IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)
e-ISSN: 2278-1676,p-ISSN: 2320-3331, Volume 13, Issue 6 Ver. I (Nov. – Dec. 2018), PP 43-55
www.iosrjournals.org

Automated Electrically Operated Dehydration Bed Modelling: Automatic Generation of Expression Approach

Enyenihi H. Johnson[1], Akaninyene J. Ntuk[2], Uduak I. Akpan[1]
[1]Department of Elect/Elect Engineering, AkwaIbom State University, IkotAkpaden, Nigeria
[2]Department of Elect/Elect Engineering, Exxon Mobil, Nigeria
Corresponding Author: Enyenihi H. Johnson

Abstract: *Almost every operational physical system in engineering is designed with a concept of mathematical principle which has theories backing them, thus making these systems theory-driven. On the contrary, these systems operate based on parameters which are seen as data. With advanced technique of data acquisition and data processing, these systems can be modeled using the acquired data and parameters thus converting the theory-driven models to data-driven models. In this paper, an automated electrically operated dehydration bed is modelled and analysed using the automatic generation of expression approach. Analysis confirmed a high level of excitation in the input data sets. The excitation orders for all the inputs are [50 50 50] meaning that model with orders higher than 50 will be problematic to estimate.*

Date of Submission: 14-01-2019 Date of acceptance: 31-01-2019

I. Introduction: Overview of Data Driven Modelling

With increased quest in engineering to improve on the developed technology, data driven modelling has become an inevitable tool of such quest, owing to the fact that data are collected for analysis and improve quest from an existing system. Of recent there has been tremendous effort in the world to preserve and store information in data form, however this act of preserving the data without really utilising it, will amount to wasted efforts. More so, every operational physical system in engineering is designed with a concept of mathematical principle which has theories backing them, thus making these systems theory-driven. On the contrary, these systems operate based on parameters which are seen as data. With advanced technique of data acquisition and data processing, these systems can be modeled using the acquired data and parameters thus converting the theory-driven models to data-driven models[1,2]

Data driven modelling process can be achieved using different methods and such methods include but not limited to:

- statistical method
- artificial neural network method
- fuzzy based system method[3]

But in consideration of the methods and means of achieving the purpose and application of data driven modelling, certain aspects of studies has greatly added and aided the processes data driven modelling and such processes are:

- computational intelligence
- data mining
- machine learning
- knowledge discovery in data
- intelligent data analysis

More importantly, [4] reveal that in applying the data-driven model approach, large data-sets can be examined and useful information are extracted easily. Such information that can be derived include clustering, classification and predictions

According to [9], data driven models are models described based on the inter-relationship between the parameters that make up a specific system, taking into consideration very limited amount of concepts and theories that describe the performance and limitations of such a system. And such parameters that make a system include the input parameter(s), the output parameter(s) and the internal variables. Conclusively data driven modelling is an approach used in studying and examining data for the purpose of deriving meaningful and useful information describing the physical system that is of interest with core emphasis and concentration on the use of Computational Intelligence (CI) method in actualisation and construction of the model. Incidentally, the whole beauty of this data –driven modelling approach is seen when the system input data is plotted against the output data and their unknown relationship is revealed, such relationship can be linear or non-linear and can be utilised for future predictions which eventually helps in the process of making decisions.

The data-driven modelling approach has been used to solve problems in different works and fields of life such as modelling of hydrology[7]. Also, [9] made a stride in a bit to improve upon the laid foundation in the field of wind turbine design by applying the data-driven modelling approach.

[10]defines data driven modelling as models that are constructed through the process of fitting equations to the data obtained from the physical process for the purpose of predictions without acknowledging the physics principle backing the physical process.

It is observed that the function of the data driven modelling is to aid and guide in the process of deriving a correct and useful mathematical function relating the inputs and output variables[6]. The illustration is as shown in figure 1.

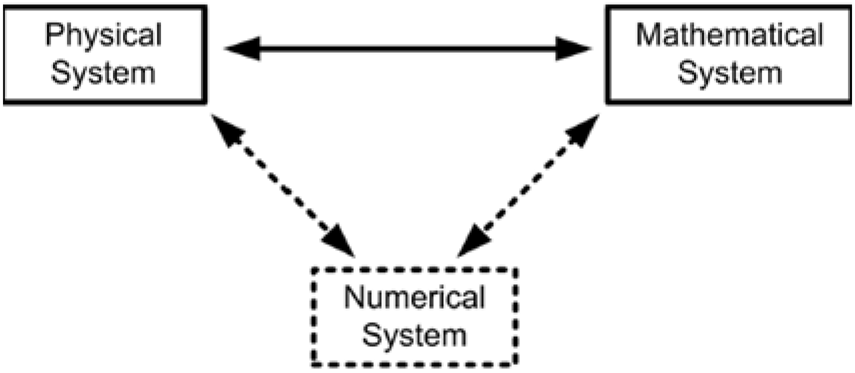

Figure 1: data driven model

But pointed out that in finding a function relating the output and input variables, the machine learning (ML) algorithm is employed. They used the illustration Figure 2 in the analogy of data driven modelling approach.

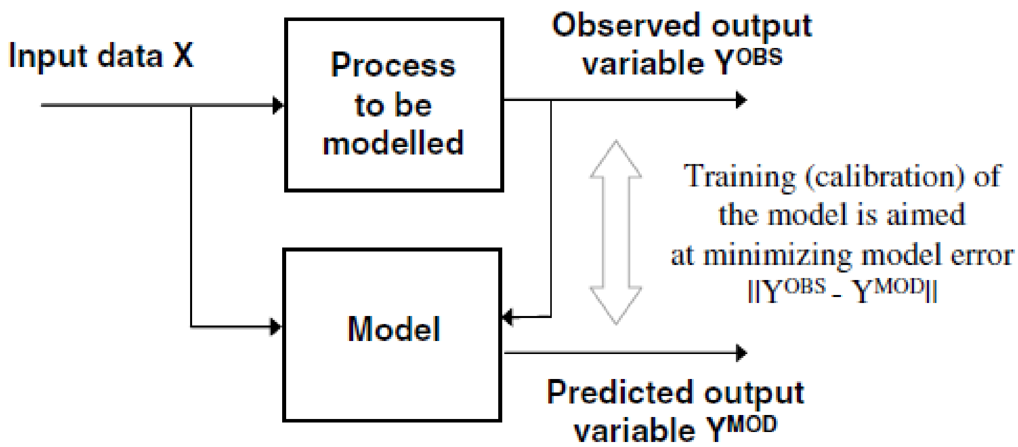

Figure 2: Analogy of data driven modelling approach

Significance of the Project

With the use of data-driven models and generation of expression approach, an expression (mathematical formula) for dew point of natural gas exiting the molecular sieve electrical dehydrator bed for the cryogenic process can be created and minimize the use of different equation of state (EoS) lumped together to achieve the same purpose. Furthermore, dew point of natural gas exiting the molecular sieve dehydrator bed for the cryogenic process can be adequately taken care of even with fluctuating input variables, through this data driven modelling approach. This can also aid in the process of building of new and efficient molecular sieve dehydrator beds for natural gas liquid extraction.

Aim and Objectives of the Project

The aim of the project is to automatically derive/ generate an expression for the dew point of a natural gas stream exiting molecular sieve dehydrator bed using Data-Driven Modelling (DDM) and generation of expression approach with three input variables (Feed Gas Flow-rate, Feed Gas Temperature and Feed Gas Pressure).

The objectives of this project are as follows:

- To obtain real process (input and output) data and Prepare the data for modelling
- To develop and define the data set type using MatLab functions
- To traine/estimate a Model with the prepared data using System Identification Toolbox or Fuzzy Logic in MatLab
- To carryout Validation of the trained Model
- To carryout Testing of the validated Model using different evaluation criteria

II. Experimental Set Up and Data Processing

This section deals extensively on the experimental set up and explanation of the plant operations of the natural gas liquid production unit, the data acquisition system and the initial data processing before model estimation in system identification toolbox.

However, generation of expression of a data set can be done in two different ways, either manually or through automatic means. The manual method of generation of expression deals with model of a system from first principles and often involves the use of mathematical formulas proposing and supporting the physical plant under study methods. Whereas the automatic generation of expression involves the use of some sort of software tools in finding the relationship which be either linear or non-line between the input and the output variables of a system without taking into consideration the mathematical formulas supporting the physical system

Acquired Data Description

The data for this analysis was recorded from a data acquisition device of agas company in Nigeria with three input variables, Feed Gas Flow-rate, Feed Gas Temperature and Feed Gas Pressure, and Output parameter as Dew point Temperature.

The Pi software used with combination of variable sensors and PROVOX and Delta V monitoring interface in the data acquisition process samples the value of the measured data at 6.00 am (once a day) and log it down for future use. The data obtained from the pi software for the purpose of this project covered a period of two and a half years.

Some of acquired data contained values that were either above or below the operational designed limits of the selected system. These constitute the inconsistent values in the data known as Outliers. Besides the outlier noticed in the obtained data, it was observed that some of the measured variables in the data missing. Accordingly, the pi software recorded such as zero for those days and for the purpose of the study, these values are termed Missing Values.

Experimental Workflow

For the experimental work of this research, the system identification technique is utilized in the modelling of the dew point of the natural gas exiting the dehydrator bed for the production of natural gas liquid. The System Identification (SID) Toolbox in MATLAB software is a convenient methodology for estimation of linear and nonlinear mathematical models of dynamic systems from measured data, most especially such systems that cannot be easily modeled from first principles. The resultant models gotten are employed for investigating system dynamics, simulating the output of a system for a given input, "predicting future outputs based on previous observations of inputs and outputs, or for control design". (Ljung, 2013)

System identification employs two key techniques in data-driven modelling. The first one is known as black-box modelling specifically useful when your primary interest is in fitting the data regardless of a particular mathematical structure of the model, and is more often than not a trial-and-error method, where estimation of parameters of various structures are carried out and the results are compared. The second is the grey-box modelling which is used when estimating the values of the unknown parameters of your model structure that has already been deduced from physical principles.

Summarily, the work flow is represented in figure 3

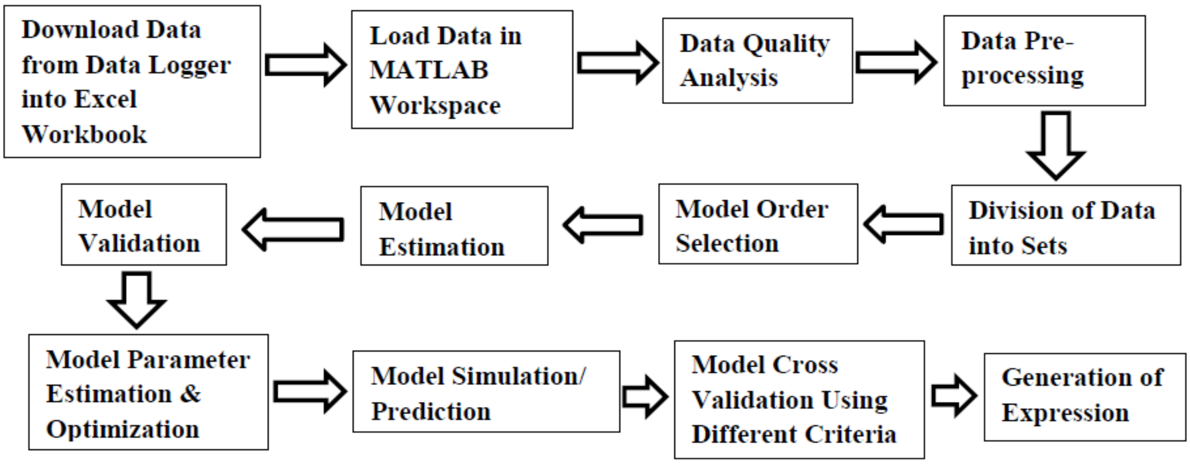

Figure 3: Block Diagram of the project Experimental Workflow

Data Acquisition Technique

In acquiring and logging of input and output variables, the process variable were measured and transmitted to the server through fibre optic cables with the aid of process variable transmitters. The data acquisition system uses a combination of Distributed Control System (DCS) designed and manufactured by PROVOX and Delta V. However; the Triplex and DCS screens obtain the displayed data from the serve for monitoring and control purposes.

The Pi software is used in the system for remote monitoring of the plant operations and it variables; nevertheless the pi software has the capability of acquiring the data from the server and logging it. The Pi software samples the values of the measured data (process variables) at sampling interval of 24 hours (once a day), specifically at 6.00 am and log it for report and future reference purposes.

Data pre-processing

The obtained raw data from the data acquisition device could not be used for the modelling of the dew point of natural gas exiting the molecular sieve dehydrator instantly as it contains some defects. Such defects include the missing values, outliers, offsets and drift and some disturbances. Therefore there was need for pre-data processing. The Data pre-processing techniques employed included:

- Detection and Removal of outliers
- Filtering
- Detrending
- Resampling
- Replacement and removal of missing values

Data Quality Analysis

Prior to commencing the estimation of models from data, the measured data was checked for the presence of any undesirable characteristics by carrying out a time plot of the data and the undesirable characteristics may include the following:

- Missing data samples
- Drifts and outliers
- Offsets and trends

Transient Plot of Data

To determine and examine the quality of the data obtained for the modelling of the dew point of natural gas exiting the molecular sieve dehydrator bed, the data was plotted and as shown in figure 4. Also, before the transient plot of the data, all inputs and output variables were defined.

Handling Missing Data

Certain factors are responsible for the missing values noticed in the acquired and logged data. These factors include:

- Shutdown of the system and other plants that serve as sources of feed the system, either for the purpose of turnaround maintenance or unscheduled shutdown.
- Failure of either the process variable sensing element or the transmitter.
- Malfunction of the pi software itself.

- Network failure within the facility instrument networking and the data acquisition system.

This is one of the pre-processing treatments that can be applied to the data set before carrying out data quality analysis. Once a data file with missing values is loaded into the MATLAB program workspace, MATLAB convert the missing values to NaN which means Not-a-Number without interfering with the structures of the variables containing the missing values. The missing values (NaN) were removed from the data set using the MATLAB program code.

Offsets and Trends

Offsets and trends are arbitrary differences between the input and output signal levels observed in measured data sets. If offsets and trends are not removed from the data sets, the model identification and estimation would most likely remain inaccurate in situation of the dehydrator bed consisting of high level of linearity. Offsets and trends are handled by detrending which is the process of removing means, from regularly sampled time-domain input-output data signals

At this point, there is need to carry out further examination of the data set in order to check the frequency, step and impulse responses of these data sets.

Moreover, these outliers and offsets cannot be removed from the data set at this point as they may affect the model estimation, until further data analysis of the initial experimental stage such as step, impulse and frequency responses are carried out, and then will the decision be taken whether or not to remove the outlier.

Experimental stage

However, further assessment of the data is required to examine static gain, time delay, poles and zeros. Such further checks on the obtained data include:

- Frequency response
- Step response and
- Impulse response
- excitation level
- presence of feedback

In order to do this, the data was presented to System Identification GUI or commands. The System Identification data objects was used to present the data in format that will be accepted by System Identification Toolbox. The System Identification has three data representations namely, time-domain, time-series and frequency domain data objects. Since the data was sampled as a function of time, it is therefore pertinent to present the data as a time-domain data which is to be represented as a time domain object.

Frequency Response of Data

The *spa* function was used to estimate frequency response with fixed frequency resolution using spectral analysis before representing this on a bode plot. However, *spa* applies the Blackman-Tukey spectral analysis method by computing the covariance and cross-covariance from the inputs and output of the data. The algorithm further computes the Fourier transforms of the covariance and the cross-covariance before it finally calculates the frequency-response function and the output noise spectrum of the system. The Bode plot shows the frequency response of the molecular sieve dehydrator bed system variables.

Step Response of Data

In carrying out the data step response analysis and plot, the code for the programming syntax was written. The step response plot of the three input variables to the output variable is as shown in figure 5.

Impulse Response of Data

In carrying out the data impulse or time response analysis and plot, the code for the programming syntax was written. The impulse or time response plot of the three input variables to the output variable is as shown in figure 3.6 below.

Determination of Excitation Level in Data Sets

In determination of the excitation level in the data set, MATLAB program code was used to calculate the level of excitation in the data sets.

The accuracy of the model depends so much on a good excitation level of the inputs to the dehydrator bed within the frequency range of the system. According to [11], the following techniques can be used to assess whether or not the inputs excites the plant appropriately:

1. Power spectral density,

2. Coherence spectrum,
3. Conditions for persistent excitation.

Though, there are several methods for estimating the power spectral density of signals, which include:

- **Periodogram**: this is a nonparametric approach known as Power spectral density estimate. The periodogram spectral estimator is given by the equation

$$\emptyset_p(\omega) = \frac{1}{N}|D|^2 \quad (1)$$

Where ***D*** is the Discrete-Time Fourier Transform of the signals and it is expressed as

$$D = \sum_{t=1}^{N} y(t)e^{-i\omega t} \quad (2)$$

Where is the available signal samples evaluated within the range *{y(1),…,y(N)}.*

- **Welch:** averaged periodograms of overlapped, windowed signal sections and is also a nonparametric method. The method consists of dividing the time domain data into overlapping segments, and then computes a modified periodogram of each of the segment, and averaged the PSD estimates, producing a Welch's PSD estimate. The Welch estimator of the PSD is given as

$$\emptyset_w(\omega) = \frac{1}{S}\sum_{j=1}^{S} \emptyset_j(\omega) \quad (3)$$

Where $\emptyset_j$ is the windowed periodogram corresponding to signal sample segment $y_j(t)$. Note that the signal segments is

$$y_j(t) = y\big((j-1)K + t\big), \quad {}^{t=1,\dots,M}_{j=1,\dots,S} \quad (4)$$

The value recommended for K in Welch method is $K = M/2$ and consequently $S = 2M/N$ segments with 50% overlap between segments. However,

$$\emptyset_j(\omega) = \frac{1}{MP}\left|\sum_{t=1}^{M} v(t)y_j(t)e^{-i\omega t}\right|^2 \quad (5)$$

P denotes the power of the temporal window $\{v(t)\}$:

$$P = \frac{1}{M}\sum_{t=1}^{M}|v(t)|^2 \quad (6)$$

- And the other methods are Multitaper**,** Yule-Walker Autoregressive (AR), Covariance, Modified Covariance, Pisarenko and Music, Eigenvector, etc.

In this project, the Welch method was employed to compute the power spectral density of the input signals to the dehydrator bed.

Checking for Feedback in Data Set

The final step in data quality assessment in this work was to check for the presence of feedback in the data. There was a strong indication of feedback in the system, possibly caused by a Pulse Regulator. Feedback calculation also confirmed that the feed gas temperature appears to have a direct effect on dew-point, that is, no delay from input to output.

III. Resultand Discussion

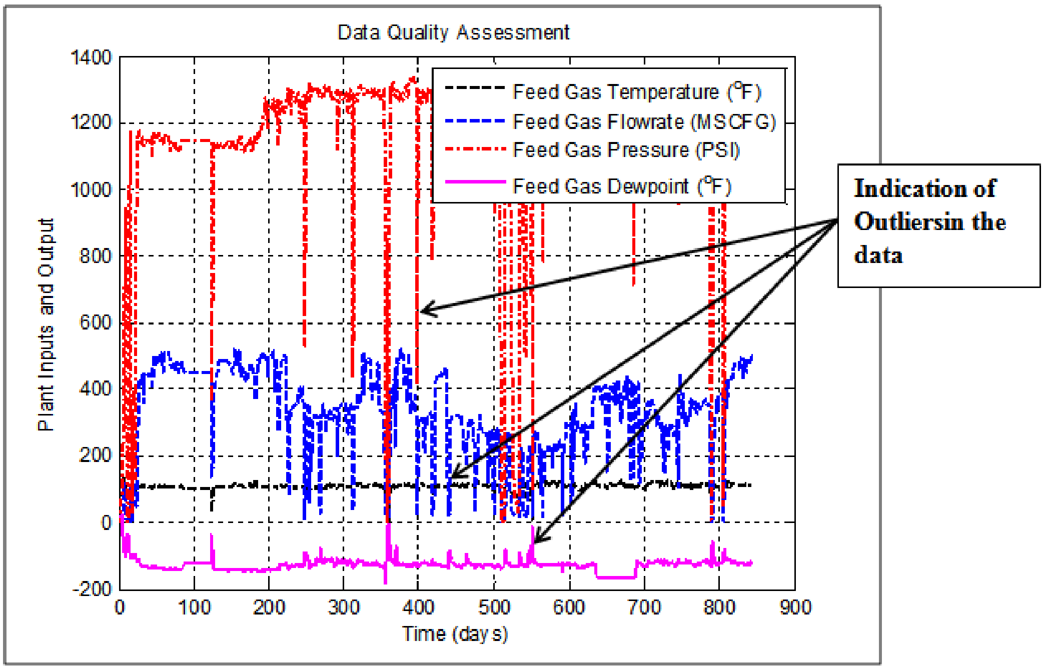

Figure 4: Time Plotting of Data for Data Quality analysis

The plot indicates that there are drifts and outliers in feed gas flow-rate, pressure (inputs) and the output (dew-point of the gas). Such outliers might be caused by signal spikes or by measurement sensor malfunctions. These drifts and outliers are handled by:

- Segmenting the data into informative portions and merge them into one multiple experiment.
- Out of range values for the plant inputs and output can be removed manually and replaced with more reliable estimates.
- Pre-filtering with appropriate filter helps remove high frequency noise (outliers) and low-frequency disturbance (drifts).

The outliers in the measured data set can adversely affect the estimated models.

Result and Discussion of Data Frequency Response Plot

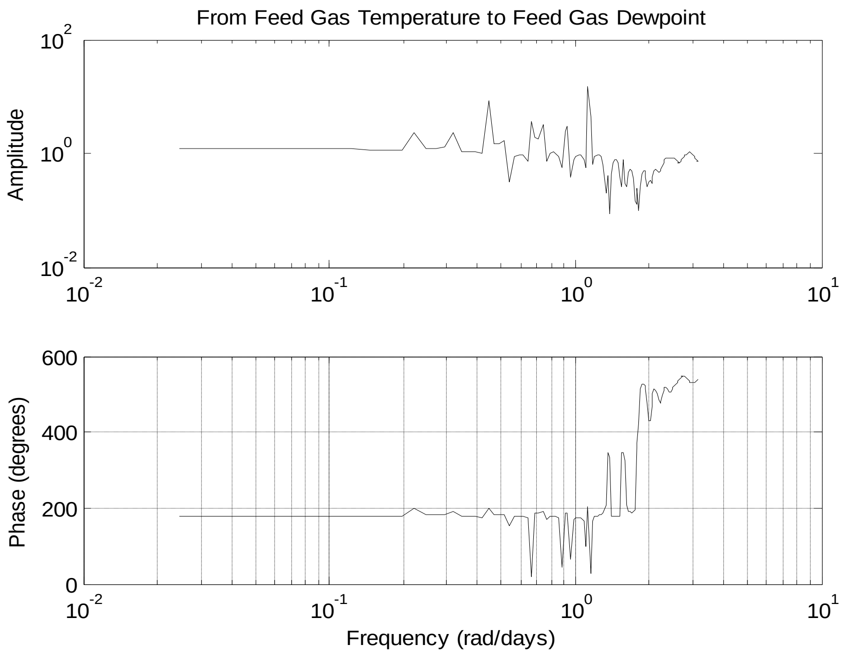

Figure 5: Frequency Response Plotting of Feed Gas Temperature to Feed Natural Gas Dew Point

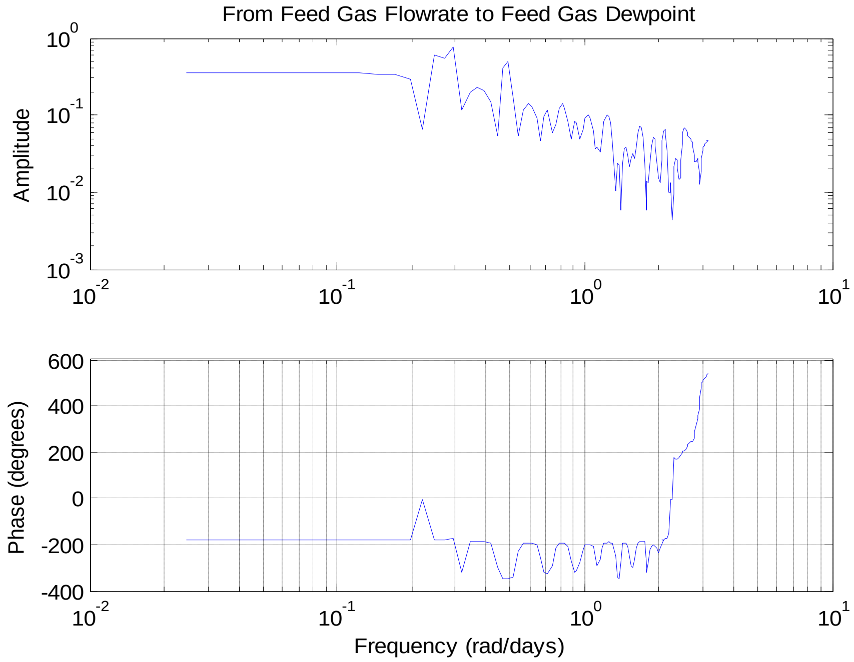

Figure 6: Frequency Response Plotting of Feed Gas Flow-rate to Feed Natural Gas Dew Point

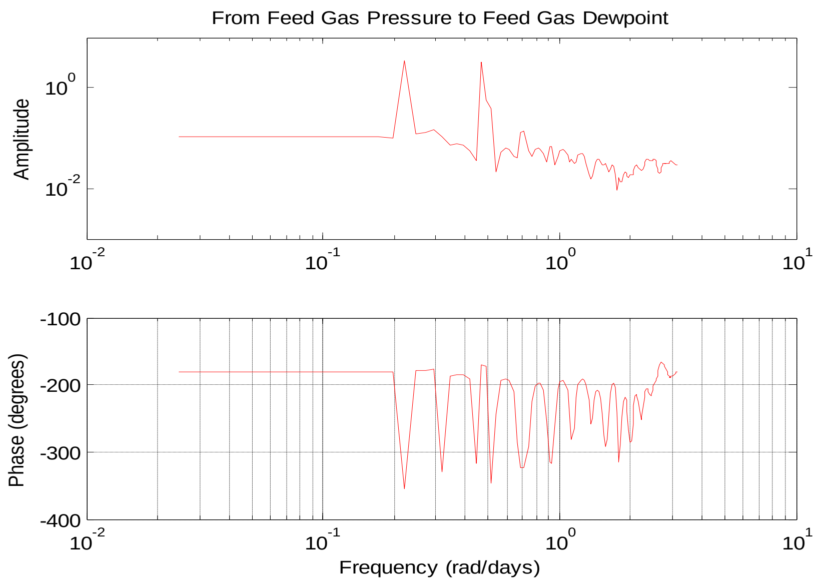

Figure 7: Frequency Response Plotting of Feed Gas Pressure to Feed Natural Gas Dew Point

Frequency response function describes the steady-state response of the dehydrator system to sinusoidal inputs. There are amplitude peaks at certain frequencies for all input-output data combinations which show that the dehydrator will become unsteady at these frequencies. The amplitude peaks at the frequencies of about 0.7 rad/day, 0.5rad/day and 0.4rad/day suggest a possible resonant behaviour (complex poles) for all the input-to-output combination. Moreover, for all pairs, there is rapid phase roll-off at frequency > 0.3rad/day, suggesting the presence of time delays in the data sampling.

Result and Discussion of Data Step Response and Plot

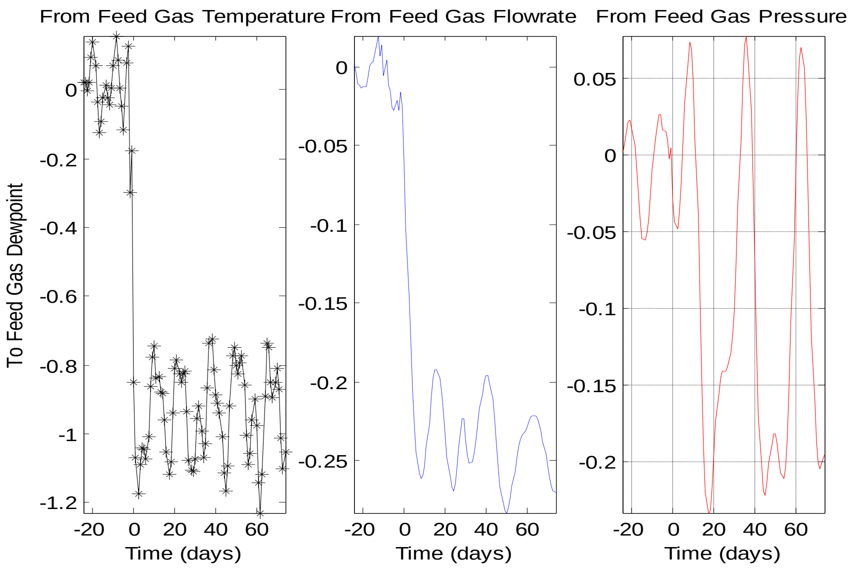

Figure 8: Step Response Plotting of Feed Gas Temperature, Flow-rate and Pressure to Feed Natural Gas Dew Point

The step response for all the input-to-output (feed gas temperature, flow-rate and pressure to feed natural gas dew point) combinations suggests an overshoot, which indicates the presence of an under-damped mode (complex poles) in the dehydrator system. The plot of step-response point out a non-zero delay in the molecular sieve dehydration system, which is in harmony with the rapid phase roll-off, established by the Bode plot in figures 3.2a, 3.2b and 3.2c above.

Result and Discussion of Data Impulse or Time Response Plot and Analysis

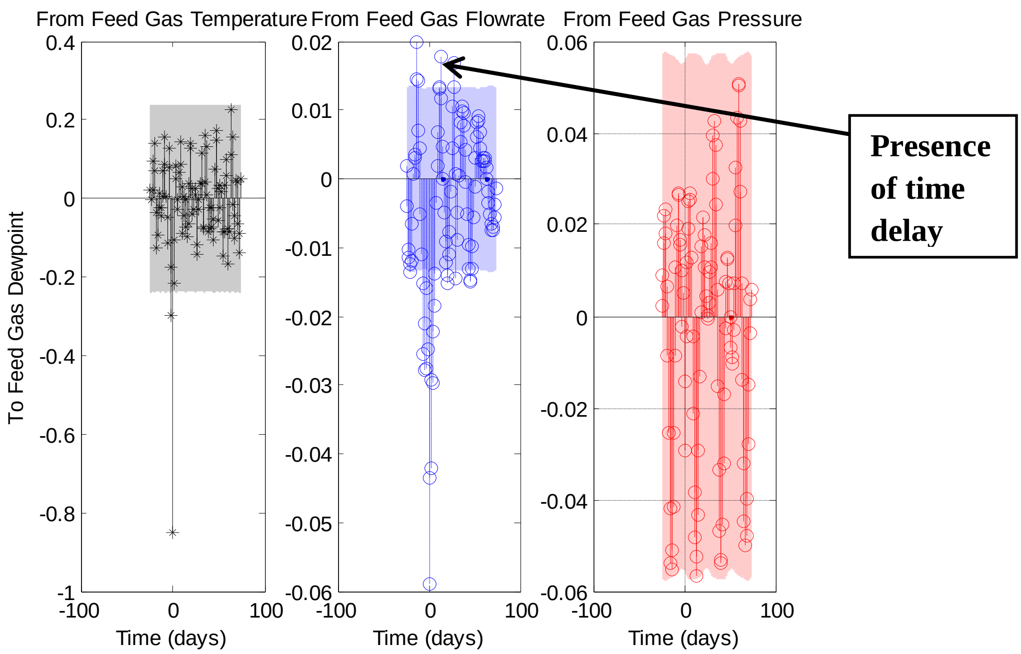

Figure 9: Impulse Response Plotting of Feed Gas Temperature, Flow-rate and Pressure to Natural Gas Dew Point

The filled portions on the plot indicate the confidence region.

The essence of the impulse plot is to confirm the time delays that have been suggested by both *bode* and *step* plots. The *impulse* command can be used to plot the impulse response. The time delay is equal to the first positive peak in the transient response magnitude that is greater than the confidence region for positive time

values.The feed gas temperature and pressure to gas dew-point registered no time delays, indicating that the dehydrator system has a direct response from the feed gas temperature and pressure inputs. Conversely, the feed gas flow rate to gas dew-point has a time delays of about 9 samples before the system response to the effect of the feed gas flow-rate input. However, these observed delays may change based on the confidence interval used, but for this data sets, standard deviation of 3 is realistic for the dehydrator system.
To further confirm the time delay as indicated in figure 3.6 above, the coding confirmed the presence of time delays in the data by calculation with a suggested model order using "*delayest*" command.

Result and Discussion of Determination of Excitation Level in Data Sets
Using a normalized frequency and segment length of 400 samples with a 350-sample overlap, the plot of two-sided PSD estimate is shown below.

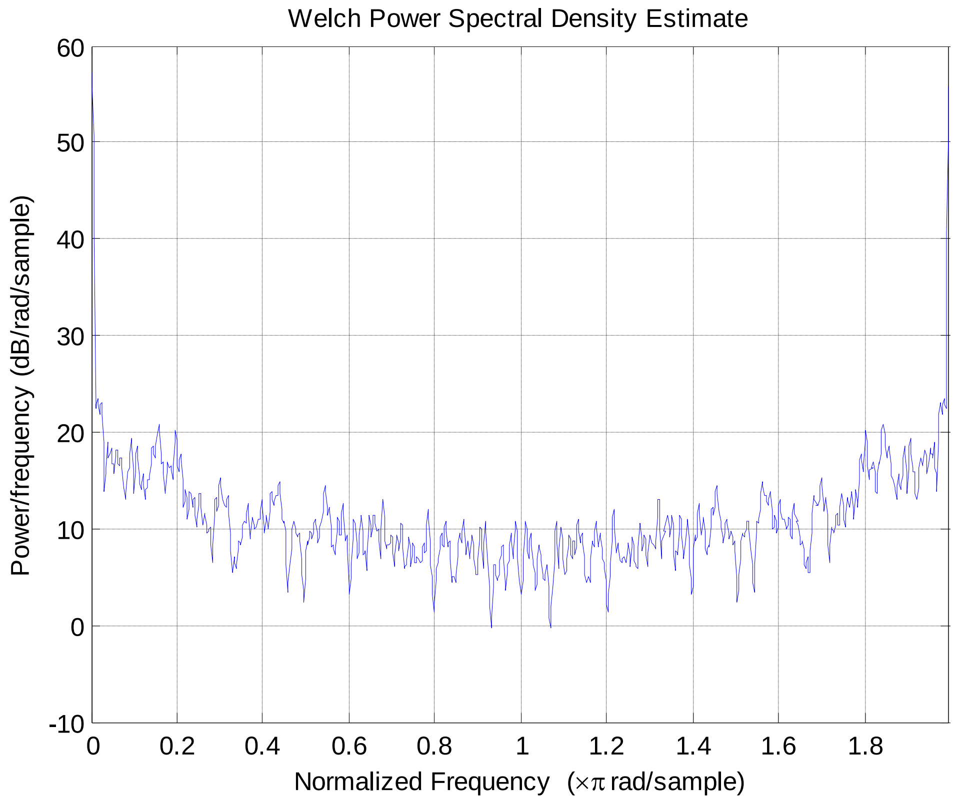

Figure 10: The Welch Method Power Spectral Density Plot of the Data Input Signals

It is obvious from the plot that the signals have sufficient energy throughout the entire frequency band, ranging from 0-23dB/rad/sample. Some of the inputs become persistently excited at PSD greater than 50dB/rad/sample.
Coherence spectrum of the input *u(t)* and output *y(t)* signals is written as

$$C_{yu}(\omega) = \frac{\emptyset_{yu}(\omega)}{[\emptyset_{yy}(\omega)\emptyset_{uu}(\omega)]^{1/2}} \quad (3.7)$$

Where $\emptyset_{yu}(\omega)$ is the cross-spectrum of the inputs and output of the signals, and $\emptyset_{yy}(\omega)$ and $\emptyset_{uu}(\omega)$ are their respective PSDs. $C_{yu}(\omega)$ can be evaluated from MATLAB using Welch method as the algorithm. Thus, a magnitude-squared coherence spectrum computed with a 400 samples segments of a hanning window and 300-sample overlap yields the following results.
The magnitude-squared spectrum measures the correlation between *u(t)* and *y(t)* whose values are just within a range of 0-1.

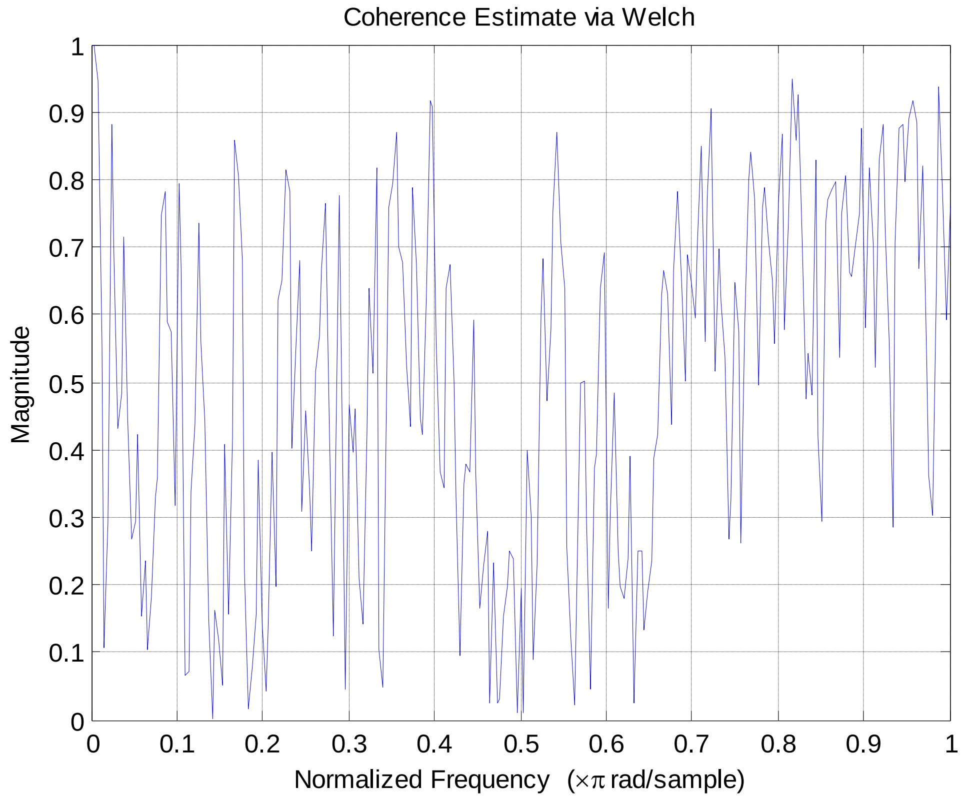

Figure 11: The Welch Coherence Estimate of the Data

The conditions for excitation in the inputs can be calculated with MATLAB function *pexcit*. Analysis confirmed a high level of excitation in the input data sets. The excitation orders for all the inputs are [50 50 50] meaning that model with orders higher than 50 will be problematic to estimate.

Data Examination and Decisions

Deciding on whether or not to remove outlier, detrend and carry out certain treatment on the data set is crucial to the estimation model process. In making this decision of how to handle the data after data quality analysis, the command known as *'advice'* in System Identification toolbox in MATLAB software was used. It helps in displaying information about the data sets and also offers advices such as:

- What are the excitation levels of the signals/data and if it affects the model orders
- Does it make meaning to remove constant offsets and linear trends from the data?
- Is there any suggestion of output feedback in the data?

General data characteristics

This data is a time domain data set with 3 input(s) and 1 output(s), 843 samples and 1 experiment(s). All inputs in the data have been denoted as 'zero order hold' ('zoh'), that is, they are assumed to be piecewise constant over the sampling interval. If the input is a sampled continuous signal and you plan to build or convert to continuous-time models, it is recommended to mark the Inter-Sample property as 'First order hold': Data.InterSample = 'foh' or Data.int = {'foh','foh', ...} for multi-input signals.

Some inputs and outputs have non-zero means. According to NageswaraRao*et al* (2012), it is generally recommended to remove the means by DAT = DETREND(DAT), except in the following cases:

- The signals are measured relative to a level that corresponds to a physical equilibrium. This could, for instance be the case if step responses are recorded from an equilibrium point. In this case, it is advisable to remove the equilibrium values rather than data means. You may do so using a TrendInfo object with DETREND command.
- Where there is an integrator in the system, and the input and output levels are essential to describe the effect of the integration.
- Where the data are to be utilized for estimation of Nonlinear Auto-Regressive with eXogenous (NARX) input models.

Excitation level in data

Input number 1 which is the feed gas temperature is persistently exciting of order 50. This means that a problems will occur when estimating models of order higher than 50, at least for model parameters associated with this input.

The excitation orders for all the inputs are [50 50 50].

Possibility of feedback in data

There is an indication that the system has a direct response from input number(s) 1 at time t to y(t).

There may be two reasons for this:

- There is direct feedback from y(t) to u(t) (like a P-regulator). In that case it is essential not to let this feedback influence the model. It is thus important always to use nk>0 for these inputs in state-space and input-output models.
- The system has a direct term (relative degree zero). Then it is essential to use nk = 0 for these inputs in all models. (Note that state-space models have nk = 1 as default, so use: MODEL = PEM(Data,n,'nk',0), where n is the model order.)

However, there is a very strong indication of feedback in the data.

Care was taken when interpreting the results of SPA and the results of output error models (Output error models result from the OE command or setting 'DisturbanceModel'= 'None' in state-space models). With feedback in data, it is recommended to estimate a model with large enough disturbance model. For example, use BJ models in place of OE models and estimate state space models using 'DisturbanceModel'= 'Estimate'.

Possibility of nonlinearity

There is an indication of nonlinearity in the data. A nonlinear ARX model of order [4 4*ones (1,3) 4*ones(1,3)] and tree partition nonlinearity estimator performs better prediction of output than the corresponding ARX model of the same order. However, consideration of using nonlinear models, such as idnlarx, or idnlhw arose. The "isnlarx" command can also be used to test for nonlinearity with more options.

Summarily, the information offered by the advice command confirmed the correctness of a greater part of the analysis performed. But the following were noted:

1. It is therefore obvious that the outlier present in the data set can neither be removed nor the offset be de-trend or filter because of the presence of nonlinearity in the dehydrator (inputs and output) system.
2. Besides, there is need to change the sampling method from zero-order-hold to first-order-hold since this a multi-input system as suggested by the advice command.
3. A Nonlinear Auto-Regressive with eXogenous input (NARX) model structure is recommended with an optimization method called Tree Partition.
4. A sampling rate of one day suggests that the data was sampled at a very slow rate. This means that the data should be resampled at higher rate in order to recover some of the lost information about the dynamics of the system.

Decision

After careful examination of the obtained results of the analysis carried out on the data set, a decision was reached not to detrend, filter and remove the outlier in the data set as these may affect the estimation of the model due to the presence of non-linearity found in the data set but resampling was done to aid in recovery of some lost information.

Resampling

This was the only data pre-processing technique that was carried out on the data sets. In this project, the resampling technique used applied an antialiasing low pass FIR filter with a 0.08 sampling interval (about 2 hours) to interpolate for the missing information about the dynamics of the system. It is also advisable to decimate your data if it was sampled at much faster rate because such data may contain high-frequency noise outside the frequency range of the system. For the purpose of model estimation and validation, the resampled data was split into two sets as displayed in table 1.

Table No 1: NARX Model Parameters

Resampled Data	**Estimation**	**Validation**
Number of Samples	5269	5269
Total	**10538**	

IV. Conclusion

In order to carry out the an efficient modeling of natural gas dew point, the obtained data needed some kind of pre-treatment and examination of the data set to ascertain the inherent attribute of the input and output variables. Such were the works carried out in this chapter three.

The data set was loaded into MATLAB work space and the presence of NaN (Not-a-Number) values indicated that these must have been periods of plant or data logging system malfunction or shutdown routine maintenance carried out on the Liquid production system. Removal of NaN values from the data set was carried out to prepare the data for further analysis. Also, codes were written to carryout data quality plot and analysis on the data set which revealed the presence of outlier in the data set.

Furthermore, the initial experimental stage of these project was executed which included computation of frequency response, step response, impulse, excitation level in inputs data, and presence of feedback in the data sets with their respective plots were shown in section 3.4 and it sub-sections 3.4.1 to 3.4.5. Nonetheless, the frequency response plot for the data set revealed a rapid phase roll-off at frequency > 0.3rad/day, suggesting the presence of time delays in the data sampling.

Additionally, the advice command was used to cross-check the pre-processing treatment earlier carried out on the data if they were appropriate.

Codes were equally written to change the interpolation method to first order hold and to resample the data at a higher rate, with a 0.08 sampling interval (about 2 hours) to interpolate for the missing information about the dynamics of the system as suggested by the advice command.

Finally in this chapter, codes were written successfully to divide the data set in to two set; one for training or estimation and the other for validation of the model. It was specified in the written MATLB program code that half of the data set be used for estimation of model.

References

[1]. Akpabio, E.J., Aimikhe, V.J.,Dynamics of Solid Bed Dehydration in a Niger Delta Natural Gas Liquids Plant, International Journal of Engineering and Technology,2012;2(12):

[2]. Andersson, L., Jönsson, U., Johansson, K. H., Bengtsson, J. "A manual for system identification,Laboratory Exercises in System Identification. KF Sigma i Lund AB. Department of Automatic Control, Lund Institute of Technology, 1998.

[3]. Åström, K.J. Eykhoff, P. System identification—A survey, Automatica,, 20097(2):23-162.

[4]. Avila, S., Benito, A., Berro, C., Blanco, S. T., Otín, S., Velasco, I. Dew-Point Curves of Natural Gas. Measurement and Modelling, Industrial & Engineering Chemistry Research,2002,45(14):179-5184.

[5]. Avila, S., Blanco, S. T., Velasco, I., Rauzy, E., Otín, S. Thermodynamic properties of synthetic natural gases: Part 4. Dew point curves of synthetic natural gases and their mixtures with water: measurement and correlation.Fluid phase equilibria,2002, 202(2):399-412.

[6]. Azar, A. T. **Adaptive neuro-fuzzy systems**. Fuzzy systems, **2010,** 85-110.

[7]. Barbour, S. L., Krahn, J. Numerical Modelling: Prediction or Process?Geotechnical News, 2010,*22,(*4): 44-52.

[8]. Bowden, G. J., Maier, H. R., Dandy, G. C. Optimal division of data for neural network models in water resources applications, 2002.38(2):1010-104

[9]. Brown, A.S., Laughton, A.P., Milton, M.J.T., Vargha, G.M., Mounce, R., Cowper, C.J., Stokes, A.M.V., Benton, A.J., Lander, D.FRidge, AMeasurement of the Hydrocarbon Dew Point of Real and Synthetic Natural Gas Mixtures by Direct and Indirect Methods, Energy & Fuels, 2009, 23(3):1640-1650.

[10]. Chen, M. Y., Linkens, D. A.Rule-base self-generation and simplification for data-driven fuzzy models, Fuzzy Sets and Systems, 2004, 142(2):243-265.

[11]. Christos, S., Dimitros, S. What is a neural network.2008 retrieved from: http//: docs. toc. com/doc/1505*/neural-networks*

Enyenihi H. Johnson. " Automated Electrically Operated Dehydration Bed Modelling: Automatic Generation of Expression Approach." IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE) 13.6 (2018): 43-55.