Building the foundation for Prudhoe Bay oil production optimisation using neural networks

Shahab D. Mohaghegh*

Department of Petroleum and Natural Gas Engineering,
West Virginia University,
Morgantown, WV 26506, USA
E-mail: Shahab@wvu.edu
*Corresponding author

Lynda A. Hutchins

BP Exploration (Alaska),
Anchorage, AK, USA
E-mail: hutchinl@Bp.com

Carl Sisk

BP Exploration (Houston),
Houston, TX, USA
E-mail: siskd@Bp.com

Abstract: Field data from the Prudhoe Bay oil field in Alaska was used to develop a neural network model of the cross-country gas transit pipeline network between the production separation facilities and central gas compression plant. The trained model was extensively tested and verified using 30% of the data that was not used during the training process. The results show good accuracy in reproducing the actual rates and pressures at the separation facilities and at the gas compression plant. The correlation coefficient for rate and pressure were 0.997 and 0.998, respectively. This development builds the foundation for building a tool to maximise total field oil production by optimising the gas discharge rates and pressures at the separation facilities.

Keywords: Prudhoe Bay; neural network; surface facility; optimisation.

Reference to this paper should be made as follows: Mohaghegh, S.D., Hutchins, L.A. and Sisk, C. (2008) ‘Building the foundation for Prudhoe Bay oil production optimisation using neural networks’, *Int. J. Oil, Gas and Coal Technology*, Vol. 1, Nos. 1/2, pp.65–80.

Biographical notes: Shahab D. Mohaghegh received a PhD in Petroleum and Natural Gas Engineering from Penn State University in 1991. Currently, he is a Professor in the Department of Petroleum and Natural Gas Engineering, West Virginia University. His current research interests include smart fields and application of intelligent systems in production optimisation and reservoir evaluation and engineering.

Lynda A. Hutchins received an MA in Natural Science (Physics) from Oxford University, an MSc in Radiation Physics from London University and has three years Post-Graduate Research in the area of ultrasonic signal processing. She has 25 years of oil industry experience, the last 17 years in Alaska. Her interests are focussed on the development and application of practical production optimisation tools for large on-shore oil fields, specifically Prudhoe Bay on the North Slope of Alaska.

Carl Sisk received a BS (EE) from University of Missouri at Rolla in 1978. He has held a variety of production operations, reservoir engineering and leadership positions with Amoco Production Company and Field of the Future Project Manager positions from 1998 to 2006 with BP America. His research interests include automation and production optimisation.

1 Introduction: background

Prudhoe Bay has approximately 800 producing wells flowing to eight remote, three-phase separation facilities (flow stations and gathering centres). High-pressure gas is discharged from these facilities into a cross-country pipeline system flowing to a central compression plant. Figure 1 illustrates the gas transit network between the separation facilities and the inlet to the compression plant.

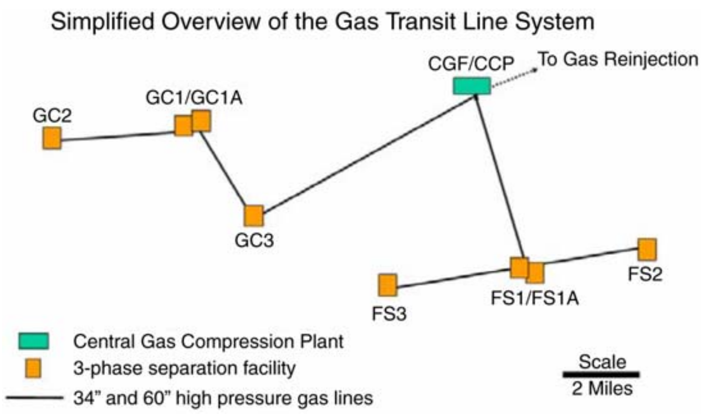

Figure 1 Prudhoe Bay gas transit line system

Fuel gas supply (at the flowstations and gathering centres) and artificial lift gas supply for the lift gas compressors at GC1 are taken off the gas transit line upstream of the compression plant. This reduces the feed gas rate and pressure at the inlet to the compression plant.

Gas feeding the central compression plant is processed to produce natural gas liquids and miscible injectant. Residue gas from the process is compressed further for reinjection into the reservoir to provide pressure support.

2 Business motivation

Ambient temperature has a dominant effect on compressor efficiency and hence total gas handling capacity and subsequent oil production. Figure 2 illustrates the range of daily average temperatures from 1990 to 2000, and the actual daily average for 2001 and 2002. Observed temperature variations during a 24-hr period can be as high as 40° Fahrenheit (°F).

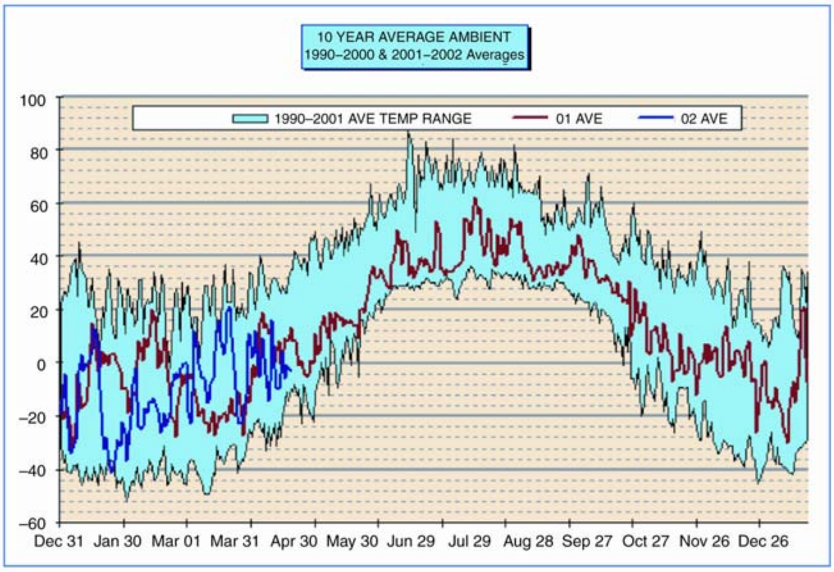

Figure 2 Historical daily average ambient temperature range

Figure 3 is a curve fit to total shipped gas rate to the compression plant versus ambient temperature for 2001. A significant reduction in gas handling capacity is observed at ambient temperatures above 0°F. Individual well Gas Oil Ratio (GOR) ranges between 800 and 35,000 scf/stb, with the lower GOR wells in the water-flood area of the field and higher GORs in the gravity drainage area. Gas compression capacity is the major bottleneck to production at Prudhoe Bay and typically field oil rate will be maximised by preferentially producing the lowest GOR wells.

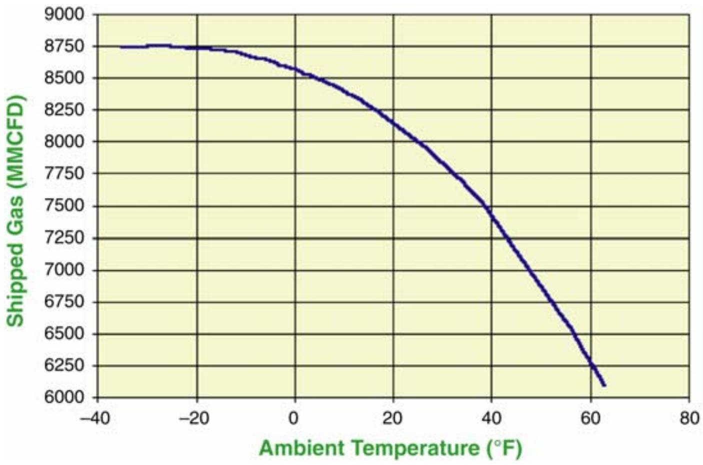

Figure 3 Shipped gas versus ambient temperature in 2001

As the ambient temperature increases from 0 to 40°F, the maximum (or 'marginal') GOR in the field decreases from approximately 35,000 to 28,000 scf/stb. A temperature swing from 0 to 40°F in one day equates to an approximate oil volume reduction of 40,000 bbls or 1000 bopd per °F rise in temperature.

The reduction in achievable oil rate, per degree Fahrenheit increase in temperature, increases with ambient temperature. This is due in part, to the increase in slope of the curve of shipped gas versus temperature, and also to the reduction in limiting or 'marginal' GOR as gas capacity decreases. At higher temperatures, larger gas cuts are required at lower GORs to stay within compression limits.

As ambient temperature rises, the inlet pressure at the compression plant increases. High inlet pressure can create backpressures at the separation facilities that will cause flaring. This is not environmentally acceptable and is avoided by cutting gas production. However, sometimes, ambient temperatures change so rapidly that significant gas cuts are necessary to avoid flaring. A similar problem occurs if a compressor experiences unexpected mechanical failure.

The ability to optimise the facilities in response to ambient temperature swings, compressor failures or planned maintenance is a major business driver for this project. Proactive management of gas production also reduces unnecessary emissions.

To maximise total oil rate under a variety of field conditions it is first necessary to understand the relationship between the inlet gas rate and pressure at the central compression plant, and the gas rates and discharge pressures into the gas transit line system at each of the separation facilities. This paper describes the development of a series of neural network models to describe these relationships.

Field oil rate will be impacted by the manner in which gas is distributed between facilities. A state-of-the-art genetic algorithm (Mohaghegh, 2000b) based optimisation tool (subject of a future paper) has been developed based on the neural network models described in this paper. The goal of the optimisation tool is to determine the gas discharge rates and pressures at each separation facility that will maximise field oil rate at a given ambient temperature, using curves of oil versus gas at each facility (based on GOR sorts). These curves are developed by iteration with a recently developed model of the surface piping network from wellhead to gathering centre/flowstation for Prudhoe Bay (Litvak et al., 2002).

Gas capacity constraints start to affect oil production at about 0°F, with increasing impact as the temperature increases. The estimated benefit of this tool for optimising oil rate during temperature swings and equipment maintenance is 1–2 MBOPD for 75% of the year.

3 Introduction

Attempts were made to develop a deterministic model of the gas transit system using commercial pipeline modelling software. However, it was extremely difficult to obtain sufficient historical data to validate the model. Development of a neural network model was undertaken to determine if this approach would provide a robust description of the observed gas rates and pressures with less stringent data requirements. For this initial test it was assumed that there was negligible hysteresis in the system. Initial results were very encouraging, suggesting that this is a valid approach, albeit limited to the data range used to train the mode.

4 Methodology

In this section the methodology of the approach used in this paper is presented. The methodology has been divided into two sections.

4.1 Data collection

The field data necessary to train the neural network models is stored in data historians from several independent control systems and had to be carefully checked for consistency. To ensure the data represented consistent field conditions (e.g. similar compressor configurations) and did not include periods where there were major equipment failures or maintenance, the data had to be carefully filtered. Consequently, the final available dataset was more limited than had been anticipated and the initial neural network model is limited to a fairly narrow range of field conditions.

The data included gas rate and gas discharge pressure from each of the eight separation facilities; fuel gas and lift gas supply rates, average hourly temperatures, and the inlet rate and pressure at the central gas compression plant. As was mentioned in the previous sections, the objective of this study is to optimise the target gas rates at each of the separation facilities in order to maximise oil production from the field. It was identified that in order to achieve this objective, the first step would be to build a representative model of the entire gas transit pipeline system. This neural network model should have two main characteristics:

- the model must accurately represent this complex dynamic system
- the model must provide fast results (close to real-time) once the required information is presented.

Any credible optimisation process would have to examine hundreds and sometimes thousands of realisations in order to achieve the global optimum values it is searching for. Therefore, both the accuracy and speed of providing the results are very important in the success of this project.

Neural network models have the capability of providing almost instantaneous results upon representation of the input data. Therefore, no matter how complex the problem, once, and if, an accurate model is built, calibrated and verified, it can serve as the ideal objective function for any optimisation routine.

Again, as was mentioned before, temperature plays a key role in this operation. The data used to build the neural network model was averaged on an hourly basis. Although the temperature was changing on an hourly basis (and that is how it was used during the model building process), it was averaged on a daily basis to illustrate the variation with time. Figure 4 shows the average daily temperature variation that was included in the dataset versus time. This figure shows the range of average daily temperature that was used during the modelling process. Moreover, the days that data could not be collected (for many different reasons) are also shown in this figure. Data from a total of 46 days was represented in the dataset. The data starts with the first day of the August and ends with the last day of September 2001.

It is important to note that the last five days of the data (25–29 September 2001) have much colder temperature than the rest of the days in the dataset. Therefore, if the network model is trained on the data from the days prior to 25 September 2001, and the data from

25 to 29 September 2001 is used in the verification dataset, then the network model would be extrapolating on the information it has not been exposed to.

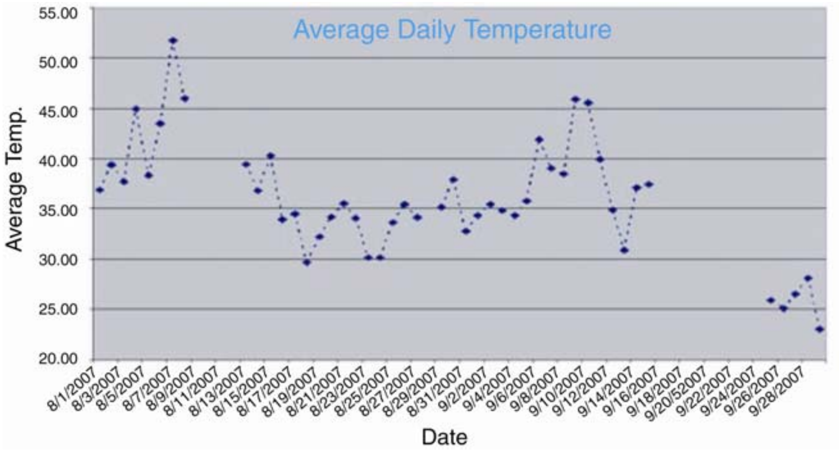

Figure 4 Average daily temperature presented in the data

The average daily temperature may be misleading in demonstrating the temperature swings within a single day. Since the model will be dealing with average temperature on an hourly basis rather than a daily basis, Figure 5 demonstrates the range of temperatures that might typically occur in one day. In this example the average hourly temperature is varied by 20° in an 8 hr period.

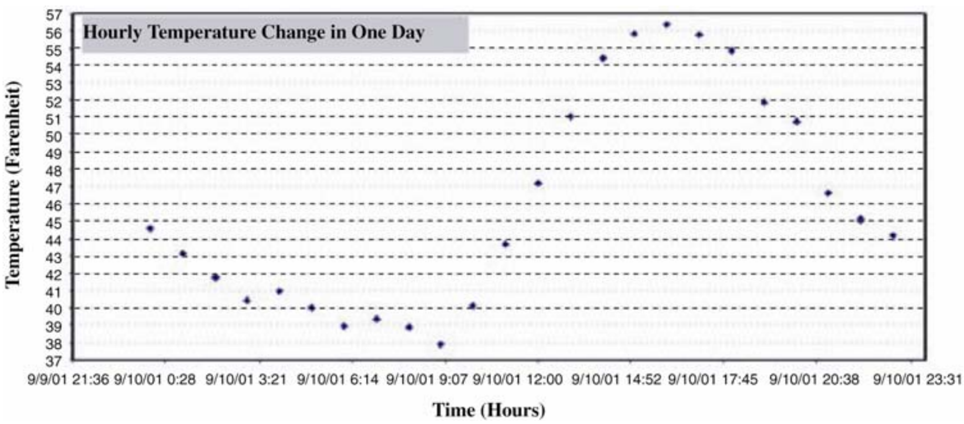

Figure 5 Average hourly temperature variation for 10 September 2001

4.2 *Training and verification of neural model*

In order to fulfil the requirements of this project, two sets of neural network models (Mohaghegh, 2000a) had to be constructed, calibrated and verified. The first set of models would represent the feed gas rate and inlet suction pressure at the Central Compression Plant and the second set of models would independently represent the set of gas discharge rates and pressures at each of the eight separation facilities.

4.2.1 Central compression plant inlet models

Prior to training, calibration and verification of the Central Compression Plant Inlet neural network model, the collected data was studied carefully. Table 1, displays the ranges of the parameters that were used during the development of the network models presented in this paper.

Table 1 Range of the parameters used in this study

Ambient temperature	*Min*	*Average*	*Max*	*SD*
	20.23	*35.85*	*57.33*	*6.44*
Gas discharge rates				
Total fuel gas	38.92	46.61	53.46	2.79
Gas-lift gas at GC1	401.23	809.30	923.09	152.29
FS1	895.75	1137.61	1304.96	76.10
FS2	428.09	704.69	769.89	63.35
FS3	382.06	786.61	1066.79	164.15
FS1A	907.18	1273.55	1530.64	136.96
GC1	456.85	964.30	1127.55	214.05
GC2	807.91	998.21	1080.00	55.59
GC3	490.54	1011.01	1131.89	112.28
GC1A	944.00	1353.91	1438.83	73.24
Feed gas rate to CCP	6473.64	7370.93	7832.34	234.48
Gas discharge pressures				
FS1	592.47	603.46	624.02	7.39
FS2	563.67	626.76	650.03	11.43
FS3	625.98	640.73	669.42	10.64
FS1A	560.75	602.37	625.66	10.47
GC1	574.95	611.38	634.87	11.54
GC2	578.68	610.47	634.62	12.16
GC3	581.66	600.94	622.99	9.66
GC1A	572.04	601.64	627.60	13.20
Inlet pressure to CCP	536.03	559.82	588.60	10.81

During the development of the network model, the dataset is divided into three segments. The majority of the data (about 70%) is used for training, hence called the training set. This is the data that is examined by the neural network continuously during training in order to build the internal knowledge and representation of the process being modelled. Another portion of the data (about 10 to 15%) is used to check the training process. This data is not used for training, and the information content of this data is not exposed to the network. This is called the calibration dataset. The calibration dataset is used, independent of the training data, to validate the generalisation capability of the network model. Furthermore, it helps the developer to make sure that the network model has not been over-trained and it does not memorise the information content in the training dataset.

The last segment of the data (the remaining 10 to 15%) is used as verification. This dataset is completely left out until the development and training of the network model is completed. The verification dataset, that has played no role whatsoever in the development and training, will provide the measure of validity of the network model. It shows how well the network has learnt the information in the dataset and how well it can mimic the process behaviour.

To fully examine the robustness of the intelligent model building processes on this dataset, three different neural network models were developed. Each of the models used a different set of the training, calibration and verification datasets. This was done in order to study the extrapolation capabilities of the intelligent model as well as its interpolative capabilities.

Figure 6 shows the spread of the data for each of the neural network models (based on the average daily temperature). As shown in this figure, in network model #1 all the temperatures are represented in the training dataset with the average being identified as 36°. In this network model, both calibration and verification datasets are subsets of the training set as far as the temperature is concerned. Therefore this is the most representative network model.

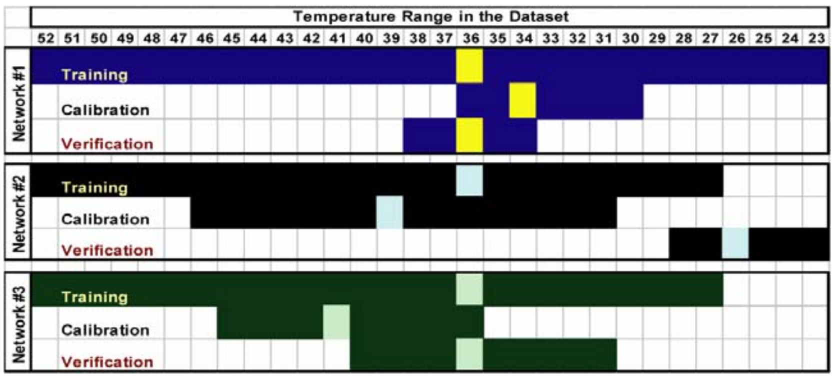

Figure 6 Dataset partitioning based on average daily temperature

In network model #2, the range of the training dataset is as low as 27° while the minimum temperature in the verification set is 23°. This is clearly an extrapolation. Can the network model predict what happens in the colder days that it has not been trained for? It should be noted that one of the greatest advantages of neural networks are that they do not break down once they face new environments. They degrade gracefully. Although the network model is not trained for these temperatures, it will not completely breakdown. It would still provide predictions, although they may not be as good as those of network model #1. Please note that in network model #2 the average temperature in the verification dataset is 26°. This value is outside of the training and calibration datasets. The average temperature in the calibration dataset is 39° while the average in the training set is 36°.

Network model #3, is probably the most practical of all the three networks models. It is hard to make that case by looking at Figure 6, and therefore we present the same information shown in Figure 6 in a table format. Table 2 is a different representation of the information presented in Figure 6. In this table not only the average daily temperature but also the corresponding dates are shown.

Table 2 Partitioning of the dataset for training the networks

	Day	*Date*	*Temp.*	*Network#1*	*Network#2*	*Network#3*
1	Wednesday	1 August 2001	36.89	Training	Training	Training
2	Thursday	2 August 2001	39.35	Training	Training	Training
3	Friday	3 August 2001	37.71	Training	Training	Training
4	Saturday	4 August 2001	44.94	Training	Training	Training
5	Sunday	5 August 2001	38.30	Training	Training	Training
6	Monday	6 August 2001	43.47	Training	Training	Training
7	Tuesday	7 August 2001	51.68	Training	Training	Training
8	Wednesday	8 August 2001	45.62	Training	Training	Training
	Thursday	9 August 2001				
	Friday	10 August 2001				
	Saturday	11 August 2001				
	Sunday	12 August 2001				
9	Monday	13 August 2001	39.44	Training	Training	Training
10	Tuesday	14 August 2001	36.79	Training	Training	Training
11	Wednesday	15 August 2001	40.27	Training	Training	Training
12	Thursday	16 August 2001	33.93	Training	Training	Training
13	Friday	17 August 2001	34.49	Verification	Training	Training
14	Saturday	18 August 2001	29.69	Training	Training	Training
15	Sunday	19 August 2001	32.21	Calibration	Training	Training
16	Monday	20 August 2001	34.18	Calibration	Training	Training
17	Tuesday	21 August 2001	35.51	Calibration	Training	Training
18	Wednesday	22 August 2001	34.10	Training	Training	Training
19	Thursday	23 August 2001	30.15	Training	Training	Training
20	Friday	24 August 2001	30.14	Calibration	Training	Training
21	Saturday	25 August 2001	33.67	Calibration	Training	Training
22	Sunday	26 August 2001	35.42	Calibration	Training	Training
23	Monday	27 August 2001	34.11	Training	Training	Training
	Tuesday	28 August 2001			Training	Training
24	Wednesday	29 August 2001	35.18	Training	Training	Training
25	Thursday	30 August 2001	37.90	Verification	Training	Training
26	Friday	31 August 2001	32.80	Training	Training	Training
27	Saturday	1 September 2001	34.36	Calibration	Training	Training
28	Sunday	2 September 2001	35.44	Calibration	Training	Training
29	Monday	3 September 2001	34.85	Calibration	Training	Training
30	Tuesday	4 September 2001	34.35	Verification	Training	Training
31	Wednesday	5 September 2001	35.76	Verification	Training	Calibration
32	Thursday	6 September 2001	41.92	Training	Training	Calibration
33	Friday	7 September 2001	39.03	Training	Calibration	Calibration

Table 2 Partitioning of the dataset for training the networks (continued)

	Day	*Date*	*Temp.*	*Network#1*	*Network#2*	*Network#3*
34	Saturday	8 September 2001	38.47	Verification	Calibration	Calibration
35	Sunday	9 September 2001	45.87	Training	Calibration	Calibration
36	Monday	10 September 2001	45.54	Training	Calibration	Calibration
37	Tuesday	11 September 2001	39.90	Training	Calibration	Verification
38	Wednesday	12 September 2001	34.89	Training	Calibration	Verification
39	Thursday	13 September 2001	30.91	Training	Calibration	Verification
40	Friday	14 September 2001	37.10	Training	Calibration	Verification
41	Saturday	15 September 2001	37.46	Training	Calibration	Verification
	Sunday	16 September 2001				
	Monday	17 September 2001				
	Tuesday	18 September 2001				
	Wednesday	19 September 2001				
	Thursday	20 September 2001				
	Friday	21 September 2001				
	Saturday	22 September 2001				
	Sunday	23 September 2001				
	Monday	24 September 2001				
42	Tuesday	25 September 2001	25.90	Training	Verification	NOT USED
43	Wednesday	26 September 2001	25.10	Training	Verification	NOT USED
44	Thursday	27 September 2001	26.50	Training	Verification	NOT USED
45	Friday	28 September 2001	28.08	Training	Verification	NOT USED
46	Saturday	29 September 2001	23.01	Training	Verification	NOT USED

As given in this table, network model #3 is the most practical case. In this case we have used the data from the earlier days to train and calibrate the network model and then used the network model to predict later days. It is important to note that in the case of network model #3, the temperatures in the later days are within the range of the training data. In reality if we have a full year of data to develop our network model we should theoretically have covered all the possible ranges of temperatures. Then once the temperature hits new highs or lows in a particular year the network has to be retrained. The portions of Table 2 that are empty represent the days that no data were collected.

Table 3 gives the appropriate statistics for each of the network models presented above. It is clear that the neural model building process was a success. In all three models the R^2 for both rate and pressure are above 0.99.

Table 3 Statistics about neural network models #1, #2 and #3

	Output	*Training*		*Calibration*		*Verification*	
		Rate	*Pressure*	*Rate*	*Pressure*	*Rate*	*Pressure*
Network 1	Cases	660		210		118	
	R^2	0.9968	0.9975	0.9919	0.9959	0.9907	0.9958
Network 2	Cases	693		192		103	
	R^2	0.9972	0.9987	0.9827	0.9943	0.6336	0.9742
Network 3	Cases	645		143		94	
	R^2	0.996	0.9977	0.9862	0.992	0.9471	0.9924

Figures 7–12 show the actual rates and pressures versus the predicted rates and pressures for training, calibration and verification datasets for all three neural network models. These figures show the degree of accuracy that can be achieved in modelling this complex system.

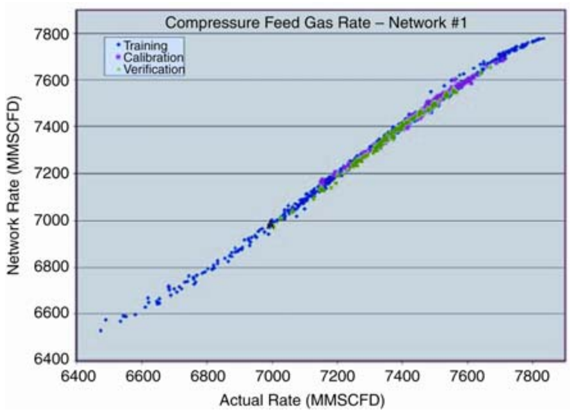

Figure 7 Cross plot for Compressor Feed Gas Rate, network model #1

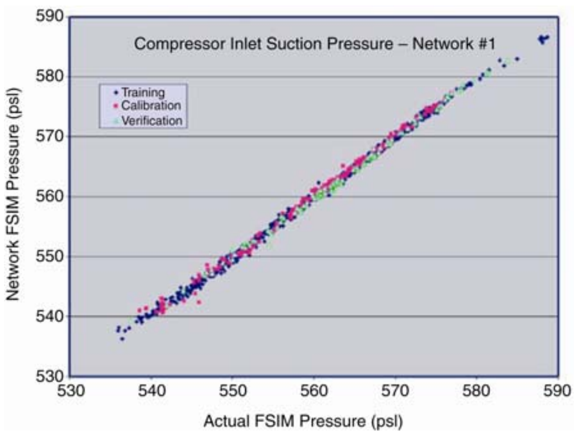

Figure 8 Cross plot for Compressor Inlet Suction Pressure, network model #1

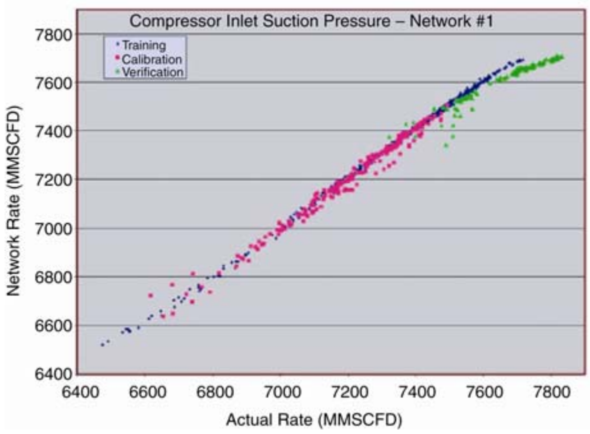

Figure 9 Cross plot for Compressor Feed Gas Rate, network model #2

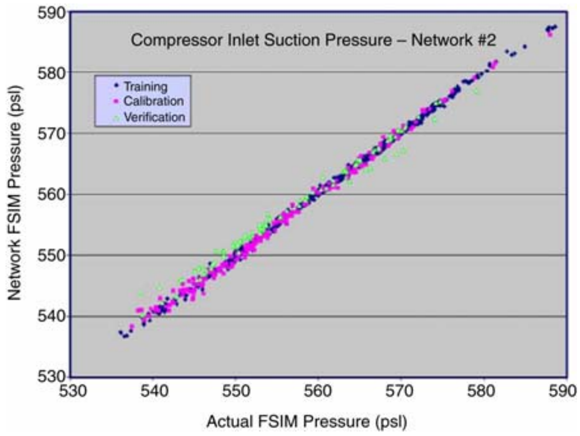

Figure 10 Cross plot for Compressor Inlet Suction Pressure, network model #2

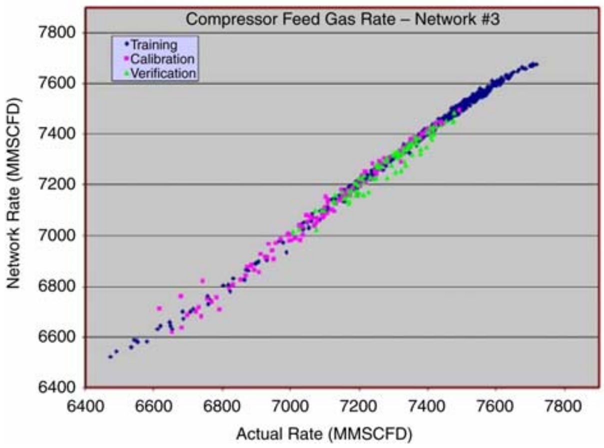

Figure 11 Cross plot for Compressor Feed Gas Rate, network model #3

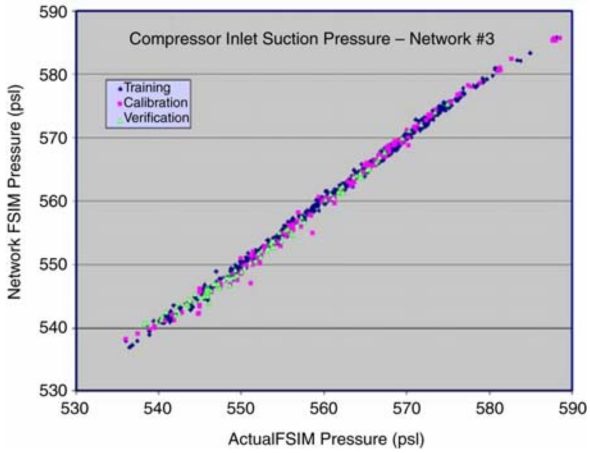

Figure 12 Cross plot for Compressor Inlet Suction Pressure, network model #3

4.2.2 Separation facility gas discharge models

A second set of neural networks was developed to model the gas discharge rates and pressures at each of the eight separation facilities. Since this is a dynamic problem where rate and pressure at each of the facilities depends on the rate and pressure at each of the other facilities as well as the corresponding rate and pressure at the inlet to the Central Compression Plant, the network model built for each of the facilities serve as a pressure-rate check for the optimisation process. This is to ensure that the pressure-rate combinations suggested by the optimisation routine for each facility does not exceed the local gas capacity or pressure limits.

Table 4 gives the statistics on each of the separation facility neural network models. High correlations were achieved for all the network models presented above. The R^2 for the verification datasets were as high as 0.975 for FS3 but never lower than 0.906 that was the case for GC2.

Table 4 Statistics for neural network models developed for the separation facilities

	Training	*Calibration*	*Verification*
Cases	*693*	*197*	*98*
R^2 for FS1	0.952	0.938	0.922
R^2 for FS2	0.933	0.918	0.909
R^2 for FS3	0.983	0.966	0.975
R^2 for FS1A	0.948	0.948	0.938
R^2 for GC1	0.963	0.954	0.969
R^2 for GC2	0.907	0.911	0.906
R^2 for GC3	0.958	0.949	0.953
R^2 for GC1A	0.932	0.940	0.927

Once these network models are constructed, calibrated and verified, one may want to use them to build Pressure-Rate curves. If such exercises are to be performed it is very important to note the following. These models are not being built based on our theoretical understanding of the system, rather by building representative functions that can approximate the collection of the data present in the dataset. The nature of the data being studied in this paper is discrete. This means that although we are modelling a continuous dynamic system, what is available to us is a discrete dataset that at best is snap shots – time freeze (on an hourly basis) of a dynamic system. Furthermore, it must be noticed that these snap shots in time do not cover all the possible situations that might occur. Rather, they represent only those situations that were covered by the limited dataset available to us. Therefore, in some instances it is possible that the data present in the dataset does not fully represent all the possible cases. One must expect to see an atypical behaviour of a Pressure-Rate curve that may or may not fit our theoretical understanding of the process.

Figures 13 and 14 illustrate typical Pressure-Rate curves that can be generated using the neural network models constructed for each of the separation facilities. These figures represent the Pressure-Rate curves at each of the eight separation facilities for a temperature of 50° Fahrenheit. The Compressor Inlet Suction Pressure is used to implicitly represent the state of the entire system.

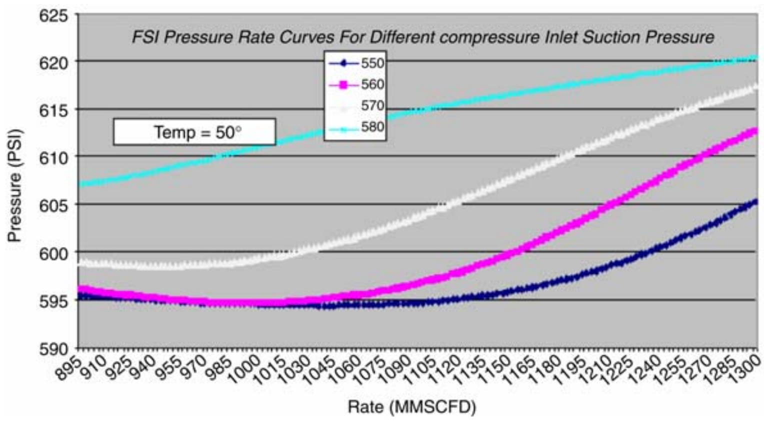

Figure 13 Rate versus pressure curves for FS1 at 50° Temperature and different Compressor Inlet Suction Pressures

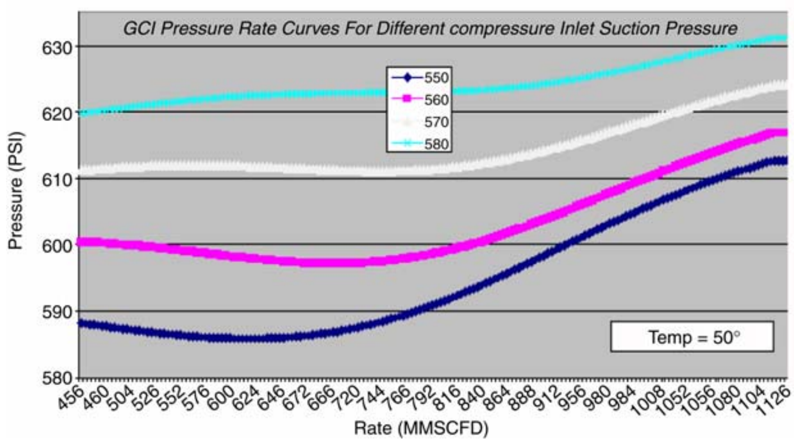

Figure 14 Rate versus pressure curves for GC1 at 50° Temperature and different Compressor Inlet Suction Pressures

5 Conclusion and future work

It is possible to represent the gas transit line system at Prudhoe Bay by a group of neural network models. However, additional data is required to retrain the network models for larger range of conditions. The next step in this development would include:

1. A rigorous data collection process will be put in place to obtain data for a broader range of conditions to retrain the network model.
2. Additional data will be collected to validate a deterministic pipeline model of the gas transit line system, which has been built using commercial pipeline simulation software. Once validated, this model will be used to generate additional data to train the neural network models. This will allow a wider range of sensitivities to be performed to generate potential solutions to the optimisation problem.

Acknowledgements

The authors would like to thank the management of BP Exploration (Alaska) Inc., Phillips (Alaska) Inc. and Exxon Mobil Corporation for their support and for granting permission to publish this paper. We also thank Hal Tucker, Mike Bolkavatz, Richard Bailey, Bryn Stenhouse and Robert Rood for valuable discussions on the applicability of these techniques to Prudhoe Bay.

References

Litvak, M.L., Hutchins, L.A., Skinner, R.C., Wood, R.C., Darlow, B.L. and Kuest, L.J. (2002) 'Prudhoe Bay e-field production optimization system based on integrated reservoir and facility simulation', *Paper SPE 77643 Presented at the 2002 Annual Technical Conference and Exhibition*, San Antonio Texas, October.

Mohaghegh, S.D. (2000a) 'Virtual intelligence applications in petroleum engineering: Part 1; artificial neural networks', *Journal of Petroleum Technology*, September, pp.64–73.

Mohaghegh, S.D. (2000b) 'Virtual intelligence applications in petroleum engineering: Part 2; evolutionary computing', *Journal of Petroleum Technology*, October, pp.40–46.

Nomenclature

GOR Gas Oil Ratio

Surface separation facilities:

FS1 FlowStation 1

FS2 FlowStation 2

FS3 FlowStation 3

FS1A FlowStation 1 Annex

GC1 Gathering Centre 1

GC2 Gathering Centre 2

GC3 Gathering Centre 3

GC1A Gathering Centre 1 Annex

Central Compression Facilities

CGF Central Gas Facility

CCP Central Compression Plant