

# **Hybrid Model for Monitoring and Optimization of Distillation Columns**

**By**

**Fahad Aljuhani**

**A Thesis**

**Submitted to the school of Graduate Studies**

**In Partial Fulfillment of the Requirements**

**For the Degree**

**Master of Applied Science**

**McMaster University**

**2015.12**

## **Abstract**

Distillation columns are primary equipment in petrochemical, gas plants and refineries. Distillation columns energy consumption is estimated to be 40% of the total plant energy consumption. Optimization of distillation columns has potential for saving large amount of energy and contributes to plant wide optimization. Currently rigorous tray to tray models are used to describe columns separation with high accuracy. Rigorous distillation models are being used as part of design, optimization and as a part of on-line real-time optimization applications. Due to large number of nonlinear equations, rigorous distillation models are not suitable for inclusion in optimization models of complex plants (e.g. refineries), since they would make the model too large. For this reason, current practice in plant-wide optimization for planning or for scheduling is to include simplified model. Accuracy of these simplified models is significantly lower than the accuracy of the rigorous models, thereby causing discrepancy between production planning and RTO decisions. This work describes reduced size hybrid model of distillation columns, suitable for use as stand-alone tool for individual column or as part of a complete plant model, either for RTO or for production planning. Hybrid models are comprised of first principles material and energy balances and empirical models describing separation in the column. Hybrid models can be used for production planning, scheduling and optimization. In addition this work describes inferential model development for estimating streams purity using real time data. Inferential model eliminates the need for Gas Chromatography GC analyzers and can be used for monitoring and control purposes. Predictions from the

models are sufficiently accurate and small size of the models enable significant reduction in size of the total plant models.

## **Acknowledgements**

All praise and thanks are for God, the One who, by His blessing and favour, good works are accomplished.

I would also like to express my sincere gratitude and appreciation to my Supervisor, Dr. Vladimir Mahalec, for his support and guidance which helped me to complete this research. His kindness, advice, comments and suggestions were very helpful throughout the entire research. Thanks to all of my colleagues for their support. Special thanks go to my colleagues Gang Fu, Omar Bageel, Pedro Castillo and Hisham Al-Dosouki.

I would like to express my appreciation to my parents for their support and endless love. I thank them for keeping me in their prayers. I dedicate this research and every success in my life to my parents. Special thanks to my wife Mashaël for being to my side all the time and for her kind and sincere support.

# Table of Contents

<b>1.</b>	<b>INTRODUCTION</b> .....	<b>1</b>
1.1	MAIN CONTRIBUTIONS .....	3
1.2	THESIS OVERVIEW.....	4
<b>2.</b>	<b>LITERATURE REVIEW</b> .....	<b>5</b>
<b>3.</b>	<b>HYBRID MODEL DEVELOPMENT</b> .....	<b>10</b>
3.1	INTRODUCTION .....	10
3.2	METHODOLOGY .....	11
3.3	TESTING SEPARATION FACTOR MODELING APPROACH.....	14
3.4	HYBRID MODEL DEVELOPMENT FOR BUTANE SPLITTER COLUMN .....	18
3.5	HYBRID MODEL DEVELOPMENT FOR DEPROPANIZER COLUMN .....	54
3.6	CASE STUDIES.....	91
<b>4.</b>	<b>INFERENCEAL MODEL DEVELOPMENT</b> .....	<b>96</b>
4.1	INTRODUCTION .....	96
4.2	METHODOLOGY .....	96
4.3	INFERENCEAL MODEL FOR BUTANE SPLITTER COLUMN .....	99
4.4	INFERENCEAL MODEL FOR DEPROPANIZER COLUMN.....	133
<b>5.</b>	<b>CONCLUSION</b> .....	<b>145</b>
<b>6.</b>	<b>REFERENCES</b> .....	<b>148</b>

## List of Illustrations

Figure 1. Hybrid model structure for distillation columns	2
Figure 2. Relationship between top impurity $X_d$ and reflux ratio for variable feed composition & $X_b$	12
Figure 3. Relationship between bottom impurity $X_b$ and $Q_r/B$ for variable feed composition & $X_d$	13
Figure 4. $SF_{\text{calculated}} / SF_{\text{actual}}$ for constant feed composition and variable $X_d$ & $X_b$	16
Figure 5. Relationship between $SF_{\text{calculated}} / SF_{\text{actual}}$ slope and $X_b$ for constant feed composition	17
Figure 6. Relationship between $SF_{\text{calculated}} / SF_{\text{actual}}$ intercept and $X_b$ for constant feed composition	17
Figure 7. $SF_{\text{calculated}} / SF_{\text{actual}}$ for variable feed composition, $X_d$ & $X_b$	18
Figure 8. Butane C4 splitter column configuration and base case operating parameters	19
Figure 9. Relationship between $X_d$ and reflux ratio for constant feed composition and variable $X_b$	23
Figure 10. Relationship between $X_d$ correlation slope (a) and $X_b$ at constant feed composition	24
Figure 11. Relationship between $X_d$ correlation intercept b and $X_b$ at constant feed composition	25
Figure 12. Relative error% for $X_d$ with slope a and intercept b function of bottom impurity $X_b$ at constant feed composition	26
Figure 13. Relationship between $X_b$ and $Q_r/B$ for constant feed composition and variable $X_d$	27
Figure 14. Relationship between $X_b$ and $Q_r/B$ with single linear regression line for constant feed composition and variable $X_d$	28
Figure 15. Relative error% for $X_b$ equation with single linear regression line at constant feed composition and variable $X_d$	29
Figure 16. Relationship between $X_b$ and $Q_r/B$ with single quadratic regression line for constant feed composition and variable $X_d$	29
Figure 17. Relative error% for $X_b$ equation with single quadratic regression line at constant feed composition and variable $X_d$	30
Figure 18. $X_b$ equation with single quadratic line for variable feed compositions and $X_d$	33
Figure 19. Relationship between C1 coefficient and feed LK% for variable feed composition & $X_d$	34
Figure 20. Relationship between C2 coefficient and feed LK% for variable feed composition & $X_d$	34
Figure 21. Relationship between C3 coefficient and feed LK% for variable feed composition & $X_d$	35
Figure 22. Relationship between $X_d$ and reflux ratio for constant feed composition and variable $X_b$	37
Figure 23. Relationship between $\ln(X_d)$ and reflux ratio for constant feed composition and variable $X_b$	38
Figure 24. Relationship between $X_b$ and $Q_r/B$ for Constant feed composition and variable $X_d$	40
Figure 25. Relationship between $\ln(X_b)$ and $Q_r/B$ for Constant feed composition and variable $X_d$	41
Figure 26. Relationship between slope a and $X_d$ at constant feed composition	42
Figure 27. Relationship between intercept b and $X_d$ at constant feed composition	42
Figure 28. Relationship between $\ln(X_b)$ and $Q_r/B$ for variable feed composition and $X_d$	45
Figure 29. Relationship between d1 and feed LK for variable feed composition & $X_d$	46
Figure 30. Relationship between d2 and feed LK for variable feed composition & $X_d$	47
Figure 31. Relationship between d3 and feed LK for variable feed composition & $X_d$	47
Figure 32. Relationship between d4 and feed LK for variable feed composition & $X_d$	48
Figure 33. Relationship between d5 and feed LK for variable feed composition & $X_d$	48
Figure 34. Relationship between d6 and feed LK for variable feed composition & $X_d$	49
Figure 35. Relative error% compared to bottom stream purity using both low and high range purity models for variable feed composition, $X_d$ & $X_b$	52
Figure 36. Depropanizer column configuration and base case operating parameters	55
Figure 37. Relationship between $X_d$ and reflux ratio at constant feed and variable $X_b$	59
Figure 38. Relationship between slope a and $X_b$ values for constant feed composition	60
Figure 39. Relationship between intercept b and $X_b$ values for constant feed composition	60
Figure 40. Relationship between $X_d$ and reflux ratio at variable feed composition and $X_b$	62
Figure 41. C1 is a quadratic function of HK for variable feed composition and $X_b$	63
Figure 42. C2 is a quadratic function of HK for variable feed composition and $X_b$	64
Figure 43. C3 is a quadratic function of HK for variable feed composition and $X_b$	64

Figure 44. C4 is a quadratic function of HK for variable feed composition and Xb	65
Figure 45. C5 is a quadratic function of HK for variable feed composition and Xb	65
Figure 46. C6 is a quadratic function of HK for variable feed composition and Xb	66
Figure 47. Relative error% for Xd complete model testing variable feed composition and Xb	68
Figure 48. Relationship between Xb and Qr/B at constant feed and variable Xd with quadratic fitting	69
Figure 49. Relationship between ln(Xb) and Qr/B at constant feed composition and variable Xd	71
Figure 50. Relationship between slope a and Xd for constant feed composition	72
Figure 51. Relationship between intercept b and Xd for constant feed composition	72
Figure 52. Relationship between ln(Xb) and Qr/B at variable feed composition and Xd	74
Figure 53. ln(Xb) slope function of LK for variable feed composition & Xd	75
Figure 54. d1 coefficient function of LK for variable feed composition & Xd	75
Figure 55. d2 coefficient function of LK for variable feed composition & Xd	76
Figure 56. d3 coefficient function of LK for variable feed composition & Xd	76
Figure 57. Relative error% for Xb model for variable feed composition and Xd	79
Figure 58. Relationship between Vd/F and condenser duty Qc at constant feed and variable Xd & Xb	80
Figure 59. Relationship between Vd/F and Xd at constant feed composition (Feed 3) and variable Xb	81
Figure 60. Relationship between Vd/F and Xd at variable feed composition and Xb	82
Figure 61. Relationship between C1 coefficient and feed LK	83
Figure 62. Relationship between C2 coefficient and feed LK	84
Figure 63. Relationship between C3 coefficient and feed LK	84
Figure 64. Relationship between $x_{C2}, V_d$ and Xd	87
Figure 65. Relationship between $x_{C3}, V_d$ and Xd	88
Figure 66. Relationship between $x_{C4}, V_d$ and Xd	88
Figure 67. Relationship between tray number and $\Delta T / \Delta x_d$ in the top section for constant feed composition and variable Xb (C4 splitter low purity range)	101
Figure 68. Relationship between Xd and trays temperature for constant feed composition and variable Xb (C4 splitter low purity range)	101
Figure 69. Xd equation as function of $T_{10}$ for variable feed composition and Xb (C4 splitter low purity range)	103
Figure 70. ProMV PLS model for Xd as function of two variables plot and statistical analysis for variable feed composition, Xd & Xb (C4 splitter low purity range)	104
Figure 71. ProMV PLS model for Xd as function of four variables plot and statistical analysis for variable feed composition, Xd & Xb (C4 splitter low purity range)	105
Figure 72. Relationship between Xd equation slope and feed HK% (C4 splitter low purity range)	107
Figure 73. Relationship between Xd equation intercept and feed HK% (C4 splitter low purity range)	108
Figure 74. Relationship between tray number and $\Delta T / \Delta x_b$ in the bottom section for constant feed composition & variable Xd (C4 splitter low purity range)	110
Figure 75. Relationship between Xd and $T_{88}$ for constant feed composition and variable Xd (C4 splitter low purity range)	110
Figure 76. Xb equation as function of $T_{88}$ for variable feed composition & Xd (C4 splitter low purity range)	111
Figure 77. ProMV PLS model for Xb as function of two variables plot and statistical analysis (C4 splitter low purity range)	112
Figure 78. Relationship between Xb equation slope a and feed LK composition (C4 splitter low purity range)	114
Figure 79. Relationship between Xb equation intercept b and feed LK composition (C4 splitter low purity range)	114
Figure 80. Relationship between tray number and $\Delta T / \Delta x_d$ in the top section for constant feed composition (C4 splitter high purity range)	116
Figure 81. Relationship between Xd and $T_{24}$ for constant feed composition and variable Xb (C4 splitter high purity range)	117

Figure 82. Relationship between $X_d$ and $T_{10}$ , $T_{12}$ & $T_{15}$ for constant feed composition and variable $X_b$ (C4 splitter high purity range)	118
Figure 83. Relationship between $X_d$ and $T_{10}$ for variable feed composition and $X_b$ (C4 splitter high purity range)	119
Figure 84. ProMV PLS model for $X_d$ as function of two variables plot and statistical analysis for variable feed composition, $X_d$ & $X_b$ (C4 splitter high purity range)	120
Figure 85. ProMV PLS model for $X_d$ as function of four variables plot and statistical analysis for variable feed composition, $X_d$ & $X_b$ (C4 splitter high purity range)	121
Figure 86. Relationship between $X_d$ equation slope and feed HK (C4 splitter high purity range)	123
Figure 87. Relationship between $X_d$ equation intercept and feed HK (C4 splitter high purity range)	123
Figure 88. Relationship between tray number and $\Delta T / \Delta x_b$ in the bottom section for constant feed composition and variable $X_d$ (C4 splitter high purity range)	125
Figure 89. Relationship between $X_b$ and selected trays in the bottom section for constant feed composition and variable $X_d$ (C4 splitter high purity range)	126
Figure 90. Relationship between $X_b$ and $T_{88}$ for variable feed composition & $X_d$ (C4 splitter high purity range)	128
Figure 91. ProMV PLS $X_b$ model as function of two variables plot and statistical analysis for variable feed composition $X_d$ & $X_b$ (C4 splitter high purity range)	129
Figure 92. ProMV PLS $X_b$ model as function of four variables plot and statistical analysis for variable feed composition $X_d$ & $X_b$ (C4 splitter high purity range)	130
Figure 93. Relationship between $X_b$ equation slope and feed LK (C4 splitter high purity range)	131
Figure 94. Relationship between $X_b$ equation intercept and feed LK (C4 splitter high purity range)	132
Figure 95. Relationship between tray number and $\Delta T / \Delta x_d$ in the top section for constant feed composition and variable $X_b$	134
Figure 96. Relationship between $X_d$ and trays 4, 5, 6, 7 & 8 for constant feed composition and variable $X_b$	135
Figure 97. Relationship between $X_d$ and $T_4$ for variable feed composition & $X_b$	136
Figure 98. ProMV PLS $X_d$ model as function of two variables plot and statistical analysis for variable feed composition, $X_d$ & $X_b$	138
Figure 99. Relationship between tray number and $\Delta T / \Delta x_b$ in the bottom section for constant feed composition and variable $X_d$	140
Figure 100. Relationship between $X_b$ and $T_{32}$ , $T_{33}$ , $T_{34}$ , $T_{35}$ , $T_{36}$ , $T_{37}$ & $T_{38}$ for constant feed composition and variable $X_d$	141
Figure 101. Relationship between $X_b$ and $T_{38}$ for variable feed composition & $X_d$	142
Figure 102. ProMV PLS $X_b$ model as function of two variables plot and statistical analysis for variable feed composition, $X_d$ & $X_b$	144

## List of Tables

Table 1. Butane C4 splitter base case operating parameters _____	20
Table 2. Feed compositions selected for C4 splitter models development _____	21
Table 3. Experimental plan for C4 splitter models development _____	22
Table 4. Slope and intercept for Xd equation at constant feed composition and variable Xb _____	24
Table 5. Slope and intercept equations as function of Xb accuracy analysis for constant feed composition _____	25
Table 6. Slope and intercept for Xb equation at specified Xd at constant feed composition _____	27
Table 7. Comparison between Xb equation linear and quadratic regression for constant feed composition _____	31
Table 8. Comparison between Xd & Xb model for constant feed composition and variable Xd & Xb ____	31
Table 9. Xb equation coefficients for variable feed composition & Xd _____	33
Table 10. C1, C2 & C3 coefficients as function of LK% equations accuracy analysis _____	35
Table 11. Xb model testing results for 144 cases (training date) & 30 cases (prediction data) of variable feed composition, Xd & Xb _____	36
Table 12. $\ln(Xd) = F(R/D)$ Slope and intercept values for variable Xb at constant feed composition ____	38
Table 13. Slope and intercept as function of Xb fitting analysis _____	39
Table 14. Slope a and intercept b values for $\ln(Xb)$ equation for variable Xd at constant feed composition _____	41
Table 15. Slope and intercept as function of Xd fitting analysis _____	43
Table 16. Comparison between Xd & Xb models accuracy for 36 cases each at constant feed composition _____	44
Table 17. List of slope and intercept equations coefficients for variable feed composition _____	46
Table 18. d1 - d6 coefficients as function of LK% fitting analysis _____	50
Table 19. Xb model testing results for 144 cases (training data) & 30 case (prediction data) of variable feed composition, Xd & Xb _____	51
Table 20. Heat of condensation and vaporization as function of Xd & Xb fitting analysis _____	54
Table 21. DeC3 base case operating parameters _____	55
Table 22. Feed compositions selected for experiments with compositions in mass basis _____	57
Table 23. Experimental plan for developing DeC3 model _____	57
Table 24. Xd equation slope and intercept at five Xb values (constant feed composition) _____	59
Table 25. Slope and intercept function of Xb fitting results _____	61
Table 26. Slope and intercept equation coefficients for variable feed composition _____	63
Table 27. C1 – C6 coefficients equations as function of HK with $R^2$ and RMSEE. _____	67
Table 28. Xd model testing results for 125 cases (training data) & 25 (prediction data) at variable feed composition, Xd & Xb _____	68
Table 29. Xb equation coefficient for variable Xd values at constant feed composition _____	70
Table 30. Xb quadratic coefficients C1, C2 and C3 as function of Xd fitting results _____	70
Table 31. $\ln(Xb)$ equation slope and intercept values for constant feed composition and variable Xd ____	71
Table 32. $\ln(Xb)$ equation slope and intercept as function of Xd fitting analysis at constant feed composition _____	73
Table 33. $\ln(Xb)$ slope and intercept equation coefficients for variable feed composition and Xd _____	74
Table 34. Slope and intercept equation coefficients as function of LK accuracy analysis _____	77
Table 35. coefficients contribution to intercept prediction _____	77
Table 36. Xb model testing results (125 cases training data) & (25 prediction data) for variable feed composition and Xd _____	78
Table 37. $VdF$ equation coefficients for variable feed composition _____	83
Table 38. Coefficients C1, C2 & C3 as function of LK% fitting analysis _____	85

Table 39. Vapor distillate model testing for variable feed composition, Xd and Xb	86
Table 40. Vd components composition as function of Xd fitting analysis	89
Table 41. C4 splitter optimization case study constraints	92
Table 42. C4 splitter optimization results comparison between hybrid model and Aspen plus	93
Table 43. DeC3 optimization case study constraints	94
Table 44. DeC3 optimization results comparison between hybrid model and Aspen plus	95
Table 45. Xd equation as function of T10, T12 & T15 accuracy analysis (C4 splitter low purity range)	102
Table 46. ProMV PLS models for Xd for 144 cases of variable feed composition, Xd & Xb & 30 prediction data set (C4 splitter low purity range)	105
Table 47. Xd =F(T <sub>10</sub> ) slope and intercept for variable feed composition & Xb (C4 splitter low purity range)	107
Table 48. Xd model with feed HK input accuracy analysis (C4 splitter low purity range)	109
Table 49. ProMV PLS models for Xb for 144 cases of variable feed composition, Xd & Xb & 30 prediction data set (C4 splitter low purity range)	113
Table 50. Xb model with slope and intercept as function of feed LK accuracy analysis for 144 cases and 30 prediction set data (C4 splitter low purity range)	115
Table 51. Xd as function of T <sub>10</sub> , T <sub>12</sub> & T <sub>15</sub> accuracy comparison (C4 splitter high purity range)	118
Table 52. ProMV PLS models for Xd for 144 cases of variable feed composition, Xd & Xb & 30 prediction data set (C4 splitter high purity range)	122
Table 53. Xd = F(T <sub>10</sub> ) slope and intercept for variable feed composition & Xb (C4 splitter high purity range)	123
Table 54. Xd model with feed HK input accuracy analysis (C4 splitter high purity range)	124
Table 55. Xb as function of T <sub>72</sub> , T <sub>76</sub> , T <sub>80</sub> , T <sub>84</sub> & T <sub>88</sub> accuracy comparison	127
Table 56. ProMV PLS Xb models accuracy analysis for 144 cases of variable feed composition Xd & Xb and 30 cases of prediction data set (C4 splitter high purity range)	130
Table 57. Xb = F(T <sub>88</sub> ) slope and intercepts for variable feed composition & Xd (C4 splitter high purity range)	131
Table 58. Xb model with slope and intercept as function of feed LK accuracy analysis	133
Table 59. Xd as function of T <sub>4</sub> , T <sub>5</sub> , T <sub>6</sub> , T <sub>7</sub> & T <sub>8</sub> accuracy comparison	136
Table 60. Xd model with single temperature variable accuracy analysis for 125 cases of variable feed composition, Xd & Xb and 25 cases of prediction data set	137
Table 61. ProMV PLS Xd model as function of two variables accuracy analysis for 125 cases of variable feed composition, Xd & Xb and 25 cases of prediction data set	139
Table 62. Xb as function of T <sub>32</sub> , T <sub>33</sub> , T <sub>34</sub> , T <sub>35</sub> , T <sub>36</sub> , T <sub>37</sub> & T <sub>38</sub> accuracy comparison	141
Table 63. Xb model with single temperature variable accuracy analysis for 125 cases of variable feed composition, Xd & Xb and 25 cases of prediction data set	143
Table 64. ProMV PLS Xb model as function of two variables accuracy analysis for 125 cases of variable feed composition, Xd & Xb and 25 cases of prediction data set	145

## ***Nomenclature***

### ***Abbreviations***

DCS	Distributed control system
RTO	Real time optimization
PLS	Partial least square
GC	Gas chromatography analyzer
DeC3	Depropanizer column
C4 Splitter	Butane (iC4/nC4) splitter column
VLE	Vapor liquid equilibrium
SF	Separation factor
LK	Light key component
HK	Heavy key component
Xd	Impurity in distillate stream (heavy key composition in distillate)
Xb	Impurity in bottom stream (light key composition in bottom)
K-value	Vapor-liquid equilibrium constant
TBP	True boiling point
LPG	Liquefied petroleum gas

### ***Subscripts***

<i>i</i>	Component or tray number
<i>r</i>	Reboiler
<i>c</i>	Condenser
<i>C2</i>	Ethane
<i>C3</i>	Propane
<i>IC4</i>	Iso-Butane
<i>NC4</i>	Normal-Butane
<i>IC5</i>	Iso-Pentane
<i>NC5</i>	Normal-Pentane
<i>D</i>	Distillate (liquid) stream
<i>Vd</i>	Vapor distillate stream
<i>B</i>	Bottom stream
<i>F</i>	Feed stream
<i>T</i>	Column top section (rectifying section)
<i>B</i>	Column bottom section (stripping section)

## ***Variables***

$SF$	Separation factor
$x_d$	Heavy key composition in distillate stream
$x_b$	Light key composition in bottom stream
$R$	Reflux flow (Kg/hr)
$D$	Distillate flow (Kg/hr)
$Vd$	Vapor distillate (Kg/hr)
$B$	Bottom flow (Kg/hr)
$F$	Feed flow (Kg/hr)
$V_r$	Boilup flow (reboiler vapor outlet) (Kg/hr)
$Q_r$	Reboiler duty (MJ/hr)
$Q_c$	Condenser duty (MJ/hr)
$\lambda_r$	Reboiler heat of vaporization (MJ/Kg)
$\lambda_c$	Condenser heat of condensation (MJ/Kg)
$\lambda'_c$	Condenser sensible and latent heat coefficient (MJ/hr)
$H_i$	Enthalpy for liquid stream $i$ (MJ/hr)
$h_i$	Enthalpy for vapor stream $i$ (MJ/hr)
$c_p$	Heat capacity (MJ/Kg-K)
$x_i$	Composition for component $i$ in mass basis
$\rho_i$	Density for component $i$ (Kg/m <sup>3</sup> )
$LK$	Light key (C3 in DeC3 & IC4 in C4 splitter)
$HK$	Heavy key (IC4 in DeC3 & NC4 in C4 splitter)
$T$	Temperature (C°)
$T_i$	Temperature at tray $i$ (C°)
$L$	Internal liquid flow (Kg/hr)
$V$	Internal vapor flow (Kg/hr)
$Z$	Objective function (\$/hr)

## **1. Introduction**

Distillation columns are primary equipment in petrochemical, gas plants and refineries. Distillation column energy consumption is estimated to be 40% of the total plant energy consumption. Optimization of distillation columns has potential for saving large amount of energy and contributes to plant wide optimization.

Currently rigorous tray to tray simulations are used to describe columns separation with very high accuracy. However this involves large number of equations in addition to being difficult to incorporate as part of the plant Distributed Control System DCS. Furthermore incorporating rigorous tray to tray simulations for planning, scheduling and control models can be difficult. Currently there are different types of models used for planning, scheduling and real-time optimization which causes discrepancies between their outputs. This calls for introducing simpler, smaller in size and yet accurate models that can describe separation.

Hybrid model is comprised of first principles material and energy balances and empirical models which describe column separation. Empirical models are built based on plant data or based on rigorous simulation data. Hybrid model is simpler, smaller in size and accurate as compared to rigorous simulation. It defines operating conditions and energy consumption requirement to meet purity specification. Compared to simplified models currently used (e.g. split fraction models), hybrid models have higher accuracy over a wide range of operating conditions and feed compositions. They can be implemented for planning, scheduling and optimization purposes. Below is the hybrid model structure.

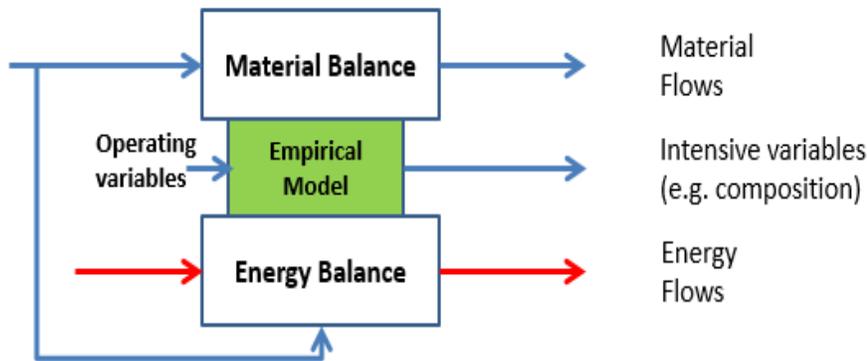


Figure 1. Hybrid model structure for distillation columns

Inferential models rely on selected operating variables to compute product composition. They have been used extensively for monitoring and control purposes. Commonly inferential models use single or multiple tray temperatures as their inputs. However, if feed composition varies, such models are often not sufficiently accurate. In order to increase accuracy of inferential models, they were developed with the use of partial least square regression to include multivariable. Model inputs are real time variables such as multiple trays temperature, reflux flow, distillate flow, bottom flow and reboiler duty. It is simple model which can be incorporated as part of DCS. It estimates accurately streams purity which eliminate the need for Gas Chromatography analyzer GC. Commonly GC analyzer is used to analyze streams purity however this requires installation and maintenance cost. In addition GC analyzers face reliability issues and hence their input is not a reliable source.

This work developed hybrid and inferential models for two types of distillation column. The first example is Butane C4 splitter which separates binary mixture of narrow boiling point components iso-butane IC4 & normal-butane NC4. The second example is

Depropanizer DeC3 column which separates multicomponent mixture of wide boiling point key components propane C3 & IC4. Both columns are part of many refineries and gas plants configuration.

### **1.1 Main contributions**

This work illustrates the development methodology for hybrid and inferential models using Aspen plus simulation data input for both columns. The final models results were compared to Aspen plus simulation results and indicated good accuracy. The main contributions for this research are:

- a) Develop hybrid model that includes mass, energy and empirical equations that describes separation for variable feed composition and streams purity specifications for both columns. The model uses off-line input data such as feed composition and streams maximum impurity specifications. It calculates the primary operating conditions such as reflux ratio and streams flows in addition to determining energy consumption requirement (condenser and reboiler duty). It is small size nonlinear model that can be used for planning, scheduling and optimization.
- b) Develop inferential model that estimates stream purity using real time data input such as tray temperature, reflux ratio and reboiler duty. It eliminates the need for GC analyzer and hence eliminates installation and maintenance cost. In addition GC analyzers are known for experiencing reliability issues. Model development followed the approaches of single input or multivariable inputs using linear PLS regression.

## **1.2 Thesis overview**

Section 2, presents a brief review of previous work on distillation column separation including separation factor model for binary separation. In addition it includes a review of previous hybrid and inferential models development.

Section 3, illustrates hybrid model development for C4 splitter and DeC3 columns in details including simulation data generation, model development, and verification. In addition to applying hybrid model in two optimization case studies to test models accuracy as compared to Aspen plus optimization tool, this section also shows examination of separation factor modeling approach for binary mixture separation (C4 splitter example).

Section 4, illustrates inferential model development for C4 splitter and DeC3 columns in details including simulation data generation, model developing, and verification.

Section 5, highlights major accomplishments and result of this research, followed by recommendation for future research

## 2. Literature review

In this section a brief review of previous work on distillation column separation such as separation factor model for binary separation. In addition, a review of the previous hybrid and inferential models development is presented

For binary separation the McCabe Thiele diagram demonstrates the graphical solution to find the number of theoretical stages (McCabe Thiele (1925)). Usage of graphical solution can be difficult for optimization problems. In addition it is difficult to implement accurately for columns with large number of stages. Later several analytical solutions for number of theoretical stages were introduced. FUG (Fenske-Underwood-Gilliland) equations introduced calculations of theoretical number of stages. The Fenske equation (Fenske (1932)) estimates the minimum number of theoretical stages at total reflux. (Underwood (1948)) The Underwood equation for an infinite number of theoretical stages estimates minimum reflux ratio. (Gilliland (1940)) The Gilliland method is an empirical correlation represented graphically. It calculates the theoretical stages number for specified separation as function of reflux ratio. It requires knowing the minimum number of theoretical stages at total reflux ratio and the minimum reflux ratio. Later (Molokanov et al. (1972)) derived analytical solution for multicomponent separation defining reflux ratio and number of theoretical trays. (Boston et al. (1974)) developed algorithm to solve the equations for multicomponent distillation columns.

Smoker (1938) developed an exact analytical solution to compute the number of theoretical stages for binary separation with the assumption of constant molar overflow. Jafarey et al. (1979 a) introduced a short cut technique with an approximate analytical solution of Smoker's (1938) equation with the use of separation factor. The separation

factor is function of both top and bottom streams purity. (Jafarey et al. (1979 a)) assumed that the number of stages in the rectifying section equals to the number of stages in the stripping section. They calculated the number of theoretical stages for design purposes; their equations can also be used to calculate the reflux ratio for specified separation. (Skogestad and Morari (1987)) introduced an analytical expression for the separation factor as function of the internal flows which gave similar results to (Jafarey et al. (1979a)) equation. It uses rectifying and stripping number of stages explicitly. It was derived with the following assumption

- Constant relative volatility
- Constant molar flows
- Feed is optimally located

This equation provides approximation results. Recently (Soliman, M. A.(2007)) introduced a new equation by adding a new term to (Jafarey et al. (1979a)). It is intended to be used for sharp separation with low constant relative volatility.

The equations for calculating the theoretical number of stages are mostly intended for column design. Once the column is in operation the concern becomes column operability and control. It is a common understanding among plants operators that an increase of the reflux ratio improves distillate purity assuming primary product is the distillate. However, this action increases column internal liquid flow and so it increases the reboiler duty. (Shinskey (1976)) introduced an empirical correlation that relates the boilup to feed ratio to the separation factor. This equation is intended for operability and control purposes rather than for column design. (Jafarey et al. (1979 b)) used their

approximation equation for operability and control purposes and the results agree with (Shinsky (1976)) empirical correlation.

The journey to establish robust models for predicting distillation column streams purity took several decades. This was achieved through the development of rigorous distillation column models. These models are based on mass and energy balance equation in addition to vapor liquid equilibrium (VLE) equations at each individual tray. Nowadays they provide tray to tray analysis such as flow, temperature, pressure, composition and VLE equilibrium data. This introduced the advantage of being used for detailed design and real time optimization RTO. However they are large in number of equations and highly nonlinear. In industrial plants they are used for process simulation, design and real time optimization RTO for various cases of operating parameters due to their high accuracy. On the other hand they are not suitable for production planning and scheduling due to the large number of equations and being highly nonlinear. In order to solve distillation column production planning and scheduling estimation models were introduced. However these estimation models in most cases don't represent the actual columns operation. This calls for introducing simpler, smaller in size and yet accurate models that can describe separation.

Hybrid models are set of equations that describes distillation column operating conditions such as streams flows and compositions, energy flows, reflux ratio and boilup ratio. They consist of mass balance, energy balance and empirical predictive equations. The main part in model development is generating empirical predictive equation for stream purities using operating conditions such as reflux ratio, boilup ratio, condenser duty and reboiler duty. The separation equations are mostly nonlinear and the

degree of nonlinearity is observed higher when operating at high purity regions. For some cases the Partial Least Square PLS technique can be used to generate a linear predictive model using multivariable. Hybrid models are small in size yet accurate as compared to tray to tray rigorous models.

The past hybrid models development application is mainly focused on Crude Distillation Units CDU due to the fact that CDU modeling play major role in refineries planning, scheduling and optimization. There several paper published for CDU hybrid model development. (Mahalec and Sanchez (2012)) developed online inferential model using trays temperatures for products yield and quality prediction. The model included CDU column and intended for use at real-time operation. Recently (Mahalec and Fu (2015)) expanded the scope to include the preflash tower and vacuum tower. In addition to presenting streams properties prediction models. (Fu and Mahalec (2015)) current work can be used for production planning, scheduling and RTO. The model doesn't rely on tray temperatures and so can be used effectively for production planning and scheduling.

The binary separation and other multicomponent distillation columns models are different from CDU models. They describe product compositions in terms of component, while CDU separates a set of components belonging to certain range of True Boiling Point TBP. For binary separation the separation factor SF is highly nonlinear equation indicating the degree of separation in which the higher SF the higher the streams purity. (Safavi, et al. (1999)) used wave - nets neural networks for hybrid model development. Their study was conducted on experimental distillation column which separates a binary mixture of water and ethanol. The wave-net used to model SF

and  $\text{Log}(\text{SF})$  for various variables. The hybrid model consists of mass and energy balance equations along with a wave-net model for  $\text{Log}(\text{SF})$  which showed higher prediction accuracy as compared to SF model. The model included 64 coefficients with training data set of 800 points while testing was conducted on 3000 data points. Then the model was used for optimization cases and showed similar results as compared to rigorous model.

The inferential models for distillation column rely mainly on tray temperature measurement to estimate streams purity. There are online measurement for composition such Gas Chromatographs GC analyzer to measure streams purity. However analyzers require installation and maintenance in addition to facing reliability issues. The inferential models can estimate composition and hence eliminate the need for GC analyzers. Inferential models estimate streams composition for various variables such as feed composition, reflux ratio and reboiler duty. The main component in the inferential model is tray temperature. The trays at the column top and bottom end provide the highest accuracy for composition estimation. However the temperature variation with respect to composition variation tend to be small. In industrial plants inferential models are built based on selecting a sensitive tray. The selection is based on tray sensitivity for composition variation when experiencing variation in feed composition, reflux ratio and reboiler duty. To enhance inferential models accuracy multivariable can be included such multiple tray temperatures, feed composition, reflux ratio and reboiler duty. The PLS regression can be used to generate inferential model function based on many tray temperature measurements. (Mejdell and Skogestad (1990)) introduced PLS application for generating distillation inferential models for control. (Mejdell & Skogestad (1991

a,b)) published the results for building inferential models for both distillation simulation and pilot plant. The models were generated using multiple temperature measurements. (Kresta et al. (1994)) outlined the procedure of building inferential model using PLS. The models included large number of operational variables and applied for two distillation examples. They showed that PLS model performs well during missing data or sensor failures. In addition it showed higher accuracy as compared to single temperature model for variation in feed composition.

### **3. Hybrid Model Development**

#### **3.1 Introduction**

This section illustrates hybrid model development for two distillation column examples. Hybrid model can be used off-line and includes mass and energy balance equations in addition to empirical models. Model input includes feed composition in addition to defining the required maximum streams impurity specifications. The streams impurities are defined as  $X_d$  (HK heavy key composition in distillate stream) and  $X_b$  (LK light key composition in bottom stream). Models were developed using Aspen Plus simulation data as input. Below are the main variables used for hybrid model development.

- Feed composition
- $X_d$  top impurity composition
- $X_b$  bottom impurity composition

The output predicts the column operating conditions such as distillate flow, bottom flow, reflux ratio and boilup ratio. In addition to defining the energy requirement (condenser and reboiler duties). The model can be used for planning and scheduling for

various specifications of feed composition and streams impurity specifications. The model is nonlinear and can be used as stand-alone application or part of the complete plant model. The models were applied for two optimization case studies. This section also included testing the approach of using separation factor for binary separation columns.

### 3.2 Methodology

Commonly, simplified distillation column model is developed at constant feed composition while in actual operation feed composition is constantly fluctuating. In addition, it is often assumed that Xd model includes with constant bottom operation (constant reboiler duty – constant Xb) and vice versa. This assumption can be misleading as operation variation at one column end affects the internal flows and so the column energy flows.

Reflux ratio is the key parameter in determining top impurity Xd.

$$\text{Reflux Ratio} = \frac{R}{D} \quad (1)$$

while boilup ratio or  $\frac{Q_r}{B}$  is the key parameter in determining bottom impurity Xb.

$$\text{Boilup Ratio} = \frac{V_r}{B} = \frac{Q_r}{\lambda_r \times B} \quad (2)$$

The equation for the distillate impurity Xd has structure as shown below

$$x_d = a \left( \frac{R}{D} \right) + b \quad (3)$$

Xd is linear function of reflux ratio as Eq. (3) however as Xd decreases the more nonlinear behaviour is observed. Xd equation if needed can be expressed with 2<sup>nd</sup>

degree polynomial or exponential regression to account for nonlinearity. For operating with variable feed composition and  $X_b$ ,  $X_d$  equation experience shifting as shown in the below example.

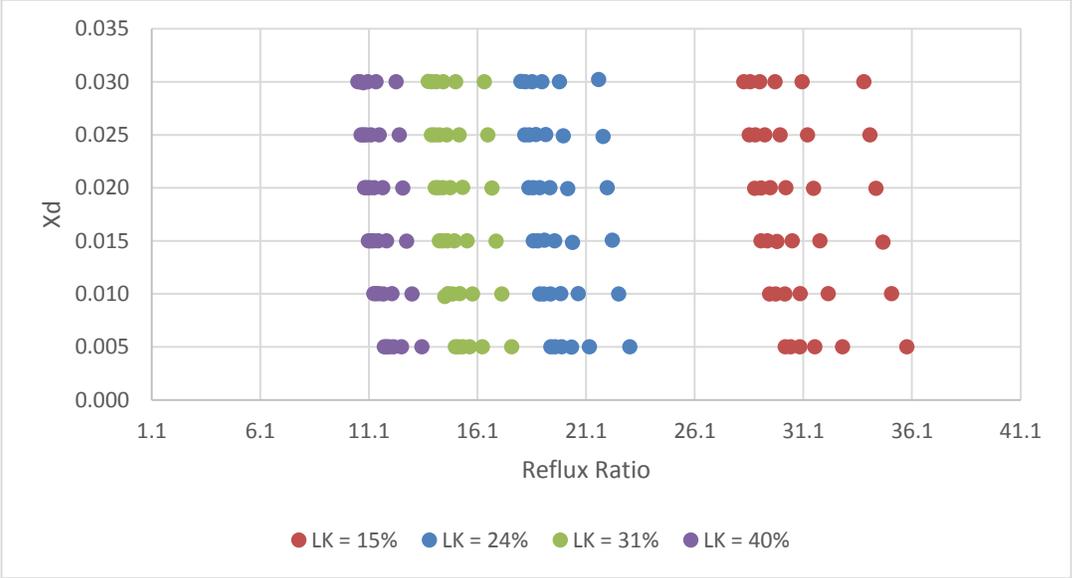


Figure 2. Relationship between top impurity  $X_d$  and reflux ratio for variable feed composition &  $X_b$

$X_d$  is function of reflux ratio in which within the same feed composition (same color) it experiences shifting due to  $X_b$  variation. In addition  $X_d$  equation experiences shifting due to feed composition variation (different color, Fig. 2). It can be described as a single operation line that moves horizontally due to feed composition and  $X_b$  variation. Hence  $X_d$  equation can be described as  $X_d = \text{Function}(\text{Reflux ratio}, X_b, \text{feed composition})$ .

$X_b$  equation on the other hand is considered function of  $\frac{Q_r}{B}$  as Eq. (4).  $\frac{Q_r}{B}$  is chosen over the boilup ratio to eliminate the error caused by variation of heat of vaporization  $\lambda_r$ . Variation in  $X_b$  contributes to changes in  $\lambda_r$ . Hence using  $\frac{Q_r}{B}$  tend to be more accurate and based on  $X_b$  calculation  $\lambda_r$  can be accurately calculated to determine boilup ratio.

$$x_b = a \left( \frac{Q_r}{B} \right) + b \quad (4)$$

$X_b$  is linear function of  $\frac{Q_r}{B}$  however as  $X_b$  range goes low the more nonlinear behaviour is observed. The equation if needed can be expressed with 2<sup>nd</sup> degree polynomial or exponential regression to account for nonlinearity. For operating with variable feed composition and  $X_d$ ,  $X_b$  equation experience shifting as shown in the below example.

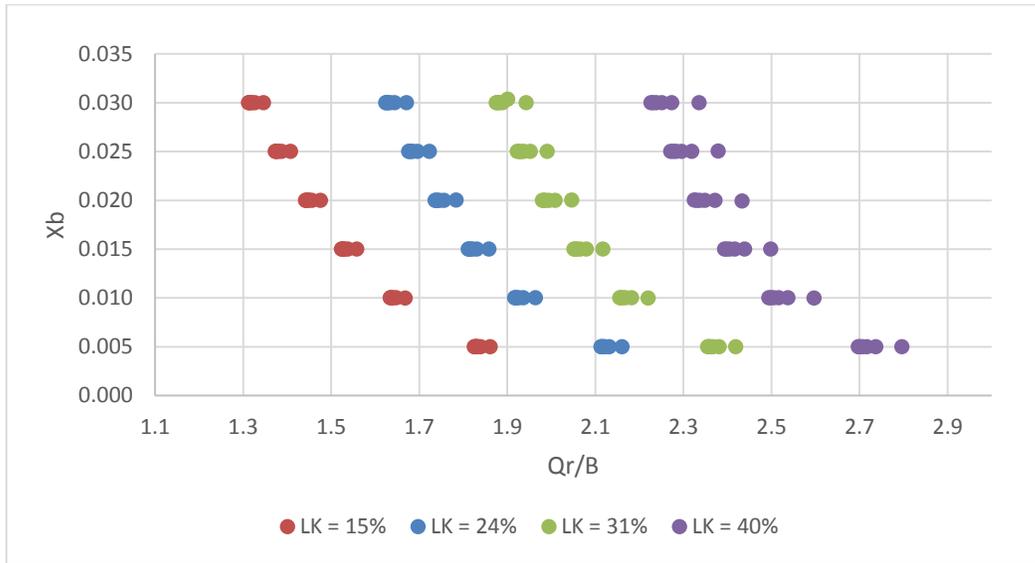


Figure 3. Relationship between bottom impurity  $X_b$  and  $Q_r/B$  for variable feed composition &  $X_d$

$X_b$  is function of  $\frac{Q_r}{B}$  in which within the same feed composition (same color) it experiences shifting due to  $X_d$  variation. In addition  $X_b$  equation experiences shifting due to feed composition variation (see different colors, Fig. 3). It can be described as a single operation line that moves horizontally due to feed composition and  $X_d$  variation.

Hence  $X_b$  equation can be described as  $X_b = \text{Function} \left( \frac{Q_r}{B}, X_d, \text{feed composition} \right)$ .

The empirical models have been developed via following two step procedure:

- At constant feed composition
  - Derive equations for prediction of  $X_d$  &  $X_b$  by best regression fitting (linear, 2<sup>nd</sup> degree polynomial or exponential)
  - Introducing additional equations that describe  $X_d$  dependence on  $X_b$  and  $X_b$  dependence on  $X_d$
- At variable feed composition
  - Derivation of model equations to account for variable feed composition

The model input is defined by the feed composition and  $X_d$  &  $X_b$  specification. The hybrid model output includes column operating conditions such as reflux ratio, boilup ratio,  $Q_c$ ,  $Q_r$  and streams flow.

### 3.3 Testing Separation Factor modeling approach

The separation factor SF is function of both  $X_d$  &  $X_b$  and it indicates the level of separation for binary mixture through the following relationship.

$$SF = \frac{(1-x_d)(1-x_b)}{x_d x_b} \quad (5)$$

Eq. (5) employs compositions expressed in mole fractions. SF is a nonlinear equation in two variables. Higher SF occurs at low  $X_d$  &  $X_b$  indicating high purity operation. (Skogestad and Morari (1987)) introduced an analytical expression Eq. (6) which describes the separation factor changes due to internal flows changes with the assumptions

- Constant relative volatility
- Constant molar flows

- Feed is optimally located

$$SF = \alpha^N \frac{(L/V)_T^{N_T}}{(L/V)_B^{N_B}} \quad (6)$$

Where

- $\alpha$  is the relative volatility,
- L & V liquid and vapor internal flows,
- N = total number of theoretical stages,
- N<sub>T</sub> & N<sub>B</sub> number of theoretical stages in the top and bottom respectively.

The formula performs better for high purity ranges. It was applied to binary separation column (C4 Splitter) for high purity range (N<sub>T</sub> = 50, N<sub>B</sub> = 49,  $\alpha = 1.221$  & LK=24%).

The internal reflux in the top section is estimated by the ratio  $\frac{R}{R+D}$  and in the bottom section by ratio  $\frac{V_{r+B}}{V_r}$  where V<sub>r</sub> is calculated from reboiler duty. All the above flows were obtained using Aspen plus simulation. The actual separation factor is calculated using AspenPlus X<sub>d</sub> & X<sub>b</sub> values. The calculated SF values are much higher than actual SF values. However the error showed relationship with X<sub>d</sub> & X<sub>b</sub> variation. The plot in Fig. 4 shows SF<sub>calculated</sub>/SF<sub>actual</sub> at constant feed composition (LK=24%) and variable X<sub>d</sub> & X<sub>b</sub>. The SF<sub>calculated</sub>/SF<sub>actual</sub> ratio varies with X<sub>d</sub> at constant X<sub>b</sub>. Slope and the intercept of Eq. (7) vary with X<sub>b</sub>.

$$\frac{SF_{calculated}}{SF_{actual}} = a x_d + b \quad (7)$$

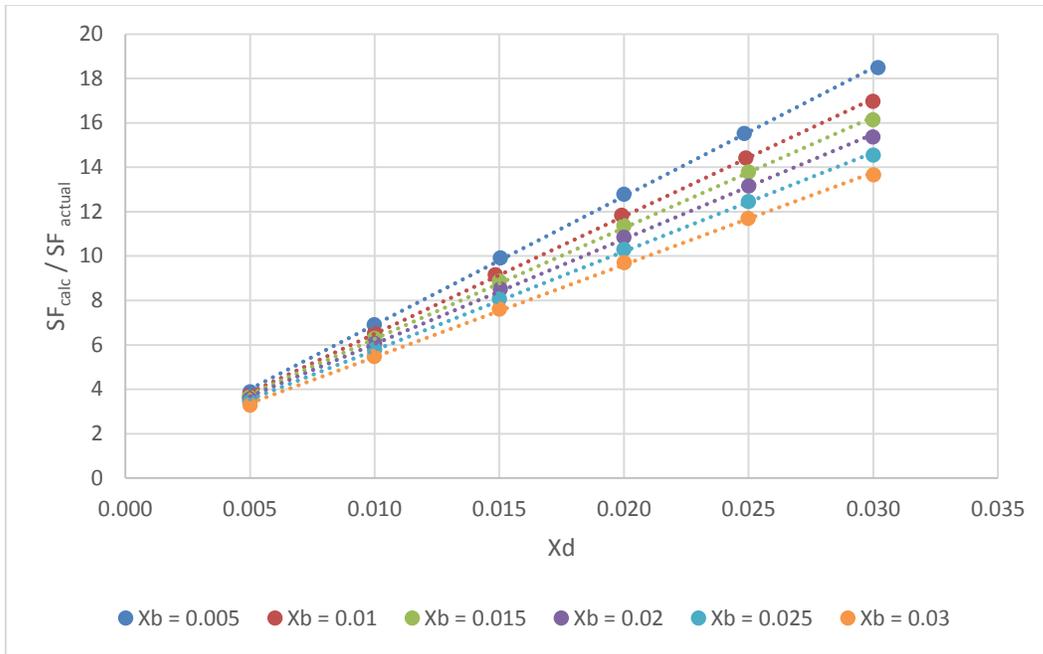


Figure 4.  $SF_{\text{Calculated}} / SF_{\text{actual}}$  for constant feed composition and variable  $X_d$  &  $X_b$

The error in  $SF_{\text{calc}}/SF_{\text{actual}}$  is lower at lower  $X_d$  &  $X_b$ ; however it is still a high error since the calculated SF is at least 3 times the actual value. The slope  $a$  and intercepts  $b$  vary with  $X_b$  values and they were modeled as 2<sup>nd</sup> polynomial function of  $X_b$ , as shown in Figs. 5 and 6.

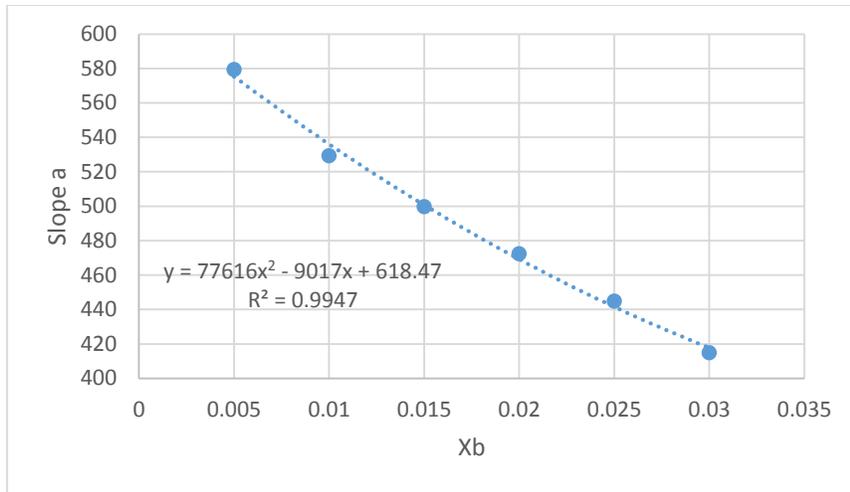


Figure 5. Relationship between  $SF_{\text{calculated}} / SF_{\text{actual}}$  slope and Xb for constant feed composition

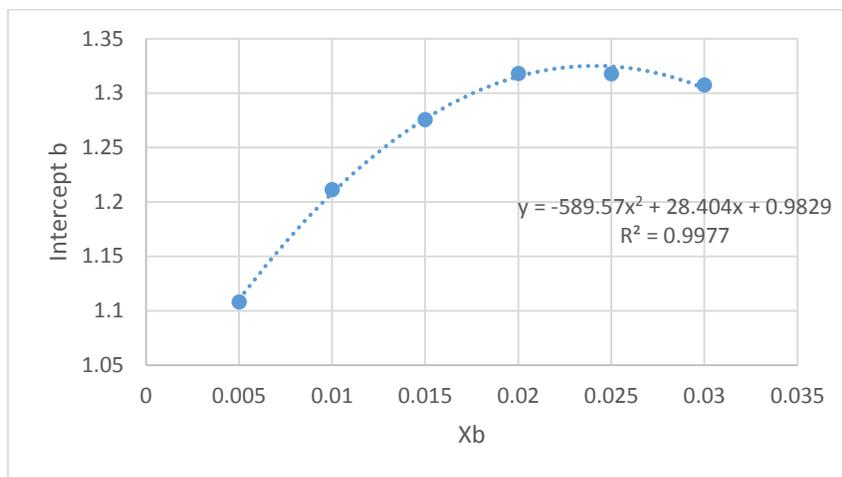


Figure 6. Relationship between  $SF_{\text{calculated}} / SF_{\text{actual}}$  intercept and Xb for constant feed composition

The same analysis were conducted for additional feed compositions (LK= 15%, 31% & 40%) and below Fig. 7 shows the  $SF_{\text{calculated}}/SF_{\text{actual}}$  at variable feed composition.

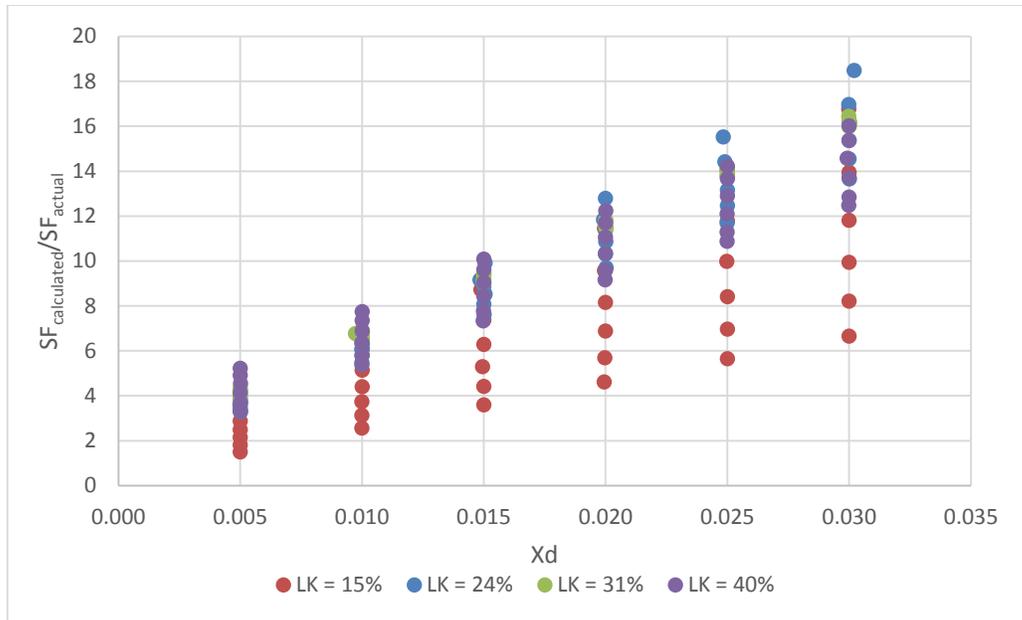


Figure 7.  $SF_{\text{calculated}} / SF_{\text{actual}}$  for variable feed composition,  $X_d$  &  $X_b$

It is possible to model  $X_d$ ,  $X_b$  and feed composition however the model will be highly nonlinear. In addition, the final outcome would be SF which is function of both  $X_d$  &  $X_b$ . This approach of modeling SF does not express each stream impurity  $X_d$  &  $X_b$  explicitly.

### 3.4 Hybrid Model Development for Butane Splitter Column

#### 3.4.1 Introduction

The Butane splitter column separates binary mixture of iso-Butane iC4 and normal-Butane nC4. This is a typical column that can be part of many oil refineries or gas plants. The iC4 product is fed to the Alkylation unit for further processing to boost the RON and then fed to Gasoline pool. While the nC4 is fed directly to LPG pool and in some cases with small portion goes to Gasoline pool. This column deals with binary separation of narrow boiling point mixture of iC4 & nC4 in which the number of trays

is high and the tray to tray temperature variation is small. Below is the Butane C4 splitter column configuration and base case operating parameter (Fig. 8 & Table. 1).

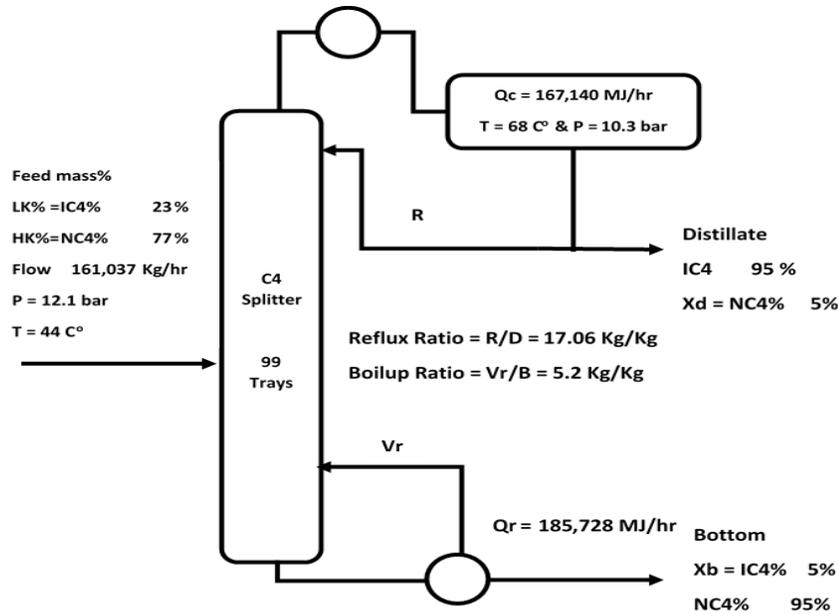


Figure 8. Butane C4 splitter column configuration and base case operating parameters

### 3.4.2 Experimental Plan and Assumption

The experimental simulation cases were conducted using Aspen plus simulation engine with thermodynamics package Peng–Robinson equation of state. Below are the assumptions considered during the model development.

- Constant feed temperature & pressure
- Constant condenser pressure
- Condenser operates at saturation temperature
- Constant pressure drop across the column
- Constant enthalpy  $H^0$ , density  $\rho$  and heat capacity  $c_p$  for F, D and B.
- Two operation ranges for distillate and bottom streams

- low purity range 0.92 – 0.97
- high purity range 0.97 – 0.995

The models of two operating ranges has been developed: high and low purity. The low purity range (high Xd & Xb) model exhibits almost linear behavior for individual Xd & Xb equations, while the high purity range (low Xd & Xb) model exhibits highly nonlinear model with exponential behavior for individual Xd & Xb equations. Having two operation ranges simplifies the model and minimizes the number of equations.

Table 1. Butane C4 splitter base case operating parameters

Feed Temperature (C °)	44	Condenser Duty (MJ/hr)	167140
Feed Pressure (bar)	12.1	Condenser Temperature (C °)	68
Feed Flow (Kg/hr)	161,037	Condenser Pressure (bar)	10.3
Feed LK% (IC4)	23 %	Reflux Ratio (Kg/Kg)	17.06
Feed HK% (NC4)	77 %	Reboiler Duty MJ/hr	185728
Xd (HK in D)	5 %	Boilup Ratio (Kg/Kg)	5.2
Xb (LK in B)	5 %	Distillate D Flow (Kg/hr)	32,634
Pressure Drop (bar)	1	Bottom B Flow (Kg/hr)	128,402
Number of Trays	99 + Condenser + Reboiler		

The model includes the following ranges of the variables:

- Feed composition (LK% (IC4%) range 15 – 40% Volume basis)

- Xd distillate impurity (NC4% in distillate stream) – Volume basis
  - low purity range 0.03 – 0.08
  - high purity range 0.005 – 0.03
- Xb bottom impurity (IC4% in bottom stream) – Volume basis
  - low purity range 0.03 – 0.08
  - high purity range 0.005 – 0.03

Feed composition, Xd and Xb are expressed on volume basis which is the preferred format in industry. However, the hybrid model is based on mass unit and it includes volume to mass basis conversion equations. All other operating parameter are expressed on mass basis. Once the above variables are specified, the model computes reboiler duty  $Q_r$ , condenser duty  $Q_c$ , reflux ratio, boilup ratio, distillate and bottom streams flow. Below are the feed compositions selected and the summary of experiments (Tables. 2 and 3).

Table 2. Feed compositions selected for C4 splitter models development

Component	Feed 1	Feed 2 (base case)	Feed 3	Feed 4
IC4 (LK)	15%	24%	31%	40%
NC4 (HK)	85%	76%	69%	60%

The model development for each operating range followed two stages

- constant feed composition and variable Xd & Xb (individual model for each feed composition)
- Variable feed composition, Xd and Xb

Table 3. Experimental plan for C4 splitter models development

Operation Range	High purity	Low purity
Feed variation	4 different feed compositions (Feed 1 – Feed 4)	4 different feed compositions (Feed 1 – Feed 4)
Xd variation	6 points over range 0.005 - 0.03 Increment 0.005	6 points over range 0.03 – 0.08 Increment 0.01
Xb variation	6 points over range 0.005 - 0.03 Increment 0.005	6 points over range 0.03 – 0.08 Increment 0.01
Number of experiments	144 (combination of Xd, Xb and feed composition points)	144 (combination of Xd, Xb and feed composition points)

The analysis first was conducted at constant feed composition with variable Xd & Xb. This is to evaluate Xb variation impact on Xd equation and vice versa. Then the individual models were combined to evaluate feed composition impact on Xd & Xb equations.

The models for each operating range were generated using 144 cases of variable feed composition, Xd & Xb. In addition, 30 cases of prediction data set for each operating

range have been evaluated. Prediction data set was selected randomly for variable feed composition,  $X_d$  &  $X_b$  in which they fall within the specified experimental ranges.

### 3.4.3 $X_d$ & $X_b$ Models for Low Purity Range

The analysis below was conducted at constant feed composition and selected feed composition  $LK\% = 24\%$  in order to illustrate the model development procedure. This is to evaluate the column behavior at variable  $X_d$  and  $X_b$ . The below plot as Fig. 9 indicates  $X_d$  equation as Eq.(3) at variable  $X_b$ .

As shown in Fig. 9,  $X_d$  dependence on  $R/D$  shifts due to the changes in  $X_b$  (which corresponds to changes in reboiler duty  $Q_r$ ). When  $Q_r$  changes to achieve specific  $X_b$  value, the internal vapor flow changes which eventually changes the top section vapor flow affecting the reflux ratio. The slope  $a$  and intercept  $b$  depend on the changes in  $X_b$ ; this relationship is expected to be nonlinear due to the fact the  $X_b$  changes cause higher magnitude of shifting at lower  $X_b$  values such as the operating equation at  $X_b = 0.03$

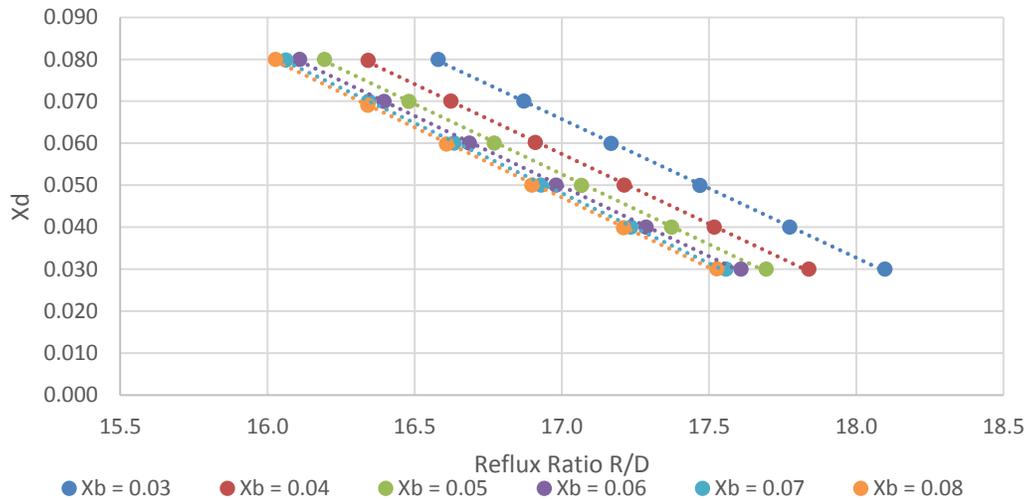


Figure 9. Relationship between  $X_d$  and reflux ratio for constant feed composition and variable  $X_b$

Table. 4 contains the slope and intercept values for each Xd equation at specified Xb. This is to evaluate the slope and intercept changes due to Xb variation.

Table 4. Slope and intercept for Xd equation at constant feed composition and variable Xb

Xb	Slope a	Intercept b
0.030	- 0.03300	0.62685
0.040	- 0.03330	0.62359
0.050	- 0.03343	0.62088
0.060	- 0.03344	0.61833
0.070	- 0.03340	0.61593
0.080	- 0.03336	0.61429

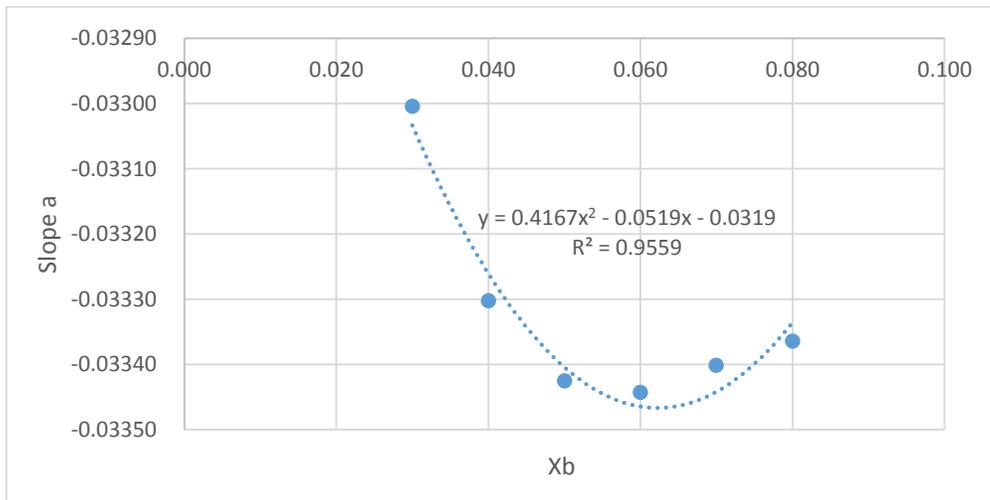


Figure 10. Relationship between Xd correlation slope (a) and Xb at constant feed composition

Eq. (8) expresses the relationship between slope  $a$  and Xb which is demonstrated in Fig. 10. The slope variation is small in magnitude; however, modeling the slope can increase Xd model accuracy.

$$\text{Slope } a = C1x_b^2 + C2 x_b + C3 \quad (8)$$

$$C1 = 0.4167 \quad C2 = -0.0519 \quad C3 = -0.0319$$

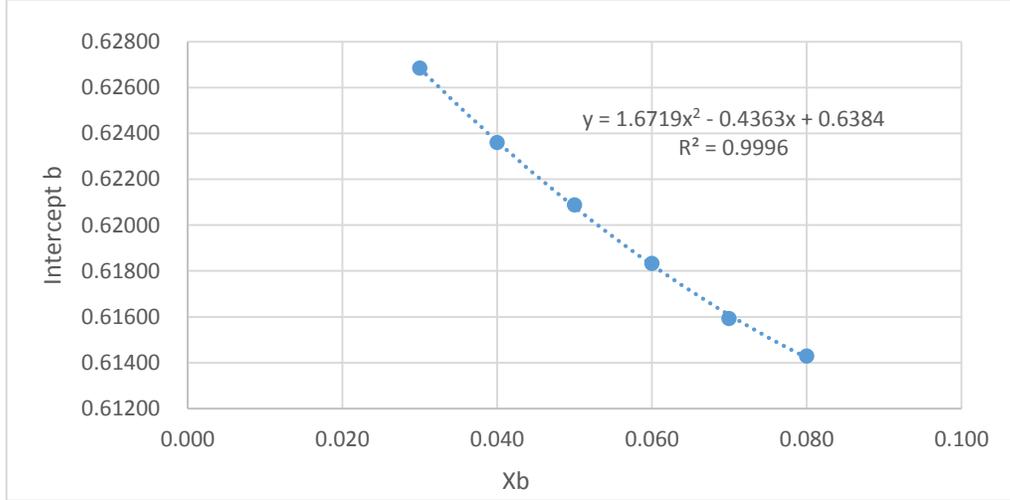


Figure 11. Relationship between Xd correlation intercept b and Xb at constant feed composition

Eq. (9) expresses the relationship between intercept  $b$  and  $X_b$  as in Fig. 11.

$$\text{Intercept } b = C4 x_b^2 + C5 x_b + C6 \quad (9)$$

$$C4 = 1.6719 \quad C5 = -0.4363 \quad C6 = 0.6384$$

The slope and intercept are both nonlinear function of  $X_b$ ; they account for the shifting caused by  $X_b$  variation. Both regression lines for slope and intercept have high  $R^2$  and low RMSEE as shown in (Table. 5).

Table 5. Slope and intercept equations as function of  $X_b$  accuracy analysis for constant feed composition

	Slope $a = F(X_b)$	Intercept $b = F(X_b)$
$R^2$ Quadratic fitting	0.9559	0.9996
RMSEE	3.14776E-05	8.59585E-05
Average Value	-0.03332	0.61998

The complete  $X_d$  model with slope and intercept function of  $X_b$  was tested for 36 case of variable  $X_d$  &  $X_b$  values as per the experimental plan. The relative error% was less than 4% while maximum absolute error is 0.0012. The model has RMSEE = 0.00062 with  $X_d$  average value = 0.055 as shown in Fig. 12.

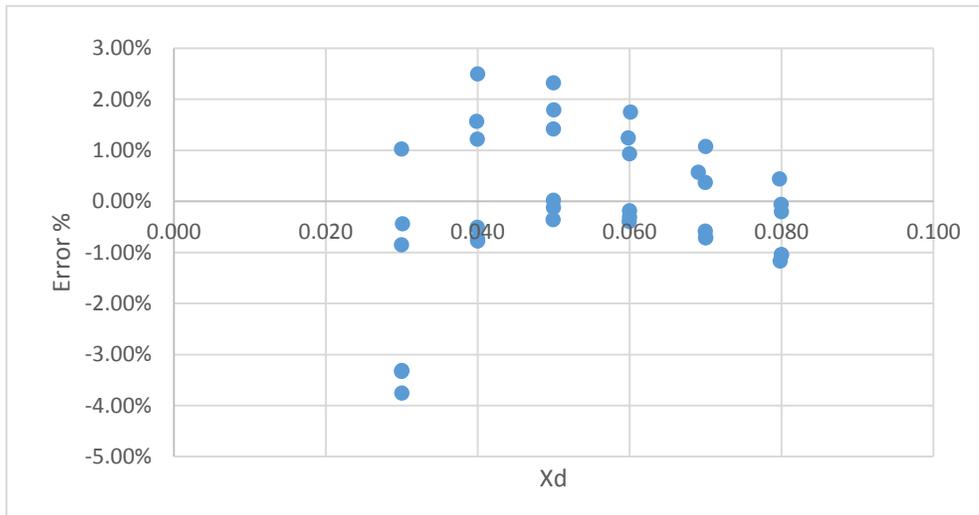


Figure 12. Relative error% for  $X_d$  with slope  $a$  and intercept  $b$  function of bottom impurity  $X_b$  at constant feed composition

Fig. 13 shows dependence of  $X_b$  equation as Eq. (4) at variable  $X_d$  and constant feed composition.  $X_d$  variation causes insignificant impact on that relationship. The changes in reflux ratio to achieve specified  $X_d$  have almost negligible impact on the bottom section of the tower. This is due to the fact that reflux flow has minimal impact on the bottom section liquid flow. Table. 6 shows the slope and intercept values for Eq. (4) for variable  $X_d$  values which indicating negligible impact from  $X_d$  variation.

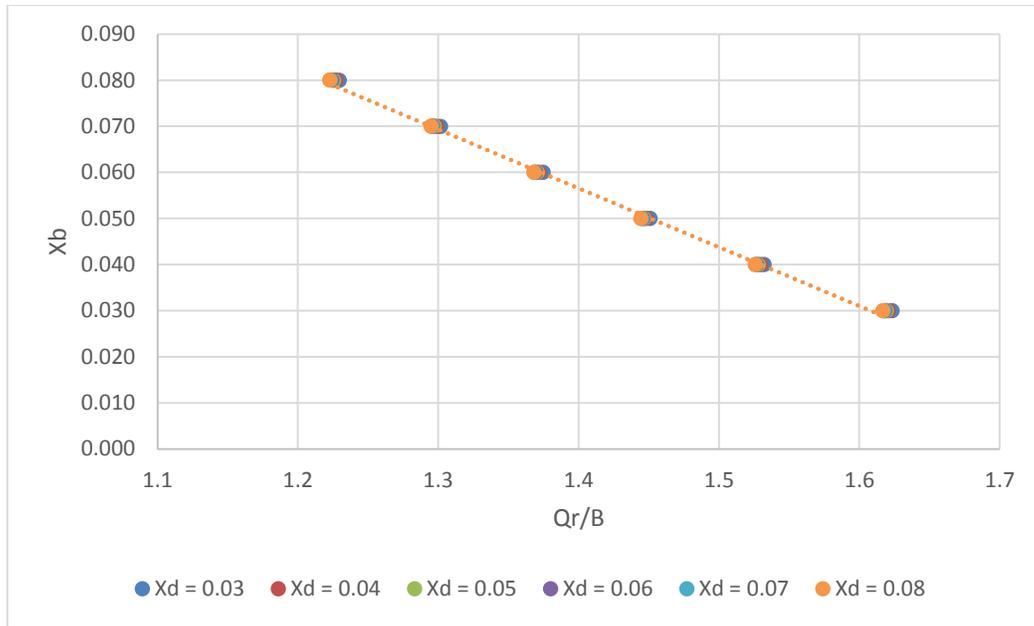


Figure 13. Relationship between  $X_b$  and  $Q_r/B$  for constant feed composition and variable  $X_d$

Table 6. Slope and intercept for  $X_b$  equation at specified  $X_d$  at constant feed composition

$X_d$	Slope a	Intercept b
0.030	-0.12761	0.23603
0.040	-0.12763	0.23572
0.050	-0.12758	0.23544
0.060	-0.12758	0.23532
0.070	-0.12758	0.23522
0.080	-0.12757	0.23513

Hence a single regression line can represent  $X_b$  equation for constant feed composition.

The relationship  $X_b = X_b(Q_r/B)$  is almost linear; however quadratic regression accounts for the slight deviation at low impurity level such as  $X_b=0.03$  which is a very important

point. It is highly recommended to ensure model accuracy at low impurity levels to accurately optimize the model toward minimum impurities. Below is a comparison between single linear as Eq. (4) and quadratic regression line as Eq. (10). Below is Eq. (4) slope and intercept values at constant feed composition which is demonstrated at Fig. 14.

*Slope a = -0.1276      Intercept b = 0.2354* for variable Xd

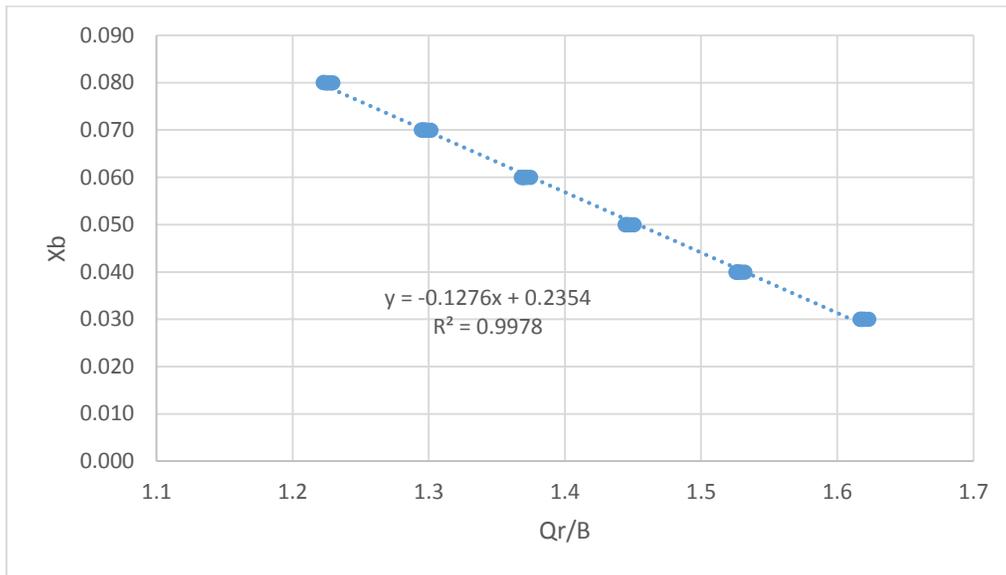


Figure 14. Relationship between Xb and Qr/B with single linear regression line for constant feed composition and variable Xd

Xb equation with single linear regression line was tested for 36 case of Xd & Xb combination as per the experimental plan. It showed has  $R^2 = 0.9978$  and RMSEE= 0.0010 with Xb average value = 0.055. The relative error% as high as 4% and the maximum absolute error = 0.0018 as shown in Fig. 15.

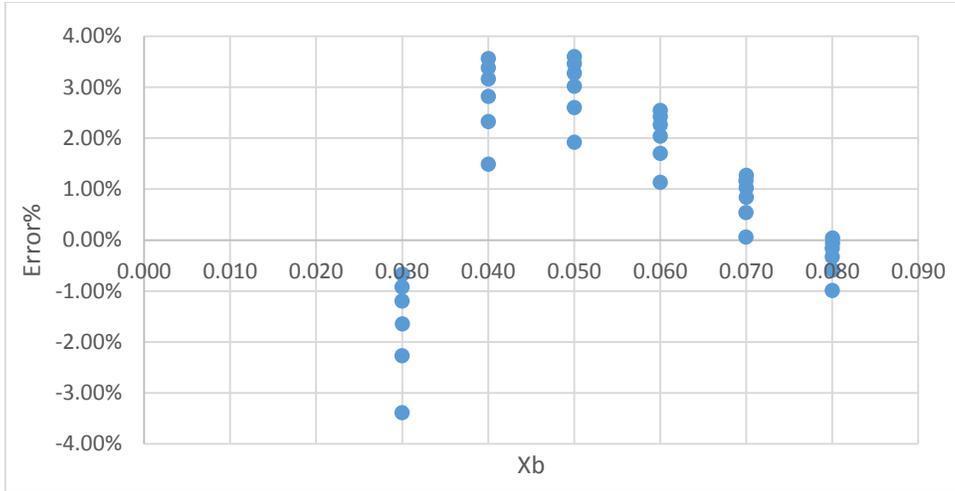


Figure 15. Relative error% for Xb equation with single linear regression line at constant feed composition and variable Xd

The quadratic regression is represented by Eq. (10) and demonstrated in Fig. 16.

$$x_b = C1 \left(\frac{Q_r}{B}\right)^2 + C2 \left(\frac{Q_r}{B}\right) + C3 \quad (10)$$

$$C1 = 0.0477 \quad C2 = -0.2632 \quad C3 = 0.331$$

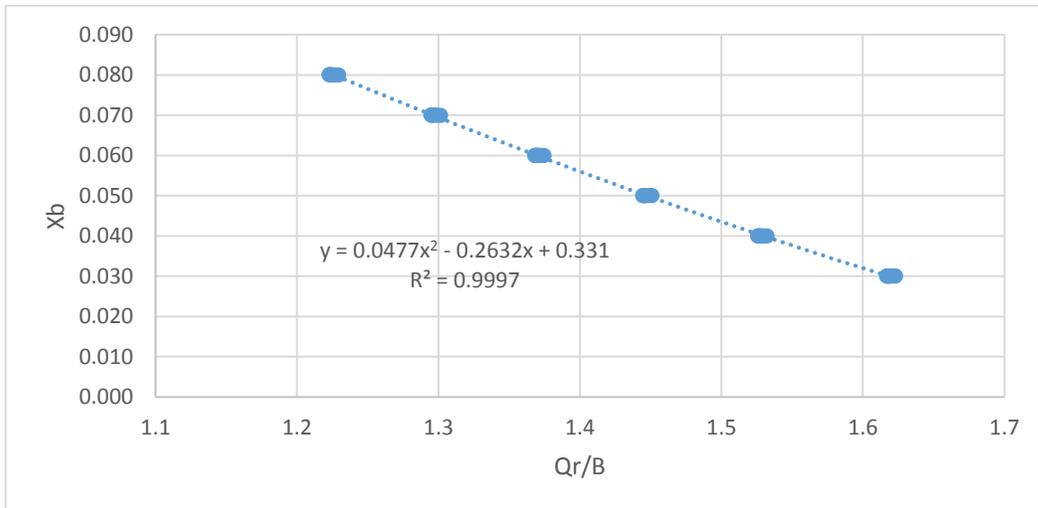


Figure 16. Relationship between Xb and Qr/B with single quadratic regression line for constant feed composition and variable Xd

Xb equation with single quadratic regression line was tested for 36 case of Xd & Xb combination. It showed  $R^2 = 0.9997$  and RMSEE= 0.000316 with Xb average value = 0.055. The relative error% not higher than 2% and the maximum absolute error = 0.0007 as shown in Fig. 17.

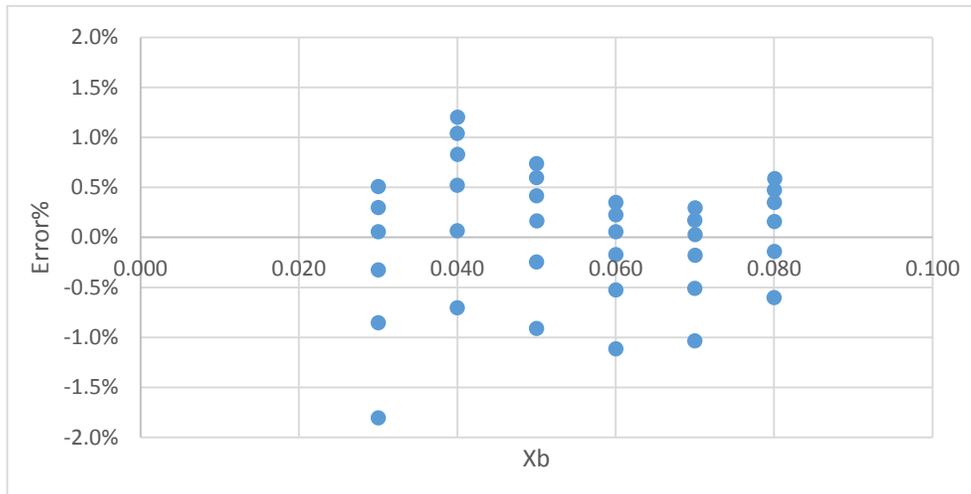


Figure 17. Relative error% for Xb equation with single quadratic regression line at constant feed composition and variable Xd

Below in (Table. 7) is a comparison between the linear and quadratic regression for Xb equation. Quadratic regression has higher  $R^2$  and lower RMSEE. In addition, RMSEE is reduced by 60% when using quadratic regression over linear regression. Hence,  $X_b = X_b(Q_r/B)$  is better represented by a single quadratic regression line for constant feed composition and variable Xd

Table 7. Comparison between Xb equation linear and quadratic regression for constant feed composition

Equation	$x_b = a \left(\frac{Q_r}{B}\right) + b$ <i>a &amp; b constants</i>	$x_b = C1 \left(\frac{Q_r}{B}\right)^2 + C2 \left(\frac{Q_r}{B}\right) + C3$ <i>C1, C2 &amp; C3 constants</i>
R <sup>2</sup>	0.9978	0.9997
RMSEE	0.000807	0.000316
Average Value	0.055	0.055

There are two models for Xd & Xb as shown earlier. Xb model has a single equations as it is not affected by Xd variation while Xd model has three equations to account for Xb variation. Xd or Xb can be calculated from the empirical equations, and the remaining compositions can be calculated from component mass balances. Below (Table. 8) is comparison between Xd and Xb models.

Table 8. Comparison between Xd & Xb model for constant feed composition and variable Xd & Xb

Model summary equations	RMSEE	Average Value
$x_d = a \left(\frac{R}{D}\right) + b$  <i>Slope a = C1x<sub>b</sub><sup>2</sup> + C2 x<sub>b</sub> + C3</i>  <i>Intercept b = C4 x<sub>b</sub><sup>2</sup> + C5 x<sub>b</sub> + C6</i>	0.00062	0.055
$x_b = C1 \left(\frac{Q_r}{B}\right)^2 + C2 \left(\frac{Q_r}{B}\right) + C3$  <i>C1, C2 &amp; C3 constants</i>	0.000316	0.055

Xb model has RMSEE = 0.000316 which is 50% less than Xd equations for the same operating range. Xb can be calculated from the above developed model while Xd can be calculated from equations (10) to (16). Xd and Xb are expressed in volume basis which is the preferred output format at industrial plants; however in order to use the component mass balance Xd & Xb required to be in mass basis. The composition conversion equation provided.

$$\text{Material balance } F = D + B \quad (11)$$

$$\text{Energy balance } H_F F + Q_r = H_D D + H_B B + Q_c \quad (12)$$

$H_F, H_D, \& H_B$  (assumed constant)

$$x_{i \text{ mass basis}} = \frac{x_{i \text{ vol basis}} \times \rho_i}{\sum x_{i \text{ vol basis}} \times \rho_i} \quad i = IC4 \& NC4 \quad (13)$$

$\rho_i$  (assumed constant)

$$LK_{\text{mass basis}} \times F = (1 - x_{d \text{ mass basis}}) D + x_{b \text{ mass basis}} B \quad (14)$$

$$HK_{\text{mass basis}} \times F = x_{d \text{ mass basis}} D + (1 - x_{b \text{ mass basis}}) B \quad (15)$$

$$x_{d \text{ mass basis}} = \frac{HK_{\text{mass basis}} \times F - (1 - x_{b \text{ mass basis}}) B}{D} \quad (16)$$

Xb empirical model was developed using eq. (10) for Feed 1 – 4 as shown in Fig. 18.

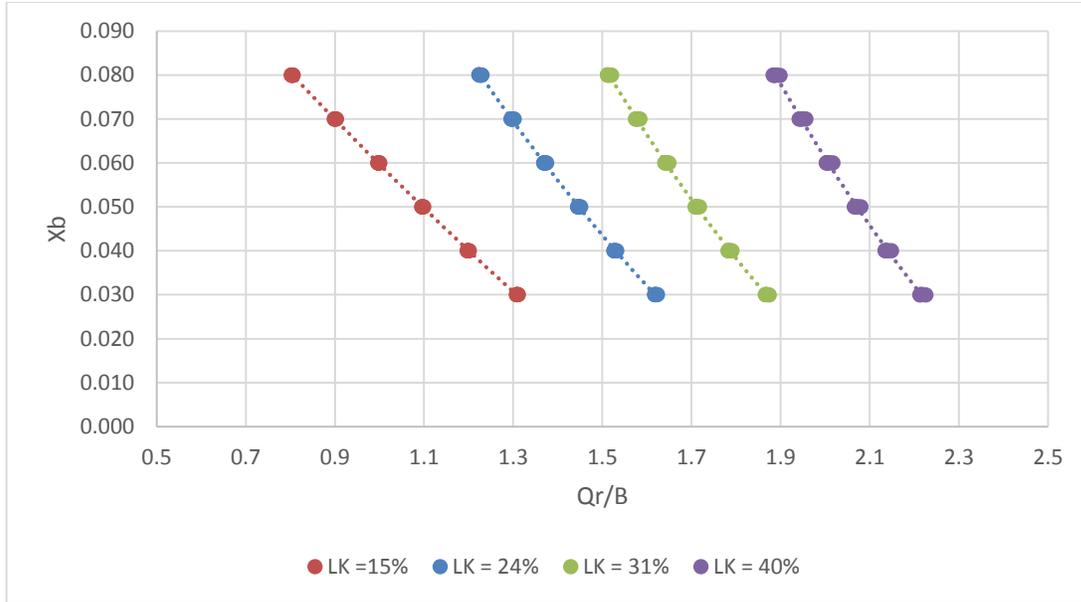


Figure 18. Xb equation with single quadratic line for variable feed compositions and Xd

Each line represents  $X_b = X_b(Q_r/B)$  for specific feed compositions in which the coefficients C1, C2 & C3 are changing due to the feed composition variation as shown (Table. 9).

Table 9. Xb equation coefficients for variable feed composition & Xd

LK%	$x_b = C1 \left(\frac{Q_r}{B}\right)^2 + C2 \left(\frac{Q_r}{B}\right) + C3$		
	C1	C2	C3
15%	0.0152388	-0.131582	0.17598701
24%	0.0476836	-0.263161	0.33098172
31%	0.0673665	-0.369633	0.48537189
40%	0.0802456	-0.481936	0.70394534

This is indicating that the coefficients are function of the changes in feed composition which can be represented by LK (iC4) in the feed. Fig. 19 - 21 show plots for individual coefficient as function of LK.

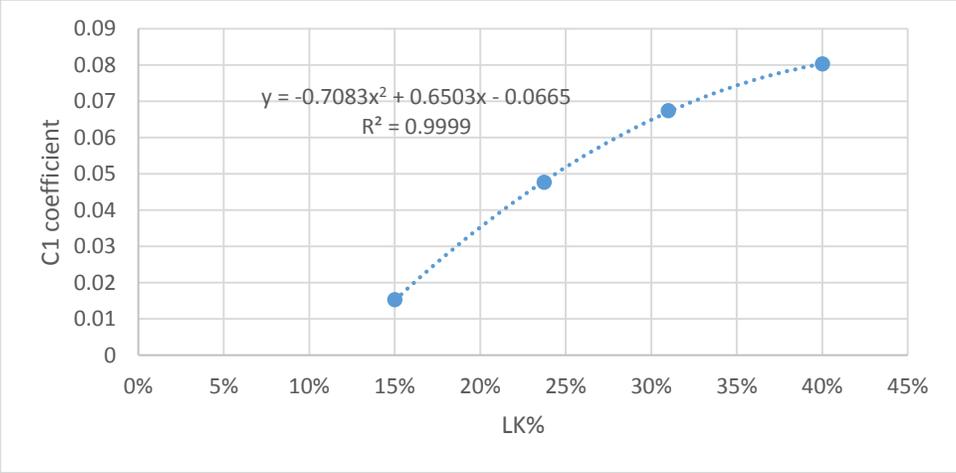


Figure 19. Relationship between C1 coefficient and feed LK% for variable feed composition & Xd

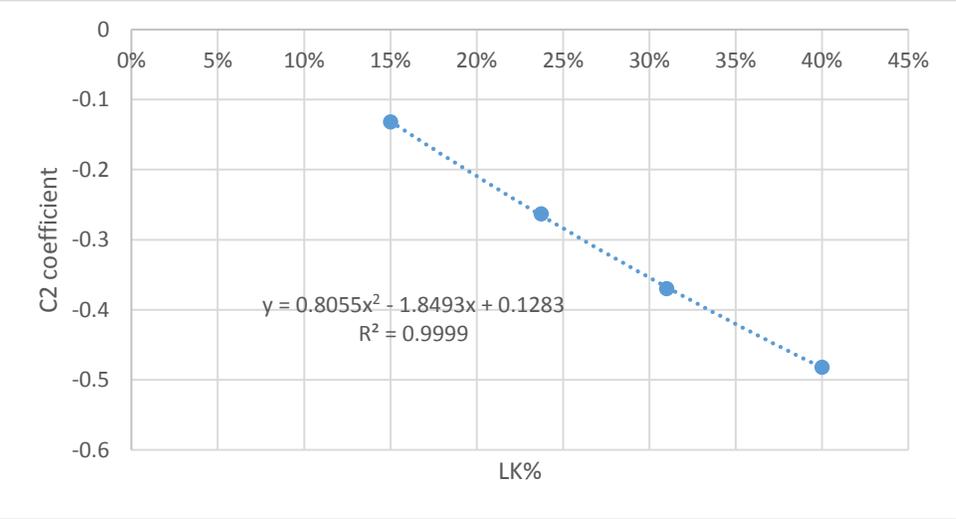


Figure 20. Relationship between C2 coefficient and feed LK% for variable feed composition & Xd

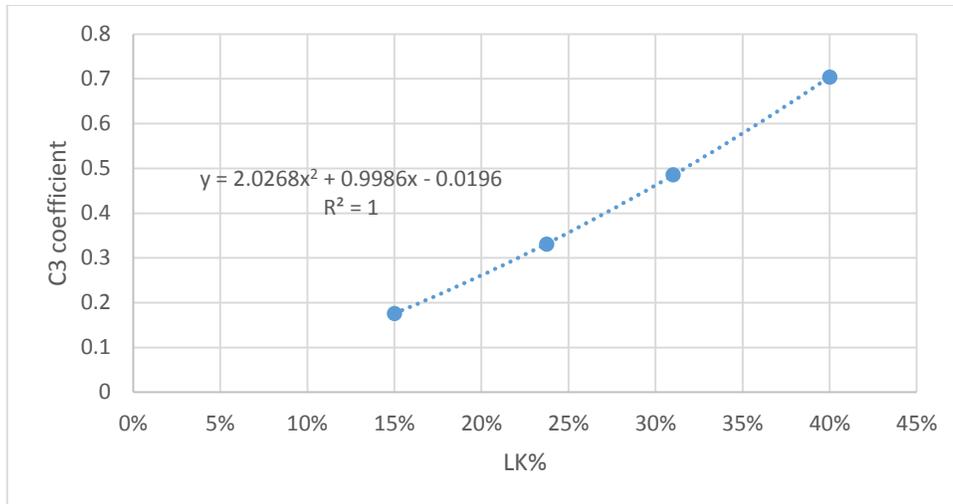


Figure 21. Relationship between C3 coefficient and feed LK% for variable feed composition & Xd

The coefficients are function of the changes in feed composition as per the below equations Eqs. (17, 18 & 19) generated from the coefficients Vs LK% plots.

$$C1 = -0.7083 (LK)^2 + 0.6503 (LK) - 0.0665 \quad (17)$$

$$C2 = 0.8055 (LK)^2 - 1.8493 (LK) + 0.1283 \quad (18)$$

$$C3 = 2.0268 (LK)^2 + 0.9986 (LK) - 0.0196 \quad (19)$$

Table 10. C1, C2 & C3 coefficients as function of LK% equations accuracy analysis

Equation	R <sup>2</sup>	RMSEE	Average value
$C1 = F(LK)$	0.9999	0.000224	0.052634
$C2 = F(LK)$	0.9999	0.001534	-0.31158
$C3 = F(LK)$	1	0.000467	0.424071

Table.10 shows the coefficients equations accuracy analysis. The  $R^2$  for the three plots are high and RMSEE is low indicating high accuracy prediction for coefficients. This is ensuring the model accuracy for variable feed composition.

The complete model for variable feed composition,  $X_d$  &  $X_b$  was tested for 144 cases training data and 30 cases prediction data set. It showed RMSEE = 0.001511 over  $X_b$  range of 0.03 – 0.08 or in other word bottom streams purity range of 0.92 – 0.97. The model showed good prediction ability with RMSEP = 0.001935 as shown in (Table. 11).

Table 11.  $X_b$  model testing results for 144 cases (training date) & 30 cases (prediction data) of variable feed composition,  $X_d$  &  $X_b$

Model Summary Equation	$x_b = C1 \left(\frac{Q_r}{B}\right)^2 + C2 \left(\frac{Q_r}{B}\right) + C3$ $C1, C2, \& C3 = F(LK)$ <i>Quadratic fitting</i>
$R^2$	0.9922
$Q^2$	0.9849
RMSEE	0.001516
RMSEP	0.001935
$X_b$ average value	0.055
Bottom purity average value	0.945
Polynomial degree	4

### 3.4.4 Xd & Xb Models for High Purity Range

The below analysis was conducted at constant feed composition and selected feed composition LK% = 31% as an example. This is to evaluate the column behavior at variable top and bottom operation. Then evaluated the column behavior at variable feed composition. Fig. 22 indicates Xd equation as Eq. (20) with exponential regression at variable Xb. While Fig. 23 indicates Xd equation as Eq. (21) where logarithmic ln applied to have a linear Xd equation with slope  $a$  and intercept  $b$ .

$$x_d = a e^{-b\left(\frac{R}{D}\right)} \quad (20)$$

$$\text{Then, } \ln(x_d) = a \left(\frac{R}{D}\right) + b \quad (21)$$

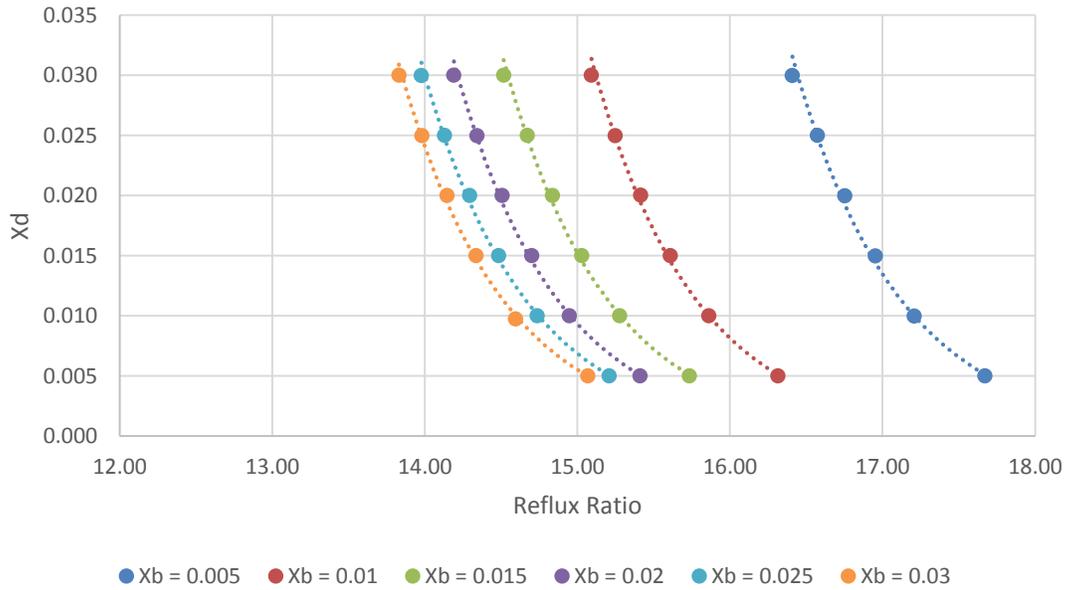


Figure 22. Relationship between Xd and reflux ratio for constant feed composition and variable Xb

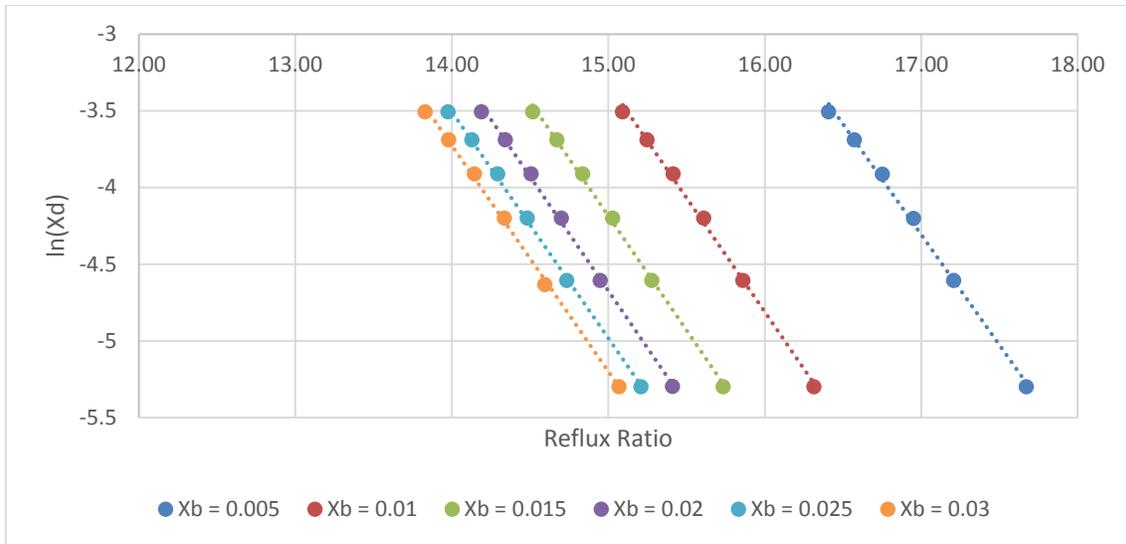


Figure 23. Relationship between  $\ln(X_d)$  and reflux ratio for constant feed composition and variable  $X_b$

$X_d$  equation experience shifting to different regions due to the changes in  $X_b$ . Slope  $a$  and intercept  $b$  are functions of  $X_b$ , which is expected to be nonlinear due to the fact that  $X_b$  changes cause higher magnitude of shifting at lower  $X_b$  values such as the operating equation at  $X_b = 0.005$ . Table. 12 show  $\ln(X_d)$  equation slope and intercept values for variable  $X_b$ .

Table 12.  $\ln(X_d) = F(R/D)$  Slope and intercept values for variable  $X_b$  at constant feed composition

$X_b$	$a$	$b$
0.005	-1.4391	20.15787
0.01	-1.48433	18.93881
0.015	-1.4912	18.18257
0.02	-1.48694	17.63032
0.025	-1.47558	17.15153
0.03	-1.47231	16.88533

Below are slope and intercept equations as Eqs. (8a & 9a) as function of Xb. Table.13 shows Eqs. (8a & 9a) accuracy analysis.

$$\text{Slope } a = C1x_b^2 + C2 x_b + C3 \quad (8a)$$

$$C1 = 225.28 \quad C2 = - 8.6593 \quad C3 = - 1.4088$$

$$\text{Intercept } b = C4 x_b^2 + C5 x_b + C6 \quad (9a)$$

$$C4 = 4195.8 \quad C5 = - 274.15 \quad C6 = 21.364$$

Table 13. Slope and intercept as function of Xb fitting analysis

	Slope a = F(Xb)	Intercept b = F(Xb)
R <sup>2</sup> Quadratic fitting	0.809	0.9972
RMSEE	0.007544417	0.058776668
Average Value	-1.474909271	18.15773975

Xd model with slope and intercept function of Xb was tested for 36 cases of Xd & Xb combinations over the range 0.005 – 0.03. The relative error is as high as 38% while maximum absolute error is 0.0114. The model has RMSEE = 0.0035284 with Xd average value = 0.0175. This is indicating high error and hence Xd model is not reliable. Xd can be calculated through component mass balance equations as explained earlier.

The below plots Figs. 24 and 25 shows Xb & ln(Xb) equations Eqs. (22) and (23) at constant feed composition and variable Xd.

$$x_b = a e^{-b\left(\frac{Qr}{B}\right)} \quad (22)$$

$$\text{Then, } \ln(x_b) = a \left(\frac{Qr}{B}\right) + b \quad (23)$$

Xb & ln(Xb) equations experience shifting due to the changes in Xd. Slope  $a$  &  $b$  are function of the changes in Xd which is expected to be nonlinear due to the fact the Xd changes cause higher magnitude of shifting at lower Xd values such as Xb equation at Xd = 0.005. Table.14 shows Eq. (23) slope and intercept values for variable Xb values.

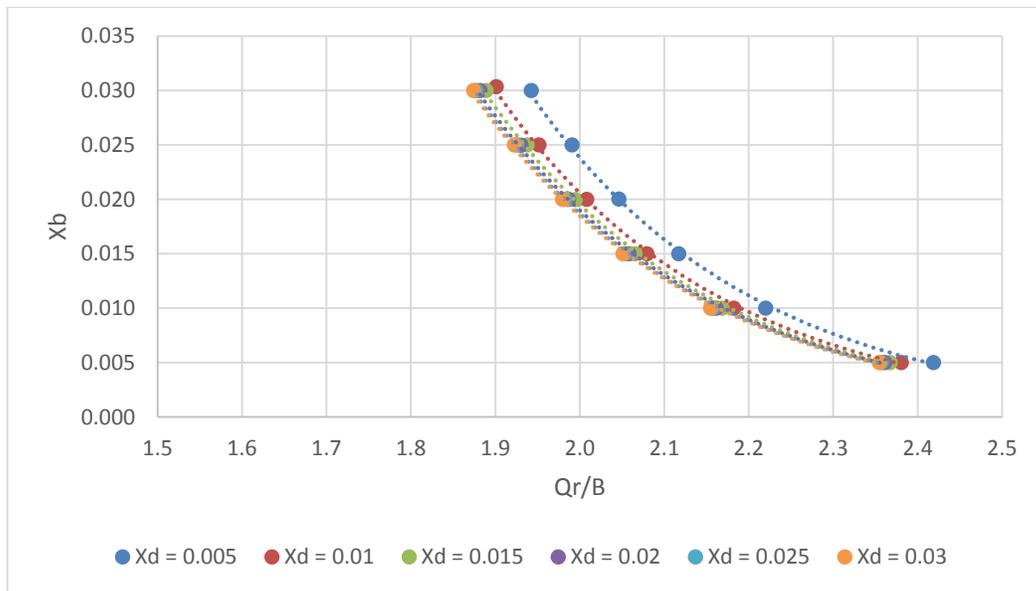


Figure 24. Relationship between Xb and Qr/B for Constant feed composition and variable Xd

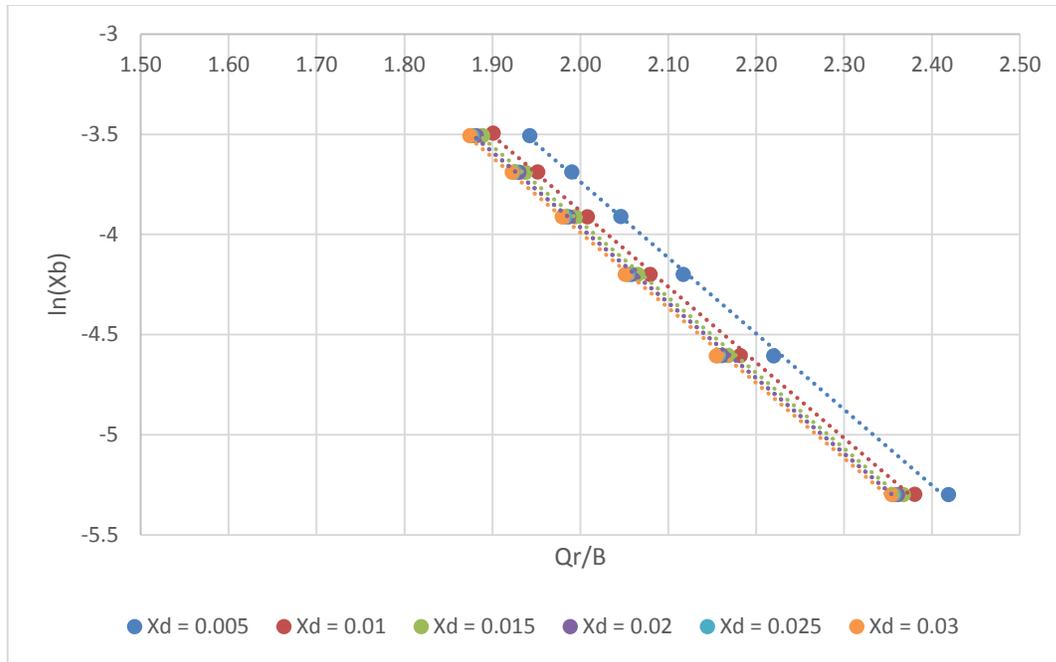


Figure 25. Relationship between  $\ln(X_b)$  and  $Q_r/B$  for Constant feed composition and variable  $X_d$

Table 14. Slope  $a$  and intercept  $b$  values for  $\ln(X_b)$  equation for variable  $X_d$  at constant feed composition

$X_d$	Slope $a$	Intercept $b$
0.005	-3.78735	3.837015
0.01	-3.78013	3.676985
0.015	-3.77048	3.604143
0.02	-3.76367	3.56313
0.025	-3.75845	3.536454
0.03	-3.75457	3.518067

The slope  $a$  & intercept  $b$  are function of  $X_d$  variation and can be modeled as shown in the below Figs. 26 and 27.

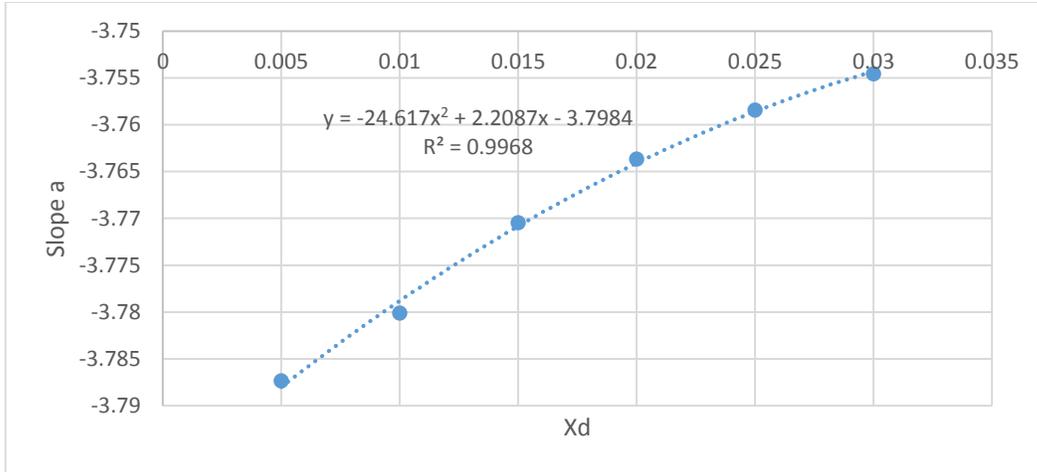


Figure 26. Relationship between slope  $a$  and  $X_d$  at constant feed composition

Below is equation Eq. (24) for slope  $a$  as function of  $X_d$  which indicates the magnitude of shifting caused by  $X_d$  variation.

$$\text{Slope } a = d1 x_d^2 + d2 x_d + d3 \quad (24)$$

$$d1 = -24.617 \quad d2 = 2.2087 \quad d3 = -3.7984$$

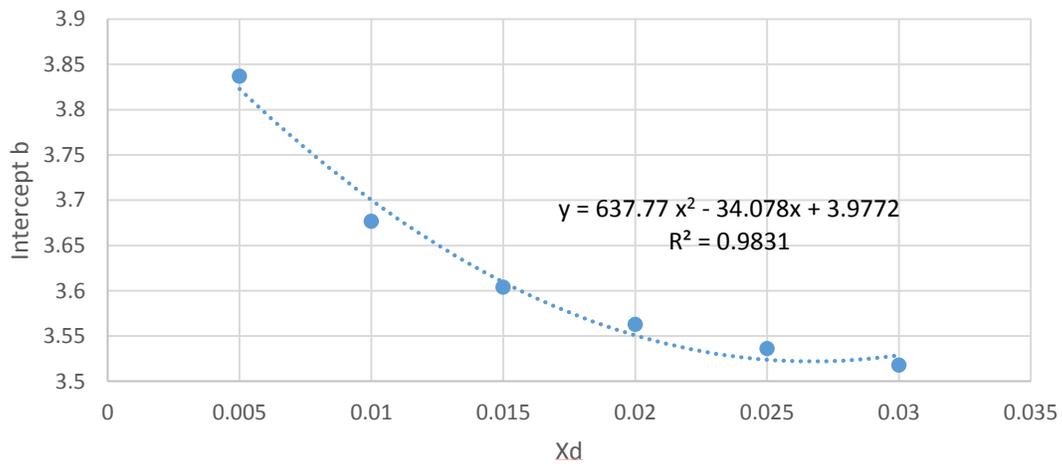


Figure 27. Relationship between intercept  $b$  and  $X_d$  at constant feed composition

Below is equation Eq. (25) for intercept  $b$  as function of  $X_d$  which indicates the magnitude of shifting caused by  $X_d$  variation. Table. 15 shows Eq. (24) and (25) accuracy analysis.

$$\text{Intercept } b = d4 x_d^2 + d5 x_d + d6 \quad (25)$$

$$d4 = 637.77 \quad d5 = -34.078 \quad d6 = 3.9772$$

Table 15. Slope and intercept as function of  $X_d$  fitting analysis

	Slope a = F( $X_d$ )	Intercept b = F( $X_d$ )
$R^2$ Quadratic fitting	0.9968	0.9831
RMSEE	0.00065471	0.014139537
Average Value	-3.76911	3.622632283

$X_b$  model with slope and intercept function of  $X_d$  was tested for 36 case of  $X_d$  &  $X_b$  combination over the range 0.005 – 0.03. The relative error as high as 6% while maximum absolute error is 0.00085 which is very small value. The model has RMSEE = 0.0004021 with  $X_b$  average value = 0.0175 and average purity stream = 0.9825. This is indicating high accuracy model. Table. 16 shows a comparison between  $X_d$  and  $X_b$  models.  $X_b$  model showed higher accuracy with much lower RMSEE as compared to  $X_d$  model. Hence  $X_b$  can be calculated using the developed equations in this report while  $X_d$  can be calculated using equation of mass and energy balance as demonstrated earlier.

Table 16. Comparison between Xd & Xb models accuracy for 36 cases each at constant feed composition

Model Summary Equation	RMSEE	Average Value	Stream purity
$x_d = a e^{-b(\frac{R}{D})}$ $\ln(x_d) = a \left(\frac{R}{D}\right) + b$ <p>Slope <math>a = C1x_b^2 + C2 x_b + C3</math>                      Intercept <math>b = C4 x_b^2 + C5 x_b + C6</math></p>	0.003528405	0.0175	0.9825
$x_b = a e^{-b(\frac{Q_r}{B})}$ $\ln(x_b) = a \left(\frac{Q_r}{B}\right) + b$ <p>Slope <math>a = d1 x_d^2 + d2 x_d + d3</math>                      Intercept <math>b = d4 x_d^2 + d5 x_d + d6</math></p>	0.0004021	0.0175	0.9825

The modeling approach requires modeling for Xb only while Xd can be calculated using mass, energy and component mass balance equation. The same analysis for developing Xb model were carried out for the other three feed composition (LK=15%, LK=24% & LK=40%). The below plot as Fig. 28 shows the raw data for 144 run for four different feed compositions and variable Xd & Xb. The first modeling part for Xb is dealing with Xb equation shifting due to Xd variation while the next step is to model Xb equation shifting due to feed variation.

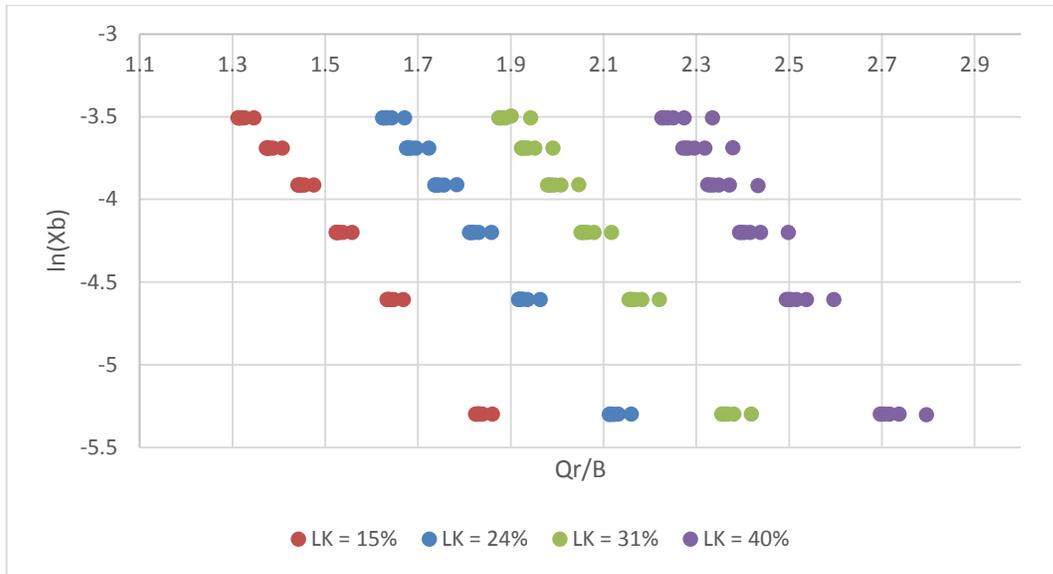


Figure 28. Relationship between  $\ln(X_b)$  and  $Q_r/B$  for variable feed composition and  $X_d$

As shown in Fig. 28 there are two dimensions of shifting due to  $X_d$  variation and feed variation.  $X_d$  variation was modeled by introducing the slope and intercept equations as function of  $X_d$ . The analysis were conducted for the four selected feed compositions and generated slope and intercept equations for each feed composition. It was noticed that  $X_d$  variation effect increases for high LK feed composition as shown in the above plot. In addition the slope and intercept equation coefficients  $d_1 - d_6$  are changing due to feed variations as shown in Table. 17.

Table 17. List of slope and intercept equations coefficients for variable feed composition

LK%	$\ln(x_b) = a \left(\frac{Q_r}{B}\right) + b$					
	Slope $a = d1 x_d^2 + d2 x_d + d3$			Intercept $b = d4 x_d^2 + d5 x_d + d6$		
	d1	d2	d3	d4	d5	d6
15%	31.6581	-1.50119	-3.50505	236.0279	-12.0062	1.2962
24%	15.6550	-0.50745	-3.68976	382.948	-19.7283	2.7423
31%	-24.6170	2.208747	-3.79843	637.7676	-34.0781	3.9772
40%	-113.479	7.511553	-3.95386	1227.263	-66.1524	5.8938

The coefficients d1 –d6 are function of light key feed composition as per the below

Figs. 29 – 34.

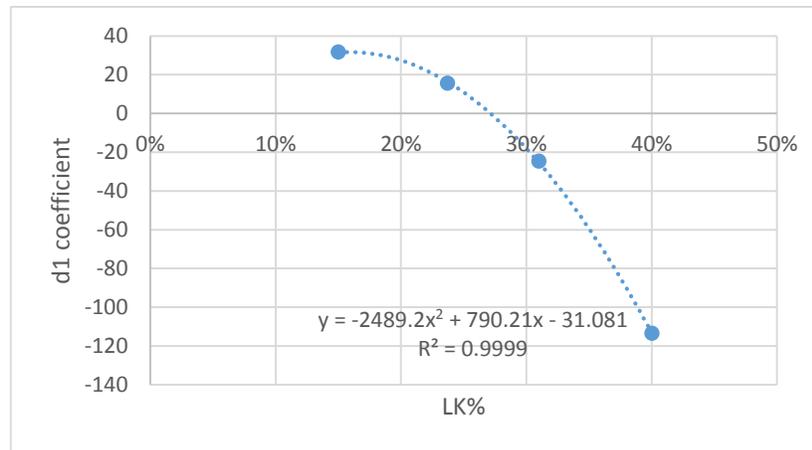


Figure 29. Relationship between d1 and feed LK for variable feed composition & Xd

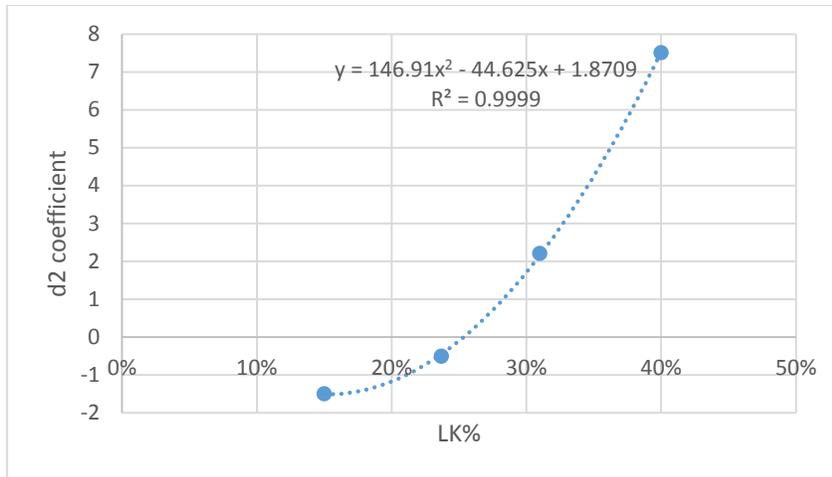


Figure 30. Relationship between d2 and feed LK for variable feed composition & Xd

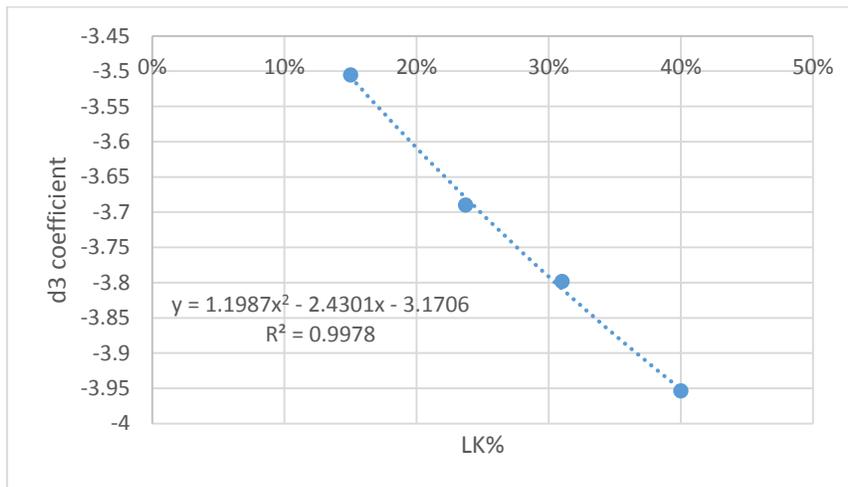


Figure 31. Relationship between d3 and feed LK for variable feed composition & Xd

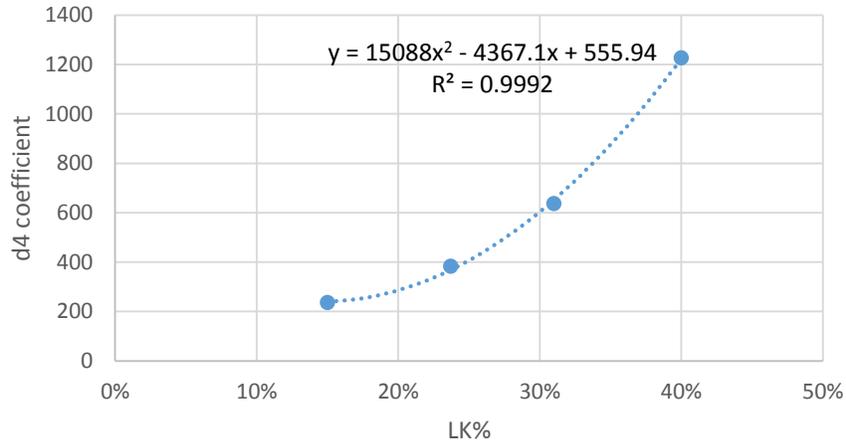


Figure 32. Relationship between d4 and feed LK for variable feed composition & Xd

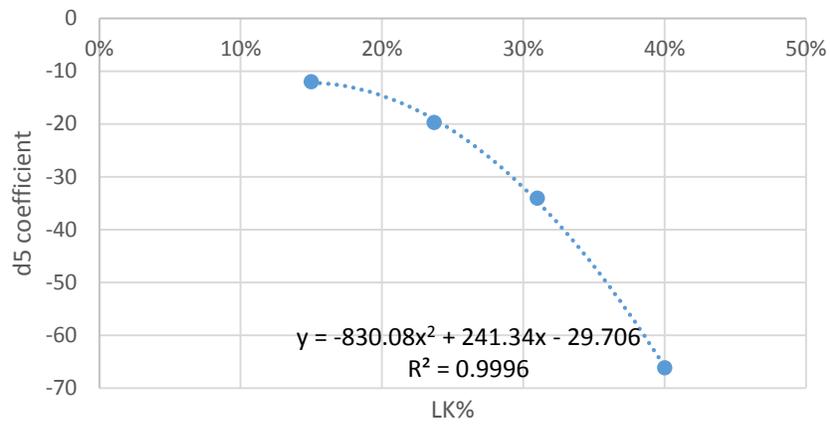


Figure 33. Relationship between d5 and feed LK for variable feed composition & Xd

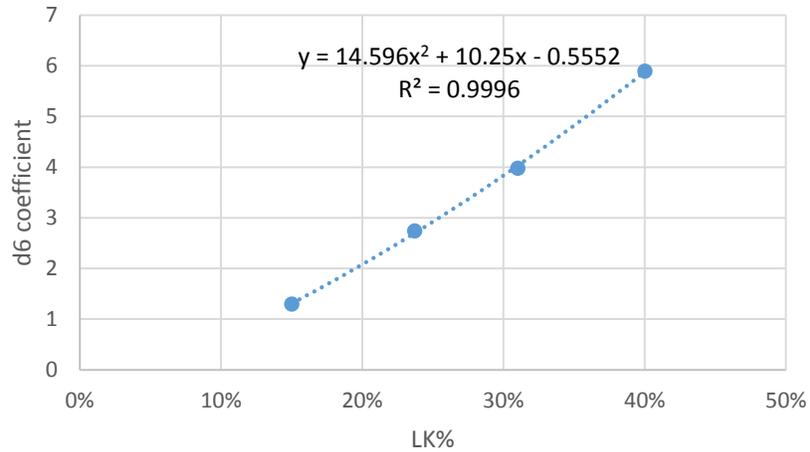


Figure 34. Relationship between d6 and feed LK for variable feed composition & Xd

As shown the coefficients are function of feed composition with high  $R^2$ . Below are the coefficient equations Eqs. (26) – (31) as function of LK. Table. 18 shows the coefficients d1 – d6 modeling equations accuracy which showed high  $R^2$  and RMSEE.

$$d1 = -2489.2 LK^2 + 790.21 LK - 31.081 \quad (26)$$

$$d2 = 146.91 LK^2 - 44.625 LK + 1.8709 \quad (27)$$

$$d3 = 1.1987 LK^2 - 2.4301 LK - 3.1706 \quad (28)$$

$$d4 = 15088 LK^2 - 4367.1 LK + 555.94 \quad (29)$$

$$d5 = -830.08 LK^2 + 241.34 LK - 29.706 \quad (30)$$

$$d6 = 14.596 LK^2 + 10.25 LK - 0.5552 \quad (31)$$

Table 18. d1 - d6 coefficients as function of LK% fitting analysis

Equation	R <sup>2</sup>	RMSEE	Average value
$d1 = F(LK)$	0.9999	0.530459	-22.6958
$d2 = F(LK)$	0.9999	0.039591	1.927916
$d3 = F(LK)$	0.9978	0.007687	-3.73678
$d4 = F(LK)$	0.9992	10.62896	621.0018
$d5 = F(LK)$	0.9996	0.434114	-32.9912
$d6 = F(LK)$	0.9996	0.035432	3.477371

Xb model consists of slope and intercepts equations to account for Xd variation. Furthermore slope and intercepts equations coefficients are function of LK which accounts for feed composition variation. Xb model was tested for 144 cases for variable feed composition and Xd in addition to 30 cases of prediction data set. It showed low RMSEE = 0.000548404 over operating range of 0.005 – 0.003 of impurity or in other words streams purity range of 0.97 – 0.995. In addition to good prediction ability with RMSEP = 0.0006074 as shown in Table. 19.

Table 19. Xb model testing results for 144 cases (training data) & 30 case (prediction data) of variable feed composition, Xd & Xb

Model Summary Equation	$\ln(x_b) = a \left(\frac{Q_r}{B}\right) + b$ <p>Slope <math>a = d1 x_d^2 + d2 x_d + d3</math></p> <p>Intercept <math>b = d4 x_d^2 + d5 x_d + d6</math></p> <p><math>d1, d2, d3, d4, d5</math> &amp; <math>d6 = F(LK)</math></p> <p>Quadratic fitting</p>
R <sup>2</sup>	0.9959
Q <sup>2</sup>	0.9934
RMSEE	0.0005484
RMSEP	0.0006074
Xb average value	0.0175
Bottom purity average value	0.9825
Polynomial degree	3

Xb model was tested for high and low purity ranges for 144 cases over each range. The bottom stream purity relative error percentage is well below 0.3% for low purity operating range and well below 0.17% for the high purity operating range. Both ranges showed low error percentages however low purity range showed higher error due to the assumption of negligible Xd impact on Xb model. Fig. 35 shows the relative error percentage for the two operating range which is compared to final bottom stream purity.

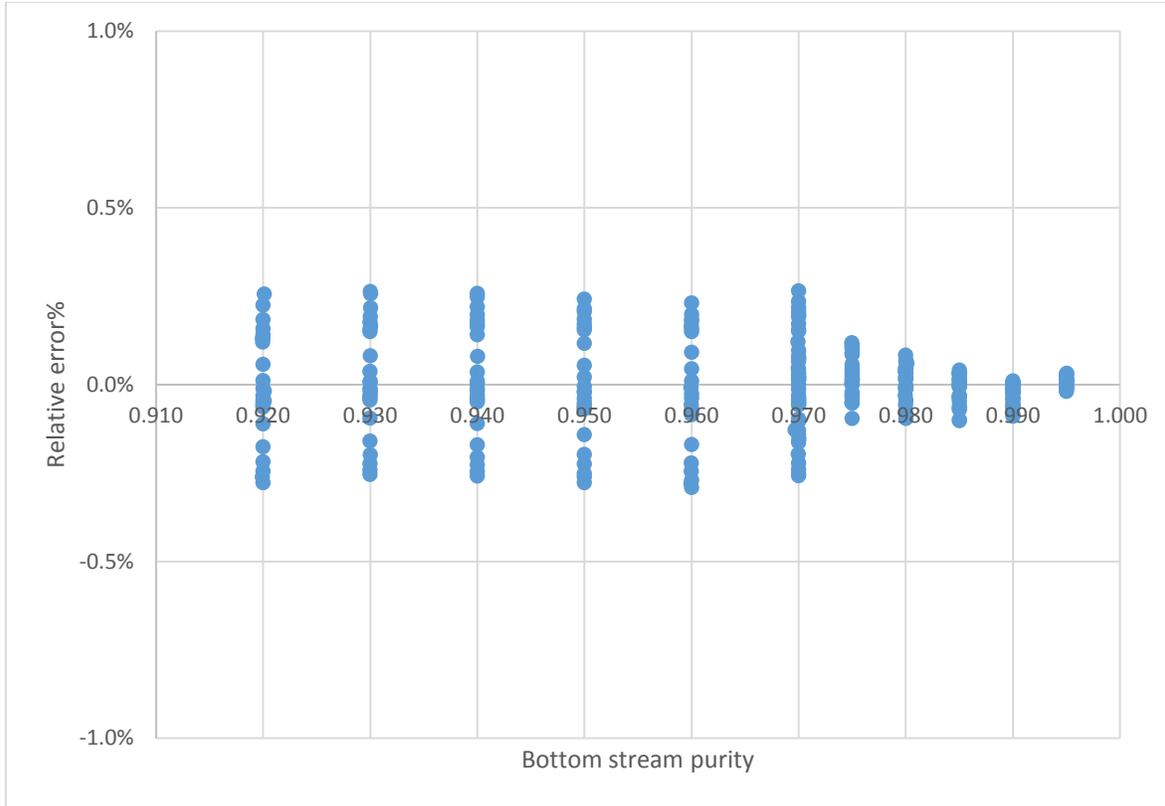


Figure 35. Relative error% compared to bottom stream purity using both low and high range purity models for variable feed composition,  $X_d$  &  $X_b$

### 3.4.5 Butane Splitter Hybrid Model Configuration

The hybrid model for C4 splitter consists of mass balance, component mass balance and energy balance equations as Eqs. (11) – (16). Then  $X_b$  models for low purity range as Eqs. (10), (17), (18) and (19). While  $X_b$  models for high purity range as Eqs. (23) – (31). The condenser duty is calculated as shown below as Eq. (32)

$$Q_c = (R + D) c_p \Delta T + \lambda_c (R + D) \quad \Delta T = T_c - T_{Top\ Tray} \quad (32)$$

Duty includes sensible and latent heat terms. The heat capacity  $c_p$  and heat of condensation  $\lambda_c$  change due to top impurity variation  $X_d$ . Furthermore the condenser

operates at saturation temperature in which  $\Delta T$  changes due to top impurity variation Xd. Hence a new term  $\lambda'_c$  introduced to evaluate  $c_p$ ,  $\lambda_c$  &  $\Delta T$  as function of Xd as Eq. (33). This yields the condenser duty equation to be as Eq. (34).

$$\lambda'_c = c_p \Delta T + \lambda_c \quad (33)$$

$$Q_c = \lambda'_c (R + D) \quad (34)$$

Eq. (35) was generated using  $\lambda'_c$  data over 288 cases of variable feed composition, Xd and Xb over the two operating ranges. Feed composition variation and Xb have negligible impact on  $\lambda'_c$  while Xd considered the main factor. The below  $\lambda'_c$  equation as Eq. (35) improves condenser duty calculations and so reflux flow.

$$\lambda'_c = 0.0267 (x_d) + 0.2821 \quad (35)$$

The same approach followed to generate equation for reboiler heat of vaporization  $\lambda_r$  changes due to bottom impurity variation Xb. Eq. (36) was generated using  $\lambda_r$  data over 288 cases of variable feed composition, Xd and Xb over the two operating ranges. Feed composition variation and Xd have negligible impact on  $\lambda_r$  while Xb considered the main factor. The below  $\lambda_r$  equation as Eq. (36) improves boilup rate  $V_r$  calculation. Boilup flow  $V_r$  is function of  $Q_r$  and  $V_r$  as Eq. (2).

$$\lambda_r = -0.0136 (x_b) + 0.2785 \quad (36)$$

The below Table. 20 illustrates  $\lambda'_c$  as Eq(35) and  $\lambda_r$  as Eq. (36) estimation equations accuracy which indicate very low RMSEE values as compared to average values.

Table 20. Heat of condensation and vaporization as function of Xd & Xb fitting analysis

	$\lambda'_c = F(x_d)$	$\lambda_r = F(x_b)$
R <sup>2</sup> Quadratic fitting	0.9997	0.9996
RMSEE	9.86233E-06	6.35852E-06
Average Value	0.283109606	0.277998696

### 3.5 Hybrid Model Development for Depropanizer Column

#### 3.5.1 Introduction

The Depropanizer column DeC3 separates Propane C3 from Butane C4 and then send distillate (vapor and liquid) and bottom streams for further processing. The C3 rich stream is fed mainly for LPG pool while C4 stream can be fed to both LPG mainly and with small portion to Gasoline pools. The DeC3 feed contains also Ethane C2 and Pentane C5 with small compositions around 2%. Butane is present in the feed in the form of iso-Butane iC4 and Normal-Butane nC4. The light key for separation is defined C3 while heavy key is defined as iC4. The column has a vapor distillate Vd which is changing in flow and composition for variable feed composition and streams purities. The DeC3 model included top impurity Xd model, bottom impurity Xb model and vapor distillate Vd model.

Below are the typical operating conditions for the DeC3 column as shown in Fig. 36 and Table. 21. The column operates with vapor distillate Vd which adds complexity to the model.

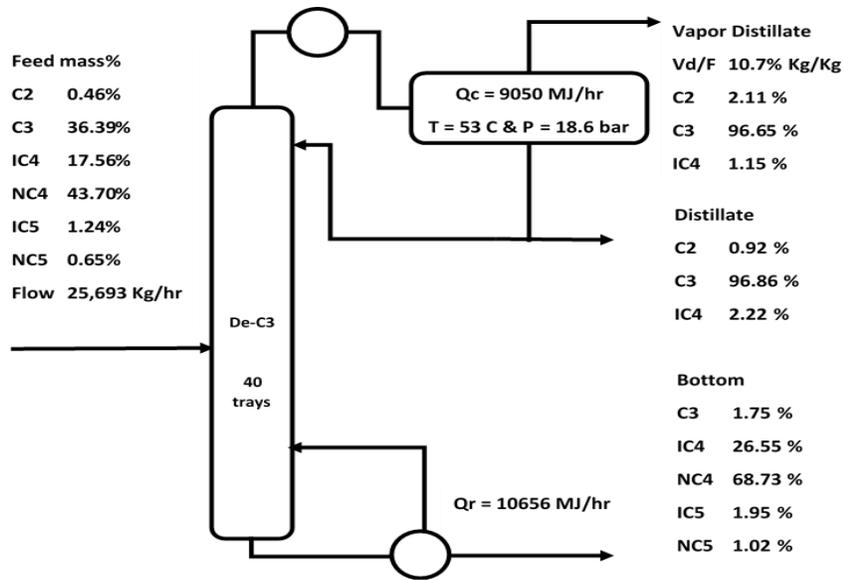


Figure 36. Depropanizer column configuration and base case operating parameters

Table 21. DeC3 base case operating parameters

Feed Temperature (C <sup>o</sup> )	77	Condenser Duty (MJ/hr)	9050.01
Feed Pressure (bar)	18.9	Condenser Temperature (C <sup>o</sup> )	53
Feed Flow (Kg/hr)	25,693	Condenser Pressure (bar)	18.6
Feed LK% (C3)	36.39 %	Reflux Ratio (Kg/Kg)	2.76
Feed HK% (IC4)	17.56 %	Reboiler Duty MJ/hr	10656.54
Xd (HK in D)	2.22 %	Boilup Ratio (Kg/Kg)	2.799
Xb (LK in B)	1.75 %	Vd / F (Kg/Kg)	10.7%
Pressure Drop (bar)	0.6	Distillate D flow (Kg/hr)	6641.81
Number of Trays	40 + Condenser + Reboiler	Bottom B flow (Kg/hr)	16312.53

### 3.5.2 Experimental Plan and Assumption

The experimental simulation conducted using Aspen Plus simulation engine with thermodynamics package Peng–Robinson equation of state. Below are the assumptions for DeC3 model.

- Constant feed temperature & Pressure (base case value)
- Constant condenser temperature and pressure (base case value)
- Constant pressure drop across the column
- Constant feed impurity compositions i.e. C2, iC5 & nC5 (base case value)
- Constant ratio of iC4 / nC4 in feed composition
- Constant enthalpy  $H^o$  and heat capacity  $c_p$  for F, Vd, D and B.

The condenser temperature is assumed constant at 53 C°. The feed impurities are assumed constant while feed composition components C3/C4 are variable. This is to evaluate the variation in C3 and C4 in the feed as they account for 98% of feed composition. The C4 in feed assumed as one component in which assumed constant ratio of iC4/nC4. The iC4/nC4 separation if required can be handled by downstream Butane Splitter column.

Below are the ranges of the variables which have been used to develop the mode.

- Feed composition for LK over range of 32.5% to 40.3% and HK over range of 15.6% to 19.5%.
- Xd distillate impurity (iC4% in D) with range (0.011 – 0.033 mass basis)
- Xb bottom impurity (C3% in B) with range (0.009 – 0.026 mass basis)
- Vapor distillate flow and composition

Changes in the above variables cause variation in the column operating conditions such as reboiler duty  $Q_r$ , condenser duty  $Q_c$ , reflux ratio, boilup ratio, distillate and bottom streams flow. Table. 22 illustrates the variable feed compositions used to develop DeC3 hybrid model while Table. 23 illustrates the experimental plan and number of simulation cases generated.

Table 22. Feed compositions selected for experiments with compositions in mass basis

Component	Feed 1	Feed 2	Feed 3 (base case)	Feed 4	Feed 5
C2	0.46%	0.46%	0.46%	0.46%	0.46%
C3 (LK)	32.54%	34.46%	36.39%	38.32%	40.27%
IC4 (HK)	19.51%	18.54%	17.56%	16.57%	15.58%
NC4	45.61%	44.66%	43.70%	42.74%	41.78%
IC5	1.24%	1.24%	1.24%	1.24%	1.25%
NC5	0.65%	0.65%	0.65%	0.65%	0.65%

Table 23. Experimental plan for developing DeC3 model

Feed variation	$X_d$ variation	$X_b$ variation	Number of experiments
5 different feed compositions (Feed 1 – Feed 5)	5 points over range (0.011 – 0.033), increment 0.0055	5 points over range (0.009 – 0.026), increment 0.004	125 cases (combination of $X_d$ , $X_b$ and feed composition points)

The models were generated using 125 cases of variable feed composition,  $X_d$  &  $X_b$  as per the experimental plan in Table. 23. In addition included 25 cases as prediction data of variable feed composition,  $X_d$  &  $X_b$ . Prediction data set was selected randomly in which they fall within the specified experimental ranges.

### 3.5.3 Model for distillate impurity, $X_d$

The analysis was conducted at two stages at constant and variable feed composition. The first stage was conducted at constant feed composition (Feed 3) with 25 combination of  $X_d$  and  $X_b$  values. It evaluated the impact from bottom operation i.e. variable  $X_b$  on  $X_d$  equation. The below plot shows  $X_d$  equation as Eq. (37) at variable  $X_b$ . It is linear relation to reflux ratio as per the below equation. Since the column has a vapor distillate the distillate stream is consider sum of vapor and liquid distillate and reflux ratio expressed with Eq. (38).

$$x_d = a \left( \frac{R}{V_{d+D}} \right) + b \quad (37)$$

$$\text{Reflux Ratio} = \frac{R}{V_{d+D}} \quad (38)$$

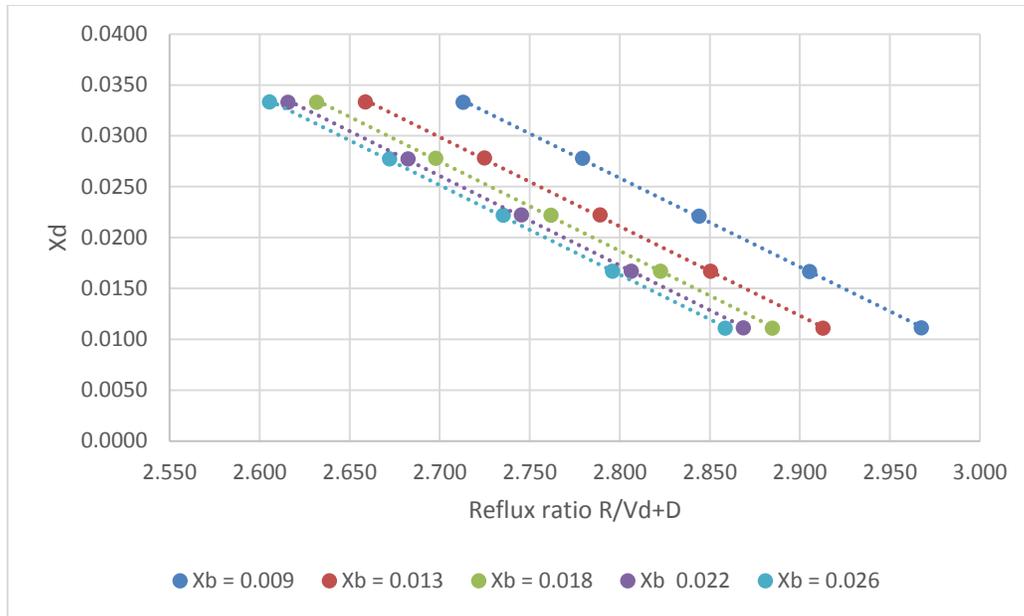


Figure 37. Relationship between  $X_d$  and reflux ratio at constant feed and variable  $X_b$

$X_d$  equation as Eq. (37) experience shifting due to  $X_b$  variation as shown in Fig. 37. This is attributed to the changes in the column energy input from reboiler duty  $Q_r$  which is causing changes in the internal column vapor flow and so changes in the top condenser duty. The shifting effect can be modeled by evaluating  $X_d$  equation slope and intercept at various  $X_b$  values as shown in Table. 24.

Table 24.  $X_d$  equation slope and intercept at five  $X_b$  values (constant feed composition)

$X_b$	slope a	Intercept b
0.00878	- 0.08738	0.27050
0.01315	- 0.08769	0.26664
0.01755	- 0.08794	0.26492
0.02193	- 0.08802	0.26371
0.02634	- 0.08809	0.26301

The slope and intercept are changing as a result of Xb variation. Although the changes are small in magnitude it can have major impact on Xd calculation considering how small the Xd values. The below plots Figs. 38 and 39 show the relationship between slope and intercept and Xb values.

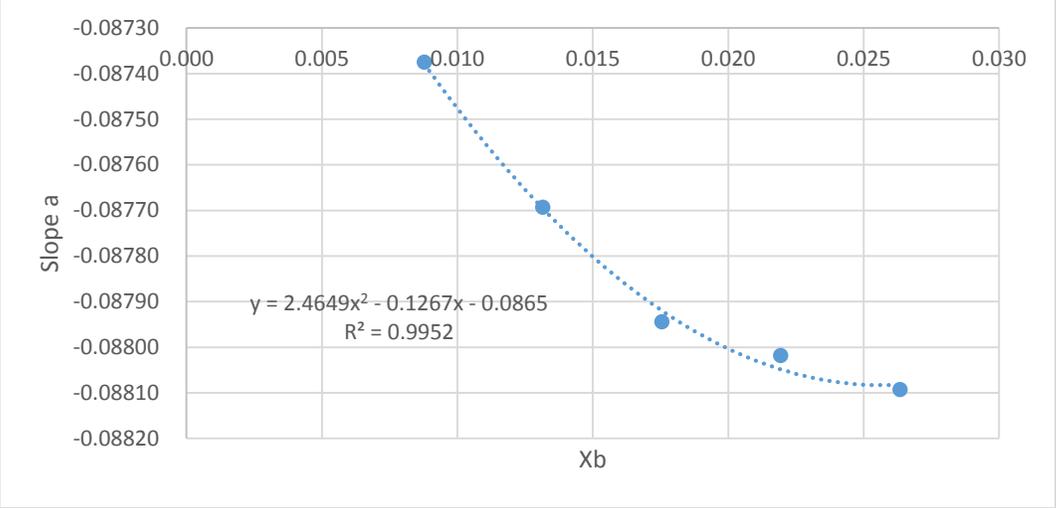


Figure 38. Relationship between slope a and Xb values for constant feed composition

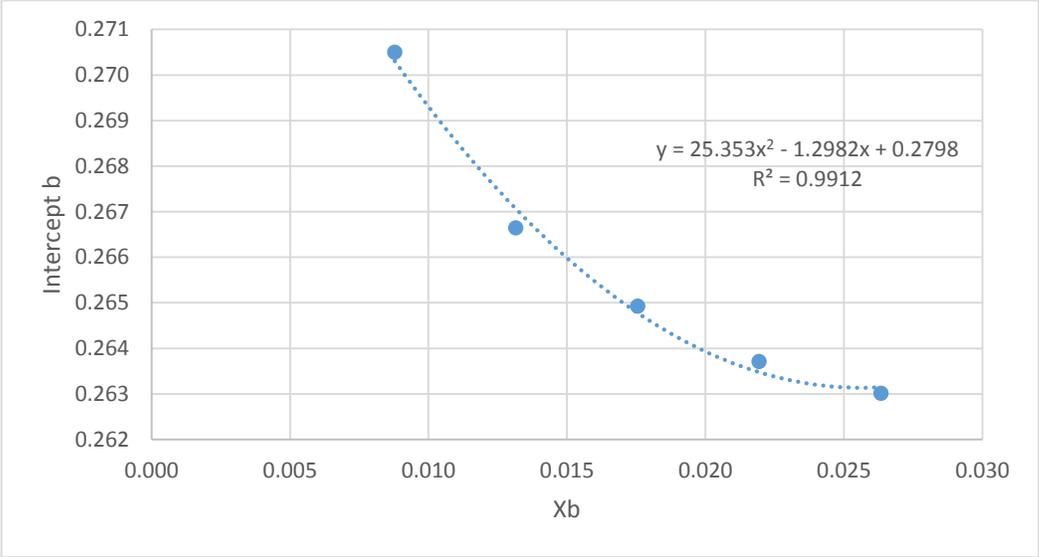


Figure 39. Relationship between intercept b and Xb values for constant feed composition

The slope and intercept are function of Xb as Eqs. (39) and (40) show high R<sup>2</sup> and low RMSEE as compared to the average value. At constant feed composition Xd model consists of Xd linear equation function of reflux ratio with slope and intercept function of Xb as shown below. Table. 25 shows slope and intercept equations Eqs. (39) and (40) accuracy analysis.

$$\text{Slope } a = C1 x_b^2 + C2 x_b + C3 \quad (39)$$

$$C1 = 2.4649363 \quad C2 = -0.1266542 \quad C3 = -0.0864563$$

$$\text{Intercept } b = C4 x_b^2 + C5 x_b + C6 \quad (40)$$

$$C4 = 25.3530114 \quad C5 = -1.298150 \quad C6 = 0.2797528$$

Table 25. Slope and intercept function of Xb fitting results

	Slope a = F(Xb)	Intercept b = F(Xb)
R <sup>2</sup> Quadratic fitting	0.9952	0.9912
RMSEE	1.8207E-05	0.00025063
Average Value	-0.08782	0.26576

The same analysis for Xd model development conducted for the remaining feed compositions Feed 1, Feed 2, Feed 4 and Feed 5. Fig. 40 illustrates the relationship between Xd and reflux ratio as Eq. (37) for various feed composition and Xb. The first stage of Xd modeling dealt with shifting caused by Xb variation. While the second stage deals with the shifting effect caused by feed composition variation.

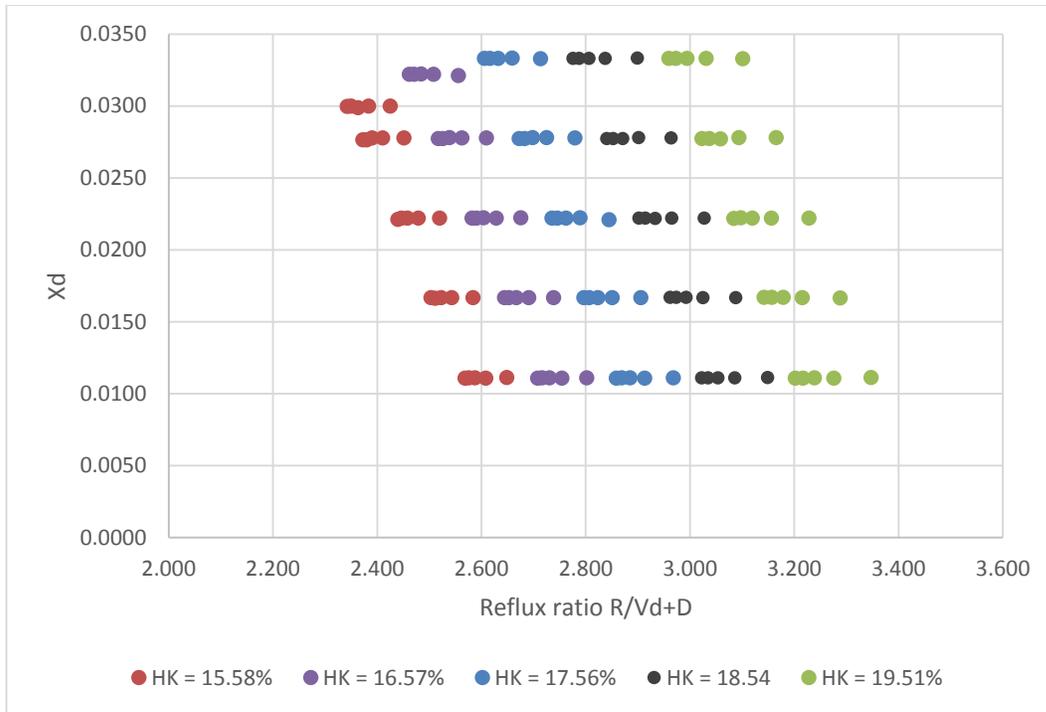


Figure 40. Relationship between  $X_d$  and reflux ratio at variable feed composition and  $X_b$

Aspen Plus simulation did not converge at high  $X_d$  for low HK% as shown for cases of  $HK = 15.58\%$  &  $16.57\%$ . Feed heavy key composition impacts the top vapor composition and so impacts the condenser saturation temperature. This indicating the max  $X_d$  value at the specified condenser temperature for the selected feed composition. The model focuses toward minimizing  $X_d$  values.

The slope and intercept equations as Eqs. (39) and (40) coefficients are changing due to feed composition variation. Table. 26 shows slope and intercept equations for each feed composition.

Table 26. Slope and intercept equation coefficients for variable feed composition

HK%	$x_d = a \left( \frac{R}{V_d + D} \right) + b$					
	Slope $a = C1 x_b^2 + C2 x_b + C3$			Intercept $b = C4 x_b^2 + C5 x_b + C6$		
	C1	C2	C3	C4	C5	C6
15.58%	-1.5305	0.0748	-0.0848	27.4843	-1.38794	0.2442
16.57%	1.3498	-0.0662	-0.0852	24.0878	-1.2351	0.2601
17.56%	2.4649	-0.1267	-0.0865	25.3530	-1.2982	0.2798
18.54%	2.9794	-0.1681	-0.0877	28.6114	-1.4304	0.3015
19.51%	3.9676	-0.2272	-0.0884	30.6257	-1.5159	0.3234

Figures 41 – 46 show the relationship between each coefficient and heavy key composition in the feed.

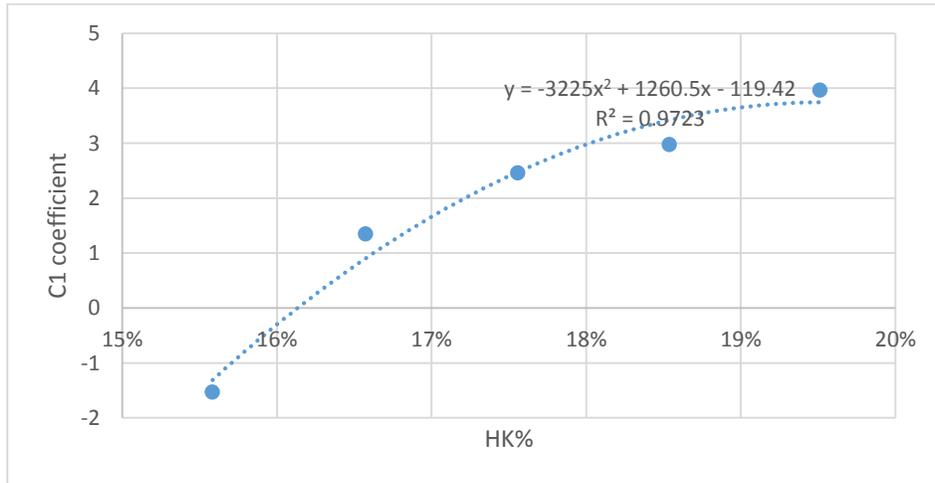


Figure 41. C1 is a quadratic function of HK for variable feed composition and Xb

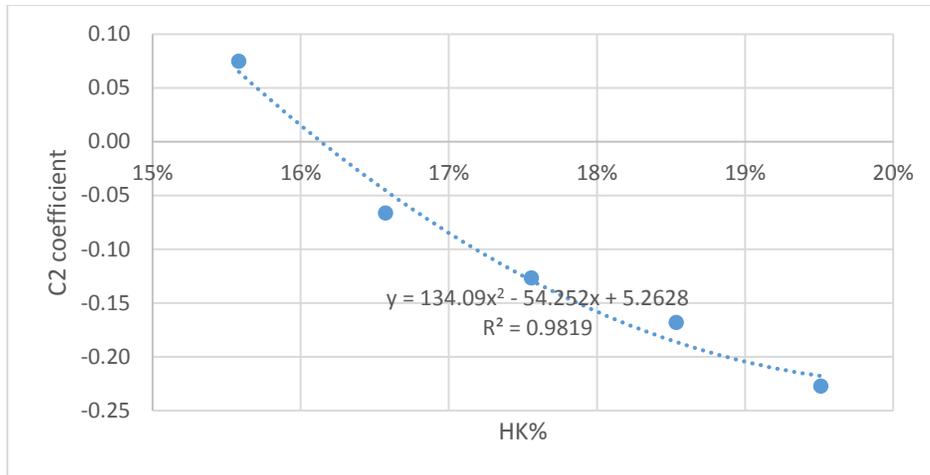


Figure 42. C2 is a quadratic function of HK for variable feed composition and Xb

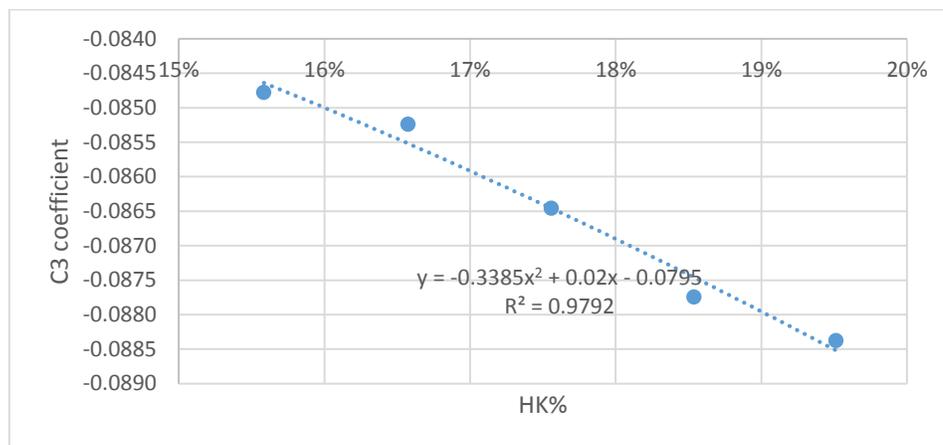


Figure 43. C3 is a quadratic function of HK for variable feed composition and Xb

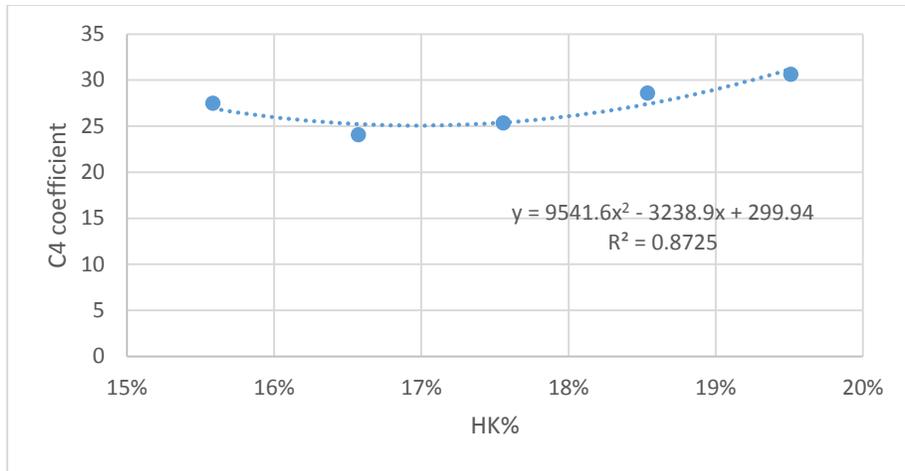


Figure 44. C4 is a quadratic function of HK for variable feed composition and Xb

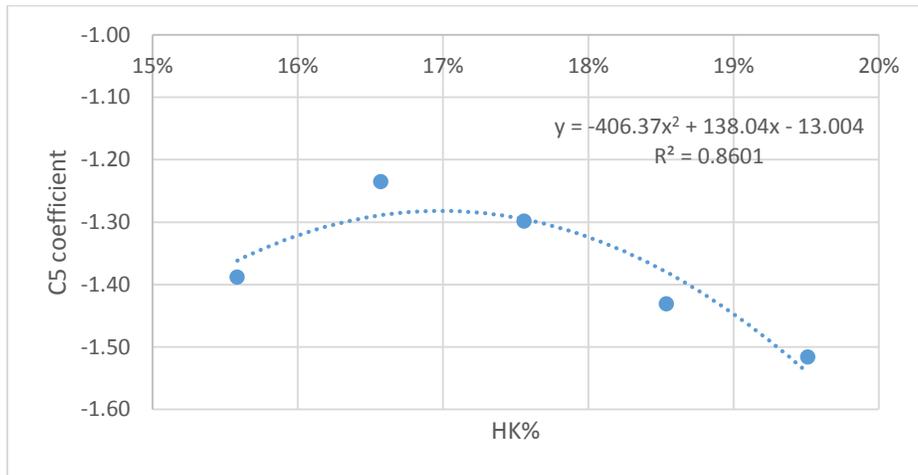


Figure 45. C5 is a quadratic function of HK for variable feed composition and Xb

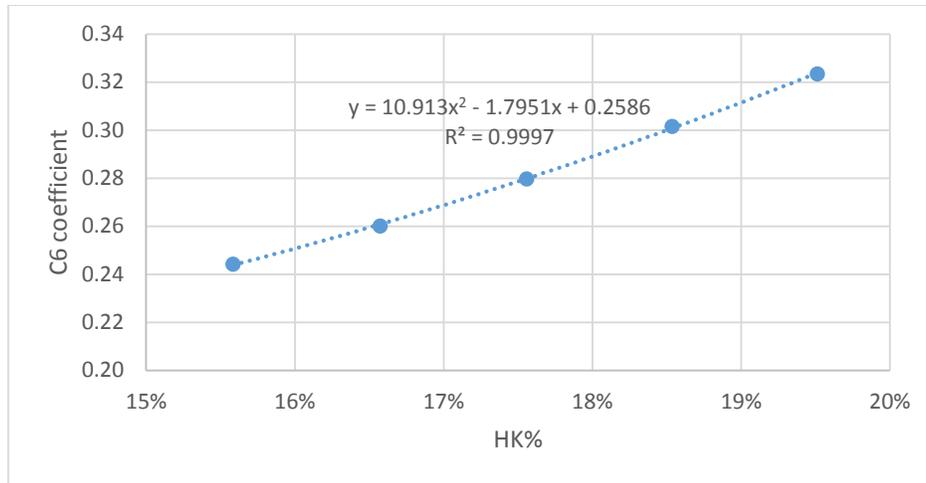


Figure 46. C6 is a quadratic function of HK for variable feed composition and Xb

Below are the generated coefficients equations as function of feed HK as Eqs. (41) – (46). Table.27 shows coefficients equations as Eqs. (41) – (46) accuracy analysis.

$$C1 = -3225 (HK)^2 + 1260.5(HK) - 119.42 \quad (41)$$

$$C2 = 134.09 (HK)^2 - 54.252(HK) + 5.2628 \quad (42)$$

$$C3 = -0.3385 (HK)^2 + 0.02(HK) - 0.0795 \quad (43)$$

$$C4 = 9541.6 (HK)^2 - 3238.9(HK) + 299.94 \quad (44)$$

$$C5 = -406.37 (HK)^2 + 138.04(HK) - 13.004 \quad (45)$$

$$C6 = 10.913 (HK)^2 - 1.7951(HK) + 0.2586 \quad (46)$$

Table 27. C1 – C6 coefficients equations as function of HK with R<sup>2</sup> and RMSEE.

Equation	R <sup>2</sup>	RMSEE	Average value
$C1 = F(HK)$	0.9723	0.314110067	1.8463
$C2 = F(HK)$	0.9819	0.0138581	-0.1027
$C3 = F(HK)$	0.9792	0.000200226	-0.0865
$C4 = F(HK)$	0.8725	0.828237984	27.2324
$C5 = F(HK)$	0.8601	0.036846179	-1.3735
$C6 = F(HK)$	0.9997	0.000496825	0.2818

The coefficient C3 and C6 are the offset introduced to slope and intercept prediction. They showed high accuracy with high R<sup>2</sup> and low RMSEE as compared to average value. The error in C1, C2, C4 and C5 have small impact on Xd model as they small contribution for slope and intercept prediction.

The model was tested for variable feed composition and Xb over 125 cases (training data) and 25 case (prediction data) and showed low RMSEE = 0.00030529 with Xd average value = 0.022. The maximum absolute error in Xd is 0.0007 which represents very small error as compared to average liquid distillate stream purity of 0.966. It showed good prediction ability with RMSEP = 0.00026749. Below is the summary of the testing results (in Table 28) and the relative error plot in Fig. 47.

Table 28. Xd model testing results for 125 cases (training data) & 25 (prediction data) at variable feed composition, Xd & Xb

Model Summary Equation	$x_d = a \left( \frac{R}{V_d + D} \right) + b$ <p>Slope <math>a = C1 x_b^2 + C2 x_b + C3</math></p> <p>Intercept <math>b = C4 x_b^2 + C5 x_b + C6</math></p> <p><math>C1, C2, C3, C4, C5</math> &amp; <math>C6 = F(HK)</math></p> <p>Quadratic fitting</p>
R <sup>2</sup>	0.9984
Q <sup>2</sup>	0.9986
RMSEE	0.00030529
RMSEP	0.000267488
Xd average value	0.022
Distillate purity average value	0.966
Polynomial degree	3

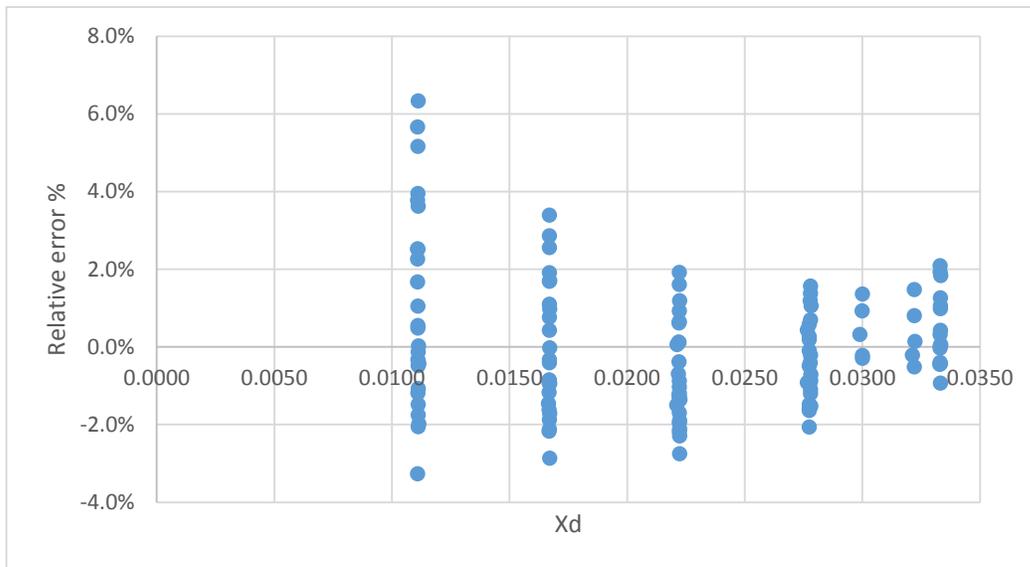


Figure 47. Relative error% for Xd complete model testing variable feed composition and Xb

### 3.5.4 Model for bottom impurity, $X_b$

The analysis was conducted at two stages at constant and variable feed composition. The first stage conducted at constant feed composition (Feed 3) with 25 combination of  $X_b$  and  $X_d$  values. It evaluated the impact from top operation i.e. variable  $X_d$  on  $X_b$  equation. The below Fig. 48 shows  $X_b$  equation with quadratic regression as Eq. (47) at various  $X_d$  values.

$$x_b = C1 \left(\frac{Q_r}{B}\right)^2 + C2 \left(\frac{Q_r}{B}\right) + C3 \quad (47)$$

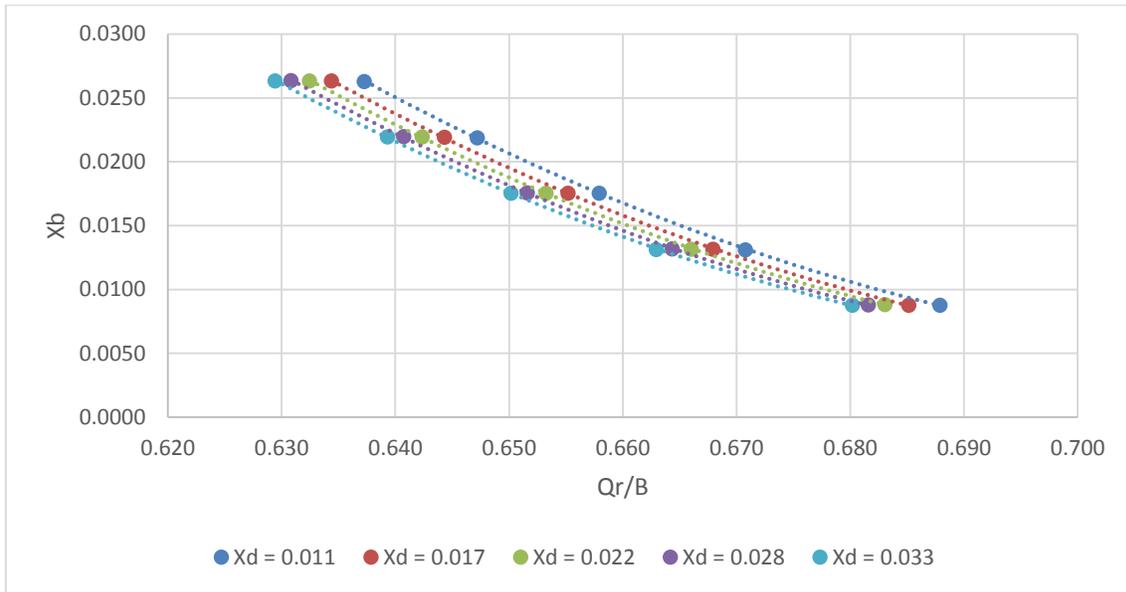


Figure 48. Relationship between  $X_b$  and  $Q_r/B$  at constant feed and variable  $X_d$  with quadratic fitting

$X_b$  equation as Eq. (47) experience shifting due to  $X_d$  variation. This is attributed to the changes in the column energy input from  $Q_c$  which is causing changes in the internal column liquid flow and so changes in the reboiler duty. The shifting effect can be modeled by evaluating  $X_d$  equation coefficients for various  $X_b$  values as per Table 29 below.

Table 29. Xb equation coefficient for variable Xd values at constant feed composition

Xd	C1	C2	C3
0.011099	2.62394	-3.82475	1.398117
0.016688	2.614914	-3.79766	1.383191
0.022191	2.626444	-3.80276	1.380888
0.027774	2.633417	-3.80389	1.378057
0.03331	2.638816	-3.80344	1.374935

The coefficients C1, C2 and C3 are changing due to Xd variation. The resulting fitting quadratic equations of coefficients as function of Xd showed low  $R^2$  and high RMSEE as shown in the below Table. 30.

Table 30. Xb quadratic coefficients C1, C2 and C3 as function of Xd fitting results

	C1 = F(Xd)	C2 = F(Xd)	C3 = F(Xd)
$R^2$ Quadratic fitting	0.8202	0.6967	0.9415
RMSEE	0.003472	0.005174	0.001942
Average Value	2.627506186	-3.806497655	1.383037342

Having a quadratic fitting for coefficients as function of Xd can lead to high error percentage. The high error in intercept  $C3 = F(Xd)$  imposes high offset affecting Xb prediction. This can eliminate the use of quadratic fitting for Xb equation as Eq. (47) due to the low quadratic coefficient models accuracy especially at the intercept prediction.

Alternatively Xb equation can be modeled through exponential fitting following Xb equation as eq. (22a) then,  $\ln(x_b) = a \left(\frac{Q_r}{B}\right)$  as eq.(23a) which is demonstrated in Fig. 49 for constant feed composition and variable Xd & Xb.

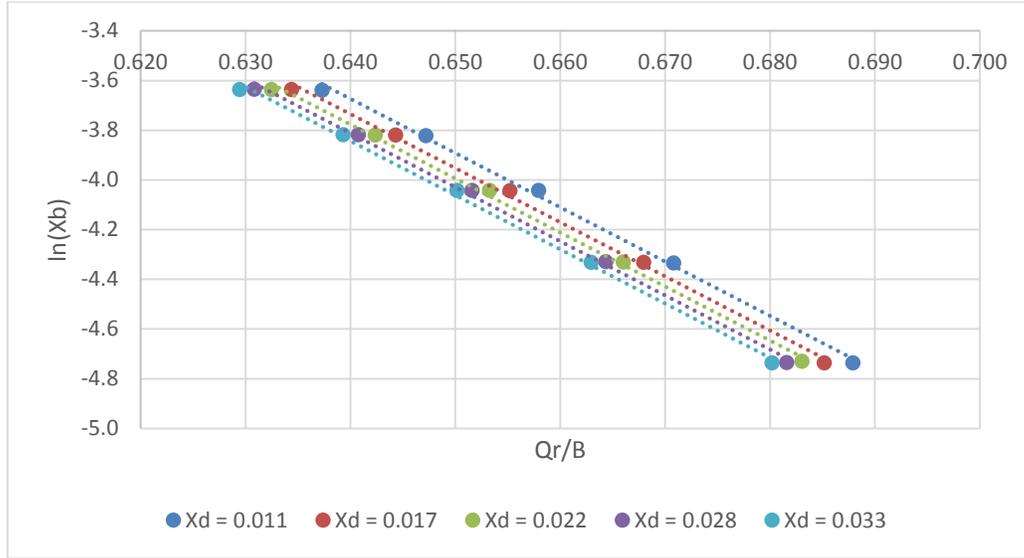


Figure 49. Relationship between  $\ln(X_b)$  and  $Q_r/B$  at constant feed composition and variable  $X_d$

The slope  $a$  and intercept  $b$  for Eq. (23a) are changing due to  $X_d$  variation as shown in Table. 31.

Table 31.  $\ln(X_b)$  equation slope and intercept values for constant feed composition and variable  $X_d$

$X_d$	slope	intercept
0.011099	-21.82111	10.29192
0.016688	-21.76966	10.19764
0.022191	-21.73688	10.13511
0.027774	-21.77360	10.1239
0.03331	-21.79442	10.10554

Then generated plots for slope and intercept as function Xd as shown in Figs. 50 and 51.

Table. 32 shows slope and intercept equations accuracy analysis.

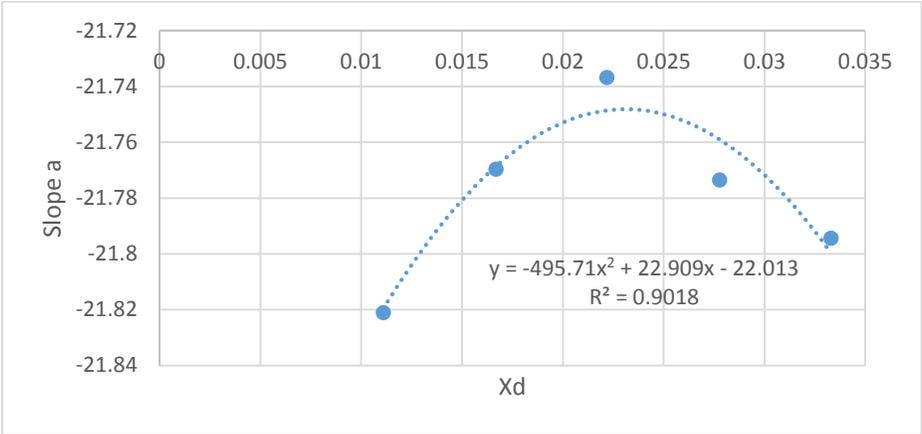


Figure 50. Relationship between slope a and Xd for constant feed composition

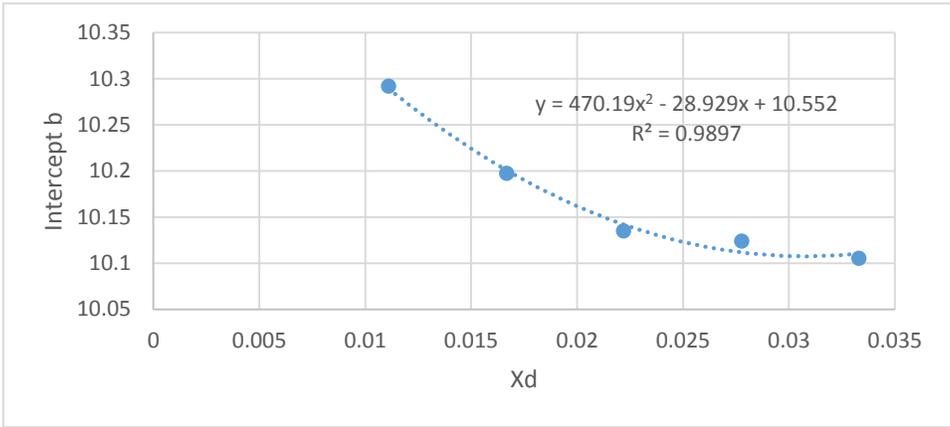


Figure 51. Relationship between intercept b and Xd for constant feed composition

Table 32. ln(Xb) equation slope and intercept as function of Xd fitting analysis at constant feed composition

	Slope a = F(Xd)	Slope = average	Intercept b = F(Xd)
R <sup>2</sup> Quadratic fitting	0.9018	NA	0.9897
RMSEE	0.008752	0.027932	0.006899
Average Value	-21.7791	-21.7791	10.170

$$\text{Slope } a = \text{average (Slope)} \quad (48)$$

$$\text{average (Slope)} = -21.7791$$

$$\text{Intercept } b = d1 x_d^2 + d2 x_d + d3 \quad (49)$$

$$d1 = 470.1852 \quad d2 = -28.9287 \quad d3 = 10.55244$$

The intercept equation as Eq. (49) is more accurate as compared to the quadratic fitting Eq. (47) intercept C3 by looking at R<sup>2</sup> and RMSEE. This is important as intercept error imposes offset for ln(Xb) equation. The slope quadratic fitting equation doesn't show high R<sup>2</sup> due to the small variation in slope values. Hence, slope can be assumed average value for variable Xd values as Eq. (48). This is minimizing the number of equations and simplifying the model without jeopardizing the model accuracy.

The same analysis for ln(Xb) equation were carried out for Feed 1, Feed 2, Feed 4 and Feed 5. The below Fig. 52 illustrates the relationship between ln(Xb) and Qr/B for variable feed composition and Xb. The first stage of ln(Xb) modeling dealt with shifting caused by Xd variation, while the second stage deals with the shifting effect caused by feed composition variation.

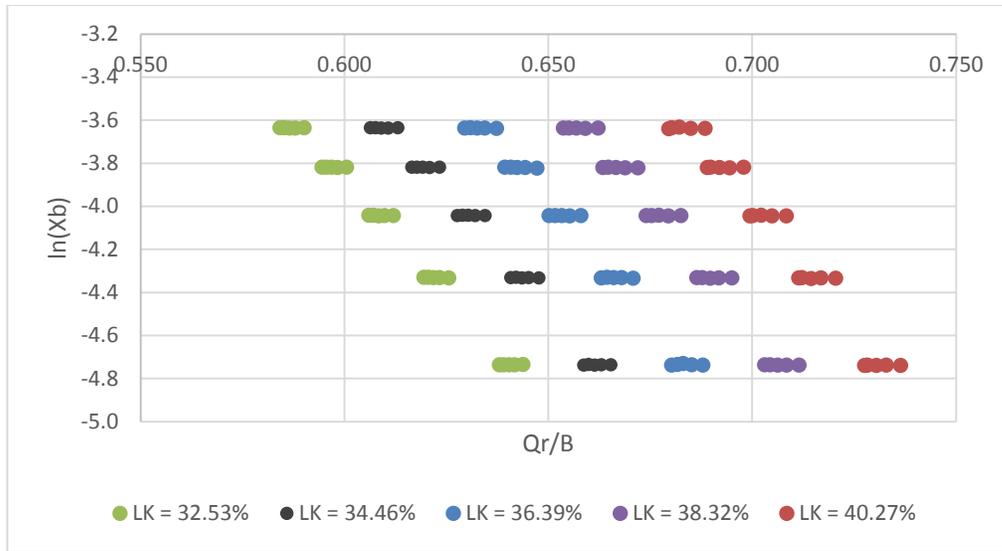


Figure 52. Relationship between  $\ln(X_b)$  and  $Q_r/B$  at variable feed composition and  $X_d$   $\ln(X_b)$  equation as Eq. (23a) experiences shifting due to  $X_d$  variation and feed composition. The  $X_d$  variation was modeled by introducing the intercept equations as Eq. (49) function of  $X_d$  and slope as average value as Eq. (48). The below Table. 33 illustrates  $\ln(X_b)$  equation with slope and intercept function for each feed composition.

Table 33.  $\ln(X_b)$  slope and intercept equation coefficients for variable feed composition and  $X_d$

LK%	$\ln(x_b) = a \left( \frac{Q_r}{B} \right) + b$			
	<i>Slope a = average (Slope)</i>	<i>Intercept b = d1 x<sub>d</sub><sup>2</sup> + d2 x<sub>d</sub> + d3</i>		
	Average slope for variable $X_d$	d1	d2	d3
32.54%	-20.5628	195.28602	-14.0866	8.656498
34.46%	-21.1642	243.40974	-16.9196	9.524752
36.39%	-21.7791	470.18523	-28.9287	10.55244
38.32%	-22.4057	320.65999	-23.6411	11.46397
40.27%	-23.0441	484.47133	-31.715	12.56563

$\ln(Xb)$  equation slope changes due to feed variation represented by LK composition in feed. In addition the intercept equation coefficients d1, d2 and d3 are changing due to feed composition variation. Figures 53 - 56 show the slope and intercept coefficients d1 – d3 behavior with respect to variation in light key feed composition.

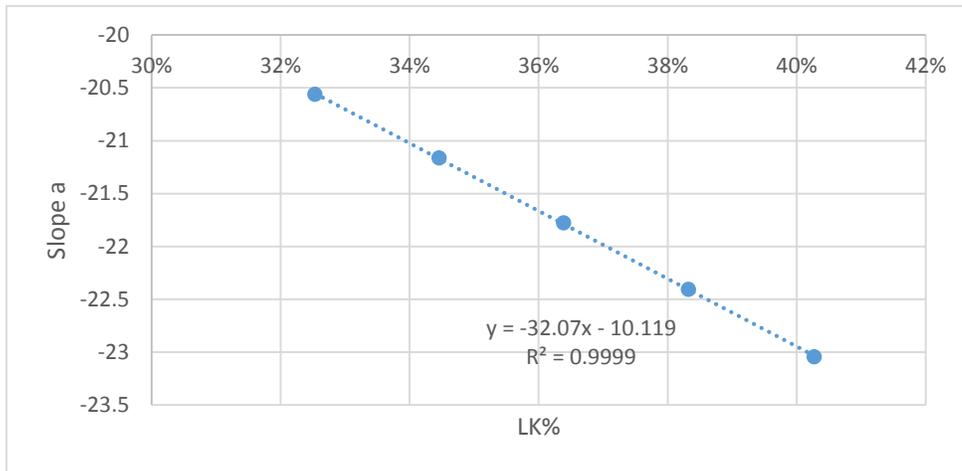


Figure 53.  $\ln(Xb)$  slope function of LK for variable feed composition & Xd

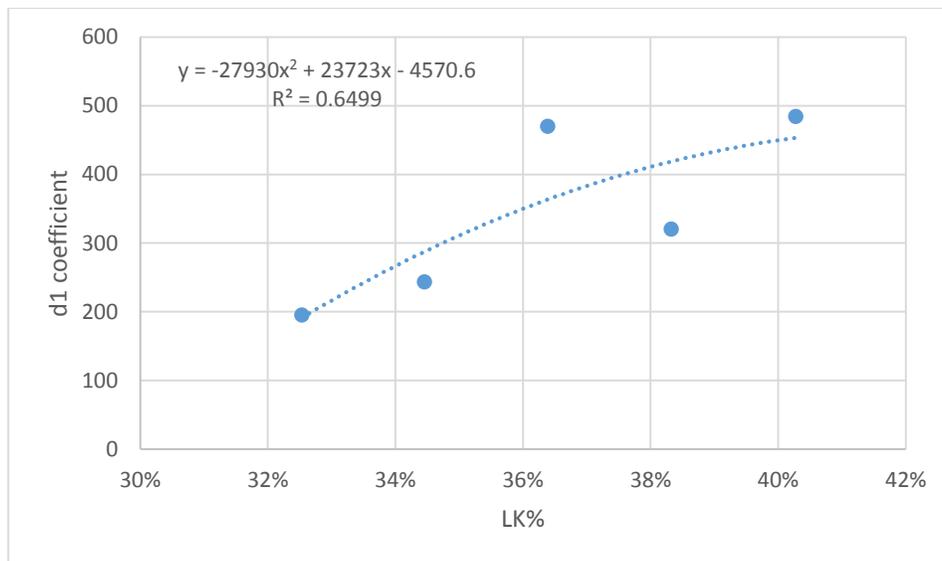


Figure 54. d1 coefficient function of LK for variable feed composition & Xd

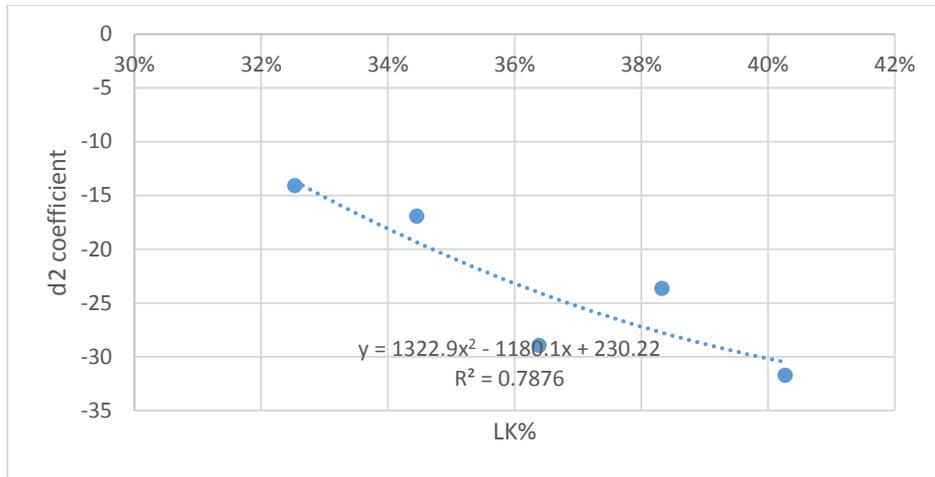


Figure 55. d2 coefficient function of LK for variable feed composition & Xd

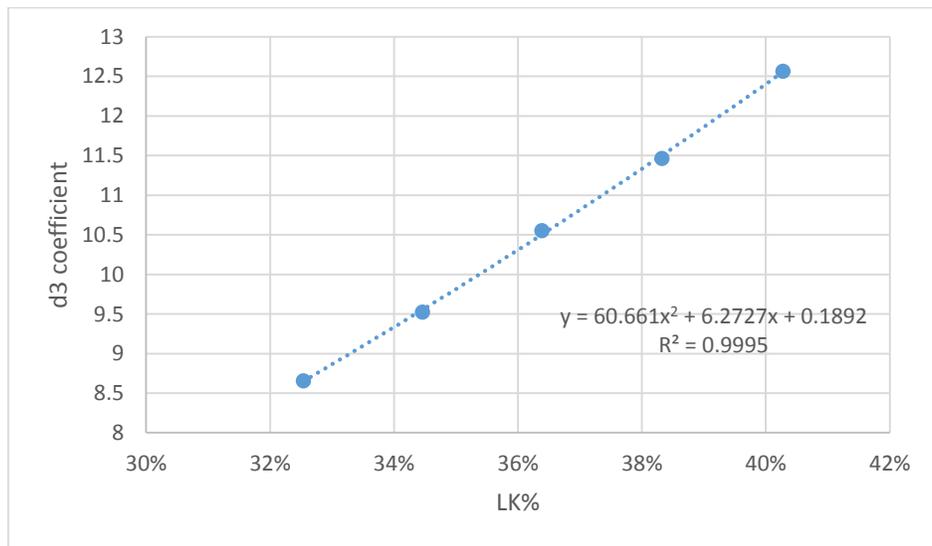


Figure 56. d3 coefficient function of LK for variable feed composition & Xd

Slope is represented by Eq. (50) and intercept equation coefficients are given by Eqs (51) - (53) as function of feed LK. Table 34 summarizes statistical accuracy of Eqs. (50) - (53).

$$\text{Slope } a = -32.07 (\text{LK}) - 10.119 \quad (50)$$

$$d1 = -27930 (LK)^2 + 23723(LK) - 4570.6 \quad (51)$$

$$d2 = 1322.9 (LK)^2 - 1180.1(LK) + 230.22 \quad (52)$$

$$d3 = 60.661 (LK)^2 + 6.2727(LK) + 0.1892 \quad (53)$$

Table 34. Slope and intercept equation coefficients as function of LK accuracy analysis

Equation	R <sup>2</sup>	RMSEE	Average value
<i>Slope a = F(LK)</i>	0.9999	0.007735	-21.79
<i>d1 = F(LK)</i>	0.6499	69.223	342.80
<i>d2 = F(LK)</i>	0.7876	3.111	-23.06
<i>d3 = F(LK)</i>	0.9995	0.031	10.55

The slope equation as Eq. (50) has high R<sup>2</sup> and low RMSEE. The coefficients d1 & d2 for intercept equation as Eq. (51) – (52) have low R<sup>2</sup> and high RMSEE, while Eq. (53) which represents offset in intercept prediction has high R<sup>2</sup> and low RMSEE as compared to average values. High accuracy for offset modeling contributes for whole model high accuracy. Although Eqs. (51) and (52) predictions are not very accurate they actually have minimal contribution in intercept prediction. Table. 35 shows an example (Xd = 0.011 & LK% = 36.39%) for each coefficient contribution.

Table 35. coefficients contribution to intercept prediction

<i>Intercept b = d1 x<sub>d</sub><sup>2</sup> + d2 x<sub>d</sub> + d3</i>		
<i>d1 = F(LK)</i>	<i>d2 = F(LK)</i>	<i>d3 = F(LK)</i>
<i>d1 x<sub>d</sub><sup>2</sup></i>	<i>d2 x<sub>d</sub></i>	<i>d3</i>
0.0448	-0.2669	10.50254

The contribution from d1 and d2 is in small magnitude and hence coefficient modeling inaccuracy has minimal impact on intercept prediction. Furthermore higher accuracy prediction for offset coefficient (d3) contributes for higher accuracy intercept prediction which indeed increase Xb prediction model accuracy.

Xb model for variable feed composition and Xd was tested for 125 cases (training data) and 25 (prediction data). It showed low RMSEE = 0.0004134 with average Xb = 0.0175. The maximum absolute error in Xb turn to be 0.0013 which represents very small error as compared to average bottom stream purity of 0.953 (iC4+nC4). It showed good prediction ability with RMSEP = 0.0004217. Table. 36 shows Xb model testing results summary while the relative error percentage is shown in Fig. 57.

Table 36. Xb model testing results (125 cases training data) & (25 prediction data) for variable feed composition and Xd

Model Summary Equation	$x_b = a e^{-b\left(\frac{Qr}{B}\right)}$ $\ln(x_b) = a \left(\frac{Qr}{B}\right) + b$ <i>Slope a = -32.07 (LK) - 10.119</i> <i>Intercept b = d1 x<sub>d</sub><sup>2</sup> + d2 x<sub>d</sub> + d3</i> <i>d1, d2 &amp; d3 = F(LK)</i> <i>Quadratic fitting</i>
R <sup>2</sup>	0.996
Q <sup>2</sup>	0.9951
RMSEE	0.0004246
RMSEP	0.0004217
Xb average value	0.0175
Bottom purity average value	0.953
Polynomial degree	2

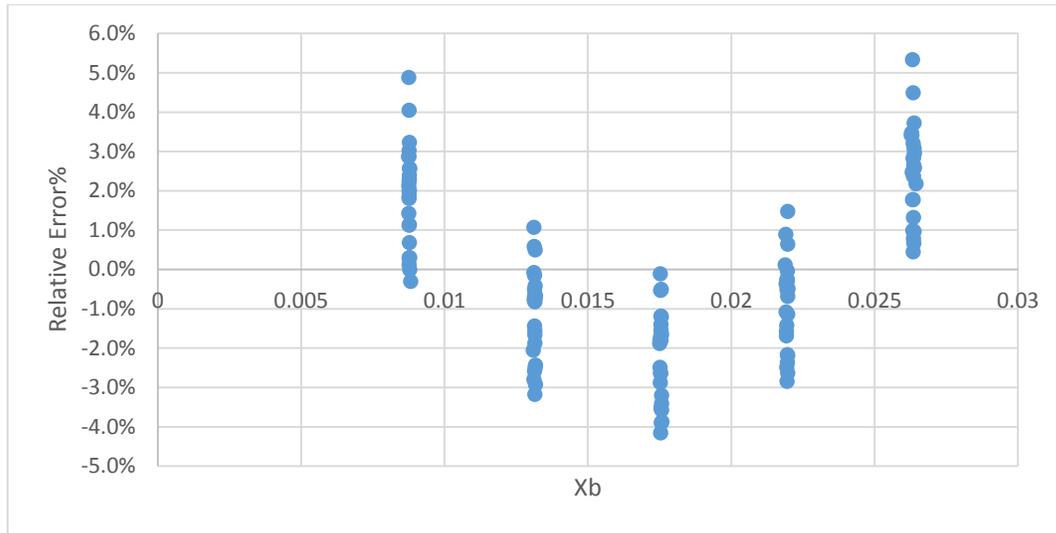


Figure 57. Relative error% for Xb model for variable feed composition and Xd

### 3.5.5 Vapor Distillate Flow and Composition Model

The DeC3 column operates at constant condenser temperature and pressure which eliminate their impact on vapor distillate  $V_d$  flow. Condenser duty  $Q_c$  is the main factor affecting  $V_d$  flow. Condenser duty varies mainly to meet specified  $X_d$  value. In addition changes in reboiler duty to meet specified  $X_b$  causes  $Q_c$  variation. Hence  $V_d$  is affected by the column top ( $X_d$  specification) and bottom operation ( $X_b$  specification).

The vapor distillate is modeled as  $\frac{V_d}{F}$  ratio which helps in applying the model for variable feed flow. The below Fig. 58 shows the relationship between  $\frac{V_d}{F}$  and  $Q_c$  for constant feed composition and flow (Feed 3). The  $\frac{V_d}{F}$  equation experience shifting due to  $X_b$  variation i.e. reboiler duty variation. Along each  $\frac{V_d}{F}$  equation line  $Q_c$  varies to meet  $X_d$  specification.

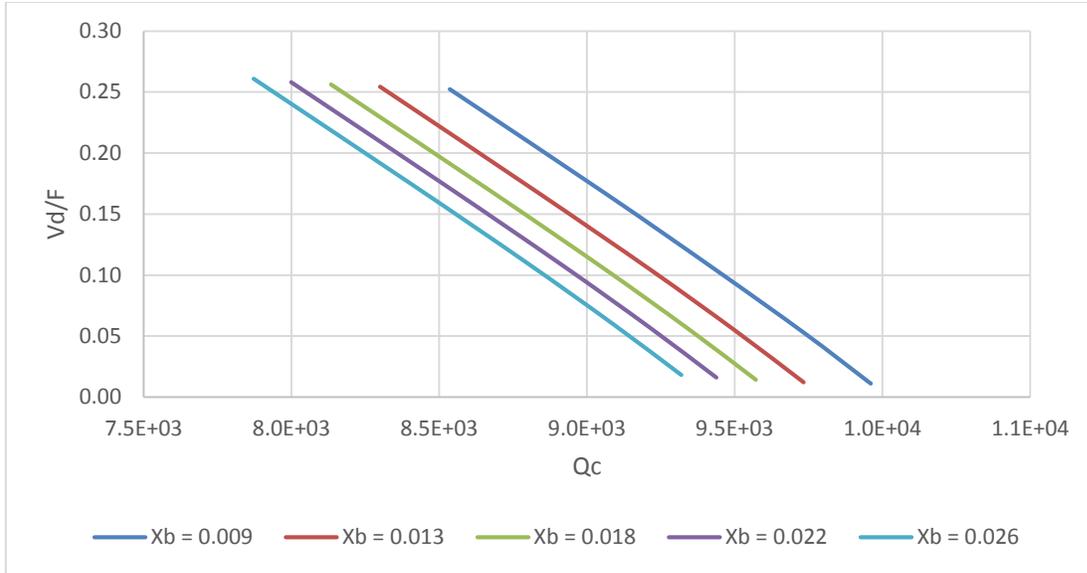


Figure 58. Relationship between  $V_d/F$  and condenser duty  $Q_c$  at constant feed and variable  $X_d$  &  $X_b$

$X_d$  model for constant feed composition includes reflux ratio in addition to modeling the shifting effect of  $X_b$  variation,  $X_d = F(\text{reflux ratio}, X_b)$ .  $\frac{V_d}{F}$  can be modeled as function of  $X_d$  in which it includes top and bottom operation. This approach is minimizing the number of equations in  $\frac{V_d}{F}$  model. The below Fig. 59 illustrates the relationship between  $\frac{V_d}{F}$  and  $X_d$  which is represented by a quadratic line for constant feed composition and variable  $X_b$ .

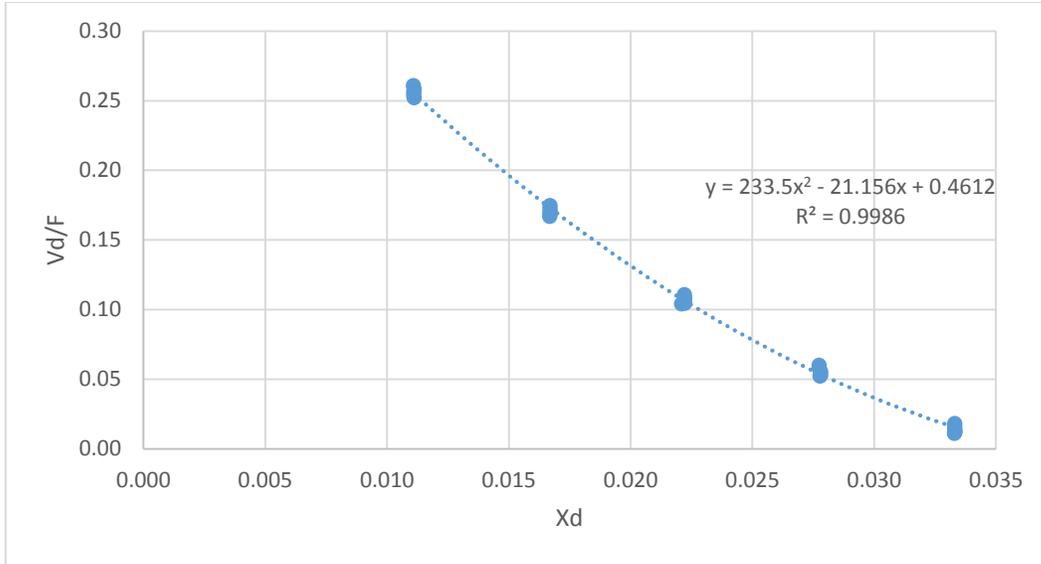


Figure 59. Relationship between Vd/F and Xd at constant feed composition (Feed 3) and variable Xb

The quadratic regression line as Eq. (54) has high  $R^2 = 0.9986$  and low RMSEE = 0.003212441 for average  $\frac{V_d}{F} = 0.12$ .  $\frac{V_d}{F}$  equation as function of Xd experience negligible shifting effect due to Xb variation considering the high  $R^2$  and low RMSEE values.  $\frac{V_d}{F}$  equation Eq. (54) for constant feed composition (Feed 3) and variable Xb is shown below.

$$\frac{V_d}{F} = C1 x_d^2 + C2 x_d + C3 \quad (54)$$

$$C1 = 233.5 \quad C2 = 21.156 \quad C3 = 0.4612$$

The same analysis were carried out for Feed 1, Feed 2, Feed 4, and Feed 5 and the results are shown in Fig. 60.

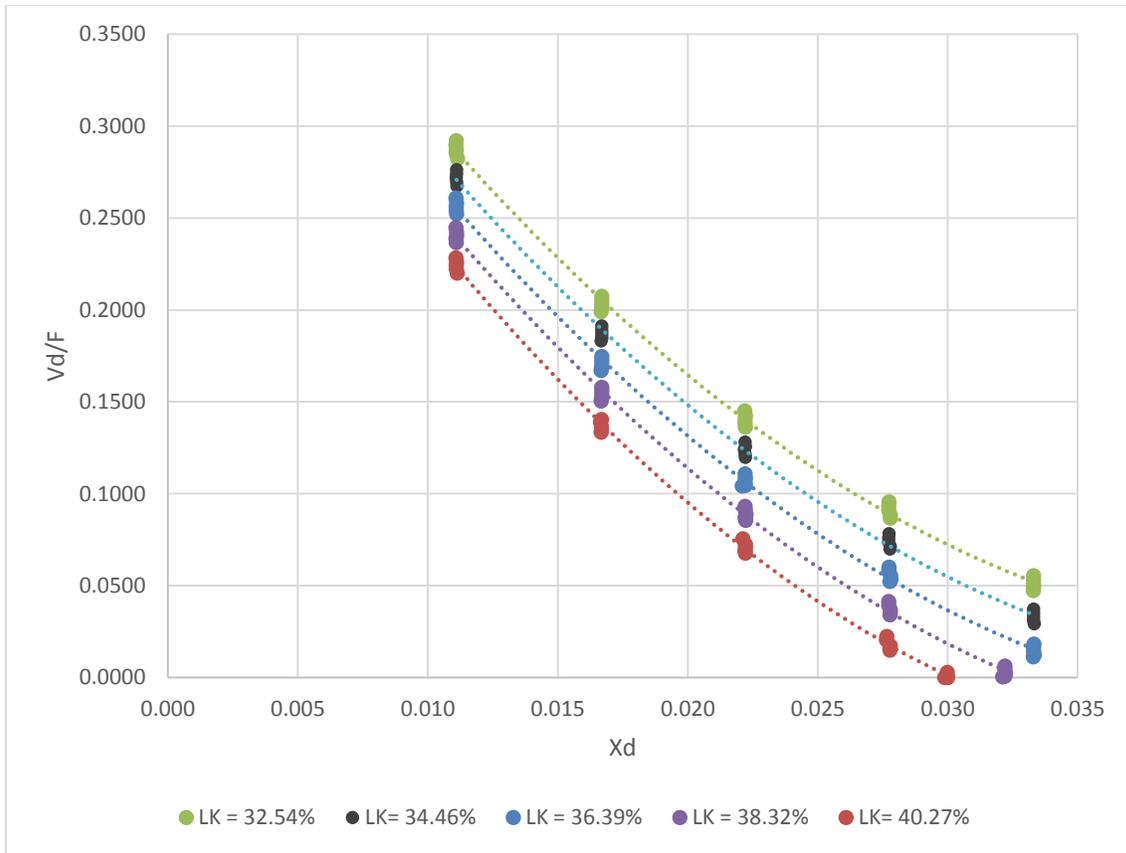


Figure 60. Relationship between  $V_d/F$  and  $X_d$  at variable feed composition and  $X_b$

As shown  $\frac{V_d}{F}$  equation Eq. (54) experience shifting due to feed composition variation.

The ratio  $\frac{V_d}{F}$  approaches zero for high LK% &  $X_d$ . This indicating the minimum  $V_d$

value for various LK &  $X_d$  values.  $\frac{V_d}{F}$  equation coefficients  $C_1$ ,  $C_2$  and  $C_3$  change due

to variation in light key composition in the feed as shown in Table. 37.

Table 37.  $\frac{V_d}{F}$  equation coefficients for variable feed composition

	$\frac{V_d}{F} = C1 x_d^2 + C2 x_d + C3$		
LK%	C1	C2	C3
32.54%	234.5657341	-20.9527952	0.4898151
34.46%	234.1046803	-21.0464816	0.4755149
36.39%	233.4988669	-21.1558787	0.4611803
38.32%	241.1670522	-21.6051492	0.4495287
40.27%	265.4391580	-22.6872915	0.4427909

The coefficients C1, C2 and C3 can be modeled as function of LK% as shown in Figs.

61 - 63.

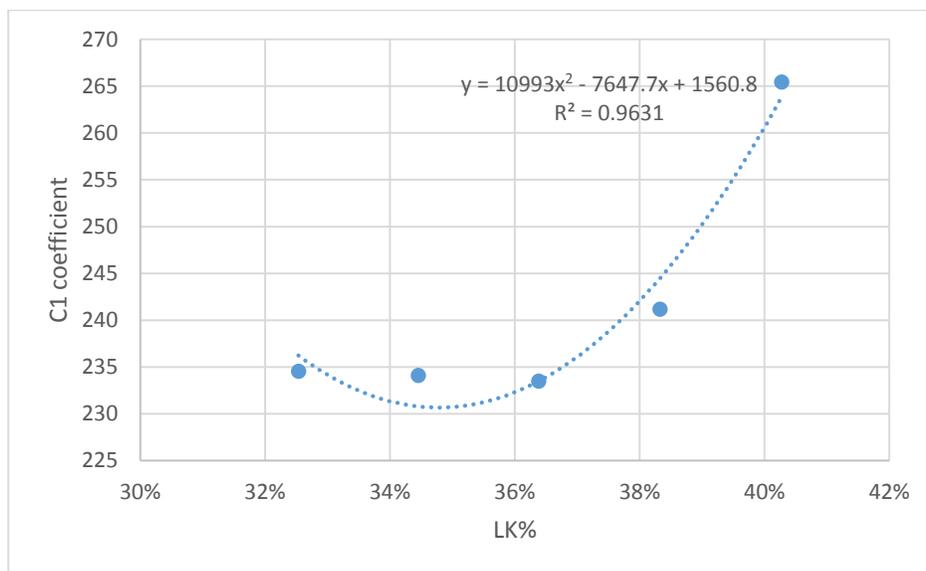


Figure 61. Relationship between C1 coefficient and feed LK

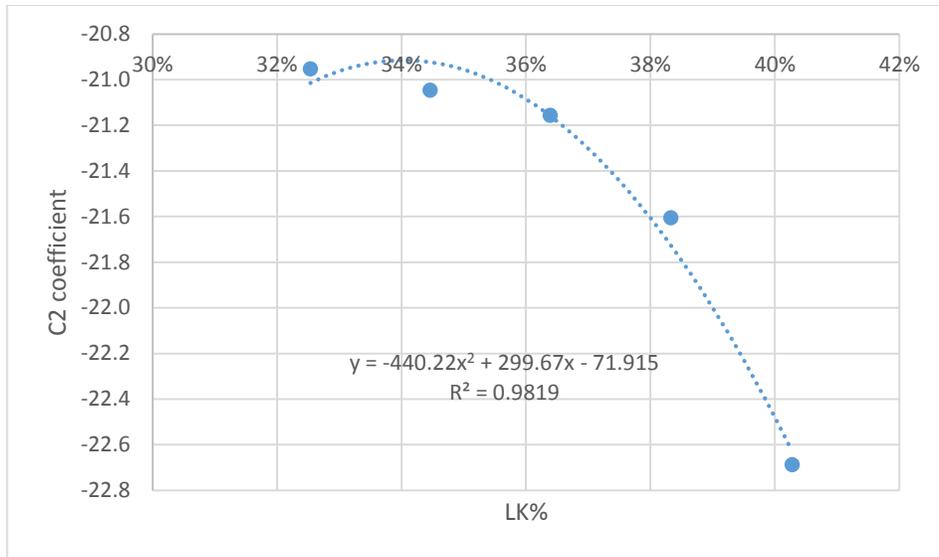


Figure 62. Relationship between C2 coefficient and feed LK

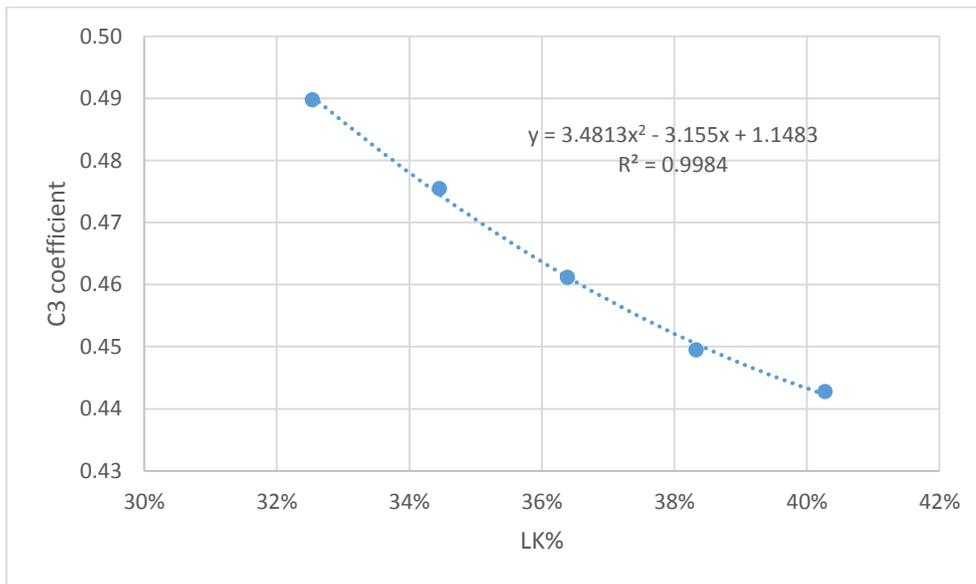


Figure 63. Relationship between C3 coefficient and feed LK

The coefficients are function of the changes in feed composition as per the below equations Eqs. (55), (56) and (57) generated from the coefficients Vs LK% plots. Table 38 shows the coefficients equations Eqs. (55), (56) and (57) accuracy analysis.

$$C1 = 10993 (LK)^2 - 7647.7 (LK) + 1560.8 \quad (55)$$

$$C2 = -440.22 (LK)^2 + 299.67 (LK) - 71.915 \quad (56)$$

$$C3 = 3.4813 (LK)^2 - 3.155 (LK) + 1.1483 \quad (57)$$

Table 38. Coefficients C1, C2 & C3 as function of LK% fitting analysis

Equation	R <sup>2</sup>	RMSEE	Average value
$C1 = F(LK)$	0.9631	2.334870156	241.7550983
$C2 = F(LK)$	0.9819	0.085942937	-21.48951924
$C3 = F(LK)$	0.9984	0.000689708	0.463766011

The R<sup>2</sup> for the three plots is high as shown indicating high accuracy. The intercept coefficient C3 indicated high R<sup>2</sup> which is minimizing  $\frac{V_d}{F}$  prediction error. The model was tested for variable feed composition, Xd and Xb with 125 cases (training data) and 25 (prediction data). It resulted in vapor distillate Vd/F ratio RMSEE = 0.00312 & RMSEP = 0.0027 for average Vd/F average value = 0.122, which resulted in Vd flow prediction error of RMSEE = 80.1 Kg/hr & RMSEP = 70.66 Kg/hr with Vd average flow = 3127.5 Kg/hr. The error is small when compared to average flow. Table 39 illustrates the  $\frac{V_d}{F}$  model accuracy analysis for 125 cases (training data) and 25 (prediction data) of variable feed composition, Xd and Xb.

Table 39. Vapor distillate model testing for variable feed composition, Xd and Xb

Model Summary Equation	$\frac{V_d}{F} = C1 x_d^2 + C2 x_d + C3$ <p><i>C1, C2 and C3 = F(LK)</i> <i>Quadratic fitting</i></p>
R <sup>2</sup>	0.9987
Q <sup>2</sup>	0.9988
RMSEE	0.003115548
RMSEP	0.002704461
Vd/F average value	0.12164852
Polynomial degree	4

The vapor distillate Vd consists of C2, C3 and traces of iC4. The components compositions  $x_{C2,Vd}$ ,  $x_{C3,Vd}$  &  $x_{iC4,Vd}$  are determined by flash calculations at the condenser. Each component mole composition in Vd is related to its composition in distillate D in mole basis by K values as shown below for  $x_{iC4,Vd}$ .

$$\frac{x_{i,Vd}}{x_{i,D}} = K_i \quad (58)$$

Hence, individual component composition in liquid distillate D has major impact on Vd component composition. For  $x_{iC4,Vd}$  it can be directly calculated as Eq. (59) in which Xd is already modeled. A conversion factor for mole to mass basis is required.

$$\frac{x_{iC4,Vd}}{x_d} = K_{iC4} = 0.5159 \times \frac{Kg}{mole} \text{ conversion factor} \quad (59)$$

However for  $x_{C2,Vd}$  &  $x_{C3,Vd}$  it requires obtaining their compositions in distillate which are eventually changing due to  $X_d$  variation. Compositions also need to be converted from mole to mass basis, which is accounted for in the data used to generate Fig. 64. Alternatively, empirical models were generated to include flash calculations along with mole to mass basis conversion. Figures 64 – 66 show  $x_{C2,V}$ ,  $x_{C3,Vd}$  &  $x_{IC4,Vd}$  as function of  $X_d$  (iC4% in D) for 125 cases of variable feed composition,  $X_d$  and  $X_b$ .

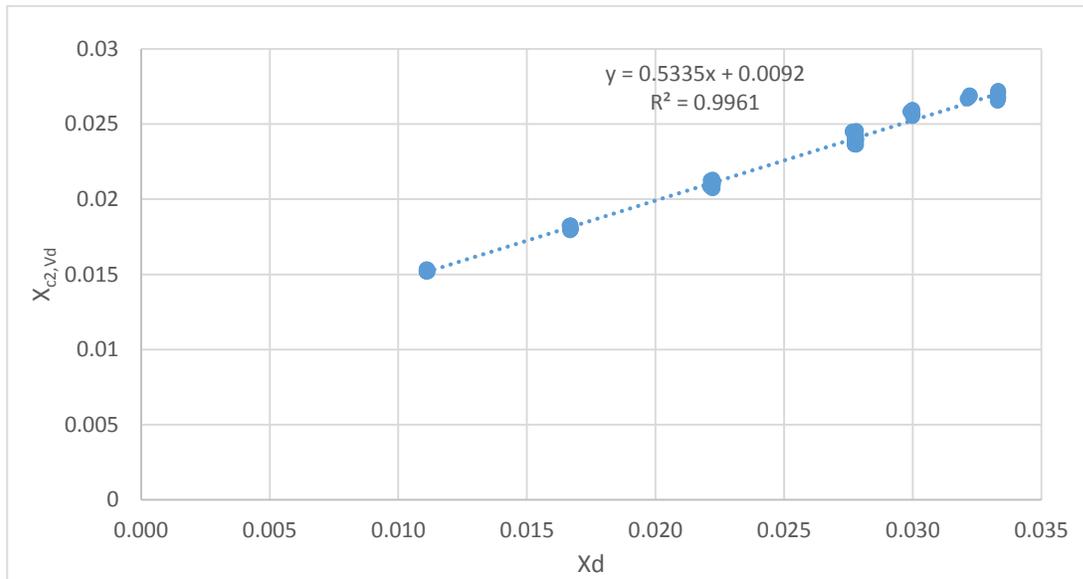


Figure 64. Relationship between  $x_{C2,Vd}$  and  $X_d$

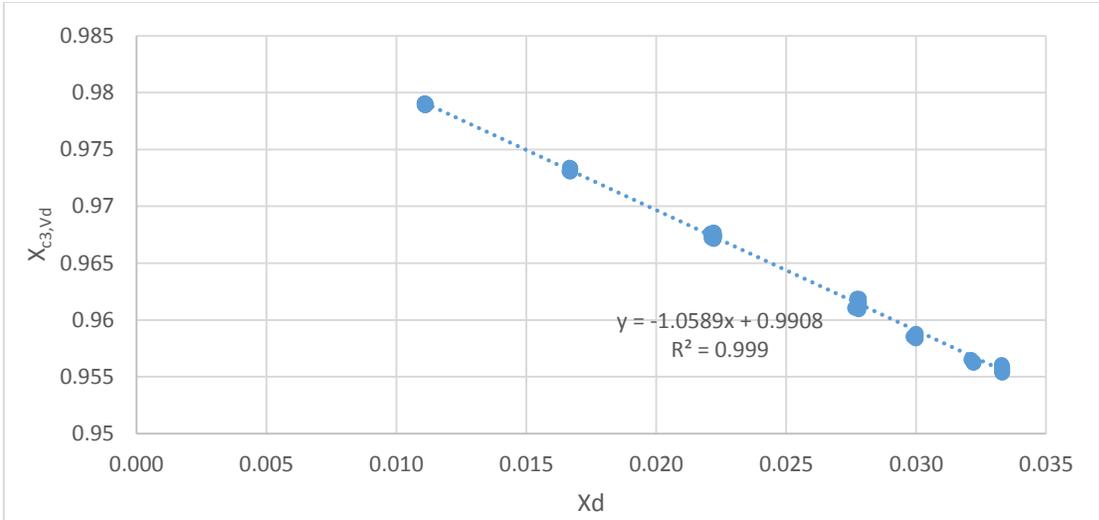


Figure 65. Relationship between  $X_{C3,Vd}$  and  $X_d$

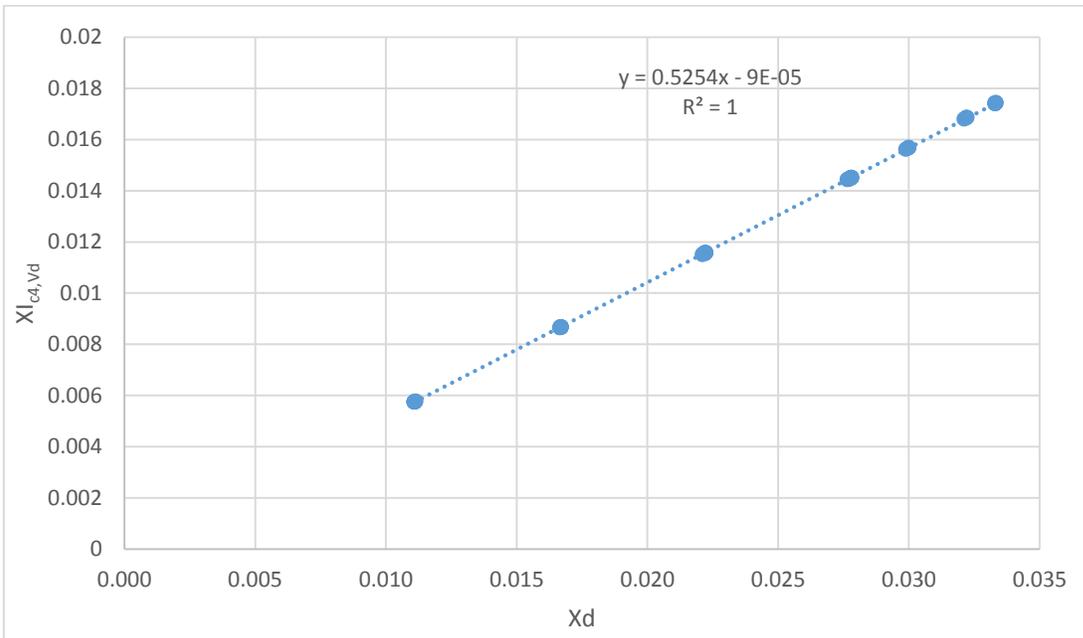


Figure 66. Relationship between  $X_{IC4,Vd}$  and  $X_d$

Eqs. (60), (61) and (62), represent the relationships between vapor phase compositions and the distillate impurity  $x_d$  for the partial condenser (mass basis) for variable feed composition,  $x_d$  and  $x_b$ .

$$x_{C2,Vd} = 0.5335 x_d + 0.0092 \quad (60)$$

$$x_{C3,Vd} = -1.0589 x_d + 0.9908 \quad (61)$$

$$x_{IC4,Vd} = 0.5254 x_d - 9E - 05 \quad (62)$$

Eq. (62) has an intercept of almost zero and slope equals to the K value with mole to mass basis conversion factor. The above equations simplify the flash calculations to determine Vd components composition. Eqs. (60), (61) and (62) were tested for 125 cases (training data) and 25 cases (prediction data) of variable feed composition,  $x_d$  &  $x_b$ . The equations showed low RMSEE & RMSEP as compared to average values as shown in Table. 40.

Table 40. Vd components composition as function of  $x_d$  fitting analysis

Equation	R <sup>2</sup>	Q <sup>2</sup>	RMSEE	RMSEP	Average value
$x_{C2,Vd} = F(x_d)$	0.9961	0.9969	0.000252	0.000207	0.021002
$x_{C3,Vd} = F(x_d)$	0.999	0.9976	0.000259	0.0011596	0.967507
$x_{IC4,Vd} = F(x_d)$	1	1	1.1422E-05	2.3874E-05	0.011491

### 3.5.6 Depropanizer Column Hybrid Model Configuration

The hybrid model for DeC3 consists of mass balance, components mass balance, energy balance and prediction models for  $x_d$ ,  $x_b$  and  $V_d/F$ . The heat of condensation for

condenser  $\lambda_c$  data were generated from Aspen plus in which changes in distillate impurity Xd cause changes in  $\lambda_c$ . Hence  $\lambda_c$  is correlated to Xd as Eq. (70) with linear regression and  $R^2 = 0.9928$ . The same applied to reboiler heat of vaporization  $\lambda_r$  as function of bottom stream impurity Xb as Eq. (72) with linear regression  $R^2 = 0.9819$ . The heat estimation equation for  $\lambda_c$  &  $\lambda_r$  improve the model accuracy. DeC3 hybrid model mainly consists of equations Eqs. (63) – (72) in addition to Xd, Xb and Vd/F empirical models.

$$F = Vd + D + B \quad (63)$$

$$H_F F + Q_r = h_{Vd} V_d + H_D D + H_B B + Q_c \quad (64)$$

$$H_F, h_{Vd}, H_D, H_B \text{ (assumed constant)}$$

$$x_{i,F} F = x_{i,Vd} V_d + x_{i,D} D + x_{i,B} B \quad i = C2, C3, IC4, NC4, IC5 \text{ and } NC5 \quad (65)$$

$$Vd = Flow_{C2,Vd} + Flow_{C3,Vd} + Flow_{IC4,Vd} \quad (66)$$

$$D = Flow_{C2,D} + Flow_{C3,D} + Flow_{IC4,D} \quad (67)$$

$$B = Flow_{C3,B} + Flow_{IC4,B} + Flow_{NC4,B} + Flow_{IC5,B} + Flow_{NC5,B} \quad (68)$$

$$Q_c = \lambda_c (R + D) \quad (69)$$

$$\lambda_c = 0.094381x_d + 0.275983 \quad (70)$$

$$V_r = \frac{Q_r}{\lambda_r} \quad (71)$$

$$\lambda_r = 0.273681x_b + 0.228564 \quad (72)$$

### 3.6 Optimization Case Studies

Hybrid models as mentioned earlier included mass and energy balance equations in addition to empirical models. As illustrated earlier the empirical models for C4 splitter (high and low purity ranges) were tested for 144 cases of training data and 30 prediction data for each operating range. The empirical models for DeC3 were tested for 125 cases of training data and 25 prediction data. Empirical models results were compared to Aspen Plus results which indicated high accuracy. Models showed high  $R^2$  &  $Q^2$  and low RMSEE & RMSEP as compared to average value. This section tests hybrid models accuracy for optimization application. Two case studies are illustrated below for the C4 splitter high purity range and DeC3 columns.

#### 3.6.1 Butane C4 Splitter Optimization Case Study (High Purity Range)

The optimization case study was carried out for selected feed composition (Feed 2 - LK=24%) and feed flow = 161,039 Kg/hr. The objective function, assumed streams prices and constraints are described below.

$$Z = (2.44 - 4.4(x_d - 0.05))D + 2.09 B - 2.09 F - 14.21 Q_r \quad (73)$$

*D stream price = 2.44 \$/kg*

*$x_d$  reduction from design value 0.05 increase D price by 4.4 \$/0.01  $\Delta x_d$*

*B & F streams prices = 2.09 \$/kg*

*$Q_r$  energy cost 14.21 \$/GJ*

The objective function included price improvement factor when  $X_d$  is reduced from design value of 0.05.  $X_d$  reduction is assumed to increase distillate value prior to be fed to Alkylation unit and eventually to Gasoline pool. While  $X_b$  reduction assumed to have negligible impact on bottom stream value. Both bottom and feed streams assumed to have the same price due to the fact that bottom stream is fed to LPG pool and so the

feed will be fed to LPG when the column is out of service. The constraints included the following equations eq. (11 – 16), (23 -31) & (34 – 35). Table. 41 shows constraints on the variables used for this optimization problem formulation.

Table 41. C4 splitter optimization case study constraints

Variable	Min	Max
Qc (GJ/hr)	39.68	61.19
Qr (GJ/hr)	44.34	65.19
Reflux flow (Kg/hr)	567,159	852,025
Reflux ratio	10.59	35.86
Xd	0.005	0.03
Xb	0.005	0.03

The optimization problem was solved using GAMS global nonlinear solver Antigone with fast execution time (0.02 sec). Table. 42 compares the hybrid model optimization results to Aspen plus optimization solver.

Hybrid model indicated high accuracy as compared to Aspen plus results with relative error below 0.1% for all parameters. The optimum point was calculated by both hybrid model and Aspen plus is  $X_d = 0.005$  &  $X_b = 0.007$ . The objective function showed accurate result as compared to Aspen plus. Hybrid models present the advantage of being small size model in addition to being accurate as compared to rigorous tray to tray models.

Table 42. C4 splitter optimization results comparison between hybrid model and Aspen plus

Variable	Hybrid Model	Aspen plus	Relative Error%
D (Kg/hr)	36,431	36,424	0.02%
B (Kg/hr)	124,607	124,615	-0.01%
Qc (GJ/hr)	237.87	237.71	0.07%
Qr (GJ/hr)	256.23	256.23	0.00%
Reflux flow (Kg/hr)	806,257	805,697	0.07%
Reflux ratio	22.131	22.120	0.05%
Xd	0.005	0.005	0.00%
Xb	0.007	0.007	0.00%
Objective function (\$/hr)	16323	16316	0.04%

### 3.6.2 DeC3 Optimization Case Study

The optimization case study was carried out for selected feed composition (Feed 3 - LK=36.39% & HK=17.56%) and feed flow = 25,693 Kg/hr. The objective function, assumed streams prices and constraints are described below.

$$Z = 1.3 V_d + (1.83 - 4.11(x_d - 0.0222))D + (2.44 - 3.08(x_b - 0.0175))B - 1.9F - 14.21Q_r \quad (74)$$

$$V_d \text{ stream price} = 1.3 \text{ \$/kg}$$

$$D \text{ stream price} = 1.83 \text{ \$/kg}$$

$$B \text{ stream price} = 2.44 \text{ \$/kg}$$

$$F \text{ stream price} = 1.9 \text{ \$/kg}$$

$$x_d \text{ reduction from design value } 0.0222 \text{ increase } D \text{ price by } 4.11 \text{ \$/0.01 } \Delta x_d$$

$$x_b \text{ reduction from design value } 0.0175 \text{ increase } B \text{ price by } 3.08 \text{ \$/0.01 } \Delta x_b$$

$$Q_r \text{ energy cost } 14.21 \text{ \$/GJ}$$

The objective function included price improvement factor when  $X_d$  and  $X_b$  are reduced from their design values of 0.0222 & 0.0175 respectively. The constraints included the following equations eq. (37 – 46), (23), (48 – 57) & (63, 64 & 69). Table. 43 illustrates variables constraints used for this optimization problem formulation.

Table 43. DeC3 optimization case study constraints

Variable	Min	Max
Qc (GJ/hr)	1.754	2.432
Qr (GJ/hr)	2.438	2.677
Reflux flow (Kg/hr)	24,323	28,197
Reflux ratio	2.342	3.347
$X_d$	0.011	0.033
$X_b$	0.009	0.026

The optimization problem was solved using GAMS global nonlinear solver Antigone with fast execution time (0.1 sec). Table. 44 compares the hybrid model optimization results to Aspen plus optimization solver.

Hybrid model has high accuracy as compared to AspenPlus; it computes the same optimum point of  $X_d=0.033$  &  $X_b=0.009$ . The relative error is below 1% for most of the variables. The highest relative error% is for  $V_d$  prediction however the absolute error is 101 Kg/hr which imposes small error to the column mass input by 0.4%. This caused negligible impact to the overall model accuracy as shown most of parameters indicated relative error below 1%. The objective function showed lower value by 1.4% as compared to Aspen plus. Hybrid models present the advantage of being small size

model in addition to being sufficiently accurate as compared to rigorous tray to tray models.

Table 44. DeC3 optimization results comparison between hybrid model and Aspen plus

Variable	Hybrid Model	Aspen plus	Relative Error%
D (Kg/hr)	9,326	9,338	-0.13%
B (Kg/hr)	15,923	16,012	-0.56%
Vd (Kg/hr)	444	342	29.61%
Vd/F (Kg/Kg)	0.0173	0.0133	29.61%
Qc (GJ/hr)	9.96	9.93	0.30%
Qr (GJ/hr)	10.83	10.87	-0.43%
Reflux flow (Kg/hr)	26,510	26,258	0.96%
Reflux ratio	2.713	2.712	0.04%
Xd	0.033	0.033	0.00%
Xb	0.009	0.009	0.00%
Objective function (\$/hr)	7528	7637	-1.43%

## 4 Inferential Models

### 4.1 Introduction

This section illustrates inferential model development for two distillation column examples. Inferential model can be used on-line with real time data input for monitoring purpose. It estimates streams composition accurately. Hence eliminates the need for GC analyzer which eliminate analyzer installation and maintenance cost. In addition inferential model can used for control purposes. Model input can include trays temperature, reflux flow, reboiler duty, feed composition, distillate flow and bottom flow. The model is linear and can be configured part of the plant DCS. There are several configurations for inferential model such as

- Single tray temperature variable
- Multivariable model including feed composition input
- Multivariable model excluding feed composition input

This section illustrates each model configuration, accuracy and applicability.

### 4.2 Methodology

The main input is sensitive tray temperature in the rectifying section to predict  $X_d$  as Eq. (75) and in the stripping section to predict  $X_b$  as Eq. (76).

$$x_d = a T_i + b \quad i = \text{sensitive tray in the top section} \quad (75)$$

$$x_b = a T_i + b \quad i = \text{sensitive tray in the bottom section} \quad (76)$$

The sensitive tray selection plays major role in the model accuracy. It is highly recommended to select a tray that shows strong linear relation for variable  $X_d$ ,  $X_b$  and feed composition. While indicating significant temperature variation (sensitivity)  $\frac{\Delta T C^o}{\Delta x_d}$

or  $\frac{\Delta T C^o}{\Delta x_b}$ . Trays at column ends represents strong correlation for composition however temperature sensitivity is in small magnitude. Generally the higher  $\frac{\Delta T C^o}{\Delta x_d}$  or  $\frac{\Delta T C^o}{\Delta x_b}$  the more nonlinear Xd and Xb equations.

Sensitive tray selection for Xd inferential model is carried out at two stages constant feed composition and variable Xb and then for both variables. The most sensitive tray location has been determined by following the same steps as those followed for Xb model development.

- At constant feed composition and variable Xb
  - Generating data for  $\frac{\Delta T C^o}{\Delta x_d}$
  - Testing the highest  $\frac{\Delta T C^o}{\Delta x_d}$  tray temperate for linearity  $x_d = a T_i + b$  (usually nonlinear)
  - Testing trays above highest  $\frac{\Delta T C^o}{\Delta x_d}$  toward column end to check for linearity
  - Selecting a sensitive tray with significant  $\frac{\Delta T C^o}{\Delta x_d}$  that shows high  $R^2$  and low RMSEE. In addition to being not impacted by Xb variation.
- At variable feed composition and Xb.
  - Testing Xd inferential model with single sensitive tray for accuracy

The model of single temperature input presents simple and linear relation. The temperature set point can be changed to meet new impurity specification. However in case of low accuracy single temperature model it is recommended to introduce new

variables. Partial Least Square PLS can be used to generate single linear relation for Xd or Xb as function of multivariable. This work used ProMV software by ProSensus to generate PLS models. The reflux flow, reboiler duty and feed composition are the main contributors for Xd & Xb variation. In addition their variation cause internal flows changes leading to trays temperature variation. The PLS can include additional trays temperature with high sensitivity to improve model accuracy. The following variables were introduced to improve inferential models accuracy

- $T_i$  (sensitive tray and additional tray)
- $\frac{R}{R+D}$
- $\frac{V_r+B}{V_r}$
- Feed composition

Installing high number of tray temperature measurements require installation and maintenance cost. In addition each tray temperature measurement requires welding to the column shell which can increase the risk of hydrocarbon leaks during operation. The aim is to build models with minimum number of tray temperature measurements without compromising model accuracy. The additional temperature measurement other than the sensitive tray contributes with smaller magnitude for Xd & Xb prediction. The ratio  $\frac{R}{R+D}$  contributes mainly for Xd prediction. In addition to indicating the change in the column internal liquid flow which impact both Xd & Xb prediction. Furthermore causes internal reflux ratio variation which impact tray temperatures in the rectifying section. It provides estimation for the internal reflux ratio at each tray in the rectifying section.

$$\frac{L_{i-1}}{V_i} \cong \frac{R}{R+D} \quad (77)$$

The ratio  $\frac{V_r+B}{V_r}$  contributes mainly for Xb prediction. In addition to indicating the change in the column internal vapor flow which impact both Xd & Xb prediction. Furthermore causes internal reflux ratio variation which impact tray temperatures in the stripping section. It provides estimation for the internal reflux ratio at each tray in the stripping section.

$$\frac{L_{i-1}}{V_i} \cong \frac{V_r+B}{V_r} \quad (78)$$

All the above variables can be obtained from online operating conditions. However feed composition values can be obtained from lab analysis or GC analyzer. Lab analysis are conducted few times a day and hence can't be considered reliable source. In addition GC analyzer installation and maintenance associate with high cost in addition to being unreliable. Hence models with feed composition input associate with high cost and unreliable input.

Below are the inferential models generated for Xd & Xb. The models generated for C4 splitter (high and low purity ranges) and DeC3 columns. The experimental plan, assumptions and simulation cases used in this chapter are the same in the section 3 for both columns.

### 4.3 Inferential Model for Butane Splitter Column

The C4 splitter column operates for narrow boiling points components iC4 & nC4.

Hence the tray to tray temperature variation are not significant which indicate low  $\frac{\Delta T}{\Delta x_d} C^\circ$

and  $\frac{\Delta T}{\Delta x_b} C^\circ$ . The temperature variation in most cases are less than 1 C° however causes

significant composition variation. The inferential models generated for C4 splitter assume high accuracy temperature measurement in industrial plants to avoid model inaccuracy. The models were generated using 144 cases of variable feed composition,  $X_d$  &  $X_b$  as per the experimental plant. While the prediction data set included 30 cases of randomly selected feed compositions,  $X_d$  &  $X_b$  which fall within the specified experimental ranges.

#### 4.3.1 $X_d$ Model for Low Purity Range (0.03 – 0.08)

The below analysis was conducted at constant feed composition and variable  $X_b$ . The selected feed composition is  $HK\% = 76\%$  as an example. This is to select a sensitive tray that shows significant  $\frac{\Delta T C^\circ}{\Delta x_d}$ , shows linear behavior and not impacted by  $X_b$  variation. Fig. 67 shows the column top section trays temperature variation for  $X_d$  variation  $\frac{\Delta T C^\circ}{\Delta x_d}$ .  $X_d$  variation extended over the range of 0.03 – 0.08. The below plot indicates the analysis for 36 cases of variable  $X_d$  and  $X_b$ . As shown the maximum  $\frac{\Delta T C^\circ}{\Delta x_d}$  value is 0.537  $C^\circ$  per  $X_d$  variation of 0.01 at tray 15. It is small magnitude and hence temperature measurement accuracy plays major role in the model accuracy. The trays numbering follows top to bottom pattern where tray 1 is the first tray below the condenser while tray 51 is the feed tray.

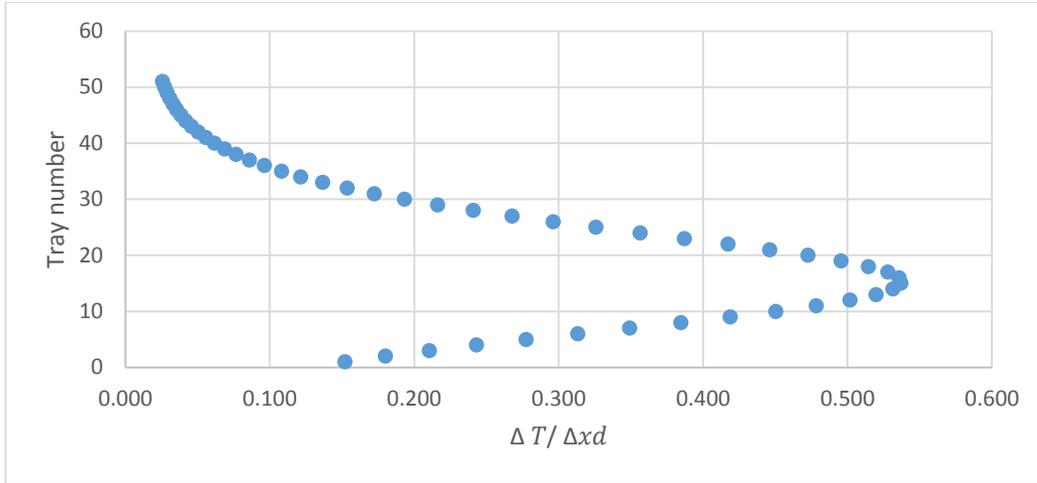


Figure 67. Relationship between tray number and  $\Delta T / \Delta x_d$  in the top section for constant feed composition and variable Xb (C4 splitter low purity range)

Tray 15 showed the highest  $\frac{\Delta T C^o}{\Delta x_d}$  which is expected to show nonlinear behavior. In addition trays 12 & 10 were selected for comparison. Generally the trays above tray 15 show lower  $\frac{\Delta T C^o}{\Delta x_d}$  and so show higher linear behavior. Fig. 68 shows Xd equation as Eq. (75) as function of  $T_{10}, T_{12},$  &  $T_{15}$ .

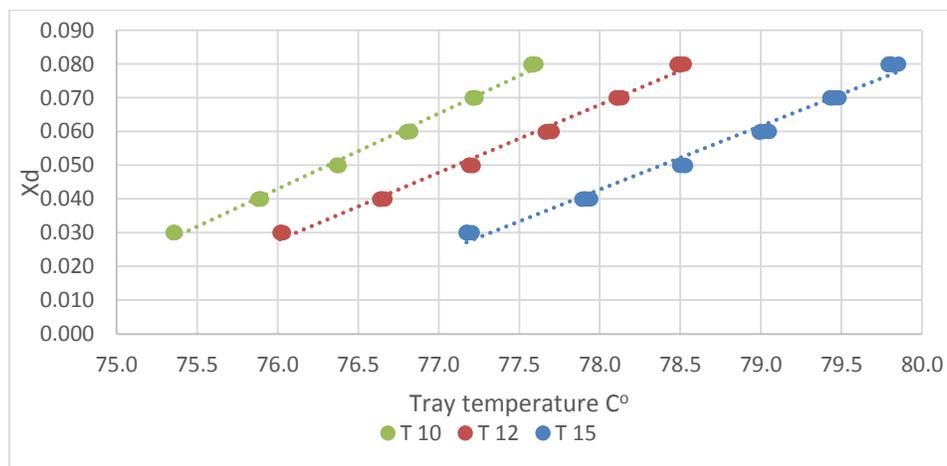


Figure 68. Relationship between Xd and trays temperature for constant feed composition and variable Xb (C4 splitter low purity range)

As the tray number decrease toward the column end the more linear behavior is observed. The horizontal spread at each Xd value is caused by Xb variation. As shown the lower tray number the less impact is caused by Xb variation. Xd equation was tested for  $T_{10}$ ,  $T_{12}$ , &  $T_{15}$  as shown in Table. 45.

Table 45. Xd equation as function of T10, T12 & T15 accuracy analysis (C4 splitter low purity range)

Equation	R <sup>2</sup>	RMSEE	Average Xd	$\frac{\Delta T C^o}{\Delta x_d}$
$x_d = 0.0224 T_{10} - 1.6566$	0.9954	0.00116	0.055	0.450
$x_d = 0.0201 T_{12} - 1.5002$	0.9918	0.00155	0.055	0.502
$x_d = 0.0188 T_{15} - 1.4258$	0.9839	0.00217	0.055	0.537

Tray 10 shows higher R<sup>2</sup> and lower RMSEE as compared to average Xd value. However it shows lower  $\frac{\Delta T C^o}{\Delta x_d}$  with small magnitude as compared to  $T_{12}$ , &  $T_{15}$ . Hence selecting  $T_{10}$  indicates high accuracy Xd model for constant feed composition and variable Xb.

Next step is testing Xd equation with  $T_{10}$  for variable Xd (0.03 – 0.08), Xb (0.03 – 0.08) and feed compositions Feed 1, Feed 2, Feed 3 and Feed 4. Fig. 69 shows the relationship between Xd and  $T_{10}$  for 144 cases of variable Xd, Xb and feed composition.

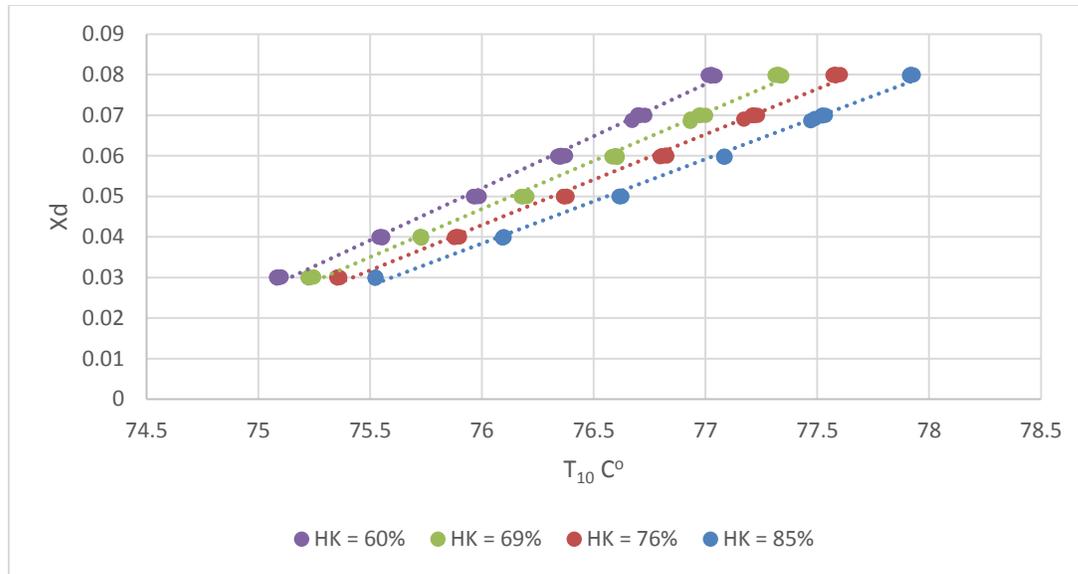


Figure 69. Xd equation as function of T<sub>10</sub> for variable feed composition and Xb (C4 splitter low purity range)

Xd equation as Eq. (75) experience shifting due to feed composition variation. In order to test Xd equation a single linear regression was generated for the above data.

$$x_d = 0.0205 T_{10} - 1.5139 \quad \text{Variable feed composition, Xd \& Xb} \quad (79)$$

$$R^2 = 0.888, \quad \text{RMSEE} = 0.00569 \quad \text{Xd average value} = 0.055$$

Eq. (79) shows low accuracy and hence additional variables are required to improve the model accuracy. The first linear Xd equation Eq. (80) with two variables T<sub>10</sub> &  $\frac{R}{R+D}$  was generated using multivariate software (ProMV). The PLS model was generated using 144 cases of variable feed composition, Xd & Xb and tested for 30 cases of prediction data set.

$$x_d = a_1 T_{10} + a_2 \frac{R}{R+D} + b \quad (80)$$

$$a_1 = 0.022175 \quad a_2 = -0.267881 \quad b = -1.389575$$

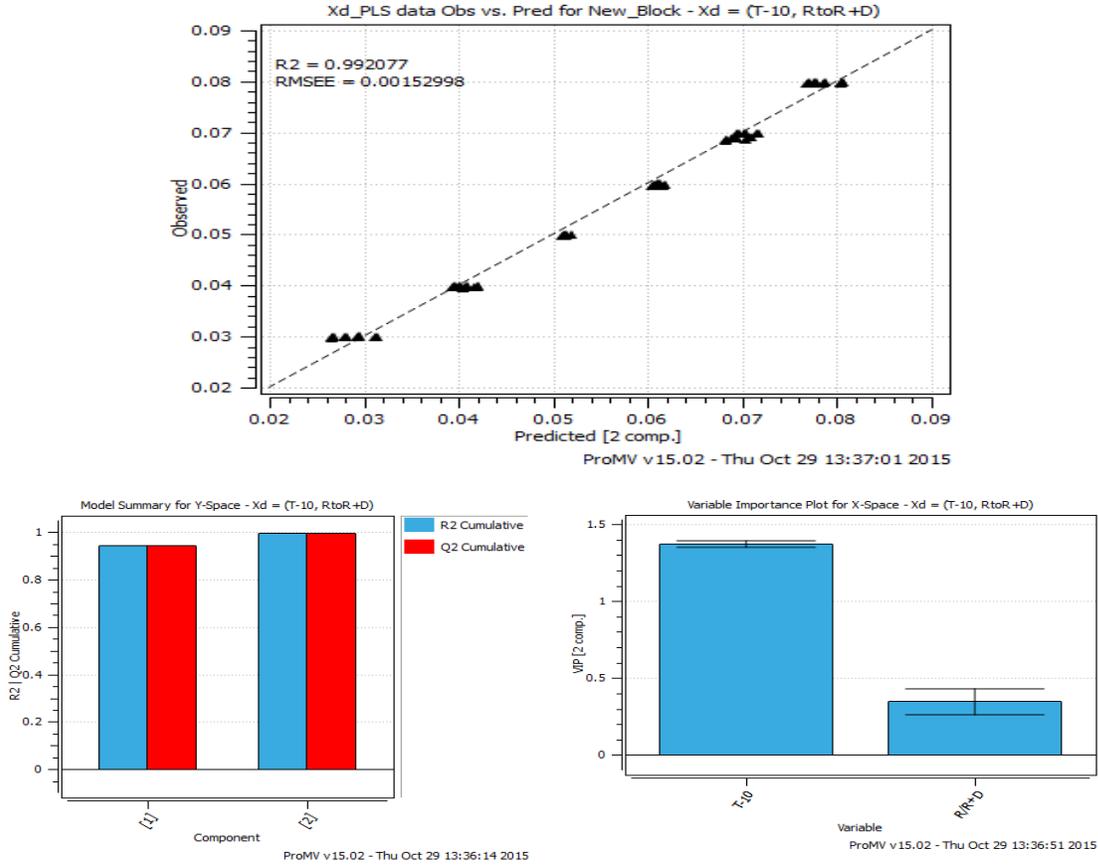


Figure 70. ProMV PLS model for Xd as function of two variables plot and statistical analysis for variable feed composition, Xd & Xb (C4 splitter low purity range)

The second Xd equation Eq. (81) with four variables  $T_{10}$ ,  $T_{15}$ ,  $\frac{R}{R+D}$  &  $\frac{V_r+B}{V_r}$  was generated using multivariate software (ProMV). Table. 46 illustrates accuracy comparison for Eqs. (80) and (81) which were tested for 144 (training data) and (30 prediction data) of variable feed composition, Xd and Xb.

$$x_d = a_1 T_{10} + a_2 T_{15} + a_3 \frac{R}{R+D} + a_4 \frac{V_r+B}{V_r} + b \quad (81)$$

$$a_1 = 0.046583 \quad a_2 = -0.020224 \quad a_3 = 0.019989 \quad a_4 = 0.005981 \quad b = -1.945769$$

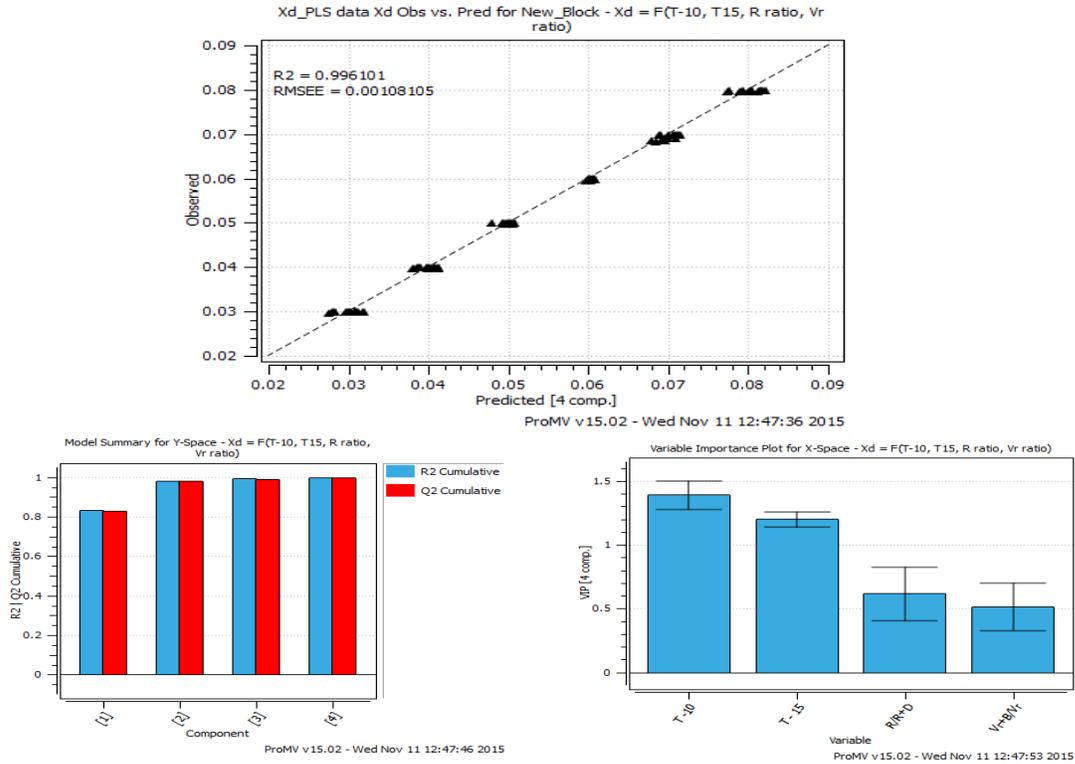


Figure 71. ProMV PLS model for  $x_d$  as function of four variables plot and statistical analysis for variable feed composition,  $x_d$  &  $x_b$  (C4 splitter low purity range)

Table 46. ProMV PLS models for  $x_d$  for 144 cases of variable feed composition,  $x_d$  &  $x_b$  & 30 prediction data set (C4 splitter low purity range)

Equation	$x_d = F\left(T_{10}, \frac{R}{R+D}\right)$	$x_d = F\left(T_{10}, T_{15}, \frac{R}{R+D}, \frac{V_r+B}{V_r}\right)$
$R^2$	0.992076	0.996101
$Q^2$	0.992067	0.995658
RMSEE	0.00152998	0.00108105
RMSEP	0.001262402	0.00070
Average value	0.055	0.055
Components #	2	4

Eq. (81) variables showed importance as per the variable importance plot VIP in Fig.70. It indicated strong correlation and prediction ability with high  $R^2$  &  $Q^2$  and low RMSEE & RMSEP as compared to average  $X_d$  value. Eq. (81) variables showed importance as per the variable importance plot VIP in Fig. 71. The model was built using 4 components to explain the model nonlinearity and shifting effect caused by feed variation. As shown in Fig. 71,  $T_{10}$  has higher importance than  $T_{15}$  and it contributes more to the inferred value of  $X_d$  since coefficient  $a_1$  is higher. Introducing  $T_{15}$  improves the model robustness as it showed higher  $\frac{\Delta T C^0}{\Delta x_d}$ . The ratio  $\frac{V_r+B}{V_r}$  explains the horizontal shifting in  $T_{15}$  due to  $X_b$  variation as shown earlier. The model indicated strong correlation and prediction ability. In addition indicated higher accuracy with small magnitude as compared to Eq. (80). Hence it is recommended to use  $x_d = F(T_{10}, T_{15}, \frac{R}{R+D}, \frac{V_r+B}{V_r})$  as Eq. (81) as  $X_d$  inferential model for C4 splitter low purity range.

Below is the inferential model with feed composition as input. It is shown for information purposes and to compare its accuracy to PLS model. As shown previously  $X_d$  equation is linear function of  $T_{10}$ . However it experiences shifting due to feed composition variation. The slope and intercept values for each  $X_d$  equation for variable feed composition are shown in Table 47.

Table 47.  $X_d = F(T_{10})$  slope and intercept for variable feed composition &  $X_b$  (C4 splitter low purity range)

HK%	Slope a	Intercept b
60%	0.025649	-1.89737
69%	0.023734	-1.75693
76%	0.02236	-1.65643
85%	0.020801	-1.54249

Figs. 72 and 73 show the relationship between slope  $a$  and intercept  $b$  and the feed HK%. Eqs. (82) and (83) were generated to express slope  $a$  and intercept  $b$  as function of feed HK.

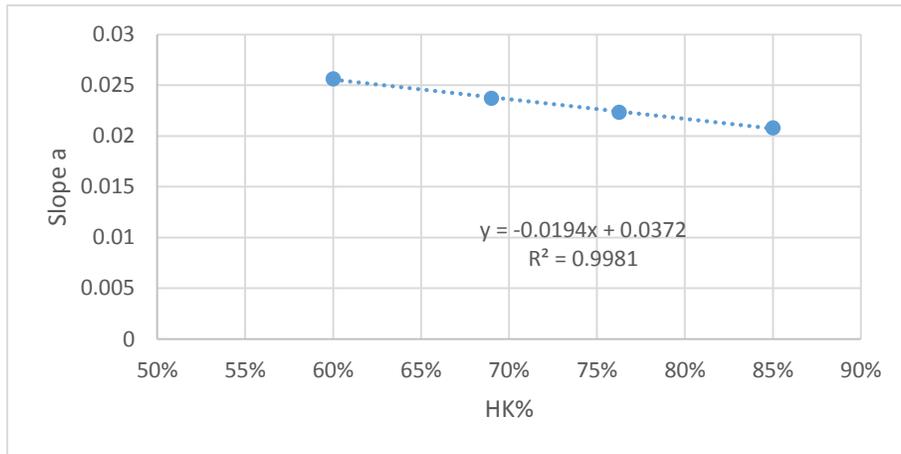


Figure 72. Relationship between  $X_d$  equation slope and feed HK% (C4 splitter low purity range)

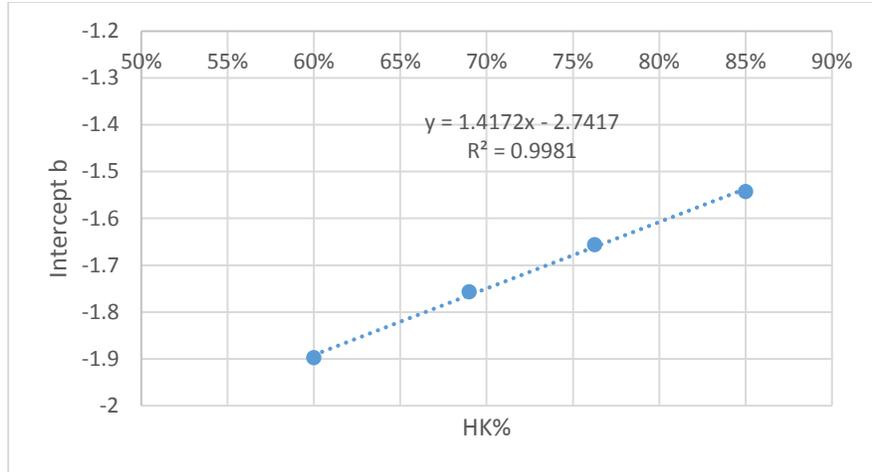


Figure 73. Relationship between Xd equation intercept and feed HK% (C4 splitter low purity range)

$$\text{Slope } a = -0.0194 (KH) + 0.0372 \quad (82)$$

$$R^2 = 0.9981$$

$$\text{Intercept } b = 1.4172 (HK) - 2.7417 \quad (83)$$

$$R^2 = 0.9981$$

The model (Eqs. (75), (82) and (83)) was tested for 144 cases (training data) and 30 cases (prediction data) of variable feed composition, Xd & Xb. This model showed strong correlation and slightly lower accuracy than Eq. (81) as shown in Table 48. However the model requires feed HK composition input which is not a reliable input in practice. In addition, it is a nonlinear model with polynomial of 2<sup>nd</sup> degree. It is recommended to implement the PLS model with four variables eq. (81) considering its accuracy, linearity and input reliability.

Table 48. Xd model with feed HK input accuracy analysis (C4 splitter low purity range)

Equation	$x_d = a T_{10} + b$ <i>Slope a = -0.0194 (HK) + 0.0372</i> <i>Intercept b = 1.4172 (HK) - 2.741</i>
R <sup>2</sup>	0.9954
Q <sup>2</sup>	0.9925
RMSEE	0.0011588
RMSEP	0.001221756
Average value	0.055
Polynomial degree	2

#### 4.3.2 Xb Model for Low Purity Range (0.03 – 0.08)

The below analysis was conducted at constant feed composition and variable Xb. The selected feed composition is LK% = 24% as an example. This is to select a sensitive tray that shows significant  $\frac{\Delta T C^o}{\Delta x_b}$ , shows linear behavior and is not impacted by Xd variation. Fig. 74 shows the column bottom section trays temperature variation for Xb variation  $\frac{\Delta T C^o}{\Delta x_b}$ . Xb variation extended over the range of 0.03 – 0.08. The below plot indicates the analysis for 36 cases of variable Xd and Xb. The trays numbering follows top to bottom pattern where tray 51 is the feed tray and tray 99 is last tray above the reboiler.

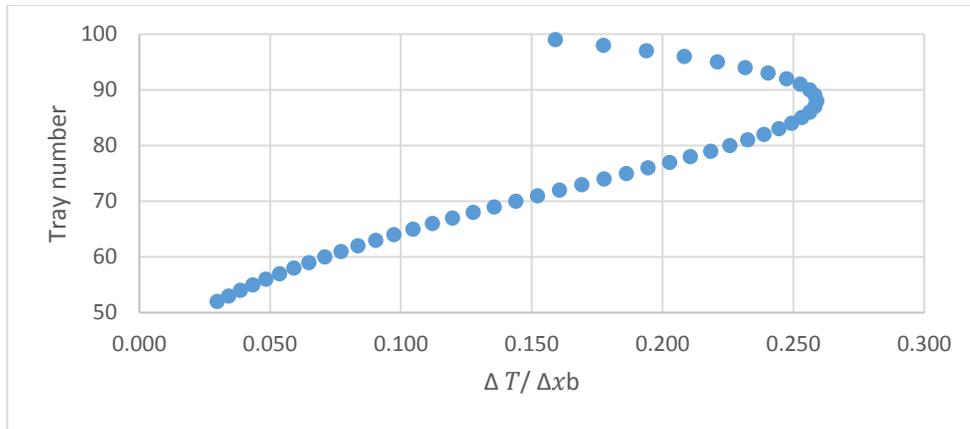


Figure 74. Relationship between tray number and  $\Delta T / \Delta x_b$  in the bottom section for constant feed composition & variable  $X_d$  (C4 splitter low purity range)

As shown the maximum  $\frac{\Delta T C^o}{\Delta x_b}$  value is 0.259 C° per  $X_b$  variation of 0.01 at tray 88. It is very small magnitude and hence temperature measurement accuracy plays major role in the model accuracy. The main drawback of this model is that small error from temperature measurement can cause significant error in  $X_b$  prediction. Fig. 75 indicated  $X_b$  equation Eq. (76) as function of sensitive tray  $T_{88}$  for constant feed composition and variable  $X_d$ .

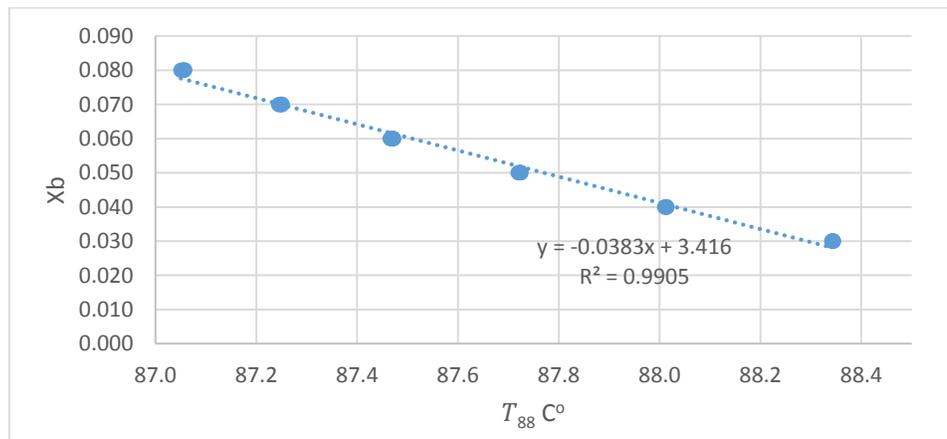


Figure 75. Relationship between  $X_d$  and  $T_{88}$  for constant feed composition and variable  $X_d$  (C4 splitter low purity range)

The model was generated with the assumption of having a very accurate temperature measurement. For constant feed composition,  $X_b$  as function of  $T_{88}$  showed strong correlation. Choosing tray below  $T_{88}$  toward the column end will indicate stronger linear correlation however the temperature sensitivity will be very low. Fig. 76 shows  $X_b$  equation as Eq. (76) for variable feed composition and  $X_d$ .

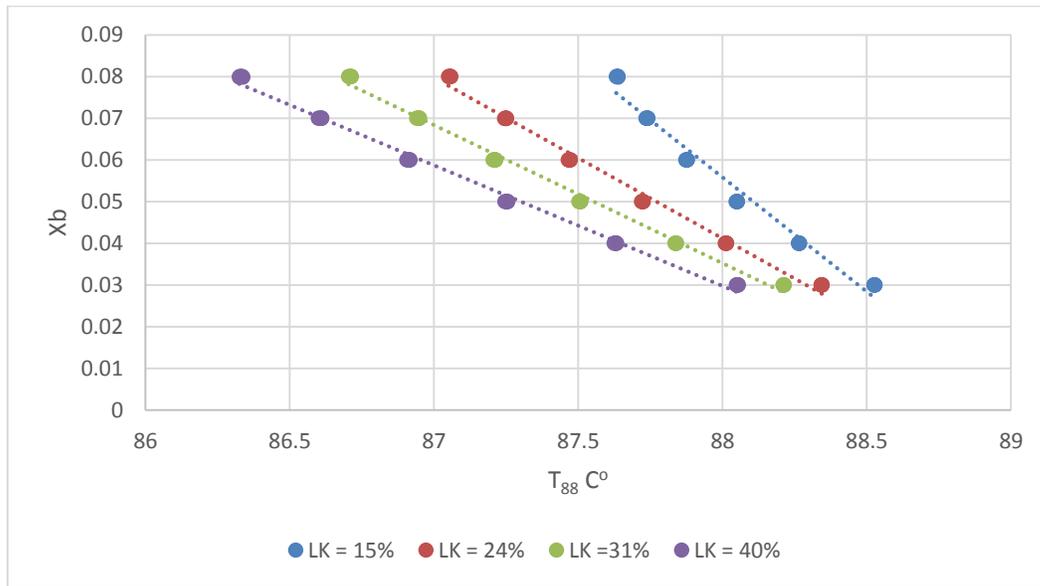


Figure 76.  $X_b$  equation as function of  $T_{88}$  for variable feed composition &  $X_d$  (C4 splitter low purity range)

$X_b$  equation experience shifting due to feed composition variation. The equation exhibits nonlinear behavior at low LK. In order to test  $X_b$  equation a single linear regression as Eq. (84) was generated for the above data.

$$x_b = -0.0237 T_{88} + 2.13780 \quad \text{Variable feed composition \& } X_d \quad (84)$$

$$R^2 = 0.6407$$

$$RMSEE = 0.010234$$

$$X_b \text{ average value} = 0.055$$

It showed low accuracy hence a linear Xb equation with two variables as Eq. (85) was generated using a multivariate software (ProMV).

$$x_b = a_1 T_{88} + a_2 \frac{V_r + B}{V_r} + b \quad (85)$$

$$a_1 = -0.028434 \quad a_2 = 0.1843898 \quad b = 2.324372$$

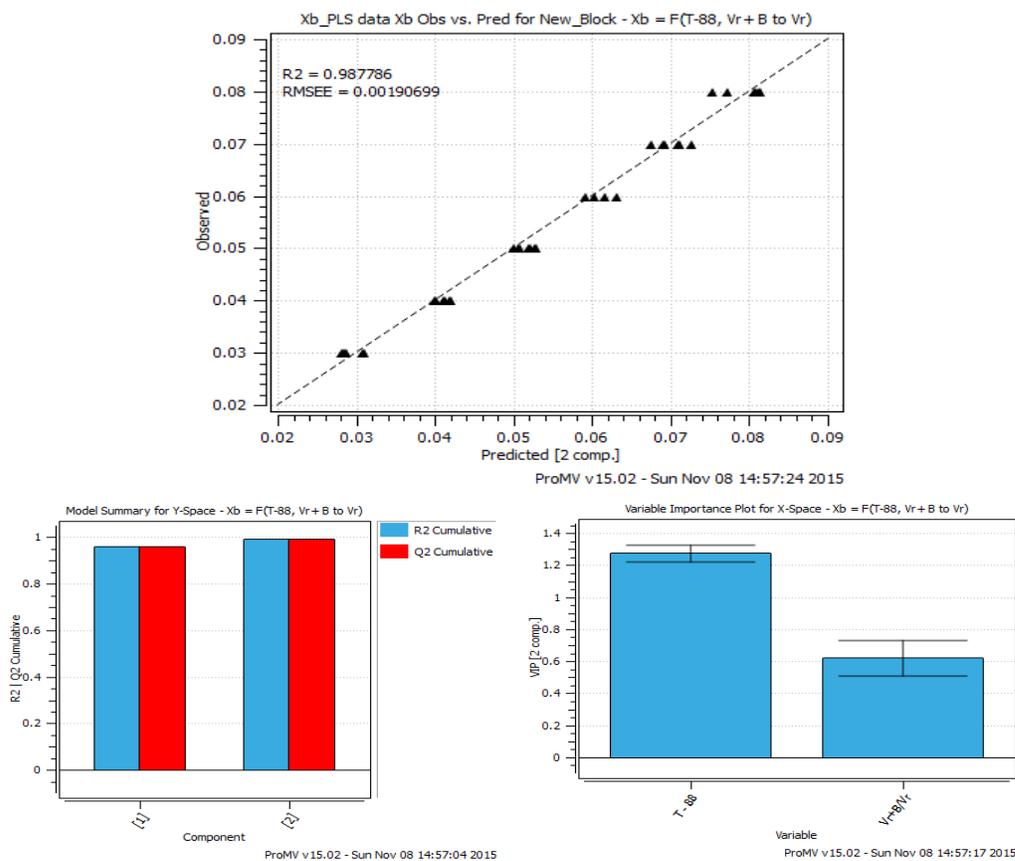


Figure 77. ProMV PLS model for Xb as function of two variables plot and statistical analysis (C4 splitter low purity range)

Eq. (85) variables indicated importance as per the variable importance plot VIP in Fig. 77. The model generated with two components with high  $R^2$  &  $Q^2$  and low RMSEE as compared to average Xb value. The model accuracy can be increased by adding other

tray temperature closer to the column end however the temperature sensitivity will be very small. Table. 49 illustrates Eq. (85) model accuracy analysis for 144 cases and 30 cases of prediction set.

Table 49. ProMV PLS models for Xb for 144 cases of variable feed composition, Xd & Xb & 30 prediction data set (C4 splitter low purity range)

Equation	$x_b = a_1 T_{88} + a_2 \frac{V_r + B}{V_r} + b$
R <sup>2</sup>	0.98778
Q <sup>2</sup>	0.98775
RMSEE	0.00190699
RMSEP	0.00157
Average value	0.055
Components #	2

Below is inferential model with feed composition as an input. As shown earlier Xb equation slope and intercept vary with feed composition LK. Figs. 78 and 79 illustrate the relationship between slope and intercepts as function of feed LK. Then slope and intercepts equations were generated as Eqs. (86) and (87).

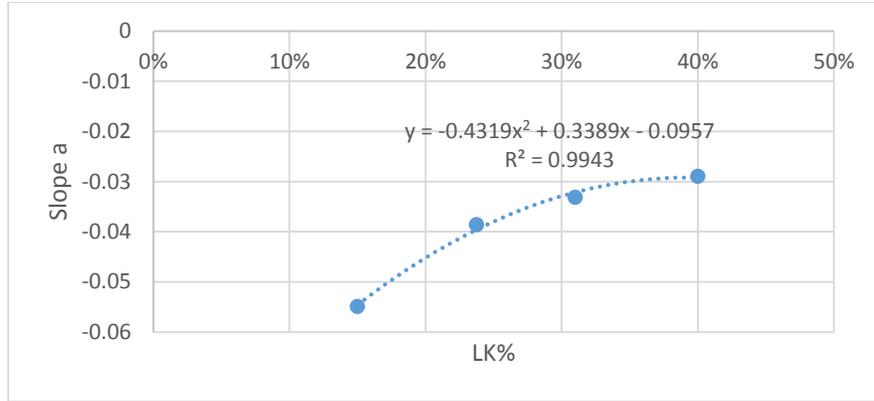


Figure 78. Relationship between Xb equation slope a and feed LK composition (C4 splitter low purity range)

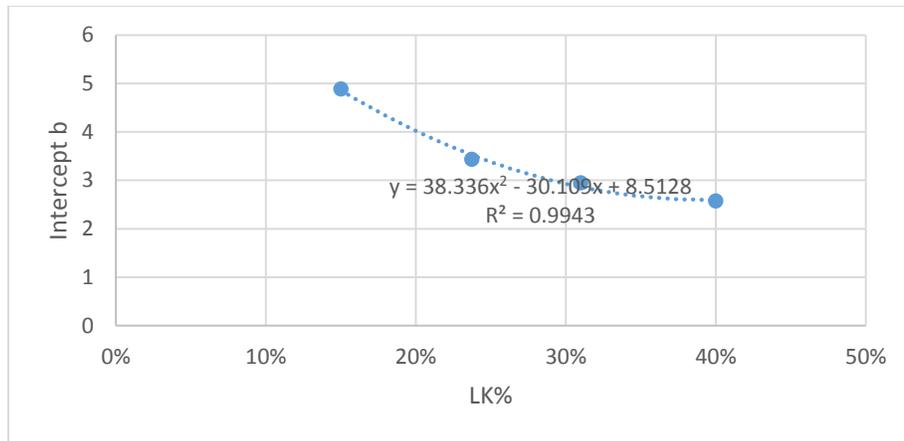


Figure 79. Relationship between Xb equation intercept b and feed LK composition (C4 splitter low purity range)

$$\text{Slope } a = -0.4319(LK)^2 + 0.3389(LK) - 0.0957 \quad (86)$$

$$R^2 = 0.9943$$

$$\text{Intercept } b = 38.336(LK)^2 - 30.109(LK) + 8.5128 \quad (87)$$

$$R^2 = 0.9943$$

The slope and intercepts are 2<sup>nd</sup> degree polynomial to express Xb equation nonlinearity at low LK. The model was tested for 144 cases of variable feed composition, Xd & Xb and 30 prediction data set as shown in Table. 50. This model showed slightly lower accuracy as compared to Xb PLS model as Eq. (85). It is nonlinear with polynomial of 3<sup>rd</sup> degree. It is recommended to implement the linear PLS model Eq. (85) considering its accuracy, linearity and input reliability.

Table 50. Xb model with slope and intercept as function of feed LK accuracy analysis for 144 cases and 30 prediction set data (C4 splitter low purity range)

Equation	$x_b = a T_{88} + b$ <i>Slope</i> $a = -0.4319(LK)^2 + 0.3389(LK) - 0.0957$ <i>Intercept</i> $b = 38.336(LK)^2 - 30.109(LK) + 8.5128$
R <sup>2</sup>	0.9844
Q <sup>2</sup>	0.9811
RMSEE	0.002144679
RMSEP	0.002387821
Average value	0.055
Polynomial degree	3

#### 4.3.3 Xd Model for High Purity Range (0.005 – 0.03)

The below analysis was conducted at constant feed composition and variable Xb. The selected feed composition is HK% = 76% as an example. This is to select a sensitive tray that shows significant  $\frac{\Delta T C^o}{\Delta x_d}$ , shows linear behavior and not impacted by Xb variation. The below plot shows the column top section trays temperature variation for

$\Delta T C^o$   
 $\Delta x_d$  Xd variation. Xd variation extended over the range of 0.005 – 0.03. Fig. 80 indicates the analysis for 36 cases of variable Xd and Xb. The trays numbering follows top to bottom pattern where tray 1 is the first tray below the condenser while tray 51 is the feed tray.

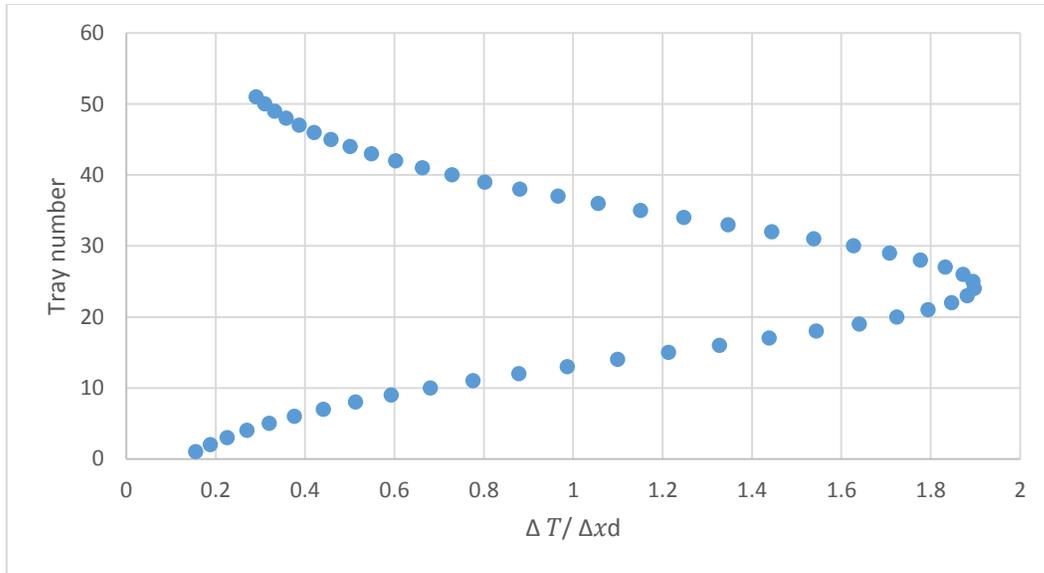


Figure 80. Relationship between tray number and  $\Delta T / \Delta x_d$  in the top section for constant feed composition (C4 splitter high purity range)

As shown the maximum  $\frac{\Delta T C^o}{\Delta x_d}$  value is 1.897 C° per Xd variation of 0.01 at tray 24 which is significant magnitude. Tray 24 however showed nonlinear behavior and is impacted by Xb variation as shown in Fig. 81. The horizontal spread at each Xd value is caused by Xb variation.

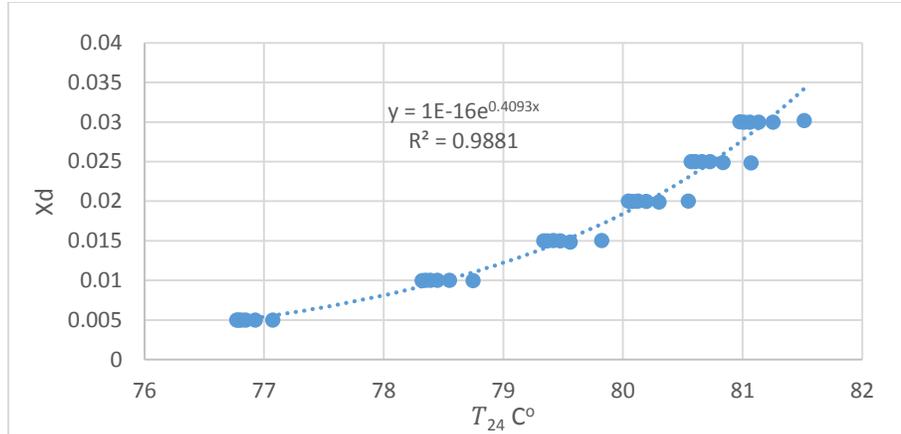


Figure 81. Relationship between  $X_d$  and  $T_{24}$  for constant feed composition and variable  $X_b$  (C4 splitter high purity range)

$X_d$  as function of  $T_{24}$  is represented by exponential regression for constant feed composition and variable  $X_d$ .  $X_b$  equation if represented as function of  $T_{24}$  experience shifting due to variation in feed composition and  $X_b$ . This leads to the same structure of nonlinear offline hybrid model.

As the tray number decrease toward the column end the more linear behavior is observed. Generally the trays above tray 24 show lower  $\frac{\Delta T C^o}{\Delta x_d}$  and more linear behavior with less impact caused by  $X_b$  variation. Trays 10, 12 & 15 were selected for comparison as shown in Fig. 82.

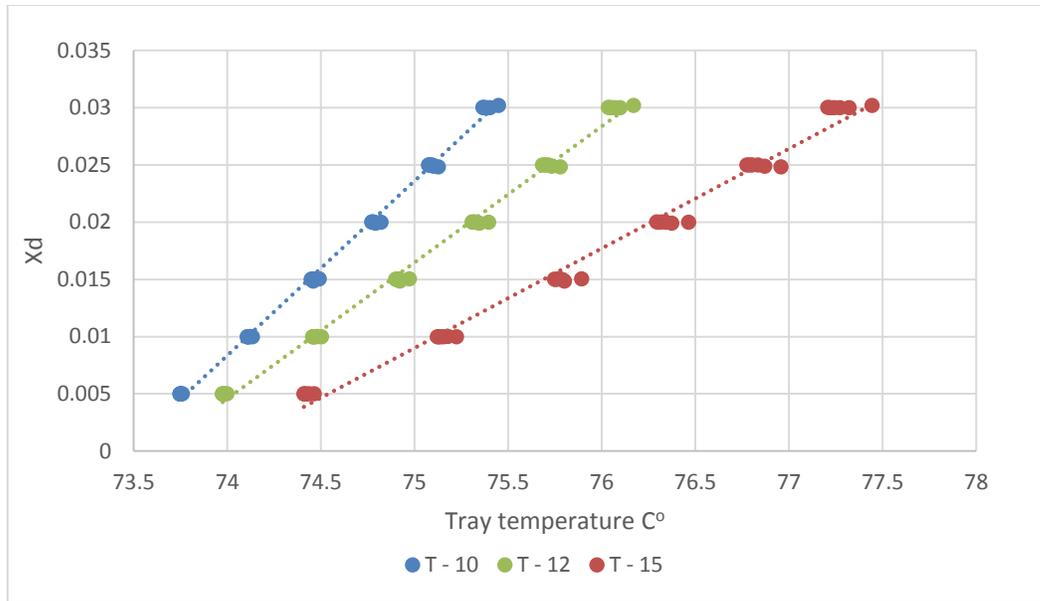


Figure 82. Relationship between X<sub>d</sub> and T<sub>10</sub>, T<sub>12</sub> & T<sub>15</sub> for constant feed composition and variable X<sub>b</sub> (C4 splitter high purity range)

X<sub>d</sub> equations were generated from the above plot and tested individually for X<sub>d</sub> as function of T<sub>10</sub>, T<sub>12</sub>, & T<sub>15</sub> as Eq. (75) as shown below.

Table 51. X<sub>d</sub> as function of T<sub>10</sub>, T<sub>12</sub> & T<sub>15</sub> accuracy comparison (C4 splitter high purity range)

Equation	R <sup>2</sup>	RMSEE	Average X <sub>d</sub>	$\frac{\Delta T C^\circ}{\Delta x_d}$
$x_d = 0.0152 T_{10} - 1.1193$	0.9974	0.000434225	0.0175	0.68063808
$x_d = 0.0119 T_{12} - 0.8758$	0.995	0.000607586	0.0175	0.87861684
$x_d = 0.0087 T_{15} - 0.6432$	0.9882	0.000933161	0.0175	1.21365196

Tray 10 shows higher  $R^2$  and lower RMSEE as compared to average  $X_d$  value. However it shows lower  $\frac{\Delta T C^o}{\Delta x_d}$  as compared to  $T_{12}$ , &  $T_{15}$ . Hence selecting  $T_{10}$  indicates high accuracy  $X_d$  model for constant feed composition and variable  $X_b$ .

Next step is testing  $X_d$  equation with  $T_{10}$  for variable  $X_d$  (0.005 – 0.03),  $X_b$  (0.005 – 0.03) and feed compositions Feed 1, Feed 2, Feed 3 and Feed 4. Fig. 83 shows the relationship between  $X_d$  and  $T_{10}$  for 144 cases of variable  $X_d$ ,  $X_b$  and feed composition.

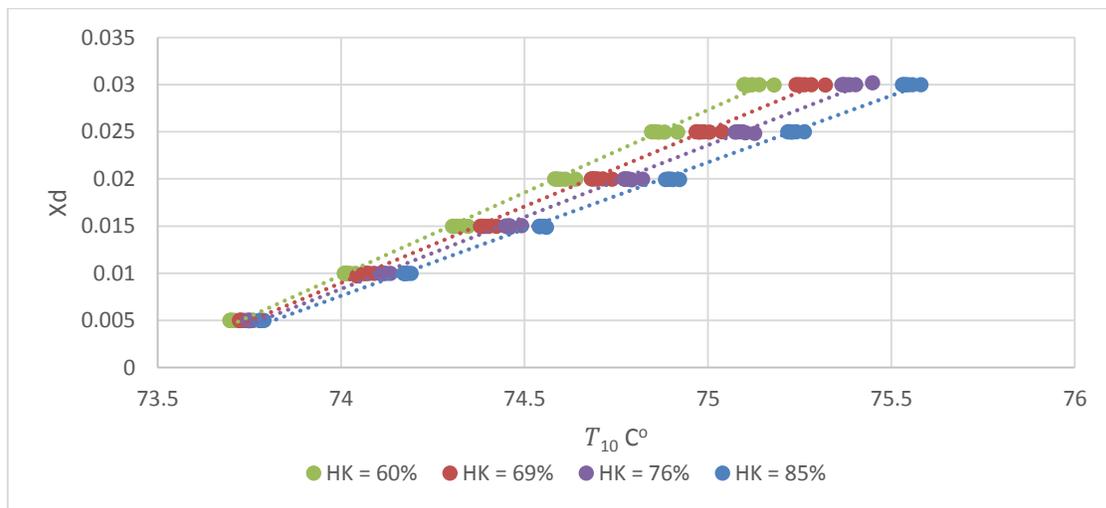


Figure 83. Relationship between  $X_d$  and  $T_{10}$  for variable feed composition and  $X_b$  (C4 splitter high purity range)

$X_d$  equation experience shifting due to feed composition variation. In order to test  $X_d$  equation a single linear regression was generated for the above data as Eq. (88).

$$x_d = 0.0151 T_{10} - 1.1102 \quad \text{Variable feed composition, } X_d \text{ \& } X_b \quad (88)$$

$$R^2 = 0.9618, \quad \text{RMSEE} = 0.00166983 \quad \text{Xd average value} = 0.0175$$

Eq. (88) shows low accuracy and hence additional variables are required to account for feed variation. The below linear Xd equation as Eq. (80a) with two variables  $T_{10}$  and

$\frac{R}{R+D}$  was generated using multivariate software (ProMV) as shown in Fig. 84.

$$x_d = a_1 T_{10} + a_2 \frac{R}{R+D} + b \text{ as eq. (80a)}$$

$$a_1 = 0.015397 \quad a_2 = -0.084348 \quad b = -1.050944$$

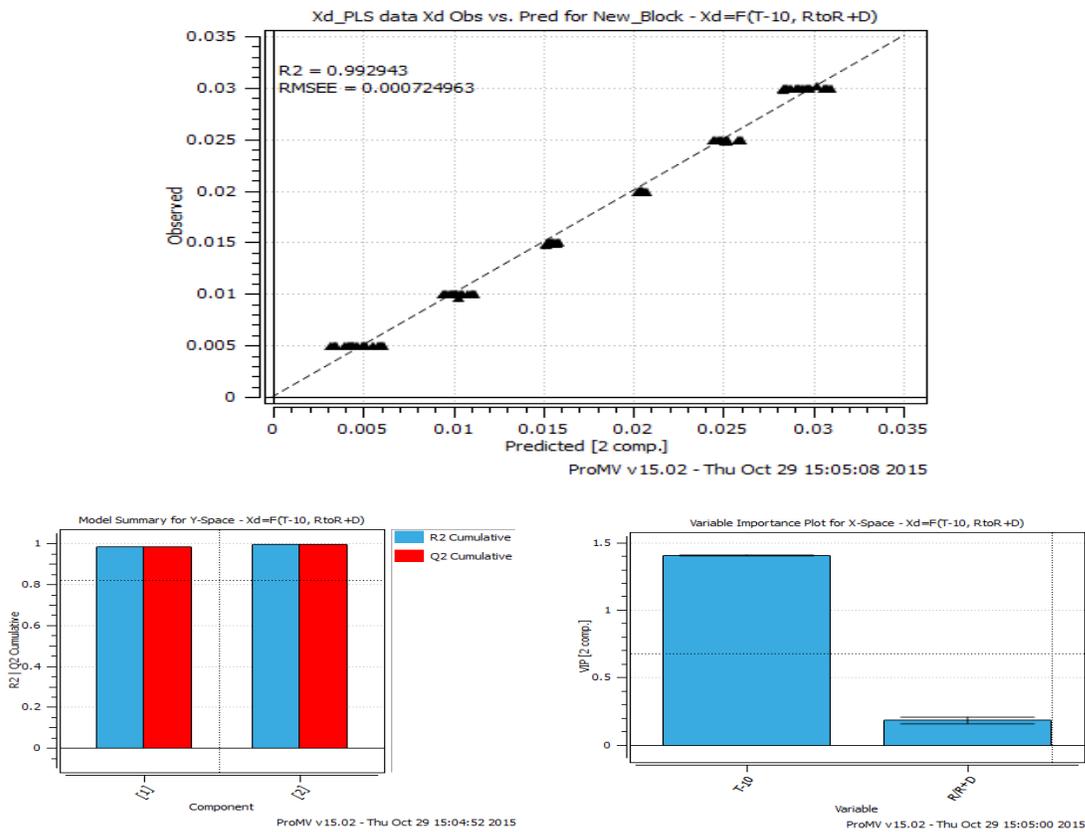


Figure 84. ProMV PLS model for  $X_d$  as function of two variables plot and statistical analysis for variable feed composition,  $X_d$  &  $X_b$  (C4 splitter high purity range)

Below is  $X_d$  equation as Eq. (81a) with four variables  $T_{10}$ ,  $T_{15}$ ,  $\frac{R}{R+D}$  &  $\frac{V_r+B}{V_r}$  was generated using multivariate software (ProMV) as shown in Fig. 85.  $T_{15}$  has higher

sensitivity with  $\frac{\Delta T C^o}{\Delta x_d} = 1.21$  which is significant and hence adding  $T_{15}$  improves the model accuracy. Tray 16 to 24 have higher sensitivity however they exhibit nonlinear behavior which impact the model accuracy

$$x_d = a_1 T_{10} + a_2 T_{15} + a_3 \frac{R}{R+D} + a_4 \frac{V_r+B}{V_r} + b \text{ as eq. (81a)}$$

$$a_1 = 0.032841 \quad a_2 = -0.010035 \quad a_3 = -0.005804 \quad a_4 = 0.001784 \quad b = -1.666533$$

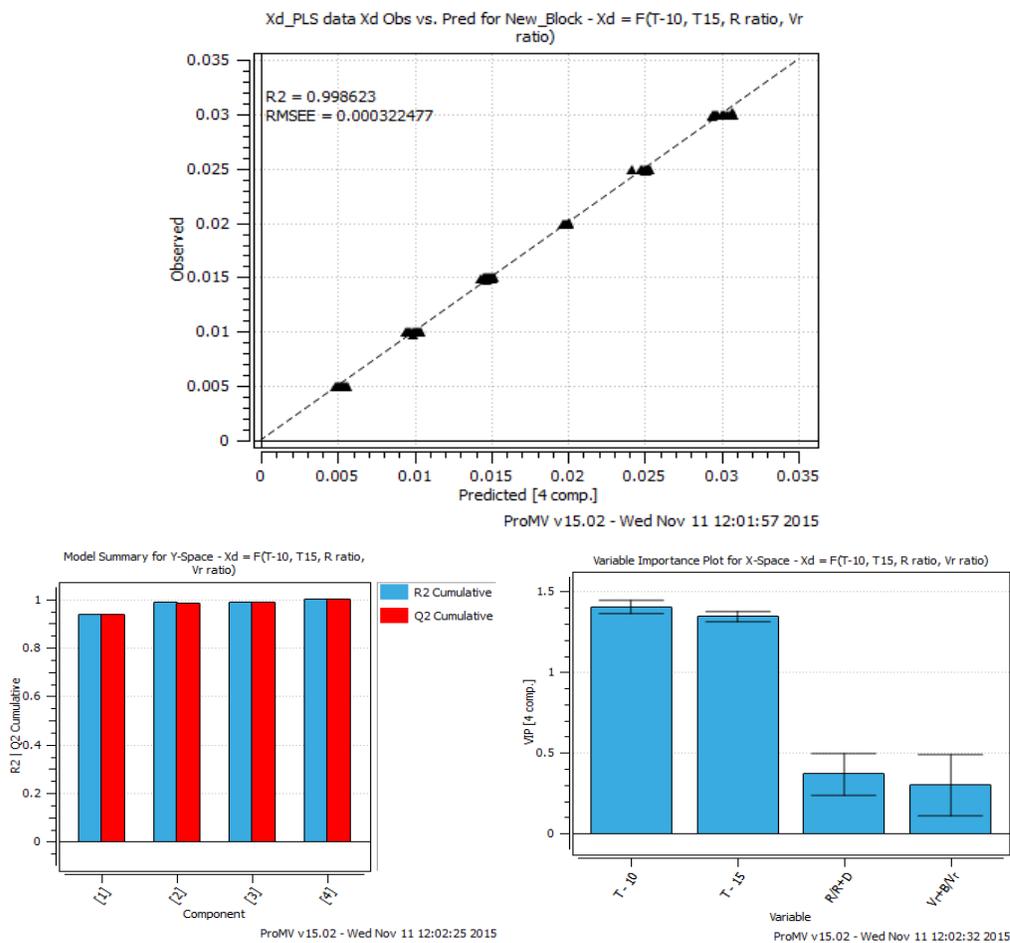


Figure 85. ProMV PLS model for  $X_d$  as function of four variables plot and statistical analysis for variable feed composition,  $X_d$  &  $X_b$  (C4 splitter high purity range)

Table 52. ProMV PLS models for Xd for 144 cases of variable feed composition, Xd & Xb & 30 prediction data set (C4 splitter high purity range)

Equation	$x_d = F(T_{10}, \frac{R}{R+D})$	$x_d = F(T_{10}, T_{15}, \frac{R}{R+D}, \frac{V_r+B}{V_r})$
R <sup>2</sup>	0.992943	0.998623
Q <sup>2</sup>	0.992866	0.998583
RMSEE	0.000724963	0.000322477
RMSEP	0.000430	0.000277
Average value	0.0175	0.0175
Components #	2	4

Eq. (80 a) indicated strong correlation and high prediction ability as shown in Table. 52. However Eq. (81a) indicated higher accuracy with higher R<sup>2</sup> & Q<sup>2</sup> and lower RMSEE & RMSEP as shown in Table. 52. All inputs can be obtained online from the column operating conditions. It is recommended to use equation Eq. (81a) as inferential model for high purity range.

Below is Xd model with feed composition as an input. As shown earlier  $x_d = a T_{10} + b$  experience shifting due to feed composition variation. The slope a and intercept b values are changing due to feed HK variation. The below Table. 53 and Figs. 86 and 87 express the relationship.

Table 53.  $X_d = F(T_{10})$  slope and intercept for variable feed composition &  $X_b$  (C4 splitter high purity range)

HK%	Slope a	Intercept b
60%	0.017528	-1.28726
69%	0.016208	-1.19044
76%	0.015238	-1.11927
85%	0.01415	-1.03948

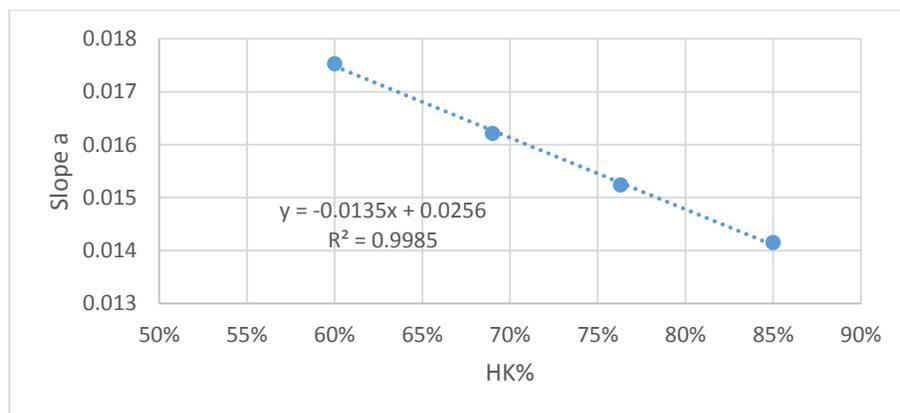


Figure 86. Relationship between  $X_d$  equation slope and feed HK (C4 splitter high purity range)

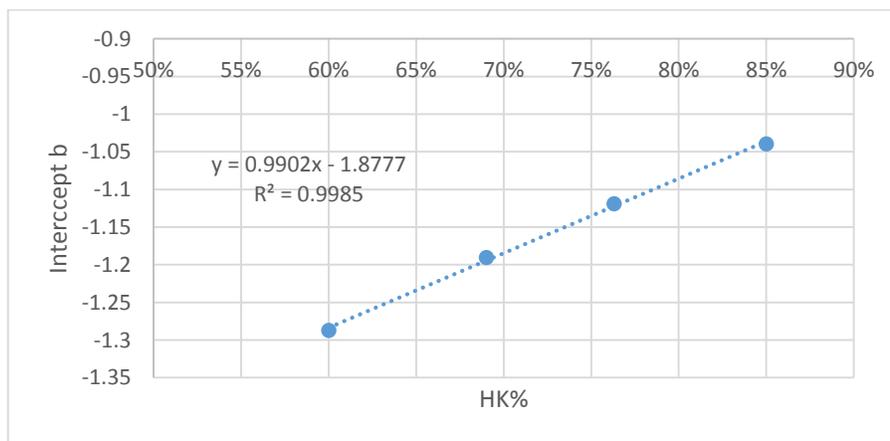


Figure 87. Relationship between  $X_d$  equation intercept and feed HK (C4 splitter high purity range)

$$\text{Slope } a = -0.0135 (KH) + 0.0256 \quad (89)$$

$$R^2 = 0.9985$$

$$\text{Intercept } b = 0.9902 (HK) - 1.8777 \quad (90)$$

$$R^2 = 0.9985$$

Eq. (89) and (90) express the slope  $a$  and intercept  $b$  as function of feed HK composition with linear regression. The model (Eqs. (75), (89) and (90)) with  $T_{10}$  sensitive tray, was tested over 144 cases of variable feed composition,  $X_d$  &  $X_b$  and 30 cases of prediction data set as shown in Table. 54.

Table 54.  $X_d$  model with feed HK input accuracy analysis (C4 splitter high purity range)

Equation	$x_d = a T_{10} + b$ $\text{Slope } a = -0.0135 (HK) + 0.0256$ $\text{Intercept } b = 0.9902 (HK) - 1.8777$
$R^2$	0.9973
$Q^2$	0.9967
RMSEE	0.000443468
RMSEP	0.000401
Average value	0.0175
Polynomial degree	2

This model showed slightly lower accuracy as compared to  $X_d$  PLS model. It is recommended to select  $X_d$  PLS model with four variables Eq. (81a) considering its accuracy, linearity and input reliability.

#### 4.3.4 Xb Model for High Purity Range (0.005 – 0.03)

The below analysis was conducted at constant feed composition and variable Xb. The selected feed composition is LK% = 24% as an example. This is to select a sensitive tray that shows significant  $\frac{\Delta T C^o}{\Delta x_b}$ , shows linear behavior and not impacted by Xb variation. The below plot shows the column bottom section trays temperature variation for Xb variation  $\frac{\Delta T C^o}{\Delta x_b}$ . Xb variation extended over the range of 0.005 – 0.03. Fig. 88 indicates the analysis for 36 cases of variable Xd and Xb. The trays numbering follows top to bottom pattern where tray 51 is the feed tray and tray 99 is the last tray above the reboiler.

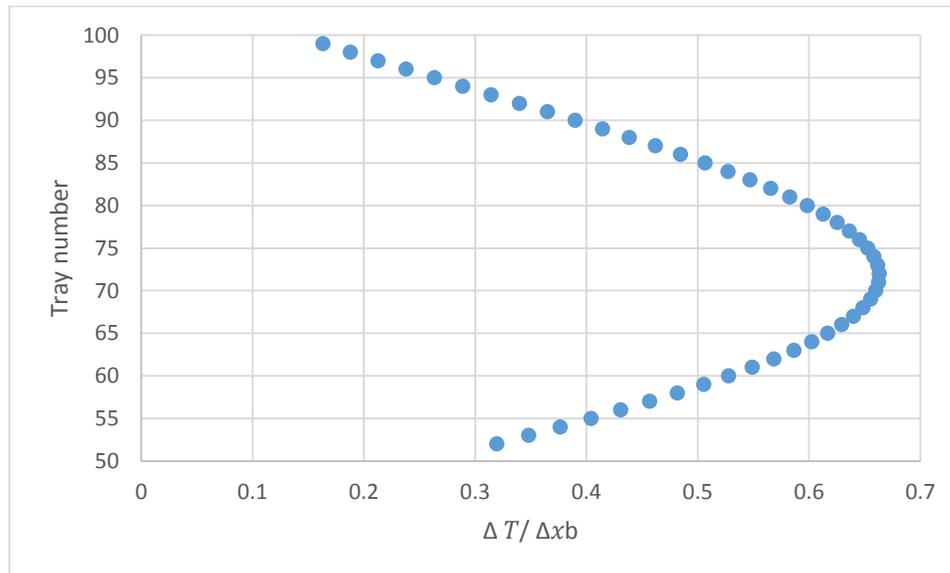


Figure 88. Relationship between tray number and  $\Delta T / \Delta x_b$  in the bottom section for constant feed composition and variable Xd (C4 splitter high purity range)

As shown the maximum  $\frac{\Delta T C^o}{\Delta x_b}$  value is 0.663 C° per Xb variation of 0.01 at tray 72. It is small magnitude and hence temperature measurement accuracy plays major role in the model accuracy. Tray 72 showed nonlinear behavior and it is impacted by Xd variation

as shown in Fig. 89. The horizontal spread at each  $X_d$  value is caused by  $X_b$  variation. As the tray number increase toward the column end in the bottom section the more linear behavior is observed. Generally trays below tray 72 show lower  $\frac{\Delta T C^o}{\Delta x_b}$  and more linear behavior with less impact caused by  $X_d$  variation. Trays 72, 76, 80, 84 & 88 were selected for comparison as shown in Fig. 89.

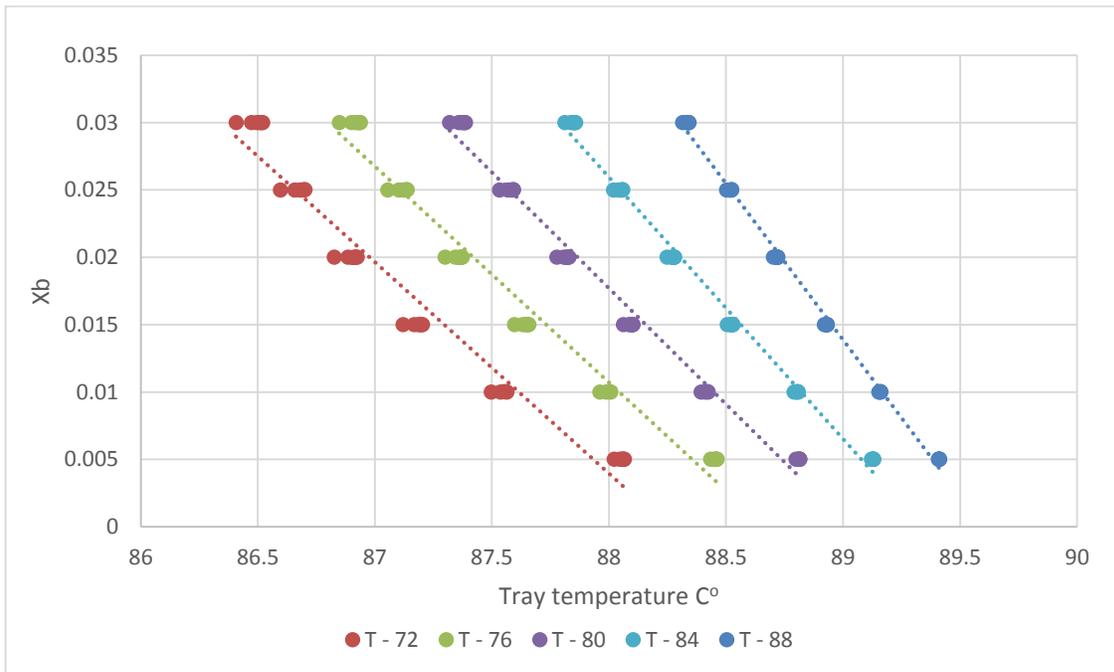


Figure 89. Relationship between  $X_b$  and selected trays in the bottom section for constant feed composition and variable  $X_d$  (C4 splitter high purity range)

$X_b$  equation as Eq. (76) was generated from the above plot and tested individually for  $X_b$  as function of  $T_{72}, T_{76}, T_{80}, T_{84}$  &  $T_{88}$  as shown in Table. 55.

Table 55. Xb as function of T<sub>72</sub>, T<sub>76</sub>, T<sub>80</sub>, T<sub>84</sub> & T<sub>88</sub> accuracy comparison

Equation	R <sup>2</sup>	RMSEE	Average value	$\frac{\Delta T C^o}{\Delta x_d}$
$x_b = -0.0157 T_{72} + 1.3852$	0.9625	0.012243732	0.0175	0.66315508
$x_b = -0.016 T_{76} + 1.4226$	0.9755	0.012215893	0.0175	0.64555604
$x_b = -0.0172 T_{80} + 1.5323$	0.9853	0.012469212	0.0175	0.59862696
$x_b = -0.0194 T_{84} + 1.7349$	0.9922	0.012164164	0.0175	0.52744516
$x_b = -0.0232 T_{88} + 2.0783$	0.9965	0.012047217	0.0175	0.43873788

Tray 88 shows higher R<sup>2</sup> and lower RMSEE as compared to average Xb value. However it shows lower  $\frac{\Delta T C^o}{\Delta x_d}$  as compared to the other trays. Hence selecting T<sub>88</sub> indicates high accuracy Xb model for constant feed composition and variable Xd.

Next step is testing Xb equation with T<sub>88</sub> for variable Xd (0.005 – 0.03), Xb (0.005 – 0.03) and feed compositions Feed 1, Feed 2, Feed 3 and Feed 4. Fig. 90 shows the relationship between Xb and T<sub>88</sub> for 144 cases of variable Xd, Xb and feed composition.

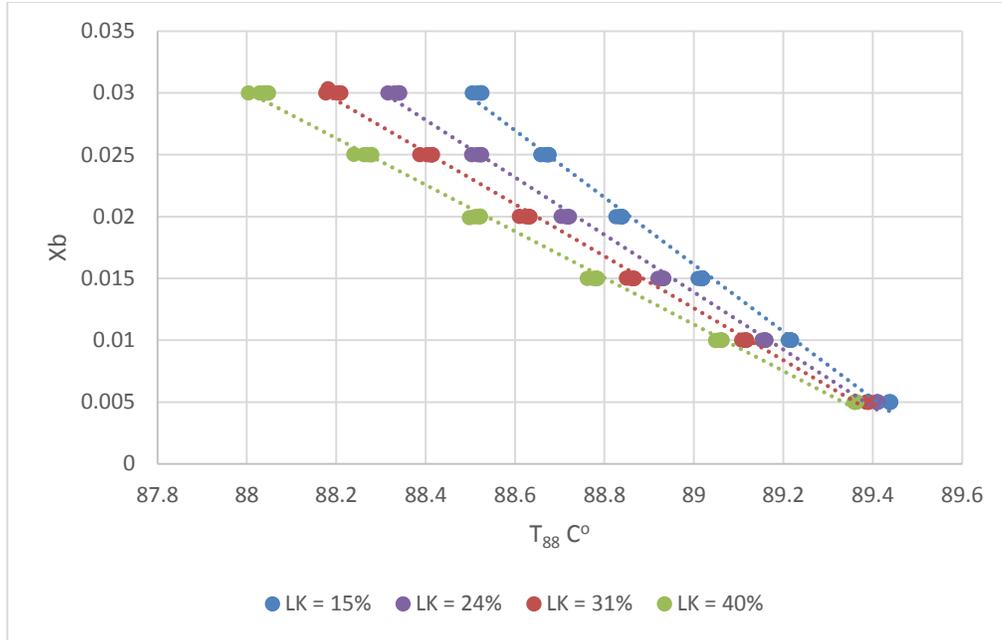


Figure 90. Relationship between  $X_b$  and  $T_{88}$  for variable feed composition &  $X_d$  (C4 splitter high purity range)

Eq. (76) with  $T_{88}$  as sensitive tray experienced shifting due to feed composition variation. In order to test  $X_b$  equation a single linear regression was generated for the above data as Eq. (91).

$$x_b = -0.0203 T_{88} + 1.8226 \quad \text{Variable feed composition, } X_d \text{ \& } X_b \quad (91)$$

$$R^2 = 0.9151, \quad \text{RMSEE} = 0.0024886 \quad X_b \text{ average value} = 0.0175$$

Eq. (91) shows low accuracy and hence additional variables are required to improve model accuracy. The below linear  $X_b$  equation as Eq. (85a) with two variables was generated using multivariate software (ProMV) as shown in Fig. 91.

$$x_b = a_1 T_{88} + a_2 \frac{V_r + B}{V_r} + b \text{ as eq. (85a)}$$

$$a_1 = -0.019387 \quad a_2 = 0.081740 \quad b = 1.645454$$

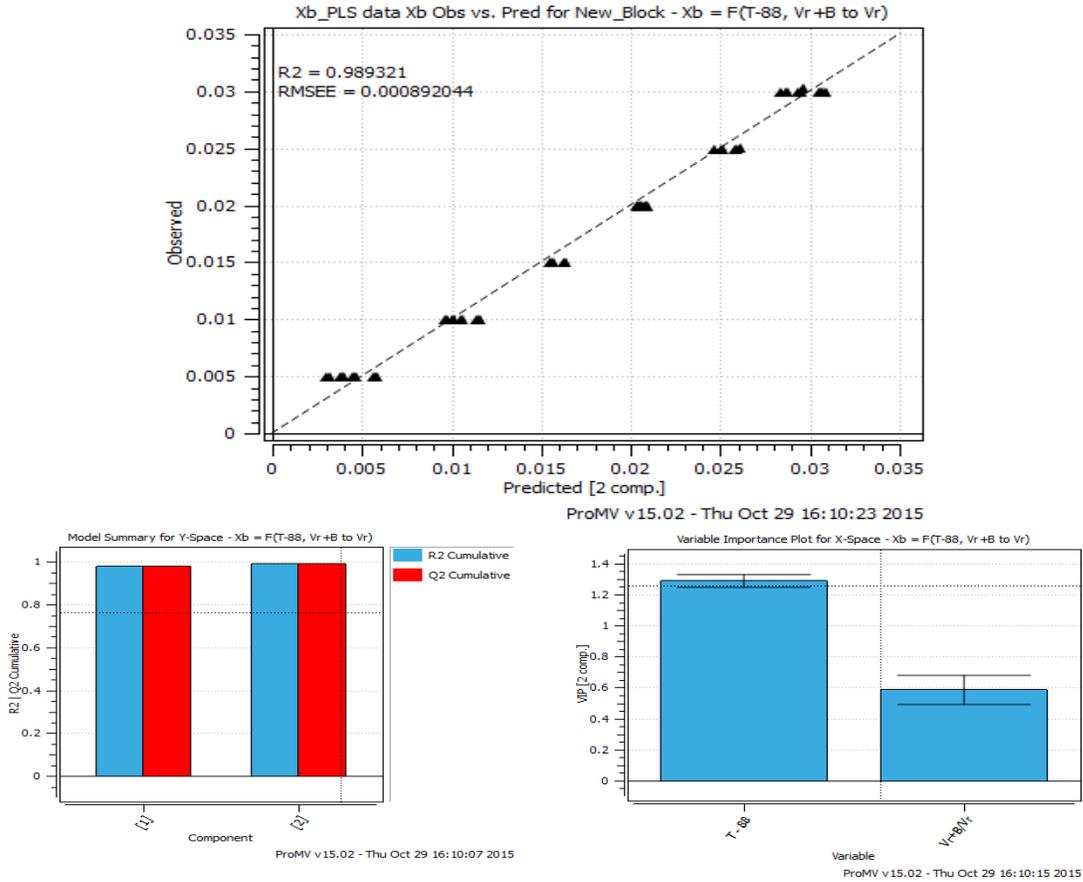


Figure 91. ProMV PLS Xb model as function of two variables plot and statistical analysis for variable feed composition Xd & Xb (C4 splitter high purity range)

Below is Xb equation as Eq. (92) with four variables  $T_{88}$ ,  $T_{77}$ ,  $\frac{R}{R+D}$  &  $\frac{V_r+B}{V_r}$  was generated using multivariate software (ProMV) as shown in Fig. 92.  $T_{72}$  has the highest sensitivity with  $\frac{\Delta T C^o}{\Delta x_d} = 0.663$  in which adding  $T_{72}$  improves model accuracy.

$$x_b = a_1 T_{88} + a_2 T_{72} + a_3 \frac{R}{R+D} + a_4 \frac{V_r+B}{V_r} + b \quad (92)$$

$$a_1 = -0.024744 \quad a_2 = 0.004815 \quad a_3 = -0.174678 \quad a_4 = 0.111301 \quad b = 1.83396$$

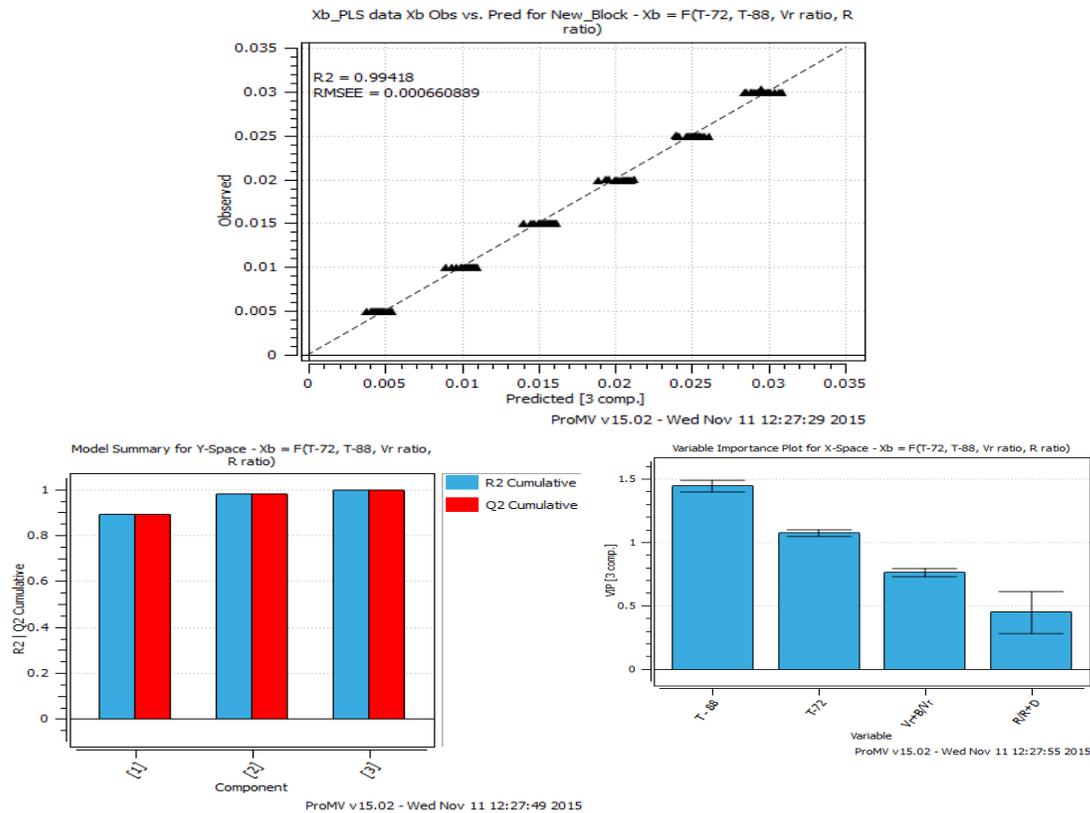


Figure 92. ProMV PLS Xb model as function of four variables plot and statistical analysis for variable feed composition Xd & Xb (C4 splitter high purity range)

Table 56. ProMV PLS Xb models accuracy analysis for 144 cases of variable feed composition Xd & Xb and 30 cases of prediction data set (C4 splitter high purity range)

Equation	$x_b = F\left(T_{88}, \frac{V_r + B}{V_r}\right)$	$x_b = F\left(T_{88}, T_{72}, \frac{R}{R + D}, \frac{V_r + B}{V_r}\right)$
R <sup>2</sup>	0.9893	0.994180
Q <sup>2</sup>	0.9893	0.993718
RMSEE	0.000892044	0.000660889
RMSEP	0.000693	0.000347
Average value	0.0175	0.0175
Components #	2	3

Eq. (92) showed higher accuracy with higher  $R^2$  &  $Q^2$  and lower RMSEE & RMSEP as compared to Eq. (85a) as shown in Table. 56. The Xb inferential model for high purity range can be expressed with equation  $x_b = F(T_{88}, T_{72}, \frac{R}{R+D}, \frac{V_r+B}{V_r})$  as eq. (92).

Below is the model with feed composition as an input. Xb equation  $x_b = a T_{88} + b$  experience shifting due to feed composition variation. The below Table. 57 and Figs. 93 and 94 show the relationship between slope and intercept as function of feed variation expressed by LK.

Table 57. Xb = F( $T_{88}$ ) slope and intercepts for variable feed composition & Xd (C4 splitter high purity range)

LK%	Slope a	Intercept b
15%	-0.02713	2.430931
24%	-0.0232	2.078283
31%	-0.02101	1.882431
40%	-0.01882	1.68611

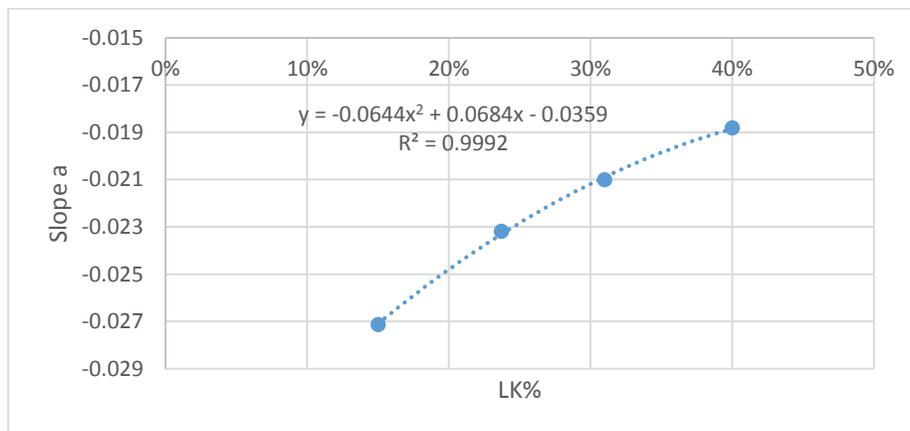


Figure 93. Relationship between Xb equation slope and feed LK (C4 splitter high purity range)

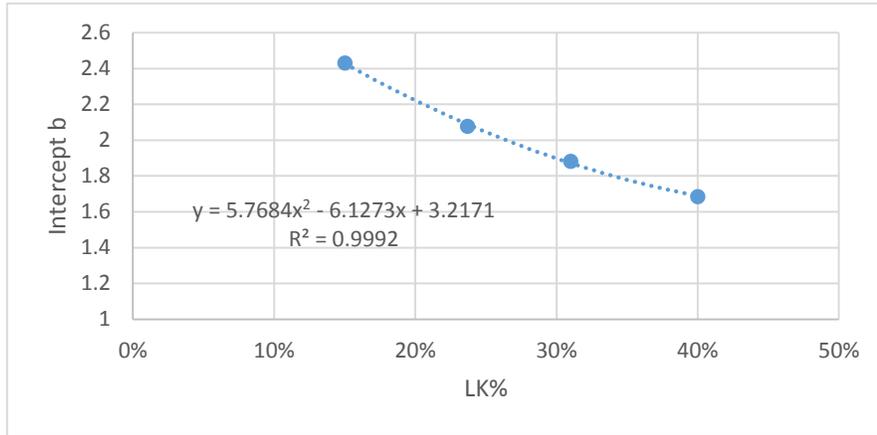


Figure 94. Relationship between Xb equation intercept and feed LK (C4 splitter high purity range)

$$\text{Slope } a = -0.0644 (LK)^2 + 0.0684 (LK) - 0.0359 \quad (93)$$

$$R^2 = 0.9992$$

$$\text{Intercept } b = 5.7684 (LK)^2 - 6.1273 (LK) + 3.2171 \quad (94)$$

$$R^2 = 0.9992$$

Eqs. (93) and (94) express the slope  $a$  and intercept  $b$  as function of feed LK composition. The model (Eq. (76), (93) and (94) with  $T_{88}$  as sensitive tray was tested for 144 cases of variable feed composition,  $X_d$  &  $X_b$ . In addition to 30 cases prediction testing. It showed slightly higher  $R^2$  &  $Q^2$  however it showed higher RMSEE & RMSEP as compared to Eq. (92) as shown in Table. 58. It is recommended to use the PLS model as Eq. (92) considering its accuracy, linearity and input reliability.

Table 58. Xb model with slope and intercept as function of feed LK accuracy analysis

Equation	$x_b = a T_{88} + b$ <i>Slope</i> $a = -0.0644 (LK)^2 + 0.0684 (LK) - 0.0359$ <i>Intercept</i> $b = 5.7684 (LK)^2 - 6.1273 (LK) + 3.2171$
R <sup>2</sup>	0.9963
Q <sup>2</sup>	0.9976
RMSEE	0.00089193
RMSEP	0.000928
Average value	0.0175
Polynomial degree	3

#### 4.4 Inferential Model for Depropanizer Column

The DeC3 column operates for wide boiling points components. Hence the tray to tray temperature variation are considered significant which indicate relatively high  $\frac{\Delta T C^o}{\Delta x_d}$  and  $\frac{\Delta T C^o}{\Delta x_b}$ . The models were generated using 125 cases of variable feed composition, Xd & Xb as per the experimental plant explained in the previous chapter. While the prediction data set included 25 cases of randomly selected feed compositions, Xd & Xb which fall within the specified experimental range.

##### 4.4.1 Top impurity Xd Model

The below analysis was conducted at constant feed composition and variable Xb. The selected feed composition is HK% = 17.56% as an example. This is to select a sensitive

tray that shows significant  $\frac{\Delta T C^o}{\Delta x_d}$ , shows linear behavior and not impacted by Xb variation. Fig. 95 shows the column top section trays temperature variation for Xd variation  $\frac{\Delta T C^o}{\Delta x_d}$ . It indicates the analysis for 25 cases of variable Xd and Xb. The trays numbering follows top to bottom pattern where tray 1 is the first tray below the condenser while tray 21 is the feed tray.

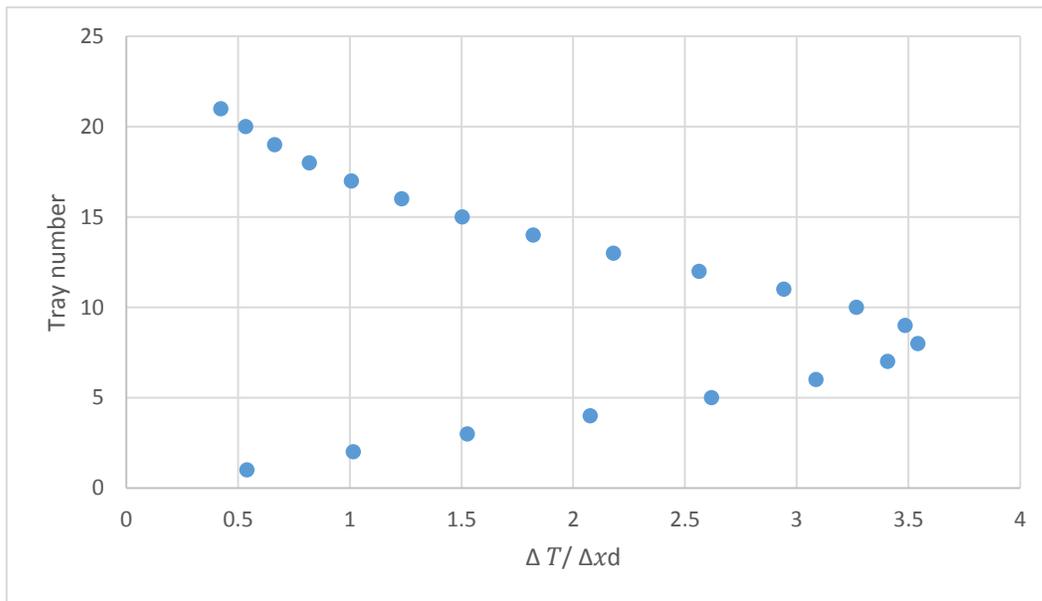


Figure 95. Relationship between tray number and  $\Delta T / \Delta x_d$  in the top section for constant feed composition and variable Xb

As shown the maximum  $\frac{\Delta T C^o}{\Delta x_d}$  value is 3.543 C° per Xd variation of 0.01 at tray 8 which is significant magnitude. As the tray number decrease the more linear behavior is observed. Generally the trays above tray 8 show lower  $\frac{\Delta T C^o}{\Delta x_d}$  and more linear behavior with less impact caused by Xb variation. Tray 8 however showed nonlinear behavior and it is impacted by Xb variation as shown in Fig. 96. The horizontal spread at each Xd

value is caused by  $X_b$  variation. Trays 4, 5, 6, 7 & 8 were selected for comparison as shown in Fig. 96.

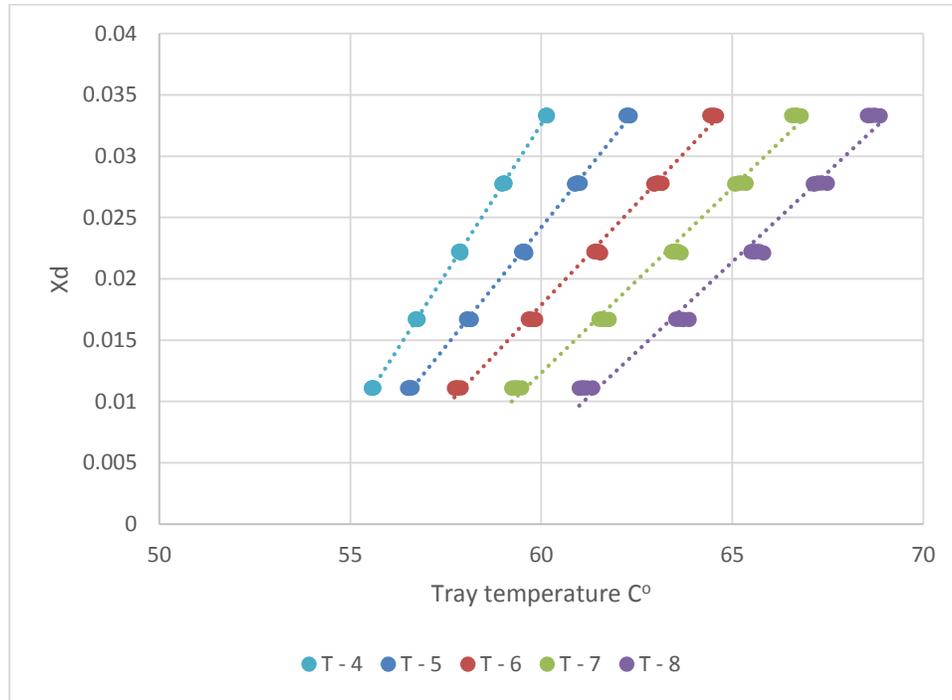


Figure 96. Relationship between  $X_d$  and trays 4, 5, 6, 7 & 8 for constant feed composition and variable  $X_b$

$X_d$  equation as Eq. (75) was generated from the above plot and tested individually for  $X_d$  as function of  $T_4, T_5, T_6, T_7$  &  $T_8$  as shown in Table. 59. Tray 4 shows higher  $R^2$  and lower RMSEE as compared to average  $X_d$  value. However it shows lower  $\frac{\Delta T C^\circ}{\Delta x_d}$  as compared to the other trays. Hence selecting  $T_4$  indicates high accuracy  $X_d$  model for constant feed composition and variable  $X_b$ . In addition it still shows significant  $\Delta T = 2.1 C^\circ$  per  $X_d$  variation of 0.01.

Table 59.  $x_d$  as function of  $T_4$ ,  $T_5$ ,  $T_6$ ,  $T_7$  &  $T_8$  accuracy comparison

Equation	$R^2$	RMSEE	Average value	$\frac{\Delta T C^o}{\Delta x_d}$
$x_d = 0.0049 T_4 - 0.2598$	0.9997	0.000123596	0.022	2.1
$x_d = 0.0039 T_5 - 0.2089$	0.9987	0.000273245	0.022	2.6
$x_d = 0.0033 T_6 - 0.1812$	0.9959	0.000494106	0.022	3.1
$x_d = 0.003 T_7 - 0.1689$	0.9908	0.00073724	0.022	3.4
$x_d = 0.0029 T_8 - 0.1686$	0.9838	0.000979785	0.022	3.5

Next step is testing  $x_d$  equation with  $T_4$  for variable  $x_d$ ,  $x_b$  and feed composition. Fig. 97 shows the relationship between  $x_d$  and  $T_4$  for 125 cases of variable  $x_d$ ,  $x_b$  and feed composition.

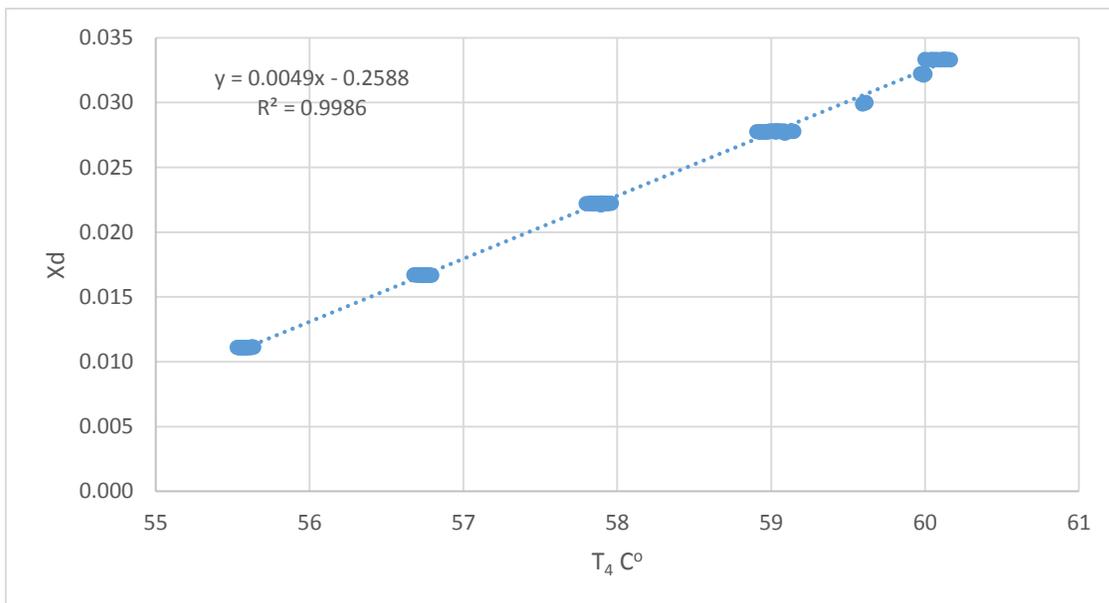


Figure 97. Relationship between  $x_d$  and  $T_4$  for variable feed composition &  $x_b$

Eq. (95) experienced negligible horizontal data spread due to feed composition variation. Hence a single linear regression can describe the relationship between  $X_d$  &  $T_4$  for variable feed composition as Eq. (95).

$$x_d = 0.0049 T_4 - 0.2588 \quad \text{Variable feed composition, } X_d \text{ \& } X_b \quad (95)$$

Table 60.  $X_d$  model with single temperature variable accuracy analysis for 125 cases of variable feed composition,  $X_d$  &  $X_b$  and 25 cases of prediction data set

Equation	$x_d = 0.0049 T_4 - 0.2588$
$R^2$	0.9986
$Q^2$	0.9985
RMSEE	0.000281336
RMSEP	0.00027611
Average value	0.022

Eq. (95) shows high  $R^2$  and low RMSEE as compared to the average value as shown in Table. 60. In addition this presents the advantage of a single linear equation in one variable that can predict  $X_d$  for variable feed composition and  $X_b$ . An alternative linear model for  $X_d$  [equation as Eq. (96)] with two variables  $T_4$  and  $\frac{R}{R+D+V_d}$  was generated using a multivariate software (ProMV) as shown in Fig. 98 for 125 cases of variable feed composition,  $X_d$  &  $X_b$ . The ratio  $\frac{R}{R+D+V_d}$  is an estimation for the internal reflux ratio which impacts both  $X_d$  and trays temperature values.

$$x_d = a_1 T_4 + a_2 \frac{R}{R+D+V_d} + b \quad (96)$$

$$a_1 = 0.004897 \quad a_2 = 0.010732 \quad b = -0.269103$$

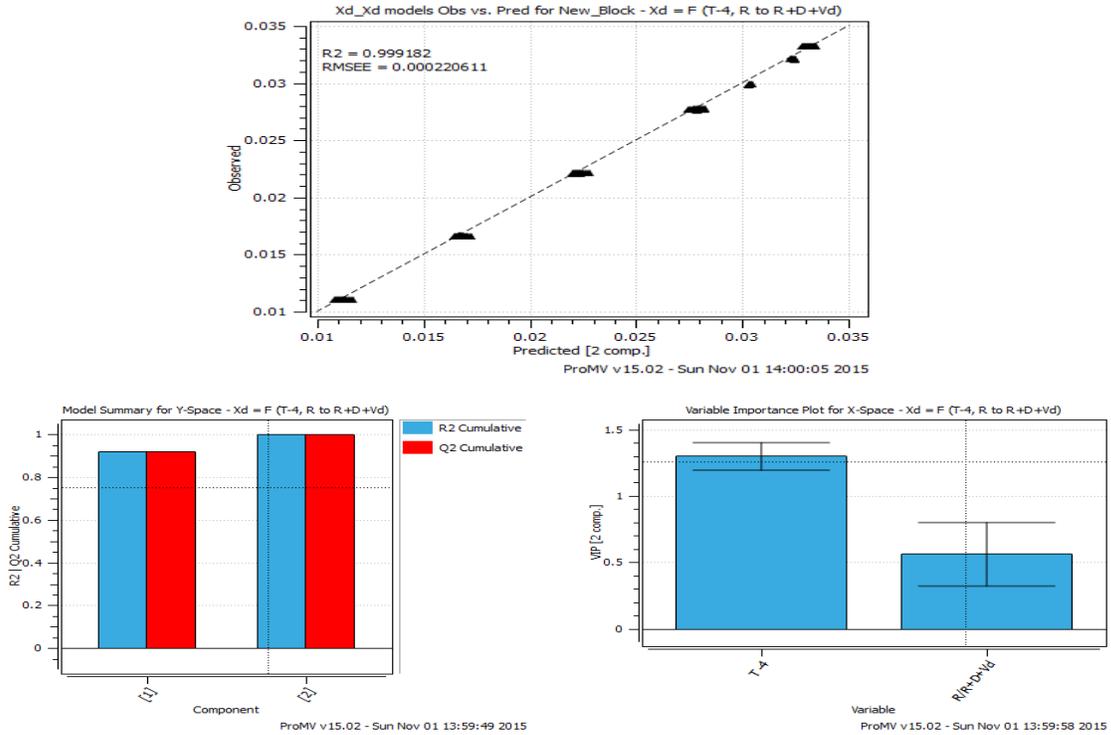


Figure 98. ProMV PLS Xd model as function of two variables plot and statistical analysis for variable feed composition, Xd & Xb

Eq. (96) is function of two variables which showed importance as per the variable importance plot VIP in Fig. 98. The model generated with two components with high  $R^2$  &  $Q^2$  and low RMSEE & RMSEP as compared to average Xd value as shown in Table. 61. This model presents the advantages of being linear and with high accuracy values. Both inputs can be obtained from the column online operating values. This model showed slightly higher accuracy as compared to  $x_d = 0.0049 T_4 - 0.2588$  where  $R^2$  is higher by 0.00058 and almost the same RMSEE. Using single temperature input model Eq. (95) presents the advantage of small model. In addition for basic control the temperature set point can be easily calculated to meet Xd specification.

Table 61. ProMV PLS Xd model as function of two variables accuracy analysis for 125 cases of variable feed composition, Xd & Xb and 25 cases of prediction data set

Equation	$x_d = a_1 T_4 + a_2 \frac{R}{R + D + Vd} + b$
R <sup>2</sup>	0.999182
Q <sup>2</sup>	0.999176
RMSEE	0.000220611
RMSEP	0.000223
Average value	0.022
Components #	2

#### 4.4.2 Bottom impurity Xb Model

The analysis below was conducted at constant feed composition and variable Xb. The selected feed composition is LK% = 36.39% as an example. This is to select a sensitive tray that shows significant  $\frac{\Delta T C^o}{\Delta x_b}$ , shows linear behavior and not impacted by Xd variation. Fig. 99 shows the column bottom section trays temperature variation for Xb variation  $\frac{\Delta T C^o}{\Delta x_b}$ . It indicated the analysis for 25 cases of variable Xd and Xb. The trays numbering follows top to bottom pattern where tray 21 is the feed tray and tray 40 is the last tray above the reboiler.

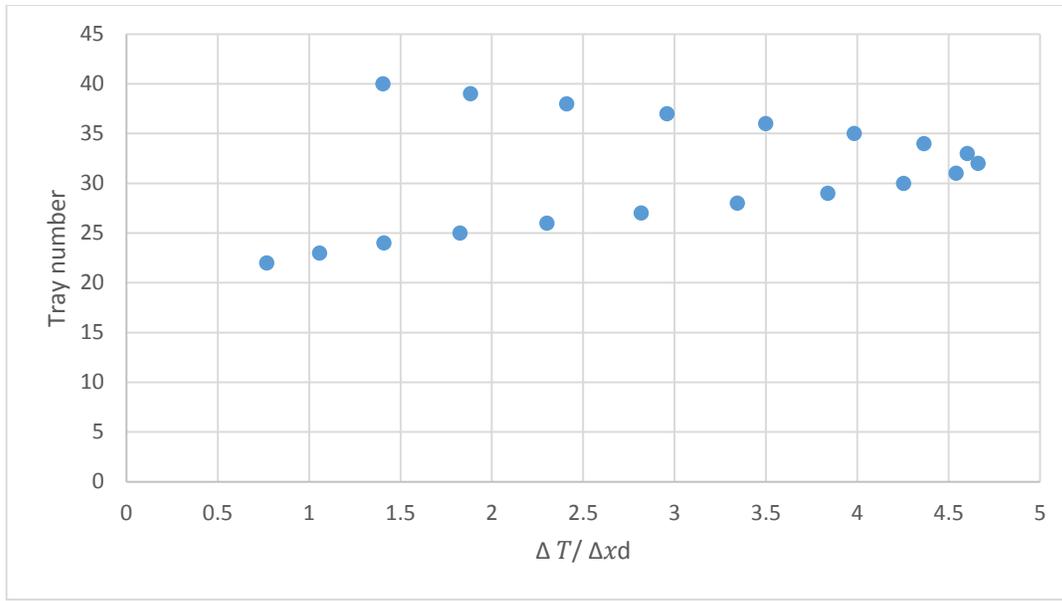


Figure 99. Relationship between tray number and  $\Delta T / \Delta x_b$  in the bottom section for constant feed composition and variable  $X_d$

As shown the maximum  $\frac{\Delta T C^o}{\Delta x_b}$  value is 4.661 C° per  $X_b$  variation of 0.01 at tray 32 which is significant magnitude. As the tray number increase toward the column end the more linear behavior is observed. Generally the trays below tray 32 show lower  $\frac{\Delta T C^o}{\Delta x_d}$  and more linear behavior with less impact caused by  $X_d$  variation. Tray 32 however showed nonlinear behavior and it is impacted by  $X_d$  variation as shown in Fig. 100. The horizontal spread at each  $X_b$  value is caused by  $X_d$  variation. Trays 32, 33, 34, 35, 36, 37 & 38 were selected for comparison as shown in Fig. 100.

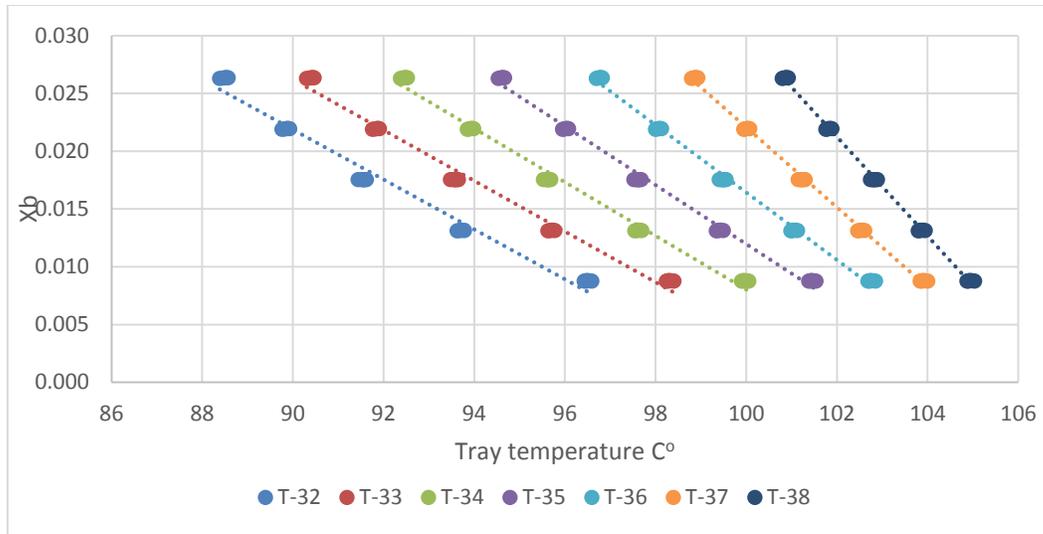


Figure 100. Relationship between  $X_b$  and  $T_{32}$ ,  $T_{33}$ ,  $T_{34}$ ,  $T_{35}$ ,  $T_{36}$ ,  $T_{37}$  &  $T_{38}$  for constant feed composition and variable  $X_d$

Eq. (76) was generated from the above plot and tested individually for  $X_b$  as function of  $T_{32}$ ,  $T_{33}$ ,  $T_{34}$ ,  $T_{35}$ ,  $T_{36}$ ,  $T_{37}$  &  $T_{38}$  as shown in Table. 62.

Table 62.  $X_b$  as function of  $T_{32}$ ,  $T_{33}$ ,  $T_{34}$ ,  $T_{35}$ ,  $T_{36}$ ,  $T_{37}$  &  $T_{38}$  accuracy comparison

Equation	R <sup>2</sup>	RMSEE	Average $X_b$	$\frac{\Delta T C^\circ}{\Delta x_d}$
$x_b = -0.0022 T_{32} + 0.2162$	0.9781	0.010916	0.0175	4.661072075
$x_b = -0.0022 T_{33} + 0.2237$	0.9848	0.010923	0.0175	4.600464228
$x_b = -0.0023 T_{34} + 0.2401$	0.99	0.010933	0.0175	4.364741488
$x_b = -0.0026 T_{35} + 0.2671$	0.9938	0.010945	0.0175	3.982689411
$x_b = -0.0029 T_{36} + 0.3081$	0.9962	0.010961	0.0175	3.497725032
$x_b = -0.0035 T_{37} + 0.3687$	0.9975	0.010984	0.0175	2.95820322
$x_b = -0.0043 T_{38} + 0.4584$	0.9978	0.011015	0.0175	2.408485028

Tray 38 shows slightly higher  $R^2$  while RMSEE values are almost the same for all the trays. However it shows lower  $\frac{\Delta T C^o}{\Delta x_b}$  as compared to the other trays. Hence selecting  $T_{38}$  indicates high accuracy Xb model for constant feed composition and variable Xd. In addition it still shows significant  $\Delta T = 2.4 C^o$  per Xb variation of 0.01.

Next step is testing Xb equation with  $T_{38}$  for variable Xd, Xb and feed compositions. Fig. 101 shows the relationship between Xb and  $T_{38}$  for 125 cases of variable Xd, Xb and feed composition.

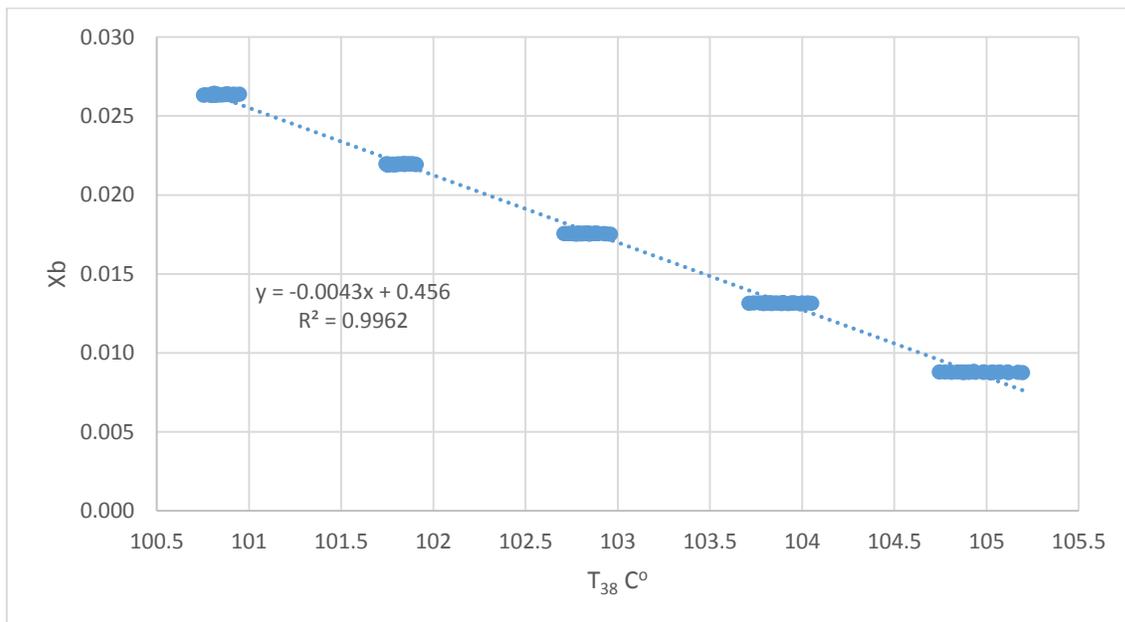


Figure 101. Relationship between Xb and  $T_{38}$  for variable feed composition & Xd

With variable feed composition Xb dependence on  $T_{38}$  has horizontal data spread due. A single linear regression was tested to describe the relationship between Xb &  $T_{38}$  for variable feed composition as Eq. (97).

$$x_b = -0.0043 T_{38} + 0.456 \quad \text{Variable feed composition, Xd \& Xb} \quad (97)$$

Table 63. Xb model with single temperature variable accuracy analysis for 125 cases of variable feed composition, Xd & Xb and 25 cases of prediction data set

Equation	$x_b = -0.0043 T_{38} + 0.456$
R <sup>2</sup>	0.9962
Q <sup>2</sup>	0.9961
RMSEE	0.000374
RMSEP	0.000331
Average value	0.0175

Eq. (97) shows high R<sup>2</sup> and low RMSEE as compared to the average value as shown in Table. 63. In addition this presents the advantage of a single linear equation in one variable that can predict Xb for variable feed composition and Xd.

Alternatively a linear Xb equation as Eq. (98) with two variables  $T_{38}$  and  $\frac{V_r+B}{V_r}$  was generated using multivariate software (ProMV) as shown in Fig. 102 for 125 cases of variable feed composition, Xd & Xb. The ratio  $\frac{V_r+B}{V_r}$  is an estimation for the internal reflux ratio which impacts both Xb & trays temperature values.

$$x_b = a_1 T_{38} + a_2 \frac{V_r+B}{V_r} + b \quad (98)$$

$$a_1 = -0.004321 \quad a_2 = -0.006723 \quad b = 0.471165$$

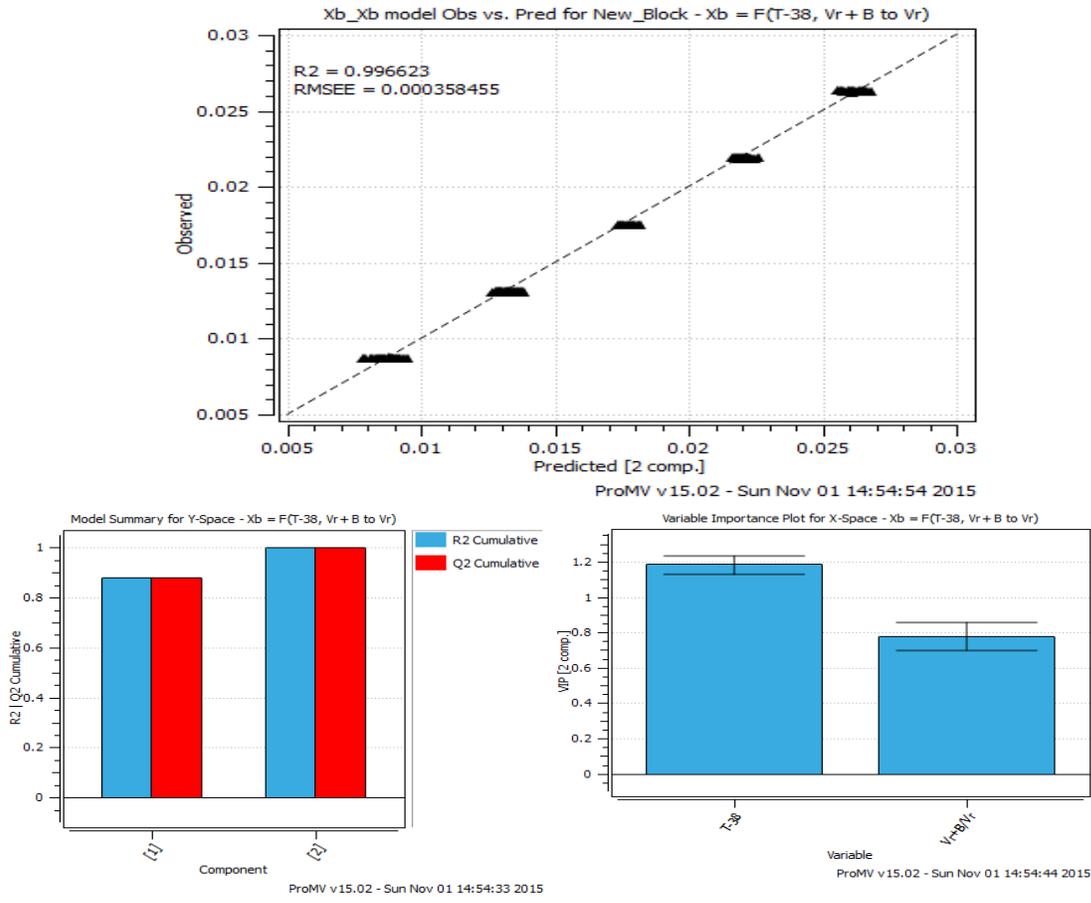


Figure 102. ProMV PLS Xb model as function of two variables plot and statistical analysis for variable feed composition, Xd & Xb

Eq. (98) is function of two variables which showed importance as per the variable importance pot VIP in Fig. 102. The model generated with two components with high  $R^2$  &  $Q^2$  and low RMSEE as compared to average Xb value as shown in Table. 64. This model presents the advantages of being linear and with high accuracy values. Both inputs can be obtained from the column online operating values.

Table 64. ProMV PLS Xb model as function of two variables accuracy analysis for 125 cases of variable feed composition, Xd & Xb and 25 cases of prediction data set

Equation	$x_b = a_1 T_{38} + a_2 \frac{V_r + B}{V_r} + b$
R <sup>2</sup>	0.9966
Q <sup>2</sup>	0.9965
RMSEE	0.000358455
RMSEP	0.000301
Average value	0.0175
Components #	2

This model showed slightly higher accuracy as compared to  $x_b = -0.0043 T_{38} + 0.456$  where R<sup>2</sup> is higher by 0.0004 and almost the same RMSEE. Using single temperature input model Eq. (97) presents the advantage of small model. In addition for basic control the temperature set point can be easily calculated to meet Xb specification.

## 5 Conclusions

This work addressed the development of simplified distillation columns models. The tray to tray rigorous models are very accurate models to describe distillation column separation however they are large in size and difficult to incorporate for planning, scheduling and optimization purposes. Currently there are different types of simplified models used for planning, scheduling and optimization which cause discrepancies between their outputs. Such situation calls for introducing simpler, smaller in size and accurate models that describes separation.

Hybrid model is small in size model that is comprised of first principles material and energy balances and empirical models which describes column separation. Empirical models were built based on rigorous simulation data input. The approach of developing empirical models using the separation factor SF for binary separation was tested and resulted in large error as compared to actual SF values generated by Aspen plus. The error was modeled and resulted in nonlinear model that calculates SF without calculating streams impurities  $X_d$  &  $X_b$  explicitly. This work illustrated the methodology for developing hybrid models for two distillation column C4 splitter and DeC3. Models were developed based on Aspen plus simulation cases for three main variables as shown below.

- Feed composition
- Top impurity  $X_d$
- Bottom impurity  $X_b$

These are the main decision variables for production planning, scheduling and optimizations problems. Hybrid models described distillation main parameters such as reflux ratio, streams flows, condenser duty and reboiler duty for variable feed compositions and impurity specifications. Models are nonlinear however they are small in size and can be solved using commercial NLP solvers or Excel GRG solver.

Extensive testing was carried to the models by comparing results to Aspen plus results for variable feed composition,  $X_d$  &  $X_b$ . The testing results showed high accuracy through high  $R^2$ ,  $Q^2$  and low RMSEE and RMSEP. Models are simple, small in size and accurate as compared to Aspen plus results. This makes hybrid models suitable for applications of production planning, scheduling and optimization. The models can be

used as stand-alone application or part of plant wide model. The models were tested for two optimization case studies and showed good accuracy with relative error less than 1% for most of the parameters as compared to Aspen plus optimization tool. Models were solved using GAMS Antigone nonlinear global solver.

In addition this work introduced inferential models that estimate streams impurities using column real time data for variable feed composition,  $X_d$  &  $X_b$ . This eliminates the need for Gas Chromatography analyzer GC. Commonly GC analyzers are used to analyze streams purity however this requires installation and maintenance cost. In addition GC analyzers face reliability issues and hence their input is not a reliable source. Furthermore this work investigated the industrial practice of inferential model development as function of single sensitive tray temperature input as compared to multivariable inputs. The single sensitive tray temperature models showed high accuracy for wide boiling point DeC3 distillation column. However for narrow boiling point C4 splitter it showed low accuracy. This attributed to the fact that wide boiling point mixture columns show significant tray temperature sensitivity for composition variation. The PLS regression was used to generate models as function of multivariable inputs such as multiple trays temperature,  $\frac{R}{R+D}$  and  $\frac{V_r+B}{V_r}$  which showed higher accuracy as compared to single tray temperature input for both columns. The accuracy improvement was with higher magnitude for C4 splitter as compared to DeC3. This is due to the fact that both single temperature and PLS models for DeC3 showed high accuracy. As shown models with feed composition input showed high accuracy. However this requires having GC analyzer to generate online feed composition input. The developed PLS models estimate streams impurity without feed composition input

which showed higher accuracy. This eliminated the need for feed composition GC analyzer. The models are linear and can be incorporated to the plants DCS for monitoring and control purposes. The models results were compared to Aspen plus and showed high accuracy through high  $R^2$ ,  $Q^2$  and low RMSEE and RMSEP.

The future work can focus on hybrid models linearization to be included in plant wide LP models. In addition the data behavior can be described as single line of operation for  $X_d$  and  $X_b$  that experience shifting. This can be dealt with by linear interpolation between various simulation cases. Hence models can be solved as MINLP or MILP when linearized. In addition future work can include incorporating both hybrid and inferential models for real time optimization RTO.

## 6 References

- Aspen Technology Inc. Aspen Plus 11.1 Unit Operation Models. Cambridge, MA: Aspen Technology Inc. 2001
- Boston, J. F., Sullivan, S. L. (1974). A new class of solution methods for multicomponent, multistage separation processes. The Canadian Journal of Chemical Engineering, 52(1), 52-63.
- Douglas, J. M., Jafarey, A., Seemann, R. (1979b). Short-cut techniques for distillation column design and control. 2. Column operability and control. Industrial & Engineering Chemistry Process Design and Development, 18(2), 203-210.
- Fenske, M. (1932). Fractionation of straight-run Pennsylvania gasoline. Industrial & Engineering Chemistry, 24(5), 482-485.

- Fu, G., Sanchez, Y., Mahalec, V. (2015). Hybrid model for optimization of crude oil distillation units. *AIChE Journal*.
- Gilliland, E. R. (1940). Multicomponent rectification estimation of the number of theoretical plates as a function of the reflux ratio. *Industrial & Engineering Chemistry*, 32(9), 1220-1223.
- Jafarey, A., Douglas, J. M., McAvoy, T. J. (1979a). Short-cut techniques for distillation column design and control. 1. Column design. *Industrial & Engineering Chemistry Process Design and Development*, 18(2), 197-202.
- Kresta, J. V., Marlin, T. E., MacGregor, J. F. (1994). Development of inferential process models using PLS. *Computers & Chemical Engineering*, 18(7), 597-611.
- Mahalec, V., Sanchez, Y. (2012). Inferential monitoring and optimization of crude separation units via hybrid models. *Computers & Chemical Engineering*, 45, 15-26.
- McCabe, W. L., Thiele, E. W. (1925). Graphical design of fractionating columns. *Industrial & Engineering Chemistry*, 17(6), 605-611.
- Mejdell, T., Skogestad, S. (1991a). Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression. *Industrial & Engineering Chemistry Research*, 30(12), 2543-2555.
- Mejdell, T., Skogestad, S. (1991b). Composition estimator in a pilot-plant distillation column using multiple temperatures. *Industrial & engineering chemistry research*, 30(12), 2555-2564.
- Mejdell, T., Skogestad, S. (1990, November). Composition control of distillation columns using multiple temperature measurements. In *AIChE annual meeting in Chicago*.

- Molokanov, Y. K., Korablina, T. P., Mazurina, N. I., Nikiforov, G. A. (1971). An approximate method of calculating the basic parameters of a multicomponent fractionation. *Chemistry and Technology of Fuels and Oils*, 7(2), 129-133.
- Safavi, A. A., Nooraii, A., Romagnoli, J. A. (1999). A hybrid model formulation for a distillation column and the on-line optimisation study. *Journal of Process Control*, 9(2), 125-134.
- Shinskey, F.G. (1977), *Distillation control*, 1st Edition, McGraw-Hill, New York.
- Smoker, E. H. (1938). Analytical determination of plates in fractionating columns. *AIChE J.*, 12, 165-172.
- Soliman, M. A. (2007). A Shortcut method for binary distillation column design.
- Skogestad, S, and Morari, M., 1987, "Shortcut Models for Distillation Columns- ISteady State Behavior, Chapter 9 in Ph.D. Thesis, "Studies on Robust Control of Distillation Columns" by S. Skogestad, California Institute of Technology, Pasadena
- Underwood, A. J. V. (1948). Fractional distillation of multicomponent mixtures. *Chem. Eng. Prog*, 44(8), 603-614.