2014 UKACC International Conference on Control
9th - 11th July 2014, Loughborough, U.K

Preprocessing of Raw Data for Developing Steady-State Data-Driven Models for Optimizing Compressor Stations

Dionysios P. Xenos
and Nina F. Thornhill
Centre for Process Systems Engineering
Department of
Chemical Engineering Imperial College London
London, UK
d.xenos@imperial.ac.uk

Matteo Cicciotti
and Ala E. F. Bouaswaig
Advance Process Control
Automation Technology, BASF SE,
Ludwigshafen, Germany
Department of Mechanical Engineering
Imperial College London
London, UK

***Abstract*—Compressors operate in parallel to increase the supply of a gas in many applications in process industries (e.g. an air separation process). To optimally distribute the load of the compressors in parallel, an optimization problem is formulated that takes into account the operational constraints of the compressors and the objective is to reduce operational costs, i.e. power consumption of the drivers. The optimization takes place when the system is in steady-state. The structure of the optimization employs steady-state data-driven models to represent the operation in steady-state. Many researchers reported that the identification of steady-states of the data plays a key role for accurate representation of the actual process by a data-driven model. However, to the best of the authors' knowledge, there is not much research on the quantification of the influence of the output of the steady-state detection methods on the data-driven models. For these reasons there is a need to examine this topic.**

I. Introduction

It is generally accepted that compressors consume large amounts of energy in various industrial sectors. For example the utility for compressed air is considered as one of the most expensive in many industries [1]. Compressors are also major energy consumers in many intensive chemical processes such as air separation [2]. A compressor station involves single-stage or multi-stage centrifugal compressors operating in parallel to increase the amount of the gas they supply. Gas turbines or electrical motors are mainly used to drive the compressors. A compressor station with multi-stage compressors driven from electrical motors can be seen in Fig. 1. One of the assumptions mainly used is that individual compressors have the same characteristics and the same performance behaviour. Regarding this assumption, a conventional practice is to distribute the load evenly among the compressors [3].

Optimization by mathematical programming, estimates optimal load sharing and it is expected to reduce the operational costs compared to operation with the conventional equal split or other non-optimized methods. Abbaspour et al. [4] considered the optimization of compressor stations with single-stage compressors and they mentioned that a compressor station has dissimilar compressors in reality. Moreover, Han et al. [5] used optimization to optimize the operation of an air-supply and gas-supply network of a terephthalic acid manufacturing plant.

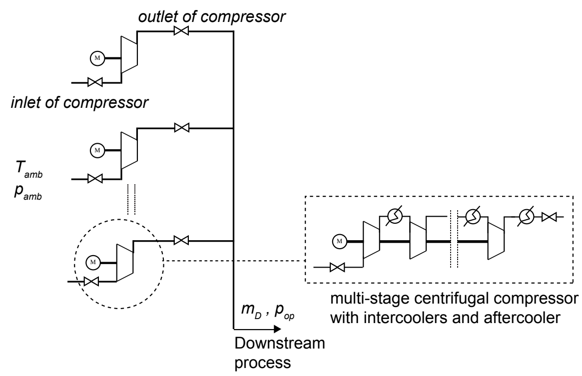

Fig. 1. Compressor station with multi-stage centrifugal compressors in parallel.

Characteristics and performance of compressors change after many hours of operations. The typical performance maps given by the vendors are not valid after a period of operation and after several episodes of maintenance. For this reason many authors focused on the update of these maps using available data from the operation. In other words historical data of available measured variables, e.g. inlet pressure and inlet mass flow rate, can be used to generate black box models, known as data-driven models. Paparella et al. [3] presented a work for updating the parameters of a data-driven model describing the speed and the efficiency of the compressors as a function of mass flow and pressure ratio. Wu et al. [6] expressed the fuel cost of the turbine drives as a function of outlet and inlet pressure, and flow of the compressor. Moreover, Han et al. [5] used hybrid models (coupled basic thermodynamic relationships and black box models) to estimate the power consumption of each compressor of a supply compressor station of a terephthalic acid manufacturing plant. DeMarco et al. [7] developed similar models as were used

978-1-4799-5011-9/14/$31.00 ©2014 IEEE

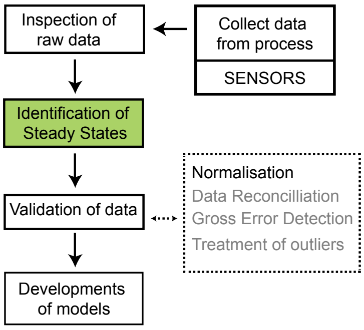

Fig. 2. Typical procedure to process raw data for developing data-driven models.

in the work from Wu et al. The former authors mentioned that operational data has to be processed to be used for the polynomial models. However, they studied data-driven models without mentioning the problem of steady-state detection of the sampled data.

Figure 2 shows a typical procedure to process raw data from operation to develop data-driven models. The highlighted block refers to the procedure of the identification of steady-states. According to [8], operational data has to be collected close to steady-state in order for polynomial regression models to reliably represent a process. There is a little research on this topic, therefore the main aim of this paper is to investigate the influence of steady-state detection in the accuracy of the models. The paper examines the operation of an air compressor station with multi-stage compressors which distribute compressed air to different end-users, such as air-separation columns and plant-site utilities for compressed air. In this particular application, the power rate of the compressor station is several tens of MWs.

The remaining sections of the paper are structured as follows. Section II presents the methodology for processing the raw data given as input data to the data-driven models. Section II-A presents the methodology of a moving window steady-state detection algorithm. The following section II-B defines a steady-state detection index, a measure of proximity which illustrates how much close is the operation to an ideal steady-state. Finally, Section II-C describes the development of data-driven models, the training and fitting of models using process data. Section III presents the development of data-driven models of compressors with processed data with different steady-state indexes. Section IV summarises the work and gives the conclusions.

II. Methodology for Developing Steady-State Models

Section II presents the procedure for detecting steady-states using raw data from the process. First, the description of a steady-state algorithm is given in Section II-A. The algorithm employs a moving data set window. A conventional steady-state detection algorithm using a moving window can detect disturbances with fast dynamics. On the other hand, when the operation changes smoothly from one steady-state to another at a slow rate, the result is that the steady-state detection cannot capture this change. In the examined case study, changes in ambient temperature influence the power consumed from the motors of the compressors. The ambient temperature changes slowly over time compared to changes in the process itself (e.g. changes in the downstream process influence the process to be analysed). The developed algorithm copes with this phenomenon. Section II-B deals with the definition of the steady-state index of the operation and Section II-C presents the methodology for the modeling of the operation using data-driven models.

A. Steady-state detection algorithm

In this paper the purpose of the steady-state detection algorithm is to identify the steady-states of the operation to develop reliable data-driven models. Data-driven models are black boxes which hold a relationship between input and output variables. Therefore, these input and output variables should be close to steady-state to hold the validity of the mass and energy balances implied in the black box. A steady-state of a process variable is a subjective concept, and a steady-state index is introduced in this work.

Figure 3 shows the application of a moving window to a data set of a process variable. A moving data set window is defined from a fixed number of data points of the process variable, n_s. The data included in the window is updated at each step τ, recent data is added and old data is discarded. The window moves every S number of data points. A process variable i has value $x_{i,t}$, where $i \in N$ (N is the number of variables) and $t \in T$ (T is the total number of data points which corresponds to the total sample time). Variable τ corresponds to a window with data $(t - n_s + 1, t)$. The sample rate of the data set, the n_s and the S are parameters which have to be tuned for a desired function of a steady-state identification algorithm regarding to the application. Kim et al. [9] studied this topic and presented results from an online steady-state detection application to a residential air conditioner. At each step τ the data set window moves $t' = t + S$ and the standard deviation of the included data points in the window, $\sigma_{i,\tau}$ of variable i is calculated from:

$$\sigma_{i,\tau} = \sqrt{\frac{1}{n_s} \sum_{t' \in (t-n_s+1,t)} (x_{i,t'} - \mu_{i,\tau})^2} \qquad (1)$$

with

$$\mu_{i,\tau} = \frac{1}{n_s} \sum_{t' \in (t-n_s+1,t)} (x_{i,t'}) \qquad (2)$$

The steady-state algorithm detects a steady-state episode of a process variable i when a particular condition holds true, for example if the $3\sigma_{i,\tau}$ is less than a predefined value, h_i (h_i is chosen by the user).

A multivariate steady-state detection algorithm involves the examination of more than one variables to conclude if the system is in steady-state (the system in the current paper is an industrial mutli-stage air compressor). According to Mansour et al. [10] a system is in steady-state when all the considered variables are in steady-state. Hence, if a process variable i is in steady-state then a variable $Y_{ss,i}$ is equal to 1, otherwise

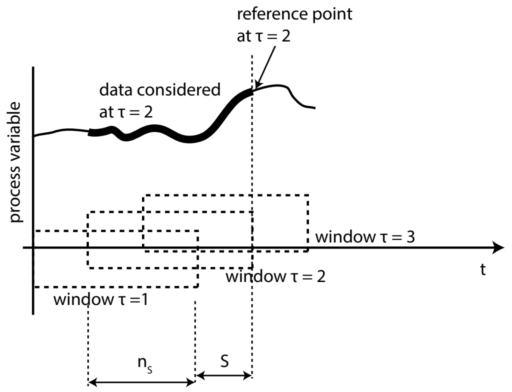

Fig. 3. Moving time data set window.

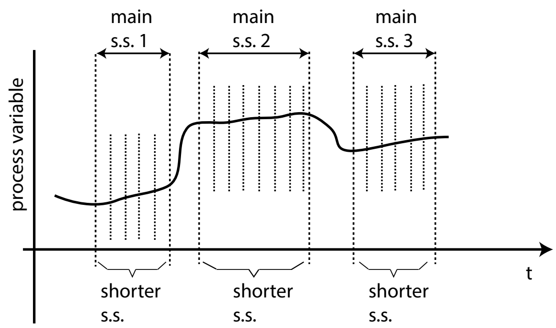

Fig. 4. Main and shorter steady-states of a process variable.

the variable takes the value 0. Finally, the steady-state of the system ($Y_{ss,system}$) is estimated from the equation below:

$$Y_{ss,system} = \prod_{i=1}^{N} Y_{ss,i} \tag{3}$$

It has been previously mentioned that some variables change slowly and a conventional steady-state detection method cannot capture these changes. For example the ambient temperature changes during the day at a slow rate, thus the power changes due to the changes in ambient temperature. The algorithm used in this paper is able (i) to identify when a steady-state episode is continuous, and (ii) when its length reaches an upper bound then the algorithm divides this original steady-state episode (main s.s) to shorter ones (shorter s.s.). The reason for this is to improve the quality of the measurements when data is noisy as will be seen in Section III. An illustrative example of the application of this method can be seen in Fig. 4. The algorithm identifies three main steady-states but as can be seen the process is changing over time in a slow rate. The algorithm divides these three main steady-state into shorter ones in which an average of the data is estimated. The average value represents the value of the steady-state of the data included in the short s.s.

The output of the steady-state detection algorithm is a matrix of data with N rows (variables) and T^* columns (number of final steady-states of the system) where $T^* \subseteq T$.

B. Steady-state detection index

The decision of the values of the bounds of the standard deviation is usually subjective. The smaller the values of these bounds are, the more conservatively the steady-state detection algorithm identifies steady-states. If these bounds are chosen very small then the algorithm may not detect any steady-state. On the other hand, if the bounds become significantly large, transients states will be considered as steady-state.

Assuming that there is a value of the bound of each variable i which describes its ideal steady-state, $h_{id,i}$ and that any other bound, which is greater than the ideal, is defined by h_i. The ratio between a value of any steady-state and the ideal case defines the Steady State Index (SSI) of variable i:

$$s_i = \frac{h_i}{h_{id,i}} \tag{4}$$

The s_i can take values greater than one where one represents the ideal steady-state. The greater the SSI is, the further the operation moves away from its ideal steady-state. When the steady-state of the system depends on more than one variable then the definition of the SSI of the system for two process variables is given by the following relationship:

$$s_{sys} = \frac{p_{ss,1} \cdot s_1 + p_{ss,2} \cdot s_2}{p_{ss,1} + p_{ss,2}} \tag{5}$$

where the s_1 and s_2 are SSIs of two process variables respectively and parameters $p_{ss,1}$ and $p_{ss,2}$ are weight factors. The relationship can be generalised when there are more than two variables involved.

C. Development of models

The top diagram in Fig. 5 illustrates a generic multi-stage compressor driven by an electrical motor. The development of a black box model of this compressor system includes: the multi-stage compressor, the cooling system and the gearbox. The power consumption of the motor (P_{el}) depends on the position of the Inlet Guide Vanes (IGVs), actuators which operators can control, (θ), the mass flow entering the compressor (m_a), the ambient conditions (T_{in}, p_{in}) and the pressure at the exit of the compressor (p_{out}). It was observed that the mass flow depends on the position of the IGVs and inlet conditions.

The lower panel in Fig. 5 shows the procedure for the development of the models of the operation of a multi-stage compressor. The output variable, power consumed from the motor of the compressor, can be used in an optimization framework such as the one presented in [11]. Figure 5 shows that two black boxes are used to predict first the mass flow rate and the power respectively.

The connection between inputs and outputs of a black box is described by a regression model. The inputs are variables which can be measured from the process, such as the inlet temperature of air, and the position of the IGVs. The output is the desired variable to be predicted, mass flow or power.

The first step is the steady-state identification. After this step, the processed data has to be normalised. This is because different variables have different units, for example power is measured in kW and pressure in bar.

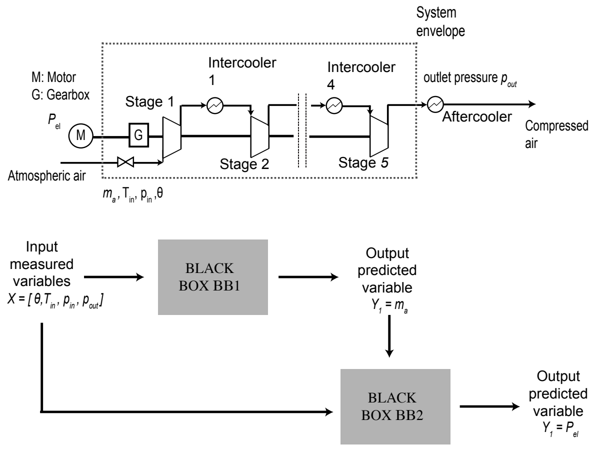

Fig. 5. Regression models of a multi-stage compressor.

The third step involves the division of the steady-state set of data into two smaller sets. The first set, the calibration set, is used to estimate the parameters of the regression model while the validation set examines if the generated model provides good predictions. The data set used to calibrate the model defines the domain of validity of the model, which is called regression domain. Extrapolation of the model outside its valid domain may give values which are not true. Brooks et al. [12] presented a study which characterizes the domain of regression models.

To develop the polynomial regression models, $X = x_k = [\theta, T_{in}, p_{in}, p_{out}]$ and $Y = y_l = [m_a, P_{el}]$ ($k \in K$ number of input variables and $l \in L$ number of output variables) are defined and the models are given by:

$$y_1^* = b_0 + b_1 \cdot x_1^* + b_2 \cdot x_2^* + b_3 \cdot x_3^* + b_4 \cdot x_1^* \cdot x_2^* + b_5 \cdot x_1^* \cdot x_3^* + b_6 \cdot x_2^* \cdot x_3^* + b_7 \cdot x_1^{*2} + b_8 \cdot x_2^{*2} + b_9 \cdot x_3^{*2}, \quad (6)$$

$$y_2^* = b_{10} + b_{11} \cdot x_1^* + b_{12} \cdot x_2^* + b_{13} \cdot x_3^* + b_{14} \cdot x_4^* + b_{15} \cdot y_1^* \quad (7)$$

where $y_1^* = y_1/y_{(1,max)}, y_2^* = y_2/y_{(2,max)}, x_1^* = x_1/x_{(1,max)}, x_2^* = x_1/x_{(2,max)}, x_3^* = x_3/x_{(3,max)}, x_4^* = x_4/x_{(4,max)}$ are the scaled variables of the regression model.

The $x_{(k,max)}, y_{(l,max)}$ are the maximum variables of their respective calibration and validation sets. The input variables and the output variable are scaled with the maximum value of each respective data set. The parameters of the models, $b_m, m = 1 \ldots 14$ are calculated with regression methods [13].

III. Case Study and Results

This section presents the results from the steady-state identification of the operation of industrial centrifugal multistage compressors. Section III-A gives the description of the case study to be examined and the demonstration of a step test of an industrial multi-stage compressor. The analysis of the step test is useful to inspect the process data and identify upper bounds for ideal steady-states of compressors with similar specifications. Section III-B presents the analysis of three

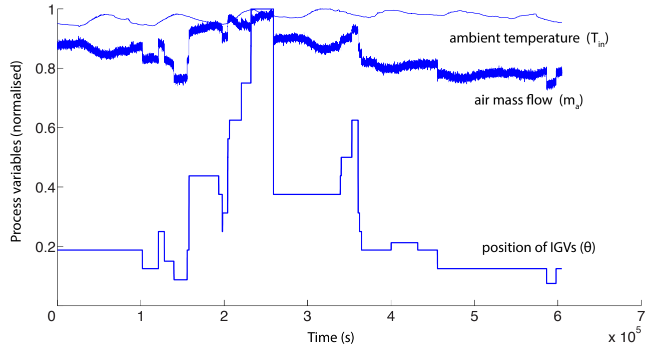

Fig. 6. Several process variables of the operation of an industrial centrifugal compressor over time.

industrial compressors working in parallel in normal operation. Finally the results are presented and discussed.

A. Case study and analysis of operation of an industrial compressor

The industrial case study is an air separation plant with air compressors. The air compressors of the plant provide compressed air to several air separation columns. The focus of the work is on three industrial multistage air compressors which operate in parallel providing compressed air to one separation column.

A fourth compressor similar to the other three is examined during a step test. In this test the compressor is disconnected from the downstream process and receives different constant set points of the position of the IGVs for defined time periods. The outlet conditions of the compressor are kept constant. These step tests took place for one week and data with 10 s sample interval, i.e 0.1 Hz sample rate, was collected. The investigation of the behaviour of the process variables of the operation of this compressor helps to inspect which data should be taken into account for the development of data-driven models and helps to identify upper bounds for the ideal steady-states and other bounds for non-ideal cases.

Several measurements can be acquired from the Distributed Control System (DCS), for example the mass flow at the exit of the compressor, the temperature and pressure of the environment and the power consumed by the motors. Figure 6 shows three process variables, position of the IGV (θ), mass flow (m_a) and ambient temperature (T_{in}), over time during the step tests. The values on the vertical axes are normalised. All of these variables influence the power consumption, which is plotted in normalised units in Fig. 7. Regarding to Fig. 6 the mass flow is considerably influenced by the θ. Indeed, the mass flow changes in a nonlinear way with the change of the θ. When the θ changes from 0.19 to 0.43 between $1 \cdot 10^5$ and $2 \cdot 10^5$ s, the mass flow changes significantly compared to smaller changes in the mass flow from larger changes of the θ between $2 \cdot 10^5$ and $3 \cdot 10^5$ s. Moreover, when θ is fixed, the mass flow changes at a slow rate following inversely the change in the ambient temperature.

The steady-state algorithm used a moving data set window with $n_s = 10$ and $S = 1$ for each variable. Two variables

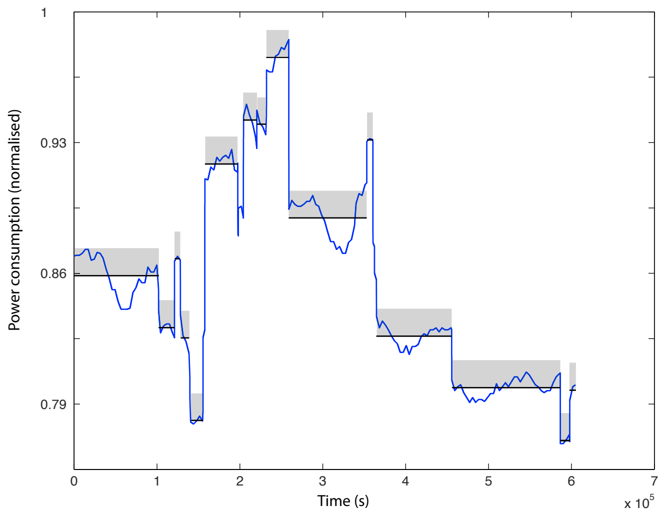

Fig. 7. steady-states of operation from the steady-state detection algorithm.

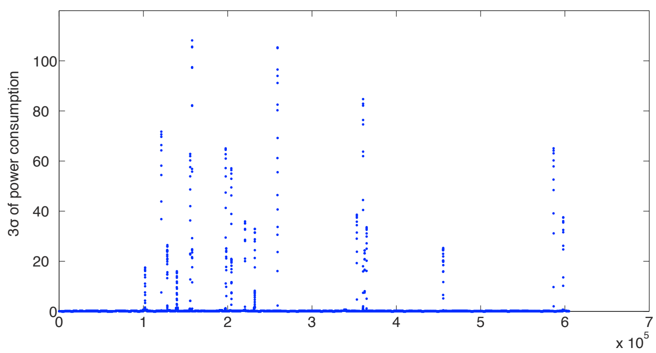

Fig. 8. 3σ of the power per data set window of the fourth compressor during step tests.

were taken into account in Eq. (3), power consumption P_{el} and θ. The algorithm used a minimum bound which defines the condition for ideal steady-state, $s_{sys} = 1$ and weight factors $p_{ss,1}$ and $p_{ss,2}$ equal to one in Eq. (5). Figure 7 shows the power consumption of the motor of the fourth compressor and the steady-states identified. Figure 7 shows that the algorithm captures all the episodes when the position of the IGVs is constant, however the power changes gradually from one state to another due to changes in ambient conditions.

Figure 8 shows the 3σ of every time window of the power consumption of the motor of compressor $j = 4$ over the whole time space of one week with sample rate 0.1 Hz. The values presented on the vertical axes in Fig. 7, 8 and 9 are scaled due to confidentiality restrictions. It is important to notice here that the condition for steady-state operation in the steady-state identification algorithm which has to hold true is $h_i < 3\sigma$. The values of the ideal steady-state indexes of P_{el} and θ were assumed $h_1 = 0.6$ and $h_2 = 0.035$ respectively. On the other hand, Fig. 9 shows the 3σ of the power consumption of compressor $j = 1$ when the compressor works in normal operation connected with the downstream process. As can be

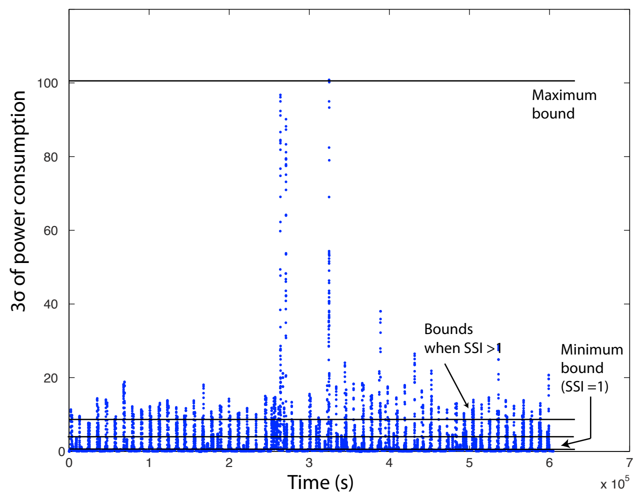

Fig. 9. 3σ of power per data set window of compressor 1.

seen in Fig. 9, the values of the 3σ of the power are more frequently above the bound of $s_1 = 1$ than in those in the case of the step test of compressor $j = 4$. This observation is explained from the fact that the disturbances coming from the downstream operation cause these transients.

B. Analysis of three industrial compressors in parallel supplying compressed air a process

The previous analysis helps to determine values of the ideal steady-state and various steady-state indices (see Table I) to examine how much the processed data influence the accuracy of the models of compressors $j = 1, 2$ and 3. The data is collected from the operation for one week with the same sample rate as in the case of compressor 4. The flow and power measurements of compressor 2 were noisy compared to the other measurements of the other compressors. The accuracy of a model was investigated using the steady-state detection methodology with shorter steady-states and without. The processed data used to estimate the mass flow error (Flow error) in the black box BB1, the error when estimating the power with BB2 using the estimated flow from BB1 (Power error 1) and the error when estimating power with BB2 using the flow measurement (Power error 2). The error is given from the calculation of the Root Mean Square Error (RMSE) of the validation of the fitted model divided by the mean value of the validation data set.

The results can be seen in Table I. The $h_{id,1}$ and $h_{id,2}$ used for the description of the ideal SSI of compressor 1 and 3 are chosen similarly to those of compressor 4. Due to the noise in the measurements of flow and power of compressor 2, the chosen values of its bounds were assumed higher. Moreover, the maximum upper bounds were chosen equal to the maximum 3σ (another option was to choose the 3σ of the 3σ of each variable). The overview of the results shows that the closer the index SSI is to the ideal case the more accurate models are the developed models. By comparing Case 1 (ideal) with the Case 4 (including all transitions) of Power error 2 of compressor 1 the improvement is 22.4% (considering shorter s.s algorithm). The algorithm which estimates the

TABLE I. SSIs AND RMSE ERRORS IN THE VALIDATION OF THE MODELS OF THREE DIFFERENT COMPRESSORS IN FOUR CASES

Compressor 1	**Case 1 (ideal)**	**Case 2**	**Case 3**	**Case 4**
SSI (σ_{sys})	1	1.83	7.17	211.6
h_1	0.6	1.6	8	219.8
h_2	0.035	0.035	0.035	1.99
Flow error	0.78	0.79	0.82	0.81
Power error 1	0.64	0.72	0.77	0.77
Power error 2	0.46	0.57	0.57	0.59
Compressor 2	**Case 1 (ideal)**	**Case 2**	**Case 3**	**Case 4**
SSI (σ_{sys})	1	1.06	1.17	23.63
h_1	36	40	48	117.4
h_2	0.05	0.05	0.05	2.2
Flow error	0.47	0.50	0.49	0.49
Power error 1	0.41	0.40	0.42	0.41
Power error 2	0.35	0.34	0.35	0.34
Compressor 3	**Case 1 (ideal)**	**Case 2**	**Case 3**	**Case 4**
SSI (σ_{sys})	1	1.83	7.17	150.50
h_1	0.6	1.6	8	144.6
h_2	0.035	0.035	0.035	2.1
Flow error	2.06	1.97	2.01	2.14
Power error 1	1.41	1.30	1.44	1.52
Power error 2	0.78	0.77	0.82	0.83
Without shorter s.s				
Compressor 1	**Case 1 (ideal)**	**Case 2**	**Case 3**	**Case 4**
Flow error	0.83	0.85	0.87	0.90
Power error 1	0.66	0.76	0.77	0.78
Power error 2	0.51	0.62	0.63	0.63
Compressor 2	**Case 1 (ideal)**	**Case 2**	**Case 3**	**Case 4**
Flow error	0.59	0.59	0.60	0.60
Power error 1	0.72	0.73	0.74	0.75
Power error 2	0.70	0.72	0.72	0.73
Compressor 3	**Case 1 (ideal)**	**Case 2**	**Case 3**	**Case 4**
Flow error	2.04	2.01	2.06	2.08
Power error 1	1.40	1.33	1.45	1.45
Power error 2	0.76	0.77	0.82	0.82

shorter steady-states helps to improve significantly the results in the case of compressor 2. The improvement in the error of Power error 2 in the ideal steady-state is 50%. This is because when the shorter steady-states are estimated the data of each shorter steady-state is averaged. Moreover, the results show also improvement in the accuracy of the models in the case of compressor 3, particularly in the power estimated directly from the mass flow measurements.

The results show that the steady-state algorithm effectively identifies the steady-states of the operation of compressors which operate connected to a downstream process. The recommendation is to use parameters estimated from step tests of each compressor. The estimation of parameters depends on the quality of the measurements (regarding to the signal noise they have), in the case of compressor 2 different parameters has to be used than in the case of the other compressors 1 and 3. Moreover the feature of the estimation of shorter steady-states improves the accuracy when a measurement is noisy. The processed data coming from the steady-state identification of these parameters is suitable for use in the optimization framework in [11] to find the best distribution of load. This is because the errors on the data-driven models used in the optimization are smaller than the standard deviation of the power measurement itself, hence it will be possible to reduce the uncertainty in the optimization and establish a more robust optimization framework for achieving energy savings.

IV. Conclusions

The paper presented an analysis and investigation of the steady-state identification of process data to develop data-driven compressor models for optimizing compressor stations. The structure of the optimization employs steady-state data-driven models to represent the operation in steady-state. Many researchers reported that the identification of steady-states of the data plays a key role for accurate representation between data-driven models and actual process. The paper described the development of a steady-state algorithm which highlighted the influence of the steady-state process in the model accuracy.

Acknowledgment

Financial support from the Marie Curie FP7-ITN project "Energy savings from smart operation of electrical, process and mechanical equipment - ENERGY SMARTOPS", Contract No: PITN-GA-2010-264940 is gratefully acknowledged. The authors would like to thank BASF SE for providing a case study and technical support.

References

[1] R. Saidur, N.A. Rahim, M. Hasanuzzaman, 2010, A review on compressed-air energy use and energy savings, Renewable and Sustainable Energy Reviews, Vol. 14, pp. 1135-1153.

[2] S. Mitra, I.E. Grossmann, J.M. Pinto, N. Arora, 2012, Optimal production planning under time-sensitive electricity prices for continuous power-intensive processes, Computers and Chemical Engineering, Vol. 38, pp. 171-184.

[3] F. Paparella, L. Domínguez, A. Cortinovis, M. Mercangöz, D. Pareschi, S. Bittanti, 2013, Load sharing optimization of parallel compressors, European Control Conference (ECC), July 17-19, 2013, Zürich, Switzerland.

[4] M. Abbaspour, K.S. Chapman, P. Krishnaswami, 2005, Nonisothermal compressor station optimisation, Journal of Energy Resources Technology, Transactions of the ASME, Vol. 127, pp. 131-141.

[5] I.-S. Han, C. Han, C.-B., Chung, 2004, Optimisation of the air-and gas-supply network of a chemical plant,Trans IChemE, Part A, Chemical Engineering Research and Design, Vol. 82 (A10), pp. 1337-1343.

[6] S. Wu, R.Z. Rios-Mercado, E.A. Boydm, L.R. Scott, 2000, Model relaxations for the fuel cost minimization of steady-state gas pipeline networks, Mathematical and Computer Modelling, Vol. 31, pp. 197-220.

[7] F.C.G. DeMarco, G.P. Elias, 2011, Fuel consumption model on natural gas compression stations driven by two-shaft gas turbine, PSIG Annual Meeting, Napa Valley, California.

[8] S. Cao, R.R. Rhinehart 1995, An efficient method for on-line identification of steady state, Journal of Process Control, Vol. 5, No. 6, pp. 363-374.

[9] M. Kim, S.H. Yoon, P.A. Domanski, W.V. Payne, 2008, Design of a steady-state detector for fault detection and diagnosis of a residential air conditioner, International Journal of Refrigeration, Vol. 31, pp. 790-799.

[10] M. Mansour, J.E. Ellis, 2008, Methodology of on-line optimisation applied to a chemical reactor, Applied Mathematical Modelling, Vol. 32, pp. 170-184.

[11] D.P. Xenos, M. Cicciotti, A.E.F. Bouaswaig, N.F. Thornhill, R. Martinez-Botas, 2014, Modeling and optimization of industrial centrifugal compressor stations employing data-driven methods, Proceedings of ASME Turbo Expo 2014: Turbine Technical Conference and Exposition GT2014, June 16-20, 2014, Düsseldorf, Germany.

[12] D.G. Brooks, SS. Carroll, W.A. Verdini, 1998, Characterizing the domain of a regression model, The American Statician, Vol. 42, No.3, pp. 187-190.

[13] R. Rosipal, and N. Kramer, 2006, Overview and recent advances in partial least squares, Subspace, Latent Structure and Feature Selection: Statistical and Optimization Perspectives Workshop (SLSFS 2005), Revised Selected Papers (Lecture Notes in Computer Science 3940). Berlin, Germany: Springer-Verlag, pp. 3451.