprehensive data workflow and management system for a long-term, multi-site biometeorological research program at our study sites. This system was designed and implemented between 2008 and 2011, to accommodate all biometeorological and ecological data collected by our research group. The system was also designed to facilitate collaborative data management amongst multiple group members. In this chapter, the data workflow and management system is critically evaluated by discussing the successes and limitations of this system, and assessing the long-term sustainability prospects of collaborative data management using this tool.

In the interest of providing thorough documentation to current and future researchers using this tool, this chapter has been made comprehensive in nature. This chapter is not submitted to an academic journal.

Chapter 3: Assessing the suitability of roving eddy covariance systems to produce reliable time-integrated carbon exchange estimates at multiple sites

This paper evaluates the roving eddy covariance (rEC) approach that was applied to our study sites, for its usefulness in deriving reliable time-integrated carbon exchange estimates both at our site and

generally. Monte Carlo analyses are used to quantify the net ecosystem exchange (NEE) uncertainty associated with different rEC measurement schedules (number of sites, rotation period, schedule timing), as well as the effects of different gap-filling models, and the application of a variety of proposed NEE uncertainty mitigation strategies. This chapter has been written for submission to Agricultural and Forest Meteorology.

Chapter 4: Implications of commonly-used footprint and friction velocity filtering methods on data availability and carbon exchange estimates in a fetch-limited temperate forest

This paper uses a data-driven, factorial analyses approach to characterize the differences among, and interaction between four analytical footprint models, and three u_*^{Th} estimation methods. These effects were evaluated in terms of their consequences for EC data quantity and distribution at our fetch-limited temperate forest, as well as subsequent implications for annual carbon exchange estimates for the TP39 study site. In this paper, we compare NEE estimates across different footprint models and footprint stringency specifications and evaluate the performance of these filtering approaches. A novel, data-driven approach is introduced, which uses internal gap-filling model performance metrics as a means of evaluating the consistency of filter-passing EC-measured data, and thus the ecosystem-representativeness of different footprint filtering approaches. The filtering approaches deemed most suitable for our site are applied to NEE time series, and subsequently gap-filled to produce best estimates of annual NEE, ecosystem respiration (RE) and gross ecosystem productivity (GEP) values. Differences between the new values and previous EC-and biometric-based estimates at this stite are critically investigated, and the nature of environmental control on carbon exchange – as inferred from our results – is discussed.

This chapter has been written for submission to Agricultural and Forest Meteorology.

Chapter 5: Conclusions

This chapter summarizes the findings of the three papers included in this thesis, discusses the results in the context of contributions to scientific understanding, and recommends directions for future research that follow from these works.

1.4 References

Arain, M., Restrepo-Coupe, N., 2005. Net ecosystem production in a temperate pine plantation in southeastern Canada. Agric. For. Meteorol. 128, 223–241.

Aubinet, M., 2008. Eddy covariance CO2 flux measurements in nocturnal conditions: an analysis of the problem. Ecol. Appl. 18, 1368–1378.

Aubinet, M., Chermanne, B., Vandenhaute, M., Longdoz, B., Yernaux, M., Laitat, E., 2001. Long term carbon dioxide exchange above a mixed forest in the Belgian Ardennes. Agric. For. Meteorol. 108, 293–315.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. Fluxnet: a new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82, 2415–2434.

Baldocchi, D.D., 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Glob. Chang. Biol. 9, 479–492.

Barr, A.G., Black, T.A., Hogg, E.H., Griffis, T.J., Morgenstern, K., Kljun, N., Theede, A., Nesic, Z., 2007. Climatic controls on the carbon and water balances of a boreal aspen forest. Glob. Chang. Biol. 13, 561–576.

Barr, A.G., Richardson, A.D., Hollinger, D.Y., Papale, D., Arain, M.A., Black, T.A., Bohrer, G., Dragoni, D., Fischer, M.L., Gu, L., Law, B.E., Margolis, H.A., McCaughey, J.H., Munger, J.W., Oechel, W.C., Schaefer, K., 2013. Use of change-point detection for friction–velocity threshold evaluation in eddy-covariance studies. Agric. For. Meteorol. 171-172, 31–45.

Benson, B.J., Bond, B.J., Hamilton, M.P., Monson, R.K., Han, R., 2010. Perspectives on next-generation technology for environmental sensor networks. Front. Ecol. Environ. 8, 193–200.

Campbell, J.L., Rustad, L.E., Porter, J.H., Taylor, J.R., Dereszynski, E.W., Shanley, J.B., Gries, C., Henshaw, D.L., Martin, M.E., Sheldon, W.M., Boose, E.R., 2013. Quantity is nothing without quality. Bioscience 63, 574–585.

Coursolle, C., Margolis, H.., Giasson, M.-A., Bernier, P.-Y., Amiro, B.D., Arain, M. a., Barr, A.G.,

Black, T. A., Goulden, M.L., McCaughey, J.H., Chen, J.M., Dunn, A.L., Grant, R.F., Lafleur, P.M., 2012. Influence of stand age on the magnitude and seasonality of carbon fluxes in Canadian forests. Agric. For. Meteorol. 165, 136–148.

Falge, E., Baldocchi, D., Tenhunen, J., Aubinet, M., Bakwin, P., Berbigier, P., Bernhofer, C., Burba, G., Clement, R., Davis, K.J., Elbers, J. A, Goldstein, A.H., Grelle, A., Granier, A., Guðmundsson, J., Hollinger, D., Kowalski, A.S., Katul, G., Law, B.E., Malhi, Y., Meyers, T., Monson, R.K., Munger, J.W., Oechel, W., Paw U, K.T., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Valentini, R., Wilson, K., Wofsy, S., 2002. Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. Agric. For. Meteorol. 113, 53–74.

Foken, T., Wichura, B., 1996. Tools for quality assessment of surface-based flux measurements. Agric. For. Meteorol. 78, 83–105.

Göckede, M., Rebmann, C., Foken, T., 2004. A combination of quality assessment tools for eddy covariance measurements with footprint modelling for the characterisation of complex sites. Agric. For. Meteorol. 127, 175–188.

Gu, L., Falge, E., Boden, T., Baldocchi, D.D., Black, T.A., Saleska, S.R., Suni, T., Verma, S.B., Vesala, T., Wofsy, S.C., Xu, L., 2005. Objective threshold determination for nighttime eddy flux filtering. Agric. For. Meteorol. 128, 179–197.

Hamilton, M.P., Graham, E.A., Rundel, P.W., Allen, M.F., Kaiser, W., Hansen, M.H., Estrin, D.L., 2007. New approaches in embedded networked sensing for terrestrial ecological observatories. Environ. Eng. Sci. 24, 192–204.

Janssens, I. A., Lankreijer, H., Matteucci, G., Kowalski, A. S., Buchmann, N., Epron, D., Pilegaard, K., Kutsch, W., Longdoz, B., Grunwald, T., Montagnani, L., Dore, S., Rebmann, C., Moors, E.J., Grelle, A., Rannik, U., Morgenstern, K., Oltchev, S., Clement, R., Gudmundsson, J., Minerbi, S., Berbigier, P., Ibrom, A., Moncrieff, J., Aubinet, M., Bernhofer, C., Jensen, N.O., Vesala, T., Granier, A., Schulze, E.-D., Lindroth, A., Dolman, A. J., Jarvis, P.G., Ceulemans, R., Valentini, R., 2001. Productivity overshadows temperature in determining soil and ecosystem respiration across European forests. Glob. Chang. Biol. 7, 269–278.

Kao, R., Gibson, C., Gallery, R., Meier, C., Barnett, D., Docherty, K., Blevins, K., Travers, P., Azuaje, E., Springer, Y., Thibault, K., McKenzie, V., Keller, M., Alves, L., Hinckley, E., Parnell, J., Schimel, D., 2012. NEON terrestrial field observations: designing continental-scale, standardized

sampling. Ecosphere 3, 1–17.

Kljun, N., Kormann, R., Rotach, M.W., Meixner, F., 2003. Comparison of the langrangian footprint model LPDM-B with an analytical footprint model. Boundary-Layer Meteorol. 106, 349–355.

Lee, X., Finnigan, J., U, K.T.P., 2004. Coordinate systems and flux bias error, in: Lee, X., Massman, W.J., Law, B. (Eds.), Handbook of Micrometeorology. Kluwer Academic Publishers, Dordrecht, Germany, p. 250.

Massman, W.J., Lee, X., 2002. Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges. Agric. For. Meteorol. 113, 121–144.

McLaren, J., Arain, M., Khomik, M., Peichl, M., Brodeur, J.J., 2008. Water flux components and soil water-atmospheric controls in a temperate pine forest growing in a well-drained sandy soil. J. Geophys. Res. 113, 1–16.

Moffat, A., Papale, D., Reichstein, M., Hollinger, D., Richardson, A., Barr, A., Beckstein, C., Braswell, B., Churkina, G., Desai, A., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. 147, 209–232.

Neftel, A., Spirig, C., Ammann, C., 2008. Application and test of a simple tool for operational footprint evaluations. Environ. Pollut. 152, 644–52.

Olajuyigbe, S., Tobin, B., Saunders, M., Nieuwenhuis, M., 2012. Forest thinning and soil respiration in a Sitka spruce forest in Ireland. Agric. For. Meteorol. 157, 86–95.

Paw U, K.T., Baldocchi, D.D., Meyers, T.P., Wilson, K.B., 2000. Correction of eddy-covariance measurements incorporating both advective effects and density fluxes. Boundary-Layer Meteorol. 97, 487–511.

Peichl, M., Arain, M.A., 2006. Above- and belowground ecosystem biomass and carbon pools in an age-sequence of temperate pine plantation forests. Agric. For. Meteorol. 140, 51–63.

Peichl, M., Arain, M.A., Brodeur, J.J., 2010a. Age effects on carbon fluxes in temperate pine forests. Agric. For. Meteorol. 150, 1090–1101.

Peichl, M., Brodeur, J.J., Khomik, M., Arain, M.A., 2010b. Biometric and eddy-covariance based estimates of carbon fluxes in an age-sequence of temperate pine forests. Agric. For. Meteorol. 150, 952–965.

Porter, J.H., Hanson, P.C., Lin, C.-C., 2012. Staying afloat in the sensor data deluge. Trends Ecol. Evol. 27, 121–9.

Porter, J.H., Nagy, E., Kratz, T.K., Hanson, P., Collins, S.L., Arzberger, P., 2009. New eyes on the world: advanced sensors for ecology. BioScience 59, 385–397.

Richardson, A.D., Black, T.A., Ciais, P., Delbart, N., Friedl, M.A., Gobron, N., Hollinger, D.Y., Kutsch, W.L., Longdoz, B., Luyssaert, S., Montagnani, L., Munger, J.W., Moors, E., Piao, S., Reichstein, M., Saigusa, N., Tomelleri, E., Vargas, R., Varlargin, A., 2010. Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philos. Trans. R. Soc. London. Ser. B Biol. Sci. 365, 3227–3246.

Richardson, A.D., Hollinger, D.Y., Aber, J.D., Ollinger, S. V., Braswell, B.H., 2007. Environmental variation is directly responsible for short- but not long-term variation in forest-atmosphere carbon exchange. Glob. Chang. Biol. 13, 788–803.

Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agric. For. Meteorol. 169, 156–173.

Saunders, M., Tobin, B., Black, K., Gioria, M., Nieuwenhuis, M., Osborne, B.A., 2012. Thinning effects on the net ecosystem carbon exchange of a Sitka spruce forest are temperature-dependent. Agric. For. Meteorol. 157, 1–10.

Schwalm, C.R., Williams, C.A., Schaefer, K., Anderson, R., Arain, M.A., Baker, I., Barr, A., Black, T.A., Chen, G., Chen, J.M., Ciais, P., Davis, K.J., Desai, A., Dietze, M., Dragoni, D., Fischer, M.L., Flanagan, L.B., Grant, R., Gu, L., Hollinger, D., Izaurralde, R.C., Kucharik, C., Lafleur, P., Law, B.E., Li, L., Li, Z., Liu, S., Lokupitiya, E., Luo, Y., Ma, S., Margolis, H., Matamala, R., McCaughey, H., Monson, R.K., Oechel, W.C., Peng, C., Poulter, B., Price, D.T., Riciutto, D.M., Riley, W., Sahoo, A.K., Sprintsin, M., Sun, J., Tian, H., Tonitto, C., Verbeeck, H., Verma, S.B., 2010. A model - data intercomparison of CO_2 exchange across North America: Results from the North American Carbon Program site synthesis. J. Geophys. Res. 115.

Son, Y., Jun, Y.C., Lee, Y.Y., Kim, R.H., Yang, S.Y., 2004. Soil carbon dioxide evolution, litter decomposition, and nitrogen availability four years after thinning in a Japanese larch plantation. Commun. Soil Sci. Plant Anal. 35, 1111–1122.

Strasser, C., Cook, R., Michener, W., Budden, A., 2012. Primer on data management: What you always wanted to know (but were afraid to ask). DataOne.

Valentini, R., Matteucci, G., Dolman, A.J., Schulze, E.D., Rebmann, C., Moors, E.J., Granier, A., Gross, P., Jensen, N.O., Pilegaard, K., Lindroth, A., Grelle, A., Bernhofer, C., Grünwald, T., Aubinet, M., Ceulemans, R., Kowalski, A.S., Vesala, T., Rannik, U., Berbigier, P., Loustau, D., Gudmundsson, J., Thorgeirsson, H., Ibrom, A., Morgenstern, K., Clement, R., 2000. Respiration as the main determinant of carbon balance in European forests. Nature 404, 861–5.

Van de Boer, A., Moene, A., Schüttemeyer, D., Graf, A., 2013. Sensitivity and uncertainty of analytical footprint models according to a combined natural tracer and ensemble approach. Agric. For. Meteorol. 169, 1–11.

2   A system for collaborative data workflow and management in a long-term, multi-site biometeorology measurement program

2.1 Abstract

Modern investigations of biosphere-atmosphere interactions have benefited from developments in measurement equipment, as well as data storage, analysis and transfer technology. These technologies have permitted the long-term and continuous collection of many variables at high or very high frequencies and, as a result, have led to dramatic increases in biometeorological research data production. In order to address the issue of the so-called 'data deluge', there has been heightened focus on developing and implementing proper data management operations through all stages of the biometeorological data life cycle. Though such actions are necessary to ensure integrity and quality of data-derived results, implementing effective and sustainable data workflow and management systems presents numerous data- and resource-related challenges. In this descriptive paper, we document the background, design and implementation of a comprehensive data workflow and management system for a long-term, multi-site biometeorological research program. The biometeorological data life cycle is first presented to provide context to the work, and is followed by an assessment of the requirements and challenges associated with implementing effective data management and processing workflows in biometeorological research programs. Following this, we describe the design and operation of our data workflow and management system, which was designed to facilitate collaborative data management amongst multiple group members. We discuss the successes and limitations of this system, and assess the long-term sustainability prospects of collaborative data management.

2.2 Introduction

In the field of biometeorology, advances in instruments and systems over the past 20 years have provided researchers new tools for investigating atmosphere-biosphere interactions, fostered new ways in which these tools may be applied, and have allowed researchers to pose new hypotheses and evaluate them experimentally (Baldocchi, 2003; Benson et al., 2010). The development of instruments such as fast-response anemometers, gas analyzers and data loggers provide researchers with continuous (or at least, continual) ecosystem exchange measurements over a wide range of time scales. Simultaneously, increasing amounts of automated meteorological instruments provide more comprehensive and higher resolution information about the state of the environment where these exchange measurements are taken (Campbell et al., 2013; Peters, 2010; Porter et al., 2012). At larger scales, the establishment of research programs and networks at national and international levels has provided the infrastructure needed to study these interactions in more places and at longer time scales,

improving understanding of ecosystem-atmosphere energy and mass exchange dynamics across both space and time (Baldocchi et al., 2001; Coursolle et al., 2012a; Kao et al., 2012; Keller et al., 2008; Melaas et al., 2013).

Accordingly, measurement intensification and the application of systems over longer periods and more sites has greatly increased the quantity and diversity of data that must be managed at the level of the research group unit (principal investigator, research students and staff). Those responsible for this data must be able to manage large quantities of data in various formats, and ensure that it is downloaded, processed and stored in an efficient and reliable manner (Campbell et al., 2013). At a larger scale, the development of e-science and extra-institutional research networks has increased the need for sharing scientific data, which may be used for such activities as research collaboration, synthesis studies, data mining and data re-visitation (Baker and Barton, 2009; Uhlir and Schröder, 2008). As a result, research groups in biometeorology must devote considerable resources to data stewardship activities, in order to ensure that their data is collected, managed, archived, accessed and shared correctly (Campbell et al., 2013; Hamilton et al., 2007).

Determining best-practices for a research group's data management activities depends on the data need and the existing research environment. In an ideal situation, all data management activities are overseen by a permanent, highly-skilled and dedicated data manager, who is responsible for designing and implementing data management plans and systems. This individual ensures that data operations are consistent, and that top-quality data products are shared with end-users for their analyses. However, dedicated data managers are not always employed within research groups – a result of high position turnover, funding limitations (by both monetary amounts and term length), or a lack of compelling incentive for administrators to employ dedicated data managers and carry-out proper data management (Barton et al., 2010).

In response to this, it may be necessary for data management to occur in a collaborative framework, where each member of the research group is tasked with independently managing all data pertaining to their own research and making their final data products available to internal and external end-users. In this setting, data stewardship arises from the combined individual efforts of all members, either with or without the presence of a coordinating framework. Though this framework allows for division of the data management workload, there are a number of potential drawbacks to this approach, which may have negative consequences for data quality within the research group:

The regular turnover of research group members and their discrepancies in data management skills can

lead to inconsistent quality across the research group's data products. Furthermore, data quality inconsistency can also arise from varying levels of incentive across members towards developing data beyond what is necessary for their individual research (Peters, 2010).

Second, unless methods and formats are standardized across members, data products will suffer from inconsistencies and incompatibilities, making it difficult to assimilate different outputs into a single, coordinated product. Particularly, standardized methods should restrict members from managing data through manually-driven means (e.g. spreadsheets), as such approaches are prone to manipulation errors that are commonly latent, untraceable and unreproducible in nature (Cook et al., 2001; Hook et al., 2010; Strasser et al., 2012).

Third, in cases where members are carrying out data operations on separate, distributed computer systems, problems can arise from poor data and methodology coordination. Version discrepancies can arise in situations where processing methodology and data are shared among a number of systems, which increases the likelihood of processing errors and data loss due to mismanagement. Such situations also result in higher operational redundancy, and lag times for assimilating data products.

Finally, the lack of a permanent overseer of data management activities has negative consequences for its sustainability and future development. Without a centralized individual to facilitate effective communication between members and to provide leadership for data management activities, there exists less capacity for the collaborative management framework to function optimally in the present or the future.

Though a collaborative framework may be necessary in many settings to yield an acceptable amount of dedicated data management time for all members (Buneman et al., 2011; Ives et al., 2008), the potential for one or many of these aforementioned problems to arise should be addressed within the research group, as they may prove detrimental to productivity, data quality and quantity, and researcher reputation. In such cases, it is imperative that research group members collaborate to develop a data management plan that considers all data operations, and seeks to employ a data workflow and management system (DWMS) that incorporates best-practices at all stages (Lyon, 2007; Sallans and Donnelly, 2012; Strasser et al., 2012). An effective DWMS coordinates and standardizes data management activities, and facilitates communication of data management details between the research-group members and with the external scientific community, to promote proper data use and leverage multi-institutional collaborative learning opportunities.

Addressing such a need, this article details the development and application of a collaborative, research group-maintained DWMS used in biometeorological research. The role of data management in the biometeorological data life cycle is first introduced, along with design considerations for creating a DWMS to implement data management best-practices. The DWMS is then presented, with all components described in detail within the context of the data life cycle in biometeorological research. A discussion of the DWMS follows, with an assessment of its benefits, shortcomings and long-term success potential, as well as recommendations for future development and refinement.

2.3   Data management in biometeorology

2.3.1   The biometeorological data life cycle

Data is an integral part of the scientific process – it consists of the observations that form the basis for the scientific method. As such, effective data management is critical to measurement-based research, as it enables the development of raw, collected data into refined data products that may be used to evaluate specific research hypotheses, from which knowledge is ultimately gained. The data life cycle serves as a context for data management in scientific research, as it outlines the flow and development of data in scientific research activities. Data is collected (created), processed, analyzed, stored and then shared with other users, where it is subsequently incorporated into new data life cycles. The adaptation of the data life cycle framework in biometeorological research is shown in Figure 2.1; specific activities are listed for each stage, and are described further below.

1   Creation
- Design study
- Select site and location
- Set up site, tower, equipment
- Measure primary data
- Collect ancillary data
- Transfer data to file system
- Create metadata

2   Processing
- Read and organize data from field into master files
- Calculate time-averages from high-frequency data
- Assess appropriate data filtering criteria (u , footprint)
- Conduct quality control (inspect, clean, calibrate data)
- Create intermediate data products (if necessary)
- Fill data gaps or flag unfilled data
- Compile and store master files for each quality level
- Compile and save data products
- Create metadata for data and processes

3   Analysis
- Integrate data from various sources/years
- Calculate time-integrated data values (days-years)
- Conduct appropriate statistical analyses
- Use data to run/calibrate/evaluate ecosystem models
- Use data in model-data fusion activities
- Output and save derived data products
- Produce figures and publications

4   Preservation
- Actively back-up all instances of data and software
- Organize/maintain backup media
- Compile, save, back-up all metadata and documentation
- Standardize all data products for sharing and archiving
- Archive previously distributed data products

5   Access
- Make data products available within research group (manual collection, network or internet-accessible share)
- Upload selected data products to external users (Internet share, ftp server, upload to DIS)
- Make metadata available
- Set data fair-use guidelines
- Communicate data opportunities for potential users

6   Reuse
- Incorporate analysis products in filesystem
- Use data products in synthesis studies
- Make data available for data revisitation/ cross-analysis initiatives
- Use data products as a teaching and learning tool

Figure 2.1*: A diagrammatic representation of the biometeorology data life cycle. The cycle is initiated with the act of data creation, and re-initiates when shared data products are incorporated into new data collection procedures. Though data analysis represents the third stage of the data life cycle, it does not need to immediately succeed the data processing stage, as it may be undertaken at any subsequent point in time. Preservation and access (sharing) operations are applied to data products of both processing and analysis stages.*

Data creation (box 1, Figure 2.1) is initiated by research design, where the theoretical and operational details of the project are addressed and formalized. Considerations made during the design process include: research site selection, equipment setup, and the meteorological, gas exchange and ancillary data variables that are to be measured. In addition, research design also necessitates the development of a data management plan, which specifies the means and media by which field-collected data will be captured, stored, collected and processed, and ultimately shared with data users (Van Den Eynden et al., 2011). A data management plan should consider not only the equipment, protocols and documentation needed to deliver high-quality data products, but also the responsibilities and time commitments of individuals involved in data management. Once instruments are employed in the field, measured data is collected from all systems at appropriate intervals, and transferred to the desired file systems for processing.

Processing collected data makes up a considerable fraction of the operational and time requirements in the biometeorological data life cycle (box 2, Figure 2.1), as considerable intermediate data operations are required to refine raw data into end-user-sharable data products. The prevalence of measurement errors and non-ideal measurement conditions require raw and calculated data to undergo a number of quality assurance (QA) operations (Foken et al., 2004; Göckede et al., 2008; X. Lee et al., 2004; Mauder et al., 2008; Zeri and Sá, 2010). Quality-assured data may also need to be gap-filled to suit end-user requirements for continuous variables (Falge et al., 2001; Moffat et al., 2007). The succession of operations results in a hierarchy of data quality 'levels', which culminate in final data products that are made available to end-users for analysis. Creating thorough metadata is critical at this stage of the data life cycle, as documents describing both data operations and products serve to improve data reproducibility and end-user usability.

Analyses conducted on processed data are specific to the goals and objectives of the study, as well as the quantity and quality of available data (box 3, Figure 2.1). Analyses typically involve the integration and detailed assessment of meteorological, remotely-sensed, flux and ancillary data products, at timescales of days to years. Processed data may be used to drive or assess ecological models, or incorporated into model-data fusion approaches. Data products from this step are typically incorporated into published results and figures, and are available for insertion into new data life cycles.

Preservation activities ensure short- and long-term file organization, archival and backup for all data-associated systems (box 4, Figure 2.1). Preserved files include all levels of data (raw, intermediate and final), as well as processing scripts, documentation and metadata. Data backups protect against

losses due to user error, software bugs, data corruption, or hard disk malfunction. Files may be backed up using automated or manual means, to a number of possible storage media types, including magnetic and optical discs, as well as online and hard copy repositories. In addition to the file system of the data processing computer, backups may also be carried out on data collection devices in the field, to preserve data as it is created. Appropriate file organization is another important consideration for all devices that collect or store data; well-organized file systems help to streamline and simplify manual and automatic data operations, creating a more robust data workflow and management system. Depending on file system storage capacities, user and data manager needs, and file access preferences, the file system may archive all files on the original file system, or selected data files only.

Data and metadata products are shared with end-users located within and (commonly) outside of the research group (box 5, Figure 2.1). End-users within a given research-group may access data and metadata products via a variety of methods, including manual transfer on removable media, or through Internet- and network-accessible shares. Sharing data at the research-network scale can be accomplished by the transfer of data to interested parties via physical or digital means, or by uploading files to an external, centralized data server, which is operated at an multi-institutional national or international level. Interested end-users with the proper data permissions may then download files shared on the centralized data server for use in their own analyses. The format, type and amount of data and metadata shared with local and external end-users may be specifically tailored to individual and organizational requirements.

Data that is shared with end-users can be re-used in novel applications (box 6, Figure 2.1), which include independent analyses, large-scale synthesis studies, and cross analysis and data revisitation efforts (Hook et al., 2010). Shared data products may also be used in training applications within the research group, or in classroom teaching and learning activities. The provision of high quality metadata is integral to the success of data re-use activities, as it is critical to ensure that end-users understand and use the shared data properly.

2.3.2   Design considerations for a biometeorology DWMS

In the context of the biometeorological data life cycle at the research-group scale, a successful DWMS provides the framework and the tools necessary to efficiently and effectively move data through each of the previously-described stages. Such a system consists of a number of integrated components,

including: a) hardware devices (instruments, data loggers, data servers, communication devices, etc.) that create, move and transform data, b) software routines that provide instructions for hardware operations, and c) the various forms of media, upon which the data is stored and moved. For successful and effective operation, an appropriate DWMS must be capable of implementing data management best-practices at each stage of the data cycle. Incorporating such guidelines into a comprehensive data management plan improves the performance of the DWMS, and maximizes the consistency, quality, longevity and usability of its generated data products. In biometeorological research, a number of established data management best practices should be used to guide DWMS development. A summary of these guidelines, and their benefit to DWMS operation is provided below, alongside a review of considerations for operating such a system in a collaborative-management framework.

2.3.2.1 General design

DWMS design should be initiated by careful consideration of the desired functionality, its operating environment, and the resources (equipment, time, funding, people, etc.) available for system development and maintenance. Consultation between potential DWMS operators and data end-users is an important step in the design stage, as it helps to define system expectations and constraints, and set developmental time lines (Zowghi and Coulin, 2005).

Important general considerations for biometeorological DWMS design include provisions for long-term system consistency, flexibility and scalability; an effective DWMS should accommodate expansion in any or all of its components, permitting the straightforward integration of new instruments, data inputs, software operations, and data products. DWMS hardware and software should address needs for long-term device compatibility, as well as archival and backup requirements. All components of the DWMS should support data diversity, acknowledging that data may be collected from a wide range of devices on various media types, and may be comprised of an assortment of data formats, quality levels and measurement frequencies.

2.3.2.2 Field data collection and organization

A primary consideration for implementing best-practices in field operations is the standardization of hardware, data files and operational activities. Standardizing instruments, collection hardware and storage media across all measurement systems serves to streamline data collection operations, while reducing the likelihood of data collection errors. Similarly, the use of standard collection software, data formats and storage structures improves the efficiency and effectiveness of data processing functions, reducing the occurrences of data incompatibilities, and diminishing the need for additional, customized

processing steps (Cook et al., 2001; Strasser et al., 2012). To complement hardware and software considerations, field data collection protocols should also be explicitly documented and standardized, to minimize (or eliminate) variations in field data collection methods, which may have negative consequences for data processing and management operations.

Furthermore, effective field data collection and organization requires comprehensive documentation of all standardized elements of field operations, as well as thorough logging of all field activities (Strasser et al., 2012). Such actions ensure communication between all involved field measurement personnel, and provide a means of tracking potential operator or equipment errors.

2.3.2.3   Data operations and software

Prior to developing and implementing DWMS software, it is important to define the scope of data operations that are required. This includes soliciting the research group's data-specific needs, and any constraints or special considerations that will need to be addressed (Wiegers, 2003). Software design involves specifying the general (architectural) structure of the program, as well as the operational details of each of its components and their interfaces (IEEE Computer Society, 2004a). Common basic principles should be adhered to during software design, namely: data abstraction through parameterization and specification; cohesion within modules and coupling between them; modularization of large programs into smaller, independent operations; encapsulation of software elements where needed; separation of the operator interface from internal processes; and construction of sufficient and concise programs (Abelson et al., 1996; IEEE Computer Society, 2004b; Kazman et al., 1994). In biometeorological DWMS applications, the implemented software should allow the operator to perform all necessary organizational, quality assurance and data calculation operations that are needed to create sharable data products from raw input data (Campbell et al., 2013; Cook et al., 2001). Although software design details will vary with the specific characteristics of a research program, a number of design considerations should be commonly shared among all implemented biometeorological DWMS software:

DWMS software operations should be as modular as possible, where independent and separate data processes may be individually executed, modified, updated or replaced without significantly affecting the rest of the software. Software should also be extensible and compatible, permitting it to incorporate new data sources, types and operations into the system with minimal modifications. Furthermore, DWMS software should provide utilities that check for errors in datasets and processes (either manually or automatically), and subsequently report all problems encountered in DWMS software

operation via log files or real-time (command-line) warnings. DWMS checking and reporting functionality allows operators to promptly identify and remediate problems with data collection, storage or processing. Finally, software should be usable, robust and intuitive, to ensure that it can be easily learned and used correctly by DWMS operators.

To achieve such functionality, all eligible data processing operations should be constructed and implemented using a scripted programming language, (Fortran, R, SAS, SPSS, MATLAB, etc.), as this approach offers numerous advantages over manual, graphically-driven data processing (Campbell et al., 2013; Hook et al., 2010). The textual commands of scripts and functions comprehensively document the operations carried out during data processing, which provides the DWMS operator with methodological consistency, traceability and reproducibility (Borer et al., 2009; Hook et al., 2010). The accessible and editable nature of processing scripts increases the efficiency with which methodological changes may be made and recalculations performed.

2.3.2.4   File organization and documentation

Data files should be structured and stored on the DWMS file system in a logical and consistent manner that clearly separates different data products and quality levels (Hook et al., 2010). The DWMS file system should exist at a single, centralized location, to ensure that data organization and processing methods are standardized across DWMS operators. Using a single, shared system promotes operational consistency, and eliminates potential divergence of software versions, data processes, metadata and data sets that occur over time in distributed systems. A centralized system also eliminates process and data redundancies, which saves operator time and reduces the chances that data will be unintentionally overwritten. Centralization provides a foundation to perpetually build upon collectively-gained group knowledge and experience, and it ensures that a legacy of documentation and consistent data is left to future users.

Data should also be be organized clearly and consistently within individual files; files containing similar information should be formatted as similarly as possible to increase end-user familiarity and maximize interoperability and consistency within the DWMS software (Cook et al., 2001). A consistent and descriptive directory and file naming convention should also be implemented, based on common identifiers such as data type, quality level, measurement date and site (Hook et al., 2010). This organization enables automated file retrieval by DWMS software, and makes file searching and identification much easier for end-users and data managers.

To further improve usability, information regarding data file contents should be provided in the form of

a descriptive header at the top of the file, or in complementary metadata files, which give information on the units, data quality levels and other measurement details of the included variables (Borer et al., 2009; Hook and Christensen, 2005). Additionally, the structure of the DWMS file system and the expected contents of its directories should be thoroughly documented and made available to DWMS operators (Karasti and Baker, 2008; Van Den Eynden et al., 2011).

2.3.2.5 Data backup, archival and sharing

Effective data preservation requires regular backup and archiving of the DWMS file system. All files (programs, data, documentation, etc.) that are likely to be used in future operations should be archived on the DWMS file system to provide prompt access to data. To protect against potential data loss, backup operations should be implemented on all data and programs at every stage of the data life cycle. If feasible, data logged onto a collection device should be backed up to an additional location in real-time, to avoid data loss from logger malfunction. All levels of data and supporting documentation (internal and shared) on the DWMS file system should be backed up to one or more additional media devices; preferably, this data should be backed up and stored at locations that are remote to the central file system. All files associated with DWMS operation should be backed up in the same manner, especially all data processing scripts and functions.

Shared data products should be made available through means that allow convenient access for the intended end user. For example, a network-accessible server may be used to provide data access to research-group (local) end-users, while external end-users may need to access data through external (FTP, SSH) servers, or via content that is uploaded to an external data server. Data may also be shared among local end-users by external media such as flash drives, but care should be taken to ensure that harmful software (i.e. viruses and malware) is not transferred to the DWMS file server. In many cases, providing access to data through shared Internet collections is an effective way to serve both local and external end-users, while ensuring long-term and reliable data accessibility (Borer et al., 2009). To maximize long-term compatibility and interoperability, data and metadata products should be provided to end-users in non-proprietary, stable formats, such as CSV, PNG, ASCII text formats (Hook et al., 2010). Additionally, data managers should archive all previously shared data products, in order to trace any errors that may be discovered in previous versions, and communicate potential problems to affected end-users and data stakeholders.

2.3.3   Considerations for collaborative DWMS application in biometeorology

Operating a DWMS in an environment where multiple users share data management responsibilities on a centralized system presents a number of challenges beyond those considered for more traditional settings, where only one or a limited number of experienced and dedicated data managers are responsible for dealing with data. In such cases, the following requirements for DWMS operations should be implemented:

Executing data operations within the DWMS must be straightforward, understandable and accessible for all operators. By providing a simple and intuitive graphical user interface (GUI) to execute programs, data managers are able to perform consistent data operations regardless of their experience and skill with scripted languages, thus removing a barrier to successful DWMS operation. In addition, restricting inexperienced users or users-in-training to GUI-driven processes only (no process modification) reduces the potential for many common processing errors, including: syntax errors, operations executed improperly or in an incorrect order, and the introduction of programming bugs. Furthermore, the GUI may be used as a training tool, through which the trainee can visualize and understand processing operations while simultaneously learning how to decipher the detailed operations contained within the scripted programs.

Where possible, eligible data operations should be automated, so to be executed at regular intervals without the requirement of operator supervision. In cases where consistent rules are applied to data, the process may be executed automatically by the DWMS data processing software; this saves time for the data manager, ensures that operations are executed at proper intervals, and provides end-users access to data more promptly. Effective automation requires that processes have proper error-checking mechanisms implemented, where a data manager can review program activities via log files and inspect the data output from the automated processes.

Effective communication between DWMS operators is a crucial aspect of successful operation of a collaborative DWMS. Roles and expectations in DWMS operation should be clearly stated and documented for all involved individuals; this facilitates timely operation of data management activities, and reduces operational redundancy among data managers. Incoming DWMS operators should be provided with proper training to ensure that all aspects of the DWMS and their responsibilities are clearly understood. Proper training includes hands-on instruction by existing

DWMS operators and access to thorough DWMS documentation.

Thorough documentation for all system components should be made available to DWMS operators. Documentation should include a description of each component of the DWMS and instructions for its operation and modification, as well as thorough commenting in all data processing scripts. These documents should be shared with the entire group in a manner that allows collaborative editing, and should be actively updated by operators as changes are made to any aspect of the DWMS.

2.4 DWMS design and implementation

2.4.1 Rationale and motivation

The following section outlines the design features and operational details of a DWMS created for use in biometeorological research by the Hydrometeorology and Climatology Research Group at McMaster University in Ontario, Canada. The DWMS development project was initiated in mid-2008, motivated by a critical evaluation of existing data operations and previously-prepared data sets. The evaluation produced concerns with previously-prepared data products, which included processing and matrix organization errors, missing or inconsistent data operations, and minimal documentation for data management and processing. Without the services of a dedicated data manager, a group-wide initiative was undertaken with the goal of centralizing, standardizing, streamlining and documenting all components of the data life cycle for our field observation sites (see section 2.4.3.1), to develop a DWMS capable of being maintained in a collaborative graduate-student-managed framework.

The following subsections describe the components, methodologies, operations and functionalities of the DWMS in the context of the biometeorological data life cycle (see section 2.3.1). The creation of a data management plan to guide DWMS development is described first, followed by the details of DWMS design and implementation to facilitate the advancement of data through the biometeorological data life cycle. Design considerations and advanced functionalities of the DWMS are then presented, ending in a description of all types of supporting documentation that is used and generated by the DWMS.

2.4.2 DWMS needs assessment and data management plan

A data management plan was developed by the research group in mid- to late-2008 as the formative stage of DWMS development. The purpose of this plan was to characterize the scope and operations of data management in the research group, and to design the systems (media, instruments, hardware, software, etc.), operations (processing methods, transfer and sharing, etc.), protocols (data formats, manual operations), and documentation (manuals, metadata, etc.) of the DWMS. The biometeorology DWMS considerations presented in section 2.3.2 were used as guiding principles for DWMS development, defining the desired characteristics of the completed system.

DWMS components were created and integrated in response to the research group's data management needs, and the data product requirements of internal and external end-users. The DWMS was designed to preserve the nature of data management responsibilities within the research group, where each member was tasked with managing the data for particular systems and sites, though data collection could be carried out by any member during a field site visit. This management structure required that data collection methods be standardized, documented, and streamlined, and that downloaded data be collected to a centralized location. In addition, data processing and archival operations needed to be centralized and standardized to ensure compatibility between data products originating from different systems and users. It was pertinent that data processing operations be repeatable, streamlined and easy to implement, so that data could be readily updated to allow field quality-assurance inspections, and to regularly provide end-users with up-to-date data. To serve end-users' diverse data needs, it was critical that the DWMS was capable of successfully integrating data from a variety of sites and systems, and that it could conveniently and robustly export these data into products in an assortment of desired data formats.

The data management plan also recognized that in order to remain viable and sustainable, it was important for the DWMS to operate within the system and human resources that were presently available, and those that were anticipated to exist in the future. The most important of these considerations was the need for the DWMS to be operable and sustainable without the services of a dedicated data manager; the DWMS needed to function in an environment where research group members (i.e. approximately 6 graduate students) can maintain and implement all aspects of data management in a consistent and organized manner. Ensuring that this need was met in the DWMS required the incorporation of the considerations outlined in section 2.3.3. A central file system and

operational software were deemed necessary for development, so that data managers (i.e. research group members) could incorporate and manage previous and existing data files from all sites and systems. For other components of the DWMS, the existing operations and systems were to be retained or modified, in order to optimize DWMS functionality and minimize time and equipment costs.

2.4.3   DWMS design and implementation

The DWMS was developed between late-2008 and early-2011, during which time it was expanded to incorporate progressively more of the research group biometeorological data operations. A schematic diagram of the DWMS is presented in Figure 2.2, which depicts the integrated components and operations that advance data through all stages of the biometeorological data life cycle.

A variety of media and methods are used to transfer raw field measurements from their associated data loggers to the DWMS central file system. Using automated and operator-guided means, customized software routines in the DWMS central file system process raw data into quality-controlled, sharable products for end-users. Selected data products are then made available to internal (research group-based) and external end-users through a variety of different data-sharing protocols. New data and software products that have been created from end-user analyses may then be incorporated into the central file system for further processing, if desired. Data is archived and preserved throughout the data life cycle, protecting against data loss at any stage of data development.

In the context of Figure 2.1 and the biometeorological data life cycle, these components and their integration into a collaboratively-managed DWMS is detailed in the subsections below.
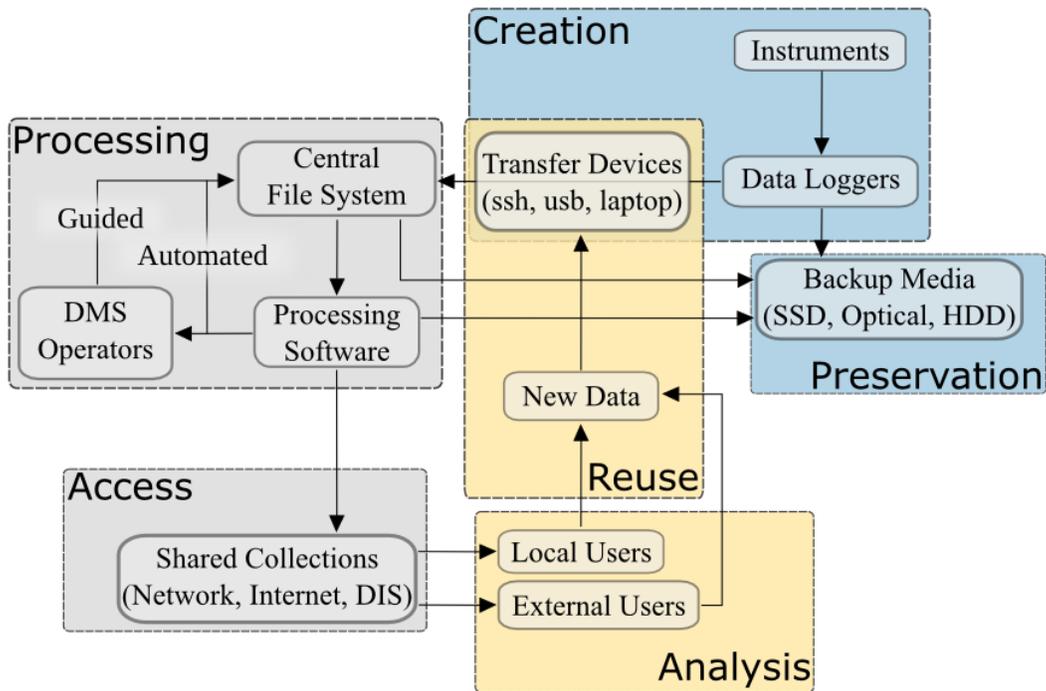
Figure 2.2*: Schematic diagram of the biometeorological data workflow and management system (DWMS) designed and implemented in this study. Data flow (solid arrows) initiates with instrument measurement and continues through each stage of the data life cycle (large, dashed-outline coloured boxes) via an assortment of associated system components (solid-outline rounded boxes).*

2.4.3.1   Data creation: Study sites, measurement, field data collection and upload

Study sites and instrumentation description

Field data collection is carried out at the Turkey Point Flux Station (TPFS) research sites, located in proximity to the North Shore of Lake Erie in Norfolk County (42.71 °N, 80.36 °W), in southern Ontario, Canada. TPFS is comprised of an age-sequence of planted and managed eastern white pine (*Pinus strobus* L.) forests, planted in 1939, 1974, 1989 and 2002, herein referred to as TP39, TP74, TP89 and TP02. Complete site descriptions may be found in Arain and Restrepo-Coupe (2005), Peichl and Arain (2006), and Peichl et al. (2010). EC and primary meteorological measurements commenced at these sites in 2002, and have continued to the present, with the exception of TP89, which was discontinued in the summer of 2008.

Primary meteorological systems consist of instrument assemblages specified by Canadian Carbon Program (formerly Fluxnet-Canada Research Network) measurement guidelines and standards (Fluxnet-Canada, 2003). Table 2.1 provides a detailed summary of instrument types, logging devices and operational periods for each site. Common meteorological variables (air temperature, T_a; relative humidity, RH; photosynthetic photon flux density, PPFD; wind speed, WS; wind direction, WD; net radiation, R_n) are measured above the forest canopy, from measurement towers installed at each site. Soil temperature (T_s), soil volumetric water content (VWC) and soil heat flux (G) are measured in separate, replicated depth profiles beneath the forest floor (see Table 2.1 for instrument details). All data is collected and logged to data logging devices (models CR10x, CR23x, CR1000, Campbell Scientific Inc., Edmonton, AB, CA) as half-hourly averages. Prior to 2008, these data loggers (at all sites but TP39) were operated in a standalone fashion, but are now connected to field desktop personal computers (PCs), which download and store data files at short, regular intervals. A full description of meteorological instrumentation at each site is provided by Peichl et al. (2010).

A number of supplementary meteorological systems have been installed and operated at sites to provide additional measurements (Table 2.2). A water table measurement system was implemented in 2010 to monitor groundwater table height at TP39; a water level pressure sensor (model PLS, OTT, Kempten, DE) was positioned 8.53 m below ground level, in a 10 cm diameter steel-cased water monitoring well, which was installed in 2009. Half-hourly water table height and temperature data were recorded by a standalone data logger (CR10x). Sap velocity measurement systems have been installed and operated at TP39 and TP74 since 2008 and 2009, respectively. Thermal dissipation probes (model TDP-30, Dynamax, Houston, Texas, USA) were installed in representative trees at both sites, and values are

recorded by standalone data loggers (CR10x at TP39, CR1000 at TP74). Further discussion of thermal dissipation probe installation and operation at TP39 is provided in Mackay et al. (2012)

An additional site (TP_PPT), has been operated from 2008 to the present, in order to provide reference ground-based precipitation measurements for TPFS. Cumulative rainfall is measured by an all-weather accumulation precipitation gauge (model T200B, Geonor Inc., Milford, PA, USA) and is recorded

Table 2.1: *Primary meteorological system details at TPFS*

	TP39	TP74	TP89	TP02
Years of Operation	2002--	2002--	2002--2008	2002--
Thermometer/ Hygrometer	HMP45C (3) (CSI)	HMP45C (CSI)	HMP45C (CSI)	HMP45C (CSI)
Anemometer	05 103-10RE (RMY)	05 103-10RE (RMY)	05 103-10RE (RMY)	05 103-10RE (RMY)
Pyranometer	PAR-LITE (3) (KZ)	PAR-LITE (2) (KZ)	PAR-LITE (2) (KZ)	PAR-LITE (2) (KZ)
Net Radiometer	NR-LITE (KZ)	NR-LITE (KZ)	NR-LITE (KZ)	NR-LITE (KZ)
Component Radiometer	CRN1 (CSI)	----	----	----
Precipitation Gauge	52202 (RMY, 2002-2007) CS700 (CSI, 2008--)	----	----	TE525 (CSI, 2002-2008) 52202 (RMY, 2008--)
Barometer	61205V (RMY)	61205V (RMY, 2008--)	----	61205V (RMY, 2008--)
Soil Thermistor	107B (15) (CSI)	107B (12) (CSI)	107B (12) (CSI)	107B (12) (CSI)
Soil Moisture Reflectometer	CSI-615/616 (16) (CSI)	CSI-615/616 (14) (CSI)	CSI-615/616 (10) (CSI)	CSI-615/616 (10) (CSI)
Soil Heat Flux Plate	HFT3 (2) (CSI)	HFT3 (2) (CSI)	HFT3 (2) (CSI)	HFT3 (2) (CSI)
Sonic Snow Depth Ranger	SR50 (CSI)	----	----	----
Data Logging Device	CR23x (2), (CSI)	CR10x (CSI, 2002-2007) CR1000 (CSI, 2008--)	CR10x (CSI)	CR10x (CSI, 2002-2009) CR1000 (CSI, 2009--)
Data Format	CSV, no headers	CSV w/o headers (2002-2007) w/ headers (2008--)	CSV w/o headers	CSV w/o headers (2002-2007) w/ headers (2008--)
Data Backup Device	Desktop PC, USB	SM16M (CSI, 2002-2007) Desktop PC, External Hard Disk (EHD, 2008--)	SM16M (CSI)	SM16M (CSI, 2002-2007) Desktop PC, External Hard Disk (EHD, 2008--)
Data Download Method	USB or USB → Laptop	SC532A (CSI) → Laptop (2002-2007) EHD → Laptop (2008--)	SC532A (CSI) → Laptop	SC532A (CSI) → Laptop (2002-2007) EHD → Laptop (2008--)

Abbreviations:
CSI: Campbell Scientific Inc., Edmonton, AB, CA
RMY: RM Young Co, Michigan, USA
LCI: LI-COR Biosciences, Lincoln, NE, USA
KZ: Kipp and Zonen, Ltd., Delft, NL

Table 2.2: *Ancillary measurement system details at TPFS*

	TP39	TP74	TPPPT
Sap Velocity Measurement Systems			
Years of Operation	(2008--)	(2009--)	----
Thermal Dissipation Probe	TDP-30 (22) (Dynamax, Houston, Texas, USA)	TDP-30 (12) (Dynamax, Houston, Texas, USA)	----
Soil Thermometer	107B (10) (CSI)	----	----
Soil Moisture Reflectometer	CSI-615/616 (6) (CSI)	CSI-615/616 (6) (CSI)	----
Data Logging Device	CR10x (CSI)	CR1000 (CSI)	----
Data Format	CSV, w/o headers	CSV, w/ headers	----
Data Backup Device	SM16M (CSI)	SM16M (CSI)	----
Data Download Method	SC532A (CSI) → Laptop	SC532A (CSI) → Laptop	----
Water Table Measurement Systems			
Years of Operation	2009--	----	----
Water Table Level Pressure Sensor	PLS (OTT, Kempten, DE)	----	----
Data Logging Device	CR10x (CSI)	----	----
Data Format	CSV, w/o headers	----	----
Data Backup Device	SM16M (CSI)	----	----
Data Download Method	SC532A (CSI) → Laptop	----	----
Ground-based Precipitation Measurement Systems			
Years of Operation	----	----	2007--
Anemometer	----	----	05 103-10RE (RMY)
Tipping Bucket Rain Guage	----	----	TE525 (CSI)
Accumulation Precipitation Rain Guage	----	----	T200B (Geonor Inc., Milford, PA, USA)
Data Logging Device	----	----	CR10x (CSI)
Data Format	----	----	CSV, w/o headers
Data Backup Device	----	----	SM16M (CSI)
Data Download Method	----	----	SC532A (CSI) → Laptop

Abbreviations:
CSI: Campbell Scientific Inc., Edmonton, AB, CA
RMY: RM Young Co, Michigan, USA
LCI: LI-COR Biosciences, Lincoln, NE, USA
KZ: Kipp and Zonen, Ltd., Delft, NL

alongside supplementary meteorological data at half-hourly intervals by a standalone data logger (CR10x).

A CPEC system has been operated continuously at TP39 since 2002 (Arain and Restrepo-Coupe, 2005). The system consists of a CSAT-3 sonic anemometer (model CSAT-3, Campbell Scientific Inc., Edmonton, AB, CA) and infrared gas analyzer (IRGA, model LI-7000, LI-COR Biosciences, Lincoln, NE, USA), operated and logged at high frequencies (20 Hz) on a desktop PC, using custom software created by the Biometeorology & Soil Physics Group at the University of British Columbia (British Columbia, CA). Similar systems have been continuously operated at TP74 (CSAT-3; model LI-6262, LI-COR) and TP02 (CSAT-3; LI-6262) since 2008. During years 2002-2008, a single, roving OPEC system was rotated amongst TP74, TP89 and TP02 at two-week to one month intervals. The sonic anemometer (CSAT-3) and IRGA (model LI-7500, LI-COR) were operated and logged by a standalone data logger (model CR3000, Campbell Scientific Inc.), which recorded EC variables at high-frequency (20 Hz) along with 10-minute-averaged meteorological variables and calculated fluxes (Table 2.3).

An automated soil chamber system (ACS) has been operated at TP39 since 2008 to measure forest floor CO_2 efflux and its components. The ACS consists of eight spatially distributed, non-steady state chambers that close sequentially to draw head space air into an infrared gas analyzer (model LI-840, LI-COR), which is controlled and logged at 1 Hz by custom software on a mini-desktop PC (Table 2.3). The hardware and software of the ACS system were designed by the Biometeorology & Soil Physics Group at the University of British Columbia (Vancouver, BC, Canada), and are described in detail by Drewitt et al. (2002) and Jassal et al. (2005).

Each of the presently-operated main sites (TP39, TP74 and TP02) are provided Internet access via a mobile broadband router (model MBR 1210, Netgear, San Jose, CA, USA), to which each collection PC is connected. To verify the proper operation of CPEC, ACS and primary meteorological systems, sites are accessed daily using remote access software (LogMeIn, Woburn, MA, USA), where output variables from each system are checked and noted in an Internet-based, operator-edited log file.

Table 2.3: *Eddy-covariance and automated chamber system details at TPFS*

	TP39	TP74	TP89	TP02
EC Systems				
Years of Operation	2002--	2002--	2002--2007	2002--
EC Type	CPEC	OPEC (2002-2007) CPEC (2008--)	OPEC	OPEC (2002-2008) CPEC (2008--)
Sonic Anemometer	CSAT-3 (CSI)	CSAT-3 (CSI)	CSAT-3 (CSI)	CSAT-3 (CSI)
Infrared Gas Analyzer (IRGA)	LI-7000 (LCI)	LI-7500 (LCI, 2002-2008) LI-6262 (LCI, 2008--)	LI-7500 (LCI)	LI-7500 (LCI, 2002-2008) LI-7000 (LCI, 2008--)
CO_2 Profile Gas Analyzer	LI-800 (LCI) LI-820 (LCI)	LI-820 (LCI)	LI-820 (LCI)	----
Data Logging Device	Desktop PC	CR3000 (CSI, 2002-2007) Mini Desktop PC (2008—)	CR3000 (CSI)	CR3000 (CSI, 2002-2008) Mini Desktop PC (2008--)
Data Format	Half-hourly (HH) Binary Files	Single Binary + CSV w/ headers (2002-2007) HH Binary (2008--)	Single Binary + CSV w/ headers	Single Binary + CSV w/ headers (2002-2008) HH Binary (2008--)
Data Backup Device	USB Flash Drive	Compact Flash (CF) Media (2002-2007) External Hard Disk (EHD, 2008--)	CF media	Compact Flash (CF) Media (2002-2008) EHD 2008--)
Data Download Method	USB or USB → Laptop	CF → Laptop (2002-2007) EHD → Laptop (2008--)	CF → Laptop	CF → Laptop (2002-2008) EHD → Laptop (2008--)
ACS Systems				
Years of Operation	(2008--)	----	----	----
Gas Analyzer	LI-840 (LCI)	----	----	----
Data Logging Device	Mini Desktop PC	----	----	----
Data Format	HH Binary	----	----	----
Data Backup Device	USB Flash Drive	----	----	----
Data Download Method	USB or USB → Laptop	----	----	----

Abbreviations:
CSI: Campbell Scientific Inc., Edmonton, AB, CA
LCI: LI-COR Biosciences, Lincoln, NE, USA

Field data management

The nature and operational considerations of measured data across the TPFS is diverse, as systems vary in terms of their logging device, data format, transfer media type and downloading requirements (Tables 2.1 through 2.3). An operational goal of the DWMS is to ensure that despite these differences, data on all logging devices are consistently backed-up and are transferred from the devices to the DWMS central file system (CFS) consistently, reliably and effectively.

Tables 2.1 through 2.3 provide information regarding the data formats, as well as collection and backup devices used for each field measurement system. CPEC and ACS systems log data to collection PCs as similarly-formatted binary files. Specifically for the CPEC system, separate raw data files are created for each measurement device (e.g. CSAT, IRGA, thermocouple). Individual files are created for each half-hour of the day, and all files from a given day are placed into a single directory. Newly-created raw data directories are backed-up to external media (USB or hard disk drives) on a consistent basis, as described later in this section. During site visits, data on the external media is copied to a reserved directory on a field laptop computer to be transferred to the DWMS central file system.

Data that is collected by the OPEC system data logger is simultaneously written to a removable Compact Flash (CF) card. Due to limited data logger storage capacity, a limited amount of recent high-frequency binary data is retained on the device, while a single, accreting file is stored on the CF media. Much smaller ASCII files of data logger-calculated 10-minute averages are stored on both media. Data is transferred from the system via the removable CF card to an appropriate field laptop directory. The CF card is then formatted and replaced into the data logger.

Primary meteorological system data loggers at each site are connected to CPEC system PCs. Data is automatically downloaded from the data loggers at hourly intervals, and is appended to comma-separated ASCII files on the PC hard disk. Data logger files are automatically backed up to attached external media (see below), which are used to transfer files to the field laptop computer. All ancillary meteorological data is logged onto standalone data loggers. Comma-separated ASCII data logger files are downloaded to a field laptop through serial connection between the devices (model SC532A, CSI).

All data downloaded to the field laptop hard drive is subsequently uploaded to the appropriate destination on the CFS, which allows raw data to be processed and continued through the data life cycle.

Data management in the DWMS central file system

All measured data is collected to a linux-based central file system (CFS) that resides on hard disks in an accessible and centralized PC within the Hydrometeorology and Climatology Research Group laboratories at McMaster University. All subsequent data life cycle operations are initiated within the CFS, which provides data organization, documentation, processing, sharing and backup activities for all DWMS data. The CFS file structure reflects a desire to organize data in a consistent and straightforward manner, using common directory rules to ensure that data is accessible to, and readily located by DWMS operators and processing programs alike. At its primary level, the CFS file structure organizes data according to its role in the data life cycle; while at subsequent levels, data is structured according to common characteristics such as measurement system, site and year of measurement, and data quality level. Figure 2.3 provides a summarized view of the CFS file structure, where selected locations are shown in a directory tree format to allow elaboration of its eight primary level directories and the organization within each.

The **/DUMP_Data** primary directory serves as a repository for all field-downloaded raw data that will be extracted and incorporated into appropriate permanent locations on the CFS. Raw data is uploaded here from the field laptop, and placed in an appropriate sub-directory; primary and ancillary meteorological data are placed inside the /Met sub-directory, while CPEC, OPEC and ACS data are placed into the sub-directory that corresponds to collection site (e.g. /TP39, /TP74, etc.). Automated processing software (see section 2.4.3.2) extracts relevant raw data from these sub-directories, and moves the original raw data files to archival and backup locations when finished.

The **/To_Burn** primary directory is a destination for post-extraction raw data, as automated software moves data into site-specific sub-directories within. Data in each sub-directory is twice copied by the operator to optical media (e.g. CD, DVD, Blu-ray Disc) for long-term data backup. After copying, raw data files in this directory are removed from the CFS.

```
/1/fielddata
  -/DUMP_Data              ←  Repository for field data
    -/Met                      ←  All meteorological (primary and ancillary) data placed
    -/<site>                       in /Met.  Other data is grouped by site
  -/To_Burn                ←  Temporary location for data to be burned to optical disc
    -/<site>                   ←  Organized by site
  -/Raw_Data_Archive       ←  Location for 1-year data archival of flux and ACS data
    -/OPEC                     ←  all OPEC data in /OPEC
    -/<site>                   ←  CPEC, ACS data grouped by site
  -/SiteData               ←  Raw and organized OPEC, CPEC and ACS data
    -/<site>                   ←  Data grouped by site
      -/MET-DATA                 ←  CPEC and ACS data
        -/data                     ←  Raw, high-frequency binary data files
        -/hhour                    ←  Calculated half-hourly averaged files
        -/annual                   ←  Organized annual half-hourly master files
      -/OPEC                     ←  OPEC data
        -/HF_data                  ←  Raw, high-frequency binary data files
        -/hhour_data               ←  Raw, 10-minute average ASCII data files
  -/EdiRe                  ←  Directory for EdiRe-based OPEC flux calculations
    -/CardConvert          ←  Converted data files
    -/OPEC_Calculations    ←  EdiRe-calculated half-hourly fluxes
  -/Matlab                 ←  Location for processing functions and all processed data
    -/Scripts                  ←  Master directory for processing functions (BACON)
      -/mcm_data_mgmt            ←  Data management functions
      -/Flux                     ←  Flux calculation functions
      -/Met                      ←  Meteorological calculation functions
      -/Gapfilling               ←  Gap-filling functions
    -/Data                   ←  Processed data
      -/Met                      ←  Meteorological data
        -/Docs                     ←  Processing documents, .ini files
        -/Raw1/<site>              ←  Raw meteorological data files
        -/Organized2/<site>        ←  Organized annual meteorological master data files
        -/Cleaned3/<site>          ←  Threshold-cleaned annual met master data files
        -/Final_Cleaned/<site>     ←  Quality-assured, cleaned met master data files
        -/Final_Filled/<site>      ←  Filled meteorological master data files
        -/Calculated4/<site>       ←  Calculated meteorological variables
      -/Flux                     ←  Flux data
        -/<EC_type>                ←  Organized by CPEC or OPEC data
          -/Docs                     ←  Processing documents, .ini files
          -/<site>                   ←  Organized by site
            -/Cleaned                  ←  Threshold-cleaned annual flux master data
            -/Final_Cleaned            ←  Quality-assured, cleaned met master data
            -/Final_Filled             ←  Gap-filled flux data products
            -/Final_Calculated         ←  Calculated flux data products
      -/Master_Files             ←  Location of all-years-compiled site master data files
        -/Docs                     ←  Data compilation documents, .ini files
        -/<site>                   ←  Data products, grouped by site
        -/Latest_Data-To_Post      ←  Master files to share with internal end-users
    -/Figs                     ←  Location for processing figure output
    -/ubc_PC_setup             ←  .ini files for CPEC, ACS flux calculation program
  -/Documentation          ←  Metadata, protocols, manuals, ancillary site information
  -/Programs               ←  Field data logger programs, additional software
```

Figure 2.3: *DWMS centralized file system structure and contents.*
Entries enclosed in '< >' are placeholders for each site (e.g. TP39, TP74, etc.) or eddy
covariance system (CPEC, OPEC).

Post-extraction raw data from CPEC, OPEC and ACS systems are also copied to the **/Raw_Data_Archive** primary directory, where data is organized and subsequently retained in archives for one year. This archive allows the operator to promptly reprocess raw data if a processing error is discovered up to a year after extraction. To avoid excessive hard disk use, raw CPEC, OPEC and ACS data files are permanently deleted from the CFS one year after extraction. Conversely, the much smaller raw meteorological system data files are archived permanently on the CFS following extraction.

Extracted CPEC, OPEC and ACS data are managed within the **/SiteData** primary directory. Data is grouped first by site, and then into structures required by their respective flux calculation programs. CPEC and ACS data are placed into similar directory structures, where extracted data is stored alongside sub-directories that contain higher-processed calculated and organized flux data files. OPEC data is organized in a unique sub-directory that is required by OPEC flux calculation software. Calculated OPEC fluxes are outputted to the **/EdiRe** primary directory, from where they are incorporated into subsequent data products.

The **/Matlab** primary directory contains all data processing software programs, data products (intermediate and final), and figures that are produced by software operation. Processing functions are organized within the /Scripts sub-directory by their operational purpose. The /Data sub-directory is organized by data product type, as intermediate meteorological data is separated from intermediate flux data and end-user master files. The /Data/Met and /Data/Flux sub-directories are similarly structured, as data is organized according to its data quality level and the site of collection, while /Data/Flux is also organized by EC system type. Quality-controlled, calculated and filled data from a given site are compiled into master files and output for end-user sharing in the /Data/Master_Files sub-directory. Files in this sub-directory are organized by site, with a separate folder (/Latest_Data-To_Post), which contains data products to be made available to local and external end-users. The /Config subdirectory contains the operator-editable configuration files that provide situation-specific instructions to processing operations. Configuration files are organized into meteorological and flux-specific subdirectories, which specify such details as raw and output data formats, data cleaning thresholds, and site- and year-specific processing parameters. Figures produced during data inspection, cleaning and filling operations are saved to appropriate locations in the /Figs sub-directory.

Metadata, protocols, data processing logs, manuals and ancillary information for DWMS and field site operations are maintained within the **/Documentation** primary directory (see section 2.4.5), while data

logger programs and ancillary software installation files are located in the **/Programs** primary directory. Both primary directories serve as centralized locations for operational files that are created, edited and used collaboratively by DWMS operators.

2.4.3.2 Data processing & the *BACON* software package

All data processing and management operations are conducted using MATLAB software (The Mathworks Inc., Natick, MA) and custom-designed MATLAB program packages. The Biometeorological Analysis, Collection, and Organizational Node (**BACON**), is a collection of over 250 custom-made MATLAB functions and scripts (representing over 30000 lines of code), which have been implemented on the DWMS central file system to perform all necessary data processing operations.

An accompanying graphical user interface (GUI, see section 2.4.4.1) provides an accessible front-end to BACON operations (Figure 2.4). Thus, all primary functions can be executed either graphically through the GUI, or textually through the MATLAB command line. BACON is designed to process data from all measurement and logging systems (meteorological, CPEC, OPEC, ACS, etc.), and standardize the processing steps for each type of data into a clear and consistent methodological framework. The BACON package uses separate programs to carry out processes on each data type, however, these programs all share parallel operational methodologies and general file structures.

The details of data processing implemented by BACON to produce quality-controlled, sharable data sets from uploaded raw data is represented schematically in Figures 2.5, 2.6 and 2.7 for meteorological, CPEC and OPEC systems, respectively. Data operations, data flow, and input/output directory structures are very similar between systems, aside from small differences that accommodate discrepancies in raw data formats and processing requirements. In each system-specific implementation, data is sequentially processed from lower- to higher-quality levels; quality-assured, filled and calculated data products are ultimately produced, and are compiled into comprehensive master files, from which sharable data sets are created for end-users. System-specific BACON operations accommodate data measured from any TPFS field site and are expandable to incorporate new measurement sites, provided that appropriate initialization files are created (see section 2.4.4.4). The following sections introduce the general steps that are used for all BACON data processing, and details the system-specific implementation of these processing steps.
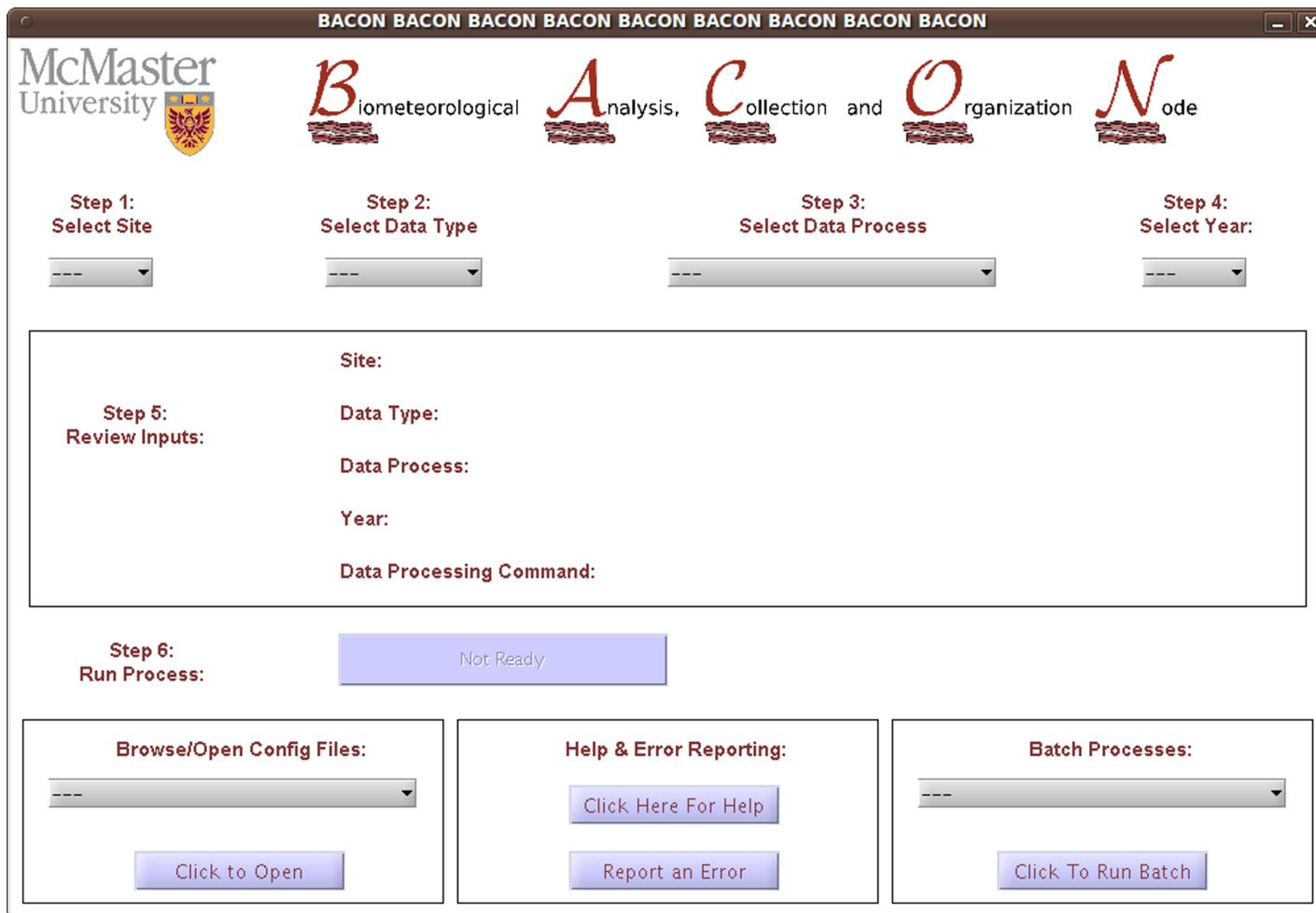
Figure 2.4: *The BACON GUI in its initial state.*
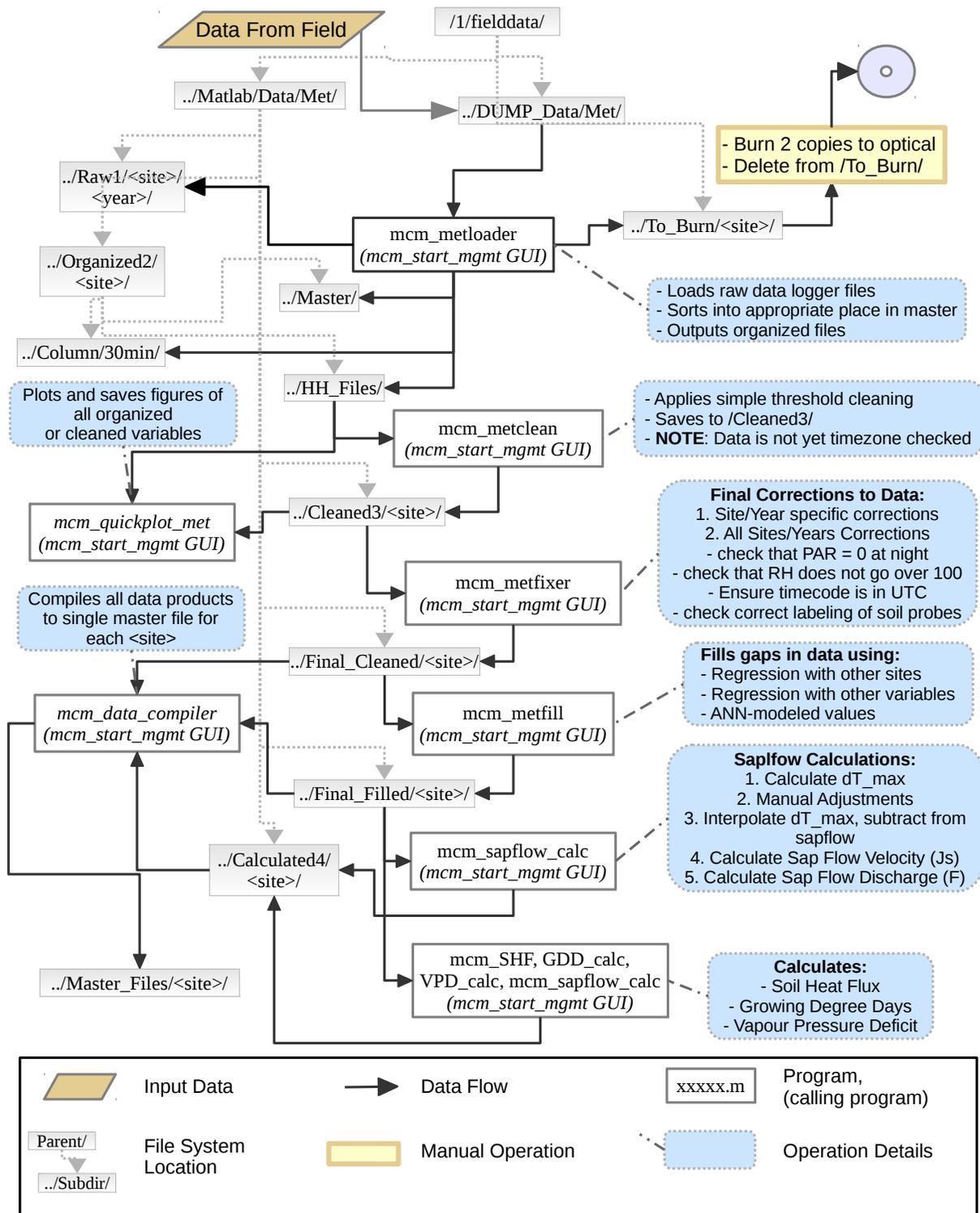
Figure 2.5: *A schematic diagram of data flow and BACON operations for meteorological data processing*
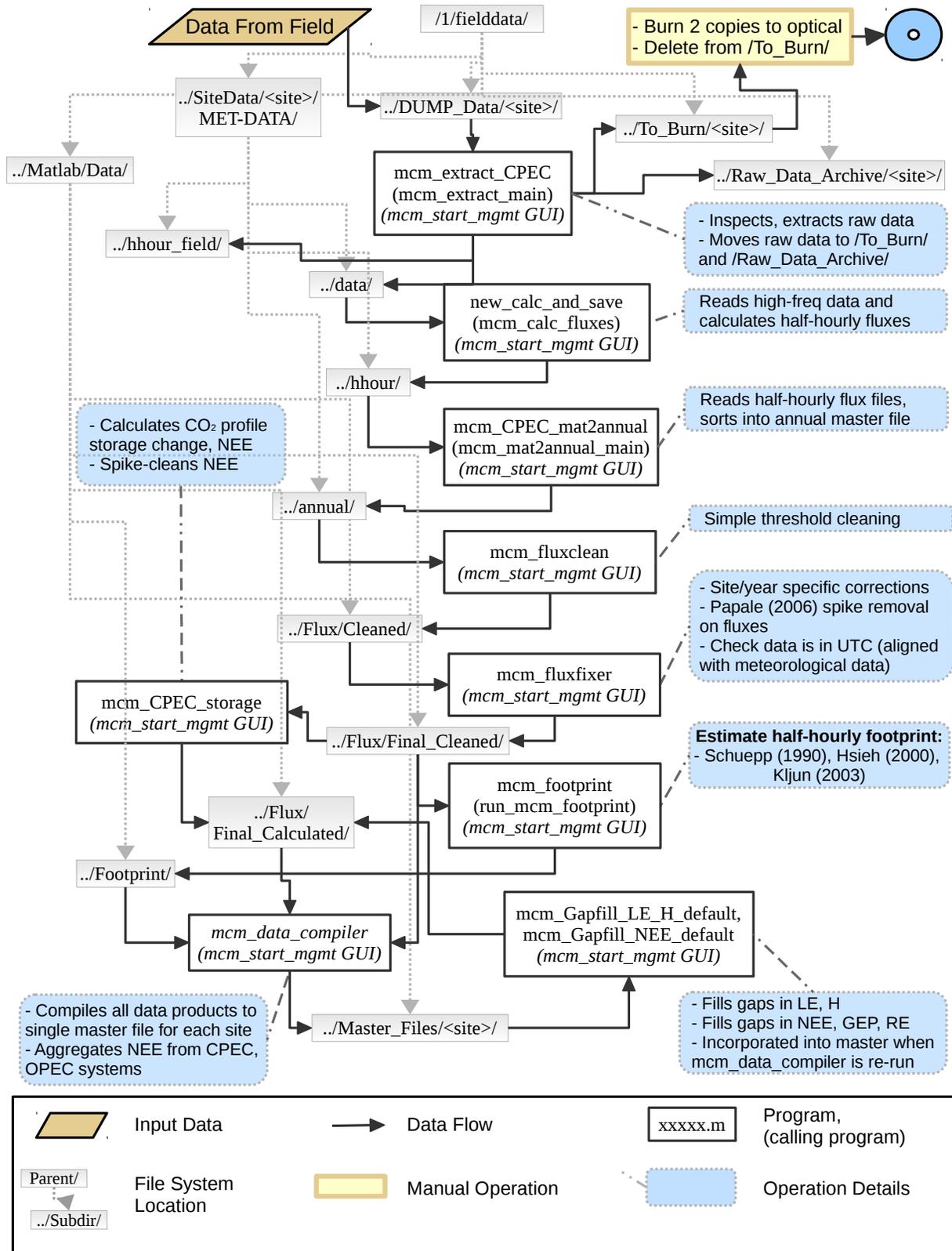
Figure 2.6: *Schematic diagram of data flow and BACON operations for CPEC data processing*
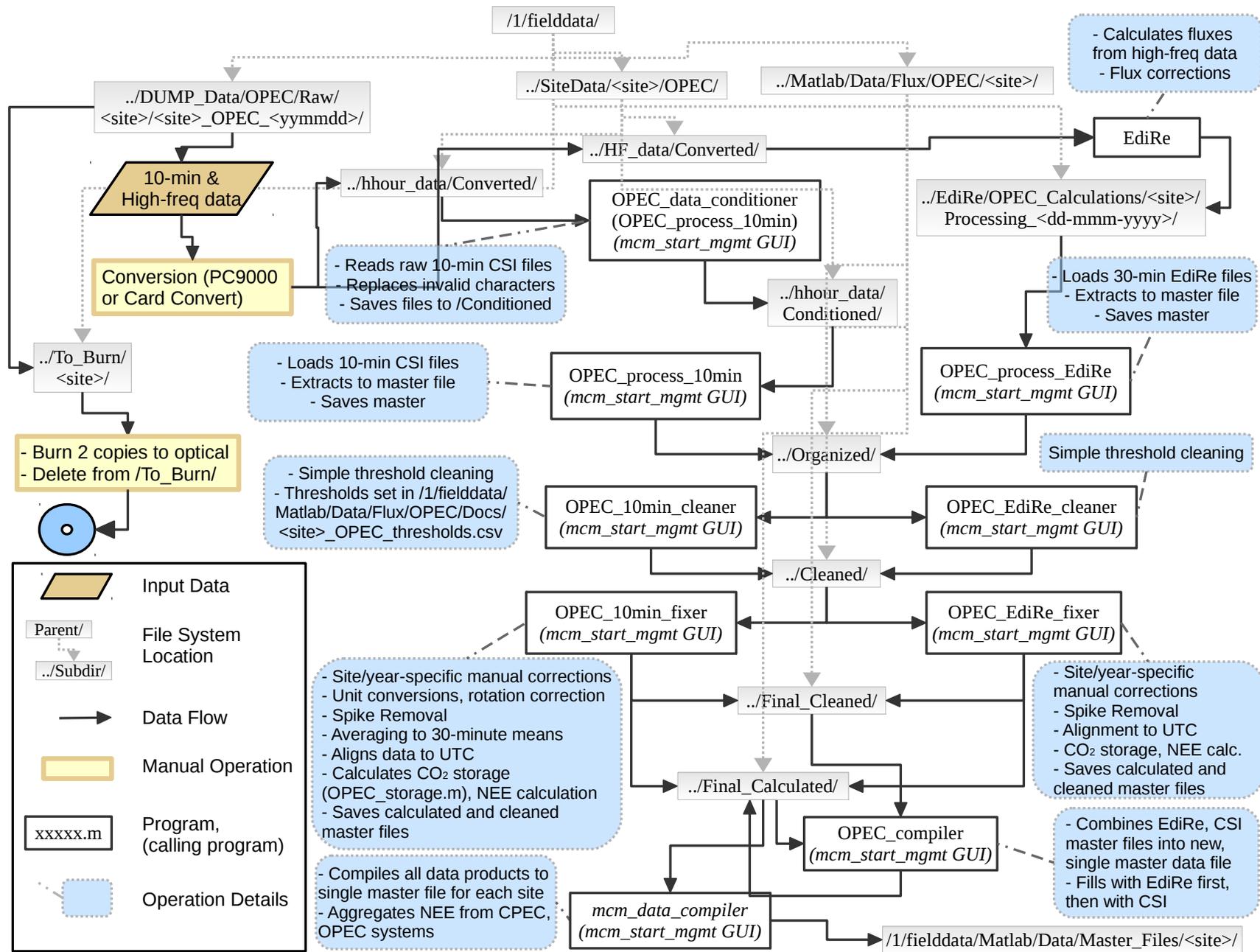
Figure 2.7: *Schematic diagram of data flow and BACON operations for OPEC data processing*

Reading raw files and incorporating new data into organized, master files or directories

In the first stage of data processing, BACON programs open raw data files that have been uploaded to the CFS (following section 2.4.3.1), compare the data with content existing in master files or master directories, and extract the desirable data to the organized master files or directories.

When organizing raw meteorological data, the objective is to read raw data logger files, identify the location of variables within the raw files, and then extract all previously unextracted data to the appropriate location in an organized master file, which collects all data from a given system over a predetermined period of time (e.g. year). Unfilled, organized master files are created at the beginning of each year, as the processing program (mcm_metloader.m, Figure 2.5) creates a matrix of Not-a-Number (NaN) values. The number of rows created in the matrix corresponds to the expected number of observations for the year, based on the time interval of variable measurement and the year of consideration. For example, 17520 rows are required to contain all possible half-hourly measurements collected in a non-leap year. The number of columns created exceeds the existing number of variables by 10 to 20%, to accommodate instrument addition.

A variable's destination column in the master file is determined by an operator-editable master template flat file. New variables may be appended to the template file, but existing variables may not be moved, in order to ensure variable consistency for columns of the master matrix within and between years. Information regarding the column position of variables in the raw data logger files are relayed by operator-editable header documents (for all data logger types; see Appendix A for example). For CR1000, CR3000 data loggers, header information is provided by the file's own header. Variable names in the master template file are set to match those in the header metadata documents, which provides a means to inform the program of the corresponding columns in the raw and master files, and accommodates changes to variable locations in the raw data logger files. Corresponding row locations in the raw and master files are determined by finding equivalent timestamps (in the form of a date vector) between both files. By ensuring that only data with an appropriate destination row and column are inserted into the master file, problems that may be encountered during manual matrix manipulation (e.g. incorrect row and column attribution) can be effectively eliminated. Command-line output informs the operator of its progress and relays information regarding successes or errors that are encountered. Specifically, a warning message is issued when a variable listed in the master template file is not updated when processing the raw data logger file.

Organizing high-frequency data files from ACS and CPEC systems requires that each raw high-frequency binary data file (created each half-hour for each EC instrument) be inspected for errors and duplicates, and then, if permitted by the user, copied into the appropriate master directory in the /SiteData primary directory. Raw data files are first checked by the data processing program (mcm_extract_CPEC.m, Figure 2.6), to ensure that empty or corrupted files and directories are flagged and not copied to the master data directories. A list of problematic files or directories are communicated to the user via warning messages on the command window interface and recorded to processing log files. The program also checks and flags redundant data files, avoiding unnecessary data copying. Desirable and acceptable binary files are copied to the correct location in the master directory, according to the date, site and system of measurement.

At the end of this operation, a separate function scans the appropriate master data directory to assess the completeness of raw binary files for a specified year. This operation creates a log file in spreadsheet format, detailing the types and numbers of data files per day in the in the master directories. Using this document, DWMS operators can inspect data completeness and manually retrieve any missing data files or directories from the field collection PCs. After organized binary files are used to calculate fluxes (section 2.4.3.2), output data files for each half-hour are read by a separate organization program (mcm_CPEC_mat2annual.m, Figure 2.6), and organized into half-hourly master files for each site and year.

Organizing raw OPEC data requires separate considerations for each of its two types of data outputs: binary data logger files with 10-minute flux and meteorological variable averages, and binary files of high-frequency (20 Hz) gas concentration, temperature and wind velocity measurements. After being manually converted from a binary to ASCII format, the 10-minute-average data logger files are organized into a master file using a processing program (mcm_process_10min.m, Figure 2.7), which operates similarly to the meteorological extraction program described above. Header files and time vectors are used to extract desired data from raw files, and insert them into an organized master file. Raw OPEC high-frequency data files are manually converted to a binary format that is readable by the flux calculation program, and placed into an appropriate source directory prior to calculation. Following flux calculation (outlined in section 2.4.3.2), an organization program (OPEC_process_EdiRe, Figure 2.7) reads files containing half-hourly averages, and extracts data into organized master files in the same manner as used for 10-minute files.

After data is organized into master files, the processing programs copy the raw data files and directories

to appropriate locations in the /Raw_Data_Archive primary directory, which stores raw data files on the DWMS file system for a year, as described in section 2.4.3.4. Raw data files and directories are then moved from the /DUMP_Data primary directory to the /To_Burn primary directory, where they are written to optical media and deleted from the DWMS file system (section 2.4.3.4).

Calculating fluxes (for CPEC, OPEC systems; ACS)

Flux calculations for CPEC and ACS systems are performed using MATLAB packages developed by the Biometeorology & Soil Physics Group at the University of British Columbia, which are evoked through BACON. The flux calculation programs (new_calc_and_save.m, Figure 2.6) read high-frequency data from the organized master directories for a period specified by the operator. All necessary EC- and chamber-specific processing is completed, and high-frequency data correction and quality control operations are implemented (following Lee et al., 2004) to produce output of half-hourly flux and environmental averages. The specific operational parameters of the flux calculation programs are set in initialization files on the DWMS file system, which may be customized by the operator (see section 2.4.5).

OPEC fluxes are calculated by the EdiRe software package (Mauder et al., 2008), developed by the Institute of Atmospheric and Environmental Science at the University of Edinburgh, UK. Organized binary files are processed and calculated in batch mode to produce output data files with half-hourly averages of flux and ancillary data. EdiRe batch processing is initiated by the user outside of BACON, after which, its products are incorporated into organized master files, as described in section 2.4.3.2.

Conducting interactive, threshold-driven data checking and cleaning

Organized master files for all data types are interactively screened for obvious outliers/erroneous data by linear threshold cleaning. Upper and lower acceptable threshold values are applied to each variable, and values above (below) the high (low) threshold are removed from the time series and replaced by NaN (data gap). Thresholds are customized to each variable, and each set of thresholds are saved in separate matrices for each year. Cleaning programs (mcm_metclean.m, Figure 2.5; mcm_fluxclean.m, Figure 2.6; OPEC_10min_cleaner and OPEC_EdiRe_cleaner.m, Figure 2.7) load the threshold matrix if it exists, otherwise prompting the operator to load and edit the previous year's threshold values, or manually enter new thresholds for each variable. The operator uses keyboard hot keys to scroll through plots of each variable and its existing thresholds, with the option of accepting or adjusting these high and low cut-off values. This operation provides a means of removing obvious erroneous data in a quick and consistent manner. Once the operator is satisfied with cleaning operations, the threshold

matrix is saved to the /Matlab/Config directory of the CFS, where it can be applied or edited during successive program runs. Cleaned variables are structured into a single, annual matrix, and are packaged into a MATLAB structure array with a data header list that contains variable and matrix location information. This master file is then saved on the CFS to be used for further quality-control operations.

Carrying out detailed manual screening, calibration, correction, spike-detection and final inspection

After threshold cleaning, data is screened and corrected in a detailed manner, to create quality-controlled data sets. Detailed screening programs (mcm_metfixer.m, Figure 2.5; mcm_fluxfixer.m, Figure 2.6; OPEC_10min_fixer.m and OPEC_EdiRe_fixer.m, Figure 2.7) allow the operator to individually scroll through and inspect time series for all variables, noting data points that are questionable, erroneous, or in need of correction/calibration. Errors in the data are corrected manually within the scripting of the BACON function, which provides complete documentation of data changes and allows data fixes to be revised or reversed in the future. In addition, system-specific quality control and quality assurance tests are incorporated into screening programs and are automatically performed on the data, informing the operator of such potential issues as instrument calibration problems, instrument errors, and the presence of error/caution flags. In particular, primary meteorological systems are tested and corrected for any user-, program- or instrument-caused data logger time shifts away from the desired UTC time code, by comparing measured shortwave radiation time series with predicted sun times (sunrise, solar noon and sunset) during cloudless days. In addition, CPEC and OPEC system timing at a given site is checked against and aligned to its primary meteorological system by cross-correlation analysis between the wind direction time series for both systems. These operations ensure that meteorological and EC systems are all properly aligned to the UTC time code.

Time series that are prone to data spikes and sharp magnitude changes (such as EC-measured fluxes and sapflow velocity measurements) are subjected to a rigorous, automated spike-filtering algorithm, adapted from Papale et al., (2006). This algorithm is used in place of an operator-assessed outlier removal approach, with the aim of reducing subjectivity and inconsistencies between operators when quality-controlling data.

After all data corrections are made, the detailed screening programs allow the operator to review their changes by interactively browsing overlaid plots of corrected and original (uncorrected) time series.

The master files of screened data created in this stage represent the quality-assured data to be shared with local and external end-users, and is the data to be used in data filling and calculation operations.

Completing necessary data calculations and producing secondary data products

In both meteorological and EC applications, a number of secondary data products are calculated from quality-controlled observational data. Performing these operations on quality-assured data ensures that calculations are carried out on screened, high-quality data, and makes possible the incorporation of properly time-aligned data across different master files (e.g. CPEC, OPEC and meteorological data). As data locations in the CFS are static, programs provided in BACON allow the operator to perform these calculations through simple command-line input or through the GUI, requiring only site and year inputs to run. Any time series-specific settings are passed to processing programs by configuration files located in the /Matlab/Config directory, and a single, operator-updated parameter file (params.m) supplies the calculation functions with operation-specific parameters for each site and year.

Calculating Net Ecosystem Exchange (NEE) for sites requires adding EC-measured, non-spike-filtered half-hourly CO_2 fluxes (F_c) to calculated CO_2 storage changes in the air column beneath the EC sensor array (ΔS_c, Finnigan, 2006). CO_2 storage changes are calculated within BACON (mcm_CPEC_storage, Figure 2.6; mcm_OPEC_storage, Figure 2.7) by vertically integrating changes in EC-measured CO_2 concentration above the canopy and meteorological measurements of the CO_2 concentration profile beneath it. Following calculation, NEE is spike-filtered according to methods defined above, and NEE and ΔS data products are saved to a directory containing derived data products.

Estimating EC flux footprint also requires the integration of EC-measured fluxes with meteorological variables. With operator initialization, the footprint calculation program (mcm_footprint, Figure 2.6) loads the required variables and executes a variety of flux footprint estimation schemes (Hsieh et al., 2000; Kljun et al., 2004; Kormann and Meixner, 2001; Schuepp et al., 1990) providing an assortment of model predictions at a half-hourly timescale.

Meteorological data calculations may also be implemented through BACON; programs are provided for soil heat flux (mcm_SHF), sapflow measurement calculations (mcm_sapflow_calc), vapour pressure deficit (VPD_calc.m) and growing-degree days (GDD_calc.m, Figure 2.5).

Filling gaps in variables, where continuous time series are required

A number of different methods are employed by BACON to fill gaps in time series. that are required to

be continuous. For meteorological variables that are highly correlated between sites for simultaneous measurements (Ta, PPFD, SW_{down}, RH, WS, atmospheric pressure and precipitation), gaps in data from the target site are filled by linear regression-modeled values from other (source) sites. The source site used to fill gaps in target data is determined by the coefficient of determination (R^2) for the given variable between sites for the year of interest. Data gaps are filled preferentially from the source site with highest R^2, where data is available. Any remaining gaps are filled from sites with next highest R^2, where data exists.

When filling gaps in meteorological variables that show poorer correlation across sites, a number of different approaches are used. Gaps in Ts measured at 2 cm and 5 cm depth are filled using a linear regression model based on air temperature data. Gaps in top-30 cm averaged VWC (VWC_{30}) are filled in a similar manner as methods described above for highly-correlated variables, except that an attempt is first made to fill data gaps for a given soil pit by a linear regression model with other soil pits from the same site. Rn at each site is filled using an artificial neural network (ANN), created with the MATLAB neural network toolbox and conditioned with T_a, PPFD, WS and RH inputs. Since performance of soil moisture and net radiation filling activities is often marginal, end-users are encouraged to use non-gap-filled, QA products where possible. All meteorological data filling operations are executed simultaneously for all sites through a single program (mcm_metfill.m, Figure 2.5); this ensures that filling operations are rerun for all sites when any revisions are made to QA data, as it may affect filled data for all sites.

Ecosystem flux gap-filling is implemented by a pair of separate BACON programs: one to fill sensible heat (H) and latent heat (LE) fluxes (mcm_Gapfill_LE_H_default.m, Figure 2.6), and one to fill NEE and its components (mcm_Gapfill_NEE_default, Figure 2.6). As a vast array of data filtering and modeling options are available for gap-filling ecosystem fluxes (see Falge et al., 2001; Moffat et al., 2007; Desai et al., 2008), the BACON programs allow the operator to specify one or a number of combinations to use in gap-filling application. An editable initialization file permits the operator to perform numerous gap-filling operations for each site, with customized settings for: years to fill, u_* threshold estimation model, gap-filling model, error estimation model and the minimum acceptable within-fetch flux proportion criteria used in footprint filtering. Separate BACON programs provide operators options for u_* threshold determination methods (including those presented in (Barford et al., 2001; Barr et al., 2013; Gu et al., 2005; Reichstein et al., 2005), as well as footprint estimation method (see above).

Fluxes of H and LE are filled using ANNs created with the MATLAB neural network toolbox. LE is modeled and filled using Rn, Ts_{5cm}, WS, VWC_{30cm} and vapour pressure deficit (VPD) as training and modeling variables. In the same manner, H is modeled and filled using PPFD, R_n, filled LE and T_a as training and modeling variables. Any remaining gaps in these variables are subsequently filled using either windowed linear regression or windowed mean diurnal variation approaches, following Amiro et al. (2006).

NEE and its components (Gross Ecosystem Productivity, GEP; Ecosystem Respiration, RE) may be filled using a number of different gap-filling models that have been created and implemented in BACON. These models include non-linear regression approaches described by Barr et al. (2004), Richardson et al. (2007), Lasslop et al.(2010) and Richardson and Hollinger, (2007), as well as the marginal distribution sampling approach of Reichstein et al. (2005). Also included is a custom-designed ANN-based method, which models RE using $T_{s,5cm}$ and VWC_{30cm} as training and modeling variables, while using inputs of PPFD, $T_{s,5cm}$, VWC_{30}, T_a, VPD and GDD to model NEE. Gap-filling models, u_* thresholds and footprint filtering methods may be combined by the operator as desired. Results of all gap-filling model runs are saved to a common directory in the DWMS file system, and may be supplied to end-users to suit their data preference or to allow for model inter-comparison analyses.

Since data from a number of different systems (CPEC, OPEC, primary meteorological) must be incorporated to gap-fill ecosystem fluxes, data for a given site must be initially compiled (using the process described below), to create an all-data, all-years master data file for each site. This intermediate all-data master file is used as input to gap-filling operations, and the compiling program is again run after gap-filling is accomplished.

Compiling data products for end-user access

Once data is quality-controlled and products have been generated, all variables from all systems at a given site are compiled into a single half-hourly 'site-wide' master file that spans all relevant variables and measurement years. A data aggregation program in BACON (mcm_data_compiler.m, Figures 2.5 to 2.7) automatically and sequentially loads all data files for each year at a given site, and copies the needed output variables to its desired location in the site-wide data master output file. From the all-data master file, any combination of variables may be extracted to produce additional data products for sharing with local and external end-users. These operations are guided by a site-specific, operator-editable configuration file (located in /Matlab/Config/Master_Files, which specifies the

49

variables from each system that will be incorporated into a site's site-wide master file, as well as the amount, size, contents, structure and format of generated data products. The desired structure of the site-wide data master and generated data products may be altered or updated at any time by modifying the configuration file and re-running the compiling program.

In cases where the same variable is recorded by different systems at a site, variables from each system are saved and separately identified in the master file. In the case where coincident time series of unfilled, quality-assured flux data exists for OPEC and CPEC systems, additional, aggregated 'all-systems' time series are generated to produce flux variables with a reduced amount of data gaps. In all cases, CPEC data is used as the base for these aggregated variables, where available OPEC data is used to fill NaNs in the CPEC data file. Both aggregated and individual variables are made available to the end-user, which they may select from to suit the needs of their application.

Data products intended for internal end-users are placed in a network-accessible location on the DWMS file system and are subsequently uploaded to an Internet-accessible shared collection by the DWMS operator. Data to be shared with external end-users are formatted by BACON to conform to the individual's or external association's submission standards, and are then zipped separately by site to allow easy uploading to the target file server. Further explanation of data sharing is provided in section 2.4.3.5.

2.4.3.3 Data analysis

Analyses outside of those included in the aforementioned data processing steps are, by design, not included in DWMS operations. To avoid the incorporation of unstable, untested or unreliable analysis products (data and methods) into the DWMS, generated data products are made available to end-users to conduct analyses outside of the DWMS. With its ability to provide consistent, high-quality and timely data products, the DWMS facilitates and augments end-user analysis capabilities. In cases where analyses are shown to generate reliable results and provide added-value to DWMS operations, their routines are incorporated into data processing operations.

2.4.3.4 Data preservation

Data preservation within the DWMS is accomplished through a number of backup and archival operations that occur both within the DWMS file system, and at distributed locations. The objectives of DWMS preservation activities are to a) ensure that data is structured in a logical and intuitive manner, and b) that all essential data files exist on at least two separate media at all times.

Data collected on ACS, CPEC and primary meteorological systems are automatically backed up onto a removable media drive by the field data collection PC. Upon insertion of a formatted external drive into the PC, all CPEC and ACS data collected during the previous 14 days are copied to the removable drive, along with the complete primary meteorological data logger file. These precautionary actions create overlap in successive data downloads, reducing the likelihood of data collection gaps. At the end of each day, data collected for the preceding day is transferred to the removable drive, keeping the data backup current. After the external drive is removed and used to transfer files to the DWMS file system, a formatted drive is reinserted into the field computer, and the redundancy and backup operations continue. Auxiliary meteorological data loggers are automatically backed up using a connected data storage module (model SM16M, Campbell Scientific Inc., Edmonton, AB), while OPEC systems are backed up to attached CF media. These backup methods maintain duplication of raw data not yet uploaded to the DWMS file system, protecting against data loss associated with failure of any individual media device.

All data on the DWMS file system (system files, processing software, raw, intermediate and finished data products, etc.) are incrementally backed-up by a remotely-located, network-accessible server. Incremental backup of the system occurs nightly, and provides protection against large-scale data loss, while also permitting restoration of file system elements to their previous states if files become damaged due to corruption or mismanagement. To reduce the likelihood of time-costly large-scale data loss events, a redundant array of independent disks (RAID) is used on the DWMS file system, where identical copies of the DWMS file system are maintained on separate but parallel hard disks. Such a setup reduces the likelihood that failure of a single hard disk would require complete restoration of the file system from the network-accessible backup server.

Due to finite storage space on the DWMS file server and the redundant nature of field-collected data, raw data files and directories are copied to optical media and eventually removed from the file system after they have been read and extracted by data organization programs (see section 2.4.3.2). Processed raw data collects in site-specific directories on the file system, and are moved by the operator to an appropriate optical media (DVD, DVD-DL, Blu-Ray Disc). Two copies of each disc are made, and are stored at separate locations to reduce likelihood that catastrophic events (such as fires) cause complete loss of raw data archives. To allow for quick and easy reprocessing of recent data, all raw data files for the current year are maintained in an archive directory on the DWMS file system (see section 2.4.3.4), and are deleted after data for the entire year has passed QA screening.

2.4.3.5   Data access

All data products to be shared with internal and external end-users are stored within a single directory on the DWMS file system (see section 2.4.3.1). Data products are created from site-wide data master files in customizable structures and formats. In this manner, different data collections may be made available to different end-users, to suit their specific data needs. To optimize the usefulness of data products for end-users, a number of data sharing considerations have been made in the DWMS:

Data is shared in multiple file formats to accommodate the diversity of software that may be used in end-user analyses. Data products intended for internal end-users are primarily made available as binary MATLAB (.mat) data files. Saving data in this format offers a number of advantages, including substantial data compression improvements over traditional ASCII files, and the ability to save hierarchical file structures.   As all research group members use this software for their analyses, these data products can be shared alongside 'value-added' Matlab processing programs that allow them to structure data files to suit their own needs. For external end-users, Matlab format files are shared in parallel with files using the traditional comma-delimited (.csv) ASCII format. Data is often required to be submitted in this (or a similar) format by research networks and data repositories, as its ubiquity ensures compatibility with all types of analysis software.

The wide range of end-user requirements for data file structures and content is addressed in the DWMS by two separate approaches. The first approach is to provide end-users with comprehensive data sets containing all possible variables, at all data quality levels and measurement periods. In this case, end-users would be required to restructure data files to suit their analysis needs. In cases where internal or external end-users express interest in receiving particular data products, BACON processing software and its associated configuration data can be quickly modified to create customized data products. This DWMS functionality has been utilized on numerous occasions to provide end-users with data files that can be immediately incorporated into their analysis routines.

Internal (research-group) end-users may access data collections directly on the DWMS file system via Secure Shell File Transfer Protocol (SSH SFTP) from within the McMaster University domain, or by connecting to its Virtual Private Network (VPN). Data products are also made available via Internet-accessible collections shared through a research-group-owned Google Documents account. To facilitate on-line sharing, the BACON data aggregation program creates compressed archives of sharable data products and saves these files to a ready-to-upload folder on the CFS. To share this data, the operator simply has to upload the archives to the target Google Documents collections to make files

accessible on-line.

Selected external end-users may access data products through the Canadian Carbon Program on-line data information system (DIS). Specially formatted data products are uploaded from the DWMS file system to the DIS via SFTP. DIS data managers are responsible for incorporating data into a research-network-wide database, and making data available to prospective users under appropriate fair-use guidelines. Files are also shared with specific external end-users by granting access permission to Internet-accessible shared collections; availability and permissions for these shares are controlled by the DWMS operator.

2.4.3.6   Data reuse

Data is extensively reused within the DWMS, where intermediate data products are further processed within BACON to create higher-quality data products. In addition, ancillary site data, including forest biometric and inventory information, are uploaded and incorporated into the DWMS to permit data operations (such as sapflow and EC footprint calculations), which require stand-specific parameters.

Furthermore, the compartmental, expandable and flexible design of the BACON software permits operators to run multiple approaches of the same processing steps (e.g. NEE gap-filling methods) in a parallel fashion. Data products from each unique run can be compared inside of BACON, to assess the effect of different data treatments on processing output. In addition, the provision of thorough documentation through metadata, processing scripts and initialization files provide the means for data revisitation/recalculation activities within or outside of BACON.

2.4.4   DWMS design features and operational functionality

2.4.4.1   The BACON graphical user interface

A Graphical User Interface (GUI), designed and implemented in MATLAB, allows the DWMS operator to implement all BACON data operations through a graphical front-end controller on the CFS (Figure 2.4). The GUI is designed to provide an intuitive and accessible means of data processing and management, while ensuring that data management best-practices are followed. Drop-down menus allow the operator to specify the desired data process to be implemented, as well as any required input parameters (i.e. site, data type and year to be processed). Input selection is outlined as sequential steps, and the selected inputs are displayed for operator review in a dialogue box beneath the drop-down menus (step 5, Figure 2.8). The DWMS operator executes the selected operation with a click of a push

button (step 6).

Background functions implemented within the GUI provide additional functionality, which assist the operator in data processing and help to reduce errors associated with improper operation and input selection. As the operator selects inputs from drop-down menus (through steps 1-4), background functions query an initialization file (mcm_mgmt_ini.m) to determine the input options available for the subsequent drop-down menu, given the selected inputs. Thus, the GUI ensures that only appropriate selections are displayed on subsequent drop-down menus, as determined by selections made in previous drop-down menus. For example, when a site and data type (i.e. system) combination is selected from drop-down menus during steps 1 and 2, a list of data processes available for this site and data combination is retrieved and displayed as options in the third drop down menu, for use in step 3 (Figure 2.8). Additionally, the processes listed in the step 3 menu are presented in the desired order for processing the specified data, following the operational methodology outlined in Figures 3-5, and described in section 2.4.3.2. Therefore, the DWMS operator may sequentially execute the listed operations to develop data from raw files to quality-assured, sharable datasets.

To ensure that the proper inputs are being selected for a given function, a background function verifies that chosen input parameters fulfill requirements for the selected data operation each time a drop-down selection is made (Figure 2.9). In the case that all necessary inputs have not been selected, the execution button (step 6) is disabled and displays a textual error message, informing the operator of the specific missing inputs. When all inputs are satisfied, the execution button is activated, and the operator may click to run the desired process.
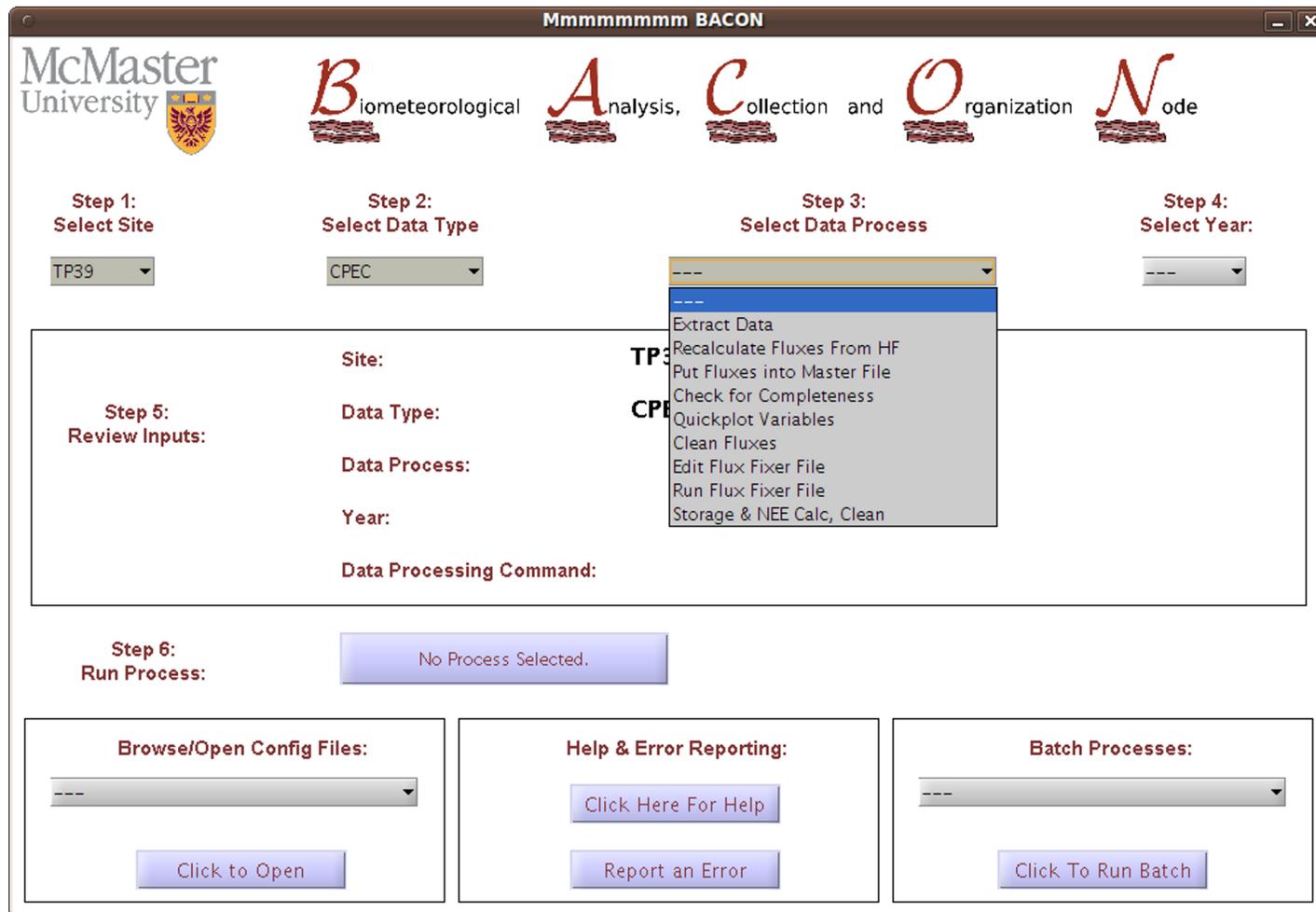
Figure 2.8: The BACON GUI during data process selection.
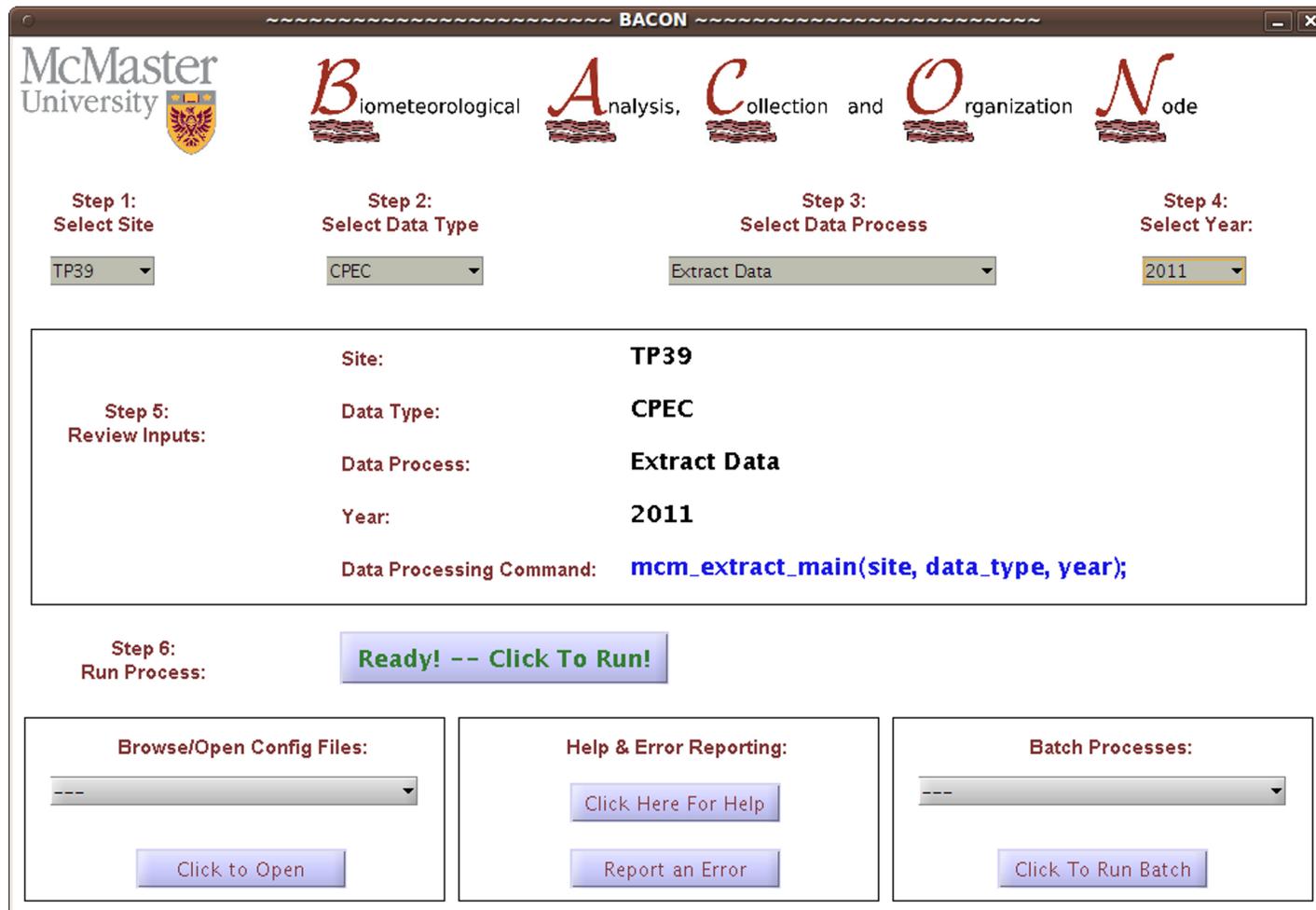
Figure 2.9: The BACON GUI in "ready to process" state.

Additional, 'value-added' DWMS and BACON processes are accessible through the bottom frames of the GUI. The bottom-left frame provides a means of opening commonly-used BACON configuration and initialization files directly from the GUI. By selecting an option from the drop-down menu and clicking the execution button, BACON opens the specified supporting documentation-containing file, directory or MATLAB script.

The bottom-middle frame provides help and error reporting capabilities for BACON. By clicking the help button at the top of the frame, an on-line version of BACON documentation (accessible at http://goo.gl/Vewv0, see section 2.4.5) is opened in the browser of the CFS PC. DWMS operators are able to report potential data errors and programming bugs by clicking on the error reporting button at the bottom of the frame. When this occurs, the DWMS operator is prompted to enter a description of the problem in a pop-up dialog box; this error description is then emailed alongside a descriptive log file to internally-specified DWMS administrators. The attached log file contains a history of the BACON operations and commands executed during the current session, and lists each input variable and its value at the time of error reporting. This provides the DWMS administrator with a means of thoroughly investigating and recreating the operator's processing environment, in order to trace and correct any system errors. The bottom-right frame allows the DWMS operator to initiate a selected BACON batch operation from the GUI (see below).

2.4.4.2 Automated and batch processing in BACON

Though operator control and guidance is necessary for most DWMS processing steps, a select number of time-intensive processes may be run automatically and inconspicuously without operator supervision. In particular, standalone batch processing options on the BACON GUI allow the operator to initiate automated extraction and organization of raw CPEC and ACS data for all sites , as well as automated calculation of half-hourly fluxes from raw data for the past 30 days at all sites (following section 2.4.3.2). All operations and errors that occur during batch processing are recorded to detailed log files, which are automatically emailed out to DWMS operators at its conclusion (using the sendmail.m command in MATLAB). In addition, automated extraction operations are followed by the production of spreadsheet-type logs that detail the completeness of organized data files at each site for a given year (as discussed in section 2.4.3.2), which are also emailed to DWMS operators. Log files provide the operator with a detailed description of all data operations, and the current state of raw and organized data within the DWMS. As a result, operators may initiate these time-intensive processes to run without supervision, and review the processing results at a later time via log files.

To further streamline DWMS operations and ensure data processing remains up-to-date, these batch programs (mcm_automated_extraction.m, mcm_automated_calc_fluxes.m) are automatically executed in the background operations of the CFS PC on a weekly basis. The Unix scheduling program *cron* is used to execute the batch programs at midnight on consecutive weeknights, after which log files are emailed to data operators for review.

Automated scheduling is also used to execute a number of additional BACON operations at regular intervals. A support documentation-compiling function (mcm_documentation_compiler.m) is executed automatically at monthly intervals to compile all supporting data files (e.g. documentation, metadata, manuals, configuration and initialization files; see section 2.4.5) within the CFS into a single, zipped archive. This archive is maintained with previous archives in the /Documentation primary directory of the CFS to provide operators quick access to backup versions of supporting documentation used in the DWMS.

2.4.4.3   Data quality monitoring in BACON

Regularly and thoroughly monitoring all recently-collected data is an essential quality-assurance component of biometeorological research, as it allows the researcher to identify and address many of the instrumental and operational problems that are common in this type of fieldwork. Damaged or improperly calibrated instruments, data logger malfunctions, site power problems and data download gaps are examples of measurement problems that can be diagnosed and promptly remediated by collecting, plotting and inspecting data at short intervals.

Remote access to field collection PCs permit daily checks, where the DWMS operator checks the general state of the field site and its measurement systems, and records observations to an Internet-shared log file. As both the remote client software and the log file are web-browser-accessible, operators can assess field site status from any location with an active Internet connection. Daily checks serve to identify large-scale, conspicuous measurement problems, and should be complimented by more thorough data inspection of collected datasets.

By streamlining data collection and processing operations, the DWMS and BACON provide the operator a means of inspecting collected data in a timely manner. To serve this specific requirement, GUI-accessible variable plotting programs allow the operator to plot annual time series of all measured variables for a given site, system and year; a pop-up dialog box prompts the operator to choose the data quality level to be plotted – either organized or threshold-cleaned data. To avoid excessive plot production, and improve inspection efficiency, the number of figures produced for meteorological

variables is condensed by dividing each figure into six subplots with different variables. Each figure is saved in the CFS /Matlab/Figs directory and manually shared to Internet-accessible Google Documents collections as raster (.png) and MATLAB vector figures (.fig); raster images allow for quick browsing at low resolution, while MATLAB vector figures support higher-resolution inspection. Figures saved on the CFS may then be accessed and inspected by any or all DWMS operators on the file system, or at remote locations.

2.4.4.4 DWMS and BACON expansion

The DWMS provides expansion capabilities to accommodate the addition of new sites, measurement systems, processes, and years of operation, with relatively little setup time required. As the DWMS was designed and implemented for use with long-term biometeorological measurements, the BACON software package and the CFS directory structure are capable of processing and storing data from any measurement year. All required year-specific directories are created at the time of BACON processing; a custom sub-function (jjb_check_dirs.m) checks the CFS for desired output directories, and prompts the user to approve of their creation if they do not exist.

New sites (and their measurement systems) are incorporated into DWMS operations by adding site and system details to the BACON initialization file (mcm_mgmt_ini.m), as well as a number of additional system-specific initialization files. In the same manner as adding years, a custom MATLAB function (via jjb_check_dirs.m) automatically creates all necessary CFS directories when needed during BACON operation.

The ease at which additional measurement systems may be added to sites depends on the nature of the newly installed system: Systems that operate similarly to those already managed by the DWMS (i.e. CSI-based meteorological systems, similar CPEC, ACS or OPEC systems), are incorporated into the DWMS by modifying system-specific initialization files, as discussed above. Incorporating systems that are novel to the DWMS would require new sets of BACON processing programs, CFS directory structures, and initialization files to be developed and implemented.

The expansion capability of the DWMS was assessed when a new TPFS measurement site (temperate deciduous forest; site label: TPD), was initiated in early 2012. Site instrument setup included a primary meteorological system similar to those described for other TPFS sites (section 2.4.3.1), except for the use of a different data logger (model CR3000, Campbell Scientific Inc., Edmonton, AB, CA). The implemented CPEC system used a familiar sonic anemometer (model CSAT3, Campbell Scientific Inc.), and a different IRGA (model Li-7200, LI-COR Biosciences, Lincoln, NE, USA), though data

collection systems and formats were comparable to other TPFS sites. To incorporate this site into the CFS and BACON processing, data collected from this site was placed into a newly-created subdirectory in the /SiteData primary directory, and all necessary meteorological and flux-related configuration files were created. Following this preparation – which required less than 8 working hours by DWMS operators -- all BACON data processing functions were successfully applied sequentially to the new data, confirming that TPD had been seamlessly incorporated into the existing data management operations.

2.4.4.5   Retroactive data incorporation and reprocessing

In addition to accommodating subsequent measurement years, the DWMS is also capable of incorporating previously-measured data in a retroactive manner through simple modifications to initialization files. This provides a means of incorporating and reprocessing any data that was collected before the current DWMS was implemented. Following implementation of the current DWMS in 2008, raw data files that were collected during previous measurement years (2002-2007) were extracted from archives stored on optical media, and placed in the /DUMP_Data primary directory on the CFS. After all required initialization files were created, these files were processed using BACON, and incorporated into existing data directories and data products on the CFS. As data from this period previously existed on the CFS only as undocumented, end-product spreadsheets, retrieving and incorporating these raw data through the current version of BACON ensured operational equivalence and consistency among all data products on the CFS for all years.

Processing these data through BACON also revealed a considerable number of problems in the previous end-product spreadsheets. Identified and corrected issues included: undetected and uncorrected instrument and measurement errors (calibrations, improper sensor labeling); inconsistent data processing methodologies (data cleaning, normalization, calculation of derived values), and incorrect spreadsheet manipulation (row and column mismatches, missing data). By reprocessing the raw 2002-2007 data in BACON, these problems were effectively documented, addressed, and corrected. The methods used to detect and correct these problems led to the implementation of error-checking subroutines within BACON data processing operations, which serve to identify such data problems in all newly-processed data.

As all quality levels of data (except for raw data) are retained on the CFS, data can be reprocessed in BACON at any time by selecting and running the desired process on the GUI. Data organization on the CFS allows simple and efficient recalculations in the case that processing operations are modified or

augmented in the future. The data that is generated from this process overwrites the previous data on the CFS, which allows for subsequent data processes to be re-run to incorporate this newly-created data. Raw data files may be reprocessed by re-adding them to the /DUMP_Data primary directory prior to running the necessary BACON operations.

2.4.5 DWMS supporting documentation

DWMS operations require and create a wide array of supporting documentation, including metadata documents, processing scripts and functions, initialization and configuration files, operational protocols, as well as instrument and system manuals. As displayed in Figure 2.10, supporting documentation is maintained at a number of locations within the CFS.

BACON processing scripts and functions are located in the /Matlab/Scripts directory, in subdirectories that correspond to their BACON functionality. Commenting included in each MATLAB file provides documentation of its overall functionality, its procedural methodology, and allows DWMS operators to observe and trace the changes made to the script over time.

Parameter and configuration files specific to BACON operations are stored in the /Config and /Scripts subdirectories of the /Matlab primary directory. Configuration files are organized in the /Config subdirectory according to the process that uses them. Descriptive names are given to these directories to assist the DWMS operator in identifying the appropriate configuration file. This organizational structure maintains consistency between directories used for data products and support supporting documentation, which allows operators and processing functions to efficiently locate desired supporting documentation.

All protocol documents, site metadata, instrument manuals, calibration information and processing logs are located and maintained within the /Documentation primary directory of the CFS. The /CCP_Metadocs subdirectory includes thorough site-specific metadata documents that were created according to Canadian Carbon Program (CCP) standards. These standards and all other CCP protocol documents are maintained in the /CCP_Protocols subdirectory, while research group-created protocols

```
/1/fielddata
-/Matlab
    -/Scripts                              ← All BACON processing scripts and functions
        -/mcm_data_mgmt                    ← All top-level BACON processing scripts and functions

            Site-specific calculation parameters stored in params.m;  BACON processing and GUI settings in
            mcm_mgmt_ini.m;  ACS, CPEC data settings in mcm_get_fluxsystem_info.m;  Manual corrections
            to data in mcm_metfixer.m, mcm_fluxfixer.m, OPEC_10min_fixer.m and OPEC_EdiRe_fixer.m

        -/Met                              ← Meteorological calculation scripts/functions
        -/Flux                             ← Secondary flux processing & calculation functions
        -/Footprint                        ← Footprint calculation functions
        -/Gapfilling                       ← Gap-filling functions
        -/Curve_Fitting                    ← Curve fitting & optimization functions
        -/Error                            ← Flux random error calculation functions
        -/genfuns                          ← General processing functions
    -/Config
        -/Met
            -/Organizing-Header_OutputTemplate  ← Header documents; Templates for data organization
            -/Cleaning-Thresholds          ← Thresholds files for meteorological cleaning operations
            -/Cleaning-BadDataTrackers     ← Tracker files for problematic time series
            -/Filling-SoilPitKeepLists     ← Tracker file for soil temperature average calculation
            -/Calculating-SapflowParameters ← Site-specific parameter files for sapflow processing
        -/Flux
            -/CPEC/Cleaning-Thresholds     ← Threshold files for cleaning all CPEC data (all sites)
            -/OPEC/                        ← Header documents and Master file templates for
                                              organization;  Threshold files for cleaning OPEC data
            -/Gapfilling/Defaults          ← Initialization files that specify gap-filling preferences
        -/Master_Files                     ← Initialization files for all-years Master files; Specification
                                              of format and contents of end-user-shared datasets
    -/ubc_PC_setup
        -/Site_Specific/<site>             ← Initialization files for CPEC and ACS flux calculation

-/Documentation
    -/CCP_Metadocs                         ← Metadata documents for each site's EC and primary
                                              meteorological data, as uploaded to the Canadian Carbon
                                              Program DIS
    -/CCP_Protocols                        ← Canadian Carbon Program protocol manuals
    -/IRGA_Calibrations                    ← Calibration records for all IRGAs
    -/Manuals                              ← Instrument manuals
    -/Protocols                            ← Research-group protocols and systems documentation
        -/Biomass                          ← Biomass measurement protocols
        -/Data_Management                  ← DMS protocols, organizational documents
        -/Fine_Root_Biomass                ← Fine Root Biomass measurement protocols
        -/Sapflow                          ← Sapflow measurement protocols and metadata
    -/BACON_Documentation_Archive          ← Destination for automated support file backup
    -/Logs
        -/cron_processing                  ← Log files for Linux automated scheduler
        -/mcm_auto_extractor               ← Log files for automated flux data extraction
        -/mcm_automated_calc_fluxes        ← Log files for automated flux calculation
        -/mcm_checkfiles                   ← Inventory lists for downloaded and processed flux data
        -/mcm_data_compiler                ← Log files for entire-site data compiling
        -/mcm_documentation_compiler       ← Log files for automated documentation collection
        -/mcm_Gapfill_main                 ← Log files for gap-filling operations
```

Figure 2.10: *Summary of DWMS centralized file system structure and contents for documentation, support and metadata.*

are located in the /Protocols subdirectory. All processing log files created during BACON operation are saved to folders in the /Logs subdirectory that correspond to the name of their creating processing functions. As described in section 2.4.4.2, compiled and compressed backups of supporting documentation are regularly and automatically generated by BACON, and saved to the /BACON_Documentation_Archive.

A research group-authored manual for TPFS field and BACON operations is collaboratively edited and maintained as a shared, Internet-accessible Google Document. The TPFS Field and Data Management Manual (FDMM), which may be accessed on-line at http://goo.gl/jk3Fe, provides documentation, description, operational instructions, and troubleshooting for all DWMS components. The accessibility of this manual provides DWMS operators with important instructional and organizational material at all times during field-site visits, and all other data management-related activities.

2.5 Discussion

2.5.1   Critical evaluation of DWMS components and operations

Field data management

The nature of field data management operations at TPFS reflects the integration of idealized best-practices with the extant systems and operations at each site, and the equipment and human resources available to the research group. An example of this compromise is the fact that field measurement systems are not completely standardized within and between sites; although standardized methodological protocols guide the general site configuration, a number of different instruments, data collection devices and formats are used. For the most part, these discrepancies have developed from the replacement and upgrade of discontinued components at sites with newer, improved models. Such an issue is expected in any long-term measurement program, as measurement devices and collection components have limited lifespans, and equipment manufacturers regularly develop new products to improve on older models.

Despite differences in the composition and contents of raw data, field collection systems are capable of connecting to each system, retrieving data, and later uploading it to the CFS for processing. Disparate raw data formats, structures and contents are accommodated in the DWMS by the development and incorporation of metadata into data processing activities. Whether embedded in data files, or created

by a DWMS operator, these metadata files allow for all existing and future raw data files to be incorporated into management operations, regardless of their characteristics. This functionality is critically important to DWMS success for any similar research program, as raw data file changes are generally an inevitable part of long-term measurement activities.

As part of an effective DWMS, all field systems and operations are extensively documented by the research group. Protocol documents guide the data collector through all aspects of managing data in the field, which includes connecting to logging devices, downloading data, naming files appropriately, filing the data on the transfer device, and uploading the data to the CFS. Operator-initiated quality assurance operations, such as daily field system checks and weekly data checks help to ensure high-quality measurement data. In addition, detailed field visit and data download logs provide a means of tracking past activities and changes to field measurement systems.

By accommodating all forms of raw data and standardizing its collection and processing methodology, the DWMS enables any member of the research group to collect and incorporate data into subsequent DWMS data activities. To streamline this process further, future DWMS development will aim to automate data download, leveraging Internet connectivity at the field sites to regularly and automatically upload data directly to the CFS via its FTP server. Automating these tasks ensures that data is downloaded from systems at regular, short intervals, reduces the field time requirements of research group members, and eliminates the potential of operator-caused data download errors (Campbell et al., 2013).

DWMS software (BACON)

The BACON software package standardizes, streamlines and automates data processing and organization activities within the DWMS, allowing operators to manage data collectively on a centralized system. The software is implemented in a scripted language and data operations are executed according to explicitly-stated commands, therefore eliminating errors associated with manual matrix manipulation such as spreadsheet processing (Borer et al., 2009). In addition, the scripted programs clearly describe the operations applied to the data, and effectively document the history of changes made to the programs – two types of supporting documentation that are typically not supplied by spreadsheet or other manual operations. The centralized implementation of the software ensures that all DWMS operators process data in a consistent manner, and that any changes made to programs will be applied to all necessary data with no version inconsistencies.

BACON is compartmentalized into separate, independent functions, so that changes can be made to

any operation with only minimal changes to other software components. However, editing scripted BACON programs or changing the way they are implemented requires medium- to high-level knowledge of the software and MATLAB programming; thus, the processing changes that may be available to operators via a spreadsheet program are restricted to experienced program users. Although such restrictions may inhibit short-term program development and innovation, it promotes long-term software usability and consistency, by restricting program modifications to users with a thorough understanding of both the software and its programming language. This functionality is aligned with general elements of software design best-practices, as software operations should be inaccessible to common users, and should be modified only by qualified operators, after appropriate planning and discussion with other data managers (IEEE Computer Society, 2004a).

Processing data to ultimately develop a finalized product is a relatively straightforward task with BACON, and especially so through the use of the GUI. The GUI is easy to access and implement, as it guides the operator through each processing step, and provides feedback to ensure that operations are being executed in the correct manner. While requiring minimal operator-input, the provided functions allow the operator to sequentially organize raw data into consistent structures and formats, visually and interactively inspect and clean data, fill and calculate required variables, and share the data products with end-users.

In addition, eligible data processes are run in batch mode and/or are automated within BACON to further streamline processing operations and ensure they are executed at regular intervals. The implementation of this software package is essential to DWMS success, as effective biometeorological data management and processing requires the careful application of a great deal of complex processing operations, which are mostly automated within the software. By providing operators with robust, comprehensive but easy-to-use programs, the software helps to reduce the likelihood that data management will be negatively affected by shortcomings in the operator's software or methodological knowledge. An important consequence of this accessibility, however, is that DWMS operators may be able to process data before having extensive knowledge of the details of the operations taking place. In such situations, the thorough supporting documentation provided by the DWMS – and especially comments within the processing scripts – provide a means for operators to gain knowledge of the processing steps that are implemented. By reducing encapsulation restrictions and providing operators-in-training read-only access to processing functions, BACON and its GUI may function as a training tool, where trainees can simultaneously process data while observing the details of data

processing methodology.

While most data operations are explicitly defined in processing programs, a select number require the application of 'operator judgment'. Data processes such as detailed manual data cleaning and calibration (see section 2.4.3.2) often require the operator to assess the quality of a selected number of data points by subjective means; this assessment may be based on previous knowledge of the site's environment and equipment, theoretical or experimental knowledge of the variable of study, or the coincident values of other variables. A concern with such subjective data management methods is that the relative ease in which they may be applied in BACON (where permitted), may misrepresent the level of care that must be taken when imparting operator judgment on data. To address this concern, it is important that operators receive appropriate training for all data management methods, and that operators-in-training are supervised by experienced data managers. As all manual actions are recoded in supporting documentation or in processing scripts, such data operations can be reviewed and potentially modified as desired, after which the data may be reprocessed to reflect the desired changes to subjective means.

DWMS central file system

As presented in section 2.3.3, effective DWMS operation in a collaborative framework requires a number of data organization features to be realized. Included in these requirements are the concepts of data centralization, logical and consistent data structure, standardized file formatting, naming and organization, the provision of data headers, and thorough data documentation (see Cook et al., 2001; Borer et al., 2009; Hook et al., 2010; Van Den Eynden et al., 2011).

The DWMS file system effectively organizes and maintains data, documentation and programs at a centralized file system that is accessible to all data operators. The file structure on the CFS is highly compartmentalized; files are sorted into top level directories according to their type (documentation, raw data, archived data, processed data, etc.), and are further named and grouped into subdirectories according to the details of its measurement site, system and year, as well as data quality level (Figure 2.3). The logical and consistent file organization and naming structure serves a number of purposes: Primarily, it increases the effectiveness and compatibility of processing software, as programs use consistent rules and notations to load, process and save data files, thus reducing the number, diversity and complexity of required program inputs. Additionally, in cases where the DWMS operator needs to extract a specific data file, they can be located efficiently based on the characteristics of the sought data.

Although raw data files are maintained in their native formats, and end-user data products are made available in non-proprietary, universally-accessible (ASCII) format, most intermediate data products (e.g. master files) are stored in proprietary MATLAB binary format. This procedure is in contrast to general best-practices guidelines presented by such authors as Hook et al. (2010) and Borer et al. (2009), who raise concerns for the eventual incompatibility of legacy formats with newer program versions, or the abandonment of the format altogether. Though the history of computing has provided a number of examples of such problems, the estimated risk for this application is greatly outweighed by the advantages offered by storing data in this proprietary binary format. An advantage of this approach is a greater than ten-fold reduction in file size, greatly reducing the hard disk (and backup) requirements for the numerous intermediate data products that are created. Additionally, the propriety format supports hierarchical file structures, which permits numerical data arrays and their metadata to be contained within a single file, as opposed to generating and managing separate flat files for data and metadata. As BACON software operates on the MATLAB platform, data that is produced within its environment in proprietary binary format will remain readable by its operations, and readable by any newer version of MATLAB that is implemented. In the case that migration from a MATLAB environment is required in the future, functions available for C/C++ programming languages may be used to convert these files to another desired format.

Currently, all data, documentation and software files are maintained permanently on the CFS, except for raw field-collected data, which are retained in the file system for a limited time (one year) before being removed to ensure sufficient free hard disk space. Research activities and data operations result in a net hard drive-use increment of 50-100 GB per year – a rate of data increase that is sustainable considering the current 3 TB capacity of the CFS. Considering the current rate of technological advances in hard disk capacity, it is very likely that the current PC-based CFS system will provide adequate functionality to accommodate current measurement activities and any potential expansions for the foreseeable future.

Data preservation and sharing

As is the case with any data-intensive research operation, long-term storage and backup is a critical issue in data management, which must be addressed on a recurring basis as data requirements change and storage technology advances. Currently, all essential data, programs and supporting data are backed up at regular, short intervals (i.e. hourly to daily) on all DWMS file systems (field collection systems, data loggers and the CFS). With this approach, essential files are kept redundant on separate

media at all stages of the data life cycle. On the CFS, the use of RAID and incremental backups to a remote server ensures that files on the CFS are protected from independent hard drive failure as well as catastrophic data loss events. Wherever possible, automated backup systems have been implemented in favour of manual operations such as optical discs and flash drive backups, as they are generally less vulnerable to read/write errors, interoperability conflicts, media mishandling problems and have better error-checking functionality.

In field-based operations, backup media are currently physically connected to the source file system, and therefore, must be maintained in its close proximity. This setup protects field-measured data from system failure or hard disk corruption, but leaves it vulnerable to large-scale, catastrophic data-destroying events, such as floods and fires. To address this problem, data is collected from the field at frequent, regular intervals. Future plans call for Internet access to the field sites, and automated data download to the CFS.

As described in section 2.4.3.4, raw CPEC and ACS data files are excluded from automated CFS backup operations; instead, these files are manually recorded to optical media and are deleted after their contents have been incorporated into the CFS and verified. In this manner, all relevant raw data is maintained both on the CFS (in an organized manner), as well as optical media (in unorganized field-downloaded structures). These steps are taken to avoid unnecessary backup of disk space-intensive collections whose data is effectively redundant with extracted and organized CFS data repositories. As a result, a total of 100-200 GB of optically-archived data is created annually.

Sharing high-quality data with end-users represents the culmination of DWMS operation, as the products of careful data management and processing are used in analyses to generate new information and knowledge. Described in section 2.4.3.5, the flexible and extensible nature of BACON processing makes it possible to readily create and share customized data products to internal and external end-users. This functionality reduces end-users' likelihood for making data manipulation errors, and reduces the time required for data preparation and analysis setup.

DWMS supporting documentation

In a collaboratively managed DWMS, creating and maintaining support documentation is critical to operational success, as it facilitates communication between research group members and helps to coordinate and standardize DWMS operations. To support collaborative maintenance of high-quality support documentation, these documents should be sufficiently thorough, protected against unintentional branching of its versions, and easily accessed by all DWMS operators.

Shown in Figure 2.10, DWMS supporting documentation encompasses a variety of different files, from those providing BACON-specific information (processing scripts, metadata, and configuration and initialization files) to those describing general DWMS operations (manuals, protocol documents, and log files for manual and automatic processes). Through this documentation, DWMS operators have access to information regarding all aspects of DWMS processes. Furthermore, its use as training and informational resource serves to make data operations more transparent and understandable. However, a significant challenge with this implementation is ensuring that the documentation itself is accessible and navigable. Support documents that cannot be easily located, accessed and understood by all research group members risk being ignored and neglected. These concerns have been addressed in the current DWMS by ensuring that documentation is centralized, organized effectively, and can be accessed as simply and freely as possible.

As outlined in section 2.4.5, the DWMS's supporting documentation exists at centralized locations in the CFS and in Internet-shared collections; within these locations, support documentation files are structured into different subdirectories and collections that correspond to their specific function within the DWMS. A well-organized and centralized documentation repository permits the user to find desired documentation promptly by navigating to a single, consistent location. This structure also increases the likelihood that operators are accessing the most up-to-date version of support documentation. By maintaining the TPFS Field and Data Management Manual (FDMM, see section 2.4.5) as an on-line document, DWMS operators are able to view and modify it from any device with Internet connectivity. Resultantly, the FDMM can be accessed and modified as-needed in the field via Internet-connected PCs or personal smart phones, or in any other setting where it is required. Access to all supporting documentation is further enhanced by their integration into BACON, as DWMS operators can access these documents through GUI operations.

2.5.2 Evaluating the collaborative DWMS framework for biometeorological research

The viability of a collaborative framework in biometeorological data management

As has been displayed in the preceding sections, a collaboratively-operated DWMS is a feasible and viable method of managing data in a biometeorological research programs where a dedicated, long-term data manager does not exist. By making it possible for research group members to assume

roles as DWMS operators, the substantial time requirement for effective data management can be spread out among all members.

A successful and long-term sustainable collaborative DWMS is viable only with continual commitment by all research group members to ensure that they partake in data management best-practices. Achieving research group-wide support for large-scale data initiatives can prove difficult, as not all members share the same awareness or motivation for committed, long-term data management operations (Ives et al., 2008). Indeed, there may often be little incentive for researchers to develop data beyond what is immediately needed for their own research and publications. To address this concern, it is critical for individuals in leadership roles to stress the importance of thorough data management, and to actively educate research group members on the benefits of applying best-practices – or conversely, demonstrating the negative consequences of improper data management (Strasser et al., 2012). Beyond this, the most effective means of ensuring researcher engagement in the DWMS is for the system to provide an overall benefit to their personal data management needs. Functionality that streamlines and automates data procedures provides immediate incentive for researchers to use the system and participate in DWMS operations (Campbell et al., 2013). This functionality has been incorporated into the current DWMS by ensuring that data operations and products accommodate the needs of each research group member, and that these services are provided in a manner that is more straightforward, reproducible and time-efficient than can be achieved by working outside of the DWMS. These advantages of DWMS operation are demonstrated and reiterated regularly, so that research group members are made aware of the benefits of their participation.

Training incoming research group members to be effective DWMS operators is a critical requirement in a collaborative framework, as data management responsibilities are regularly passed-down from senior members upon the completion of their research requirements. Making sure that incoming members have access to thorough and complete documentation is paramount to ensuring a smooth transfer of responsibilities. Manuals must include detailed protocols for operating field instruments and collection systems, collecting data, as well as managing data through all components of the DWMS and the data life cycle. Documentation and processing scripts should describe the data processing operations that are used in the DWMS, and include appropriate background and methodological references, while logs should describe all DWMS-related activities. Creating and maintaining these many information sources requires a significant time commitment on behalf of all research group members, and it is difficult to ensure that details within documents are relevant and accurate. Overcoming these

70

challenges requires diligence on behalf of all DWMS operators, as well as original solutions that streamline documentation and training processes. To this end, a number of such solutions have been implemented in the current DWMS. Operations manuals and logs are maintained as on-line documents, which promotes collaborative development and editing of material, and encourages research group members to update information whenever possible. (section 2.4.5). Instructional videos for field and BACON operations have also been created and shared on-line, to provide DWMS operators with an information-dense, and quickly browsable audio-visual manual. The BACON GUI also serves as an important training tool, as its organization, step-wise implementation and integration with support documentation provides an interface that educates as it operates.

Perhaps the most important consideration for long-term collaborative DWMS success is the prompt and effective administration of changes to system components. As biometeorology research programs are dynamic in nature, it is expected that both trivial and significant DWMS modifications and upgrades will be needed over time, to accommodate changes in data collection, processing, archival and sharing operations. The purpose of a collaborative DWMS is to allow data management responsibilities to be shared amongst all research group members. However, it is likely that in most settings, DWMS administration will need to be undertaken by a select number of research group members with adequate skill and DWMS knowledge to make important changes to system components. Though the collaborative DWMS framework requires all operators to make small changes to the system when needed to maintain functionality, it should also discourage or restrict unskilled operators from making large changes that may have substantial implications for the entire DWMS operation. Therefore, it is crucial that there exist capable and committed administrators that can make higher-level changes when needed, and communicate these updates to the rest of the research group. Administration is particularly crucial for DWMS data processing operations (e.g. BACON), as the complexity of programs to manage, process and share data will likely be beyond the skill and experience of most DWMS operators. In the TPFS DWMS, the system is administered by a few, senior research group members, who are responsible for high level changes to data management operations. The use of a GUI to run all data processing operations allows effective control of the DWMS modifications that are available to operators. The GUI provides DWMS operators the opportunity to modify settings such as configuration and site parameter files, but does not permit them to make fundamental or large-scale changes to the file system structure or processing programs, which may only be carried out by system administrators.

Despite its apparent successes, however, it is apparent that such a framework will only be successful if it can be operated in environments where: a) all research group members are engaged in the data management process; b) appropriate training programs and support documentation are developed and maintained; and c) operations can be administered by a group member with adequate skills and experience. Considering that research group member turnover rates are relatively high, and that members vary greatly in their motivations, skills and capabilities, it is paramount that those in leadership positions work diligently to ensure that these requirements are met within the research group at all times.

Considerations for DWMS use with other research groups and measurement programs

In a general sense, most biometeorology research programs share similar data management requirements, as a result of similar research goals, methodologies and measurement equipment. In this respect, the general DWMS framework and operational methodologies developed in this research should be immediately transferable to other groups conducting similar research. However, the feasibility of applying this study's entire DWMS to another research program will undoubtedly be determined by how the specific needs, resources and implementational details differ between the prospective program and that which is presented in this study. A wide variety of potential inter-program differences can be expected, each with specific challenges to implementation and proposed adaptation solutions.

Differences in data collection and transfer devices (data loggers, logging computers, transfer media, etc.) will result in field protocol discrepancies between research groups. In response, field protocols will need to be modified to serve the situation of the prospective research group, though these protocols will continue to be guided by the same general principles (e.g. data redundancy, documentation, archiving, etc.) that have been established in the current DWMS.

The current field data collection protocol requires regular, weekly to bi-weekly visits to field sites to collect data from logging devices. Collecting data at this interval keeps data integrity by ensuring that logging devices do not exhaust their writable memory, and allows DWMS operators to organize and check data regularly. At more remote field measurement sites, however, data may not be able to be downloaded as regularly. To avoid the potential problems listed above, it would be necessary to equip sites with Internet connectivity to broadband, satellite or mobile networks, to allow system connection, inspection and data download either by DWMS operators, or by automated means (Benson et al., 2010;

Campbell et al., 2013; Peters, 2010).

The number of potential data characteristic differences that may exist represents a barrier to data extraction and checking efforts, as formats would not be immediately understood by the DWMS and its processing software. For meteorological data, such differences can be accommodated straightforwardly by updating the associated metadata and configuration files that are already required by BACON processing. Since much of meteorological measurement has been standardized in terms of its raw data format, (e.g. CSV), collection devices (e.g. CSI data loggers) and instruments, there is high confidence that the adaptation process would be straightforward.

The modifications and resources required to incorporate different flux (OPEC and CPEC) systems into the current DWMS will be determined by the characteristics of the systems' raw data files. Data characteristics such as file encoding and organization will depend on whether data is logged onto a data logger or a PC, as well as what type of logging software is used. Generally, flux systems record binary-encoded data, which is either written to a single, accumulating flat file, or into separate, time-partitioned (e.g. half-hourly) files. Data files of either type can be incorporated into DWMS operations by modifying the existing extraction/checking operations. For systems that create accumulating raw files, the OPEC operations would be modified for this purpose, while the CPEC processes would be modified in the case of time-partitioned files. By making minor processing script modifications and providing BACON with necessary configuration and metadata files that describe the flux systems' raw data file configuration, data from these systems can be incorporated into the existing DWMS' extraction and checking processes in a relatively straightforward manner.

The current DWMS is built upon a presumed static CFS structure, which is leveraged by BACON to access and write data files during processing. The migration of the current DWMS to one or more systems with different organization and sharing characteristics would have implications for most of its operations. Since BACON relies on a fixed file structure (data and configuration files), it would need to be reconfigured by: a) modifying all processing scripts to read and write to the updated directories, or; b) removing static references and instead require BACON to load a configuration file that contains file address information for all data types and levels for each site. The latter approach is suggested, as it would provide more flexibility for DWMS operators to quickly and easily change directory locations at any desired time in the future.

The data processing operations used in BACON may need to be modified to accommodate the methodological preferences of the prospective research group. New BACON processes may be

directly substituted for existing processes when they are mutually exclusive and operate on data at the same quality level; such operations would require only minimal modifications to the DWMS, as directory locations could be maintained, while processing lists on the GUI would need to be updated. If possible, any process that is unique and additional to those currently in BACON should be incorporated by inserting the operations within existing processing scripts. This approach requires minimal DWMS modifications, as it operates within the extant processing methodology and file structure. In cases where novel processing scripts would need to be constructed and incorporated into the existing sequence of BACON operations, it would be necessary to create new file system directories to accommodate these data, and to update data directory references for all processing scripts that are run immediately before and after these processes.

Since the DWMS has been designed to work within a fully collaborative management framework, a wide range of research group time- and work-partitioning strategies can be used with it to deliver acceptable data management. Though it is likely that operations would be more time efficient with fewer DWMS operators, responsibilities can be spread among any number of operators to reduce the workload for any particular user. This approach is viable with the provision that collectively, DWMS operators ensure that all protocols are followed, supporting documents are regularly updated, and that data is processed, checked and shared at acceptable frequencies.

Operating the DWMS requires a baseline understanding of the theoretical background and operational details of the DWMS, as well as basic knowledge of the software that is used during these operations. It is possible that members of any given prospective research group may not meet the research program-specific requirements. To overcome this problem, it is important that research group-maintained manuals and tutorials should be available to members at all times, to ensure that new DWMS operators can gain the required knowledge prior to operating the DWMS, and that existing operators can review and update documents where needed. By creating a research environment that strongly emphasizes thorough documentation and support, research groups can help to ensure sustained, proper DWMS operation.

2.6 Conclusions

In this study, we presented the guiding principles, motivating factors, developmental methods and operational details for a comprehensive, collaborative biometeorological data workflow and management solution. The TPFS DWMS is an invaluable tool for all data operations within the

research group. Complemented by support documentation, the DWMS supports data management at all life cycle stages – effectively streamlining, coordinating and standardizing data-related activities. The DWMS explicitly prescribes the activities, systems, management protocols and processing operations undertaken in the field and the laboratory. These characteristics reduce the data management time requirement for researchers, and reduces the likelihood of data errors due to mismanagement.

Data and software are centralized in the DWMS, enabling multiple data managers to collaboratively process and share data at a single location to create consistent products. Processing methodologies are robust and aligned across all data types, promoting data consistency and processing transparency. The DWMS file system and software are extensible and flexible, and can be readily updated to include additional processes and data inputs into existing management operations. Furthermore, its capacity to create customized data products and share them by a variety of means increases data usefulness and accessibility for end-users. The BACON GUI substantially improves the DWMS operator's accessibility, usability and understanding of data processing operations in the DWMS. The simplicity and organization of the GUI streamlines program execution, and provides a framework for learning more about the DWMS data processing methodology.

The success of this study demonstrates that provided appropriate system design and maintenance, as well as a continual commitment of research group members to best practices, a collaboratively-operated DWMS is a feasible and viable method of managing data in a biometeorological research program where a dedicated, long-term data manager does not exist. By distributing research data management requirements in a thoughtful, streamlined manner, our results show that it is possible to achieve effective and ethical data practices in a sustainable and resource-mindful operational framework.

2.7 References:

Abelson, H., Sussman, G.J., Sussman, J., 1996. Structure and interpretation of computer programs, second. ed. The MIT Press, Cambridge, Massachusetts.

Amiro, B., Barr, a, Black, T., Iwashita, H., Kljun, N., McCaughey, J., Morgenstern, K., Murayama, S., Nesic, Z., Orchansky, a, 2006. Carbon, energy and water fluxes at mature and disturbed forest sites, Saskatchewan, Canada. Agric. For. Meteorol. 136, 237–251.

Arain, M., Restrepo-Coupe, N., 2005. Net ecosystem production in a temperate pine plantation in southeastern Canada. Agric. For. Meteorol. 128, 223–241.

Baker, D.N., Barton, C.E., 2009. EGY: Progress in global earth and space science informatics. Data Sci. J. 8, 226–232.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. Fluxnet: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82, 2415–2434.

Baldocchi, D.D., 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Glob. Chang. Biol. 9, 479–492.

Barford, C.C., Wofsy, S.C., Goulden, M.L., Munger, J.W., Pyle, E.H., Urbanski, S.P., Hutyra, L., Saleska, S.R., Fitzjarrald, D., Moore, K., 2001. Factors controlling long- and short-term sequestration of atmospheric CO_2 in a mid-latitude forest. Science 294, 1688–91.

Barr, A.G., Black, T. A., Hogg, E.H., Kljun, N., Morgenstern, K., Nesic, Z., 2004. Inter-annual variability in the leaf area index of a boreal aspen-hazelnut forest in relation to net ecosystem production. Agric. For. Meteorol. 126, 237–255.

Barr, A.G., Richardson, A.D., Hollinger, D.Y., Papale, D., Arain, M.A., Black, T.A., Bohrer, G., Dragoni, D., Fischer, M.L., Gu, L., Law, B.E., Margolis, H.A., McCaughey, J.H., Munger, J.W., Oechel, W.C., Schaefer, K., 2013. Use of change-point detection for friction–velocity threshold evaluation in eddy-covariance studies. Agric. For. Meteorol. 171-172, 31–45.

Barton, C., Smith, R., Weaver, R., 2010. Data practices, policy, and rewards in the information era

demand a new paradigm. Data Sci. J. 9, 95–99.

Benson, B.J., Bond, B.J., Hamilton, M.P., Monson, R.K., Han, R., 2010. Perspectives on next-generation technology for environmental sensor networks. Front. Ecol. Environ. 8, 193–200.

Borer, E.T., Seabloom, E.W., Jones, M.B., Schildhauer, M., 2009. Some simple guidelines for effective data management. Bull. Ecol. Soc. Am. 90, 205–214.

Buneman, P., Cheney, J., Lindley, S., 2011. DBWiki : A structured wiki for curated data and collaborative data management, in: SIGMOD '11 Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data. pp. 1335–1338.

Campbell, J.L., Rustad, L.E., Porter, J.H., Taylor, J.R., Dereszynski, E.W., Shanley, J.B., Gries, C., Henshaw, D.L., Martin, M.E., Sheldon, W.M., Boose, E.R., 2013. Quantity is nothing without quality. Bioscience 63, 574–585.

Cook, R.B., Olson, R.J., Kanciruk, P., Hook, L.A., 2001. Best practices for preparing ecological data sets to share and archive. Bull. Ecol. Soc. Am. 82, 138–141.

Coursolle, C., Margolis, H. A., Giasson, M. -A., Bernier, P.-Y., Amiro, B.D., Arain, M. A., Barr, A. G., Black, T. A., Goulden, M.L., McCaughey, J.H., Chen, J.M., Dunn, A. L., Grant, R.F., Lafleur, P.M., 2012. Influence of stand age on the magnitude and seasonality of carbon fluxes in Canadian forests. Agric. For. Meteorol. 165, 136–148.

Desai, a, Richardson, A., Moffat, A., Kattge, J., Hollinger, D., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M., 2008. Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. Agric. For. Meteorol. 148, 821–838.

Drewitt, G.B., Black, T.A., Nesic, Z., Humphreys, E.R., Jork, E.M., Swanson, R., Ethier, G.J., Griffis, T., Morgenstern, K., 2002. Measuring forest floor CO2 fluxes in a Douglas-fir forest. Agric. For. Meteorol. 110, 299–317.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C.T., Law, B.E., Meyers, T., Moncrieff, J., Moors, E., Munger, J.W., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agric. For. Meteorol. 107, 43–69.

Finnigan, J., 2006. The storage term in eddy flux calculations. Agric. For. Meteorol. 136, 108–113.

Fluxnet-Canada, 2003. Fluxnet-Canada measurement protocols: Working draft, version 1.3, Flux. Laval, QC.

Foken, T., Mathias, G., Mauder, M., Mahrt, L., Amiro, B., Munger, W., 2004. Chapter 9: Post-field data quality control, in: Lee, X., Massman, W.J., Law, B. (Eds.), Handbook of Micrometeorology. Kluwer Academic Publishers, Dordrecht, Germany, p. 250.

Göckede, M., Foken, T., Aubinet, M., Aurela, M., Banza, J., Bernhofer, C., Bonnefond, J.M., Brunet, Y., Carrara, A., Clement, R., others, 2008. Quality control of Carboeurope flux data–part I: Footprint analyses to evaluate sites in forest ecosystems. Biogeosciences 5, 433–450.

Gu, L., Falge, E., Boden, T., Baldocchi, D.D., Black, T.A., Saleska, S.R., Suni, T., Verma, S.B., Vesala, T., Wofsy, S.C., Xu, L., 2005. Objective threshold determination for nighttime eddy flux filtering. Agric. For. … 128, 179–197.

Hamilton, M.P., Graham, E. A., Rundel, P.W., Allen, M.F., Kaiser, W., Hansen, M.H., Estrin, D.L., 2007. New approaches in embedded networked sensing for terrestrial ecological observatories. Environ. Eng. Sci. 24, 192–204.

Hook, L.A., Christensen, S.W., 2005. Developing data management policy and guidance documents for your NARSTO program or project, Citeseer. Citeseer.

Hook, L.A., Vannan, S.K.S., Beaty, T.W., Cook, R.B., Wilson, B.E., 2010. Best practices for preparing environmental data sets to share and archive. Inf. Syst. 1–40.

Hsieh, C., Katul, G., Chi, T., 2000. An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows. Adv. Water Resour. 23, 765–772.

IEEE Computer Society, 2004a. Chapter 3: Software design, in: Abran, A., Moore, J.W., Bourque, P., Dupuid, R. (Eds.), Guide to the Software Engineering Body of Knowledge. IEEE Computer Society Press, p. 204.

IEEE Computer Society, 2004b. Guide to the software engineering body of knowledge. IEEE Computer Society Press.

Ives, Z.G., Green, T.J., Karvounarakis, G., Taylor, N.E., Pereira, F., Tannen, V., 2008. The O

RCHESTRA collaborative data sharing system. SIGMOD Rec. 37, 26–32.

Jassal, R., Black, A., Novak, M., Morgenstern, K., Nesic, Z., Gaumont-Guay, D., 2005. Relationship between soil CO2 concentrations and forest-floor CO2 effluxes. Agric. For. Meteorol. 130, 176–192.

Kao, R., Gibson, C., Gallery, R., Meier, C., Barnett, D., Docherty, K., Blevins, K., Travers, P., Azuaje, E., Springer, Y., Thibault, K., McKenzie, V., Keller, M., Alves, L., Hinckley, E., Parnell, J., Schimel, D., 2012. NEON terrestrial field observations: designing continental-scale, standardized sampling. Ecosphere 3, 1–17.

Karasti, H., Baker, K.S., 2008. Digital data practices and the long term ecological research program growing global. The International Journal of Digital Curation 3, 42–58.

Kazman, R., Bass, L., Webb, M., Abowd, G., 1994. SAAM: A method for analyzing the properties of software architectures, in: Proceedings of the 16th International Conference on Software Engineering. IEEE Computer Society Press, pp. 81–90.

Keller, M., Schimel, D., Hargrove, W., Hoffman, F., 2008. A continental strategy for the National Ecological Observatory Network. Front. Ecol. 6, 282–284.

Kljun, N., Calanca, P., Rotach, M.W., Schmid, H.P., 2004. A simple paramerization for flux footprint predictions. Boundary-Layer Meteorol. 112, 503–523.

Kormann, R., Meixner, F.X., 2001. An analytical footprint model for non-neutral stratification. Boundary-Layer Meteorol. 99, 207–224.

Lee, X., Massman, W., Law, B.E., 2004. Handbook of Micrometeorology. A Guide for Surface Flux Measurement and Analysis. Kluwer Academic Press, Dordrecht, Germany.

Lyon, L., 2007. Dealing with Data: Roles, Rights, Responsibilities and Relationships. Consultancy Report. UKOLN. Bath, UK.

Mackay, S.L., Arain, M.A., Khomik, M., Brodeur, J.J., Sciences, E., Schumacher, J., Hartmann, H., Peichl, M., 2012. The impact of induced drought on transpiration and growth in a temperate pine plantation forest. Hydrol. Process. 26, 1779-1791.

Mauder, M., Foken, T., Clement, R., Elbers, J.A., Eugster, W., Grunwald, T., Heusinkveld, B., Kolle, O., 2008. Quality control of CarboEurope flux data – Part 2 : Inter-comparison of eddy-covariance software. Biogeosciences 5, 451–462.

Melaas, E.K., Richardson, A.D., Friedl, M. A., Dragoni, D., Gough, C.M., Herbst, M., Montagnani, L., Moors, E., 2013. Using FLUXNET data to improve models of springtime vegetation activity onset in forest ecosystems. Agric. For. Meteorol. 171-172, 46–56.

Moffat, A., Papale, D., Reichstein, M., Hollinger, D., Richardson, A., Barr, A., Beckstein, C., Braswell, B., Churkina, G., Desai, A., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. 147, 209–232.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., Yakir, D., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3, 571–583.

Peichl, M., Arain, M.A., 2006. Above- and belowground ecosystem biomass and carbon pools in an age-sequence of temperate pine plantation forests. Agric. For. Meteorol. 140, 51–63.

Peichl, M., Arain, M.A., Brodeur, J.J., 2010a. Age effects on carbon fluxes in temperate pine forests. Agric. For. Meteorol. 150, 1090–1101.

Peichl, M., Brodeur, J.J., Khomik, M., Arain, M.A., 2010b. Biometric and eddy-covariance based estimates of carbon fluxes in an age-sequence of temperate pine forests. Agric. For. Meteorol. 150, 952–965.

Peters, D.P.C., 2010. Accessible ecology: synthesis of the long, deep, and broad. Trends Ecol. Evol. 25, 592–601.

Porter, J.H., Hanson, P.C., Lin, C.-C., 2012. Staying afloat in the sensor data deluge. Trends Ecol. Evol. 27, 121–9.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., Valentini, R., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Chang. Biol. 11, 1424–1439.

Richardson, A.D., Hollinger, D.Y., 2007. A method to estimate the additional uncertainty in gap-filled

NEE resulting from long gaps in the CO2 flux record. Agric. For. Meteorol. 147, 199–208.

Richardson, A.D., Hollinger, D.Y., Aber, J.D., Ollinger, S. V., Braswell, B.H., 2007. Environmental variation is directly responsible for short- but not long-term variation in forest-atmosphere carbon exchange. Glob. Chang. Biol. 13, 788–803.

Sallans, A., Donnelly, M., 2012. DMP Online and DMPTool: Different strategies towards a shared goal. Int. J. Digit. Curation 7, 123–129.

Schuepp, P.H.P., Leclerc, M.Y.M., Macpherson, J.I., Desjardins, R.L., 1990. Footprint prediction of scalar fluxes from analytical solutions of the diffusion equation. Boundary-Layer Meteorol. 50, 355–373.

Strasser, C., Cook, R., Michener, W., Budden, A., 2012. Primer on data management: What you always wanted to know (but were afraid to ask). DataOne.

Uhlir, P., Schröder, P., 2008. Open data for global science, in: Fitzgerald, B. (Ed.), Legal Framework for e-Research: Realising the Potential (SydUPLawBk 39). Sydney University Press, Syndey, Australia, p. 188.

Van Den Eynden, V., Corti, L., Woollard, M., Bishop, L., Horton, L., 2011. Managing and sharing data: best practice for researchers, third edit. ed. UK Data Arhive, University of Essex, Wivenhoe Park, Colcester, Essex.

Wiegers, K., 2003. Software Requirements 2, 2nd ed. Microsoft Press.

Zeri, M., Sá, L.D. A. A., 2010. The impact of data gaps and quality control filtering on the balances of energy and carbon for a Southwest Amazon forest. Agric. For. Meteorol. 150, 1543–1552.

Zowghi, D., Coulin, C., 2005. Requirements elicitation: A survey of techniques, approaches, and tools, in: Aurum, A., Wohlin, C. (Eds.), Engineering and Managing Software Requirements. Springer, pp. 19–46.

3   Assessing the suitability of roving eddy covariance systems to produce reliable time-integrated carbon exchange estimates at multiple sites

3.1 Abstract

As an alternative to traditional, continuous eddy covariance system operation at a single site, roving eddy covariance (rEC) approaches rotate a single system through numerous proximal sites to maximize flux measurement site coverage given limited resources. A drawback of the rEC approach is the creation of long gaps in annual net ecosystem exchange (NEE) time series, which challenge gap-filling models and potentially lead to large uncertainties in time-integrated ecosystem carbon exchange estimates. In this study, we used simulated NEE time series from two temperate forest sites to assess the feasibility of an rEC approach to produce reliable time-integrated NEE estimates. To do this, we quantified the uncertainty associated with different rEC measurement schedules (number of sites, rotation period, schedule timing), as well as the effects of different gap-filling models, and the application of a variety of proposed NEE uncertainty mitigation strategies.

Results showed that NEE estimate uncertainty was least when implementing the shortest rEC rotation period among the fewest number of sites, and was greatest for the opposite case. The length of rotation had a much larger effect on NEE measurement uncertainty than the number of sites involved. Uncertainty was minimized by implementing gap-filling approaches that incorporated multiple years of information into parameterization processes. When the optimal gap-filling model was applied to the optimal rEC scenario (2-site, 15-day rotation), the resulting total uncertainty was between 35 and 63% of the annual NEE flux magnitude at our study sites, even though approximately 70% of half-hours required filling. Annual uncertainty could be further reduced by alternating the rEC schedule between years, thus ensuring that no time of year experienced rEC gaps for all years. Applying this strategy – which must be implemented at the time of rEC operation – reduced the uncertainty by an average of 28 and 73 g C m^{-2} y^{-1} in comparison to standard NLR-HL output for our TP39 and TP74 sites, respectively.

The results of this study suggest that the rEC approach has potential to expand the ecosystems that can be investigated by research programs with limited resources, while still providing time-integrated NEE estimates with an acceptable amount of uncertainty. The methodology provided in this study provides a framework with which the rEC approach may be further investigated for suitability at a greater number and variety of study sites.

3.2 Introduction

The eddy covariance (EC) method is currently used at over 500 sites to quantify ecosystem-atmosphere exchanges of energy, momentum and trace gases, such as CO_2 (FLUXNET, 2013). This global network

of EC systems provides near-continuous information on the magnitude of net ecosystem carbon exchange (NEE), as well as the biological and environmental controls on the components of this exchange. EC carbon flux measurement data serves a wide range of interests in ecological studies, including: the parameterization and validation of ecosystem models; investigation of ecosystem exchange dynamics; and the management of ecosystem carbon resources through long-term accounting of ecosystem carbon budgets (Baldocchi, 2003; Barr et al., 2013; Moffat et al., 2007).

EC measurement networks (e.g. FLUXNET, Baldocchi et al., 2001; NEON, Kao et al., 2012 and Keller et al., 2008) have expanded over the past decade, with a goal of increasing carbon exchange quantification and understanding across ecosystem types, time, and space. This desire to increase EC measurement networks is met with the acknowledgement that such systems are costly to purchase, require personnel for their installation, maintenance and calibration (Foken, 2008), and pose logistical challenges, such as the implementation of long-term, year-round measurement programs in remote locations. Therefore, it may be relevant for research networks and groups to explore and assess means of maximizing ecosystem exchange information collection given their available resources. This may include activities such as initiating new study sites, strategically decommissioning existing ones, or implementing occasional (or part-time) measurement programs.

In this study, we investigate the roving eddy covariance (rEC) approach as a non-traditional alternative to continuous year-round flux measurements. With rEC, a single, portable EC system is rotated at regular intervals (e.g. weeks to a month) amongst two or more proximal measurement sites. The benefit of such an approach is that it permits the use of a single EC system to measure ecosystem exchange at multiple sites, effectively increasing the number of ecosystems that may be surveyed by a research program. The predominant drawback of an rEC method, however, is that NEE measurements at a given site are limited in both quantity and temporal consistency, as EC measurement periods are flanked by continuous long gaps that are equal to, or longer than the measurement period itself.

In comparison to the short data gaps (half-hours to a few days) common to typical EC operation (for causes, see: Aubinet, 2008; Goulden et al., 1996; Moncrieff et al., 1996; Papale et al., 2006; Ruppert et al., 2006; Falge et al., 2001; Moffat et al., 2007), long NEE gaps are generally more detrimental to the amount and quality of extractable ecosystem carbon exchange information for a given site (Moffat et al., 2007; Richardson and Hollinger, 2007). As gap length increases, so increases the likelihood that underlying ecosystem exchange characteristics will change during the period. This includes changes in respiration sensitivity to temperature (Reichstein et al., 2005) and soil moisture (Borken et al., 2006);

intra-annual effects of foliar nutrient status on photosynthetic capacity (Huang et al., 2011); and, changes in phenology, including bud-burst, leaf area index evolution and senescence (Barr et al., 2004). Without parameterizable data during these periods, methods that characterize environmental controls on carbon exchange and fill NEE gaps (gap-filling models) are prone to greater uncertainty and bias error, leading to errors in time-integrated NEE estimates. For example, a study by Richardson and Hollinger (2007) found that depending on the ecosystem and time of year, a single month-long gap could increase uncertainty in annual NEE estimates by as much as 35 g C m^{-2} y^{-1}. This effect is substantial considering that a site synthesis study by Moffat et al. (2007), estimated that the total annual uncertainty due to typical NEE gaps (comprising between 10 and 60% of the total annual time series) was approximately +/- 25 g C m^{-2} y^{-1}.

Considering the effects of single, long gaps on NEE estimate uncertainty, it is anticipated that the numerous, regular long data gaps introduced by the implementation of an rEC program would represent an extreme test of gap-filling model performance. This fragmented data presents challenges for obtaining reliable time-integrated (daily to yearly) carbon exchange estimates for each site – values that are desirable for applications such as estimating ecosystem carbon budgets; evaluating process-based model predictions; and, comparing EC estimates with other, lower-frequency (e.g. biometric) carbon exchange estimates (Barr et al., 2013; Falge et al., 2001; Moffat et al., 2007; Richardson and Hollinger, 2007). Therefore, assessing the appropriateness of an rEC approach requires that the impacts of their associated data gaps on time-integrated NEE estimates be understood and quantified.

Between the years 2003 and 2007, we employed an rEC approach among three forested, age-sequence sites, in order to maximize the site coverage of a single EC system. The rEC system was moved between sites at intervals of 2 weeks to 1 month throughout this time, creating NEE measurement periods that were bounded by 1-2 month long data gaps. To quantify the NEE estimate uncertainty associated with rEC operation at our sites, and to assess the larger potential for rEC application at other sites, we assessed both the feasibility and value of this approach in an objective manner. Considering these needs, the objectives of this study were as follows:

1. Quantify the effect of rEC operation on bias & uncertainty of ecosystem NEE estimates in a temperate white pine forest age sequence, and compare the effect of rEC gaps to that of shorter gaps created by standard data loss causes (i.e. operational problems, footprint and u$_*$Th filtering);

2. Investigate a number of possible rEC system rotation schedules, to identify those that are most appropriate for rEC application;

3. Evaluate the robustness of various commonly-applied gap-filling methods to produce reliable estimates of rEC-measured NEE at inter-annual, annual and intra-annual timescales; and,

4. Quantify and evaluate potential strategies for mitigating the negative effects of rEC data loss on NEE estimates.

To address the study objectives, we employed a Monte Carlo (MC) approach, where a variety of rEC measurement scenarios were applied to hundreds of simulated annual half-hourly NEE time series. A number of gap-filling models were applied to the simulated rEC time series, and the differences between gap-filled estimates and the synthetic NEE at different time scales (daily to inter-annually) were quantified. Using the approach, a thorough estimate of both the mean (bias) and spread (uncertainty) of error caused by different rEC scenario and gap-filling model application could be quantified. Both gap-filling model error and statistical metrics were used to compare and assess the performance and robustness of gap-filling models. The effects of rEC application were also compared with those simulated for standard-type gaps (operational problems, friction velocity filtering, etc.). We then evaluated the performance of a set of strategies hypothesized to improve gap-filling model performance and reduce overall error in time-integrated NEE estimates produced in rEC measurement programs.

3.3 Methods

3.3.1 Study sites and rEC operation

The focus of our study was the Turkey Point Flux Station (TPFS) -- an age-sequence of four planted, managed eastern white pine (Pinus *strobus* L.) forests, within a 10 km radius near the north shore of Lake Erie in Southern Ontario, Canada (42.71°N, 80.36°W). The forests use a naming convention relating to their year of establishment: TP39 (planted in 1939, 74 years old in 2013), TP74 (planted in 1974, 39 years old), TP89 (planted in 1989, 24 years old), TP02 (planted in 2002, 11 years old). All forests are actively managed and were planted on deep, sandy, Brunisolic Gray Brown Luvisolic soils (Gleyed at TP89; Presant & Acton, 1984) of former marginal agricultural lands (TP89, TP02) or cleared oak-savanna ecosystems (TP74, TP39). Detailed site measurements conducted by Peichl and Arain (2006) reported mean tree heights to be 2.8, 11.1, 13.1 and 21.8 m and tree densities to be 1683, 1325, 1633 and 421 trees ha^{-1} at TP02, TP89, TP74 and TP39. Maximum respective leaf area indexes were estimated as 1.0, 12.8, 5.9 and 8.0 m^2 m^{-2}. Complete, detailed description of site characteristics may be found in Peichl and Arain (2006), and Peichl et al. (2010a, 2010b).

At the three youngest sites (TP74, TP89, TP02), an rEC system was implemented between the years 2003-2007. The rEC system used an open-path eddy covariance (OPEC) system, comprising a 3-dimensional sonic anemometer (model CSAT-3, Campbell Scientific Canada Corp., Edmonton, AB), and an open-path IRGA (model LI-7500, LI-COR, Lincoln, NE, USA). Both components were affixed to a portable boom, which could be readily mounted to the top of triangular measurement towers located at each site at heights of 3, 15 and 16 m at TP02, TP89 and TP74, respectively. The system was connected to an electrical power supply and a data logger (model CR5000, Campbell Scientific Canada Corp., Edmonton, AB), which were housed in waterproof containers at the tower base. Data was logged as both high frequency (20 Hz) raw measurements, as well as calculated 10-minute average fluxes. The CO_2 concentration profile was estimated for each of these sites by supplementing the half-hourly average concentration measured by the OPEC IRGA with a measurement made at half the OPEC system height, using a separate closed-path IRGA (model LI-820, LI-COR).

The rEC system was moved between sites at semi-regular intervals of 2 weeks to 1 month. Since the system was rotated between three sites, the EC data coverage at a given site consisted of 2 weeks to 1 month of data, followed by a 1- to 2-month data gap, as the system was cycled through the other sites.

Rotation intervals were occasionally longer during the winter months, increasing the duration of measurement or data gap at the sites during such periods. The open-path IRGA was calibrated monthly, and internal chemicals were changed on a bi-annual basis. The use of rEC was discontinued in early 2008, as permanent CPEC systems were installed at TP74 and TP02, and the TP89 site was retired from all measurements due to resource constraints.

3.3.2   Data

A closed-path eddy covariance (CPEC) system has been operated continuously at TP39 from 2002 to the present, producing half-hourly fluxes of momentum, sensible heat, latent heat, and CO_2. The CPEC system consists of a 3-dimensional sonic anemometer (model CSAT-3, Campbell Scientific Canada Corp., Edmonton, AB), and a closed-path infra-red gas analyzer (IRGA; model LI-7000, LI-COR, Lincoln, NE, USA) housed inside of a temperature-controlled box . In order to estimate the air column CO_2 storage flux, ΔS_c, CO_2 was measured at three points in the canopy profile: at the top (via the CPEC IRGA), as well as at the tower midpoint and near the surface using separate closed-path gas analyzers (model LI-800, LI-820, LI-COR). Concentrations at each height were recorded as half-hourly averages. A full description of this CPEC system can be found in Arain and Restrepo-Coupe, (2005). Comparable CPEC systems were installed at TP74 (20 m) and TP02 (3 m) in 2008, and have been operated to the present.

Meteorological measurement towers were installed at all TPFS sites between 2002--2003. All towers were equipped with identical instruments to provide continuous measurement of meteorological and edaphic variables, according to Fluxnet-Canada guidelines (Fluxnet-Canada, 2003) . Meteorological variables were measured as half-hourly averages from the top of each tower. Measured variables included: air temperature and humidity (model HMP45C, Campbell Scientific Canada Corp., Edmonton, AB), wind speed and direction (model R.M. Young 05103-10, Campbell Scientific Canada Corp., Edmonton, AB), net radiation (model NR-LITE, Campbell Scientific Canada Corp., Edmonton, AB), down- and up-welling photosynthetic photon flux density (model Kipp & Zonen PAR-LITE, Campbell Scientific Canada Corp., Edmonton, AB). In addition, rainfall was measured using tipping-bucket rain gauges at TP39 (model 52202, R.M. Young, Traverse City, MI) and TP02 (model TR-525USW, Texas Electronics, Inc., Dallas, TX). Soil temperature (T_s) profiles were measured at two different locations using thermistor probes (model 107B, CSI) placed at depths of 2, 5, 10, 20, 50 and

100 cm. At the same two locations, volumetric water content (VWC) profiles were measured using water content reflectometers (model CS-615/616, CSI) inserted horizontally at depths of 5, 10, 20, 50 and 100 cm. The VWC profiles were used to calculate depth-weighted, upper root-zone averages (VWC_{30}) for the top 30 cm of each profile.

All flux and meteorological data were quality-controlled following data collection. Erroneous half-hourly meteorological data were removed using static thresholds, followed by more detailed visual scrutiny, which was implemented to remove only occasional, obvious data problems. Quality control measures for half-hourly carbon flux (F_c) data included Webb-Pearman-Leuning density corrections for the OPEC system (Webb et al., 1980), coordinate rotation (Tanner and Thurtell, 1969), and outlier removal using an automated detection algorithm method, similar to Papale et al., (2006). Net ecosystem carbon exchange (NEE) was calculated at all sites as the sum of eddy covariance-estimated CO_2 flux (F_c) and ΔS_c. ΔS_c was estimated following Barr et al., (2004), using the aforementioned half-hourly CO_2 concentration profile data. Across all measurement years, operational-type gaps (caused by equipment malfunction and quality-assurance measures) resulted in the loss of an average 10% of nighttime and 8% of daytime NEE half-hourly values. (Table 3.1).

In order to improve confidence that measured NEE reflected the true ecosystem exchange, a friction velocity threshold (u_*^{Th}) was applied to all nocturnal (PPFD < 15 μmol m^{-2} s^{-1}) NEE measurements, removing data from periods where CO_2 transport by non-turbulent means (i.e. horizontal and vertical advection attributed to density flows and breezes) was non-negligible (Aubinet, 2008; Barr et al., 2013; Gu et al., 2005; Papale et al., 2006) . Nocturnal half-hourly NEE values were removed when measured u_* was below thresholds estimated using the Moving Point Test u_*^{Th} determination method described by Papale et al., (2006), which estimated u_*^{Th} from the relationship between nighttime Net Ecosystem Exchange (NEE_n) and u_*. For this methodology, data from each year was stratified into four, equally-sized 3-month seasons (n_S=4: JFM, AMJ, JAS, OND), and then sub-stratified into six temperature classes (n_T=6) within each season. Each n_S*n_T strata was split into 20 equally-sized u_* classes, and the u_*^{Th} for a strata was estimated as the lowest u_* class that had an average NEE_n value within 99% of the average NEE_n at all higher u_* classes. Seasonal u_*^{Th} was calculated as the median of predicted values across the temperature substrata. To increase estimate accuracy, the estimation procedure was bootstrapped at the seasonal level, to produce 100 separate u_*^{Th} estimates, and a single, annual u_*^{Th} value was obtained from the median of the seasonal estimates. Estimates of u*Th obtained using this method ranged between 0.42 to 0.47 m s^{-1} for TP39, and 0.38 to 0.41 m s^{-1} for TP74.

Table 3.1: *Summary statistics of annual proportions of half-hourly NEE data availability following quality- assurance measures for both daytime and nighttime at TP39.*

	Available NEE Proportion	
	NIGHT	DAY
mean	0.901	0.922
std	0.0307	0.0189
max	0.947 (2006)	0.950 (2010)
min	0.8511 (2009)	0.889 (2008)

For the purposes of estimating the rEC-related error and uncertainty associated with time-integrated carbon exchange estimates, this study used four years (2008-2011) of continuous CPEC flux measurements from two sites in the age sequence (TP39, TP74). Data from these years and sites were selected to incorporate the largest period of simultaneous, continuous measurements at multiple sites, in which the ecosystems and measurement setup remained consistent throughout. A thinning operation at the TP39 forest in early 2012 altered the ecosystem by removing approximately 30% of dominant and co-dominant trees. As a result, data from this year was not included in the study, in order to remove error associated with anticipated changes in ecosystem carbon exchange rates and dynamics. Flux data collected from the TP02 site was not included in these analyses, as a variety of instrument problems resulted in a number of long (weeks to a month) flux and meteorological data gaps during 2008 and 2009. Preliminary tests on this data suggested that such gaps would introduce considerable error into analyses, by affecting both the generation of synthetic data, and the application of gap-filling models to provide time-integrated estimates. The implications for excluding the TP02 site data from these analyses will be discussed later.

3.3.3 Synthetic NEE time series

To investigate the four study objectives in a comprehensive manner, synthetic half-hourly NEE time series were created for each year and used in analyses. Ensembles of synthetic series were generated by combining a single modeled NEE time series (NEE_m) with multiple iterations of random error time series, in order to simulate the "true signal + noise" that is measured by EC systems. These gap-free synthetic NEE time series were used in analyses instead of the measured values, as they offered a number of analytical advantages: First, the continuous nature of the synthetic series allows gaps to be inserted at any desired location, so that an ensemble of possible gap distributions can be assessed, and their effects be quantified. Second, the synthetic series represent an ensemble of measurement noise realizations. As has been shown by Richardson and Hollinger (2007), the magnitude and distribution of inherent measurement random error affects gap-filling model performance, and thus, final gap-filled values. Therefore, using an ensemble of synthetic series generates a distribution of gap-filled values, from which confidence intervals can be estimated. Finally, using complete time series allows for NEE estimate uncertainty to be partitioned and attributed to causes such as: random measurement noise, normal operational data gaps, and rEC data gaps.

For both study sites, all years (2008-2011) of NEE and primary meteorological variable data (PPFD, T_s T_a, VPD, VWC_{30}), were assimilated into an artificial neural network (ANN), in order to produce continuous and noise-free estimates of NEE_m (Figure 3.1, top). The ANN was developed and implemented using the neural network toolbox in MATLAB (The Mathworks Inc., Natick, MA). A general two-layer feed-forward network was used, which contained a sigmoid transfer function in a 30-node hidden layer, and a linear transfer function in the output layer. Post-hoc analyses of the ANN output supported its use, as it fulfilled the two necessary criteria for implementation in our study: a) low random and bias error in comparison to the original NEE time series, and b) model errors that were poorly-correlated to those produced by the gap-filling models being evaluated. Satisfying the latter criteria helps to ensure that gap-filling metrics obtained through MC analyses reflect the true performance of the models, and not their functional similarity to the ANN.

Time series of half-hourly random error were generated using an approach similar to that used by Richardson and Hollinger (2007). Since random error associated with eddy covariance measurement is heteroscedastic (see Hollinger and Richardson, 2005; Richardson et al., 2006), it was first necessary to estimate the potential dispersion associated with each half-hourly measurement, in order to generate random error time series that are representative of the actual nature of error. Measurement random error was approximated using the model residual approach of Richardson et al., (2008), where half-hourly error was estimated as the residuals between the measured NEE and the predicted NEE values generated using a relatively simplistic non-linear gap-filling model. In contrast to the aforementioned studies -- which determined that error was best described by a Laplacian (double-exponential) distribution -- our analyses were consistent with those of Lasslop et al., (2008), which showed EC measurement error to be normally distributed and heteroscedastic, with a standard deviation that scaled with the expected magnitudes of the component fluxes of ecosystem
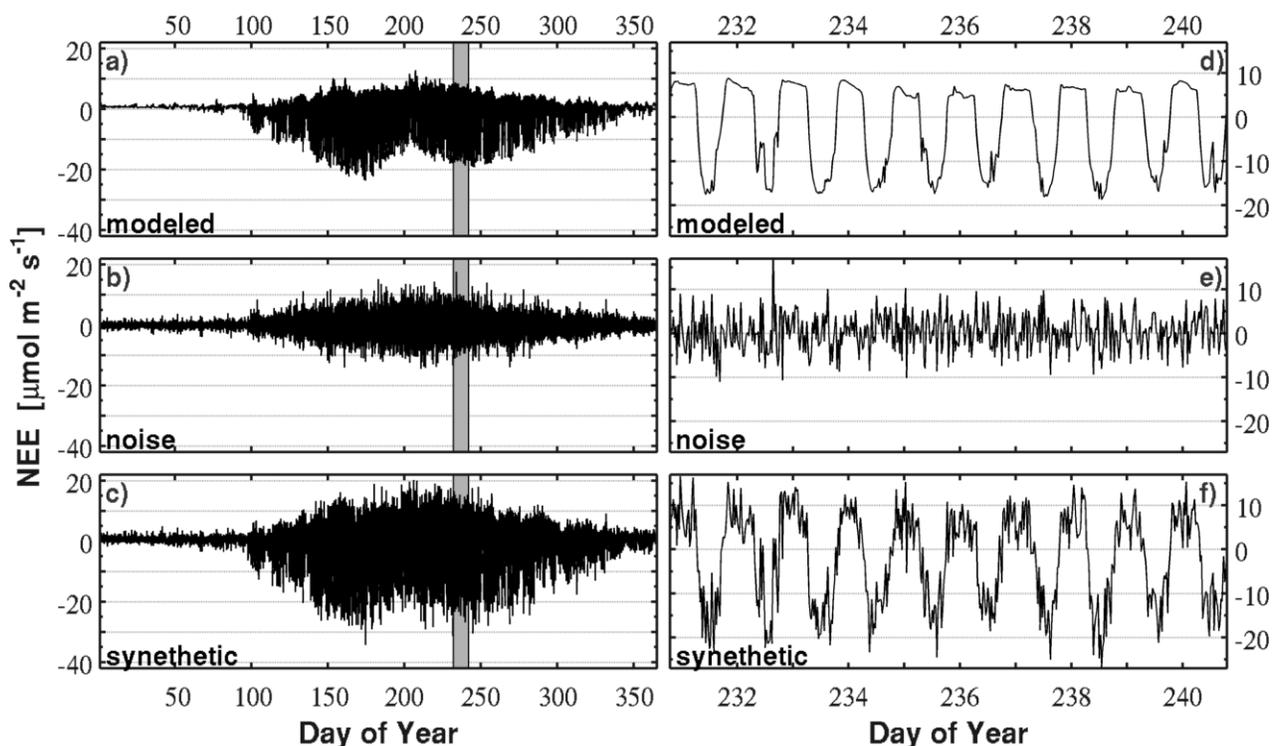
Figure 3.1: *A sample of the components used to create synthetic NEE time series. An annual synthetic NEE time series (shown in panel c) is created by adding a randomly-generated measurement noise time series (panel b) with the ANN-predicted 'modeled' data (panel a). The regions highlighted in grey are shown in more detail in panels d through f, respectively.*

respiration (RE) and gross ecosystem productivity (GEP). Our methods also differed from previous studies in the parameterization and prediction of EC measurement error. As a modification to the original method, model residuals were stratified into 20 quantiles across the ranges of observed PPFD and T_s, and the standard deviation of error values were calculated for each PPFD x T_s strata.. A multilinear regression was run to fit σ_i values to the mean PPFD and T_s of the strata, and this relationship was used to generate estimates of σ_i across all half-hours (Figure 3.2). This approach provided considerable improvement to the single-variable scaling approaches methods presented in the literature, as it better related error to the combined magnitude of component fluxes (RE, GEP), rather than the magnitude of measured NEE, which is prone to issues of equifinality.

A total of 100 half-hourly NEE error time series were generated for each year and site by drawing random values from a normal distribution with standard deviations provided by σ_i (Figure 3.1, middle). Each error time series was added to NEE_m to create a total of 100 synthetic half-hourly NEE data sets for each year and site (Figure 3.1, bottom).
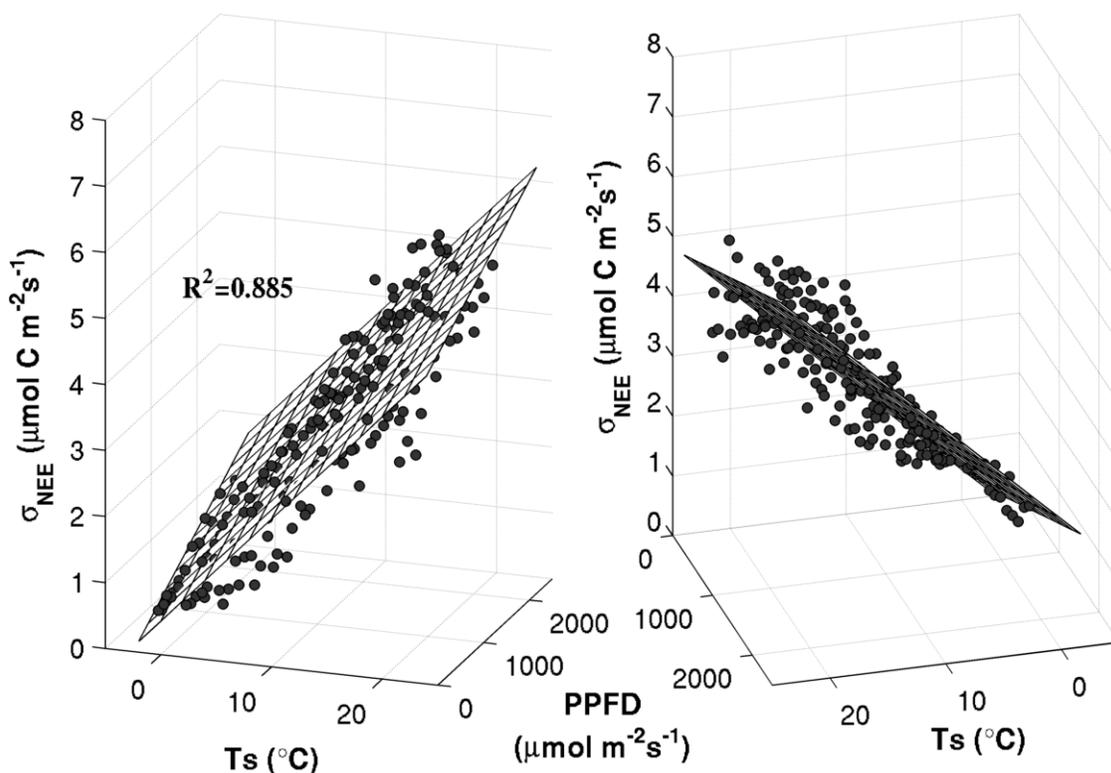
Figure 3.2: *Two perspectives of the multi-linear fit of bin-averaged NEE measurement error (σ_{NEE}) to soil temperature (T_s) and photosynthetic photon flux density (PPFD) for data measured at TP39 between the years 2008-2011.*

3.3.4   rEC scenarios

The effect of rEC application on NEE estimation bias and uncertainty was investigated by imposing each of 10 different rEC measurement scenarios on all synthetic time series. The scenarios (Figure 3.3), reflect the implementation of rEC schedules that vary across three characteristic dimensions: number of sites in rEC rotation, length of measurement cycle, and timing of measurement periods and gaps throughout the year. Scenarios 1a and 1b simulate the rotation of an rEC system between two sites at monthly intervals, which varies only in terms of the timing of measurement/gap periods. A similar approach is used for Scenarios 2a and 2b, though the measurement interval is reduced to 15 days. Scenarios 3 and 4 simulate a three-site rEC rotation operating in different phases, cycled either monthly (Scenario 3) or every 15 days (Scenario 4).

Half-hourly template files were created for each of the 10 rEC scenarios by creating column vectors consisting of ones during half-hours where the rEC schedule indicated measurements to be taken, and NaNs otherwise. rEC scenarios were imposed on the data by multiplying synthetic NEE time series with the template file, effectively removing NEE values for undesired periods.

The amount of half hourly NEE data retained after rEC scenario application was expressed on an annual basis as the proportion of remaining half-hourly data points (n_R) to the total number of half-hours in a year (n_T). This ratio is termed the available data fraction and is expressed by:

$$\phi = \frac{n_R}{n_T} \quad .$$

(3.1)

This measure spans between 0 (where no data points are retained), to a maximum of 1, and is used to express the general available fraction of all data (ϕ), as well as data available specifically for RE and GEP model parameterization (ϕ_{RE} and ϕ_{GEP}, respectively).

Generally, 15 to 30% of annual data remained after rEC scenarios were applied to the u_*^{Th}-filtered synthetic data (Figure 3.4). Typical ϕ values were approximately 0.3 for scenarios using a 2-site rotation (rEC scenarios 1 and 2), and 0.2 for those implementing 3 sites (rEC scenarios 3 and 4). Due to the application of u_*^{Th} filtering to nighttime data, the fraction of available RE-parameterizable data was considerably lower (ϕ_{RE} range of 0.11 to 0.24) than for GEP (ϕ_{GEP} of 0.23 to 0.44). The removal of nighttime data by u_*^{Th} application was more substantial for TP74 than for TP39.
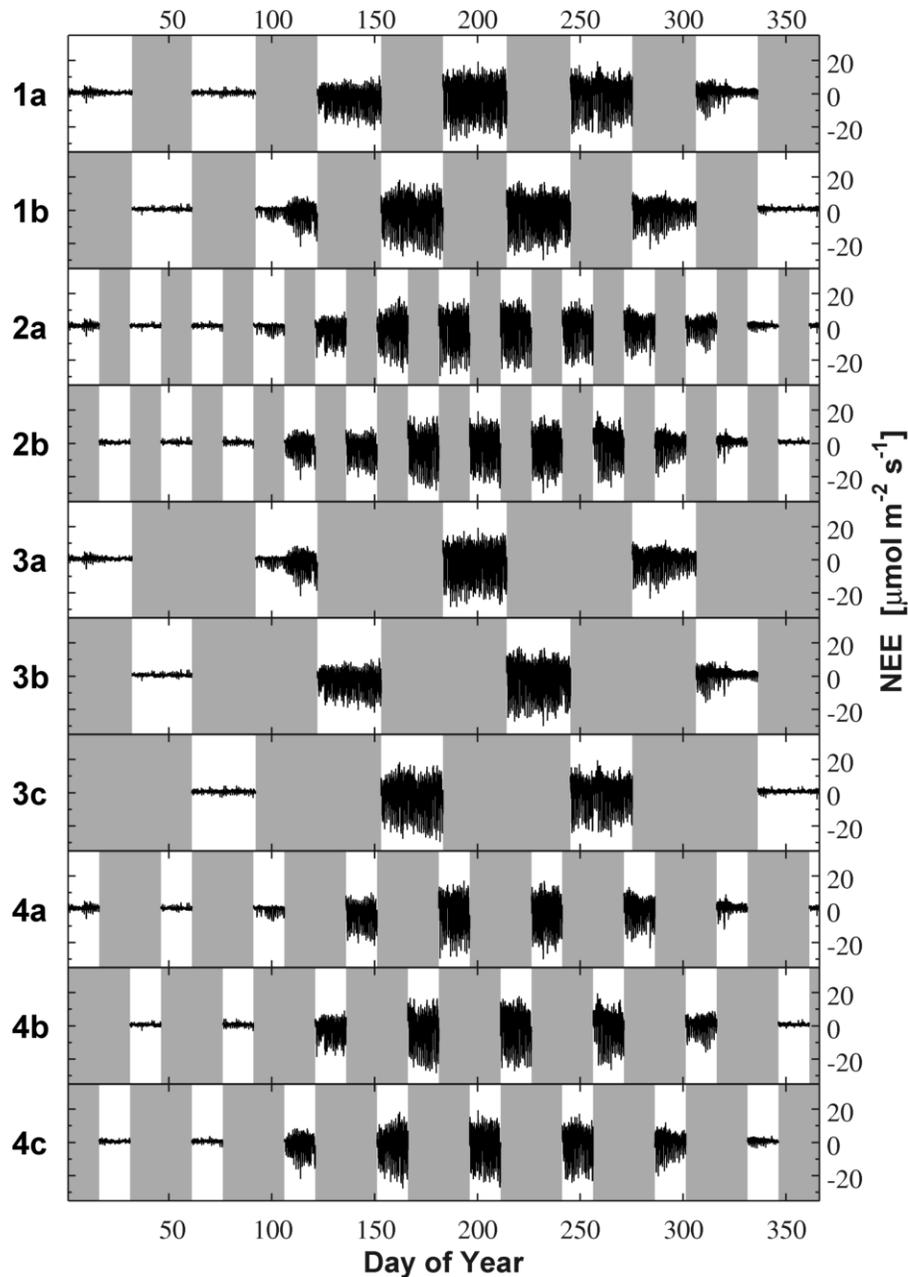
***Figure* 3.3:** *The ten rEC scenarios tested in this study, superimposed on a single year of synthetic NEE. The numbers associated with each scenario denote consistent number of sites and rEC rotation period, while sub-scenarios are indexed according to a change in rEC gap timing. Scenario 1 (a and b): two sites, one month; Scenario 2 (a and b): two sites, 15-days; Scenario 3 (a, b and c): three sites, one month; Scenario 4 (a, b and c): three sites, 15-day measurement length. Data for TP39 is shown.*
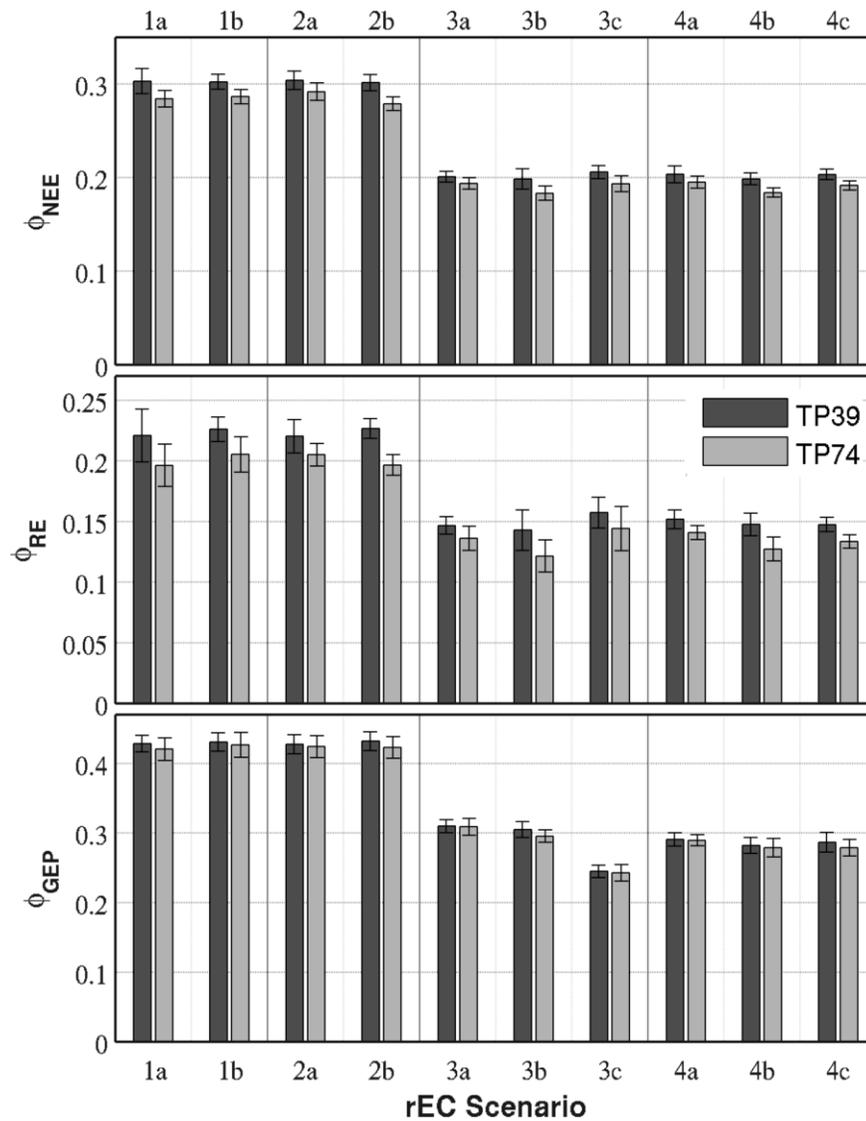
Figure 3.4: *Mean (bar) and standard deviation (error lines) of available data fraction (ϕ) of all NEE, RE- and GEP-parameterizable data for each rEC scenario applied to simulated (artificially-gapped) data for both TP39 and TP74.*

3.3.5   Operational and filtering gap scenarios

Acknowledging that the simulated rEC schedules cause data gaps that are extreme in two aspects – overall proportion of gaps, and temporal clustering – we applied two additional types of gap scenarios to the synthetic data, with the intention of partitioning error in rEC to each of these characteristics. The scenarios simulated gaps that result from more common data-loss causes, in that gaps are distributed more evenly in time for a given available data fraction (ϕ, see Eq. 3.1).

The first of the additional gap scenarios was constructed to remove data from NEE time series in a manner that reflects the nature of operationally-induced gaps at our sites. Such "operational gaps" may result from occurrences such as short-term equipment failure, data spikes, stationarity test failure, calibrations, etc.). A total of 100 operational gap scenarios were generated, with each scenario removing a larger proportion of data than the previous. The resulting scenarios spanned data fractions (ϕ) between 0.95 and 0.1, while preserving the day-to-night ratio of gaps observed in measured NEE (Table 3.1). Gaps were inserted randomly into the template files (represented as NaNs), while the length of each gap was determined by randomly drawing from a population containing a list of all actual operational gap lengths experienced at TP39 between 2004 and 2011. This operation was repeated continually until the desired ϕ value was reached. Daytime and nighttime gaps were removed separately, in order to preserve the distributions of daytime and nighttime gap lengths. Gap lengths were assigned in this manner and not by sampling from parametric distributions (e.g. negative binomial distribution), since they were not able to appropriately recreate the population distribution of gap lengths observed in NEE data.

A second scenario was generated to simulate the effect of data removal by flux filtering operations, such as footprint and friction velocity threshold filtering. In a similar fashion to the operational gaps, 100 "filtering gap" scenarios were created to span data fractions between 0.95 and 0.1. Data gaps were inserted randomly into template files, though the proportion of retained respiration-parameterizable data to overall ϕ was controlled, in order to preserve the relationship between these two values that occurs as a result of application of these filtering processes on real NEE data.

3.3.6   Gap-filling models

Synthetic NEE time series that were subjected to rEC, operational or filtering gap scenarios were filled using three different commonly-used gap-filling methods: The Fluxnet-Canada Research Network method (NLR-FC) described by Barr et al. (2004); a modified version of the nonlinear estimation model used by Richardson et al. (2007) for the Howland Research Forest data (NLR-HL); and the Reichstein et al. (2005) Marginal Distribution Sampling (MDS) method. These methods were selected due to their common use in studies on temperate forests, and history of use for gap-filling purposes at this site (e.g. Arain and Restrepo-Coupe, 2005; Peichl et al., 2010b). Each method is described in further detail below:

3.3.6.1   NLR-FC

In the NLR-FC method, measured RE is assumed to be equal to NEE during periods when GEP is zero (i.e. at night and during daytime periods when both air (T_a), and 5 cm soil (T_s) temperatures are less than 0°). The NLR-FC method fits RE to an empirical logistic relationship with T_s as:

$$\text{RE} = \frac{r_1}{1 + \exp[r_2(r_3 - T_s)]} \, r_w(t) \quad , \tag{3.2}$$

where RE is model-estimated ecosystem respiration, and r_1, r_2, and r_3 are fitted empirical parameters. $r_w(t)$ is an additional fitted parameter that varies according to the slope of a linear regression (forced through the origin) between modeled and measured RE in each 100-point moving window. The time-varying parameter corrects for the presence of sustained (autocorrelated) biases in modeled estimates of RE.   Values of RE are used in the place of gaps in the annual measured RE time series (caused by equipment malfunction, quality-assurance testing, filtering, etc.), as well as to estimate respiration during periods of non-zero GEP (when RE cannot be directly measured).

GEP was determined as the difference between gap-filled RE and measured NEE, and subsequently set to zero for all nighttime periods and daytime half hours when $T_a, T_s < 0$. A rectangular hyperbolic, Michaelis-Menten relationship is used to model GEP as:

$$GEP = \frac{\alpha Q_d \beta}{\alpha Q_d + \beta} \, p_w(t) \quad , \tag{3.3}$$

where Q_d is down-welling photosynthetic photon flux density (PPFD, μmol m^{-2} s^{-1}), and the fitted parameters α and β represent the quantum yield and photosynthetic capacity, respectively. To correct for sustained autocorrelated bias in model predictions, $p_w(t)$ is a time-varying parameter that varies according to the slope of a linear regression (forced through the origin) between modeled and measured GEP for each 100-point moving window.

3.3.6.2   NLR-HL

The NLR-HL gap-filling method is a modified version of the model described by Richardson et al., (2007), which was later modified and applied to TP39 by Peichl et al. (2010). RE measurements are identified from NEE using the same conditions as NLR-FC (NEE nighttime and daytime with T_a,$T_s <$ 0). RE is modeled as a function of T_s and top-30 cm integrated volumetric water content (VWC$_{30}$) according to the relationship:

$$RE = R_{10} \times Q_{10}^{\frac{(T_s - 10)}{10}} \times f(VWC_{30}) \quad , \tag{3.4}$$

where R_{10} and Q_{10} are fitted temperature response parameters that describe the relationship between RE and T_s. f(VWC$_{30}$) is a sigmoidal function that characterizes the role of VWC$_{30}$ in modifying the temperature response of RE as:

$$f(x) = \frac{1}{[1 + \exp(\theta_1 - \theta_2 x)]} \quad . \tag{3.5}$$

θ_1 and θ_2 are fitted parameters that allow this term to range between [0,1] as a function of the independent variable x (VWC$_{30}$ in this case), thus actings as a scaling function on the Ts-RE relationship. Including a VWC$_{30}$ control on the RE model was found to provide a statistically significant improvement to model performance during periods of low VWC (typically mid-summer, data not shown). The two-parameter Q_{10} temperature response model was chosen instead of three-parameter functions (e.g. logistic or Lloyd and Taylor, 1994), to increase model parsimony and address issues of over-fitting and equifinality (Richardson et al., 2007).

GEP was estimated in the same way as for NLR-FC (RE - NEE; zero during nighttime periods, and daytime half hours when $T_a, T_s < 0$). GEP is modeled by adding additional controlling variables to the formula used in the NLR-FC method (eq. 3.3) as:

$$GEP = \frac{\alpha Q_d \beta}{\alpha Q_d + \beta} \times f(Ts) \times f(VPD) \times f(VWC_{30}) \quad . \tag{3.6}$$

The first term in eq. 3.6 defines a Michaelis-Menten relationship between Q_d and GEP. The second through fourth terms describe sigmoidal-type [0,1] scaling responses of GEP to T_s, atmospheric vapour pressure deficit (VPD) and VWC_{30}, respectively. In contrast to Richardson et al., (2007), a scaling response to T_a was not applied in our model, since parameter analyses indicated that T_s and T_a were strongly correlated, and model explanatory power was not affected by the exclusion of T_a.

3.3.6.3   MDS

The Marginal Distribution Sampling approach of Reichstein et al. (2005), which builds upon methods presented by Falge et al. (2001), seeks to fill gaps in NEE time series by utilizing covariance between fluxes and meteorological variables, as well as an auto-correlation of fluxes with respect to time. The method applies a heuristic approach, where NEE gaps are filled with the average of NEE measurements recorded under similar environmental conditions (global radiation, R_g; T_a, T_a; VPD), within a specified number of days surrounding the occurrence of the gap. The algorithm defines a hierarchy of condition sets for environmental conditions and time difference that should be applied to find valid NEE measurements. Initial preference is given to the average of all NEE measurements made using data collected within 7 days of the gap, where differences in environmental conditions are small ($\Delta R_g < 50$ W m^{-2}, $\Delta T_a < 2.5$ K, $\Delta VPD < 5$ hPa). If no suitable NEE measurements meet the criteria, the search continues by broadening the time window of consideration and/or loosening environmental condition restrictions until an estimate for the half hour in question can be made. A complete description of the algorithm may be found in Appendix A of Reichstein et al. (2005). This method was modified for the purposes of this study by applying the NEE estimation algorithm to all half hours, rather than only to those in which NEE gaps existed. For half hours where measured data existed, NEE values were temporarily removed prior to algorithm implementation, in order to produce the same gap-filling conditions for all data. The purpose of this modification was to create a complete half-hourly time series of model-predicted NEE, which could be used to assess gap-filling model performance across all data.

3.3.7   rEC uncertainty mitigation strategies

In addition to quantifying the effect of rEC operation on NEE estimates, we formulated and evaluated three strategies to reduce the potential negative consequences of such a measurement program on these time-integrated values. The strategies differed between operational and data-based implementations, but shared a common purpose of increasing the information available to the gap-filling models during periods of rEC-induced gaps. It was hypothesized that providing information to the gap-filling models during these periods would improve model parameterization and performance, thus reducing bias and uncertainty of time-integrated estimates.

The first strategy sought to improve the performance and robustness of the NLR-HL model by pooling all years of data together, and parameterizing the gap-filling model based on a single functional relationship between environmental conditions and each the component fluxes (RE, GEP). This "pooled" strategy, used and presented first by Peichl et al., (2010b), extends the approach of NLR-HL, where single response functions of both RE and GEP to secondary-effect variables (VWC_{30} for RE, Ts, VPD, VWC_{30} for GEP) are derived at-once, according to Richardson, 2007). The aim of data pooling is to increase the data available for model parameterization, thus avoiding model-fitting problems that arise from a lack of sufficient data points and poor coverage of measurements throughout the entire range of controlling environmental variables.

The second strategy was implemented by filling rEC-imposed NEE data gaps with regression-corrected NEE values from another site in the rEC rotation that is actively measuring NEE during that period. This "filled" approach represents one that might be applied in cases where rEC sites are in close proximity (and thus, experience similar environmental conditions), and are functionally-similar – requirements that are met for our research setting. To approximate such a cross-site filling, gaps in synthetic data from the 'target' site (either TP39 or TP74) were filled from measured data from the 'source' (other) site. These two time series of NEE data were linearly regressed, and NEE values from the source site were scaled by this linear factor prior to insertion in the target data file. The true, measured NEE data from the source site was used, in order to avoid any erroneous correlations between synthetic data due to co-generation by similar ANNs. Thus, results of this test are likely to represent the 'worst-case' expectation for this approach. The filled 'target' NEE data set was then used in standard gap-filling operations, as described above.

A third tested strategy was the use of annually alternating rEC schedules, where the system was moved

in a manner that ensured that any given time of year did not experience rEC gaps on consecutive years. Since the NLR-HL gap-filling method leverages multiple years of information to inform model parameterization, the performance of these models are affected by the timing of rEC gaps over successive years. Whereas the original rEC scenarios assumed identical rotation through the study sites for all years, this "alternating" approach shifts the measurement/gap periods for each successive year, so that at least one measurement from each time of year is included in model parameterization. To investigate the potential improvement afforded by providing NEE data from all times of the year, four additional rEC scenarios were generated and tested: Each of these scenarios represents an amalgamation of the original scenarios s1, s2, s3 and s4, where the sub-scenarios of each are alternated for each year.

3.3.8   Monte Carlo simulations and analyses

Factorial Monte Carlo simulations were used in this study to quantify the effects of the aforementioned data treatments on time-integrated NEE estimates. Given the random nature of NEE measurement error, and the inconsistent interaction of this error with gap-filling model outputs (Richardson and Hollinger, 2007), this approach provides a means of comparing and evaluating different data treatments using an appropriate number of input data samples.

Evaluations were performed by inserting gaps into synthetic NEE data to simulate the desired measurement scenario, and subsequently applying each of the gap-filling models to the simulated data to produce continuous NEE estimates. For evaluation of the rEC scenarios and the proposed mitigation strategies, each of the 100 four-year synthetic NEE time series were first subjected to simulated operational gaps (see section 3.3.5), and were then u_*^{Th}-filtered according to the measured u_* values (using u_*^{Th} values of 0.5 and 0.4 m s^{-1} for TP39 and TP74, respectively). Following this, each of the 10 rEC keyfiles were applied in a factorial fashion to each synthetic time series, in order to simulate the effect of rEC system operation. Data from each combination of synthetic data and rEC scenario were then used as input data to the gap-filling models described in section 3.3.6, resulting in a total of 12000 annual time series of gap-filled, half-hourly NEE. These values were summed at daily, weekly, monthly, seasonal and annual increments. An additional step was necessary for the evaluation of the "filled" mitigation strategy: gaps in the NEE time series (following application of the rEC key file), were filled by a linear regression-adjusted value from the other site, where data was available, prior to

execution of the gap-filling models.

For operational and filtering gap tests, each of the 100 gap scenarios (representing a given ɸ value), was applied to 10 randomly-assigned synthetic NEE time series. The number of synthetic NEE data sets used for each gap scenario was reduced to 10 in the interest of computational efficiency, after initial tests showed this to be an acceptable amount, given that the gap scenarios are closely spaced along ɸ values. Results of these factorial simulations yielded 12000 annual gap-filled NEE time series, which were time-integrated as described above.

All Monte Carlo simulations were executed on remote computing clusters provided by the Shared Hierarchical Academic Research Computing Network (SHARCNET). MATLAB functions were compiled as standalone C-language programs (using function mcc.m in MATLAB), which were then submitted and executed using the SHARCNET computing environment.

Simulation results were assessed using a number of evaluation metrics. Time-integrated NEE sums (weekly, monthly, annually, etc.) produced for each gap-filled simulation run (NEE_f), were expressed as a deviation from the sum of modeled NEE (NEE_m) over the same period, according to $\Delta NEE = NEE_f - NEE_m$. Comparisons between rEC scenarios, mitigation strategies and/or gap-filling model results were made by assessing the average bias (via central tendency) and uncertainty (via 95% confidence interval) of all ΔNEE within a treatment group of interest. Due to the presence of outliers in the results, both statistics were calculated non-parametrically, as the median and 95% quantile range were used for each, respectively.

The ability of gap-filling models to predict the point-to-point variability in NEE was evaluated by comparing values of absolute root mean square error (RMSE), which was calculated for each year of gap-filled half-hourly data as:

$$RMSE = \sqrt{\left(\frac{1}{N-2}\sum (p_i - o_i)^2\right)} \quad, \tag{3.7}$$

where p_i are the gap filling model-predicted half-hourly NEE estimates, o_i are the observed NEE values, and $\bar{p}$ and $\bar{o}$ are mean values of these respective time series. RMSE was used as a complementary metric to ΔNEE, as it provides a relatively robust indcation of the magnitude and distribution of individual errors in gap-filling model estimates. Using (initially) gap-free modeled and synthetic data instead of real observations provided an opportunity to expand RMSE calculations to assess model goodness-of-fit for three different data subsets. These were:

a) "internal" RMSE values, where p_i and o_i consisted of all half-hours where observed data was available;

b) "external" RMSE, calculated using all half-hours where gaps existed in the observed data; and

c) "modeled" RMSE, calculated using all half-hourly predicted NEE for p_i, and using all original modeled (noise-free; NEE_m) in place of o_i.

Used together, these RMSE estimates provided complementary information: the internal value represented the only metric that would be available during genuine gap-filling model application (where values for data gaps would be unknown), while the external value provided an assessment of the model's performance for only data points where information is unavailable. The modeled value assessed the ability of each model to simulate the true NEE signal in the presence of measurement noise and data gaps.

Box plots (box and whisker plots) were used throughout this study to demonstrate the distribution of time-integrated NEE estimates and RMSE across different gap-filling models and rEC scenarios. This study uses standard box plot notation to display group distributions. The median of all values within a group is represented by a horizontal line, while triangular markers above and below the median indicate the standard error on the median estimate, calculated as:

$$SE = q_2 \pm 1.57 (q_3 - q_1) / \sqrt{(n)} \quad , \tag{3.8}$$

where n is the number of values in the group, q_1 and q_3 indicate the first and third quartile of the data, q_3 - q_1 is the interquartile range, and q_2 is the median.

Boxes around the median indicate the extents of the upper and lower quartiles. The extent of the whiskers are calculated as:

$$\begin{aligned} w_{low} &= q_1 - 1.5 (q_3 - q_1) \\ w_{high} &= q_3 + 1.5 (q_3 - q_1) \end{aligned} \quad . \tag{3.9}$$

Outliers – indicated as circles outside the extent of the whiskers – were not excluded from analyses, since these values indicated notable failures of gap-filling models.

3.4 Results

3.4.1   Annual sums and interannual variability of synthetic NEE

As outlined in section 3.3.8, annual NEE values obtained from synthetic time series were used as a baseline to quantify error in gap-filled NEE derived after application of rEC-, filtering- or operational-type scenarios (Table 3.2). Though both forests were consistent annual sinks of carbon, the magnitude of uptake was greater at TP74 (range of -311 to -486 g C m^{-2} y^{-1}), than at TP39 (range of -188 to -293 g C m^{-2} y^{-1}). This difference is consistent with well-established coniferous forest lifecycle carbon exchange patterns, where productivity in intermediately-aged forests (such as TP74) increases more rapidly over this period than does respiration, which generally approaches productivity values at maturity (Coursolle et al., 2012). The 95% confidence interval (twice the uncertainty bounds) of these estimates were comparable between sites, ranging between 26 and 34 g C m^{-2} y^{-1}; this uncertainty is due to the inherently noisy nature of NEE measurements. On an interannual scale, trends for TP39 and TP74 were very similar for years 2008, 2009 and 2011, but were entirely opposite for 2010, where the extreme maximum annual NEE$_m$ value for this period occurred at TP39 (-188 g C m^{-2} y^{-1}), while TP74 experienced its extreme minimum value (-486 g C m^{-2} y^{-1}). This discrepancy is attributed to differing responses of the forests to a mid growing-season dry event. During this period, productivity was predominantly negatively affected at the TP39 forest, while respiration was predominantly suppressed at TP74.

Table 3.2: Median and uncertainty bounds of annual NEE estimates at TP39 and TP74, as derived from synthetically produced data. Uncertainty bounds are derived from the 95% confidence interval on estimates.

	NEE (g C m^{-2} y^{1})			
Site	**2008**	**2009**	**2010**	**2011**
TP39	-242 (+/- 15)	-256 (+/- 13)	-188 (+/- 17)	-293 (+/- 17)
TP74	-311 (+/- 14)	-335 (+/- 13)	-486 (+/- 15)	-391 (+/- 13)

3.4.2 Effects of rEC and operational gaps on time-integrated NEE estimates

As a first step of investigation, we compared annual gap-filled NEE error (ΔNEE) of the rEC scenarios to those obtained for operational and filtering gap simulations (Figure 3.5). To facilitate comparison, ΔNEE for each scenario or simulation was plotted as a function of its available data fraction value (ϕ).

Analyses of results from operational and filtering gap simulations showed minimal bias error (within +/- 5 g C m^{-2} y^{-1}; Figure 3.5, panel c) associated with these types of gaps, and a general increase in NEE uncertainty (greater ΔNEE spread) with reducing data availability (Figure 3.5, panel b). The 95% confidence interval (CI_{95}) for operational gap scenarios was slightly larger for operational gaps (range of 34 to 150 g C m^{-2} y^{-1} for ϕ between 0.9 and 0.1) than for filtering gaps (range of 33 to 140 g C m^{-2} y^{-1}).

To establish a baseline relationship between data availability and annual (non-rEC) NEE uncertainty, we combined the operational and filtering gap results, and fit trend lines through the relationship with ϕ. A second-order polynomial showed good fit to the data, and estimated the relationship as:

$$CI_{95} = 131.29(1-\phi)^2 - 6.35(1-\phi) + 36.47, \qquad (R^2 = 0.973) \tag{3.10}$$

A simpler, linear relationship was also generated, though was valid only over the interval of $0.2 \leqslant \phi \leqslant 0.9$:

$$CI_{95} = 101.9(1-\phi) + 19.1, \qquad (R^2 = 0.963) \tag{3.11}$$

The slope of the linear model suggests that for conventional NEE operation (over the appropriate range of ϕ), uncertainty increased at a rate of 1.02 g C m^{-2} y^{-1} per 1% of data removal. At ϕ values less than about 0.2, however, uncertainty increased at an accelerated rate.

Results from rEC simulations at both TP39 and TP74 also showed increasing uncertainty of NEE estimates with ϕ reduction, and the magnitude of uncertainty for a given ϕ value was greater than those experienced for operational and filtering gap simulations. Uncertainties for separate rEC scenarios varied widely (Figure 3.5, panel b), and were 45 to 328 % larger than those established for operational/filtering gaps at a given ϕ value. Much less variability in uncertainty was observed for the 2-site rEC rotation scenarios, in comparison to those simulating 3-site schedules, as the rate of

uncertainty increase with data loss appeared to be greater for rEC scenarios than for simulations using standard-type gaps. Median NEE biases (Figure 3.5, panel c) showed variation among rEC scenarios. Bias for most scenarios was within +/- 10 g C m^{-2} y^{-1}, and all but two extreme values were within +/- 25 g C m^{-2} y^{-1}. The spread of rEC simulation NEE biases was greater for 3-site rotation scenarios (lower ϕ) than the 2-site cases. These results demonstrate that although NEE estimate error for a given ϕ was consistently greater for rEC application than for traditional-type gaps, there is considerable variation among rEC scenarios in the magnitude of this error.

To further investigate these differences, we compared NEE estimate uncertainty and bias between each rEC scenario. In terms of annual NEE estimates, total uncertainty (CI$_{95,TOT}$; Figure 3.6, panel b) was generally less for simulations of 2-site rEC application (scenarios 1 and 2; CI$_{95}$ range of 134 to 223 g C m^{-2} y^{-1}), as compared to the 3-site simulations (scenarios 3 and 4; 162 to 433 g C m^{-2} y^{-1}). Uncertainty was also generally less for 15-day rotation schedules (s2 and s4) than for corresponding month-long application (s1 and s3). For both sites, uncertainty was highest for the three iterations of scenario 3, corresponding to the rEC application experiencing the most, and longest gaps. Results also showed a general reduction in NEE estimate uncertainty for scenarios structured to begin measurements later in the year (scenarios 1b, 2b, 3b, 4b) in comparison with scenarios where measurements began on the first day of the year. Overall, uncertainty was less for TP39 than for TP74, though the differences between the two sites were minimal for most scenarios, with the exception of scenario 3, where NEE estimate uncertainty was as much as 50% greater at TP74. These results imply that rEC scenario characteristics such as number of sites, rotation length and rotation timing influence the resulting error on rEC NEE estimates.

To provide insight into the components that contribute to the uncertainty values shown in Figure 3.6, total uncertainty (CI$_{95,TOT}$) for each scenario and site was decomposed into two elements: uncertainty due to the effect of random error associated with ϕ (CI$_{95,ϕ}$), and uncertainty due to the temporal structure of rEC gaps (CI$_{95,rEC}$). This was done by first estimating the former according to Eq. 3.11, and then subtracting this value (in quadrature) from CI$_{95,TOT}$. Results showed that the contribution by CI$_{95,rEC}$ ranged between a factor of 1 to 3.5 times larger than CI$_{95,ϕ}$, and was minimized at both sites with the application of rEC scenario 2b (Tables 3.3 and 3.4).
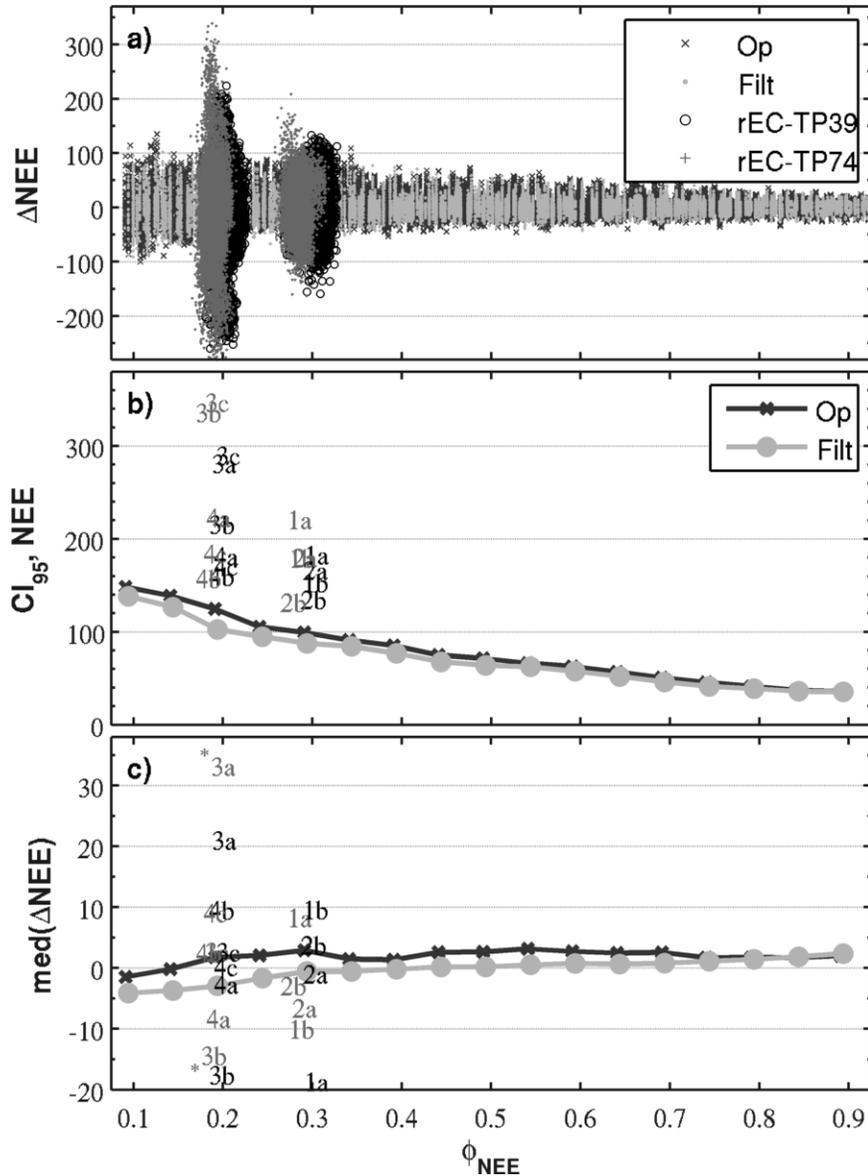
Figure 3.5: *Distribution, uncertainty and bias of annual ΔNEE for all rEC simulations at TP39 and TP74 as a function of ϕ. Results are presented alongside values obtained for operational and filtering gap scenarios. The compiled output of each simulation run is shown is panel a), while the 95% confidence interval and median of these values are shown for each individual rEC scenario in panels b) and c), respectively. TP39 results are shown in black text, while TP74 in grey. Bin-averaged values for operational and filtering simulations are shown in panels b) and c). Asterisks accompanying labels in panel c) indicate values well above (41.7) or below (-75.5) the range of the figure. NEE is reported in g C m^{-2} y^{-1}*
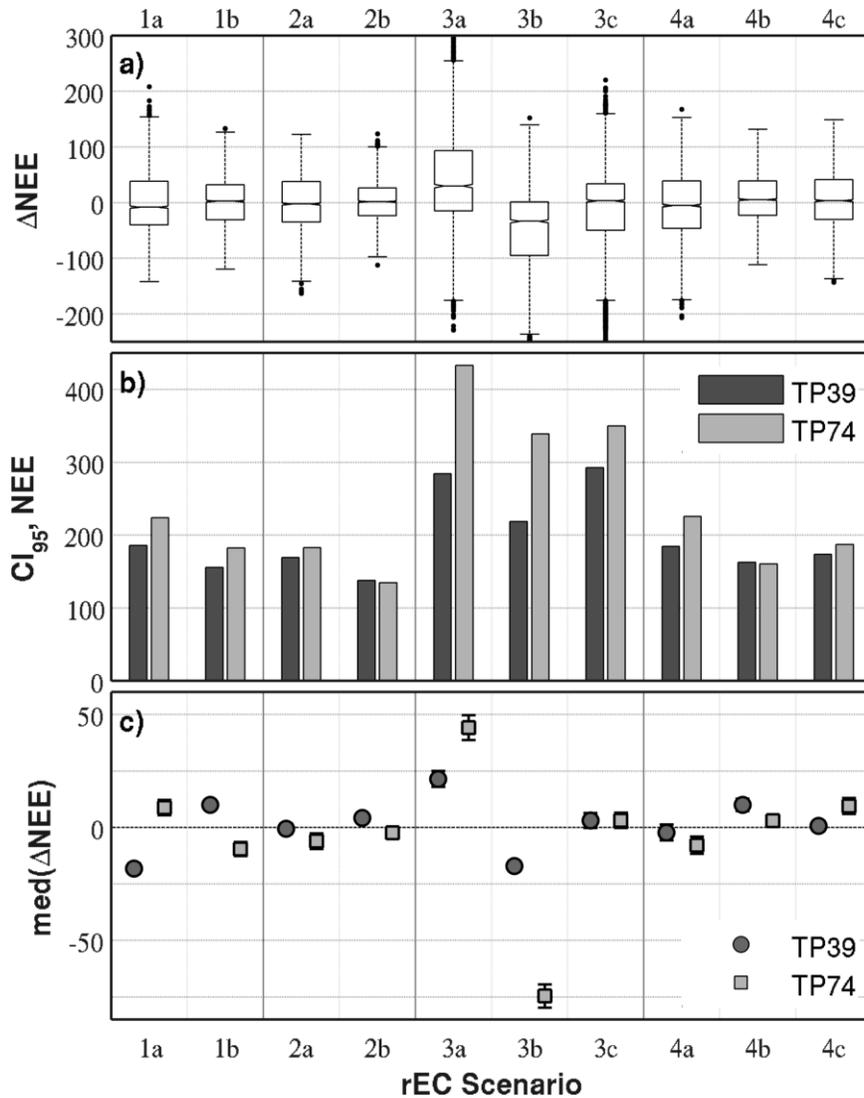
Figure 3.6: *Annual ΔNEE distribution (a), uncertainty (b) and bias (c) displayed for each rEC scenario. Data from both sites are aggregated for panel a), and are separated for panels b) and c). Standard errors for median estimates are indicated by error bars in panel c). NEE is reported in g C m^{-2} y^{-1}*

To gain a better understanding of the effect of rEC rotation schedules on NEE estimates and their uncertainty, ΔNEE was calculated at a weekly timescale, and subsequently grouped according to its rEC scenario and ensemble-averaged over each week of the year (Figure 3.7). Across all rEC scenarios and both sites, a general, temporally-autocorrelated trend in weekly ΔNEE was observed: Values of weekly ΔNEE tended to be relatively small (near zero) during the wintertime, moderately negative (model underestimation of NEE) in the spring and autumn months, and strongly positive (model overestimation of NEE) in the summer season. Considering this inherent bias tendency, the timing of rEC gaps exerted a large influence on the magnitude of weekly biases in two observable ways. First, the magnitude of biases were substantially higher during weeks experiencing rEC-induced gaps, as opposed to those during typical measurement. This result is anticipated, since gap-filled NEE would consist entirely of modeled data during periods of rEC-induced gaps – thus the effect of gap-filling model bias would be summative for all half-hours of these periods. Secondly, a positive correlation was observed between the length of rEC gaps and the magnitude of weekly ΔNEE, suggesting that gap-filling model performance for periods of rEC-induced gaps (in terms of bias error) worsens with the length of the gap.

Table 3.3: *Decomposition of total annual NEE uncertainty ($CI_{95,TOT}$) into components representing uncertainty due to random gap error ($CI_{95,\phi}$) and due to temporal structure of rEC-induced gaps ($CI_{95,rEC}$), for all rEC scenarios at TP39. Results are shown for the average of all gap-filling models, and individually for NLR-HL. All values of CI are in units of g C m^{-2} y^{-1}.*

rEC Scenario	Φ	$CI_{95,\phi}$	ALL		NLR-HL	
			$CI_{95, TOT}$	$CI_{95, rEC}$	$CI_{95, TOT}$	$CI_{95, rEC}$
1a	0.303	95.81	185.62	158.98	146.80	108.22
1b	0.302	95.95	155.65	122.56	184.21	155.47
2a	0.304	95.67	168.94	139.25	154.34	119.37
2b	0.301	96.10	137.63	98.52	108.19	40.89
3a	0.201	115.22	283.85	259.41	145.11	86.25
3b	0.199	115.70	218.40	185.24	143.99	81.30
3c	0.206	114.24	292.38	269.14	164.98	116.58
4a	0.203	114.72	184.31	144.25	187.89	147.47
4b	0.199	115.70	162.46	114.05	150.34	92.25
4c	0.203	114.74	173.34	129.92	165.16	116.42

Table 3.4: *Decomposition of total annual NEE uncertainty ($CI_{95,TOT}$) into components representing uncertainty due to random gap error ($CI_{95,\phi}$) and due to temporal structure of rEC-induced gaps ($CI_{95,rEC}$), for all rEC scenarios at TP74. Results are shown for the average of all gap-filling models, and individually for NLR-HL. All values of CI are in units of g C m^{-2} y^{-1}.*

rEC Scenario	Φ	$CI_{95,\phi}$	ALL		NLR-HL	
			$CI_{95, TOT}$	$CI_{95, rEC}$	$CI_{95, TOT}$	$CI_{95, rEC}$
1a	0.284	99.20	223.53	200.31	203.42	177.59
1b	0.286	98.79	182.07	152.94	170.76	139.28
2a	0.292	97.83	182.53	154.09	134.66	92.53
2b	0.279	100.16	134.39	89.60	134.30	89.47
3a	0.194	116.69	432.92	416.90	229.71	197.86
3b	0.183	118.85	338.68	317.14	310.67	287.03
3c	0.194	116.74	349.51	329.44	145.75	87.27
4a	0.195	116.43	225.63	193.27	221.64	188.60
4b	0.184	118.71	160.09	107.41	134.82	63.91
4c	0.192	117.14	187.09	145.88	185.30	143.58
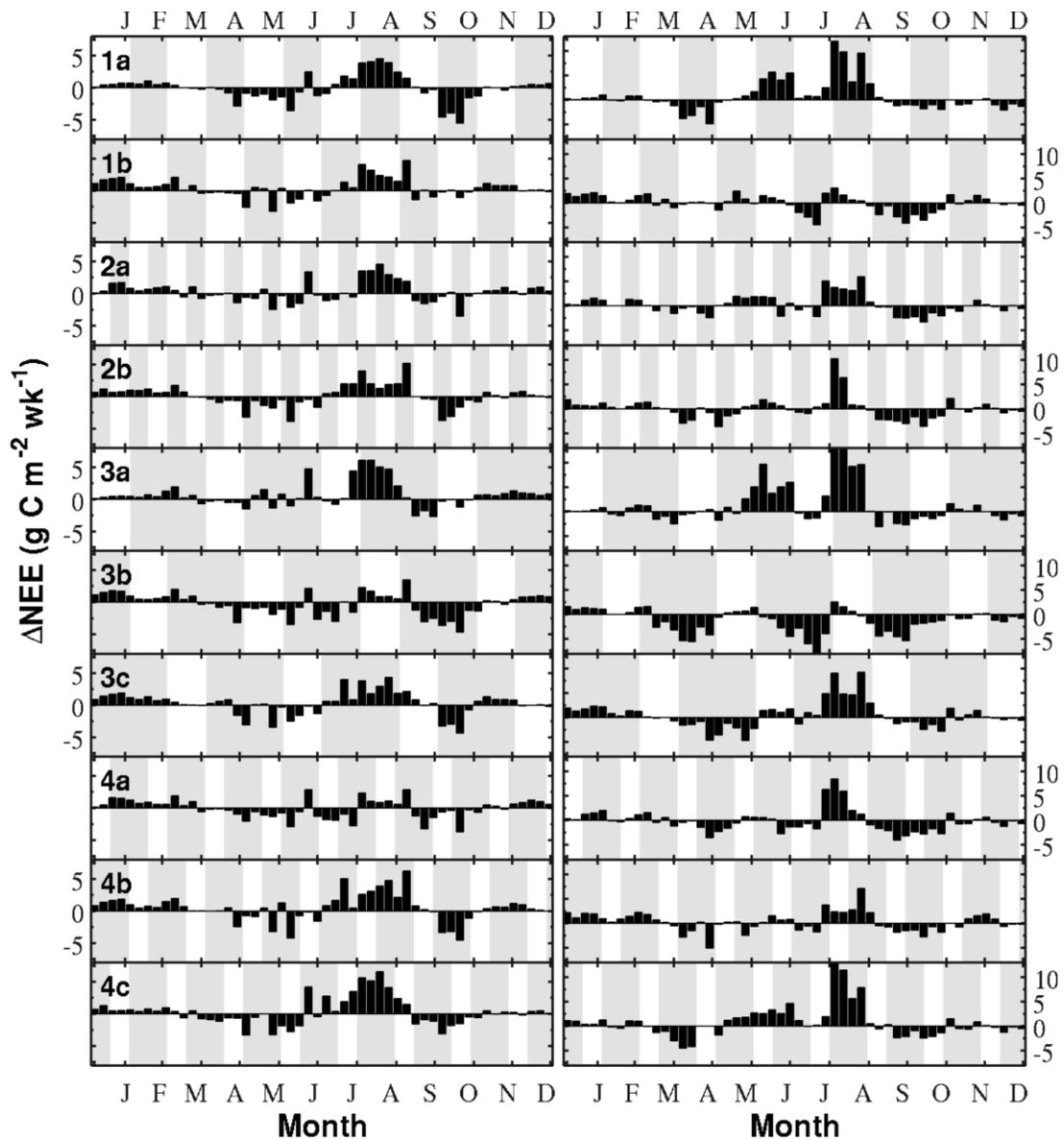
Figure 3.7: *Annual ensemble median of weekly ΔNEE resulting from NLR-HL gap-filling application across all years (2008-2011) and simulation runs. Results are shown for all rEC scenarios, applied at both TP39 (left panels) and TP74 (right panels). Greyed backgrounds indicate rEC gap periods.*

3.4.3 Evaluating gap-filling model performance

To evaluate the performance of the three selected gap-filling models (see section 3.3.6), we employed three different measures: uncertainty and bias of annual gap-filled NEE; point-to-point (half-hourly) goodness-of-fit to the measured data; and, preservation of trends in interannual variability. The uncertainty associated with application of each gap-filling model to simulated data from individual rEC scenarios was quantified by the magnitude of the 95% confidence interval in ΔNEE (Figure 3.8, panel a). Results showed that application of the MDS gap-filling was favourable across both sites for scenarios 1, 2 and 4, with 95 % confidence intervals in the general range of 100-200 g C m^{-2} y^{-1}. NLR-FC results were comparable to MDS for TP39 data, though NLR-FC uncertainties were generally higher for TP74 NEE. Both these methods experienced substantial increases (100 to 200 g C m^{-2} y^{-1}) in CI$_{95}$ when applied to rEC scenario 3, indicating a notable performance drop for this scenario. Though NLR-HL uncertainties were moderately higher for scenarios 1, 2 and 4, they were relatively stable across all rEC scenarios, with the exception of scenario 3b at the TP74 site. This indicates that NLR-HL is more robust to the range of rEC scenarios simulated in this experiment.

In terms of estimate bias (Figure 3.8, panel b), the application of each of the gap-filling models to simulations of different rEC scenarios generally produced annual NEE estimates within 50 g C m^{-2} y^{-1} of the 'true' sum established by NEE$_m$. Application of the MDS method resulted in a consistent negative bias (increased C uptake) in estimated annual NEE, ranging between -25 to -50 g C m^{-2} y^{-1}. Biases associated with application of the NLR methods were less consistent, though they were smaller in magnitude than MDS, and tended around zero. In accordance with uncertainty results, biases were generally highest for scenario 3; this was especially true for application to TP74 data.

In addition to their effects on annual sums, gap-filling models were evaluated on their ability to predict the point-to-point variability in half-hourly NEE values, in order to better quantify the capacity of each model to describe the short-term controls on NEE. As outlined in section 3.3.8, models were evaluated against three different reference data groups to create separate RMSE values that relate to internal (model-known), external (model-unknown) synthetic data, as well as the entire annual time series of modeled (noise-free) NEE$_m$ data.
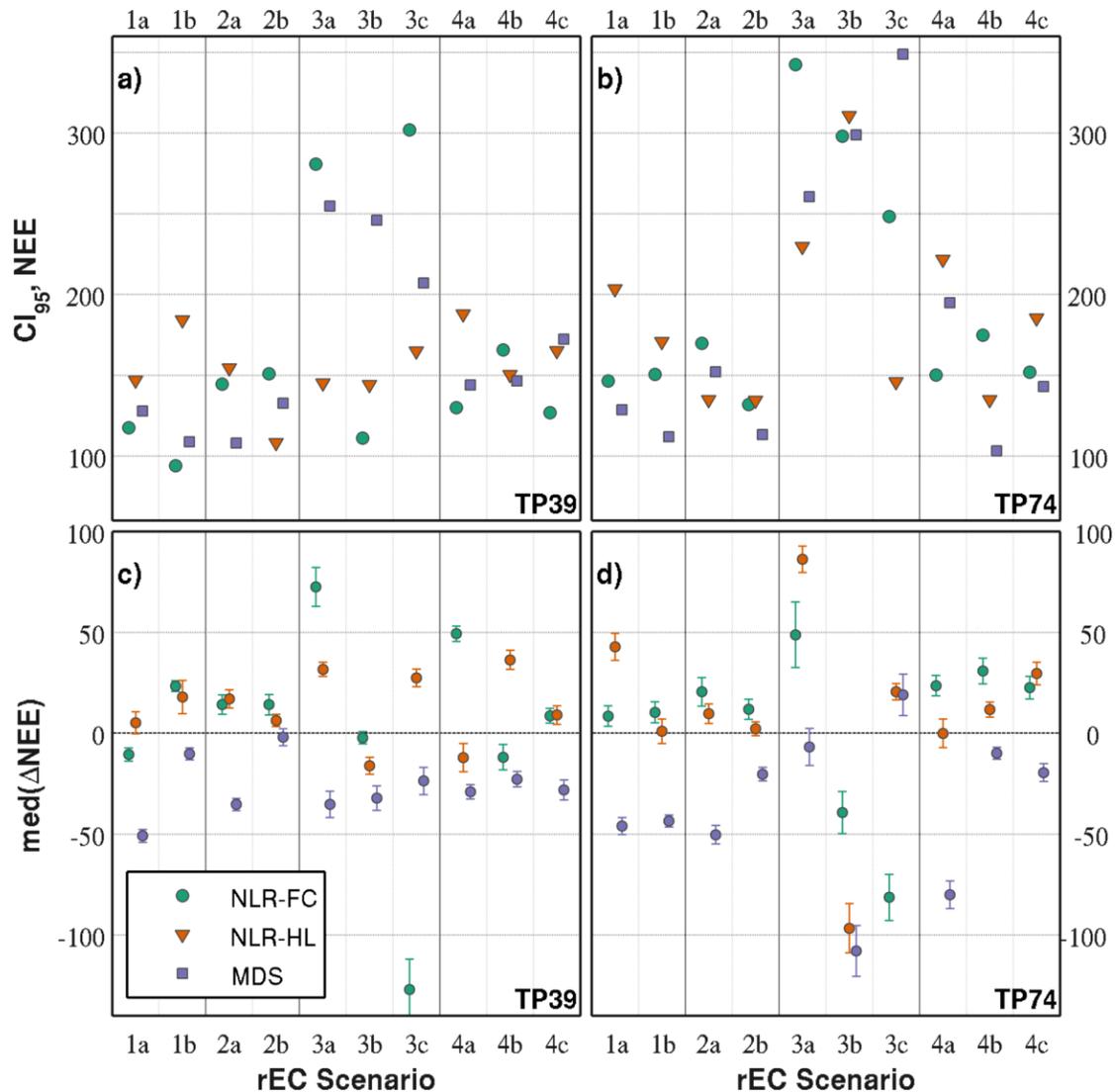
Figure 3.8: *Uncertainty (top panels) and median bias (bottom panels) associated with application of each gap-filling model across each rEC scenario. Results are shown separately for TP39 (left panels) and TP74 (right panels). NEE is reported in g C m^{-2} y^{-1}*

Across gap-filling models and rEC scenarios (Figure 3.9, panels a, b), median internal RMSE values ranged between 2.78 and 3.09 µmol CO_2 m^{-2} s^{-1}. Between models, internal RMSE was lowest using NLR-HL on TP39 data, and MDS for TP74. Internal RMSE measures were generally lower for scenarios 1 and 3 (1-month rotations), implying that model parameterization and performance improves for given periods with the length of measurement period.
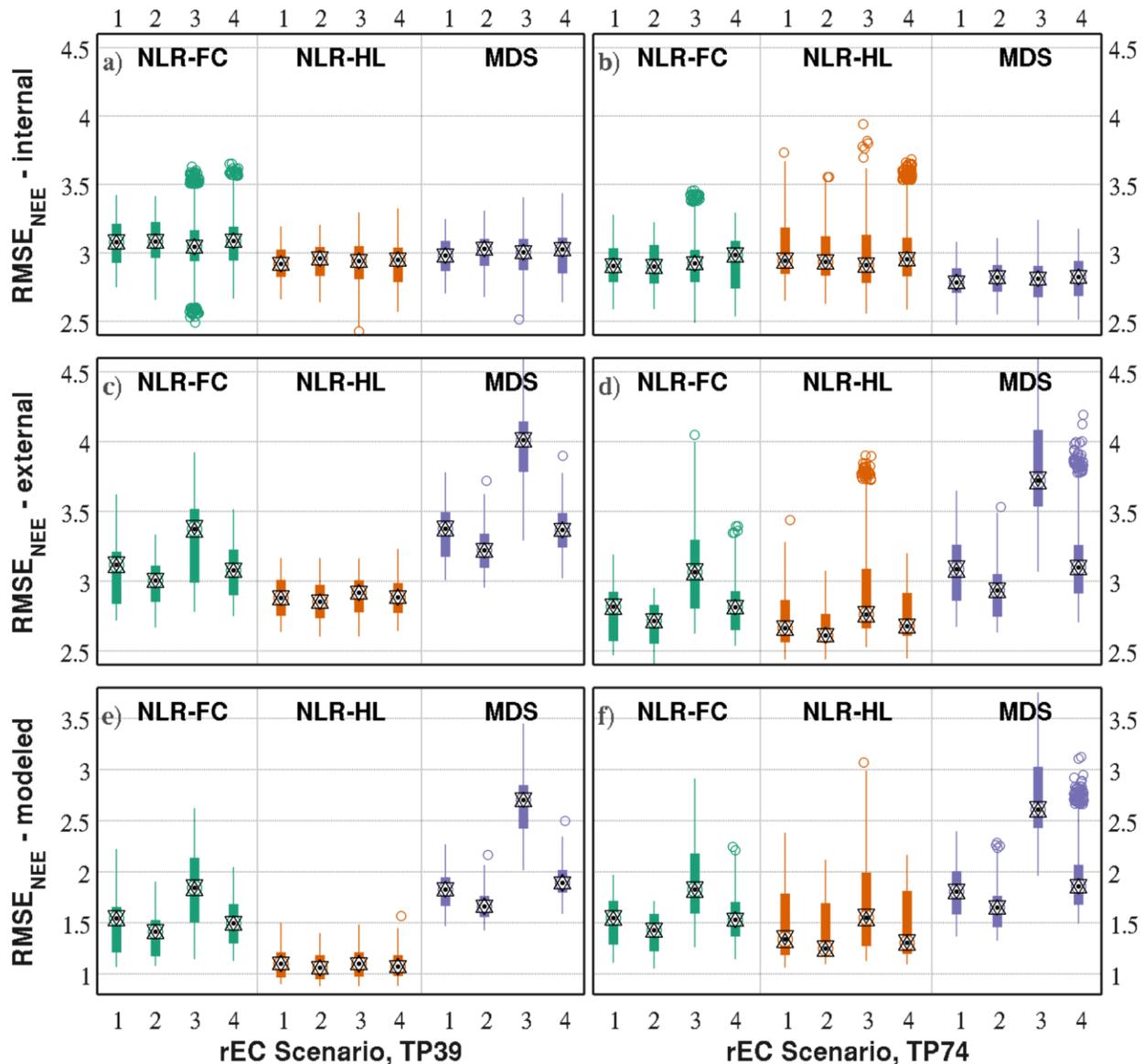
Figure 3.9: *Box plots of root mean square error (RMSE) values for each gap-filling model, calculated using different reference data groups: 'internal' – model-visible data; 'external'– model-invisible (gap) data; and, 'modeled' – entire NEE_m (noise-free) time series. Results are reported in units of µmol CO_2 m^{-2} s^{-1} for TP39 (left panels) and TP74 (right panels).*

As a benefit of using gap-free synthetic data for these analyses, it was possible to compare the internal RMSE values with estimates of 'external' RMSE (Figure 3.9, panels c, d), which expresses the goodness-of-fit of model predictions to all points that have gaps at the time of parameterization and filling. In comparison to internal measures, external RMSE for the MDS approach increased significantly across all scenarios, highlighting a strong performance decrease when modeling NEE not included in its parameterization scheme. For the NLR-FC method, RMSE was generally consistent between internal and external measures, except for a significant increase for rEC scenario 3. In comparison, external NLR-HL RMSE values showed either no change or a significant improvement from internal values, indicating no diminishing performance between predictions of points that are included and not included in parameterization.By evaluating RMSE in reference to the 'modeled' NEE_m time series, it was possible to investigate the performance of each gap-filling model to the noise-free, original signal that it was seeking to reproduce. It is relevant to note that RMSE values were lowest for this measure, as the effect of random noise on goodness-of-fit was removed. When evaluated with this metric, the NLR-HL model performed significantly better than the other methods across all rEC scenarios, further supporting the use of this model to gap-fill rEC-measured data.

The relative consistency of NLR-HL model performance across all RMSE measures provoked further investigation, in order to confirm that these results reflected a genuine performance advantage, rather than one that was induced by correlation of this specific model to the ANN used to generate the NEE_m time series. To investigate this effect, the same Monte Carlo simulation methodology was applied, however, measured NEE data was substituted for NEE_m. As a result, RMSE estimates for internal, external and all data points reflected the true goodness-of-fit of model output. Simulation results (not shown), were consistent with those achieved using NEE_m, confirming that our results reflected a genuine performance advantage of the NLR-HL model.
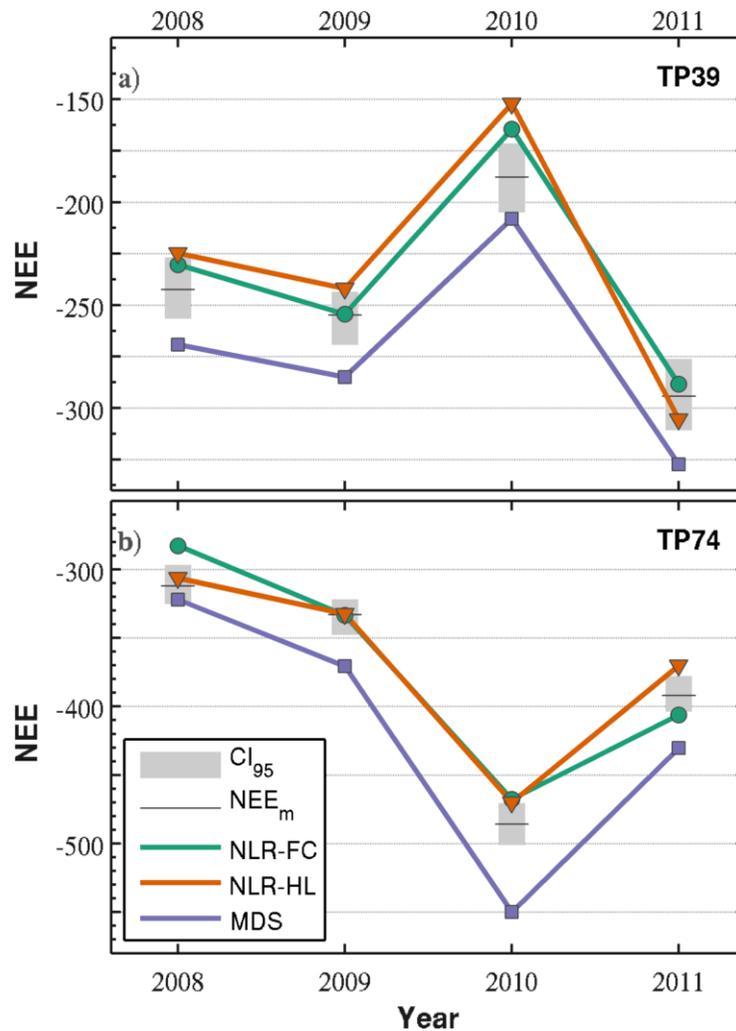
Figure 3.10: *Interannual variability of the median annual gap-filled NEE estimate for each gap-filling model for TP39 and TP74. Results are shown alongside the 'modeled' value of NEEm, and the 95% confidence interval estimated from synthetic time series. NEE is reported in g C m^{-2} y^{-1}*

Gap-filling models were also evaluated to ensure that annual NEE estimates preserved year-to-year trends and interannual variability observed in the original input time series (Figure 3.10). For both sites, the median annual values follow the same general year-to-year trend as those observed in NEE$_m$ and synthetic NEE time series. Most median values fell within the 95% confidence interval of the synthetic NEE data, with the clear exception of MDS model output, which produced annual estimates with a persistent negative bias (consistent with Figure 3.8).

3.4.4 Evaluating strategies to mitigate rEC gap effects

The three rEC gap effect mitigation strategies, described in section 3.3.7, were evaluated for their ability to improve gap-filling model uncertainty, bias and point-to-point variability prediction, while also preserving interannual trends in NEE estimates. Results are shown only for the application to the NLR-HL model, due to the fact that this was the sole model for which all strategies could be applied, and also because previous analyses indicated that this model was most robust to rEC-imposed effects on NEE estimates.

Results showed that variability in annual NEE estimates for a given site and rEC application were similar between the original NLR-HL results and those for each of the applied strategies (Figure 3.11, panels a, b). However, application of the "pooled" approach resulted in an increased number of extreme (outlier) values in comparison to the original. Conversely, application of both the "filled" and the "alternating" strategies reduced the overall number of outliers; in this way, application of these two strategies appear to stabilize the NLR-HL method from producing extreme annual NEE estimates.

The 95% confidence intervals associated with the application of these strategies (Figure 3.11, panels c, d), showed that when averaged across rEC scenarios, annual NEE uncertainty increased when either the "pooled" or "filled" strategies were applied. Confidence intervals increased by an average of 3.9 and 15.2 g C m^{-2} y^{-1} for pooled and filled applications for TP39, respectively, and 3.6 and 18.3 g C m^{-2} y^{-1} for TP74. In contrast, application of the "alternating" rEC mitigation strategy reduced uncertainty on annual estimates by an average of 27.6 and 72.8 g C m^{-2} y^{-1} at TP39 and TP74, respectively.

Comparing the median annual NEE bias of estimates across strategies and rEC scenarios (Figure 3.11, panels e, f), application of the 'pooled' rEC strategy provided small but consistent improvement to the biases associated with the original NLR-HL model (with the exception of rEC scenario 1 for TP74). In comparison to the bias associated with the original NLR-HL model, application of the 'pooled' strategy reduced the absolute value of bias by an average of 5.8 and 3.6 g C m^{-2} y^{-1} across rEC scenarios for TP39 and TP74, respectively. When the 'alternating' strategy was applied, the absolute bias across rEC scenarios increased moderately by 7.1 and 11.5 g C m^{-2} y^{-1} for these sites. In contrast, implementing the 'filled' strategy introduced large and consistent negative bias to NEE estimates across all rEC scenarios, and absolute values of bias are increased relative to the NLR-HL approach by 28.0 and 74.0 g C m^{-2} y^{-1} at TP39 and TP74.
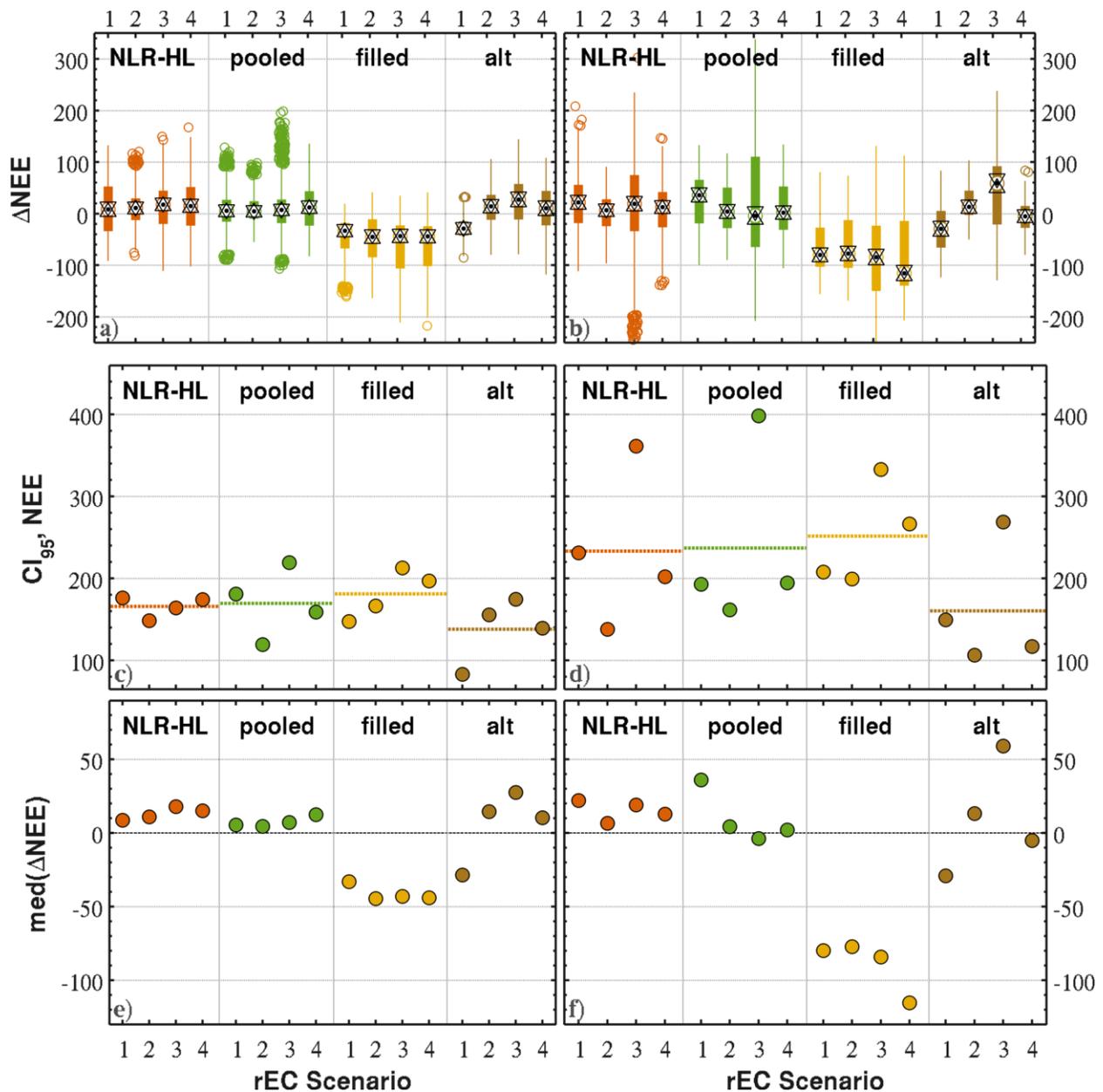
Figure 3.11: *Annual ΔNEE distribution (top panels), uncertainty (middle panels) and bias (bottom panels) associated with the application of rEC mitigation strategies to the NLR-HL gap-filling model. Results are aggregated for each rEC scenario, and shown for TP39 (left panels) and TP74 (right panels). Dashed horizontal lines in panels c and d indicate the mean uncertainty for a given strategy treatment, across all rEC scenarios. NEE is reported in g C m^{-2} y^{-1}*

rEC strategies were also evaluated for their ability to improve the point-to-point goodness-of-fit for the NLR-HL model output (Figure 3.12). Similar to investigations presented earlier, these evaluations used both internal and external RMSE measures to evaluate model performance, to explore discrepancies that may arise between the evaluations that can be provided within the model (internal RMSE) and those that better reflect the ability of the model to predict values (external RMSE).

Comparing internal RMSE estimates across strategy applications (Figure 3.12, panels a, b) indicated that data subjected to 'pooling' and 'alternating' strategies was comparable to the original results of the NLR-HL model, though the 'alternating' strategy reduced the number of extreme positive outliers in TP74 results. Internal RMSE results obtained by application of the 'filled' strategy showed significant improvement (RMSE reduction) when compared to values for the original NLR-HL model and the other mitigation strategies. Similar results were obtained through comparison of external RMSE values (Figure 3.12, panels c, d), as only the application of the 'filled' strategy showed a significant improvement in gap-filling performance. As for internal measures, application of the 'filled' strategy reduced external RMSE consistently across all rEC scenarios, and the effect was more pronounced at TP74. The results indicate that in our setting, the 'filled' strategy is effective at reducing point-to-point error in gap-filled rEC time series.

We also investigated the impact of these rEC mitigation strategies on interannual variability (Figure 3.13). In contrast to RMSE results, application of the 'filled' strategy generated greatly unsatisfactory results, as little to none of the interannual NEE trend was preserved. In correspondence with Figure 3.11, this result demonstrates that although using the filled strategy reduces the point-to-point error associated with NEE gap-filling, the error is systematically biased, which leads to marked cumulative errors when NEE values are integrated over time.

Overall, the 'alternating' strategy was the most effective at maintaining the true interannual trends; deviations from the annual NEE_m values were small, and for TP39, these deviations were smaller than observed for the original NLR-HL output. Application of the 'pooled' strategy also resulted in satisfactory preservation of interannual trends – an unexpected result given that for this strategy, RE and GEP models were forced to implement a single functional form for all years. Since the use of static parameters for functional relationships did not lead to considerable differences between years, it can be inferred that most year-to-year differences in annual estimates are due to interannual variation in meteorological drivers of ecosystem carbon exchange.
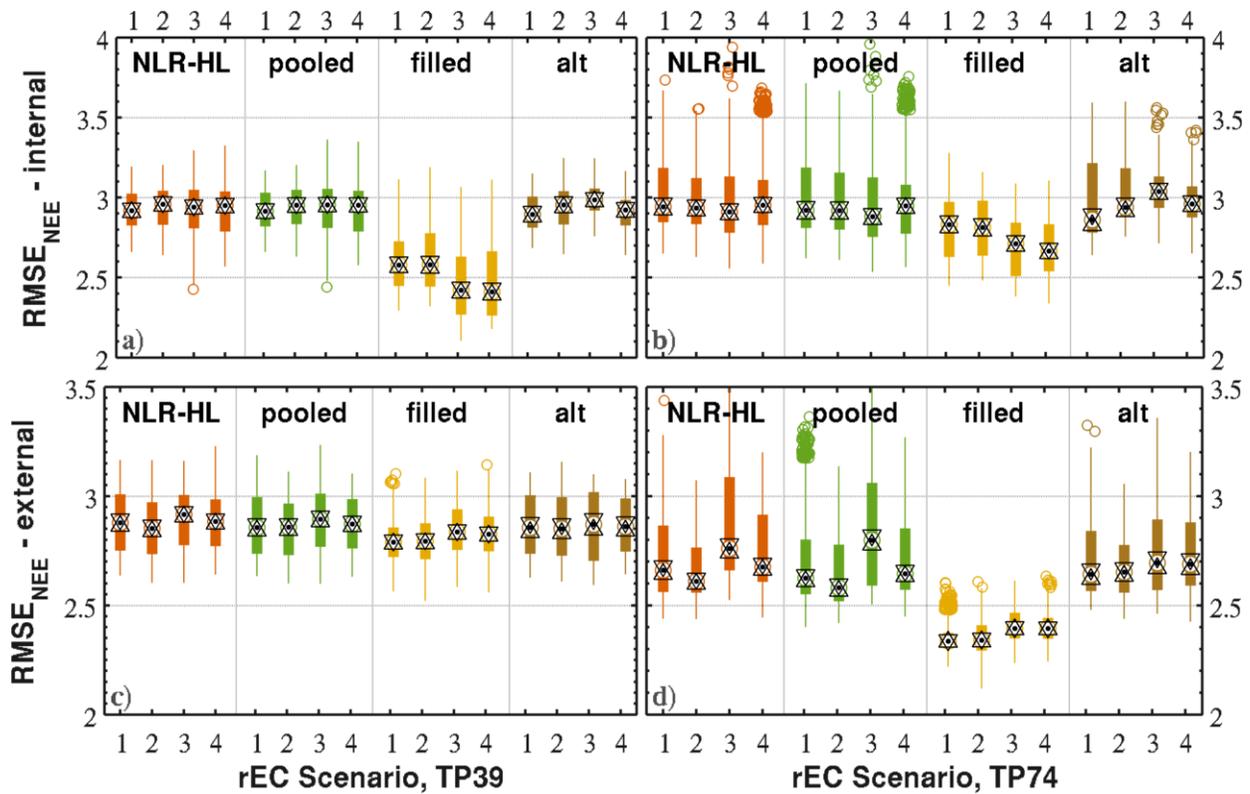
Figure 3.12: *Box plots of root mean square error (RMSE) values for the NLR-HL gap-filling model, and each of the three tested rEC mitigation strategies. Top and bottom panels show estimates made using different reference data groups: 'internal' – model-visible data; and, 'external'– model-invisible (gap) data. Results are shown in units of μmol m^{-2} s^{-1} for TP39 (left panels) and TP74 (right panels).*
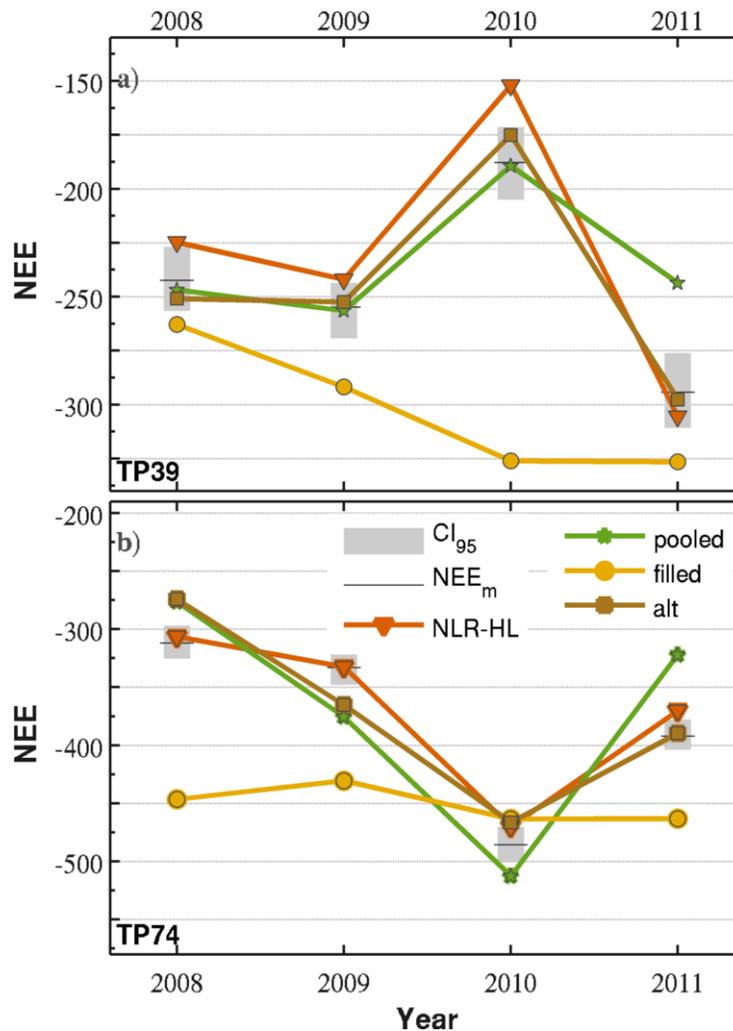
Figure 3.13: Interannual variability of the median annual gap-filled NEE estimate for NLR-HL and each rEC mitigation strategy applied to this model for TP39 and TP74. Results are shown alongside the 'true' value of NEE_m, and the 95% confidence interval estimated from synthetic time series. NEE is reported in $g\ C\ m^{-2}\ y^{-1}$

3.5 Discussion

3.5.1 Effects of rEC application on time-integrated NEE estimates-comparison to effect of standard-type gaps

In addition to the rEC simulations run in this study, two standard-type gap scenarios were also developed and simulated, in order to provide points of reference for uncertainties associated with rEC scenarios. Both reference ('operational' and 'filtering') scenarios were similar to each other – and different from rEC scenarios – in that the gaps were much smaller in length, and were distributed nearly evenly throughout the year (with slightly greater amount of gaps during RE-parameterizable half-hours). The uncertainty associated with these simulation runs showed a negative correlation with the amount of available data, as uncertainty increased with the amount of gaps. This general relationship is consistent with findings of other studies that investigated NEE data gaps and gap-filling performance (Moffat et al., 2007; Richardson and Hollinger, 2007); however the rate of uncertainty increase with the number of gaps $(1-\phi)$ differs in comparison. For example, Richardson and Hollinger, (2007) found random uncertainty (2σ in their study, similar to CI_{95} in ours) to be less sensitive to data gaps (for the range of $0.4 \leqslant \phi \leqslant 0.8$), reporting a relationship of: $2\sigma = 21.3 + 15.9(1-\phi)$. In comparison, the relationship derived from operational and filtering gap simulations for our sites is $CI_{95} = 19.1 + 101.9(1-\phi)$ for a range of $0.2 \leqslant \phi \leqslant 0.9$. While these relations share similar estimates of baseline (gap-free) uncertainty, the rate of uncertainty increase in our study is approximately 6.5 times greater. This discrepancy may be due to a number of factors: smaller investigated span of ϕ, and the generalization across numerous forest types in the Richardson and Hollinger (2007) study; as well as discrepancies in methods used to describe and simulate random error.

By quantifying annual NEE estimate uncertainty associated with standard-type (operational and filtering) gaps across a wide range of ϕ (Figure 3.5), it is possible to separate the total uncertainties reported for each rEC scenario application into that which is attributable to the number of gaps present in the time series, and that which is a result of the gap structure imposed by rEC operation (Tables 3.3 and 3.4). This process revealed that random uncertainty due solely to the amount of gaps $(1-\phi)$ accounts for 97 and 116 g C m^{-2} y^{-1} of the uncertainties estimated for 2-site (scenarios 1, 2) and 3-site (scenarios 3, 4) rEC scenarios, respectively. Removing these amounts (in quadrature) from the uncertainties recorded for rEC simulations ($CI_{95,TOT}$), the resulting gap structure-attributable

uncertainties ($CI_{95,rEC}$) range between 89.6 to 417 g C m^{-2} y^{-1}, with most values (excluding scenario 3) between 100 and 200 g C m^{-2} y^{-1}. These results compare favourably to estimates made by Richardson and Hollinger (2007), who developed a month-specific relationship between longest gap length during the month, and the total increase to annual NEE estimates for Howland forest. When their relationship is applied additively to our rEC scenarios, the amount of uncertainty attributable to rEC gaps ranges between 110-158 g C m^{-2} y^{-1} – comparable in magnitude to the uncertainty we have estimated in our study.

By testing 10 different rEC scenarios, we were able to investigate the effects of rEC schedules across three dimensions of system operation: number of sites, measurement period length, and the rotation timing. Among these operational factors, the length of rEC system measurement period had the greatest influence on overall NEE uncertainty, and scenarios implementing a 15-day rotation period experienced less overall uncertainty than those implementing month-long periods (Figure 3.6). From the perspective of gap-filling model parameterization, shorter rEC rotation periods result in better temporal spacing of input information, which reduces the degree to which underlying ecosystem exchange dynamics may change between measurements, thus reducing model deviation from the 'true' flux. The use of a three-site rEC rotation instead of two has little influence on uncertainty when a 15-day rotation period is used; however, when applied to month-long rotations, uncertainty in annual NEE estimates is increased 2- to 3-fold. This effect was demonstrated in Figures 3.6 and 3.7. This interactive effect between rotation period length and number of sites indicates that gap-filling model performance diminishes significantly and drastically as gap lengths are increased from one to two months. There was also a small effect associated with the timing of rEC rotation, as NEE estimate uncertainty was smaller (both on weekly and annual scales) when measurement periods coincided with periods of rapid changes in phenology (e.g. growing season start and end), or controlling environmental variables (mid-summer soil moisture deficit). This result suggests that the performance of rEC measurement programs could be optimized by modifying the rotation frequency and timing to maximize data capture at all sites during these periods.

Though the general uncertainty associated with rEC implementation was shown to be substantial, there also existed clear differences in outcomes between rEC scenarios, study sites and gap-filling models. Overall, application of rEC scenario 2 (two sites, two-week rotation period) resulted in the most favourable and consistent annual NEE estimate performance, especially when the NLR-HL gap-filling model was applied. Considering application of this scenario at TP39 and TP74 during the study years,

results could be expected to have minimal overall bias (<8 g C m^{-2} y^{-1}), and the estimated total uncertainty (95% confidence interval) could be expected between 108-154 g C m^{-2} y^{-1} for TP39, and around 135 g C m^{-2} y^{-1} for TP74. When normalized by the 4-year mean 'true' NEE magnitudes for these sites (-245 and -381 g C m^{-2} y^{-1} for TP39 and TP74, respectively), the noise-to-signal ratio ranged between 44 to 63% for TP39 and 35% for TP74. Due to relatively large annual NEE magnitudes, these ratios are surprisingly comparable with values obtained over a variety of forests where only single (or no) long gaps were inserted (Richardson and Hollinger, 2007), and virtually assure that the direction of net annual NEE will be preserved with rEC operation at our sites. Results also provide evidence that interannual variability is preserved, and that deviations from the 'true' annual NEE are correlated across all years of individual rEC simulation runs. This is especially true for the NLR-HL model, which incorporates information across all years during parameterization. Therefore, despite persistent uncertainty in absolute value, evidence suggests that relative differences between years will be preserved.

3.5.2   Gap-filling model performance and selection

The general variability in gap-filled NEE estimates observed across rEC scenarios and gap-filling models demonstrates the challenges associated with obtaining reliable time-integrated NEE estimates from rEC operation. The large, frequent gaps introduced by rEC operation present a formidable challenge for the gap-filling models required to fill these gaps to create accurate, continuous and time-integrable estimates. rEC gaps increase uncertainty in gap-filled time-integrated NEE estimates by two means: a) by negatively affecting gap-filling model parameterization and the accuracy of predicted values, since the underlying dynamics of ecosystem exchange may change significantly during the period of a long gap; and b) by increasing the number of gaps that need to be filled, therefore increasing the number of half-hours containing biased estimates.

Overall, the NLR-HL model was chosen as the most favourable gap-filling model for rEC application at our sites, as a result of its relative consistency across different rEC scenarios, its preservation of point-to-point and interannual variability, and the possibility for mitigating strategies to be applied to it, in order to reduce uncertainty. Contributing to the consistency of the NLR-HL model is the fact that it incorporates information across all years to parameterize the response of RE and GEP to all secondary controlling environmental variables (all but T$_s$ for RE; PPFD for GEP). Though generalizing these

responses over all years may lead to small year-to-year errors, this approach helps to ensure that the basic form and magnitude of these relationships are properly modeled.

The NLR-FC model is a more simplistic version of NLR-HL, in that both RE and GEP depend only on a single variable (T_s and PPFD, respectively). A time-varying factor is used to correct predicted values for linear offsets from measured values over all periods of the year, which in effect accounts for changes in exchange capacity, phenology, and the influence of other environmental variables. The magnitude of this time-varying factor is interpolated between periods in which data exists, in order to create a gradually-changing value, as would be expected throughout the course of a year. However, in situations where gaps are large, the interpolation of this correction factor over one or two months can lead to large inaccuracies. This effect is especially pronounced for results of rEC scenario 3, where the presence of 2-month-long gaps led to large uncertainty estimates (> 250 g C m^{-2} y^{-1} for most scenario 3 outcomes), that were also largely biased in their central tendency, effectively increasing the inaccuracy of NEE estimates.

Both the NLR-HL and NLR-FC models showed a strong tendency to overestimate NEE (lower carbon sequestration estimates) during the peak summer season (Figure 3.7), an effect that was augmented when coincident with rEC-induced gaps. This likely reflects an overestimation of the RE response to T_s at its maximum values. In reality, the magnitude of respiration plateaus during the mid-summer in response to decreases in soil moisture availability in the sandy soil of our sites. If data from this period is not available to the gap-filling models during parameterization, then these models are not be able to account for this feature, and simply extrapolate from the increase of respiration with temperature during non-water limited periods (spring and autumn).

In terms of annual NEE estimates, the MDS model performed reasonably well for all scenarios, with the exception of rEC scenario 3. For the remaining scenarios, estimate uncertainty is comparable or marginally lower than the other models examined. Consistent with the findings of Moffat et al. (2007), however, MDS results showed a consistent negative bias among estimates; this should be considered as additive to estimate uncertainty when quantifying the absolute deviation of model predictions from the 'true' value. The decreasing performance of MDS with increasing gap length is a consequence of the look-up-table algorithm used by this method. The structure of the algorithm is to continuously expand the time window and loosen restrictions on acceptable environmental conditions until acceptably similar conditions can be found. As a result, values in the middle of long gaps must be filled with the average of measurements up to months removed from the half-hour of interest. Treating such

temporally-distant values as a representative replacement likely extends beyond the period for which assumption of temporal correlation is valid.

The shortcomings of the MDS model in filling long rEC-induced gaps is demonstrated further by the discrepancy between 'internal' and 'external' measures of RMSE shown in Figure 3.9, which shows a greatly reduced capacity of the model to predict point-to-point variability during periods where information does not exist. The poor performance of this model to provide accurate NEE estimates at sub-annual timescales has implications for its usefulness beyond estimating annual and inter-annual NEE sums, for applications such as parameterizing, comparing or validating ecosystem models that operate at hourly to daily time steps.

The significant differences between 'internal' and 'external' RMSE estimates also have important implications for all gap-filling operations, as they demonstrate that 'internal' metrics obtained by comparing the model-predicted values to existing measurements (typically the only performance information available during gap-filling), are not necessarily reflective of model performance for data gaps. Thus, caution should be taken when interpreting 'internal' estimates, and more robust model validation practices should be used.

3.5.3   Evaluating rEC gap-effect mitigation strategies

Of the three strategies tested for reducing the uncertainty of gap-filled NEE estimates in rEC applications, our results indicate that alternating rEC schedules for successive years is the best approach for improving the accuracy of these estimates. Though the 'pooled' method showed minimal improvements in bias error, and the 'filled' strategy improved point-to-point prediction, both demonstrated limitations in other important metrics of performance. Although applying the 'alternating' rEC strategy to these sites slightly increases the potential overall bias of estimates (Figure 3.11), this effect is outweighed by an uncertainty reduction that is a factor of 4 to 6 times greater, thus providing a large reduction in overall deviation of estimates from 'true' values. Though half-hourly NEE values from TP39 and TP74 were strongly correlated ($r = 0.82$), using data from one site to fill in gaps in the other ('filled' strategy) produced results that were worse than the original NLR-HL model outcomes across all measures. As interannual trends showed for 2010 (Figure 3.10), despite the fact that these two ecosystems are in close proximity, are the same species, and grow on similar soil conditions, they differ in terms of carbon exchange dynamics and response to environmental drivers. These forests

differ in terms of understorey growth, surrounding land-uses that are included in the flux footprint, and responses to mid-summer dry periods, which contribute to the observed differences between the sites.

These results indicate that for the study sites in question, actions must be taken at the time of rEC operation, in order to minimize negative effects on NEE estimates; the post-collection strategies examined here do not provide any substantial benefit to NEE estimates made at any timescale.

3.5.4 Practical considerations and challenges for rEC operation and NEE uncertainty estimation at TPFS

Before the simulation results of this study can be considered representative of expected outcomes of past or future rEC measurement programs at TPFS, a number of additional, practical considerations must be made.

First, it is important to consider the impact that footprint filtering would have on gap-filling model performance and NEE uncertainty, had it been applied to these sites. While each site in the TPFS age-sequence occupies a space in a larger forested region, the age-specific, monoculture sites themselves may all be considered fetch-limited, due to the fragmented nature of land plots in the region. Though a footprint filtering scheme was not applied in this study – fluxes from a given site were inclusive of possible contamination from outside sources – it is reasonable to anticipate that such an application could be required when comparing fluxes among age-sequence forests. Applying such a half-hourly footprint filtering scheme would result in approximately as many data points removed as for u_*^{Th} filtering. Therefore, it can be expected that ϕ values for each rEC scenario would be reduced between 0.05 and 0.1 further than is reported in this study. This increased data reduction is likely to have negative effects on gap-filling model performance – especially since the rEC program implemented at TPFS used a 3-site rotation, which was a mix of rEC scenarios 3 and 4.

A second consideration to be made when extending these findings to actual rEC implementation at TPFS concerns the consistency of NEE measurement random error estimation. For the purposes of this study, random NEE measurement error was modeled and estimated for all synthetic data prior to rEC simulation runs, using the original measured NEE time series at both TP39 and TP74. This approach increased consistency in NEE estimates, and ensured that differences in values were due to the application of rEC gaps. However, in real rEC application, the amount of data available will also affect estimates of the relationship between T_s, PPFD and σ_i. This has implications for the weights used in

parameterization, and adds increased uncertainty to estimates. In order to stabilize gap-filling model parameterization, it may be necessary to remove the estimation and consideration of variable NEE measurement error altogether, by reverting to the more robust ordinary least squares cost function, in place of weighted-error sum of squares or mean absolute weighted error. As demonstrated by Wijk et al. (2008), changes to the gap-filling model parameterization cost function can impose systematic biases on gap-filled NEE estimates.

A third TPFS-specific consideration is the applicability of rEC results and recommendations obtained using TP39 and TP74 data to the other two (younger) age-sequence sites. In contrast to the relatively stable and consistent nature of the more mature TP39 and TP74 sites, the nature of ecosystem carbon exchange is much more dynamic at TP89 and TP02. This consideration is particularly relevant for the years in which rEC was implemented at TP02 – the first five years after site establishment. During this period, NEE was influenced both by rapidly growing seedlings, and extensive and dynamic understorey growth and senescence – all of which responded quickly to environmental controls, due to their shallow rooting depth on sandy soils. As a result, it is anticipated that the presence of rEC-induced gaps at TP02 (and to a lesser extent, TP89), would increase the uncertainty associated with time-integrated NEE estimates during this period. However, the noise-to-signal ratios of these two sites would differ significantly, as the magnitude of annual NEE is substantially lower at TP02 than TP89.

3.5.5   General considerations for rEC implementation

An important general consideration for rEC program implementation is the length of time (i.e. number of years) in which the rEC program will be carried out. For the purposes of our study, we simulated a four-year measurement program, which is similar in length to that which was carried out at our sites (2003-2007). As demonstrated in the results, acquiring multiple years of measurement data at each rEC site affords opportunities to improve gap-filling model performance and reduce uncertainty in estimates. Specific examples of this include the use of multiple years of data in NLR-HL model parameterization, and the application of annually-alternating rEC schedules to improve overall temporal measurement coverage. Though not explicitly tested in this study, it is anticipated that increasing the number of rEC measurement years will lead to more constrained and accurate estimates of time-integrated NEE. Conversely, if the total length of the rEC program is less than that simulated in this study, it is quite possible that one of the other gap-filling models tested (NLR-FC or MDS)

would become the favourable choice. Since both of these models incorporate data and parameterize/operate on a year-by-year basis, the results reported in this study for these two models are expected to be representative of their performance for any given length of rEC program.

The results of this study suggest that rEC-induced NEE estimate uncertainty can be minimized by the application of a 2-site rotation at a period of 15 days. Investigations show that reducing the time period of rEC rotation provides the most substantial benefit to time-integrated NEE estimates; however, there are a number of pragmatic and logistical challenges to rotating an rEC system at 15-day (or shorter intervals). A short rotation period increases the demand for personnel to remove, transport and re-install equipment at a new site, which increases cost, and the need for establishing sites in close proximity, so that such a rotation is logistically feasible. Furthermore, increased rEC system transport increases not only the amount of overall measurement half-hours lost due to take down and set-up (an additional source of gaps not included in this study), but also increases the risk of problems associated with instrument damage and poor/inconsistent instrument setup. Such practical concerns should be considered at the time when rEC measurement schedules are being designed.

3.6 Conclusions

In this study, we quantified the potential effects of rEC program operation on time-integrated NEE estimate uncertainty. Results showed that the uncertainty in these estimates varies greatly across the numerous dimensions of implemented rEC schedules (number of sites, rotation period, schedule timing), measurement sites and gap-filling methods. As anticipated, NEE estimate uncertainty was minimized by applying the shortest rotation period among the fewest number of sites (15-day rotation between two sites, respectively), and was greatest for the opposite case (one month rotation between three sites). Surprisingly, however, the length of rotation had a much larger effect on NEE measurement uncertainty than the number of sites involved, since uncertainties associated with the 15-day, 3-site application were only minimally larger in magnitude than that involving two sites. These findings suggest that rEC measurement programs should minimize rotation period among sites – a consideration that has implications for the logistical feasibility of such a program.

Among the gap-filling models investigated, the NLR-HL model showed the best overall performance, as it minimized uncertainty, introduced minimal bias, and preserved both intra- and inter-annual variability. The advantage provided to this model is the incorporation of multiple years of information into parameterization processes, thus improving robustness of estimates. Our results showed that when

NLR-HL was applied to the optimal rEC scenario (2-site, 15-day rotation), the resulting total uncertainty was between 35 and 63% of the NEE flux magnitude at our study sites – an encouraging result given that approximately 70% of annual half-hours required filling, and that numerous month-long gaps existed in the data.

Of the rEC uncertainty-reducing strategies that were tested in this study, only the act of alternating annual rEC schedules was able to reduce NEE estimate uncertainty. Applying this strategy – which must be implemented at the time of rEC operation – reduced the uncertainty by an average of 28 and 73 g C m^{-2} y^{-1} in comparison to standard NLR-HL output for our TP39 and TP74 sites, respectively.

The results of this study suggest that the rEC approach has potential to expand the ecosystems that can be investigated by research programs with limited resources, while still providing time-integrated NEE estimates with an acceptable amount of uncertainty. Despite the promise of these results, questions remain as to the applicability of rEC outside of the investigated sites, especially those that demonstrate more rapidly changing exchange characteristics – whether intra-annually (e.g. temperate deciduous forests) or inter-annually (e.g. young, rapidly-growing forests). Futhermore, the methodology provided in this study should be viewed as a mean by which the rEC approach may be investigated at a greater number and variety of study sites.

3.7 References

Arain, M., Restrepo-Coupe, N., 2005. Net ecosystem production in a temperate pine plantation in southeastern Canada. Agric. For. Meteorol. 128, 223–241.

Aubinet, M., 2008. Eddy covariance CO2 flux measurements in nocturnal conditions: an analysis of the problem. Ecol. Appl. 18, 1368–1378.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82, 2415–2434.

Baldocchi, D.D., 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Glob. Chang. Biol. 9, 479–492.

Barr, A.G., Black, T. A., Hogg, E.H., Kljun, N., Morgenstern, K., Nesic, Z., 2004. Inter-annual variability in the leaf area index of a boreal aspen-hazelnut forest in relation to net ecosystem production. Agric. For. Meteorol. 126, 237–255.

Barr, A.G., Richardson, A.D., Hollinger, D.Y., Papale, D., Arain, M.A., Black, T.A., Bohrer, G., Dragoni, D., Fischer, M.L., Gu, L., Law, B.E., Margolis, H.A., McCaughey, J.H., Munger, J.W., Oechel, W.C., Schaefer, K., 2013. Use of change-point detection for friction–velocity threshold evaluation in eddy-covariance studies. Agric. For. Meteorol. 171-172, 31–45.

Borken, W., Savage, K., Davidson, E.A., Trumbore, S.E., Hole, W., 2006. Effects of experimental drought on soil respiration and radiocarbon efflux from a temperate forest soil. Glob 12, 177–193.

Coursolle, C., Margolis, H.A., Giasson, M.-A., Bernier, P.-Y., Amiro, B.D., Arain, M.A., Barr, A.G., Black, T.A., Goulden, M.L., McCaughey, J.H., Chen, J.M., Dunn, A.L., Grant, R.F., Lafleur, P.M., 2012. Influence of stand age on the magnitude and seasonality of carbon fluxes in Canadian forests. Agric. For. Meteorol. 165, 136–148. doi:10.1016/j.agrformet.2012.06.011

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C.T., Law, B.E., Meyers, T., Moncrieff, J., Moors, E., Munger, J.W., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agric. For. Meteorol. 107, 43–69.

FLUXNET, 2013. Introduction to Fluxnet [WWW Document]. URL http://fluxnet.ornl.gov/introduction (accessed 8.2.13).

Fluxnet-Canada, 2003. Fluxnet-Canada measurement protocols: Working draft, version 1.3, Flux. Laval, QC.

Foken, T., 2008. Experimental methods for estimating the fluxes of energy and matter, in: Micrometeorology. Springer-Verlag, Berlin Heidelberg, pp. 105–152.

Goulden, M.L., Munger, J.W., Fan, S.-M., Daube, B.C., Wofsy, S.C., 1996. Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. Glob. Chang. Biol. 2, 169–182.

Gu, L., Falge, E., Boden, T., Baldocchi, D.D., Black, T.A., Saleska, S.R., Suni, T., Verma, S.B., Vesala, T., Wofsy, S.C., Xu, L., 2005. Objective threshold determination for nighttime eddy flux filtering. Agric. For. … 128, 179–197.

Hollinger, D.Y., Richardson, A.D., 2005. Uncertainty in eddy covariance measurements and its application to physiological models. Tree Physiol. 25, 873–85.

Huang, S., Arain, M.A., Arora, V.K., Yuan, F., Brodeur, J., Peichl, M., 2011. Analysis of nitrogen controls on carbon and water exchanges in a conifer forest using the CLASS-CTEMN+ model. Ecol. Modell. 222, 3743–3760.

Kao, R., Gibson, C., Gallery, R., Meier, C., Barnett, D., Docherty, K., Blevins, K., Travers, P., Azuaje, E., Springer, Y., Thibault, K., McKenzie, V., Keller, M., Alves, L., Hinckley, E., Parnell, J., Schimel, D., 2012. NEON terrestrial field observations: designing continental-scale, standardized sampling. Ecosphere 3, 1–17.

Lasslop, G., Reichstein, M., Kattge, J., Papale, D., 2008. Influences of observation errors in eddy flux data on inverse model parameter estimation. Biogeosciences 5, 1311–1324.

Lloyd, J., Taylor, J., 1994. On the temperature dependence of soil respiration. Funct. Ecol. 8, 315–323.

Moffat, A., Papale, D., Reichstein, M., Hollinger, D., Richardson, A., Barr, A., Beckstein, C., Braswell, B., Churkina, G., Desai, A., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. 147, 209–232.

Moncrieff, J., Malhi, Y., Leuning, R., 1996. The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water. Glob. Chang. Biol. 2, 231–240.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., Yakir, D., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3, 571–583.

Peichl, M., Arain, M.A., 2006. Above- and belowground ecosystem biomass and carbon pools in an age-sequence of temperate pine plantation forests. Agric. For. Meteorol. 140, 51–63.

Peichl, M., Arain, M.A., Brodeur, J.J., 2010a. Age effects on carbon fluxes in temperate pine forests. Agric. For. Meteorol. 150, 1090–1101.

Peichl, M., Brodeur, J.J., Khomik, M., Arain, M.A., 2010b. Biometric and eddy-covariance based estimates of carbon fluxes in an age-sequence of temperate pine forests. Agric. For. Meteorol. 150, 952–965.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., Valentini, R., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Chang. Biol. 11, 1424–1439.

Richardson, A.D., Hollinger, D.Y., 2007. A method to estimate the additional uncertainty in gap-filled NEE resulting from long gaps in the CO2 flux record. Agric. For. Meteorol. 147, 199–208.

Richardson, A.D., Hollinger, D.Y., Aber, J.D., Ollinger, S. V., Braswell, B.H., 2007. Environmental variation is directly responsible for short- but not long-term variation in forest-atmosphere carbon exchange. Glob. Chang. Biol. 13, 788–803.

Richardson, A.D., Hollinger, D.Y., Burba, G.G., Davis, K.J., Flanagan, L.B., Katul, G.G., William Munger, J., Ricciuto, D.M., Stoy, P.C., Suyker, A.E., Verma, S.B., Wofsy, S.C., 2006. A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes. Agric. For. Meteorol. 136, 1–18.

Ruppert, J., Mauder, M., Thomas, C., Luers, J., 2006. Innovative gap-filling strategy for annual sums of CO2 net ecosystem exchange. Agric. For. Meteorol. 138, 5–18.

Tanner, C.B., Thurtell, G.W., 1969. Anemoclinometer measurements of Reynolds stress and heat transport in the atmospheric surface layer. Research and Devel- opment Tech. Report ECOM 66-G22-F to the US Army Electronics Command. Madison, WI.

Van Wijk, M.T., van Putten, B., Hollinger, D.Y., Richardson, A.D., 2008. Comparison of different objective functions for parameterization of simple respiration models. J. Geophys. Res. 113, 1–11.

Webb, E., Pearman, G., Leuning, R., 1980. Correction of flux measurements for density effects due to heat and water vapour transfer. Q. J. R. Meteorol. Soc. 106, 85–100.

4 Implications of commonly-used footprint and friction velocity filtering methods on data availability and carbon exchange estimates in a fetch-limited temperate forest

4.1 Abstract

When operated above forested sites, the simplifying assumptions of the eddy covariance (EC) method are not always met, leading to measurements that may not be representative of the true ecosystem-atmosphere carbon exchange. It is important to identify and filter out non-representative measurements from EC time series to ensure accurate carbon exchange estimates. Among the most relevant and unresolved filtering approaches are the application of the friction velocity threshold (u_*^{Th}) filtering to remove low-turbulence periods, and footprint filtering to identify potential source-area contamination by non-target surfaces. In this study, we use a data-driven, factorial analyses approach to characterize the differences among, and interaction between four analytical footprint models, and three u_*^{Th} estimation methods at a fetch-limited temperate forest. Methods were assessed in terms of their consequences for EC data quantity and distribution, as well as their effect on gap-filled ecosystem carbon exchange estimates. In general, u_*^{Th} methods varied in terms of their estimates, due to the use of dissimilar data aggregation statistics (i.e. mean, median, maximum), as well as their different sensitivities to the nature of input friction velocity values, which was influenced by the preceding footprint filtering application. In addition, our results showed substantial, fundamental differences between analytical footprint models – both generally, and in response to different assigned filtering stringencies. Our results also showed that footprint filtering had a significant effect on ecosystem carbon exchange estimates at our site. Over the years 2006 – 2011 , annual net ecosystem exchange was reduced between 59 and 207 g C m^{-2} y^{-1} by application of our desired footprint filter in comparison to no footprint filtering. Among the footprint models and stringencies investigated, differences in performance metrics and annual carbon exchange estimates were of a comparable magnitude to the effect of applying footprint filtering altogether. A decrease in annual ecosystem respiration and net ecosystem exchange with increasing footprint model stringency indicated differences in carbon-exchange dynamics between the study site of interest, and the surrounding non-target ecosystems. This difference was primarily driven by respiration discrepancies, since footprint filtering activities had the greatest effect on nighttime data availability and thus, respiration model parameterization. Our findings support further investigation and evaluation of footprint and u_*^{Th} filtering methods at our forest, and generally across EC measurement sites. Comparisons with other estimation methods – such as Lagrangian or large eddy simulation for footprint models – are needed to further evaluate and compare these methods, and understand their implications for EC-derived ecosystem understanding.

4.2 Introduction

The eddy covariance (EC) method has proven an unparalleled tool for studying the magnitudes and controls of ecosystem-atmosphere carbon exchanges at hundreds of sites across the world (Baldocchi et al., 2001). When operated in restrictive, ideal conditions, the true net ecosystem carbon exchange (NEE) is well-represented by the sum of EC-measured turbulent carbon dioxide (CO_2) exchange, F_c, and the change in CO_2 storage in the air column beneath the instrumentation, ΔS_c (Barr et al., 2013; Finnigan, 2006). These ideal requirements include stationarity of the data, homogeneity of the underlying surface, fully-developed atmospheric turbulence, and the absence of horizontal and vertical advection (Aubinet, 2008; Baldocchi, 2003; Foken and Wichura, 1996). In practical, long-term application, however, the simplifying assumptions of the EC method are not always met – especially when measurements are being made above complex, laterally-limited surfaces (Göckede et al., 2004; Rebmann et al., 2004; van Gorsel et al., 2009). In such conditions, NEE estimates produced by EC systems are not representative of true ecosystem fluxes: the failure to compensate for these effects (by data correction or removal) introduces systematic error and uncertainty into time-integrated NEE estimates and relationships derived between NEE and environmental controls (Foken and Wichura, 1996; Xuhui Lee et al., 2004; Moncrieff et al., 1996; Paw U et al., 2000).

Among the most important quality control considerations for forested sites is the identification of periods in which EC-fluxes are likely to be underestimated due to poorly developed atmospheric turbulence (Aubinet, 2008; Aubinet et al., 2000; Gu et al., 2005; Massman and Lee, 2002). Especially at night, the development of stable layers within and above a forest decouples it from the atmosphere above and increases the likelihood that CO_2 is transferred by non-turbulent mechanisms such as horizontal and vertical advection (Aubinet et al., 2003; Feigenwinter et al., 2008; Staebler and Fitzjarrald, 2004). Most commonly, these periods are identified using a friction velocity threshold (u_*^{Th}), where NEE measurements are filtered out for all periods where friction velocity (u_*) is below this critical value. A number of methods have been developed to estimate the value of u_*^{Th} using EC measurement time series (Barr et al., 2013; Goulden et al., 1996; Gu et al., 2005; Papale et al., 2006), and recent comparisons have shown considerable discrepancies between u_*^{Th} estimates produced by these methods (Barr et al., 2013). By controlling the number of potentially non-representative flux measurements that are excluded prior to gap-filling model parameterization, the value of u_*^{Th} exerts a significant control on time-integrated sums of NEE and its component fluxes (ecosystem respiration, RE; gross ecosystem productivity, GEP; Barford et al., 2001; Papale et al., 2006).

In settings where the surface of interest is limited in fetch and bounded by dissimilar ecosystems, it is often necessary to consider the potential contamination of EC flux measurements by non-target surfaces.

Since this technique relies on transport by turbulent atmospheric eddies, the source area of the measurement varies with changes in the prevailing wind direction, the underlying surface roughness, and the nature of atmospheric turbulence and stratification (Foken and Leclerc, 2004; Kljun et al., 2002; Leclerc et al., 2003; Schuepp et al., 1990; van de Boer et al., 2013). The variability of the so-called "flux footprint" is a reasonable concern for forest measurement studies in a number of ways: Where heterogeneity exists within the surface of interest, estimates of time integrated NEE sums and their relationships with environmental variables commonly need to be considered in the context of the prevailing footprint climatologies (Chen et al., 2011, 2009; Göckede et al., 2004). Footprint variability is also a concern where carbon exchange characteristics are sought for a specific forest stand that is areally-restricted (ie. fetch-limited), and bounded by dissimilar land uses or ecosystems. Constraining for ecosystem-representative carbon exchange estimates requires footprint filtering, where periods of non-representative measurements are predicted and removed prior to analyses.

Many types of footprint models have been developed to estimate the areal extent and source strength distribution of EC-measured fluxes; types of these approaches include analytical solutions (Hsieh et al., 2000; Kljun et al., 2004; Kormann and Meixner, 2001; Schuepp et al., 1990), Lagrangian models (Kljun et al., 2002; Leclerc and Thurtell, 1990), and large eddy simulations (Leclerc et al., 1997). For footprint filtering applications, these models can be used to estimate the within-fetch integrated footprint function (F_{in}) for a given time period; periods (typically half-hours) are filtered out when F_{in} falls below a selected footprint threshold (fp^{Th}). There currently exists no universally accepted and applied footprint model or fp^{Th} level, and numerous studies have shown considerable variability between these models in terms of their estimated footprint, and the response of these footprints to environmental influences (Kljun et al., 2003; van de Boer et al., 2013). This inter-model variability has important implications for fetch-limited ecosystems, as the selection of footprint models and fp^{Th} can have significant influence on which (and how many) data points remain after footprint filtering is applied. Thus, it is expected that the nature of, and confidence in carbon exchange information extracted from EC-measured data will vary according to footprint filtering specifications.

Furthermore, the reliability of flux partitioning methods (to extract component fluxes of ecosystem respiration, RE, and gross ecosystem productivity, GEP), and gap-filling models (to provide

time-integrated carbon exchange estimates), has been shown to degrade with increasing data gap size and length (Moffat et al., 2007; Richardson and Hollinger, 2007). In a fetch-limited setting where data removal may be substantial, the adequacy of data for these operations is an added constraint when considering appropriate footprint filtering specifications.

In the absence of difficult-to-implement artificial tracer experiments, a number of studies have successfully shown footprint models and source area appropriateness may be assessed using data-driven approaches. This includes the use of 'natural tracer' experiments, where distinct adjacent surfaces and multiple EC tower installations can be used to deduce source area contributions of EC measurements (Foken and Leclerc, 2004; Leclerc et al., 2003; Neftel et al., 2008; van de Boer et al., 2013). Integral turbulence characteristics and internal consistency tests (Göckede et al., 2004; Neftel et al., 2008; Thomas and Foken, 2002), have been suggested and applied to provide information on EC measurement quality and the degree to which they are influenced by multiple heterogeneous surfaces.

Considering an absence of universally-accepted footprint and friction velocity filtering methods, and the incomplete understanding of their applicability for all sites, it is important to investigate these filtering methods in a site-specific context. Doing so provides information on the sensitivity of ecosystem estimates to changes in source area data characteristics, and allows for further investigation of the performance and interactions of footprint and u_*^{Th} filtering methods.

To this end, this study develops and uses a data-driven approach to compare footprint- and friction velocity-filtering methods in terms of their classification of ecosystem-representative data in a fetch-limited temperate forest. Building on these approaches, this study also presents and implements methods to evaluate a number of analytical footprint models and footprint threshold levels against multiple, competing criteria: a) maximizing the ecosystem-representativeness of time-integrated ecosystem carbon exchange estimates; and, b) minimizing gap-filled estimate uncertainty as a result of insufficient data. Specifically, this study addresses the following questions:

1. How is the quantity and distribution of NEE time series data affected by the application of different footprint and friction velocity filtering approaches?

2. What is the sensitivity of annual ecosystem carbon exchanges (NEE, RE, GEP) to changes in footprint- and friction velocity-filtering applications in this fetch-limited temperate forest?

3. What footprint filtering approaches best satisfy the simultaneous requirements for maximizing ecosystem representativeness while minimizing gap-filling model error?

4. What are the implications of various footprint filtering specifications on derived ecosystem carbon exchange estimates and environmental relationships?

In this study, we applied four commonly-used analytical footprint models to estimate the effective upwind source area for six years of half-hourly flux measurements at a fetch-limited, monoculture forest. For each footprint model (fp_m), a range of footprint threshold (fp^{Th}) values were applied to investigate the effect of footprint model stringency on footprint filtering. A factorial experiment was carried out, where EC carbon flux measurements were filtered according to each $fp_m * fp^{Th}$ combination, and the effect of each filtering specification was characterized for subsequent processing steps. The effect of footprint filtering on u_*^{Th} determination was investigated across three different friction velocity threshold estimation approaches. For each $fp_m * fp^{Th} * u_*^{Th}$ method filtering combination, a gap-filling model was run to produce continuous half-hourly estimates and annual sums of NEE, RE and GEP.

The impact of footprint and friction velocity filtering on estimated carbon exchanges were subsequently explored, as two different non-linear gap-filling models were run on the filtered data to produce annual estimates of NEE, RE, and GEP, as well as statistical metrics of goodness-of-fit and bias error. Analyses of variance (ANOVAs) were performed on factorial results to estimate controls on annual estimates, and the statistical performance metrics obtained for each filtering combination run. Using gap-filling model performance metrics as an evaluation criteria, we then determined the most appropriate filtering parameters (fp_m, fp^{Th} level) for our fetch-limited forest site. Following this, we explored the implications of this selection by comparing carbon exchange estimates and environmental relationships derived using the ideal filtering parameters to those obtained using different combinations.

4.3 Description of filtering methods

4.3.1   Footprint models

Four different commonly-used analytical footprint models were selected to provide half-hourly diagnostic estimates of effective upwind source area for this study. These methods seek to solve the advection-diffusion equation through analytical means, and are based on semi-empirical relationships and Monin-Obukhov similarity theory. As such, these models are well-supported by surface-layer-based experimental comparisons with field measurements (e.g. tracer experiments), or simulation experiments such as Lagrangian stochastic or large eddy simulations (Kljun et al., 2003).

These four specific footprint models were chosen based on the following three criteria: a) their relatively simplistic nature, which, unlike Lagrangian or large eddy simulation (LES) models, allow the models to be run over many simulations in a realistic computing time frame; b) their frequent use as diagnostic tools in flux measurement research; and, c) the availability of pre-constructed footprint model code for these models (supplied either by the original author, or researchers that have used the model to create published results), thus increasing the likelihood that the model outputs are comparable to those published in other studies. These models are summarized in Table 4.1, and discussed in further detail below:

Table 4.1: Description of footprint models compared in this study

Model	Denoted as	Inputs	Output
Schuepp et al., (1990)	SP90	u, u_*, H, Ta, RH, APR, z_m, h_{tree}, z_0, r	1-dimensional crosswind-integrated cumulative flux contribution
Kljun et al., (2003)	KL03	z_m, z_0, h_{BL}, σ_w, u_*, r	1-dimensional crosswind-integrated cumulative flux contribution
Hsieh et al., (2000)	HS00	u_*, H, Ta, z_m, h_{tree}	1-dimensional crosswind-integrated cumulative flux contribution
Kormann and Meixner, (2001)	KM01, KM01-2D	z_m, z_0, u, u_*, σ_v, L	2-dimensional (along- and cross-wind) footprint flux contribution; 1-dimensional crosswind-integrated cumulative flux contribution

The footprint modeling approach of Schuepp et al. (1990; SP90) builds upon an analytical solution to the advection-diffusion equation provided by Gash (1986). This approach uses an approximate characterization of the diffusion equation developed by Calder (1952) to estimate a one-dimensional footprint as the relative contribution of upwind sources at distance x, to fluxes measured at point $(0,z)$ (Eq. 9 in Schuepp et al., 1990). By integrating this function in the upwind direction, the proportional contribution of all sources within a specified distance may be estimated; this value is defined as the cumulative normalized footprint of the flux measurement (CNF; Eq 13 in Schuepp et al. (1990)). Since this function is restricted to thermally neutral stratification and a uniform wind field (constant U and K), a momentum stability correction function, taken from Dyer (1974) and Dyer and Hicks (1974) is applied to the cumulative footprint function. The correction is expressed as:

$$\Phi_m = [1 - 16\frac{(z-d)}{L}]^{\frac{-1}{4}} \tag{4.1}$$

, where z is measurement height, d is displacement height, and L is the Obukhov length. Though this correction factor adjusts the footprint function peak and relative upwind contributions in a manner that corresponds to to atmospheric stability conditions, its application has been criticized as mathematically baseless (Kormann and Meixner, 2001). The predicted shape and length of the footprint function is controlled by changes in measurement height, surface roughness and atmospheric stability. Comparisons with Lagrangian stochastic simulations (Leclerc and Thurtell, 1990) showed acceptable analytical model performance over short and tall vegetation, though discrepancies generally increased with measurement height.

The footprint model described by Kljun et al. (2004; KL04) is an approximate analytical model developed through parameterization of a previously-developed and tested Lagrangian stochastic dispersion model (see Kljun et al., 2003; Kljun et al., 2002). The analytical model uses a scaling procedure to collapse the crosswind-integrated flux footprint for all stability conditions (from stable to convective) and a given roughness length, into an ensemble of similar curves that can be constructed from a few, commonly-measured environmental variables (Kljun et al., 2004). As a result, model predictions of the one-dimensional cumulative footprint function are algebraically estimable and physically valid for all types of stratification, and calculations may be carried out with relatively little computational cost.

As an improvement to approximate analytical models of Gash (1986) and Horst and Weil (1992 and 1994), Hsieh et al. (2000; **HS00**) developed an approximate analytical model based on a combination

of dimensional analysis and testing against a Lagrangian stochastic dispersion model (Thomson, 1987). The resulting model provides a simplistic, accurate and computationally-efficient estimate of the one-dimensional crosswind-integrated flux footprint as a function of atmospheric stability, measurement height and roughness length. Assessment of the model-estimated footprint showed good agreement with field-based measurements, as well as with outputs from more complex analytical and Lagrangian dispersion models (Hsieh et al., 2000).

The Kormann and Meixner (2001; KM01) footprint model produces a two-dimensional analytical solution to the advection-diffusion equation by using a stationary gradient diffusion foundation with a height-independent crosswind dispersion function. The KM01 model – based on modifications of the analytical models of Van Ulden (1978) and Horst (1999) – assumes independence of vertical and crosswind dispersion, in order to reduce the continuity equation to two dimensions, where the crosswind dispersion is represented as a Gaussian plume (Kljun et al., 2003; Kormann and Meixner, 2001; van de Boer et al., 2013). The simplifying assumptions of the KM01 model allow it to be applied to long-term measurement data with minimal computing resources.

In a comparison test with a Lagrangian stochastic dispersion model, Kljun et al. (2003) found satisfactory correspondence between the KM01 model and Lagrangian footprint estimates when the receptor was located within the surface layer (z_m = 20 m in their study); however, the KM01 estimates for footprint distance and distance to peak were consistently longer, an effect attributed to the exclusion of longitudinal turbulent dispersion in the analytical model. Considerable discrepancies arose when the receptor was mounted above the surface layer (z_m = 100 m). Assessment of KM01 against a field-based artificial tracer experiment (van de Boer et al., 2013) similarly found the model to overestimate footprint distances (effective fetch), and the distance to footprint peaks.

4.3.2   Friction velocity threshold estimation methods

This study implemented and tested three commonly-used methods for estimating friction velocity threshold (u_*^{Th}). Each model is described below:

The Moving Point Test method (**MPT-P**), introduced by Reichstein et al. (2005), and developed further by Papale et al. (2006) was used in this study to estimate u_*^{Th} from the relationship between nighttime Net Ecosystem Exchange (NEE_n) and u_*. For this methodology, data from each year was stratified into four, equally-sized 3-month seasons (n_S=4: JFM, AMJ, JAS, OND), and then sub-stratified into six

temperature classes (n_T=6) within each season. The purpose of creating separate n_S*n_T strata was to reduce the confounding effect of seasonality and environmental temperature on the u_* vs NEE relationship. Within each n_S*n_T strata, data was split into 20 equally-sized u_* classes, and the u_*^{Th} for a strata was estimated as the lowest u_* class that had an average NEE_n value within 99% of the average NEE_n at all higher u_* classes. Seasonal u_*^{Th} was calculated as the median of predicted values across the temperature substrata, while a single, annual u_*^{Th} value was estimated as the largest of the seasonal estimates. For the purposes of this study, alternative annual u_*^{Th} estimates were generated by also recording estimates obtained by using the median and mean as the value selection statistic for the seasonal values. To obtain information on the potential variability in u_*^{Th} prediction, the estimation procedure was bootstrapped at the seasonal level, to produce 100 separate estimates of mean, median and maximum u_*^{Th}. As distributions of these values were normal, the mean of each of these bootstrapped sets was taken as the final annual u_*^{Th} value.

The Change Point Detection (**CPD**) method described by Barr et al. (2013), is based upon a similar theoretical approach as MPT-P, though with modifications to the stratification and u_*^{Th} detection procedures. To improve temporal resolution of the estimates, annual half-hourly data was subdivided into more seasonal strata (than MPT-P) by shortening the window size, and prescribing a 50% overlap between adjacent windows. For the purposes of this study, a total of n_S=12, 2-month "seasonal" windows were used, as a reasonable value within the range of n_S=7-17 used in Barr et al., (2013). Within each seasonal window, data was further stratified into n_T=4 temperature classes. For each n_S*n_T substrata, data was split into 50 u_* bins with at least 5 data points per bin. The value of u_*^{Th} within each substrata was estimated using the CPD technique originally developed by Solow (1987), which was later refined by Lund and Reeves (2002) and Wang (2003), and subsequently applied to u_*^{Th} determination by Barr et al, (2013). This technique attempts to identify u_*^{Th} as the statistically-optimized "change-point" (x_c) where the linear relationship between u_* vs NEE changes from non-zero (at lower u_* values) to zero (at all higher u_* values). Multiple quality assurance tests were applied to remove n_S*n_T estimates where the change-point selection showed poor goodness of fit, or where the non-zero slope was of opposite sign to the majority of the n_S*n_T estimates. Two different annual u_*^{Th} were created by using two different value selection statistics: the mean of all quality-assured n_S*n_T substrata estimates (as used in Barr et al., 2013); and, by the median of these values (our modification). Similar to the MPT-P approach, this procedure was bootstrapped 100 times at the seasonal level, and the means of these bootstrapping results were used at the final annual u_*^{Th} estimates.

The Gu et al. (2005) Moving Point Test method (**MPT-G**) was also applied in this analysis, providing a third estimate of u_*^{Th} using a methodology that is relatively unique from those of MPT-P and CPD. The method uses similar techniques to estimate two friction velocity thresholds: a low-turbulence threshold for removing flux data measured during low-u_* conditions; and, a high-end threshold, for removing data measured during exceptionally high u_* conditions, where "pressure pumping effects" may occur (Gu et al., 2005; Rogie et al., 2001). Since no pressure pumping effects were observed at our site, the remaining description concerns only the low-turbulence u_*^{Th} value, which is physically equal to the threshold estimated by MPT-P and CPD. MPT-G uses a nested loop, iterative approach to calculate seasonal ($n_S = 4$) u_*^{Th} values. In the outer loop, NEE_n is normalized by a temperature response function developed using data from all half-hours in which u_* was between the current low- and high- u_*^{Th} values (initially 0 and 9999). Data was then ranked from low- to high-u_*, and exported to the inner loop. In the inner loop, a moving window progressed point-wise from lowest to highest u_* data. At each step, the median of normalized NEE_n within the window was tested (via a t-test) for statistical equivalence to the median of all available NEE_n. The u_*^{Th} was calculated as the median value of u_* for the first moving window that was not statistically different from the median of the entire set. This u_*^{Th} value was exported to the outer loop, and used to filter the data available for developing the temperature response function ($u_* < u_*^{Th}$ removed). The process described above continued until the inner and outer loop u_*^{Th} values converged, at which point the process ended with a seasonal u_*^{Th} estimate. Estimates for annual u_*^{Th} was extracted by calculating the mean and median of the four seasonal estimates. This process was repeated 100 times by bootstrapping seasonal data, and the means of the two types of estimates were used as the final annual u_*^{Th} estimates.

4.4 Methods

4.4.1   Site description

Data used for this study were collected at the TP39 forest of the Turkey Point Flux Station research sites (referred to as CATP4 in global FLUXNET notation). TP39 is a 74-year old planted and managed eastern white pine (*Pinus strobus* L.) forest, established in 1939 after the clearing of former oak savanna lands for afforestation purposes (Peichl, Arain, et al., 2010a). The site is located near the North Shore of Lake Erie in Norfolk County (42.71°N, 80.36°W), in southern Ontario, Canada.

The TP39 forest stand is approximately rectangular (see Figure 4.1), with an average length of 650 m

(in NNW-SSE direction), and width of 610 m (in WSW-ENE direction). Three adjacent 1 ha blocks of *Q. vultina* exist within the rectangular outline of the stand, located directly east of the measurement tower. The TP39 forest (outlined in Figure 4.1) is part of the Turkey Point tract of the St. Williams Conservation Reserve. The Conservation Reserve is a stretch of managed forests stands of white pine, red pine (*Pinus resinosa*), Eastern black oak (*Quercus vultina*), and Red oak (*Quercus rubra*), placed within a region of predominantly cash-cropped agricultural fields and scattered Carolinian Hardwood forests. In its immediate vicinity, TP39 is bordered by forest stands that vary in terms of their areal extent, shape, tree age and species, and the management practices undertaken within. For the purposes of this study, the "target" surface for tower-based flux measurements is defined as the combined extents of the white pine forest of TP39 and the 79-year old forest (established 1934) located to its immediate west; all other stands and land-uses are considered out of bounds. As both included stands are comparable in terms of site and management history, soil characteristics, topography, tree age and species assemblage, amalgamating these for the purpose of this study expands the fetch available for tower-based flux measurements. Such an approach has been used in previous studies that have reported EC fluxes from this location (Peichl, Arain, et al. 2010a; Arain and Restrepo-Coupe, 2005).

Forest overstorey at this site is dominated by planted Eastern White Pine (*P. strobus*) trees, while naturally-generated Balsam fir (*Abies balsamea* L. Mill), Oak (*Quercus velutina* L., *Quercus alba* L.), Red maple (*Acer rubrum* L.) and Wild black cherry (*Prunus serotina* Ehrh.) comprise the intermediate and understorey layers (Peichl and Arain, 2006). As reported in Peichl and Arain (2006), mean co-dominant tree height at this site is approximately 22 m, with a leaf area index (LAI) of 8.0 (Chen et al., 2006) and a tree density of 421 ± 166 trees ha^{-1}. This spatial distribution of co-dominant trees within the forest has been influenced by a thinning application in 1983, where 105 m^3 ha^{-1} of wood volume was removed from the forest, leaving distinguishable, regular gaps in the canopy (Peichl and Arain, 2006). The site is located on predominantly flat topography, with a mean elevation of 184 m asl, and a maximum elevation variation of 2 metres. The soil is a well-drained Brunisolic Gray Brown Luvisol (Presant and Acton, 1984) that rests upon deep (> 10 m) eolian-deposited medium and fine sands. The water table typically resides between 6 and 8 metres below the soil surface.

The long-term wind rose included in Figure 4.1 indicates that westerly wind patterns dominate at this site, with winds most commonly originating from the southwestern quadrant. In general, the most common wind directions correspond well to the directions that have the largest fetch from the measurement tower to the forest boundary.
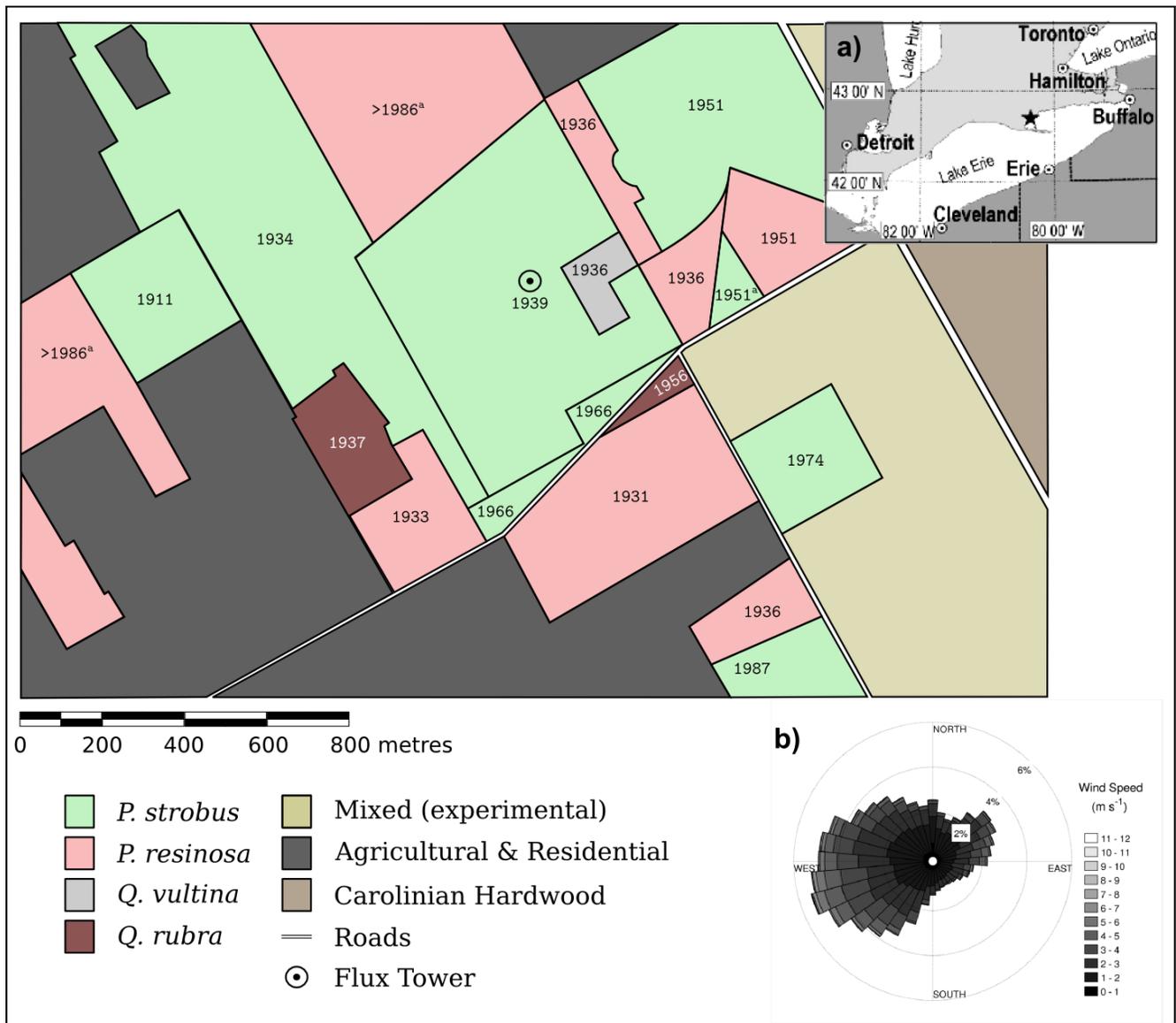
Figure 4.1: *Land use map of the region surrounding the flux tower used in this study. Managed forest plots are labeled with their date of establishment. Inset map a) shows the location of the study region in southern Ontario. Inset map b) displays the long-term (2006-2011) mean wind rose.*
[a] establishment date unknown, but estimated using historical land use records.

4.4.2   Ecosystem flux and meteorological measurement

Meteorological and flux smeasurements for the forest stand were collected from a 26 m walk-up tower located near the centre of the TP39 forest plot (Figure 4.1). The tower-based meteorological system consisted of instrument assemblages specified by Canadian Carbon Program (formerly Fluxnet-Canada Research Network) measurement guidelines and standards (Fluxnet-Canada, 2003). Primary meteorological variables were measured continuously throughout the year from instruments mounted on booms extending 1-3 m horizontally away from the tower, at 28 m above the forest floor. The measured variables included: air temperature and relative humidity (Ta, RH; model HMP45C, Campbell Scientific Inc. (CSI), Edmonton, AB, CA); wind speed and direction (u, WD, model 05 103-10RE, RM Young Co. (RMY), Michigan, USA); down- and up-welling Photosynthetic Photon Flux Density (PPFD, $PPFD_u$, model PAR-LITE, Kipp and Zonen Ltd. (KZ), Delft, NL); net radiation (Rn, model NR-LITE, KZ); down- and up-welling short and long wave radiation (model CRN1, CSI); atmospheric pressure (model 61205V, RMY). Secondary meteorological variables (e.g. Vapour Pressure Deficit, VPD), were calculated from primary variables during data processing (see below).

Precipitation was measured from two monitoring locations: The primary ground installation was located at a nearby station (2 km from the tower), where an all-weather accumulation precipitation gauge (model T200B, Geonor Inc., Milford, PA, USA) was operated in the centre of a 4 ha clearing. Data was recorded at half-hourly intervals by a standalone data logger (model CR10x, CSI). A secondary system consisted of a tipping-bucket rain gauge (model 52202, RMY; model CS700, CSI) affixed to the top of the measurement tower; measurements from this system were used as a validity check for precipitation events recorded at the primary location.

Soil temperature (T_s) profiles were measured at two different locations using thermistor probes (model 107B, CSI) placed at depths of 2, 5, 10, 20, 50 and 100 cm. At the same two locations, volumetric water content (VWC) profiles were measured using water content reflectometers (model CS-615/616, CSI) inserted horizontally at depths of 5, 10, 20, 50 and 100 cm. The VWC profiles were used to calculate depth-weighted, upper root-zone averages (VWC_{30}) for the top 30 cm of each profile. Soil heat flux plates (model HFT3, CSI) were installed at 3 cm depth at each location to estimate heat exchanges (G) across the atmosphere-soil boundary; 2 cm soil temperature was used to correct the estimates for soil heat storage changes above the sensor.

Half-hourly averages (and precipitation sums) of all atmospheric and edaphic variables were recorded

by data loggers (models CR23x, CSI), and downloaded automatically by an on-site desktop computer. Raw data was collected from the field at weekly intervals, and subsequently processed by custom programs in MATLAB (The Mathworks, Inc, Natick, MA) that organized, cleaned and quality-assured data. For micrometeorological data used in analyses operations (i.e. Ta, RH, Ts, VWC_{30}, PPFD) small data gaps were gap-filled using a linear regression with values measured at the nearby (<1 km distance) 40 year-old age-sequence study site.

A closed path eddy covariance (CPEC) system was operated continuously at TP39 to measure mass and energy exchanges throughout the study period. The system consisted of a CSAT-3 sonic anemometer (model CSAT-3, CSI) and infrared gas analyzer (IRGA, model LI-7000, LI-COR Biosciences, Lincoln, NE, USA), which were operated and logged at high frequencies (20 Hz) on a desktop PC, using custom software created by the Biometeorology & Soil Physics Group at the University of British Columbia (British Columbia, CA). High-frequency flux data were collected from the field computer at weekly to bi-weekly intervals; this data was used by customized MATLAB programs to calculate half-hourly fluxes. Outliers in flux data were removed according to Papale et al. (2006). Net ecosystem carbon exchange (NEE), was calculated as the sum of eddy covariance-estimated CO_2 flux (F_c), and air column storage flux (ΔS_c). ΔS_c was estimated following Barr et al., (2004), using half-hourly CO_2 concentrations measured at the top (z = 28 m), middle (z = 14 m) and bottom (z = 2 m) of the tower.

4.4.3   Flux footprint model standardization

All footprint models (outlined in section 4.3.1) were evaluated using MATLAB software (The Mathworks, Natick, MA). The 1D footprint models (SP90, KL04, HS00) were standardized to output half-hourly estimates of the crosswind-integrated, within-bounds cumulative flux footprint proportion (F_{in}), for years 2006-2011. This was calculated by integrating the footprint function in an upwind direction between the measurement tower and the forest bounds. The distance from the tower to the forest bounds was estimated using a look-up table of measurement fetch distances computed for every 0.1 degree of wind direction. The fetch distance look-up table was developed using remotely sensed imagery and Geographical Information Systems.

The KM01 model was configured to output flux source strength (2-D footprint function) at a 2 m grid resolution, within a 2 x 2 km grid centred around the flux measurement tower location. The 2-D KM01 model was run at a half-hourly time step for years 2006-2011. The proportion of flux footprint

contained within the forest bounds (F_{in}) was calculated for each half-hour as the ratio of summed source strength in all grid cells located within the forest bounds, to the total sum of the footprint function. A 2 m resolution digital map of the forest bounds was used to delineate cells within and without the forest. In addition to its 2D footprint output, the KM01 model was configured to estimate half-hourly, one-dimensional, crosswind-integrated F_{in} (modified model denoted as KM01-1D). This was calculated by summing the 2D footprint estimate over the entire lateral (crosswind) extent, for cells spanning the distance from the measurement tower to the upwind extent of the forest bounds. Creating this 1D output allowed for a more direct comparison with the other (1D) models, particularly since it enables an investigation of forest boundary-induced discrepancies between the 1D and 2D estimates of within-bounds flux proportion.

As demonstrated in Figure 4.2, the orientation and shape of a footprint plume relative to the boundaries of the surface of interest may lead to non-negligible over- or under-estimation of crosswind-integrated F_{in} for a given period of interest, depending on boundary shape, and dynamic variables such as wind direction, atmospheric stability and turbulence conditions. By comparing the half-hourly F_{in} estimates between the 1D and 2D KM01 models, it is possible to estimate the sign and magnitude of this effect, and subsequently correct one-dimensional estimates for this error.
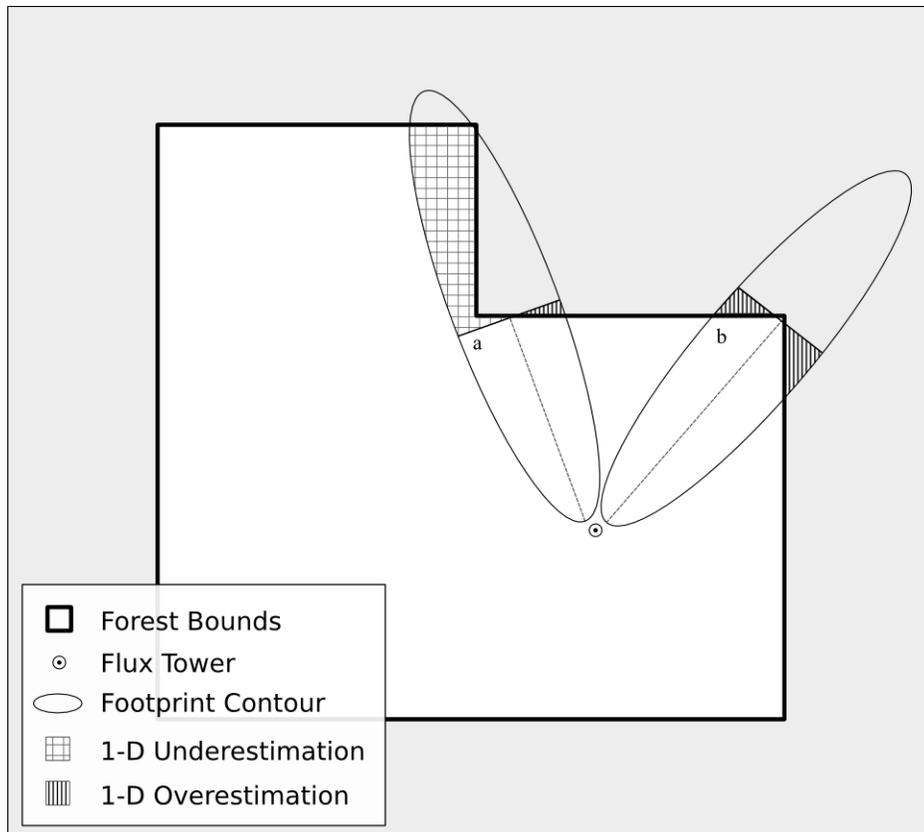
Figure 4.2: *Schematic diagram demonstrating over- and underestimation of crosswind integrated (1D) F_{in} resulting from the orientation of the footprint plume to the forest boundary. Two theoretical flux footprints are shown here to display the discrepancy between the two-dimensional (ellipse) and crosswind integrated (mid-line and perpendicular chord) estimates of F_{in}. In comparison to the 2D footprint estimate, the crosswind integrated estimate of a) shows an overall underestimation of F_{in}, while b) demonstrates an overall overestimation.*

At the highest level of data aggregation (Figure 4.3), half-hourly F_{in} estimates of KM01 and KM01-1D are strongly correlated (R = 0.958 for F_{in} > 0.05; R^2 = 0.916; mean slope of 1.03); however non-random discrepancies are observed between the two estimates. The systematic modifying effect of wind direction on this relationship is apparent when data is stratified into 5° wind direction bins (Figure 4.4).
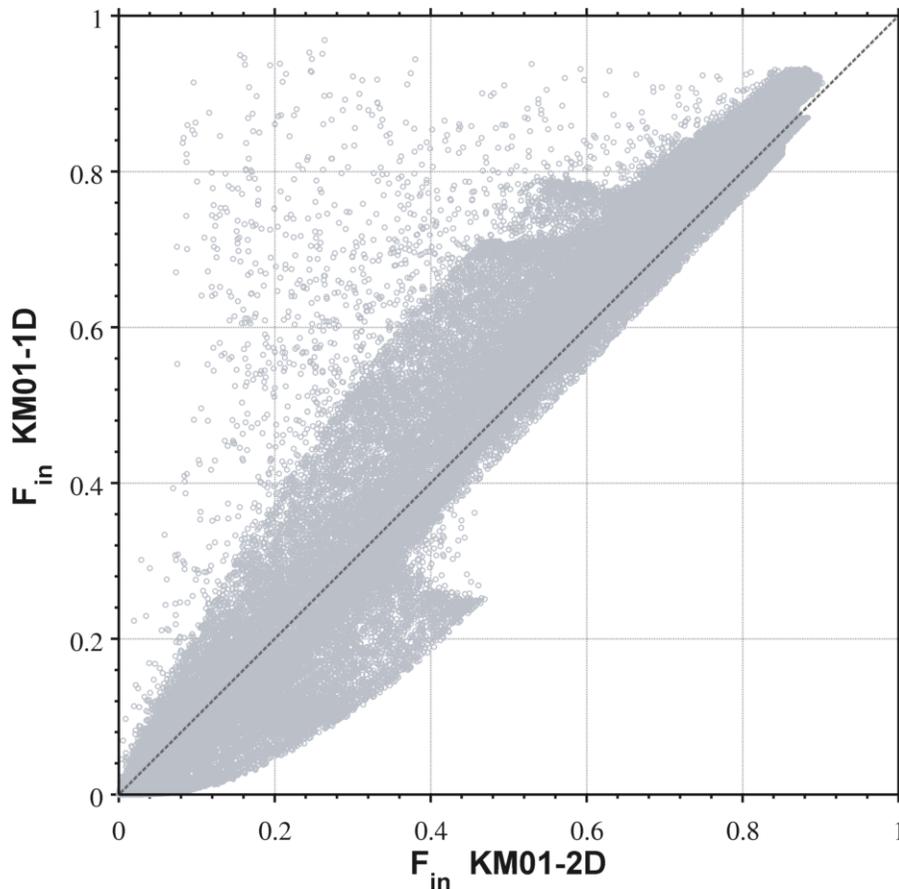
Figure 4.3: *Scatter plot comparing half-hourly F_{in} estimated using the crosswind integrated KM01-1D and two-dimensional KM01-2D footprint models. The dashed line indicates a 1:1 relationship.*

Within each strata, the KM01-KM01-1D relationship is highly-correlated and strongly linear (mean R = 0.972; mean R^2 = 0.934), with intercepts (-0.1 to 0.1) and slopes (0.8 to 1.2) that vary systematically and considerably with wind direction. Though the presence of outliers in the KM01 versus KM01-1D relationship degrades linear regression R^2 values at higher F_{in} values, the improvement resulting from wind direction stratification becomes increasingly evident, as R^2 improves from 0.49 to 0.62 with stratification for periods with F_{in} > 0.5. Using these wind direction-specific linear relationships, a look-up table of correction factors was generated, and each half-hourly one-dimensional footprint estimate was adjusted to align with the 2D model output. Application of the correction resulted in changes to F_{in} that ranged from a 25% reduction to a 22% increase, with a mean reduction of 4% (Figure 4.5). Overall, nearly 80% of all half-hours were corrected to a smaller F_{in} value, implying that applying the (uncorrected) 1D foootprint model commonly leads to overestimation of F_{in}.
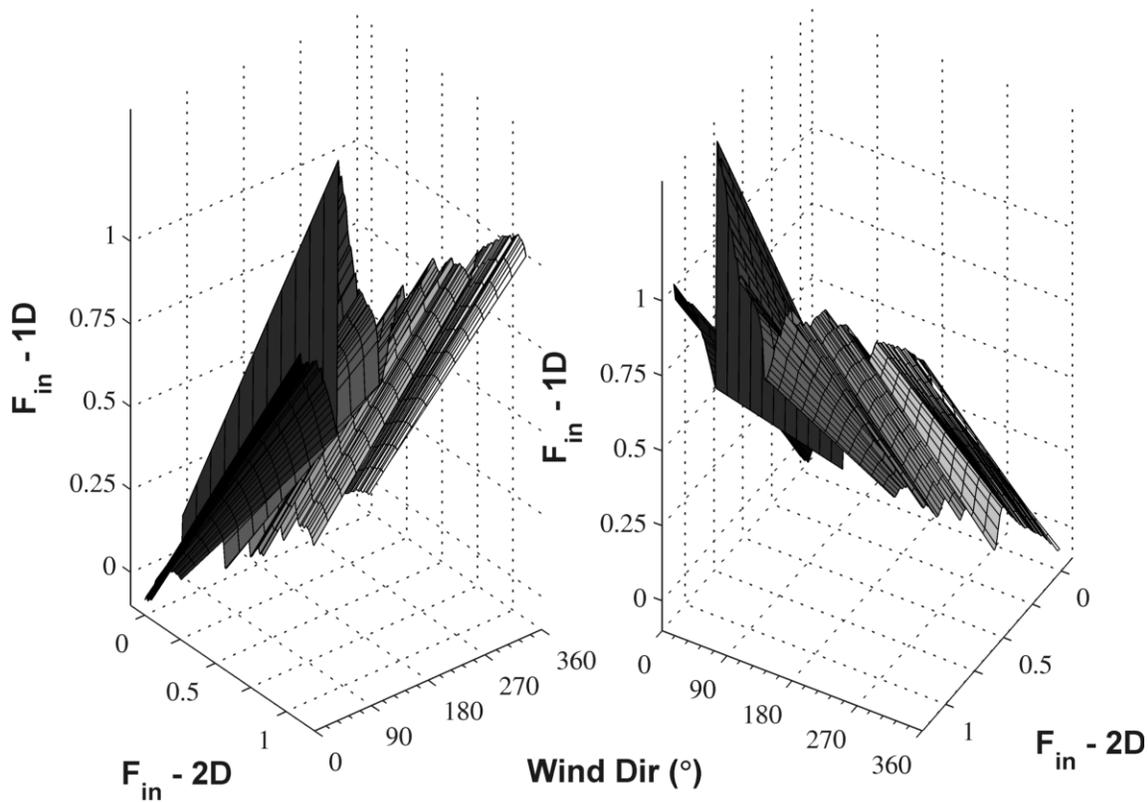
Figure 4.4: *Two perspectives of the fitted linear relationship between KM01-1D and KM01-2D F_{in} estimates, calculated for each 5° wind-direction bin.*

Acknowledging that the other 1D footprint estimation models (SP90, KL04, HS00) are subject to similar boundary-induced discrepancies, the correction method developed for KM01-1D results were applied to these model outputs as well. Though each of these 1D models assumes different (or no) analytical representations of the crosswind footprint component (and thus, their crosswind footprint plume shape likely differ), sensitivity testing with the KM01 models revealed that the crosswind width of the plume had minimal effect on the wind-controlled relationship between the 1D and 2D estimates (data not shown). This is further supported by the high coefficient of determination for the 1D-2D relationship for each wind-direction strata (Figure 4.4), despite a wide range of predicted plume widths (controlled by σ_w) contained within. As a result, it was determined that application of this correction factor would have a net positive effect on the inter-comparability of the footprint model estimates.
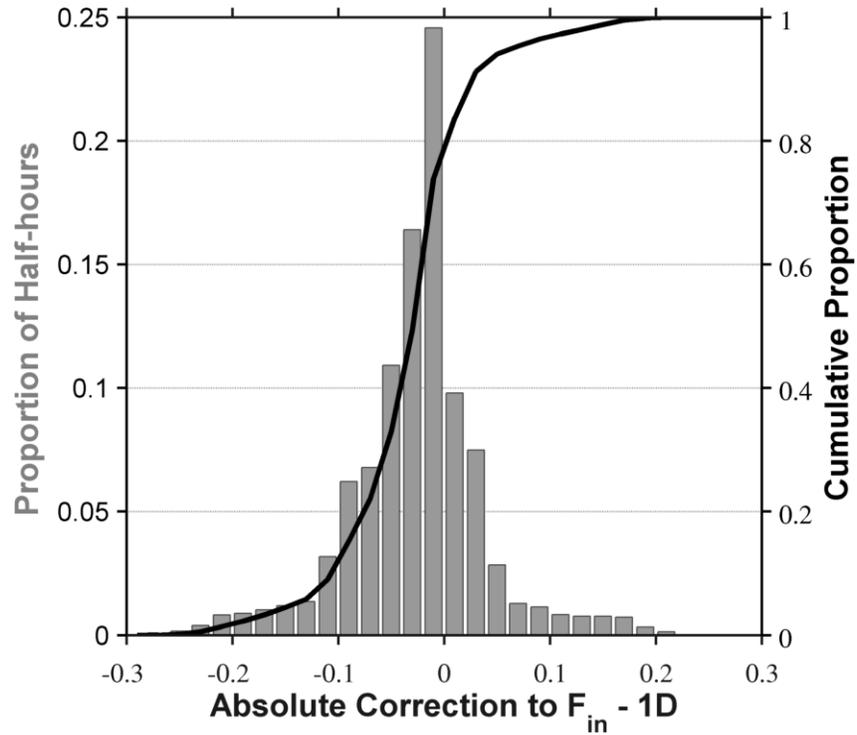
*F**igure 4.5:** Histogram (bars, left axis) and cumulative proportion (black line, right axis) of corrections applied to KM01-1D footprint estimates to align with KM01-2D predictions. Corrections are stated as absolute values.*

4.4.4   Data Filtering

The corrected footprint model F_{in} estimates were used to generate model-specific filtering keyfiles by comparing half-hourly values to a specified minimum within-bounds footprint threshold (fp^{Th}). Key files consisted of annual column vectors of half-hourly entries corresponding to flux measurements. Key file entries were set to 1 during periods where the model-estimated F_{in} was greater than fp^{Th}, and to NaN otherwise. Flux data were effectively footprint filtered by multiplying relevant data columns with the footprint key file – thereby replacing data entries for unacceptable half-hours with NaN values, and excluding them from further operations. Multiple keyfiles were created for each footprint model, corresponding to the selection of different fp^{Th} values during filtering. A total of 10 keyfiles were created for each model: one for the no-footprint case (nofp, $fp^{Th} = 0$), and one for each 0.05 increment between $fp^{Th} = 0.5$ to 0.9, inclusively.

The annual u_*^{Th} estimates of each method were used to create u_* filtering keyfiles. Similar to the footprint filtering approach, the keyfiles consisted of annual files of half-hourly entries that corresponded to flux measurements. Key file entries were set to 1 for half hours where measured u_* was greater than the predicted threshold, or to NaN for all other periods. Multiplying a column vector of flux measurements by this key file effectively removed all half hours with u_* below the indicated minimum threshold.

4.4.5   NEE gap-filling and random error estimation

Annual datasets of quality-controlled and filtered NEE were filled using two common non-linear regression gap-filling models: The Fluxnet-Canada Research Network method (NLR-FC), described by Barr et al. (2004); a modified version of the nonlinear estimation model used by Richardson et al. (2007) for the Howland Research Forest data (NLR-HL). These methods were selected due to their common use in studies on temperate forests, and history of use for gap-filling purposes at this site (e.g. Arain and Restrepo-Coupe, 2005; Peichl et al., 2010b). Each method is described in further detail below:

In the NLR-FC method, measured RE is assumed to be equal to NEE during periods when GEP is zero (i.e. at night and during daytime periods when both air (T_a), and 5 cm soil (T_s) temperatures are less

than 0°). The NLR-FC method fits RE to an empirical logistic relationship with T_s as:

$$RE = \frac{r_1}{1 + \exp[r_2(r_3 - T_s)]} \, r_w(t) \tag{4.2}$$

where RE is model-estimated ecosystem respiration, and r_1, r_2, and r_3 are fitted empirical parameters. $r_w(t)$ is an additional fitted parameter that varies according to the slope of a linear regression (forced through the origin) between modeled and measured RE in each 100-point moving window. The time-varying parameter corrects for the presence of sustained (autocorrelated) biases in modeled estimates of RE.   Values of RE are used in the place of gaps in the annual measured RE time series (caused by equipment malfunction, quality-assurance testing, filtering, etc.), as well as to estimate respiration during periods of non-zero GEP (when RE cannot be directly measured).

GEP was determined as the difference between gap-filled RE and measured NEE, and subsequently set to zero for all nighttime periods and daytime half hours when $T_a, T_s < 0$. A rectangular hyperbolic, Michaelis-Menten relationship is used to model GEP as:

$$GEP = \frac{\alpha Q_d \beta}{\alpha Q_d + \beta} \, p_w(t) \tag{4.3}$$

Where Q_d is down-welling photosynthetic photon flux density (PPFD, µmol m^{-2} s^{-1}), and the fitted parameters α and β represent the quantum yield and photosynthetic capacity, respectively. To correct for sustained autocorrelated bias in model predictions, $p_w(t)$ is a time-varying parameter that varies according to the slope of a linear regression (forced through the origin) between modeled and measured GEP for each 100-point moving window.

The NLR-HL gap-filling method is a modified version of the model described by Richardson et al., (2007), which was later modified and applied to TP39 by (Peichl et al., 2010a). RE measurements are identified from NEE using the same conditions as NLR-FC (NEE nighttime and daytime with $T_a, T_s < 0$). RE is modeled as a function of T_s and VWC_{30} according to the relationship:

$$RE = R_{10} \times Q_{10}^{\frac{(T_s - 10)}{10}} \times f(VWC_{30}) \tag{4.4}$$

, where R_{10} and Q_{10} are fitted temperature response parameters that describe the relationship between

RE and T_s. $f(VWC_{30})$ is a sigmoidal function that characterizes the role of VWC_{30} in modifying the temperature response of RE as:

$$f(x) = \frac{1}{[1+\exp(\theta_1 - \theta_2 x)]}$$

(4.5)

θ_1 and θ_2 are fitted parameters that allow this term to range between [0,1] as a function of the independent variable x (VWC_{30} in this case), thus acting as a scaling function on the Ts-RE relationship. Including a VWC_{30} control on the RE model was found to provide a statistically significant improvement to model performance during periods of low VWC (typically mid-summer, data not shown). The two-parameter Q_{10} temperature response model was chosen instead of three-parameter functions (e.g. logistic or Lloyd and Taylor, 1994), to increase model parsimony and address issues of over-fitting and equifinality (Richardson et al., 2007).

GEP was estimated in the same way as for NLR-FC (RE - NEE; zero during nighttime periods, and daytime half hours when $T_a, T_s < 0$). GEP is modeled by adding additional controlling variables to the formula used in the NLR-FC method (Eq. 2) as:

$$GEP = \frac{\alpha Q_d \beta}{\alpha Q_d + \beta} \times f(Ts) \times f(VPD) \times f(VWC_{30})$$

(4.6)

The first term in eq (5) defines a Michaelis-Menten relationship between Q_d and GEP. The second through fourth terms describe sigmoidal-type [0,1] scaling responses of GEP to T_s, atmospheric vapour pressure deficit (VPD) and VWC_{30}, respectively. In contrast to Richardson et al., (2007), a scaling response to air temperature (T_a) was not applied in our model, since parameter analyses indicated that T_s and T_a were strongly correlated, and model explanatory power was not affected by the exclusion of T_a.

Parameters for each gap filling model relationship were estimated using the Nelder-Mead downhill simplex algorithm (function fminsearch in MATLAB), which sought parameter combinations that minimized the Weighted Error Sum of Squares (WESS) cost function, defined by:

$$\Omega = WESS = \sum \left[\left[\frac{(p_i - o_i)^2}{\sigma_i^2} \right]\right] \tag{4.7}$$

, where p_i and o_i are model-predicted and observed half-hourly fluxes, and σ_i is the estimated standard deviation for each value. The WESS cost function was selected for use in optimization, based on investigation of EC measurement errors using the model-residual approach of Richardson et al., (2008). Consistent with findings of Lasslop et al., (2008), analyses showed EC measurement error to be normally distributed and heteroscedastic, with a standard deviation that scaled with the expected magnitudes of the component fluxes (RE and GEP). In such situations, WESS is the maximum likelihood estimator for model parameters. The use the WESS cost function differs from many recent studies, which assume EC measurement error to be best described by a double-exponential distribution (see: Hollinger and Richardson, 2005; Richardson et al., 2006; and, Barr et al., 2013); in which case the Mean Absolute Weighted Error (MAWE) is the maximum likelihood estimator for model parameters.

σ_i was estimated for each half-hour following the model residual approach of Richardson et al., (2008). A reduced parameter version of the NLR-HL model was used to estimate fluxes for all filter-passing half hours, from which observed values were subtracted to generate residuals. As a modification to the original method, model residuals were stratified into 20 quantiles across both Q_d and T_s, and the standard deviation of error values were calculated for each Q_d x T_s strata.. A multilinear regression was run to fit σ values to the mean Q_d and T_s of the strata, and this relationship was used to generate estimates of σ_i across all half-hours. This approach provided considerable improvement to the single-variable scaling approaches previously presented in the literature, as it better related error to the combined magnitude of component fluxes (RE, GEP), rather than the magnitude of measured NEE, which is prone to issues of equifinality. See section 3.3.3 for a complete description of the method.

Investigating this random error estimation method revealed both its output and performance to be highly consistent across most specifications of flux filtering (footprint and friction velocity), though performance declined for cases where large proportions of data were removed (e.g. > 85% of data removed). In these cases, unrealistic error estimates caused the general failure of gap-filling model parameterization. To address these issues, a single time series of estimated σ_i was generated and used across all model runs, in order to standardize gap-filling model performance metrics. This σ_i time series was created from non-footprint-filtered flux data, which was friction velocity threshold filtered according to the moving point threshold method of Papale et al. (2006).

4.4.6　　　　Factorial experiments and analyses

The study objectives were addressed by executing a factorial analysis using TP39 flux and meteorological data collected between 2006 and 2011. Each year of flux data was subjected to all possible footprint-filtering combinations ($fp_m * fp^{Th}$), thus creating 50 footprint-filtered time series per year (5 models * 10 fp^{Th} levels). Each set of footprint-filtered data was subsequently inputted into each of the three u_*^{Th} estimation techniques. As noted in section 4.3.2, multiple metrics (value selection statistics) were used to extract annual estimates from substrata estimates produced by each of the u_*^{Th} estimation methods. The mean and median of these substrata results were taken for each method, and the maximum was also recorded for MPT-P. Each realization of these filtered data sets were then filled using the NLR-HL model. As a result, 150 annual data series of filtered and gap-filled half-hourly NEE, RE and GEP were generated for each year (1800 realizations in total). Combinations are displayed in Table 4.2.

Table 4.2: Data years, filtering and gap-filling method combinations used in factorial analyses.

fp_m (5)	fp^{Th} (10)	u_*^{Th} techniques (3)	Years (6)
KM01, KM01-1D	0 (control)	MPT-P, CPD	2006, 2007, 2008,
SP90, KL04,HS00	0.5:0.05:0.9	MPT-G	2009, 2010, 2011

The amount of data removed or retained by a specific filtering application was described by the available data fraction (ϕ), which was calculated as:

$$\phi = \frac{n_a}{n_p} \tag{4.8}$$

, where n_a is the number of data points available to the gap-filling model for parameterization, and n_p is the amount of data points that would be available for parameterization in a case where no filtering was applied, and no gaps exist. Thus, the value of ϕ spans from 0, where no data points are retained by filtering methods for parameterization, to a maximum of 1, where all possible parameterizable data points are kept and used. Φ may be expressed generally for all NEE, or used more specifically to describe separately the data availability for respiration (ϕ_{RE}) and photosynthesis gap-filling models (ϕ_{GEP}).

Binary time series (gap = 0; no gap = 1) were generated for each factorial combination, to allow

similarity comparisons between model filtering output via the phi statistic, which is defined as:

$$r_\Phi = \frac{(P_{00}P_{11} - P_{01}P_{10})}{\sqrt{P_1 Q_1 P_2 Q_2}}$$

(4.9)

, where values of P in the numerator indicate the proportions of data points where the compared time series have values that correspond to the the listed subscripts (e.g. P_{00} is the proportion of data points where both time series of interest have a filtering-induced gap). In the denominator, P (Q) represents the marginal proportion of non-gaps (gaps) in the time series indicated by the subscript. The value of phi is equivalent to the calculation of Pearson's correlation coefficient (r) on the binary time series. Though the attainable maximum and minimum values of the r_Φ statistic are constrained by the marginal proportions of the binary variables being correlated (Ferguson 1941; Guilford, 1965), these effects are notable only in cases where the marginal proportions are extremely high or low.

In the first stage of investigation, analyses of variance (ANOVAs) were used to quantify the influence of different data treatments to values of u_*^{Th}. Variance in estimated u_*^{Th} values were attributed to factors representing main data treatments, namely: footprint model (fp_m), footprint threshold (fp^{Th}), u_*^{Th} method and year; all first-level interaction terms were also included in the ANOVA model. The measurement year was included as a factor in analyses in order to quantify the influence of inter-annual variability in u_*^{Th} determination. Post-hoc multiple comparison tests (function multcompare in MATLAB, with Bonferroni correction) were applied to ANOVA results to compare population marginal means between factors, and test differences in marginal means for statistical significance ($\alpha = 0.05$). ANOVA residuals demonstrated homoscedasticity and were approximately normal, supporting the use of this tool for the given analysis.

A second stage of investigation was undertaken to quantify and describe the interacting effects of fp^{Th} and u_*^{Th} values on time-integrated carbon exchange estimates (NEE, RE, GEP). For each footprint model, NLR-HL was used to fill flux data that was filtered for incremental combinations of fp^{Th} (nofp, 0.5, 0.55 , ..., 0.85, 0.9) and u_*^{Th} (0.2, 0.22 , ..., 0.68, 0.7).

The third stage of investigation used ANOVAs to quantify the influence and first-level interactions of different filtering factors on annual gap-filled NEE, RE and GEP sums. The measurement year was also included as a factor in analyses in order to quantify the influence of inter-annual variability in u_*^{Th} determination. Post-hoc multiple comparison tests (function multcompare in MATLAB, with Bonferroni correction) were applied to ANOVA results to compare population marginal means between factors (footprint model, fp^{Th} level, u_*^{Th} method, year), and test differences in marginal means for

statistical significance ($\alpha = 0.05$). ANOVA residuals demonstrated homoscedasticity and were approximately normal, supporting the use of this tool for the given analysis.

Preliminary analyses indicated that gap-filling model statistics and gap-filled sums were confounded by interannual variability, as well as a non-linear relationship with ϕ. Therefore, directly comparing these quantities across multiple years and filtering treatments required they be standardized to remove the effects of year and ϕ. This was accomplished by first estimating 'expected values' for statistics and sums, according to each year and ϕ by running gap-filling models for a series of semi-randomly filtered input data sets that spanned a wide range of ϕ values. As a result, comparing model statistics and gap-filled sums across differing filtering treatments required that they be corrected by subtracting an 'expected value' from each of these quantities to remove interactions between year and ϕ. The expected sums and statistical values were generated from year-specific regression curves fit between each of the estimated quantities and ϕ (Figures 4.6 and 4.7, respectively). Each input data set was created by randomly removing the required number of data points from the flux time series, while maintaining the proportion of respiration-parameterizable data (ϕ_{RE}) found in the original footprint- and u_*^{Th} -filtered datasets (Figure 4.8). The procedure was repeated 10 times, and all results were aggregated to constrain final regression curves.

For each filtered data set, the standardized statistics and sums were expressed using both original (unstandardized) values, and in the form of:

$$\delta x = x_{abs} - x_{exp} \tag{4.10}$$

, where x denotes the statistic or sum of interest, and δx expresses the difference between the absolute value of this quantity obtained from the filtered data set (x_{abs}) and the ϕ-predicted 'expected' value of this quantity (x_{exp}). Thus, δx is a measure of the deviation in carbon exchange sums or gap-filling model statistics for filtered datasets as compared to the results expected if filtering had been applied by random decimation of data.
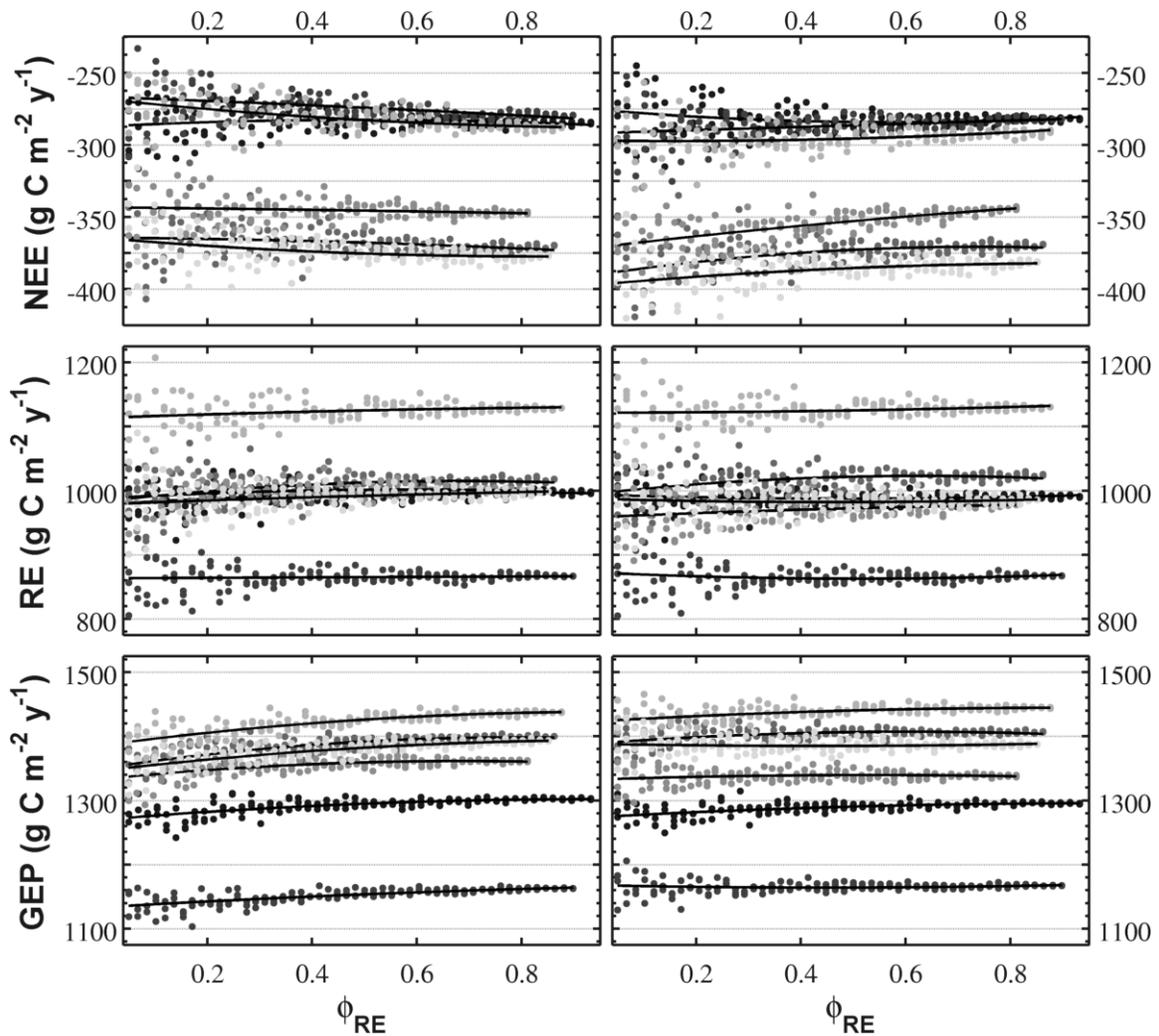
Figure 4.6: *Year-specific expected values for annual carbon exchange estimates as a function of available data fraction for respiration-parameterizable data (ϕ_RE). Results are separated by those for the NLR-FC gap-filling model (left panels) and for the NLR-HL gap-filling model (right panels).*
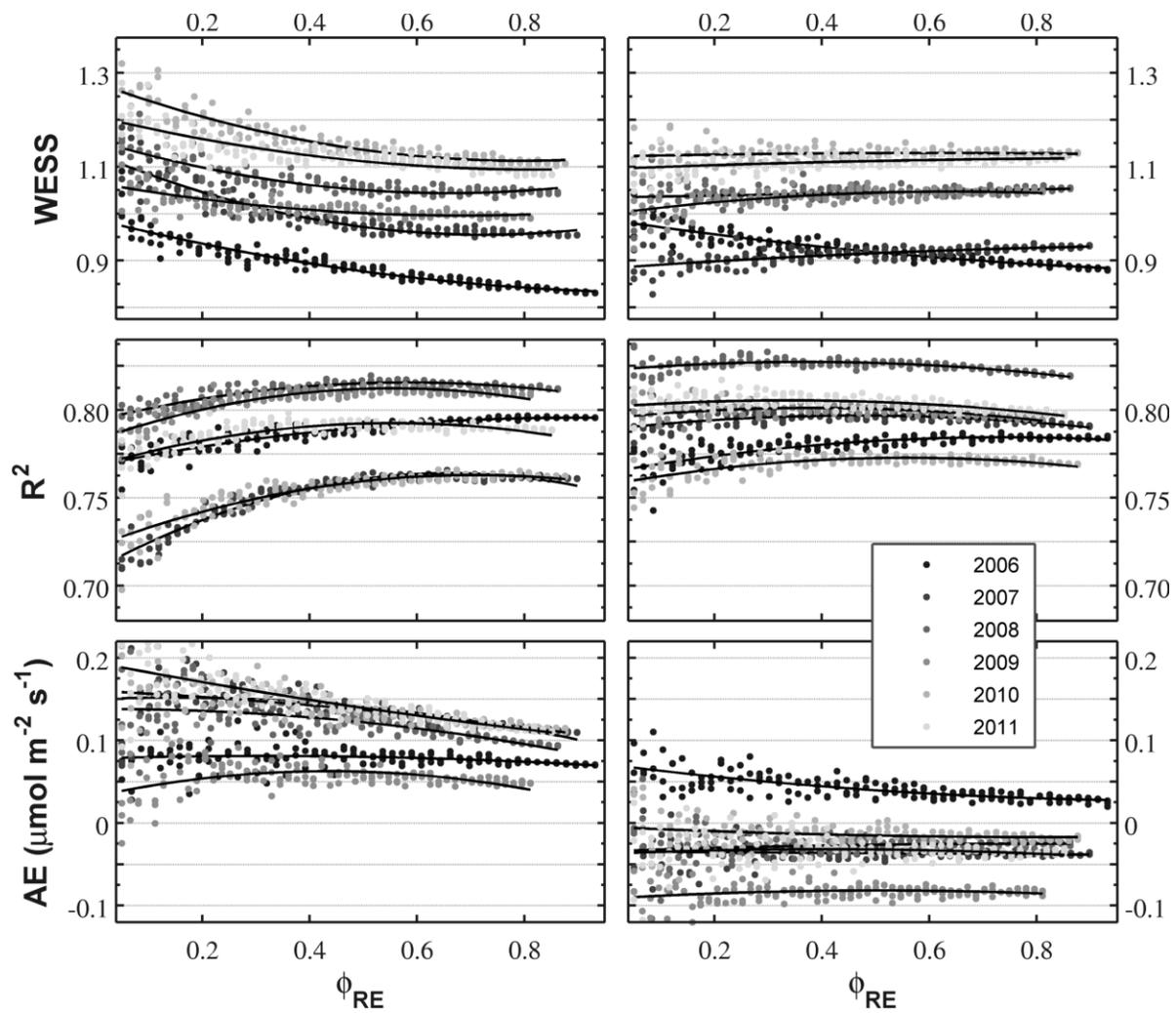
Figure 4.7: *Year-specific expected values for gap-filling model statistical performance metrics as a function of available data fraction for respiration-parameterizable data (φ$_{RE}$). Results are separated by those for the NLR-FC gap-filling model (left panels) and for the NLR-HL gap-filling model (right panels).*
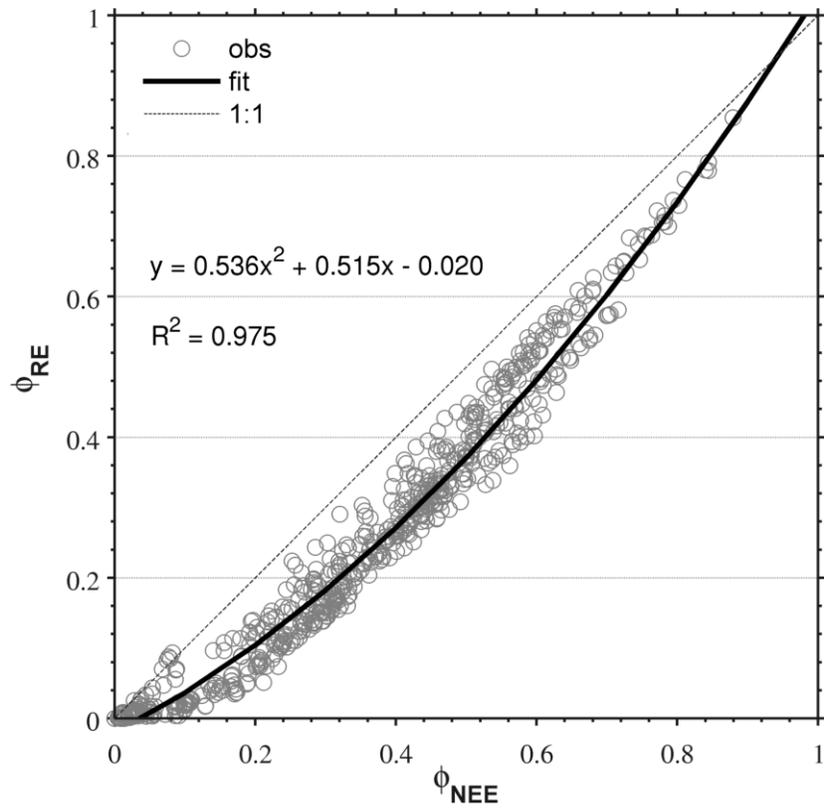
Figure 4.8: *Relationship between available data fraction for RE-parameterizable data (ϕ_{RE}) and all NEE data (ϕ_{NEE}) obtained for all factorial combinations of footprint model, fp^{Th} and u_*^{Th} estimation method over all years. A second-order polynomial curve is fitted to the data.*

To provide an additional metric for footprint model performance, we used the vertical wind speed integral turbulence characteristic (σ_w/u_*) for each half-hour, as described by Göckede et al. (2004). Based on assumptions of flux-variance similarity (Stull, 1988; Wyngaard et al., 1971), values obtained from measurement may be compared with parameterized standard functions to assess data quality; differences between measured and parameterized values indicate not fully formed turbulence or disturbances in turbulent flow fields (Foken and Leclerc, 2004; Göckede et al., 2004; Thomas and Foken, 2002). Parameterized functions are calculated according to Thomas & Foken, (2002) as:

$$
\begin{aligned}
\frac{\sigma_w}{u_*} &= 1.3(1-2\zeta)^{(1/3)} \quad, \qquad -3 < \zeta < -0.2 \\
\frac{\sigma_w}{u_*} &= 0.21\ln\left(\frac{z_+ f}{u_*}\right) + 3.1 \quad, \qquad -0.2 < \zeta < 0.4
\end{aligned}
\tag{4.11}
$$

, where σ_w is the standard deviation of vertical wind velocity, ζ is the stability parameter (calculated as *(z-d) /L*, where z is measurement height, d is displacement height, and L is the Obukhov length), f is the Coriolis parameter, and z_+ is a normalizing factor equal to 1 m.

All factorial experiments and analyses were carried out using MATLAB software. Simulations were compiled as standalone C language programs and executed using the Shared Hierarchical Academic Research Computing Network (SHARCNET).

4.5 Results and discussion

4.5.1   Footprint model filtering comparisons

As a first means of footprint model evaluation, a comparison was made of the quantity and distribution of gaps created by footprint application for each tested value of fp^{Th}. Post-filtering ϕ differed most between models at the least stringent footprint threshold value tested ($fp^{Th} = 0.5$; Figure 4.9). The crosswind-integrated (SP90, KL04 and HS00) footprint models retained between 65 and 90% of the original data, while both versions of the KM01 model (1D and 2D), retained less than 40% of data at the least stringent fp^{Th} setting. Overall, this discrepancy lessened as fp^{Th} increased (though unevenly and at different rates). At the highest footprint threshold ($fp^{Th} = 0.9$), all models retained less than 15% of data; ϕ values declined rapidly above $fp^{Th}=0.7$ for SP90 and above $fp^{Th}=0.8$ for KL04 and HS00. The KM01 model removed the greatest proportion of data across the entire tested fp^{Th} range, while HS00 removed the least for all but the most-stringent fp^{Th} values. ϕ values for KL04 displayed a negative bias compared to other model output (especially at lower fp^{Th} values), as a result of the application of an automatic rejection criteria to all half-hours with $u_* < 0.2$ m s^{-1}. This feature, in combination with relative less sensitivity to fp^{Th}, caused ϕ for the KL04 model to decrease more gradually with fp^{Th} than was observed with the SP90 and HS00 models.

Model filtering outputs showed similar trends when examined for differences between day and nighttime ϕ (Figure 4.9c). All models removed a larger proportion of nighttime data than daytime data; however, the difference between these two values were relatively smaller for the crosswind-integrated footprint models (day/night ϕ difference of 0.05 to 0.1), than for the KM01 model (ϕ difference up to 0.2). When examined over the course of a year for a given fp^{Th} value (Figure 4.9b), KL04 and HS00 exhibited very similar intra-annual trends, with relative ϕ peaks in the winter and mid-summer months. In comparison, ϕ of KM01 and SP90-filtered data reached a single maximum in mid-summer, with the seasonal difference much more prevalent for SP90.
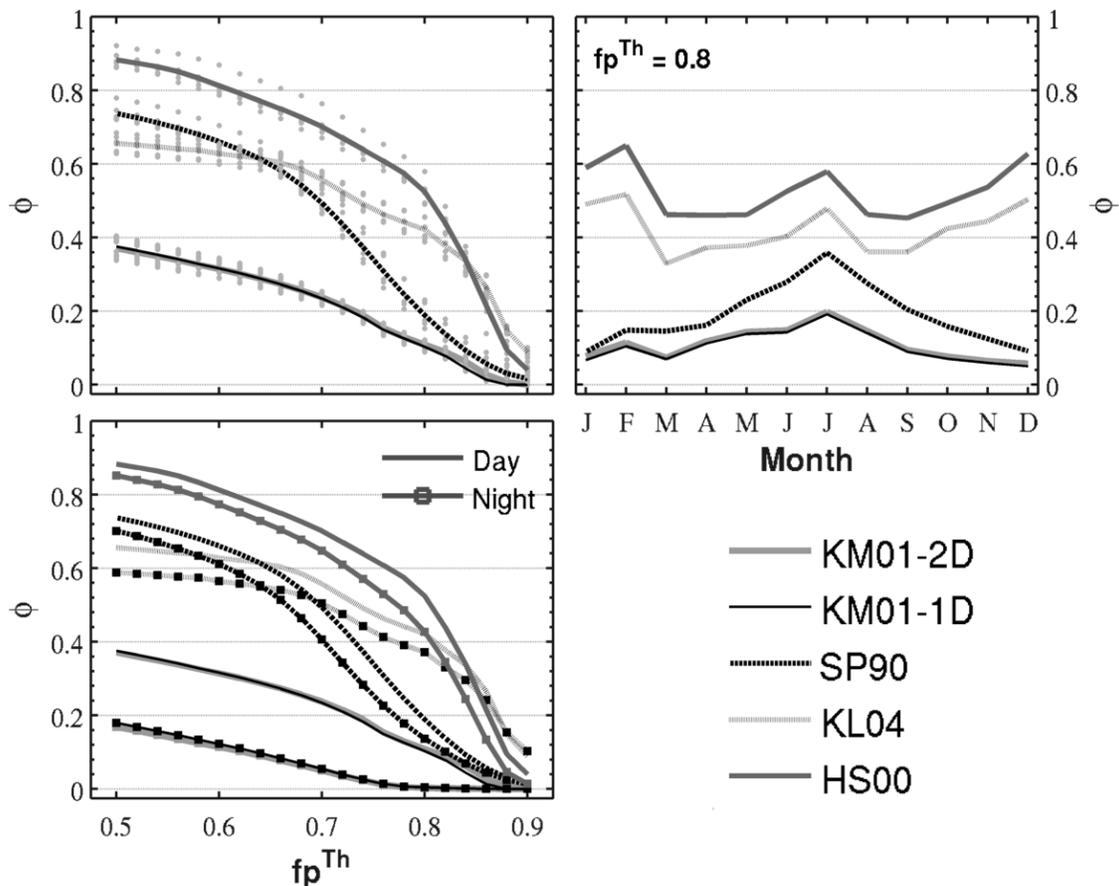
Figure 4.9: *Available NEE data fraction (ϕ) resulting from the application of each footprint model (fp_m). Panel a) shows ϕ as a function of the assigned footprint threshold (fp^{Th}) for individual years between 2006-2011 (dots) and the all-years mean (line). Panel c) displays the daytime (line, no markers) and nighttime (line with boxes) difference in the all-years average value for each fp_m and fp^{Th}, while panel b) shows monthly variability in ϕ resulting from application of each fp_m at fp^{Th}=0.8.*

The significantly greater data removal by KM01 is consistent with results obtained by previous investigations (Kljun et al., 2003; van de Boer et al., 2013), which suggested that the KM01 model overestimates footprint length across all conditions. Sensitivity tests conducted on this model (not shown), revealed a strong sensitivity to friction velocity – periods with low friction velocities (< 0.2 m s⁻¹) are particularly prone to footprints that extend greatly upwind, with considerable portions of the total flux signal emanating from far beyond the bounds of the forest. Even during periods where u* was relatively high (0.35-0.5 m s⁻¹), footprints still commonly stretched well past the forest bounds. Tests also showed this model to be sensitive to the height of measurement above the surface, as the model became less selective for simulations where the sensor was placed a few metres from the ground.

Footprint model outputs were also investigated for their sensitivity to changes in atmospheric stratification (Figure 4.10). Results showed that models performed similarly for unstable and neutral conditions ($\zeta < 0$), as median F_{in} values during these conditions ranged between a low of 0.72 for KM01 models to a high of 0.84 for HS00. The amount of unstable- and neutral-period data retained by application of these models across fp^{Th} values generally reflected trends in F_{in}, as ϕ was highest for HS00 and lowest for both KM01 models. Due to its automatic removal of data for periods where $u_* <$ 0.2 m s^{-1}, ϕ for KL04 was negatively biased, especially at low fp^{Th} values. Models showed a much more varied response during stable ($\zeta > 0$) conditions. The KM01 models were highly sensitive to stable conditions, and rejected data for most or all half-hours across fp^{Th}. HS00 and KL04 models, conversely, were relatively little affected by changes to stable stratification, and produced ϕ values similar to that observed for neutral and unstable conditions. SP90 showed a moderate response to changes in stratification. The strong sensitivity of the KM01 model to stratification found in our study contradicts the findings by van de Boer et al. (2013), who found the KM01 model to be relatively insensitive to such changes. A potential explanation for this disagreement arises from differences in measurement heights between these two studies (between z = 2.5 and 6 m in van de Boer et al., 2013; z = 28 m in our study). An investigation of this model by Kljun et al., (2003) found the magnitude of stable/unstable discrepancies increased with height from the surface.

We furthered the temporal characterization of model filtering differences by investigating half hour-specific ensemble ϕ averages (Figures 4.11 and 4.12).The day-night distinction was strongest for KM01 results, as most nighttime data was removed, and nighttime ϕ values differed distinctly from the daytime. The tendency for ϕ to be higher during all daytime hours explained the mid-summer peak observed in Figure 4.9b, as a longer daytime resulted in relatively more filter-passing half-hours. Comparing Figure 4.11 to Figure 4.12, increasing fp^{Th} from 0.6 to 0.8 had the largest overall effect on SP90 filtering, where ϕ was substantially reduced. Also apparent in SP90 results was a consistent tendency to remove data during day-night transitional periods. This effect was attributed to the model's use of a momentum stability correction (Eq. 1), which approaches infinity asymptotically as the Obukhov length (*L*) approaches *16(z-d)* during the transition from unstable to stable conditions and vice versa. This characteristic has important negative implications for the use of methods that require measurements during this transitional period to estimate nocturnal and daytime respiration rates (Griffis, 2003; van Gorsel et al., 2008). An increase in fp^{Th} also had a substantial effect on the daytime-nighttime ϕ ratio of the HS00 model output, as a considerably higher proportion of nighttime

data was removed at $fp^{Th} = 0.8$ than for $fp^{Th} = 0.6$.
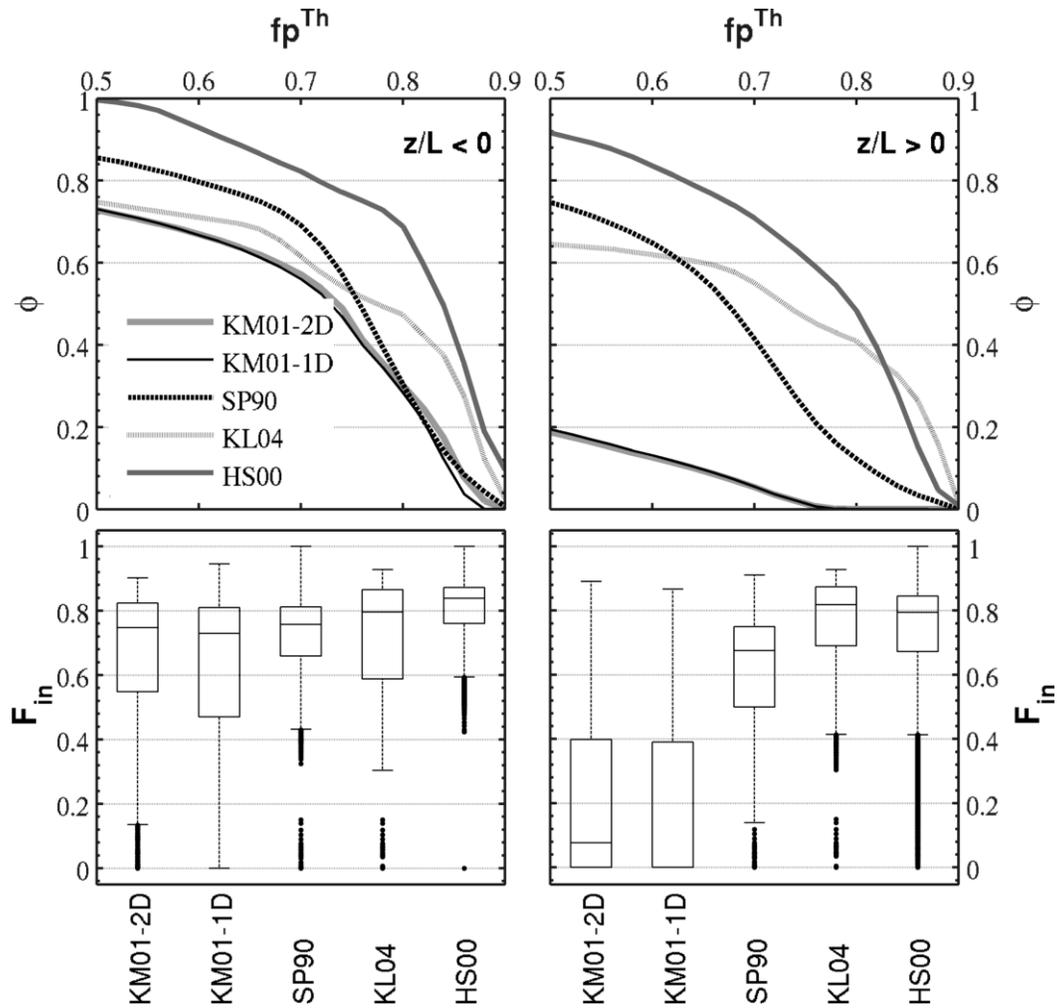
Figure 4.10: *Footprint model results presented for unstable (z/L < 0; left panels) and stable (z/L > 0; right panels) periods of atmospheric stratification. Top panels show available data fraction (ϕ) resulting from application of each model across all fp^{Th}. Bottom panels use box plots to show the general distributions of within-fetch flux footprint proportion (F_{in}) for each model in unstable (left) and stable (right) conditions.*
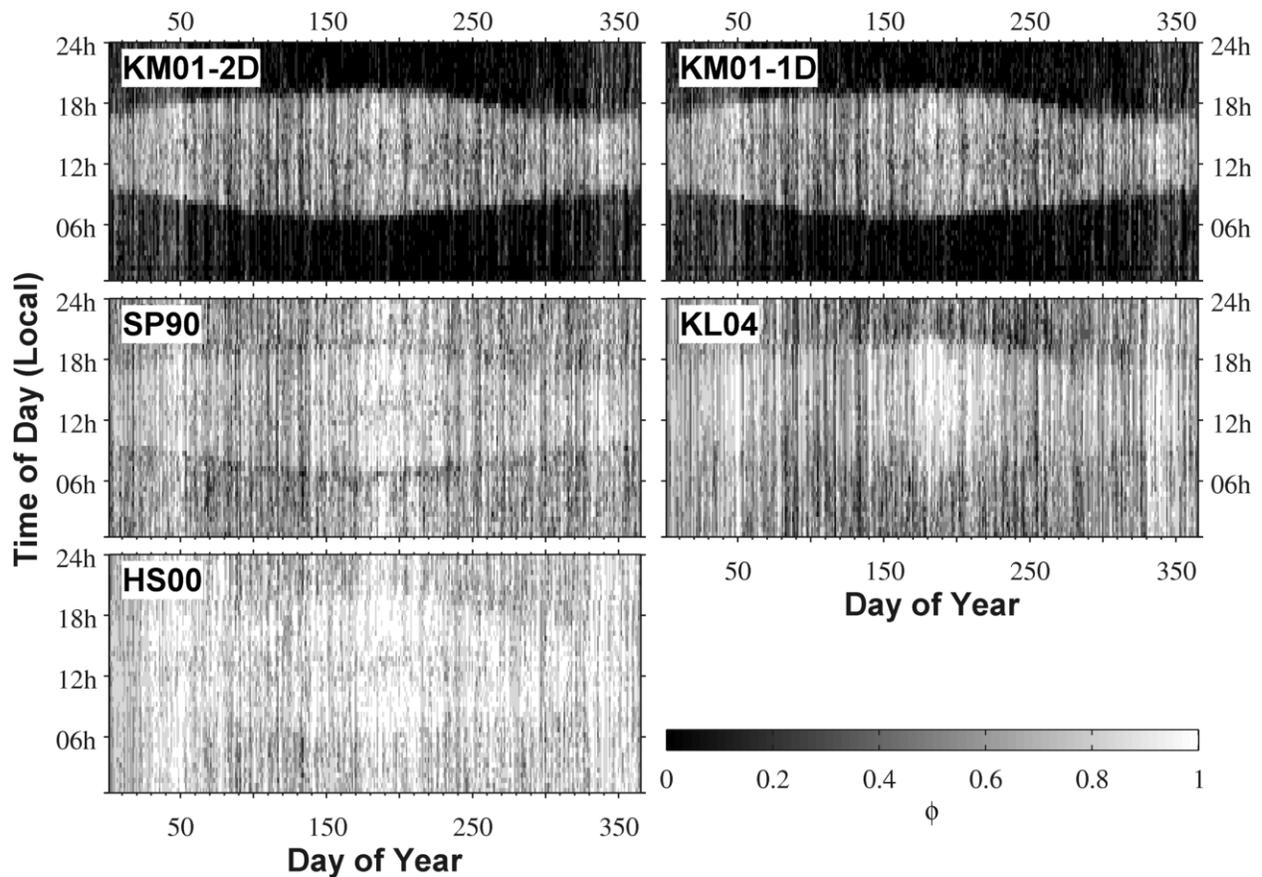
Figure 4.11: *Half-hourly structure of data retained following footprint filter application for each method at $fp^{Th}=0.6$. Each pixel represents the proportion of specific half-hours that are retained for years 2006-2011; darker colours represent an increased occurrence of footprint-filtering gaps for the given half-hour. Pixel columns span the course of each day (bottom to top), over each day of year.*
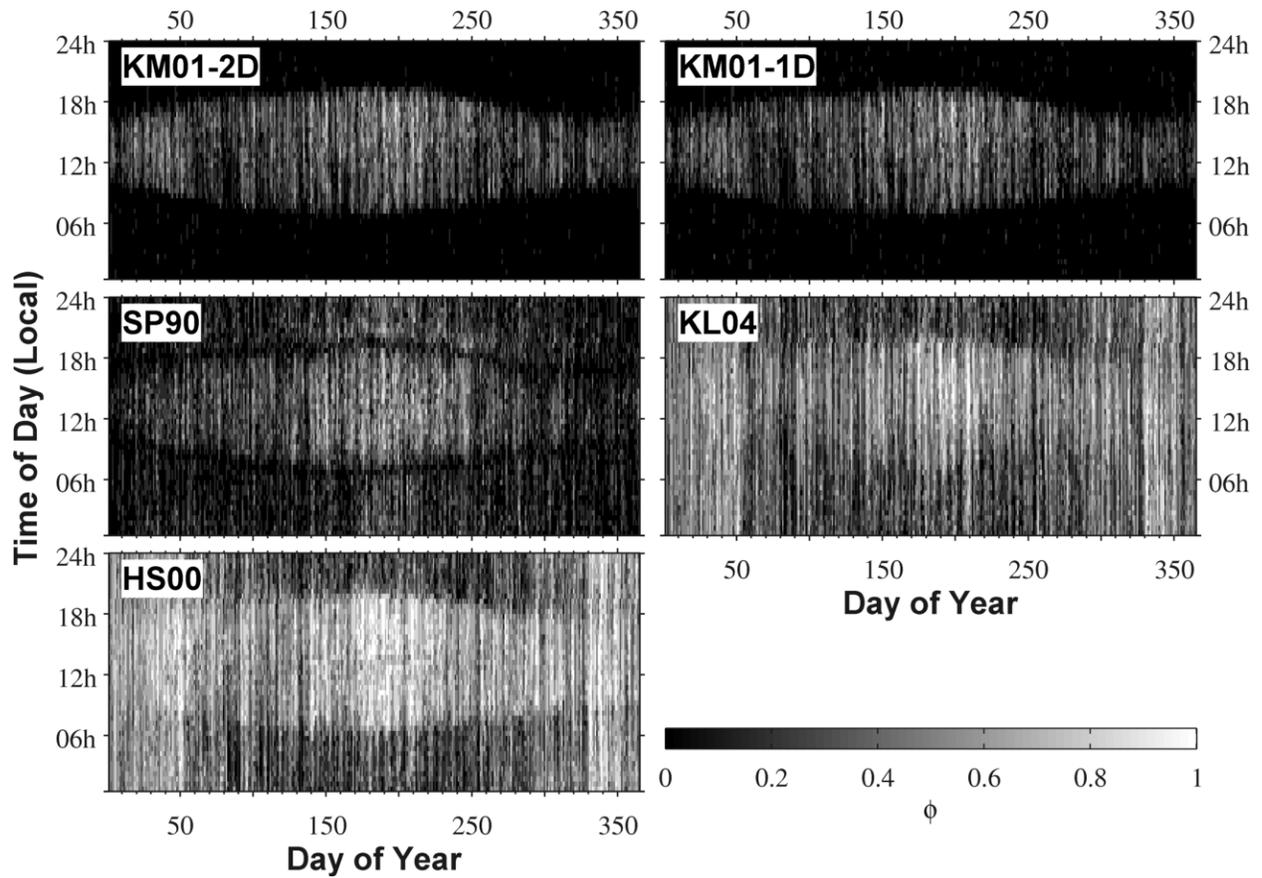
Figure 4.12: *Half-hourly structure of data retained following footprint filtering, as described in Figure 4.11, but for fp^Th=0.8.*

The binary outputs created for each footprint filtering method output were used with Eq. 2 to quantify the degree of similarity between footprint filtering methods (Figure 4.13). Values of r_Φ typically ranged between 0.2 and 0.6, and correlation between models was generally greatest at low-to-mid fp^{Th} values. Correlation scores decreased rapidly at the highest fp^{Th} values, due to the constraints imposed on maximum values of r_Φ when binary datasets consist mostly of zeros (Davenport and El-Sanhurry, 1991; Guilford, 1965). Among notable results was the anticipated strong correlation between both KM01 models (1D and 2D), and a moderate to strong correlation between KL04 and HS00 model outputs, which tended to increase with increasing fp^{Th} values. When the analysis was expanded to consider the lagged correlation between model outputs at differing fp^{Th} values (cross-correlation), the relationship between the KL04 and HS00 models further improved, as footprint-filtered data from the HS00 model was most similar to KL04 output obtained with fp^{Th} lagged by 0.1 to 0.15 (Figure 4.14). Considered along with similarities in intra-annual ϕ trends (Figure 4.9) and responses to atmospheric stratification conditions (Figure 4.10), this result provides more evidence that both footprint models remove much of the same data; though the KL04 model is more stringent for a given fp^{Th} value, likely due to its pre-filtering of half-hours with $u_* < 0.2$ m s^{-1}. The similarity in estimates between KL04 and HS00 reflects the fact that both models were developed as parameterizations of more complex Lagrangian simulations (Hsieh et al., 2000; Kljun et al., 2004; van de Boer et al., 2013).
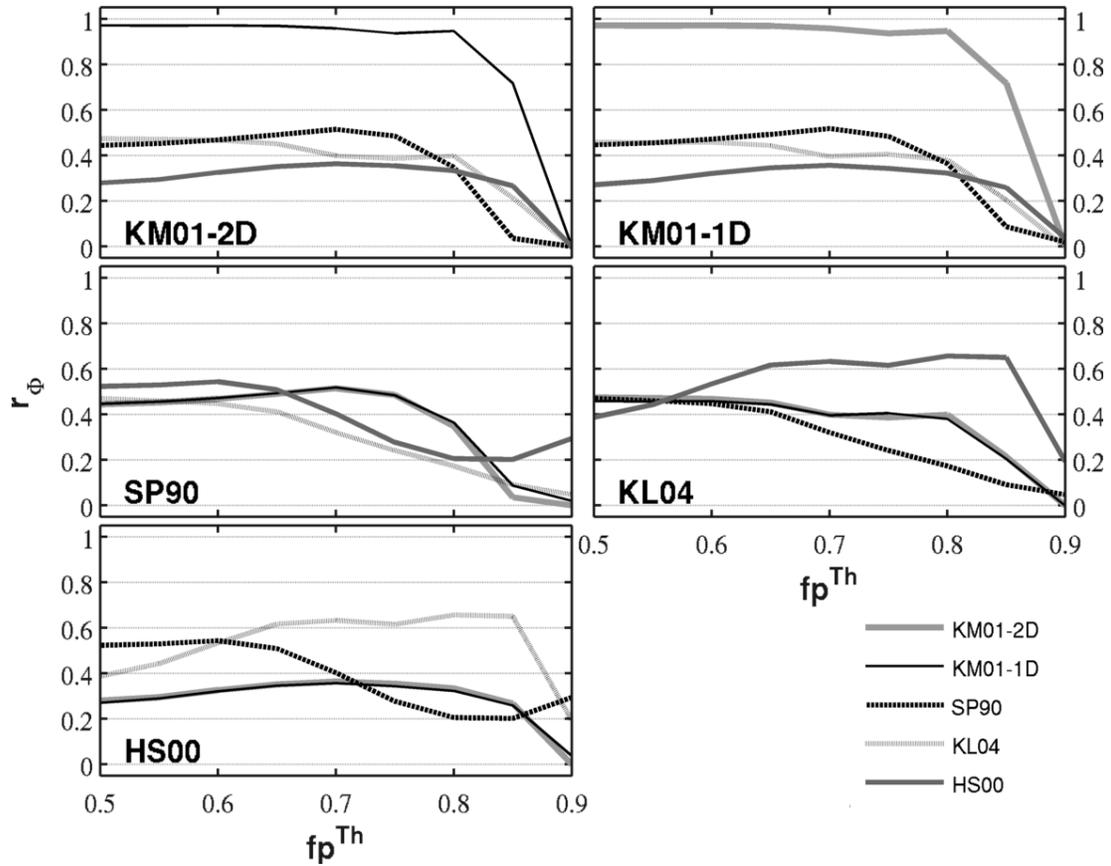
Figure 4.13: *Value of Pearson's correlation coefficient, r_Φ , calculated between the binary gap/no-gap outputs from each footprint model (indicated in the bottom-left of each panel) and all others at equal fp^{Th} values.*
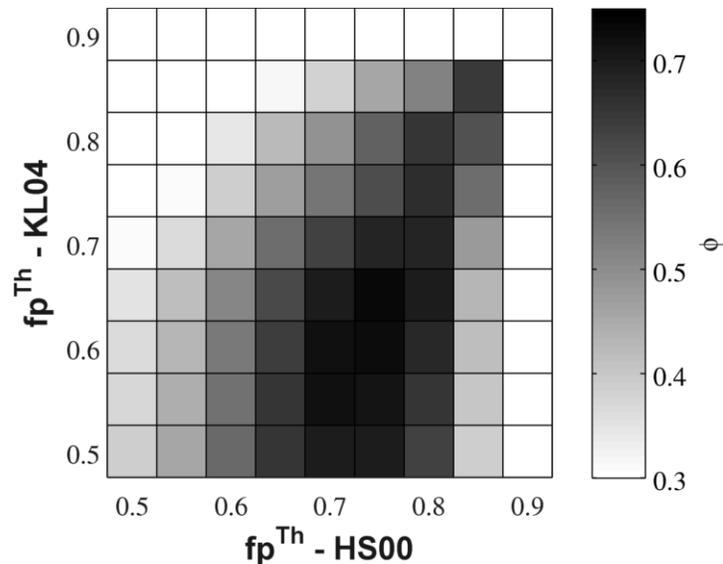
Figure 4.14: *Cross-correlation values between HS00 and KL04 model output for all levels of fp^{Th}*

4.5.2   Comparing friction velocity threshold estimation methods

To investigate differences in u_*^{Th} resulting from the selection of estimation methods and value selection statistic, factorial analyses results were aggregated across all runs, and grouped according to their u_*^{Th} method application (Figure 4.15). The classification of value selection statistic (mean, median, max) refers to the statistical metric used within a given estimation method to produce a single, annual u_*^{Th} estimate from numerous estimates produced during data stratification (see section 4.3.2).
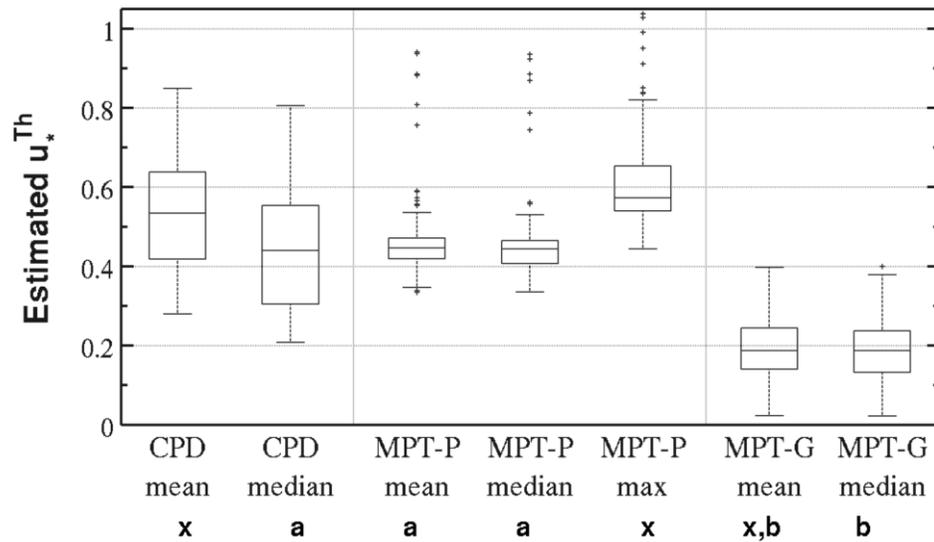
Figure 4.15: *Box plot representation of u_*^{Th} distributions calculated for all factorial combination runs, organized by u_*^{Th} determination method (CPD, MPT-P, MPT-G) and value selection method (mean, median, max).*
x indicates the standard value selection statistic for each determination method.
a,b indicate groups with statistically similar means ($\alpha=0.05$), as determined using a post-hoc comparison test of ANOVA results.

In terms of u_*^{Th} estimate variability, the CPD method showed the greatest spread in predicted values across factorial runs, indicating a greater sensitivity to input data treatments (year, footprint model, fp^{Th}). The interquartile ranges (IQR) of CPD estimates were greater than 0.22 m s^{-1}, in comparison to values between 0.05 to 0.11 m s^{-1} for the MPT-P method and 0.1 m s^{-1} for the MPT-G method. CPD estimate dispersion was slightly greater (IQR = 0.25 m s^{-1}) when the median was used instead of the mean (IQR = 0.22 m s^{-1}) as the value selection statistic. For MPT-P, estimate dispersion was comparable across mean and median (IQRs of 0.053 and 0.058 m s^{-1}, respectively), but increased considerably (IQR of 0.11 m s^{-1}) when the maximum value selection statistic was used.

u_*^{Th} estimates produced using the MPT-G method were significantly and substantially lower than those estimated by the CPD and MPT-P methods. Regardless of value selection statistic, the application of the MPT-G resulted in a u_*^{Th} estimate value of 0.19 m s^{-1}, which was between 0.24 and 0.42 m s^{-1} lower than estimates from all other methods (Table 4.4). The predominance of outliers in the MPT-P results is a consequence of a varied response between methods to situations where there was insufficient input data to provide an appropriate estimate of u_*^{Th}. This situation generally occurred when footprint filtering applications removed a very large proportion (greater than 90-95%) of nighttime data points prior to u_*^{Th} estimation. When operated in such situations, the CPD method experienced an outright algorithm failure (returning a NaN value), providing immediate feedback on the suitability of input data. MPT-P and MPT-G estimates were realized, though MPT-P estimates were strongly positively skewed, creating positive outliers , while MPT-G results were variable and unpredictable. The varying degrees of feedback that each method provides following algorithm failure is an important consideration when evaluating the suitability of any of these methods, given that latent erroneous results can introduce both random and systematic errors into u_*^{Th} estimates.

In general, the majority of CPD and MPT-P results suggested the use of a u_*^{Th} value between 0.4 and 0.6 m s^{-1}. Within and between these results, the choice of value-selection statistic had a statistically significant impact on estimated u_*^{Th} value (Table 4.4). Though the selection between median or mean had little effect on MPT-P and MPT-G results, the CPD method was considerably more sensitive to this choice, as mean-derived u_*^{Th} estimates were an average of 0.09 m s^{-1} higher than median-derived values (Table 4.4). Such a difference between mean and median values suggests that the collection of stratified u_*^{Th} estimates – which the CPD method aggregates to derive annual estimates – may not be normally distributed, or are contaminated by extreme outlier values. Further investigation of this phenomenon revealed a tendency for the CPD method to create a bimodal distribution of u_*^{Th} estimates for individual

nS*nT strata (Figure 4.16a). This bimodal distribution was caused by a consistently high u_*^{Th} estimate for the lowest of the four temperature strata used. The non-normality of CPD strata estimates led to the discrepancy between calculated mean and median values (Figure 4.16a). In the case of our study, the mean cannot be expected to provide an accurate estimation of the central tendency of u_*^{Th} for CPD estimates, and thus, should not be used. As the mean is the standard statistic used by the CPD method to estimate u_*^{Th}, this result suggests that other studies using the standard CPD method closely examine strata estimates to ensure normality, or otherwise switch to a non-parametric estimate of central tendency (i.e. median).

Use of the median value selection statistic for CPD, however, results in u_*^{Th} estimates that are, on average, comparable to those produced using the MPT-P method with either of the mean or median value selection statistics (differences of 0.028 and 0.016 m s^{-1}, respectively; Table 2). In comparison, estimates for MPT-P strata passed normality tests, and u_*^{Th} estimates were similar for both mean and median value selection statistics (Figure 4.16b). Paired t-tests between treatments (method * value selection statistic) indicated that the means of these three distributions (CPD-median, MPT-P-mean, MPT-P-median) were one of only two sets of estimates that were not statistically different from each other. Despite similarity of means, however, the variance of estimates was considerably larger for CPD than for MPT-P.

For MPT-P, the use of the maximum value selection statistic, which is the standard value for this method, produced the highest mean estimate of u_*^{Th} (0.58 m s^{-1}), and increased u_*^{Th} estimates by an average of 0.148 and 0.160 m s^{-1} relative to the use of mean and median, respectively. The maximum value selection statistic was originally used with MPT-P by Papale et al. (2006), who applied it to a multi-site synthesis analysis. In such a multi-site application, the use of the maximum value selection statistic provides extremely conservative (high) estimates of u_*^{Th} values in order to ensure consistency across sites, and increase confidence in the robustness of analysis results. However, results from this study suggest that the median is a more appropriate estimator for u_*^{Th}, as the use of the maximum value leads to estimates that are not representative of the majority. The use of such a high value may also have unintended and considerable effects on flux estimates and relationships derived using u_*^{Th}-filtered data. In addition, the results of these analyses indicate that the use of a consistent value selection statistic is necessary to permit a meaningful comparison of u_*^{Th} estimation methods, since considerable variation between methods may be explained by different approaches to selecting a representative value from a sample of estimates. Given the limitations of the CPD-mean approach and the desire for the use

of a common value selection statistic across u_*^{Th} estimation methods, this paper used the median value selection statistic for all following analyses of factorial runs and u_*^{Th} estimation methods.

Table 4.3: *Mean differences in predicted u_*^{Th} between estimation models and value-selection methods across factorial run combinations. Results from the KM01 footprint model have been excluded from calculations, due to destabilizing effects on u_*^{Th} estimation methods. Superscripts indicate groups with statistically similar means ($\alpha=0.05$), as determined using a post-hoc comparison test of ANOVA results.*

	CPD mean	CPD median	MPT-P mean	MPT-P median	MPT-P max	MPT-G mean	MPT-G median
CPD mean		0.090	0.063	0.075	-0.086	0.332	0.337
CPD median	-0.090		-0.028[a]	-0.016[a]	-0.176	0.242	0.247
MPT-P mean	-0.063	0.028[a]		0.012[a]	-0.148	0.270	0.275
MPT-P median	-0.075	0.016[a]	-0.012[a]		-0.160	0.258	0.263
MPT-P max	0.086	0.176	0.148	0.160		0.418	0.423
MPT-G mean	-0.332	-0.242	-0.270	-0.258	-0.418		0.005[b]
MPT-G median	-0.337	-0.247	-0.275	-0.263	-0.423	-0.005[b]	
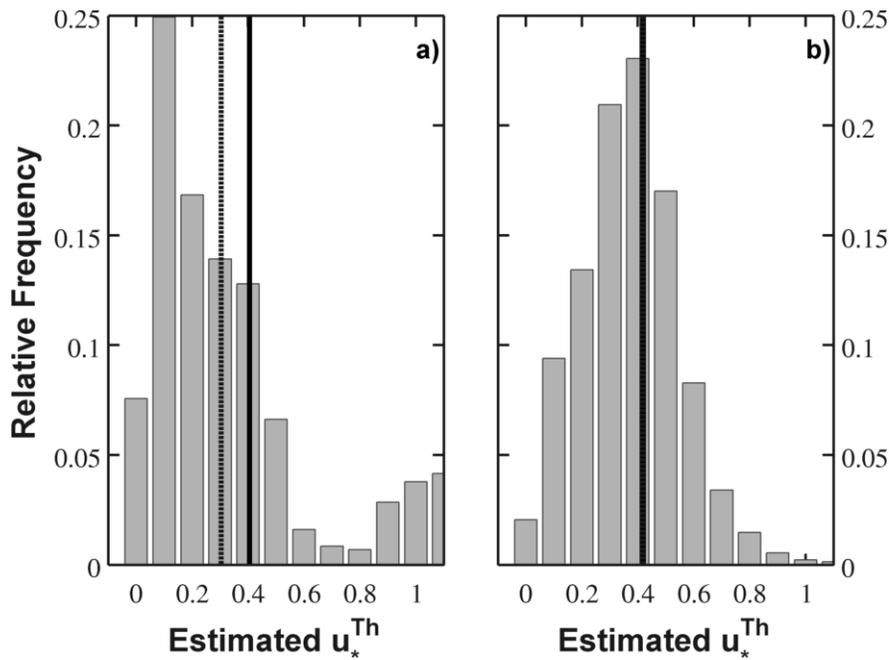
Figure 4.16: *Histograms of Monte Carlo intermediary u_*^{Th} estimates for all n_s*n_T strata for 2006 (no footprint applied), obtained using the a) CPD, and the b) MPT-P u_*^{Th} estimation methods. The solid black vertical line indicates the median value of all estimates, while the mean of estimates is represented by the vertical dashed line.*

Our results showed large and varied differences from analyses performed by Barr et al. (2013), who implemented and compared the CPD and MPT-P methods over a large collection of eddy covariance measurement sites. These analyses included our site, though their study used data measured between 2003 and 2007, while ours used 2006 to 2011 data. In their study, TP39 was identified as one of a handful of irregular sites, which commonly demonstrated an "excess" situation during CPD estimation (NEE magnitudes decreased as u_* increased up to u_*^{Th}). CPD estimates were comparable to those generated in our study, but had a high degree of associated uncertainty (CPD-estimated u_*^{Th} was 0.5 +/- 0.38 ms-1). Applying the MPT-P-mean method, however, their estimates were substantially lower than ours (0.2 m s^{-1} in their study, compared to 0.4 - 0.5 ms-1 in ours).

The causes of these discrepancies are likely numerous: investigations by Barr et al. (2013) were carried out on older data that was known to have some integrity and structural issues. This data has since been updated and improved considerably; thus it would be advisable that these analyses be repeated using the new (and expanded) data sets. Indeed, our analyses indicated that CPD operated in a manner more consistent with most other sites in Barr et al. (2013), as "deficit" situations were almost exclusively experienced. Another likely cause of these noted differences is the use of the mean value selection statistic with CPD by Barr et al. (2013) instead of the median. As shown in our study, the use of CPD-mean is invalid at our site, as assumptions of normality and unimodality are violated.

In this study, we used a single, static u_*^{Th} value for the entire year. Results of Barr et al. (2013) suggested that our site experiences a considerable seasonal variation in estimated u_*^{Th} – results that were confirmed by our own analyses. Though all the u_*^{Th} methods we implemented in this study were capable of producing seasonal estimates, we used a single annual value in the interest of simplifying the tasks of classifying effects on u_*^{Th} estimates and to reduce issues related to missing/poorly estimated seasonal values. Further work should expand this comparison to consider the effect of seasonal estimates.

4.5.3   Controls on u_*^{Th} estimates: Footprint filtering and inter-annual variability

As a first investigation of u_*^{Th} estimates controls,we aggregated and compared factorial runs results according to the following data treatment groups: footprint model, u*Th estimation method, input data year and fpTh level (Figure 4.17). Factorial analyses indicated problems estimating u*Th values using

KM01-filtered data, since the highly-restrictive model commonly removed a greater amount of data than tolerable by the u*Th estimation methods. As a result, many estimates associated with this method were NaN values, or were otherwise highly variable outliers. Due to the broad failure of this model to produce unusable data sets, all results associated with application of the KM01 model were discarded for the remainder of analyses, as it was concluded that this model cannot be applied practically to data at our site.

All groups – with the exception of MPT-G estimates – showed a general increase in estimated u_*^{Th} with fp^{Th}; this relationship was especially pronounced for high fp^{Th} values ($fp^{Th} > 0.7$). Additionally, dispersion between and within groups increased with fp^{Th}, showing a correlation between u_*^{Th} estimate variability and the amount of input data available to estimation methods.

Among footprint model groups (Figure 4.17, panel a), application of KL04 generally resulted in higher u_*^{Th} estimates for given fp^{Th} values up to a level of $fp^{Th} = 0.85$. The SP90-filtered data showed the greatest sensitivity to changes in fp^{Th} levels, as group u_*^{Th} means ranged between 0.35 for the no-footprint (control) case, to 0.73 m s^{-1} for the highest fp^{Th} level. HS00 results were least sensitive to changes in fp^{Th}, with a u_*^{Th} mean range of 0.35 to 0.58 m s^{-1}. Within-group variability of u_*^{Th} estimates was greatest for the SP90-filtered data, while lowest for data filtered with the KL04 model. Year-to-year variation was observed in u_*^{Th} estimates, but the apparent discrepancies between years was much smaller than those observed among levels of other groups (Figure 4.17, panels c, d). The relative difference between u_*^{Th} for different years was generally consistent over all fp^{Th} values, and the interaction effect of these two factors was not significant.
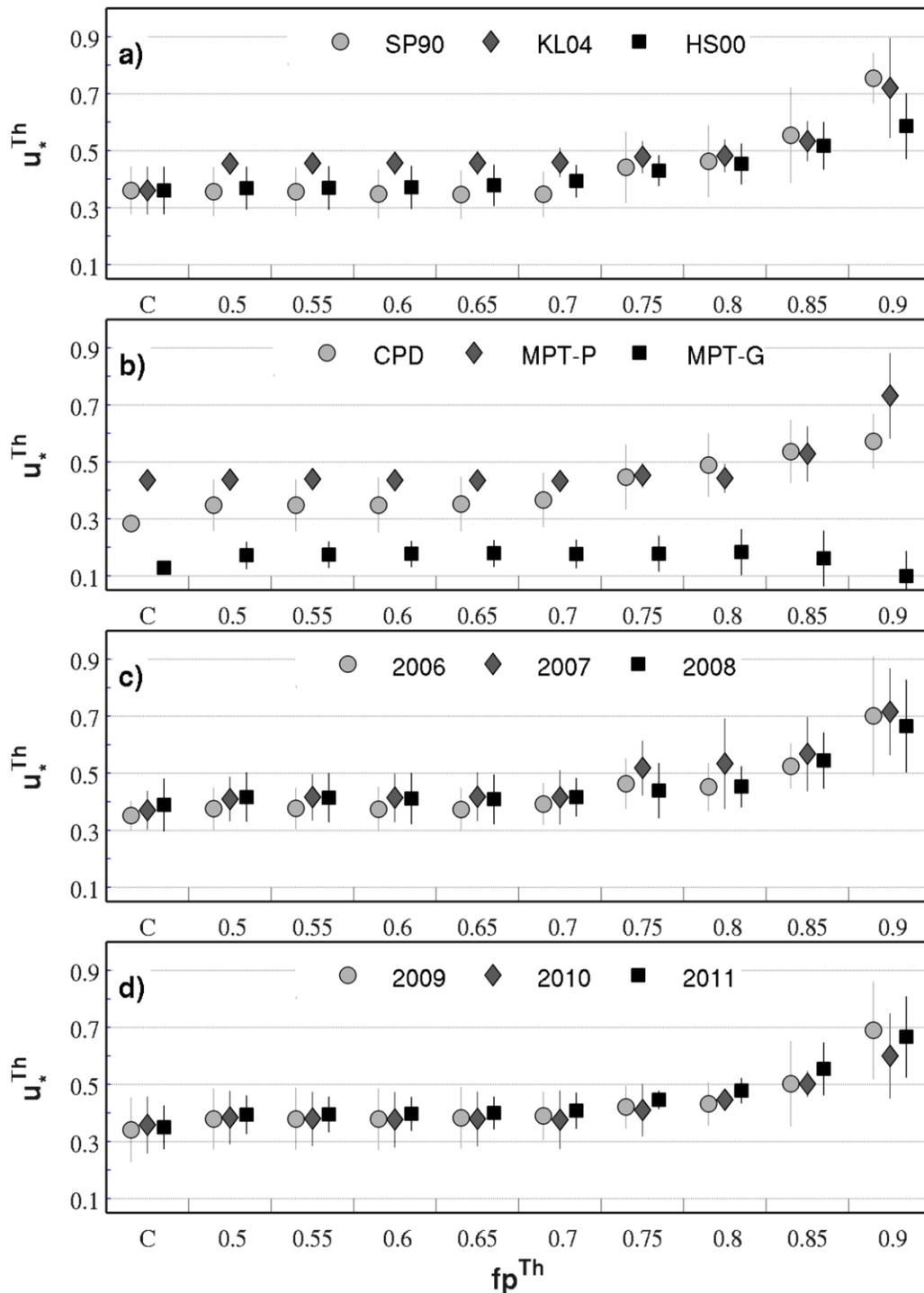
Figure 4.17: *Group mean u_*^{Th} estimates (symbols) and standard deviation (vertical bars) across dimensions of footprint model (panel a), u_*^{Th} estimation method (panel b), and year (panels c and d). Results obtained through application of the MPT-G method are only included for demonstrative purposes in panel b) – they are removed from all other panels in order to reduce unwarranted group biases and error.*

When grouping and comparing between u_*^{Th} estimation methods (Figure 4.17, panel b), MPT-P estimates were typically higher than those produced by the CPD method, though predictions of the two tended to converge for fp^{Th} between 0.75 and 0.85. The MPT-G method, in contrast, consistently produced much lower u_*^{Th} values than the other methods, while also showing little sensitivity to changes in fp^{Th}.

Consistently, the MPT-G method results differed from those of the other estimation methods, both in terms of the magnitudes of u_*^{Th} estimates produced, and the response of these estimates to differing methodological and input data treatments. In-depth investigation of MPT-G method operation and performance indicated that estimates produced by this method were closely related to the minimum values of u_* included in the input data. Since the algorithm showed a tendency to exit (returning a u_*^{Th} estimate) after a short number of outer loop repetitions (between two and five), the final estimate from this method was much more correlated to the median u_*^{Th} value of the initial inner-loop evaluation window, than to specific data treatments. Considering these questionable features, results associated with MPT-G application were discarded for other panels in Figure 4.17, and for all subsequent ANOVAs, since their inclusion confounded results, and diminished the importance of footprint filter-related treatments on u_*^{Th} values. Thus, the remainder of analyses consider only the CPD and MPT-P u_*^{Th} estimation methods.

An ANOVA was conducted on factorial results, in order to identify significant controlling factors on u_*^{Th} estimation, and diagnose interaction effects that explain inconsistencies among and between group responses in Figure 4.17. Results of this analysis are shown in Table 4.4 (ANOVA table), and results from post-hoc multiple comparison tests are displayed in Figure 4.18. ANOVA tests indicated that all four main effects (footprint method, fp^{Th}, u_*^{Th} method and year) had statistically significant impacts on the estimated value of u_*^{Th}, while four two-factor interactions were also significant. The largest mean effect (as denoted in Table 4.4 by Effect Mean Square) was contributed by the choice of u_*^{Th} method, followed in importance by the choice of footprint model. Though fp^{Th} level explained the largest amount of variability in u_*^{Th} of any factor tested (24 % of total variability explained), its mean effect on u_*^{Th} was spread among many fp^{Th} levels. Investigating the marginal means predicted for each fp^{Th} level (Figure 4.18, panel c), showed that estimates for fp^{Th} levels between 0.5 and 0.7 are not statistically different, while the estimated marginal mean u_*^{Th} values for the control (no footprint) case and $fp^{Th} > 0.7$ differ significantly and considerably from these middle values. Considered along with relatively small mean effects for fp^{Th}-associated interaction terms, these results suggest that the effect of fp^{Th} level

182

on u_*^{Th} is relatively minimal at low to mid-level fp^{Th} values, and is greater at the lowest (no footprint case) and highest values. In order to investigate interactive effects, redundant fp^{Th} levels were not amalgamated, and instead retained for the remaining analyses.

Table 4.4: Simplified ANOVA results table, displaying factors with significant effect on u_*^{Th} value, and the amount of overall variation explained by each. Values for interaction terms are shown individually and grouped. All listed factors are statistically significant for α<0.001.

Factor	Effect Sum of Squares	d.f.	Effect Mean Squares	F-score	% Variability Explained
fp_m	0.289	2	0.145	64.889	9.85
fp^{Th}	0.713	8	0.089	39.981	24.27
$u_*^{Th}{}_m$	0.286	1	0.286	128.337	9.74
Year	0.103	5	0.021	9.238	3.51
fp_m x fp^{Th}	0.124	18	0.008	3.491	4.24
fp_m x $u_*^{Th}{}_m$	0.222	2	0.111	49.901	7.57 (32.07)
fp^{Th} x $u_*^{Th}{}_m$	0.257	9	0.032	14.432	8.76
$u_*^{Th}{}_m$ x Year	0.338	5	0.068	30.318	11.50
Error	0.604	295	0.002	----	20.56
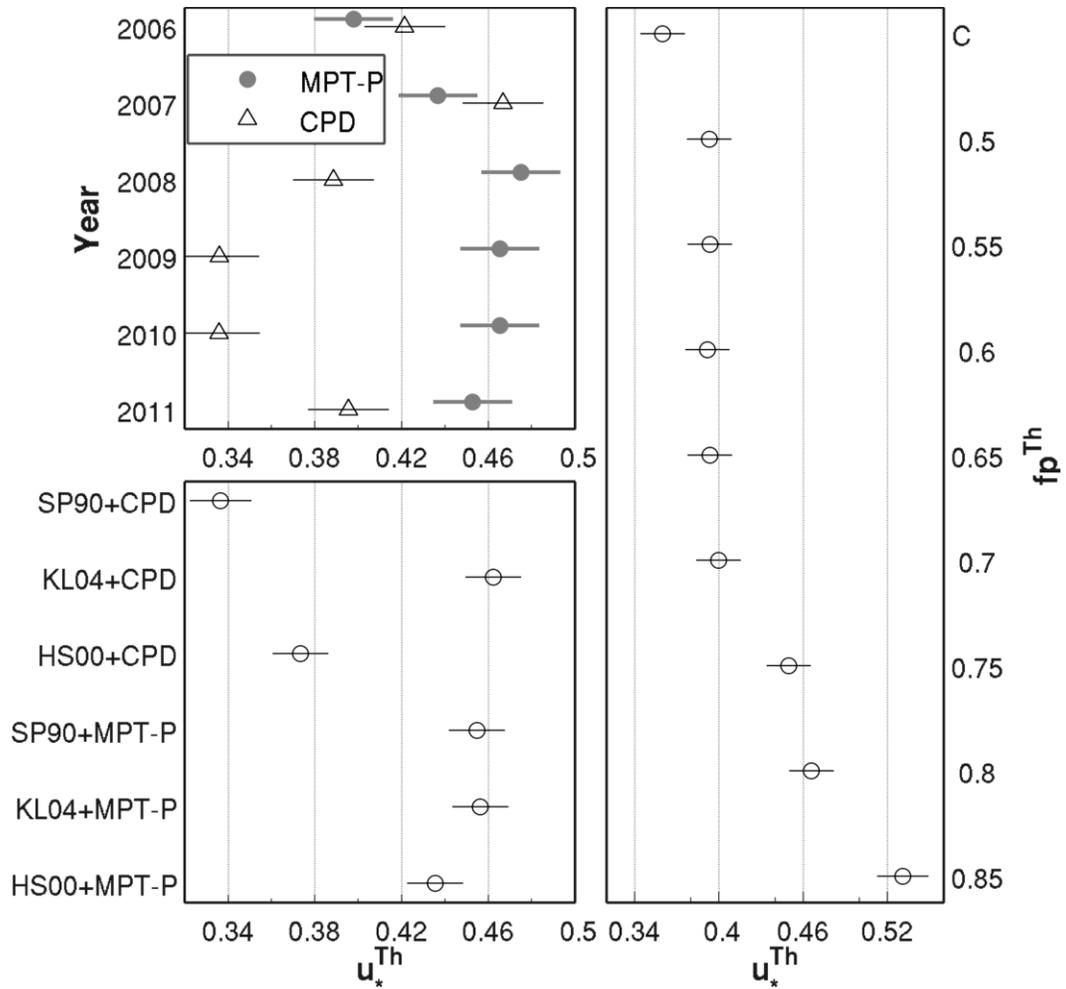
Figure 4.18: Marginal means (circle) and standard errors (line) for group-wise u_*^{Th} estimates, calculated using post-hoc multiple comparison tests from ANOVA results. Top left: Interannual u_*^{Th} marginal means for both MPT-P and CPD estimation methods. Bottom left: u_*^{Th} marginal means for footprint model x u_*^{Th} method combinations. Right: u_*^{Th} marginal means for each fp^{Th} level.

As demonstrated in Figure 4.18, u_*^{Th} estimates were subject to interannual variability. When only the primary effect was considered, the year-to-year effect was statistically significant, but its magnitude was relatively small ($+/-$ 0.04 m s^{-1}) in comparison with other first-level factors. The influence of interannual variability was further realized by a significant interaction effect between u_*^{Th} method and year (Table 3), indicating that the amount of interannual variability observed in u_*^{Th} was dependent on the u_*^{Th} method applied. CPD and MPT-P values differ significantly and considerably from each other for years 2008-2011 (Figure 4.18, panel a), both in terms of mean u_*^{Th} estimate, and year-to-year trends.

Comparing the methods, the MPT-P showed less interannual variability (CV = 6.1%) than did CPD (CV = 12.9%). This result contrasts the findings of Barr et al. (2013), who reported that the application of the CPD method reduced the magnitude of interannual variability as compared to MPT-P. Their result, however, was derived from aggregation across 38 different sites, and therefore, doesn't necessarily reflect the nature of interannual variation at our site. Reasons for the interannual variability observed at our site may be attributable to year-to-year differences in flux source areas (due to wind speed, direction and turbulence regimes), variation in canopy structure and density (e.g. leaf area index), as well as changes in the nature and magnitude of ecosystem respiration.

The strongest interaction was observed between footprint method and u_*^{Th} estimation method selections (fp_m x $u_*^{Th}{}_m$; Table 3; Figure 4.18). Applying the CPD method to SP90- or HS00-filtered data resulted in markedly lower u_*^{Th} estimates than obtained when applying the KL04 method – or when MPT-P was applied to any footprint filtered data. In an attempt to diagnose a previously unidentified influence on u_*^{Th} estimates, we further investigated the relationship between footprint filter-passing half-hourly u_* values for each footprint model and fp^{Th} combination, and the u_*^{Th} estimate produced by CPD and MPT-P. Results showed a wide difference in the distribution and mean of u_* values remaining after different footprint applications (Figure 4.19). Among footprint methods, half-hours retained by application of the SP90 model had consistently lower overall u_* values, while those produced from KL04-filtering (which applies an internal u_*^{Th} of 0.2 m s^{-1}) were consistently highest. Increases in fp^{Th} had little effect on the mean filter-passing u_* value for KL04, but coincided with mean u_* value increases for HS00. Mean u_* of SP90 filtered data increased with fp^{Th} up to 0.65, but then decreased afterwards, becoming increasingly variable as the number of filter-passing half-hours rapidly decreased beyond this point (Figure 4.19). The similarity between relative differences of mean filter-passing u_* values for each footprint method and the differences in fp_m x $u_*^{Th}{}_m$ marginal means (Figure 4.18, panel b) suggests a potential relationship between input u_* values and the u_*^{Th} values estimated by the CPD

and MPT-P methods.
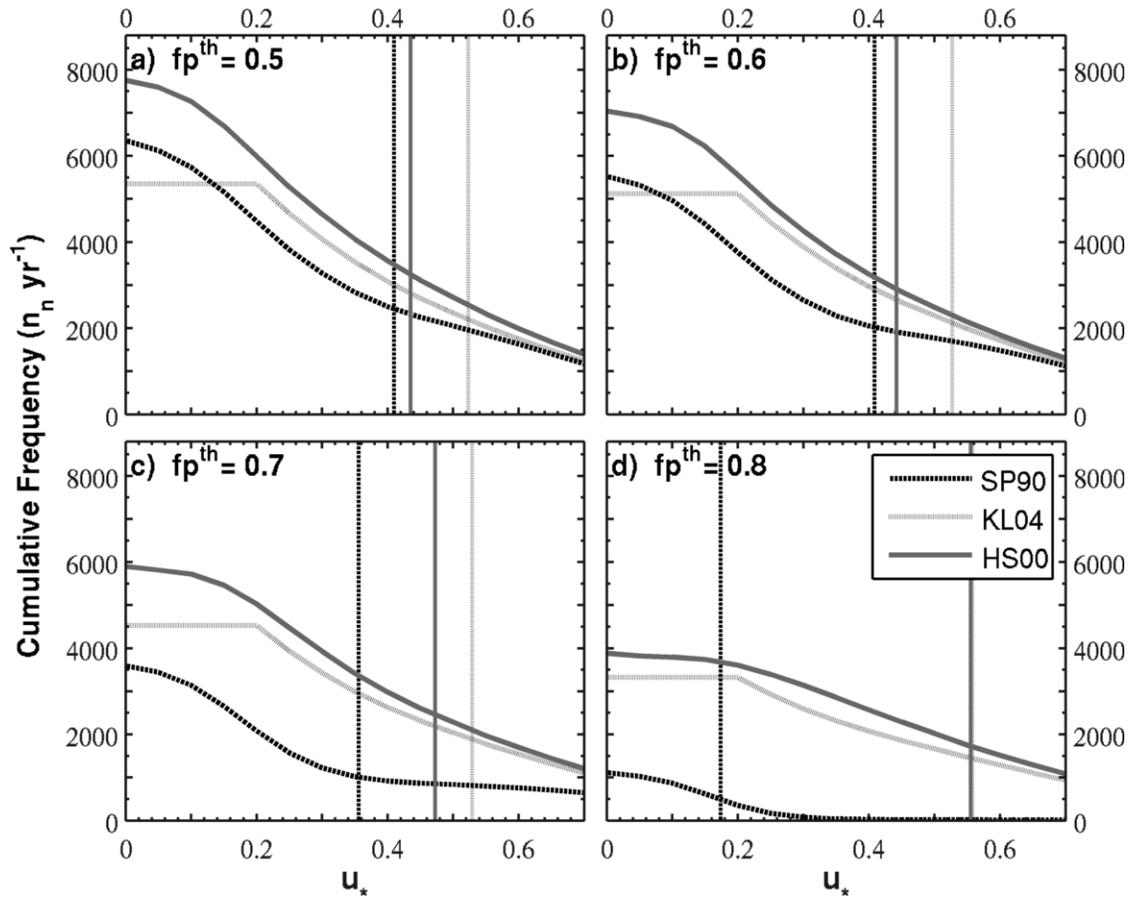
Figure 4.19: Cumulative frequency analysis of u* values among footprint filter-passing half-hours, for different footprint models and fp^{Th} values. Cumulative frequencies are shown as the number of half-hours with a value greater than that values on the abscissa. Cumulative amounts shown are average amounts across years 2006—2011. Vertical lines indicate the mean u* value for all footprint filter-passing half-hours for the specified footprint model and fp^{Th} value. Mean u* values are coincident for KL04 and HS00 in panel d).

Furthering this analysis, a linear regression was performed between mean input u_* and predicted u_*^{Th} for each method. Results revealed a differing response of CPD and MPT-P methods to variability in input u_* values (Figure 4.20). The CPD u_*^{Th} estimates showed a significant, strongly positive response ($R^2 = 0.96$, p~0) to mean input u_* value; on average, each unit increase in mean input u_* caused a 1.8 fold corresponding increase in the estimated u_*^{Th}. The relationship between mean input u_* and output u_*^{Th} was also significant and positive for MPT-P, though the relationship was weaker ($R^2 = 0.47$, $p < 0.001$), with a much smaller slope (0.43) than observed for CPD. Considered alongside the differences in input u_* values produced by each footprint model (Figure 4.19), this differential response between u_*^{Th} estimation models explains the strong fp_m x $u_*^{Th}_m$ interaction observed (Table 4.4, Figure 4.18). The mechanism for the apparent sensitivity of the CPD method to mean input u_* value is presently unclear. Further analyses are required to determine whether the CPD method is providing appropriate estimates of u_*^{Th} (that coincidentally scale with the mean value of input u_*), or whether this phenomenon represents a limitation of the CPD method itself. This finding does, however, prompt an important note (and caution) regarding inferences drawn from generalized data: Results shown in Figure 4.15 indicated that CPD-median and MPT-P-median estimates were equivalent in terms of mean estimated value; however subsequent analyses demonstrated fundamental differences between the estimation methods.
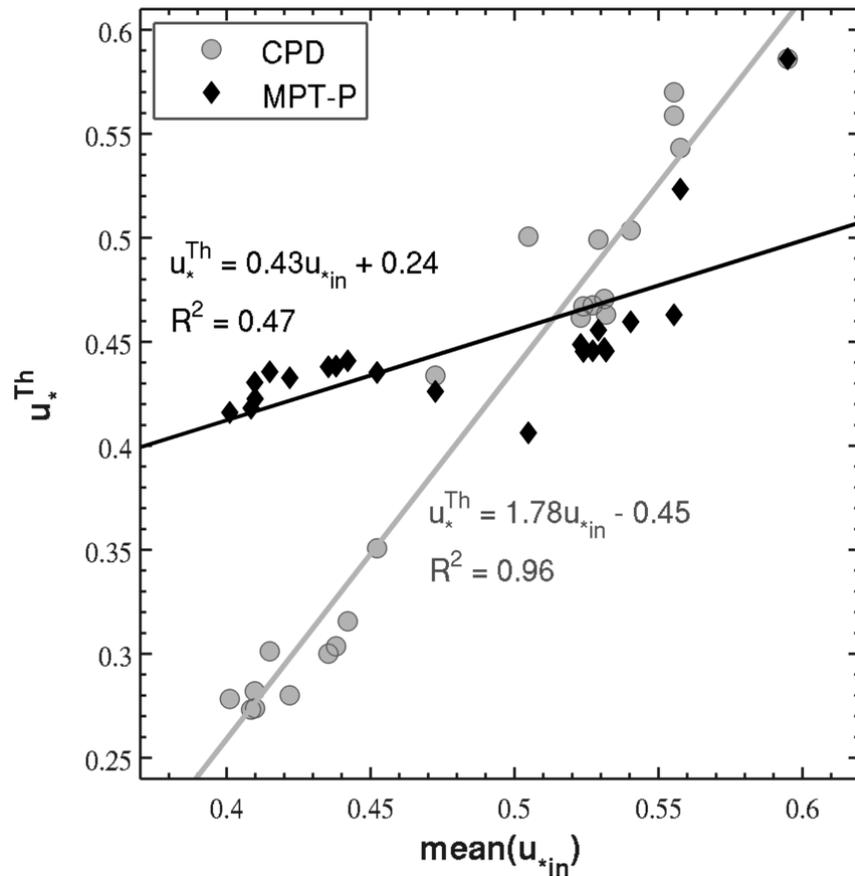
Figure 4.20: *Scatter plot of the mean input u_* for each footprint filtering combination versus u_*^{Th} estimates produced by the CPD (grey circles) and MPT-P (black diamonds) methods. Both trend lines are statistically significant for $\alpha < 0.001$.*

4.5.4 Combined effects of footprint and u_*^{Th} filtering on annual carbon exchange estimates

The effects of footprint filtering methods and u_*^{Th} estimates on annual ecosystem carbon exchanges (NEE, RE, GEP) were investigated through a factorial simulation, which calculated gap-filled sums for a wide range of possible u_*^{Th} values (0.1 to 0.7), for all fp^{Th} levels and footprint models (see section 4.4.6). The results are presented as surfaces in Figure 4.21, which depict the mean deviation of annual carbon exchange sums (ΔNEE, ΔRE, ΔGEP) for each footprint model, fp^{Th} and u_*^{Th} combination. Deviations were calculated as the difference between the estimate of a given pixel (fp^{Th}, u_*^{Th} combination), and the value corresponding to the bottom-left pixel of the surface (no footprint, u_*^{Th} = 0.1); this approach was used to standardize results against interannual variability. Superimposed on these surfaces are actual u_*^{Th} estimates for each u_*^{Th} method, year and fp^{Th} combination (see section 4.5.3), which provides a means of quantifying the implications of these treatments on annual carbon exchange sums.

The general effect of u_*^{Th} on annual carbon exchange values is demonstrated as north-south variation in the panels of Figure 4.21. Across footprint models, increasing u_*^{Th} generally resulted in corresponding increases to annual RE estimates. The rate of RE increase with u_*^{Th} was strongest between 0.2 and 0.4 m s^{-1}, and lessened above this value. The net effect of increasing u_*^{Th} from 0.1 to 0.5 m s^{-1} was to increase annual RE by as much as 250 g C m^{-2} y^{-1} for a given footprint model and fp^{Th} value.

A notable deviation from general trends was observed when the SP90 footprint model was applied to data with a fp^{Th} greater than 0.7. In these cases, the proportion of data removed by footprint- and u_*^{Th} filtering applications was beyond the tolerance limits of the NLR-HL gap-filling model, resulting in RE (and subsequently, NEE) estimates with large uncertainty and a strong negative bias. The extensive data removal in these cases were attributed to combined effects of stringent SP90 footprint filtering (Figure 4.9), and its tendency to retain low-u_* values (Figure 4.19), which were removed as u_*^{Th} increased (demonstrated in Figure 4.22). The effect of a u_*^{Th} increase was also positive for annual GEP, though the rate of increase and strength of the effect were both less than for RE. Net increase in GEP as a result of u_*^{Th} increase was similar across all fp models and as much as 120 g C m^{-2} y^{-1}. As a result of the partially-offsetting increases in both RE and GEP, NEE increased with u_*^{Th} at a lower rate and to a lesser amount than found for RE. An increase of u_*^{Th} from 0.1 to 0.5 m s^{-1} generally had a net positive effect on NEE (meaning less ecosystem carbon sequestration) in the order of 120 g C m^{-2} y^{-1}.
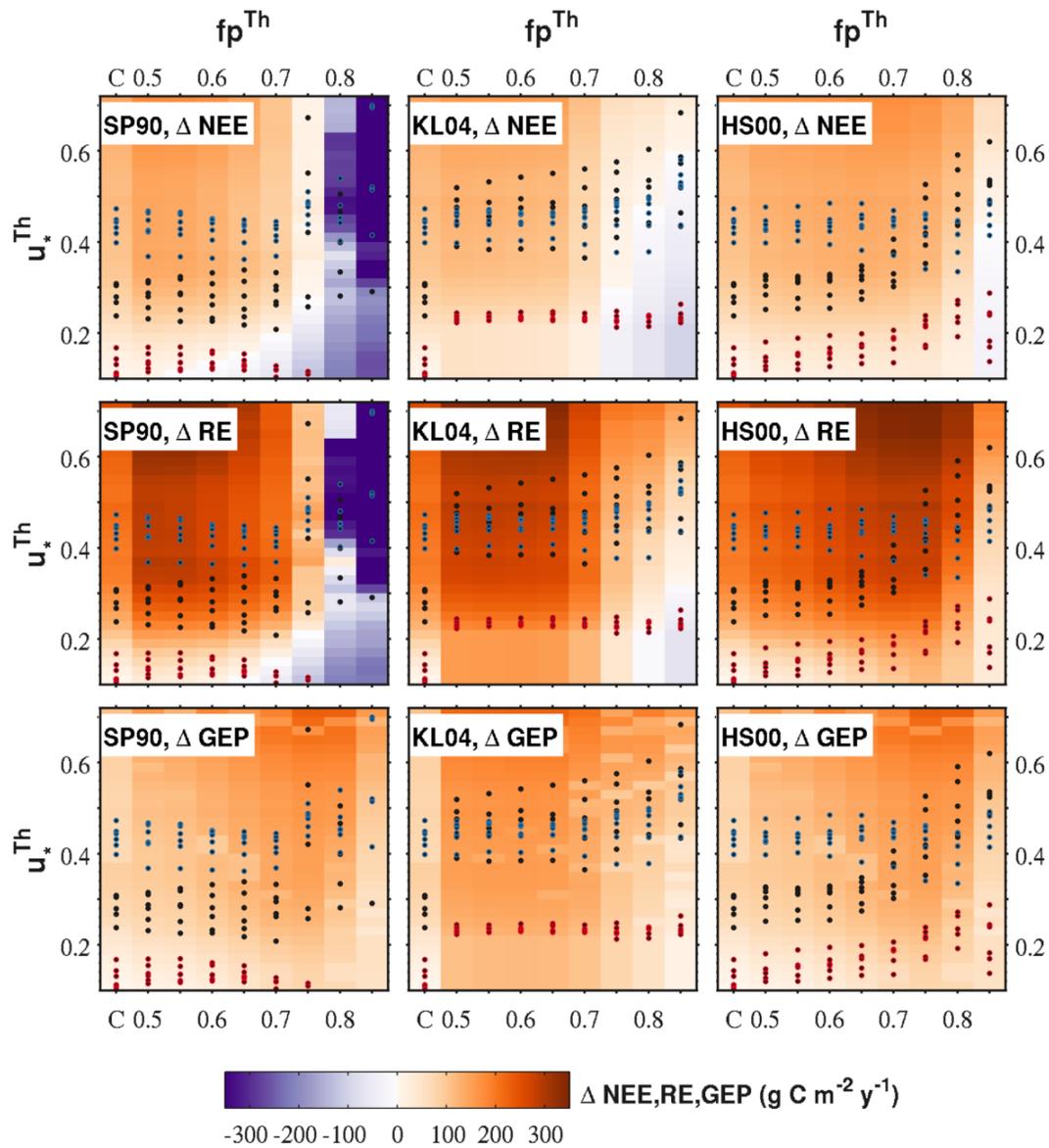
Figure 4.21: Pseudocolour surfaces representing mean deviation of annual carbon exchange sums (ΔNEE, ΔRE, ΔGEP) across all fp^{Th} and u_*^{Th} combinations, for each footprint model. Deviations represent the difference between estimates for a given fp_m, fp^{Th}, u_*^{Th} combination and the value corresponding to the bottom-left pixel (no footprint, $u_*^{Th} = 0.1$). Actual u_*^{Th} estimates for each footprint model, fp^{Th} level, and year are superimposed on surfaces as coloured dots (CPD, black; MPT-P, blue; MPT-G, red)
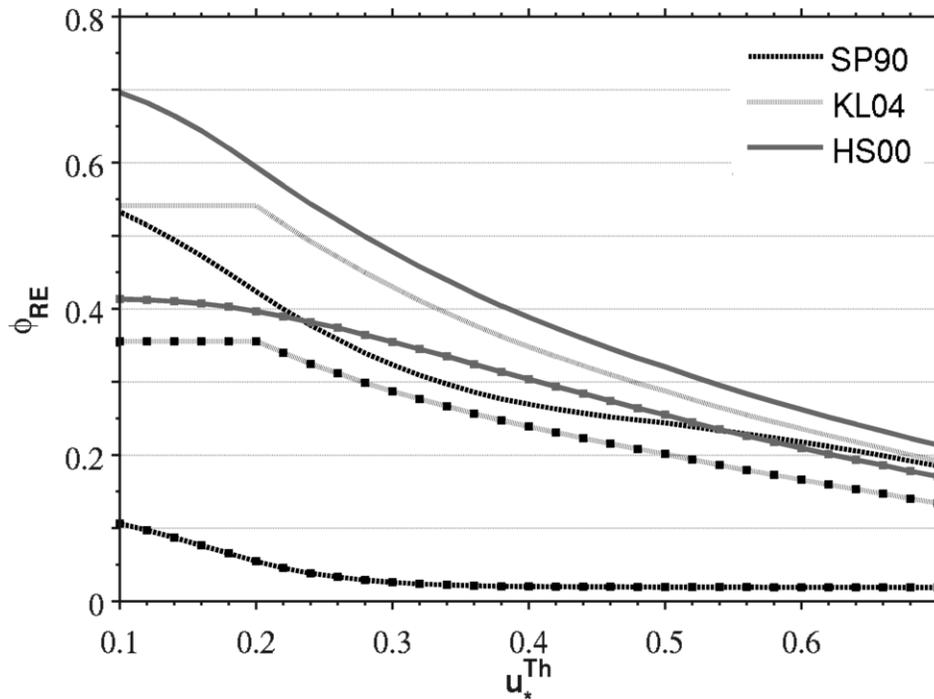
Figure 4.22: *Effect of applied u_*^{Th} value on the available data fraction of respiration-parameterizable data (ϕ_{RE}). Values are shown separately for each footprint model type at two fp^{Th} levels: $fp^{Th}=0.6$ (no markers), and $fp^{Th}=0.8$ (square markers).*

When these surfaces were considered in the context of actual u_*^{Th} estimates (included in Figure 4.21), the large discrepancies between u_*^{Th} methods resulted in NEE, RE and GEP sums that were consistently larger for MPT-P than for MPT-G, with NEE differences generally ranging between 50 and 100 g C m^{-2} y^{-1}. Estimates associated with the CPD method ranged between these two extremes, due to the strong interaction effects of this method with footprint model and fp^{Th}.

The effects of fp^{Th} value on annual carbon exchange estimates (west-to-east trends in Figure 4.21), showed a variable effect. Increasing fp^{Th} from the control case (no footprint) to moderate values (0.5 to 0.7) resulted in a strong positive increase in annual RE. For SP90 and KL04, the effect was most pronounced between the control case and $fp^{Th} = 0.5$, where RE increased by approximately 100 g C m^{-2} y^{-1}; subsequent increases in fp^{Th} through 0.7 had little effect on RE. For the HS00 model, the magnitude of RE increase was similar, but the rate was more gradual, peaking around $fp^{Th} = 0.75$. Disregarding problematic results for SP90, annual RE decreased sharply with fp^{Th} at high values (fp^{Th} > 0.7 for KL04, and > 0.75 for HS00), as estimates were as much as 200 g C m^{-2} y^{-1} lower at $fp^{Th} = 0.85$ than at the mid-value peak.

GEP demonstrated a similar, but muted relationship with fp^{Th} as observed for RE; annual values increased up to 50 g C m^{-2} y^{-1} at moderate fp^{Th} values, and decreased by approximately the same magnitude at the highest values. The combined effect of these RE and GEP trends are demonstrated in the fp^{Th} relationship with NEE (Figure 4.21, top panels): The greater amplitude of RE changes relative to GEP resulted in a pattern of annual sums that increased between low and medium fp^{Th}, reaching a peak deviation of approximately +100 g C m^{-2} y^{-1}. This trend reversed at high fp^{Th}, and *ΔNEE* became negative for fp^{Th} > 0.7 and 0.8 for KL04 and HS00, respectively. Overall, KL04 and HS00 showed similar responses to changes in fp^{Th} and u_*^{Th}, though the trends in HS00 results were shifted along the dimension of fp^{Th} by approximately 0.2.

The relationships shown here between fp^{Th}, u_*^{Th} and annual carbon exchange sums provide substantive information about the nature and representativeness of carbon fluxes measured at our site. A general increase of annual RE with u_*^{Th} is consistent with findings of other u_*^{Th} studies (Barr et al., 2013; Goulden et al., 1996; Papale et al., 2006; van Gorsel et al., 2009), and the general plateau of this relationship above approximately $u_*^{Th} = 0.3$ m s^{-1} means that differences between u_*^{Th} estimation methods above this value will have minimal implications for annual carbon exchange estimates – though this value is not attained in all scenarios.

Perhaps a more interesting and important relationship is the increase and subsequent decrease of RE

(and subsequently NEE) with increasing fp^{Th}. This effect suggests that footprint applications are capturing real differences in carbon exchange magnitudes between the TP39 forest and the larger region. With the exclusion of SP90 results, high fp^{Th} are coincident with considerable decreases in annual RE and NEE, implying that regions in proximity to the EC measurement tower (i.e. the TP39 site), have a greater net productivity than the surrounding ecosystems. This result further supports the use of a footprint filter at our site, and establishes the need for an objective method of evaluating which filtering approach is most appropriate and physically correct at our site.

When considered in the context of actual results, the carbon exchange sums that emerge from changing fp^{Th} values represented an interplay between two influences: the direct effect of fp^{Th} levels on the annual sums (demonstrated by the non-linear west-to-east trends in the surfaces of Figure 4.21); and the positive relationship between fp^{Th} and estimated u_*^{Th} value. These two effects were positive and reinforcing for low to medium fp^{Th} values (0 to 0.7), leading to strongly positive deviations in NEE, RE and GEP. Above this level, however, the direct effect of fp^{Th} transitioned to negative, and served to offset the positive indirect effect on u_*^{Th}. This effect is particularly strong for CPD estimates, due to the sensitivity of its u_*^{Th} estimates to the applied fp^{Th} (Figure 4.20).

These results demonstrate that the selection and application of a footprint method and fp^{Th} value affects final carbon exchange sums in three different ways:

1. Footprint application selectively removes flux data that is deemed non-representative of the ecosystem of interest. Considering that these excluded points are contaminated by fluxes from differing ecosystems, the resulting values used for parameterizing gap-filling models and deriving annual sums will differ according to which method and fp^{Th} is used.

2. Footprint application affects the magnitude of u_*^{Th} estimates. As seen in Figure 4.19, footprint models vary in terms of the distribution of retained u_* values. Figure 4.20 shows that the nature of input u_* influences the estimates of u_*^{Th} created (much more so for CPD than for MPT-P), and Figure 4.21 demonstrates that the value of estimated u_*^{Th} influences annual sums of carbon exchanges.

3. Footprint applications remove data from NEE time series, thus affecting the performance of the u_*^{Th} estimation methods and gap-filling models that utilize this data. Both these applications show a considerable resilience to data gaps (see Chapter 3; Barr et al., 2013); however, data removal above 90% of nighttime data (as seen for SP90 at high fp^{Th}), causes failure in both u_*^{Th}

estimation and gap-filling models, thus resulting in unrealistic and erroneous annual estimates.

4.5.5   Effect of filtering on annual carbon exchange estimates

To further investigate the relationship between EC filtering and annual carbon exchange estimates at our site, we standardized estimates for each factorial run with expected values for given ϕ and year (see section 4.4.6). Results were structured according to ϕ_{RE} to investigate the effect of increasing data removal by filtering. Results showed that filtering operations had a strong, non-negligible effect on annual values (Figure 4.23). Compared to annual values expected for the same ϕ_{RE} and year, NEE, RE and GEP were all positively biased by filtering operations, with median increases at the largest ϕ_{RE} values of 131, 265 and 126 g C m^{-2} y^{-1}, respectively (Figure 4.23, right panels). Since the highest ϕ_{RE} values corresponded to the no-footprint (fpTh = 0) case, this initial offset in carbon exchange estimate deviations was completely the result of u$_*^{Th}$ filtering application. As ϕ_{RE} decreased (primarily due to footprint filtering), deviations of annual carbon exchange sums varied in response. NEE deviation (δNEE) generally decreased with decreasing ϕ_{RE}, while both δRE and δGEP increased and then decreased over this range. As ϕ_{RE} decreased, δRE increased by an average of 62 g C m^{-2} y^{-1}, to peak in the range 0.35 < ϕ_{RE} < 0.4, while GEP peaked at a lower ϕ_{RE} (0.25) with a similar average increase (57 g C m^{-2} y^{-1}). The spread in annual estimates was negatively correlated with ϕ_{RE}, and was most pronounced for δRE values.

We also aggregated gap-filling model performance metrics (section 4.4.6) according to ϕ_{RE}, to explore potential changes in model goodness-of-fit and bias with filtering applications. Though spread existed in results, the average goodness of fit measures (WESS, R^2) reached distinct optimal values in the vicinity of 0.15 < ϕ_{RE} < 0.20 (Figure 4.24, right panels). From this peak, a gradual decrease in performance was found for increasing ϕ_{RE}, while goodness-of-fit worsened abruptly and markedly for ϕ_{RE} <0.15. Similarly, bias (approximated by AE) was not substantially negatively affected until ϕ_{RE} fell below 0.15. Results for both annual gap-filled sums and gap-filling model performance statistics suggest that, on the whole, the filtering operations applied to the EC-measured data have a considerable effect on ecosystem exchange estimates, and the correspondence of the filter-passing data to the parameterized models. Trends in annual sums across ϕ_{RE} imply a source-area control on these values, and trends in statistical performance metrics suggest that filter-passing data is most consistent with model estimates when footprint filtering is relatively stringent.
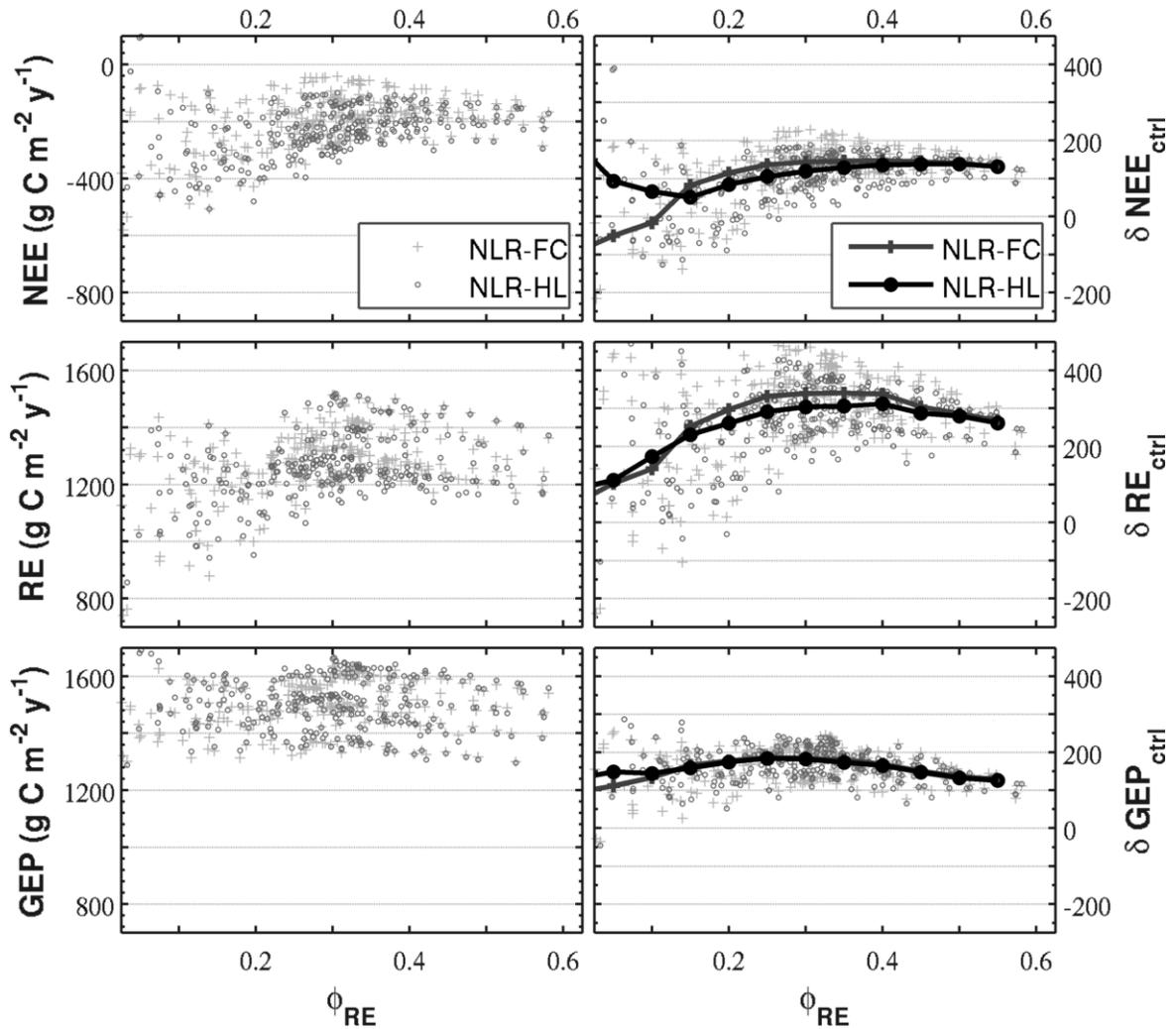
Figure 4.23: *Annual carbon exchange estimates aggregated for all factorial filtering runs and all years, presented as a function of ϕ_{RE}. The left panels show absolute results (uncorrected for interannual and ϕ_{RE} effects), while the right panels present the results as standardized deviations from expected values.*
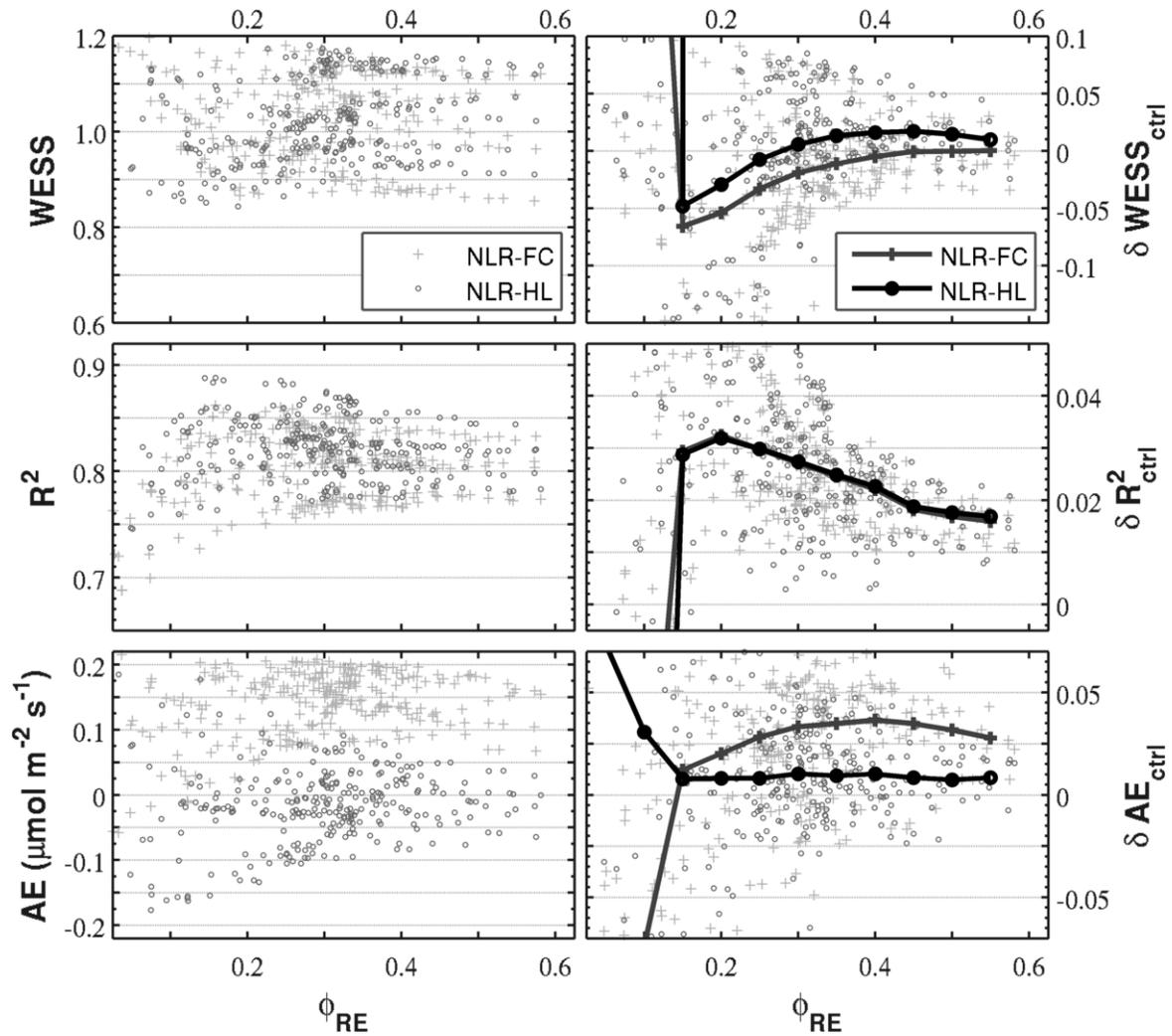
Figure 4.24: *Gap-filling model performance metrics aggregated for all factorial filtering runs and all years, presented as a function of ϕ$_{RE}$. The left panels show absolute results (uncorrected for interannual and ϕ$_{RE}$ effects), while the right panels present the results as standardized deviations from expected values. Desired statistical results are minimized WESS, maximized R^2, and an AE value of zero.*

4.5.6　　　Flux filtering and controls on annual exchange estimates

Considering the numerous factors (footprint model, fp^{Th} level, year and u_*^{Th} estimation method) that may affect the observed trends in annual NEE (Figure 4.23), we conducted ANOVAs to quantify their influence on annual estimates, and to characterize any first-level interactions between them. Results showed that interannual variability had the largest overall effect on annual gap-filled NEE estimates, explaining 50 and 41% of the total variability for NLR-FC and NLR-HL outcomes, respectively (Tables 4.5 and 4.6). The selected footprint model (fp_m) and fp^{Th} level each exhibited strong, relatively similar mean effects on annual NEE values; however, due to its greater degrees of freedom, fp^{Th} level explained more of the total variability (21 and 18 % for NLR-FC and NLR-HL) than fp_m (5.5% for each). The $fp_m * fp^{Th}$ factor was the strongest of the interaction terms; taken together, all interaction terms accounted for approximately 15% of total NEE estimate variability. Though the application of u_*^{Th} filtering was shown to have a considerable effect on NEE estimates (Figure 4.23), results showed that the selection of a particular u_*^{Th} method was insignificant to the resulting annual NEE estimates (p = 0.96 for NLR-FC; p = 0.2 for NLR-HL). All first-level u_*^{Th} interaction terms were insignificant or very small, suggesting that both of the tested methods (CPD, MPT-P) produce equivalent effects. Overall, the ANOVA findings reinforce the need to standardize annual results for the effect of interannual variability, and demonstrates the importance of footprint model and fp^{Th} level on NEE values estimated for this site.

Post-hoc multiple comparison tests were run on ANOVA results to test for significant differences in marginal means of RE, GEP and NEE sums between years, footprint models and fp^{Th} levels. Results for RE (Figure 4.25) showed interannual variability to have the largest effect on annual gap-filled estimates, as variation was as much as 280 g C m^{-2} y^{-1} between the year with the lowest value (2011) and the highest (2010). In comparison, annual RE was less sensitive to the selection of footprint model, as variability was less than 50 g C m^{-2} y^{-1} between the all models. RE estimates were highest when the HS00 model was applied and lowest for SP90. In terms of fp^{Th} level, applying a fp^{Th} of 0.5 resulted in a step increase of 75 g C m^{-2} y^{-1} compared to the no-footprint case. This effect remained with increasing fp^{Th} up to values equal or greater than fp^{Th} = 0.7, at which point the annual estimates were negatively affected. Relative to the no-footprint results, application of fp^{Th}=0.8 caused a net reduction in annual RE estimates.

Table 4.5: *ANOVA results table of effects on annual NEE estimate obtained using the NLR-FC gap-filling model. Factors with p-values listed as zero are statistically significant for α<0.0001.*

Factor	Effect Sum of Squares	% Variability Explained	d.f.	Effect Mean Squares	F-score	p-value
fp_m	157102	5.5	2	78551.1	84.47	0
fp^{Th}	597433	20.7	8	74679.1	80.30	0
$u_*{}^{Th}{}_m$	3	0	1	2.9	0.00	0.96
Year	1451125	50.4	5	290224.9	312.08	0
fp_m x fp^{Th}	219284	7.6	16	13705.3	14.74	0
fp_m x $u_*{}^{Th}{}_m$	786	0.0	2	393.2	0.42	0.66
fp_m x Year	32971	1.1	10	3297.1	3.55	0
fp^{Th} x $u_*{}^{Th}{}_m$	22191	0.8	8	2773.9	2.98	0
fp^{Th} x Year	188083	6.5	40	4702.1	5.06	0
$u_*{}^{Th}{}_m$ x Year	8540	0.3	5	1708	1.84	0.11
Error	197156	6.8				
Total	2881314		309			

Table 4.6: *ANOVA results table of effects on annual NEE estimate obtained using the NLR-HL gap-filling model. Factors with p-values listed as zero are statistically significant for α<0.0001.*

Factor	Effect Sum of Squares	% Variability Explained	d.f.	Effect Mean Squares	F-score	p-value
fp_m	127369	5.5	2	63684.7	65.78	0
fp^{Th}	408502	17.8	8	51062.7	52.75	0
$u_*{}^{Th}{}_m$	1624	0.1	1	1623.9	1.68	0.2
Year	931414	40.6	5	186282.7	192.42	0
fp_m x fp^{Th}	156084	6.8	16	9755.2	10.08	0
fp_m x $u_*{}^{Th}{}_m$	9682	0.4	2	4841	5.00	0.01
fp_m x Year	36714	1.6	10	3671.4	3.79	0
fp^{Th} x $u_*{}^{Th}{}_m$	4202	0.2	8	525.2	0.54	0.82
fp^{Th} x Year	147424	6.4	40	3685.6	3.81	0
$u_*{}^{Th}{}_m$ x Year	8050	0.4	5	1610	1.66	0.15
Error	200394		207	968.1		
Total	2295037		304			
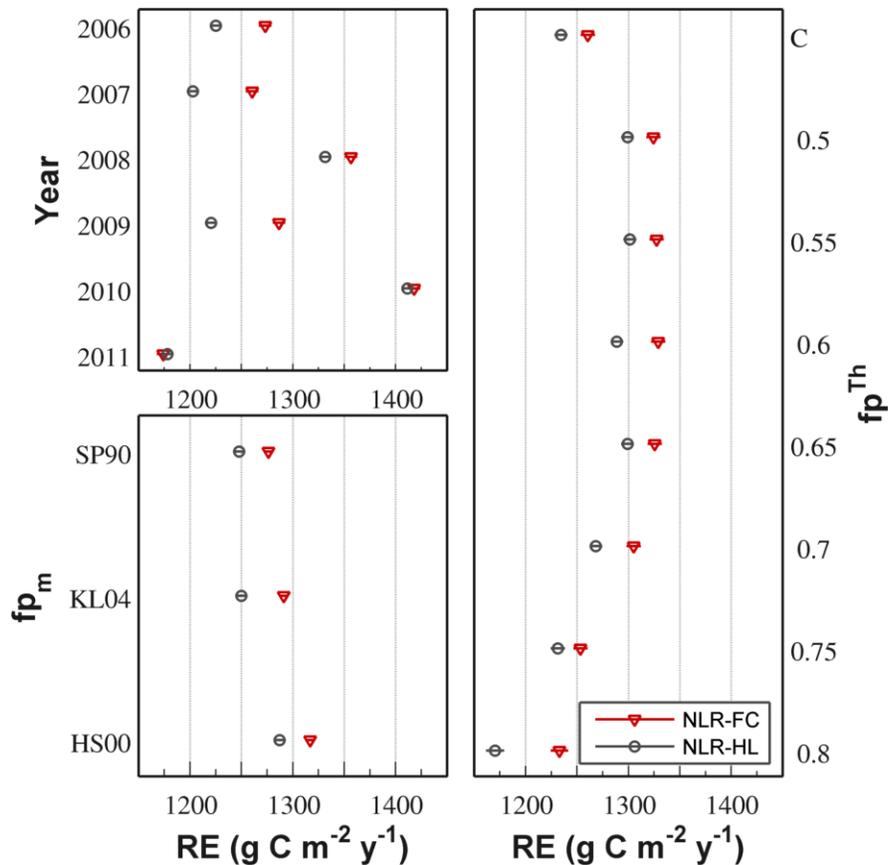
Figure 4.25: *Marginal means (symbols) and standard errors (line) for group-wise RE estimates, calculated using post-hoc multiple comparison tests from ANOVA results. Top left: Interannual RE marginal means. Bottom left: RE marginal means for footprint model selection. Right: RE marginal means for each fp^{Th} level. Note that marginal means for fp^{Th} are only shown up to 0.8, as a result of non-normality of higher-level results due to gap-filling model failure with SP90-filtered data above this fp^{Th} level.*

Interannual variability was comparably strong for GEP (Figure 4.26), however year-to-year trends in annual values were somewhat discordant (low value in 2007; high in 2011). Annual GEP was less sensitive to footprint model selection than RE, as all marginal mean estimates were within 25 g C m^{-2} y^{-1}. Variability in GEP due to fp^{Th} level was less than for RE as well. Though a similar step increase occurred from the no-footprint case to fp^{Th} = 0.5, this effect was maintained for most of the fp^{Th} range; only a small decrease occurred at fp^{Th} = 0.8, and values remained higher than the no-footprint case. The results and trends for annual NEE values (Figure 4.27) were considered in the context of its component

fluxes (RE, GEP). Inter-annual variability was less pronounced for NEE than for either RE or GEP, suggesting compensating effects of component fluxes. Footprint model-related effects were similar in magnitude to that of RE, and NEE was most negative for SP90 application, followed by KL04 and HS00. As a result of the strong RE decrease at high fp^{Th} levels and the relative consistency of GEP estimates over this range, NEE was strongly negatively influenced by increases in fp^{Th} above 0.65. Estimates for $fp^{Th} = 0.8$ were as much as 150 g C m^{-2} y^{-1} lower than the no-footprint case, indicating that the strictness with which footprint filtering is applied has large implications for annual NEE estimates.
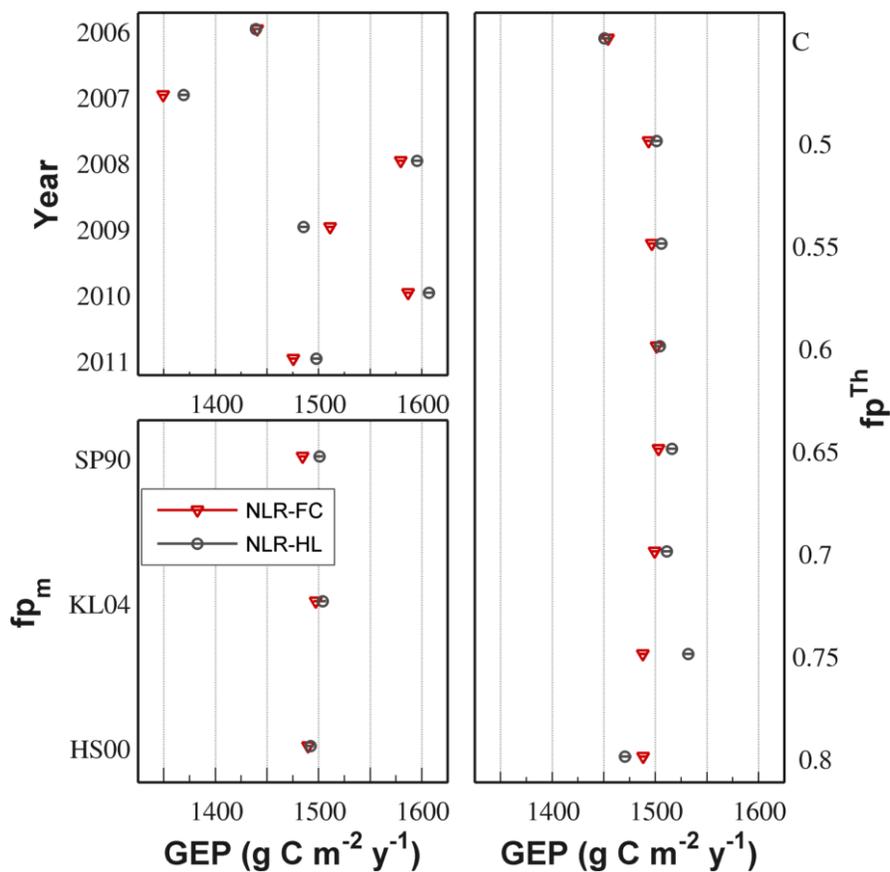
Figure 4.26: *Marginal means (symbols) and standard errors (line) for group-wise GEP estimates, calculated using post-hoc multiple comparison tests from ANOVA results. Top left: Interannual GEP marginal means. Bottom left: GEP marginal means for footprint model selection. Right: GEP marginal means for each fp^{Th} level. Note that marginal means for fp^{Th} are only shown up to 0.8, as a result of non-normality of higher-level results due to gap-filling model failure with SP90-filtered data above this fp^{Th} level.*
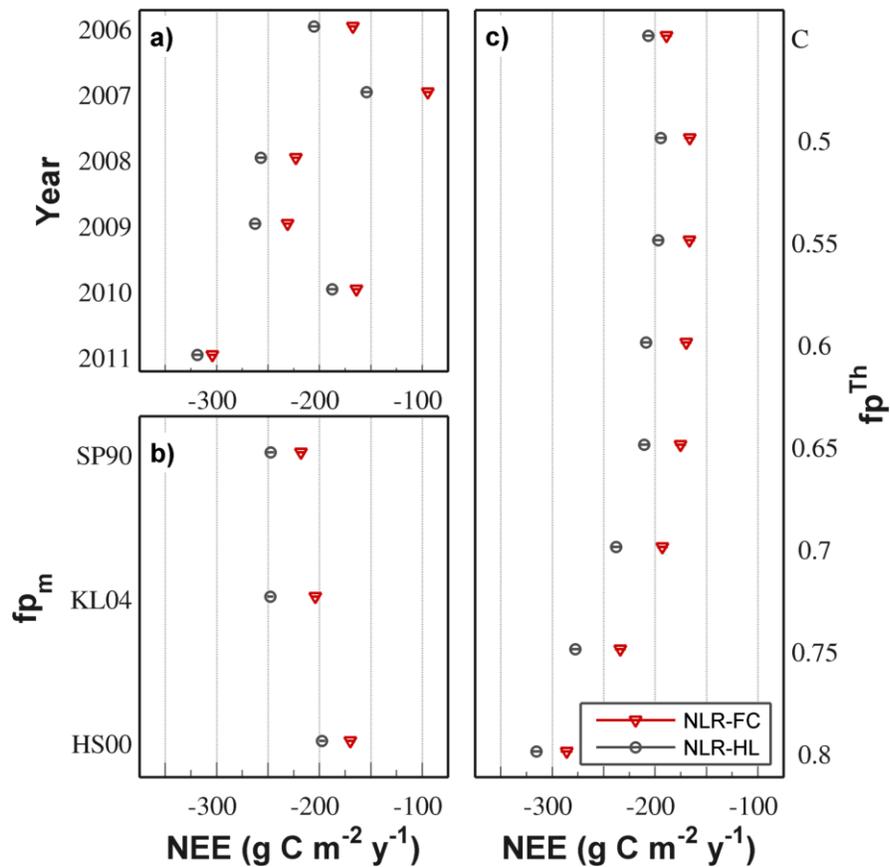
Figure 4.27: *Marginal means (symbols) and standard errors (line) for group-wise NEE estimates, calculated using post-hoc multiple comparison tests from ANOVA results. Top left: Interannual NEE marginal means. Bottom left: NEE marginal means for footprint model selection. Right: NEE marginal means for each fp^{Th} level. Note that marginal means for fp^{Th} are only shown up to 0.8, as a result of non-normality of higher-level results due to gap-filling model failure with SP90-filtered data above this fp^{Th} level.*

4.5.7   Evaluating performance of flux filtering applications

Extending the results of the previous sections, we conducted additional ANOVAs and multiple comparison tests on both carbon exchange sums and gap-filling model performance metrics, in order to identify the most appropriate footprint filtering specifications for our site. These specifications were evaluated on the constraints that methods should maximize goodness-of-fit metrics, while maintaining low annual estimate uncertainty and bias. We constrained our analyses by removing non-significant factors, and removing the effects of interannual variability and ϕ by performing tests on deviation values of sums and statistics (see section 4.4.6). Multiple comparison tests showed consistent and significant differences in performance metrics between footprint models and fp^{Th} levels across both gap-filling methods (Figure 4.28). Goodness-of-fit measures (WESS, R^2) were relatively similar across footprint models at low fp^{Th} values, but deviated at higher levels. Across both gap-fulling models and for both R^2 and WESS, KL04 results consistently had the best performance metrics. HS00 showed similar responses of goodness-of-fit metrics to fp^{Th}, but at a fp^{Th} lag of approximately 0.1. Both KL04 and HS00 showed similar, consistent bias estimates across fp^{Th}. Conversely, performance of SP90 results diminished rapidly beyond $fp^{Th} = 0.7$, suggesting a rapid increase in gap-filling model uncertainty and bias at high fp^{Th} values.

In order to investigate the effect of increasing fp^{Th} (and therefore, removing additional data points) on the uncertainty and reliability of annual estimates, we also used multiple comparison tests to compare the inter-quartile range (IQR) of annual carbon exchange estimate deviations across footprint models and fp^{Th} levels (Figure 4.29). For both gap-filling models, RE IQR demonstrated a greater sensitivity to fp^{Th} (and thus, data removal), than GEP or NEE, due to the fact that footprint filtering removed a relatively larger proportion of RE-parameterizable (nighttime) data.   IQR was typically lowest for KL04, and was relatively consistent (between 50 to 100 g C m^{-2} y^{-1}) for fp^{Th} levels less than 0.85. IQR of SP90-associated estimates increased rapidly above $fp^{Th} = 0.7$, indicating that the degree of data removal surpassed the tolerance of the gap-filling models to produce reliable estimates. Similar effects were seen for the remaining models, though at higher fp^{Th} values (greater than 0.8 for KL04; above 0.85 for HS00).

Considering the competing constraints for ecosystem-representativeness and gap-filling model reliability, it can be inferred from the combined results of these tests that the optimal footprint

specification is the KL04 applied at a fp^{Th} level of 0.8. This combination maximizes the goodness-of-fit, while ensuring consistency of gap-filling model estimates.
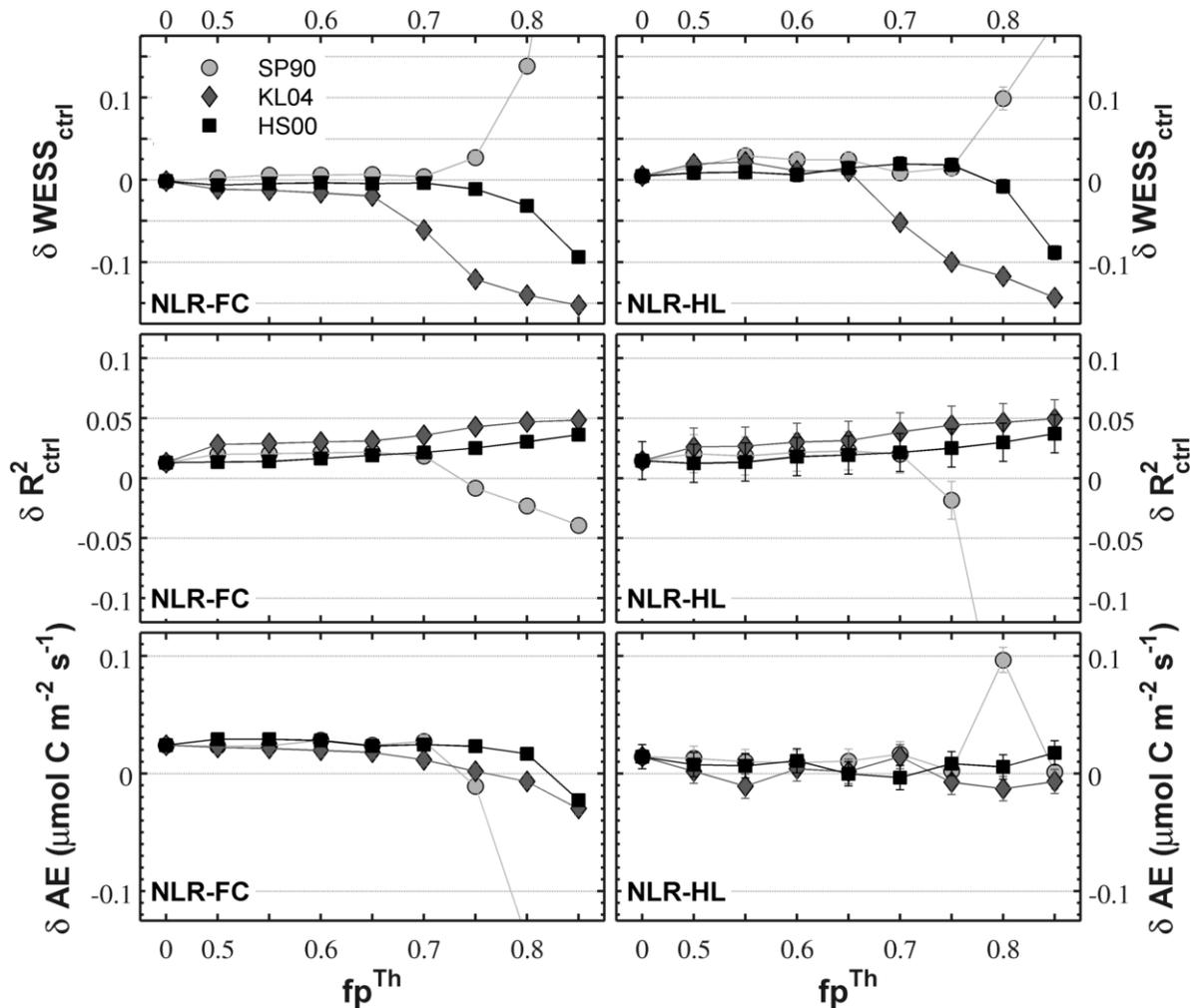
Figure 4.28: *Marginal means (symbols) and standard errors (lines) of standardized gap-filling model statistical performance metrics for each footprint model, as a function of fp^{Th} level. Group-wise means were calculated using post-hoc multiple comparison tests from ANOVA results. Results are separated by those for the NLR-FC gap-filling model (left panels) and for the NLR-HL gap-filling model (right panels).*
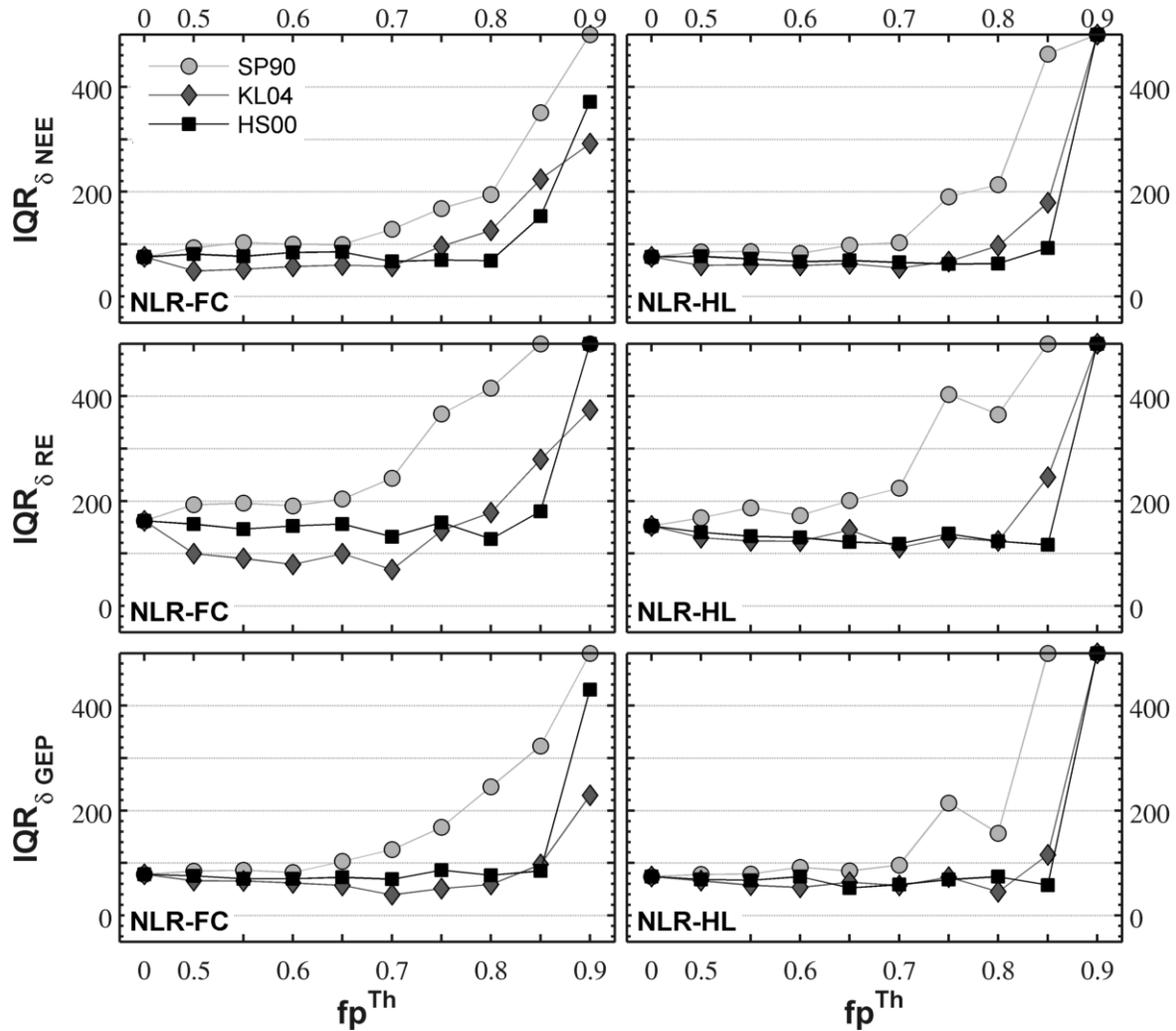
Figure 4.29: *Marginal means (symbols) and standard errors (lines) of inter-quartile range of standardized gap-filled annual carbon exchange estimates for each footprint model, as a function of fp^{Th} level. Group-wise means were calculated using post-hoc multiple comparison tests from ANOVA results. Results are separated by those for the NLR-FC gap-filling model (left panels) and for the NLR-HL gap-filling model (right panels).*

4.5.8 Effects of footprint filtering specification on estimated carbon exchanges

To characterize the implications of footprint model and fp^{Th} level selection for carbon exchange information extracted from EC-measured data, we furthered our investigation to explore the differential effects on annual estimates, intra-annual trends and half-hourly relationships derived from filtered data. Comparison of annual carbon exchange estimates showed RE was the most sensitive to changes in footprint specification (Figure 4.30). Across all footprint and gap-filling models, annual RE deviation peaked and then declined with increasing fp^{Th}, though these trends were offset in terms of fp^{Th} levels. The rapid decline of RE estimates for SP90-treated results beyond $fp^{Th} = 0.7$ corresponds with general gap-filling model failure described earlier. Applying the recommended footprint model and fp^{Th} specification of (KL04, 0.8) resulted in RE estimates that were on average 50 to 80 g C m^{-2} y^{-1} lower than the no-footprint case, and approximately 120 g C m^{-2} y^{-1} lower than HS00 applied at the same fp^{Th} level. All RE deviations in this case were considerably higher than the control estimates, which demonstrates the strong positive influence of u_*^{Th} application on annual RE estimates, since u_*^{Th} filtering was not applied to control case estimates.

Differences between GEP estimates among footprint models were comparatively less, suggesting that estimates for this component are more robust than for RE. Interestingly, the 200 g C m^{-2} y^{-1} offset introduced into RE estimates from u_*^{Th} application was halved for GEP estimates, although RE estimates were added to NEE measurements to produce GEP estimates, from which gap-filling models were parameterized and used to fill. In contrast to RE, applying the (KL04, 0.8) footprint filter specification resulted in a net GEP increase of approximately 40 g C m^{-2} y^{-1} relative to the no-footprint case. Effects of filtering specification on NEE estimates reflected the combined effects on RE and GEP. At low-to-medium fp^{Th} levels, trends in GEP estimates compensated for changes in RE in a manner that resulted in consistent estimates for fp^{Th} up to 0.65. Above this fp^{Th} level, however, decreases in RE were not matched by similar changes in GEP, and as a result, NEE estimates were markedly lower than the no-footprint case. Applying (KL04, 0.8) resulted in NEE estimates that were between 100 and 120 g C m^{-2} y^{-1} lower than for the no-footprint scenario. These results suggest that appropriate footprint application resolved ecosystem non-representativeness to a greater extent for RE-parameterizable (nighttime and non-growing season) measurements than those for GEP-parameterizable (daytime, growing season) data.
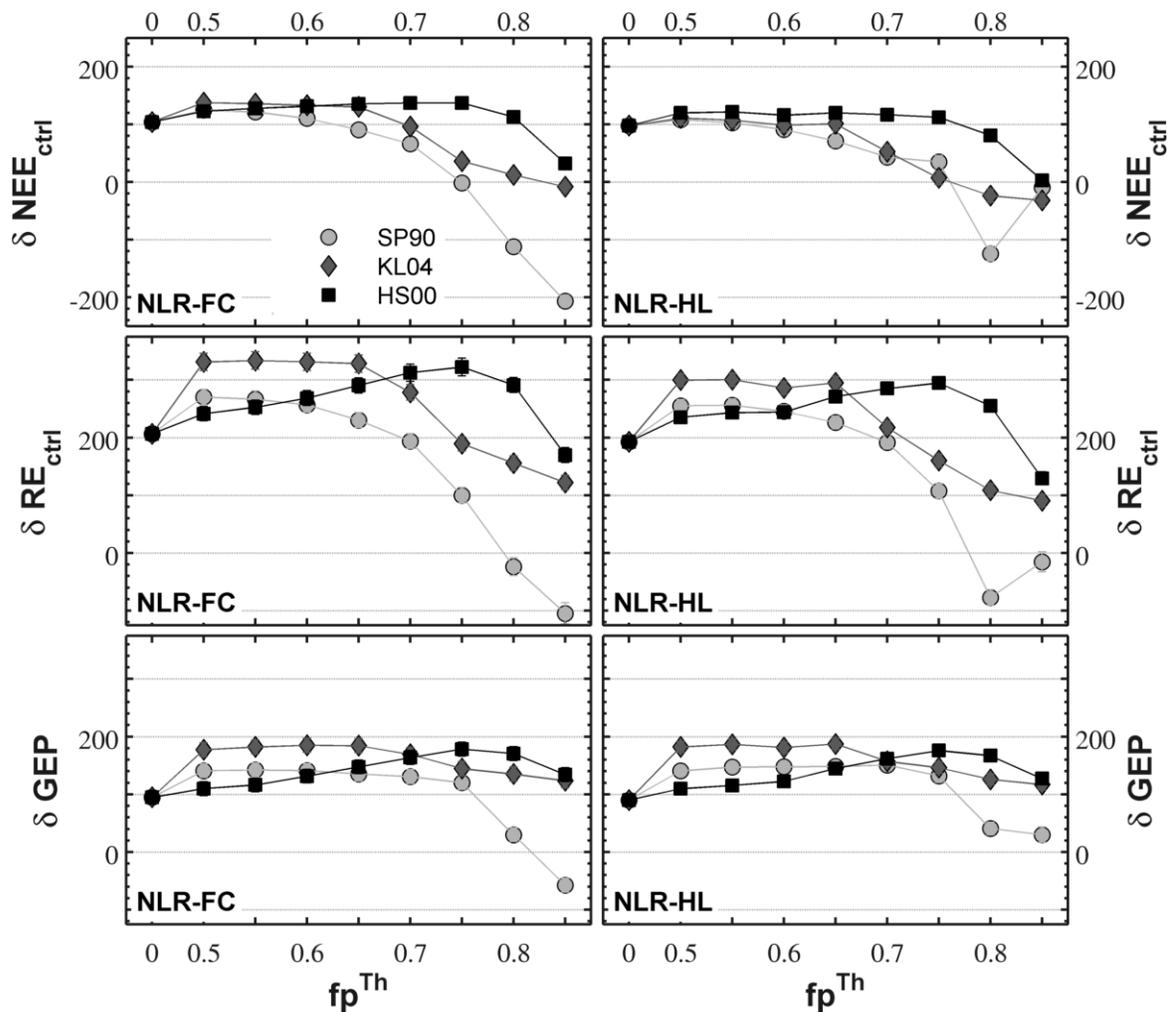
Figure 4.30: *Marginal means (symbols) and standard errors (lines) of standardized gap-filled annual carbon exchange estimates for each footprint model, as a function of fp^{Th} level. Group-wise means were calculated using post-hoc multiple comparison tests from ANOVA results. Results are separated by those for the NLR-FC gap-filling model (left panels) and for the NLR-HL gap-filling model (right panels).*

In terms of absolute values of carbon exchange, application of the (KL04, 0.8) specification resulted in annual NEE sums that ranged from a high of -249 g C m^{-2} y^{-1} (smallest sink) in 2007 to a low of -418 g C m^{-2} y^{-1} (greatest sink) in 2011 (Figure 4.31). Annual sums of RE (range: 1065 to 1249 g C m^{-2} y^{-1}) and GEP (range: 1348 to 1603 g C m^{-2} y^{-1}) showed similar interannual trends, and were distinct from those observed for NEE. Instead, trends in NEE were dictated by relative differences between GEP and RE, rather than by one of the two component fluxes. This result demonstrates that the year-to-year values of component fluxes respond differently to driving variables, whether they are biotic, phenological or meteorological in nature.

We extended these comparisons by investigating inter-annual carbon exchange effects for the preferred specification (KL04, 0.8), versus a number of other possible filtering combinations. Results showed that application of the (KL04, 0.8) footprint filter produced annual estimates for RE, GEP and NEE that were regularly lower than those predicted using other footprint models, or when no footprint was applied (Figure 4.31); differences were greatest for RE and least for GEP. On average, annual NEE estimates made using (KL04, 0.8) were 94 g C m^{-2} y^{-1} lower than with (HS00, 0.8), and 134 g C m^{-2} y^{-1} lower than when no footprint filtering was applied. Annual differences between filtering specifications were not consistent across years, as deviations between the (KL04, 0.8) specification and others generally increased with time, suggesting that the typical measurement source area may have changed with time, or the carbon exchange characteristics between the target and non-target surfaces has further diverged. The application of different u$_*$Th methods, or the use of a static value (i.e. u$_*$Th = 0.325 m s^{-1}) had minimal impact on annual estimates, suggesting that the differing u$_*$Th values estimated by these methods were more or less equivalent in effect.

Differences between footprint filtering specification were also investigated at intra-annual scales, by comparing daily-averaged and cumulative carbon exchange estimates (Figure 4.32). Results indicated that most all of the differences between model specifications were embodied by mid-summer reductions in RE, which translated to similar reductions in NEE. This difference was also observed in the annual Ts-RE logistic curves produced during parameterization of the NLR-HL model on different footprint-filtered data (Figure 4.33). RE estimates for a given Ts value were consistently lower for the (KL04, 0.8) specification, with the largest difference occurring at highest soil temperatures. In cases where footprint filtering removed more data than tolerated by the gap-filling model (Figure 4.33, panel d), response curves became irregular and inconsistent. These results suggest that the land uses surrounding our study site have a stronger respirative response to soil temperature and a higher

capacity than our site, and that incomplete or improper footprint filtering will result in misrepresentation of carbon exchange response relationships.
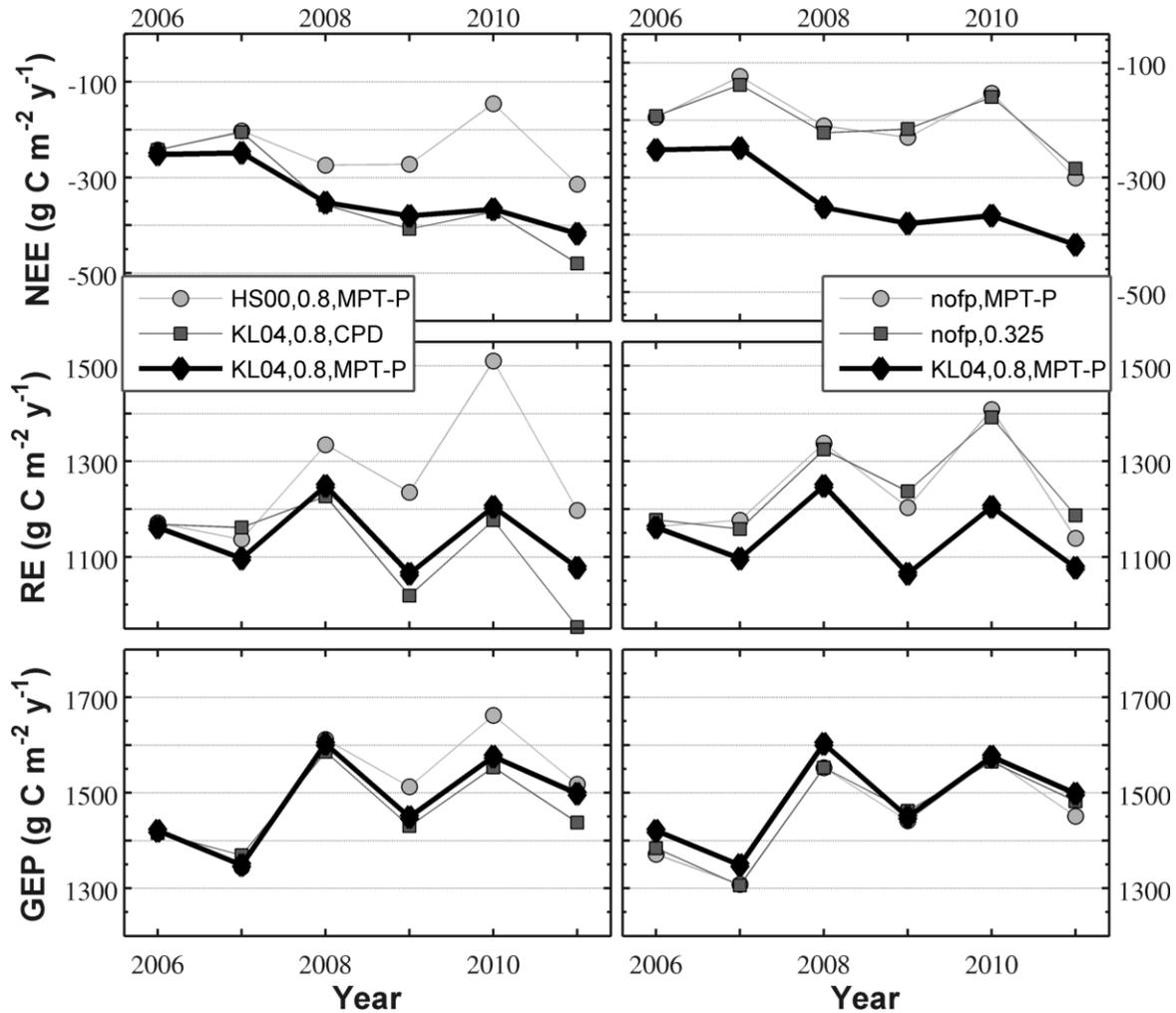
Figure 4.31: *Annual gap-filled carbon exchange estimates obtained by application of various footprint filtering and u_*^{Th} filtering approaches. The preferred specification (KL04, 0.8, MPT-P) is included in all plots (black squares and line).*
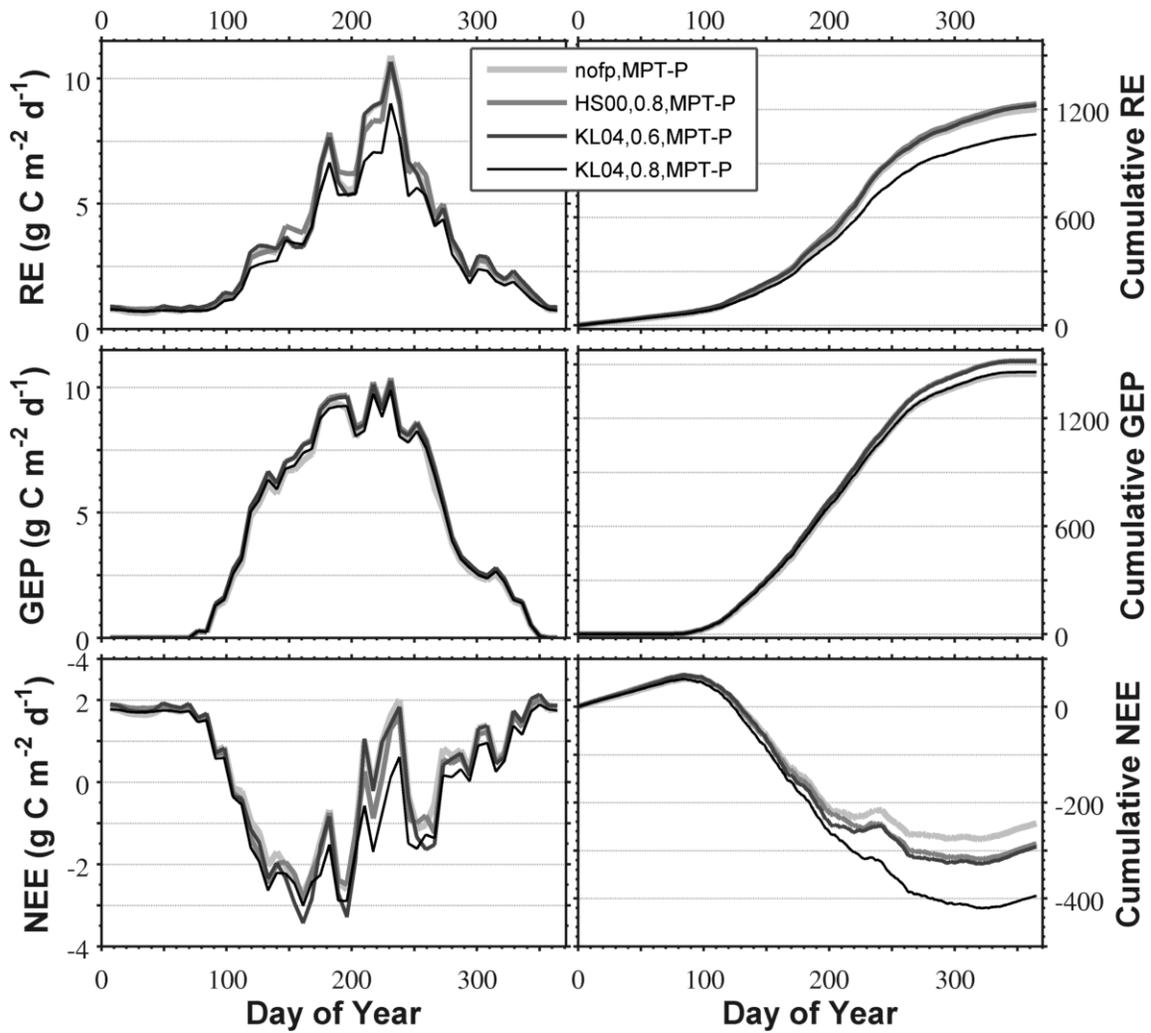
Figure 4.32: *Ensemble averages of gap-filled carbon exchange daily sums (left panels), and cumulative exchanges (right panels), obtained by application of various footprint filtering and u_*^{Th} filtering approaches.*
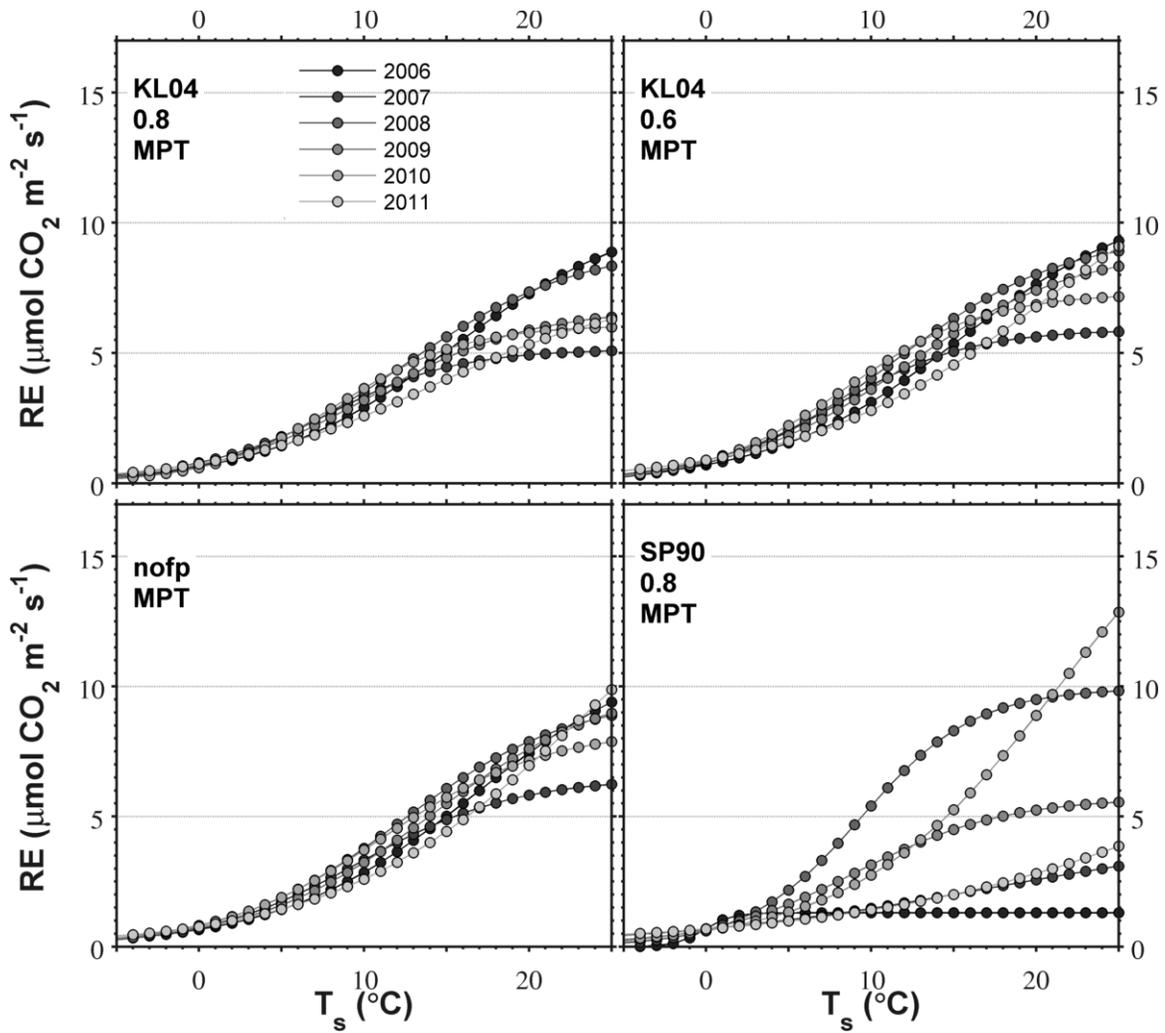
Figure 4.33: *Modeled logistic functional relationships between RE and soil temperature (T_s), as predicted for each year by the NLR-FC model for various footprint filtering and u_*^{Th} filtering approaches.*

4.5.9 Implications of footprint filtering specification on estimated carbon fluxes

In this study, we found that selection of footprint filtering parameters – both footprint model type and fp^{Th} level – had a substantial and significant effect on annual ecosystem carbon exchange estimates; data-derived environmental variable - carbon exchange relationships; and, statistical performance of gap-filling models at our fetch-limited forest site. By seeking optimized gap-filling model performance metrics, we were able to identify the Kljun et al., (2004) model with an fp^{Th} level of 0.8 as the footprint filtering parameter combination that yielded the most ecosystem-representative set of filtered EC measurement data, while maintaining a necessary amount of data to ensure gap-filling model reliability.

Through factorial experiments, we showed that applying footprint filtering at our site strongly affected the carbon exchange estimates and relationships derived from EC-measured data. These results indicate that the carbon exchange characteristics of our study forest are considerably different than the areas that surround it. This difference was greatest for RE, as a result of a greater discrepancy between target and non-target ecosystems in the exchange capacity and environmental responses of this component, as well as a tendency for more RE-parameterizable data to be contaminated by non-target surfaces.

As shown in Figure 4.1, the TP39 forest is bounded by different forests and land-use types. In the prevailing westerly wind direction, the site is bordered primarily by comparably-aged pine forests, which are expected to demonstrate similar exchange characteristics. At a distance beyond approximately 1 km from the tower, however, the landscape is dominated by cereal crop agricultural use. As a result of its intermediate distance from the tower, the agricultural area to the southwest contributes minimally to measured daytime fluxes, but is regularly incorporated into the flux footprint during nocturnal periods, when turbulence diminishes and flux footprints extend. When no footprint filter is applied to EC measurement data, signal contamination from these agricultural systems resulted in higher estimates of nighttime NEE (and thus, estimates of RE) than was experienced when stringent footprint filtering was applied. This, in turn, led to positively-biased partitioning and gap-filling models, and thus, higher estimates of RE at daily to annual timescales. This finding is supported by studies of cereal crop carbon exchange in similar climates, which have shown greater cumulative amounts and higher peak rates of respiration during the growing season than found in temperate forested ecosystems. For example, Verma et al. (2005) found peak respiration rates of $9 – 11$ μmol CO_2

m^{-2} s^{-1} in rain-fed maize and soybean fields in Nebraska, approximately double that of rates observed in temperate forests. Consistent with this, the cumulative RE reported for these fields over the course of the 150-day growing season (1154 and 826 g C m^{-2} y^{-1} for maize and soybean fields, respectively) are similar in magnitude to RE values reported for temperate coniferous forests over an entire year (Suyker, 2005). Considering the systematic source area discrepancies that occur at our site and the differing carbon exchange characteristics between these areas, it is important to identify and apply accurate footprint filtering strategies to remove the potential for bias to EC-derived NEE estimates.

When our preferred footprint parameters were applied (KL04, fp^{Th} = 0.8), annual NEE decreased between 31 and 129% (59 to 207 g C m^{-2} y^{-1}) relative to the no-footprint application. These results indicate that our target site sequesters significantly more carbon than has been previously reported (Arain and Restrepo-Coupe, 2005; Peichl et al., 2010a, 2010b). When these revised estimates are compared with biometric carbon exchange estimates made for this site (Peichl et al., 2010b), the deviations in annual RE estimates are substantially reduced in comparison to the original, non footprint-filtered EC-derived estimates. Since GEP was only minimally affected by footprint filtering application, the discrepancy reported between these methods persists for this exchange term. Consequently, as RE and GEP errors were compensatory in the original comparison, applying footprint filtering increases the differences in NEE estimates between methods, and EC-derived values indicate the forest to be a greater carbon sink by as much as 200 g C m^{-2} y^{-1}. Such a discrepancy is commonly found in comparisons of EC and biometric estimates, and may be the result of measurement errors associated with either method, or due to the fact that sampling areas are weighted differently due to EC footprint tendencies. Further investigation of these differences are required. Interestingly, updating EC-derived values in Peichl et al. (2010b) with these new estimates would sharply reduce the deviation this site showed from the other TPFS age-sequence sites in terms of annual stem volume increment versus annual NEE relationship (their Figure 2). Of equal importance to annual estimates, footprint filter application also affected shorter-interval time-integrated estimates, as well as functional relationships between environmental controls and carbon exchanges. Thus, for this site – and all others where fetch-limitations exist – the fundamental understanding of ecosystem exchange dynamics may be influenced by application of flux footprint filtering.

Of equal importance to ecosystem carbon exchange estimates at our site is the selection of a specific footprint model and a desired stringency (fp^{Th}) with which to carry out footprint filtering operations. As has been shown, footprint models differ greatly in terms of their half-hourly flux source area

estimation, which in a fetch-limited forest leads to large discrepancies in the amount and type of data removed for a given fp^{Th} level. As a result, the variation between disparate footprint models and fp^{Th} levels were commonly of similar magnitude to the effect of applying footprint filtering in the first place. This has important implications not just for our study and other similar studies that deal with fetch limitations, but for all investigations where analytical footprint models are used to quantify the effects of spatial heterogeneity on EC-measured fluxes. For example, results of footprint climatology studies that use analytical models to inform spatial flux partitioning operations (e.g. Chen et al., 2012, 2009; Foken and Leclerc, 2004; Göckede et al., 2004; Rebmann et al., 2004) are equally vulnerable to discrepancies in source area estimation. Therefore, the results of this study suggest that investigations using analytical footprint models to characterize EC flux measurement source areas consider means of evaluating footprint model performance (e.g. via tracer experiments, Lagrangian or Large Eddy Simulation comparisons or through data-driven exercises), or at least consider their results across a number of footprint models.

In this study, we used a novel approach to evaluate the filtering quality of footprint models; gap-filling model goodness of fit was used to indicate the consistency of footprint filter-passing data. This approach is based on the assumption that carbon exchange measurements will be more consistent when fluxes originate mostly (or completely) from within the surface of interest, than when source area contribution is spread among numerous diverse surfaces outside of the forest bounds. This assumption is likely to be valid in the case of our study, given that the target forest exhibits consistency in many of the traits used to characterize degree of homogeneity, in that it is a planted and managed monoculture forest growing on nearly flat, extensive sandy soils, with co-dominant trees that are of the same (or very similar) age, density and management history (Peichl and Arain, 2006). Furthermore, the forest of interest is bounded by forests of considerably different ages or species, or different land uses altogether. Though our study site may represent a best-case end-member for assessing this approach, this method may prove useful for studies in similar settings, where relatively homogeneous surfaces are bounded by distinctly dissimilar ones. Though our footprint model evaluation approach is new, it shares similarities with methods that evaluate EC flux quality and source area consistency by use of integral turbulence characteristics (Göckede et al., 2004; Thomas and Foken, 2002). Between these methods, a characteristic of EC-measured data is compared to modeled values to infer the degree of surface similarities (in roughness or energy and scalar exchanges). In fact, when integral turbulence characteristics were used to assess the footprint models in our study (Figure 4.34), results showed the

KL04 model retained data with markedly lower deviations than the other models, thus further supporting our selection of the KL04 model.
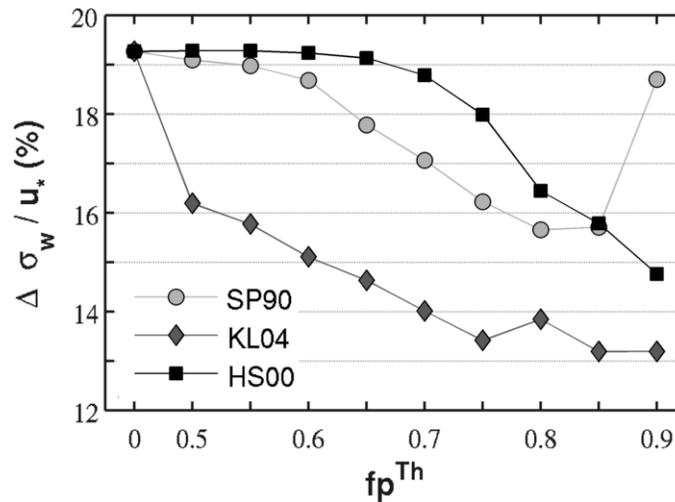
Figure 4.34: Median difference (in percentage) between measured and predicted integral turbulence characteristics for data grouped according to footprint model and fp^{Th} level.

In a fetch-limited study site such as ours, obtaining "ecosystem-representative" time-integrated carbon exchange estimates is a challenging task. Successful filtering approaches must constrain measurements to achieve maximum representativeness of the target forest fluxes, while simultaneously ensuring that enough data is retained to maintain gap-filling model consistency and reliability. To accomplish this task, we relied on indicators associated with gap-filling model output variability and goodness-of-fit to evaluate both constraints. Our resulting "best" solution – using the KL04 model with a fp^{Th} level of 0.8 – therefore, represents a compromise between these competing requirements. For any site using this approach, the optimal solution will depend on a number of factors, including: the robustness of the selected gap-filling model to measurement gaps; the complexity of the surface of interest and its carbon exchange dynamics; and, the nature of gaps that are induced in flux data by filtering operations. As a result, our gap-filled annual sums should not necessarily be viewed as the definitive ecosystem values, but rather as further-constrained estimates for our site. Without the input of multiple tower installations or artificial tracer experiments, future work would benefit from the operation of Lagrangian or large eddy simulations, to further understanding of analytical footprint model performance, and the true extents of half-hourly flux source areas.

4.6 Conclusions and outlook

In summary, our work revealed marked variation among, and important interactions between, footprint and friction velocity filtering methods at our fetch-limited forest site. Furthermore, our results show that the selection of footprint model, footprint threshold (fp^{Th}) and u_*^{Th} method have implications for annual gap-filled carbon exchange estimates. The key findings of our study can be described as follows:

- The analytical footprint models evaluated in this study vary substantially in their half-hourly flux source area estimates, as indicated by the differences in filter-passing ϕ among models. Discrepancies existed in terms of time of day, atmospheric stratification conditions, and the fp^{Th} level applied to filtering. The KL04 and HS00 models generally removed less data than KM01 and SP90, and were the most similar of those tested. When operated at our site, the KM01 and SP90 models tended to remove such a large proportion of flux data that parameterization of u_*^{Th} methods and gap-filling models commonly failed.

- Results among the tested u_*^{Th} estimation methods varied in two aspects. The first was due to the standard selection statistic (mean, median, maximum value) used to estimate a single u_*^{Th} value from a number of substrata estimates. Our results suggest that comparisons between these methods should be made using a single, consistent selection statistic. Specifically for our site, the use of the mean value selection statistic with the CPD method is invalid due to non-normality in substrata u_*^{Th} estimates. Future applications of this method to other sites should be considerate to this potential biasing characteristic, and take appropriate action to address it, where required.

- u_*^{Th} estimates from both the MPT-P and CPD methods were sensitive to changes in both the selected footprint model and fp^{Th} level. The underlying mechanism for this relationship was found to be differences in the u_* value distribution of footprint filter-passing data that was used to parameterize u_*^{Th} estimation methods. The CPD method was considerably more sensitive to input u_* values, and hence, showed a greater variability to changes in footprint model and fp^{Th}. This result exposes a potential weakness and bias source in both u_*^{Th} estimation methods, and its wider existence and causes should be explored further.

- Annual ecosystem carbon exchange estimates are sensitive to both the selected fp^{Th} level for a given footprint model, and the estimated u_*^{Th} value. Increases in estimated u_*^{Th} has a large,

215

positive effect on annual RE estimates, particularly for u_*^{Th} estimates below 0.35 m s^{-1}; above this value, estimates are relatively consistent. Conversely, high footprint model stringency (fp^{Th}) leads to relative decreases in RE estimates. Since GEP is less sensitive to both parameters, annual NEE estimates are decreased (more carbon sequestration) when footprint filtering is strictest. This result suggests that our study forest has carbon exchange dynamics that are significantly different than the surrounding area, and that implementing accurate footprint filtering models at our sites is critical to achieving ecosystem-representative carbon exchange estimates. We recommend further analysis of this effect at our site, in order to better contextualize the control of flux filtering on carbon exchange estimates, and to evaluate the performance of these differing analytical models.

- The footprint filtering specifications (footprint model and fp^{Th} level) that were applied to EC-measured data at our fetch-limited forest had a substantial, significant effect on carbon exchange estimates and data-derived dynamic relationships. Factorial experiments showed that the application of footprint filtering caused universal reductions in annual ecosystem NEE estimates (increased sequestration) that were between 50 and 130% of the magnitude of original, non-footprint filtered estimates. These findings demonstrate the sensitivity of data-derived carbon exchange information at our site to the specification of the flux footprint source area, and highlight the differences among the analytical footprint models that were tested. Using gap-filling model performance metrics to evaluate ecosystem-representativeness of different footprint specifications, we identified the footprint model developed by Kljun et al. (2004) as the most appropriate for our site. The optimal stringency of this footprint model was achieved at $fp^{Th} = 0.8$, which reflects a compromise between the need for extracting maximally-representative measurements, while retaining enough data points to ensure gap-filling model operational reliability. The footprint filter evaluation method introduced in this study is a novel and effective data-driven approach to assessing flux source area estimates at our site, and should be considered for other sites where a relatively similar 'target' surface is fetch-limited and bordered by distinctly different ecosystem types.

In conclusion, our results identify fundamental differences among methods of footprint and friction velocity filtering that are commonly used for EC measurement studies at forested sites. These differences have important consequences for carbon exchange estimates at our fetch-limited site, and the strength of our results suggest that these effects be considered at other sites with similar EC flux

filtering needs. As demonstrated in this study, the differences between analytical footprint models, and the marked consequences of their selection, suggest that further work should be undertaken to compare these models with independent – or at least, additional – information about flux footprint source areas, such as: Lagrangian and large eddy simulations, or natural or artificial tracer experiments.

4.7 References

Arain, M., Restrepo-Coupe, N., 2005. Net ecosystem production in a temperate pine plantation in southeastern Canada. Agric. For. Meteorol. 128, 223–241.

Aubinet, M., 2008. Eddy covariance CO2 flux measurements in nocturnal conditions: An analysis of the problem. Ecol. Appl. 18, 1368–1378.

Aubinet, M., Grelle, A., Ibrom, A., Rannik, Ü., Moncrieff, J., Foken, T., Kowalski, A.S., Martin, P.H., Berbigier, P., Bernhofer, C., Clement, R., Elbers, J., Granier, A., Grünwald, T., Morgenstern, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., Vesala, T., 2000. Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology. Adv. Ecol. Res. 30, 113–175.

Aubinet, M., Heinesch, B., Yernaux, M., 2003. Horizontal and vertical CO2 advection in a sloping forest. Boundary-Layer Meteorol. 108, 397–417.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82, 2415–2434.

Baldocchi, D.D., 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Glob. Chang. Biol. 9, 479–492.

Barford, C.C., Wofsy, S.C., Goulden, M.L., Munger, J.W., Pyle, E.H., Urbanski, S.P., Hutyra, L., Saleska, S.R., Fitzjarrald, D., Moore, K., 2001. Factors controlling long- and short-term sequestration of atmospheric CO2 in a mid-latitude forest. Science 294, 1688–91.

Barr, A.G., Black, T. A., Hogg, E.H., Kljun, N., Morgenstern, K., Nesic, Z., 2004. Inter-annual variability in the leaf area index of a boreal aspen-hazelnut forest in relation to net ecosystem production. Agric. For. Meteorol. 126, 237–255.

Barr, A.G., Richardson, A.D., Hollinger, D.Y., Papale, D., Arain, M.A., Black, T.A., Bohrer, G., Dragoni, D., Fischer, M.L., Gu, L., Law, B.E., Margolis, H.A., McCaughey, J.H., Munger, J.W., Oechel, W.C., Schaefer, K., 2013. Use of change-point detection for friction–velocity threshold

evaluation in eddy-covariance studies. Agric. For. Meteorol. 171-172, 31–45.

Black, T. A., Hartog, G., Neumann, H.H., Blanken, P.D., Yang, P.C., Russell, C., Nesic, Z., Lee, X., Chen, S.G., Staebler, R., Novak, M.D., 1996. Annual cycles of water vapour and carbon dioxide fluxes in and above a boreal aspen forest. Glob. Chang. Biol. 2, 219–229.

Calder, K.L., 1952. Some recent british work on the problem of diffusion in the lower atmosphere, in: U.S. Tech. Conf. Air Poll. McGraw-Hill, New York, NY, pp. 787–792.

Chen, B., Black, T.A., Coops, N.C., Hilker, T., Trofymow, J.A.T., Morgenstern, K., 2009. Assessing tower flux footprint climatology and scaling between remotely sensed and eddy covariance measurements. Boundary-Layer Meteorol. 130, 137–167.

Chen, B., Coops, N.C., Fu, D., Margolis, H. A., Amiro, B.D., Barr, A.G., Black, T.A., Arain, M.A., Bourque, C.P. -A., Flanagan, L.B., Lafleur, P.M., McCaughey, J.H., Wofsy, S.C., 2011. Assessing eddy-covariance flux tower location bias across the Fluxnet-Canada Research Network based on remote sensing and footprint modelling. Agric. For. Meteorol. 151, 87–100.

Chen, B., Coops, N.C., Fu, D., Margolis, H. A., Amiro, B.D., Black, T.A., Arain, M.A., Barr, A.G., Bourque, C.P. -A., Flanagan, L.B., Lafleur, P.M., McCaughey, J.H., Wofsy, S.C., 2012. Characterizing spatial representativeness of flux tower eddy-covariance measurements across the Canadian Carbon Program Network using remote sensing and footprint analysis. Remote Sens. Environ. 124, 742–755.

Chen, J.M., Govind, A., Sonnentag, O., Zhang, Y., Barr, A., Amiro, B., 2006. Leaf area index measurements at Fluxnet-Canada forest sites. Agric. For. Meteorol. 140, 257–268.

Davenport, E.C., El-Sanhurry, N. A., 1991. Phi/phimax: review and synthesis. Educ. Psychol. Meas. 51, 821–828.

Dyer, A., 1974. A review of flux-profile relationships. Boundary-Layer Meteorol. 7, 363–372.

Dyer, A.J., Hicks, B.B., 1974. Flux-gradient relationships in the constant flux layer. Q. J. R. Meteorol. Soc. 96, 715–721.

Feigenwinter, C., Bernhofer, C., Eichelmann, U., Heinesch, B., Hertel, M., Janous, D., Kolle, O., Lagergren, F., Lindroth, A., Minerbi, S., Moderow, U., Montagnani, L., Queck, R., Rebmann, C., Vestin, P., Yernaux, M., Zeri, M., Ziegler, W., Aubinet, M., 2008. Comparison of horizontal and vertical advective CO_2 fluxes at three forest sites. Agric. For. Meteorol. 148, 12–24.

Finnigan, J., 2006. The storage term in eddy flux calculations. Agric. For. Meteorol. 136, 108–113.

Fluxnet-Canada, 2003. Fluxnet-Canada measurement protocols: Working draft, version 1.3, Flux. Laval, QC.

Foken, T., Leclerc, M.Y., 2004. Methods and limitations in validation of footprint models. Agric. For. Meteorol. 127, 223–234.

Foken, T., Wichura, B., 1996. Tools for quality assessment of surface-based flux measurements. Agric. For. Meteorol. 78, 83–105.

Gash, J., 1986. A note on estimating the effect of a limited fetch on micrometeorological evaporation measurements. Boundary-Layer Meteorol. 35, 409–413.

Göckede, M., Rebmann, C., Foken, T., 2004. A combination of quality assessment tools for eddy covariance measurements with footprint modelling for the characterisation of complex sites. Agric. For. Meteorol. 127, 175–188.

Goulden, M.L., Munger, J.W., Fan, S.-M., Daube, B.C., Wofsy, S.C., 1996. Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. Glob. Chang. Biol. 2, 169–182.

Griffis, T., 2003. Ecophysiological controls on the carbon balances of three southern boreal forests. Agric. For. Meteorol. 117, 53–71.

Gu, L., Falge, E., Boden, T., Baldocchi, D.D., Black, T.A., Saleska, S.R., Suni, T., Verma, S.B., Vesala, T., Wofsy, S.C., Xu, L., 2005. Objective threshold determination for nighttime eddy flux filtering. Agric. For. … 128, 179–197.

Guilford, J., 1965. The minimal PHI coefficient and the maximal PHI. Educ. Psychol. Meas. 25.

Hollinger, D.Y., Richardson, A.D., 2005. Uncertainty in eddy covariance measurements and its application to physiological models. Tree Physiol. 25, 873–85.

Horst, T., Weil, J., 1994. How far is far enough?: The fetch requirements for micrometeorological measurement of surface fluxes. J. Atmos. Ocean. Technol. 11, 1018–1025.

Horst, T., Weil, J.C., 1992. Footprint estimation for scalar flux measurements in the atmospheric surface layer. Boundary-Layer Meteorol. 59, 279–296.

Horst, T.W., 1999. The footprint for estimation of atmosphere-surface exchange fluxes by profile

techniques 171–188.

Hsieh, C., Katul, G., Chi, T., 2000. An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows. Adv. Water Resour. 23, 765–772.

Kljun, N., Calanca, P., Rotach, M.W., Schmid, H.P., 2004. A simple paramerization for flux footprint predictions. Boundary-Layer Meteorol. 112, 503–523.

Kljun, N., Kormann, R., Rotach, M.W., Meixner, F., 2003. Comparison of the langrangian footprint model LPDM-B with an analytical footprint model. Boundary-Layer Meteorol. 106, 349–355.

Kljun, N., Rotach, M., Schmid, H., 2002. A three-dimensional backward Lagrangian footprint model for a wide range of boundary-layer stratifications. Boundary-Layer Meteorol. 103, 205–226.

Kormann, R., Meixner, F.X., 2001. An analytical footprint model for non-neutral stratification. Boundary-Layer Meteorol. 99, 207–224.

Lasslop, G., Reichstein, M., Kattge, J., Papale, D., 2008. Influences of observation errors in eddy flux data on inverse model parameter estimation. Biogeosciences 5, 1311–1324.

Leclerc, M.., Meskhidze, N., Finn, D., 2003. Comparison between measured tracer fluxes and footprint model predictions over a homogeneous canopy of intermediate roughness. Agric. For. Meteorol. 117, 145–158.

Leclerc, M.Y., Shen, S., Lamb, B., 1997. Observations and large-eddy simulation modeling of footprints in the lower convective boundary layer. J. Geophys. Res. Atmos. 102, 9323–9334.

Leclerc, M.Y., Thurtell, G.W., 1990. Footprint prediction of scalar fluxes using a Markovian analysis. Boundary-Layer Meteorol. 52, 247–258.

Lee, X., Finnigan, J., U, K.T.P., 2004. Coordinate systems and flux bias error, in: Lee, X., Massman, W.J., Law, B. (Eds.), Handbook of Micrometeorology. Kluwer Academic Publishers, Dordrecht, Germany, p. 250.

Lloyd, J., Taylor, J., 1994. On the temperature dependence of soil respiration. Funct. Ecol. 8, 315–323.

Lund, R., Reeves, J., 2002. Detection of undocumented changepoints: A revision of the two-phase regression model. J. Clim. 15, 2547–2554.

Massman, W.J., Lee, X., 2002. Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges. Agric. For. Meteorol. 113, 121–144.

Moffat, A., Papale, D., Reichstein, M., Hollinger, D., Richardson, A., Barr, A., Beckstein, C., Braswell, B., Churkina, G., Desai, A., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. 147, 209–232.

Moncrieff, J., Malhi, Y., Leuning, R., 1996. The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water. Glob. Chang. Biol. 2, 231–240.

Neftel, A., Spirig, C., Ammann, C., 2008. Application and test of a simple tool for operational footprint evaluations. Environ. Pollut. 152, 644–52.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., Yakir, D., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3, 571–583.

Paw U, K.T., Baldocchi, D.D., Meyers, T.P., Wilson, K.B., 2000. Correction of eddy-covariance measurements incorporating both advective effects and density fluxes. Boundary-Layer Meteorol. 97, 487–511.

Peichl, M., Arain, M.A., 2006. Above- and belowground ecosystem biomass and carbon pools in an age-sequence of temperate pine plantation forests. Agric. For. Meteorol. 140, 51–63.

Peichl, M., Arain, M.A., Brodeur, J.J., 2010a. Age effects on carbon fluxes in temperate pine forests. Agric. For. Meteorol. 150, 1090–1101.

Peichl, M., Brodeur, J.J., Khomik, M., Arain, M.A., 2010b. Biometric and eddy-covariance based estimates of carbon fluxes in an age-sequence of temperate pine forests. Agric. For. Meteorol. 150, 952–965.

Presant, E.W., Acton, C.J., 1984. The Soils of the Regional Municipality of Haldimand-Norfolk. Volume 1. Report No. 57. Guelph, Ontario.

Rebmann, C., Göckede, M., Foken, T., Aubinet, M., Aurela, M., Berbigier, P., Bernhofer, C., Buchmann, N., Carrara, A., Cescatti, A., Ceulemans, R., Clement, R., Elbers, J. A., Granier, A., Grünwald, T., Guyon, D., Havránková, K., Heinesch, B., Knohl, A., Laurila, T., Longdoz, B., Marcolla, B., Markkanen, T., Miglietta, F., Moncrieff, J., Montagnani, L., Moors, E., Nardino, M., Ourcival, J.-M., Rambal, S., Rannik, Ü., Rotenberg, E., Sedlak, P., Unterhuber, G., Vesala, T., Yakir, D., 2004. Quality analysis applied on eddy covariance measurements at complex forest sites

using footprint modelling. Theor. Appl. Climatol. 80, 121–141.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., Valentini, R., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Chang. Biol. 11, 1424–1439.

Richardson, A., Mahecha, M., Falge, E., Kattge, J., Moffat, A., Papale, D., Reichstein, M., Stauch, V., Braswell, B., Churkina, G., 2008. Statistical properties of random CO2 flux measurement uncertainty inferred from model residuals. Agric. For. Meteorol. 148, 38–50.

Richardson, A.D., Hollinger, D.Y., 2007. A method to estimate the additional uncertainty in gap-filled NEE resulting from long gaps in the CO2 flux record. Agric. For. Meteorol. 147, 199–208.

Richardson, A.D., Hollinger, D.Y., Aber, J.D., Ollinger, S. V., Braswell, B.H., 2007. Environmental variation is directly responsible for short- but not long-term variation in forest-atmosphere carbon exchange. Glob. Chang. Biol. 13, 788–803.

Richardson, A.D., Hollinger, D.Y., Burba, G.G., Davis, K.J., Flanagan, L.B., Katul, G.G., William Munger, J., Ricciuto, D.M., Stoy, P.C., Suyker, A.E., Verma, S.B., Wofsy, S.C., 2006. A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes. Agric. For. Meteorol. 136, 1–18.

Rogie, J., Kerrick, D., Sorey, M., Chiodini, G., Galloway, D., 2001. Dynamics of carbon dioxide emission at Mammoth Mountain, California. Earth Planet. Sci. Lett. 188, 535–541.

Schmid, H.P., 1994. Source areas for scalars and scalar fluxes. Boundary-Layer Meteorol. 67, 293–318.

Schuepp, P.H.P., Leclerc, M.Y.M., Macpherson, J.I., Desjardins, R.L., 1990. Footprint prediction of scalar fluxes from analytical solutions of the diffusion equation. Boundary-Layer Meteorol. 50, 355–373.

Solow, A., 1987. Testing for climate change: An application of the two-phase regression model. J. Clim. Appl. Meteorol. 26, 1401–1405.

Staebler, R., Fitzjarrald, D., 2004. Observing subcanopy CO2 advection. Agric. For. Meteorol. 122,

139–156.

Stull, R.B., 1988. Similarity theory. Springer.

Thomas, C., Foken, T., 2002. Re-evaluation of integral turbulence characteristics and their parameterisations.

Thomson, D.J., 1987. Criteria for the selection of stochastic models of particle trajectories in turbulent flows 180.1 (1987): 529-556. J. Fluid Mech. 180, 529–556.

Van de Boer, A., Moene, A., Schüttemeyer, D., Graf, A., 2013. Sensitivity and uncertainty of analytical footprint models according to a combined natural tracer and ensemble approach. Agric. For. Meteorol. 169, 1–11.

Van Gorsel, E., Delpierre, N., Leuning, R., Black, A., Munger, J.W., Wofsy, S., Aubinet, M., Feigenwinter, C., Beringer, J., Bonal, D., Chen, B., Chen, J., Clement, R., Davis, K.J., Desai, A.R., Dragoni, D., Etzold, S., Grünwald, T., Gu, L., Heinesch, B., Hutyra, L.R., Jans, W.W.P., Kutsch, W., Law, B.E., Leclerc, M.Y., Mammarella, I., Montagnani, L., Noormets, A., Rebmann, C., Wharton, S., 2009. Estimating nocturnal ecosystem respiration from the vertical turbulent flux and change in storage of CO2. Agric. For. Meteorol. 149, 1919–1930.

Van Gorsel, E., Leuning, R., Cleugh, H., Keith, H., Kirschbaum, M.U.F., Suni, T., 2008. Application of an alternative method to derive reliable estimates of nighttime respiration from eddy covariance measurements in moderately complex topography. Agric. For. Meteorol. 148, 1174–1180.

Van Ulden, A.P., 1978. Simple estimates for vertical diffusion from sources near the ground. Atmos. Environ. 12, 2125–2129.

Wang, X., 2003. Comments on "Detection of undocumented changepoints: A revision of the two-phase regression model". J. Clim. 3383–3385.

Wyngaard, J.C., Coté, O.R., Izumi, Y., 1971. Local free convection, similarity, and the budgets of shear stress and heat flux. J. Atmos. Sci. 28, 1171–1182.

5   Conclusions

As outlined in the introduction and highlighted throughout this work, the development and widespread implementation of advanced biometeorological measurement systems – including eddy covariance (EC) and complementary environmental sensor arrays – over the past two decades has greatly improved scientific understanding of momentum, mass and energy exchanges in forested ecosystems (Baldocchi et al., 2001; Kao et al., 2012). Specific to the carbon cycle, these developments have helped researchers better characterize and quantify the magnitudes, drivers and complex dynamics of forest carbon exchanges over timescales ranging from seconds to decades (Barr et al., 2007; Falge et al., 2002; Richardson et al., 2010; Richardson, et al., 2007). Such advances have, in turn, led to improved process understanding, model construction, and outcomes for practical applications such as forest monitoring and management (e.g. McLaren et al., 2008; Richardson et al., 2013; Schwalm et al., 2010; Saunders et al., 2012).

Despite the utility of biometeorological measurement systems, obtaining valid, robust and ecosystem-representative characterizations of forest carbon cycle dynamics from the raw measured data requires a considerable amount of data management, careful processing, and interpretation of the physical phenomena represented by the data. By documenting, critically assessing and improving the processes used to achieve these goals, it is possible to improve the quality of scientific understanding obtained from measurement studies. Through the preceding chapters, this thesis has addressed important considerations for research operations at the Turkey Point Flux Station (TPFS), as well as those that are meaningful to the greater research community. Specifically, this research addressed the following challenges:

- Managing and processing large and diverse quantities of biometeorological and ecological measurement data in a graduate student-run, collaborative framework;

- Operating a roving eddy covariance system to increase site coverage at the expense of measurement temporal continuity at any given site; and,

- Performing eddy covariance measurements in fetch-limited forests, which are bordered by multiple, dissimilar surface types.

Properly managing measurement data throughout the entire data life cycle is a considerable challenge in biometeorological studies, given the quantity, diversity and complexity of data, and the need to perform a variety of complicated post-collection operations. To successfully facilitate all needs in this

context, a data workflow management system (DWMS) must be robust to data and format variability, extensible to new processes and input data, and be able to document and standardize processes (and products) for all past, present and future operations. The collaboratively-run DWMS introduced in Chapter 2 provides a comprehensive solution to these various needs, and does so in the context of a research group with minimal technical support. A conceptualization of the biometeorological data life cycle was developed, and then integrated with practical considerations for data management within the research group. Over its developmental and operational lifetime, this system has served as the data management solution for many research projects, including published papers (e.g. Mackay et al., 2012; Peichl et al., 2010a; Peichl et al., 2010b), and numerous theses. Though recent software products have been developed to facilitate specific parts of the biometeorological data processing (e.g EdiRe, EddyPro), the significance and novelty of this DWMS is that it provides a larger framework that accommodates all stages of the data life cycle. In addition to standardizing and streamlining data processes within the research group, this system also facilitates data re-visitation and replication activities, thus increasing the integrity of scientific processes and results. In this setting, the most substantial challenges are not technological, but are related to operations and adequate training; leadership within the research group is a critical component to ensure that members continually partake in, learn and teach best management practices.

Further DWMS work should focus on the increased automation of data checking and analyses operations. Improving the ability of the DWMS to actively investigate data against an established rule set, and communicate potential problems to research group members, has potential to further improve data integrity, while decreasing the human time requirements of data management. Next steps should also seek to develop and assess advanced DWMS training materials, through interactive documentation, or digital (or blended) learning modules. The expansion of this system to operate in other research groups should be viewed as an ultimate goal of DWMS development, as an impetus to increase system abstraction, modularity and interoperability.

A second, more general challenge addressed by this work was the relatively high cost associated with the purchase, maintenance and operation of EC and biometeorological equipment for long-term carbon exchange measurement studies. Expanding the carbon exchange measurement network to include a greater diversity of ecosystems, therefore, either requires funding increases, or depends on the development of unique strategies to maximize information collection with a fixed set of resources. In response to this, Chapter 3 investigated the roving eddy covariance (rEC) approach as a strategy with

potential to expand the site coverage of biometeorological research programs with limited added resource requirements. To investigate the carbon exchange estimate uncertainty associated with the rotation of a single rEC system among multiple sites, this chapter extended and built upon methods previously used to quantify net ecosystem exchange (NEE) uncertainty in response to EC data gaps (Richardson and Hollinger, 2007; Moffat et al., 2007). The act of rotating an rEC system among two or three sites in schedules of two weeks or a month was shown to negatively impact time-integrated estimates of NEE in two distinct ways: First, long data gaps masked changes in carbon cycle dynamics during these periods, and thus reduced the quality of gap-filling model parameterization; and, second, the effect of model bias was augmented by increased amounts of gaps to be filled. Gap-filling errors were largest during the active growing season, and increased substantially for rotation scenarios where measurement gaps were larger than a month, suggesting that any future rEC application would benefit from minimizing measurement gap length as much as possible during the growing season. Additionally, multiple-year measurement programs should ensure orthogonality of measurement periods between years, in order to further reduce annual NEE uncertainty. Results for our forest sites indicated that the gap-filling model of Richardson et al. (2007) is the preferred method for rEC gaps, since it prescribes relatively robust relationships between environmental variables and carbon exchange components, and it is capable of incorporating multiple years worth of data to improve model parameterization. When the best-performing rEC scenarios and gap-filling models were applied to our forested sites, the resulting noise-to-signal ratio for annual NEE ranged between 35 and 63%, translating to annual errors in the magnitude of 100 to 150 g C m^{-2} y^{-1}. Though annual NEE estimate error is notable, it is encouraging that a single rEC system could be operated in the specified context to provide reasonable NEE estimates for up to three ecosystems, while generally preserving interannual trends, and still providing detailed short-term exchange information during measurement periods. In summary, results of this study generally supported the use of rEC at our site to obtain reliable carbon exchange information, and suggested that such an application may also be suitable for other measurement sites. These findings support the further investigation of non-traditional ecosystem flux measurement programs, to better assess the trade-offs between spatial and temporal measurement distributions.

Building on the encouraging results of Chapter 3, future research – both within the research group and the wider research community – should embrace investigations of non-traditional EC and biometeorological measurement programs. This should be done by further exploring the interacting

trade-offs between the three components that determine the accuracy of gap-filled carbon estimates: the complexity of the underlying carbon exchange signal to be modeled; the amount of data available to parameterize the model; and, the performance capability of the gap-filling model itself. Exploring these dimensions has the potential to maximize the information acquisition capacity of ecosystem flux measurement programs given fixed resources.

A final challenge addressed by this research was the need to determine and implement appropriate methods to filter out EC-measured carbon exchange estimates that are unrepresentative of the ecosystem of interest. In order to accurately characterize the nature of the carbon cycle in a given forest, it is necessary that EC measurements are representative of the true carbon fluxes occurring between the atmosphere and the target ecosystem. Among the various tests available for filtering EC data, much discussion and investigation remains in regards to the application of footprint and friction velocity (u_*^{Th}) filtering operations. With current work toward standardizing these filtering approaches at the research network level, there is a need to better understand the nature of the various available filtering approaches, in terms of their scientific merit, and their impact on carbon exchange information that is obtained from EC measurements. In response, Chapter 4 investigated footprint and friction velocity filtering operations at the fetch-limited TP39 study site. This was accomplished through quantification of the differences among, and interaction between four analytical footprint models and three u_*^{Th} estimation methods, in terms of their consequences for EC data quantity and distribution, as well as subsequent implications for annual carbon exchange estimates. Novel approaches were used to quantify the relative importance of footprint and friction velocity filtering to gap-filled ecosystem carbon exchange estimates at our fetch-limited TP39 forest.

A number of deficiencies and shortcomings of commonly-used friction velocity filtering methods were exposed, and it was concluded that the u_*^{Th} determination method of Papale et al. (2006) – slightly modified to aggregate substrata estimates using its median – performed best at our site. The tendencies of the investigated u_*^{Th} methods to vary with aggregation statistic (mean, median, maximum) and input data provides a topic for further investigation and re-visitation by the larger research community. Since results showed that the selection of footprint model and stringency had significant effects on annual estimates, this study also developed and presented a data-driven, internal footprint filtering evaluation approach. Such an approach may be used to evaluate the ecosystem-representativeness of different footprint filtering treatments at fetch-limited sites where no external evaluation information (e.g. multiple-tower installations, tracer experiments) are available. Reinforcing the findings of a limited

number of previous studies (e.g. van de Boer et al., 2013), results showed substantial variation in footprint model predictions for our site – prompting further work to evaluate these analytical footprint models.

Furthermore, the choice of footprint model and stringency had considerable effects on both friction velocity threshold estimates, and annual carbon exchange estimates. Our analyses identified the footprint model of Kljun et al. (2004), applied at a footprint threshold of 0.8, provided the most consistent and reasonable estimate of EC measurement source area at our forested site. Applying this footprint filtering specification to EC measurements at our site resulted in annual NEE values that were, on average, 134 g C m^{-2} y^{-1} lower than estimated when no footprint filtering was applied. Deviations in NEE estimates between footprint specifications were almost entirely due to changes in RE, which was attributed to frequent flux footprint extension beyond the forest – and into surrounding agricultural land – during calm nighttime periods. The strong control of footprint specification on carbon exchange estimates at our site highlights the importance of assessing the differences between approaches, and evaluating them against objective criteria.

In general, the results of this chapter suggest that future work should continue to investigate friction velocity filtering as a means of identifying and removing potential periods of poorly-formed turbulence, and further compare methods that attempt to establish these thresholds. More studies are needed to compare the multitude of available analytical footprint models, to assess their suitability over a wide range of operational environments; this includes comparisons of analytical footprint outputs to Lagrangian and large eddy simulations, as well as natural and artificial tracer experiments. Furthermore, the data-driven footprint model evaluation approach that was introduced should also be further investigated using these more rigorous footprint determination methods.

5.1 References

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. Fluxnet: a new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82, 2415–2434.

Barr, A.G., Black, T.A., Hogg, E.H., Griffis, T.J., Morgenstern, K., Kljun, N., Theede, A., Nesic, Z., 2007. Climatic controls on the carbon and water balances of a boreal aspen forest. Glob. Chang. Biol. 13, 561–576.

Falge, E., Baldocchi, D., Tenhunen, J., Aubinet, M., Bakwin, P., Berbigier, P., Bernhofer, C., Burba, G., Clement, R., Davis, K.J., Elbers, J. A, Goldstein, A.H., Grelle, A., Granier, A., Guðmundsson, J., Hollinger, D., Kowalski, A.S., Katul, G., Law, B.E., Malhi, Y., Meyers, T., Monson, R.K., Munger, J.W., Oechel, W., Paw U, K.T., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Valentini, R., Wilson, K., Wofsy, S., 2002. Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. Agric. For. Meteorol. 113, 53–74.

Kao, R., Gibson, C., Gallery, R., Meier, C., Barnett, D., Docherty, K., Blevins, K., Travers, P., Azuaje, E., Springer, Y., Thibault, K., McKenzie, V., Keller, M., Alves, L., Hinckley, E., Parnell, J., Schimel, D., 2012. NEON terrestrial field observations: designing continental-scale, standardized sampling. Ecosphere 3, 1–17.

Kljun, N., Calanca, P., Rotach, M.W., Schmid, H.P., 2004. A simple paramerization for flux footprint predictions. Boundary-Layer Meteorol. 112, 503–523.

Mackay, S.L., Arain, M.A., Khomik, M., Brodeur, J.J., Sciences, E., Schumacher, J., Hartmann, H., Peichl, M., 2012. The impact of induced drought on transpiration and growth in a temperate pine plantation forest. Hydrol. Process. 26, 1779-1791.

Moffat, A., Papale, D., Reichstein, M., Hollinger, D., Richardson, A., Barr, A., Beckstein, C., Braswell, B., Churkina, G., Desai, A., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. 147, 209–232.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal,

S., Valentini, R., Vesala, T., Yakir, D., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3, 571–583.

Peichl, M., Arain, M.A., Brodeur, J.J., 2010a. Age effects on carbon fluxes in temperate pine forests. Agric. For. Meteorol. 150, 1090–1101.

Peichl, M., Brodeur, J.J., Khomik, M., Arain, M.A., 2010b. Biometric and eddy-covariance based estimates of carbon fluxes in an age-sequence of temperate pine forests. Agric. For. Meteorol. 150, 952–965.

Richardson, A.D., Black, T.A., Ciais, P., Delbart, N., Friedl, M.A., Gobron, N., Hollinger, D.Y., Kutsch, W.L., Longdoz, B., Luyssaert, S., Montagnani, L., Munger, J.W., Moors, E., Piao, S., Reichstein, M., Saigusa, N., Tomelleri, E., Vargas, R., Varlargin, A., 2010. Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philos. Trans. R. Soc. London. Ser. B Biol. Sci. 365, 3227–3246.

Richardson, A.D., Hollinger, D.Y., 2007. A method to estimate the additional uncertainty in gap-filled NEE resulting from long gaps in the CO_2 flux record. Agric. For. Meteorol. 147, 199–208.

Richardson, A.D., Hollinger, D.Y., Aber, J.D., Ollinger, S. V., Braswell, B.H., 2007. Environmental variation is directly responsible for short- but not long-term variation in forest-atmosphere carbon exchange. Glob. Chang. Biol. 13, 788–803.

Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agric. For. Meteorol. 169, 156–173.

Saunders, M., Tobin, B., Black, K., Gioria, M., Nieuwenhuis, M., Osborne, B.A., 2012. Thinning effects on the net ecosystem carbon exchange of a Sitka spruce forest are temperature-dependent. Agric. For. Meteorol. 157, 1–10.

Schwalm, C.R., Williams, C.A., Schaefer, K., Anderson, R., Arain, M.A., Baker, I., Barr, A., Black, T.A., Chen, G., Chen, J.M., Ciais, P., Davis, K.J., Desai, A., Dietze, M., Dragoni, D., Fischer, M.L., Flanagan, L.B., Grant, R., Gu, L., Hollinger, D., Izaurralde, R.C., Kucharik, C., Lafleur, P., Law, B.E., Li, L., Li, Z., Liu, S., Lokupitiya, E., Luo, Y., Ma, S., Margolis, H., Matamala, R., McCaughey, H., Monson, R.K., Oechel, W.C., Peng, C., Poulter, B., Price, D.T., Riciutto, D.M.,

Riley, W., Sahoo, A.K., Sprintsin, M., Sun, J., Tian, H., Tonitto, C., Verbeeck, H., Verma, S.B., 2010. A model - data intercomparison of CO2 exchange across North America: Results from the North American Carbon Program site synthesis. J. Geophys. Res. 115.

Suyker, A.E., Verma, S.B., Burba, G.G., Arkebauer, T.J., 2005. Gross primary production and ecosystem respiration of irrigated maize and irrigated soybean during a growing season. Agric. For. Meteorol. 131, 180–190.

Van de Boer, A., Moene, A., Schüttemeyer, D., Graf, A., 2013. Sensitivity and uncertainty of analytical footprint models according to a combined natural tracer and ensemble approach. Agric. For. Meteorol. 169, 1–11.

Verma, S.B., Dobermann, A., Cassman, K.G., Walters, D.T., Knops, J.M., Arkebauer, T.J., Suyker, A.E., Burba, G.G., Amos, B., Yang, H., Ginting, D., Hubbard, K.G., Gitelson, A.A., Walter-Shea, E.A., 2005. Annual carbon dioxide exchange in irrigated and rainfed maize-based agroecosystems. Agric. For. Meteorol. 131, 77–96.