An hybrid ensemble based approach for process parameter estimation in offshore oil platforms

Piero Baraldi[a,], Enrico Zio[b,a], Francesca Mangili[a], Giulio Gola[c,d], Bent H. Nystad[c],*
[a]Politecnico di Milano, Italy
[b]Ecole Centrale Paris-Supelec, France
[c]Institutt for energiteknikk, OECD Halden Reactor Project, Norway
[d]IO Center for Integrated Operations, Trondheim, Norway
**piero.baraldi@polimi.it*

Abstract

The valve flow coefficient is commonly used as a parameter to assess the erosion state of choke valves in offshore oil platforms. In particular, the difference between the theoretical value of the valve flow coefficient and its actual value calculated during operation is retained as the valve health indicator. The actual valve flow coefficient is analytically calculated from the oil, water and gas mass flow rates. These quantities, which are allocated on a daily basis based on the measured total production from a number of wells, on physical parameters (pressures and temperatures) related to the specific well, and on a physical model of the process, can be affected by large uncertainties. Based on such values, the evaluation of the health indicator becomes unreliable and undermines the possibility of using it for prognostic purposes. Similar situations arise every time health monitoring rely on unreliable measurement taken by sensors subject to hard working condition, as often happen in the nuclear industry. This paper proposes a method to obtain more accurate daily estimates of the actual values of the oil, water and gas flow rates, from which improved estimates of the flow coefficient will follow. In this respect, an hybrid ensemble aggregating the physical model with data-driven models built using the Kernel Regression (KR) method has been used. Ensemble diversity is ensured by using different training sets;a local procedure based on the historical performance of the models is adopted to aggregate their predictions. The method is verified on real measurements performed on a number of similar offshore choke valves.

1. Introduction

In this paper, we consider the degradation of choke valves located topside at wells on the Norwegian Continental Shelf [1] [2]. The difference between the actual valve flow coefficient and its theoretical value is retained as the indicator of the choke valve health state and is used to assess the degree of erosion affecting the choke. While the theoretical value of the valve flow coefficient depends only on the choke opening, the actual valve flow coefficient is analytically calculated on a daily basis as a function of the pressure drop through the choke which is directly measured and oil, gas and water flow rates which are allocated based on the measured total production from a number of wells and on physical parameters (pressures and temperatures) related to the single well. Such flow rates are actually measured only during a number of well tests carried out throughout the valve life.

In practice, the resulting indicator of the choke valve state is very noisy and lacks the physical monotonicity of the erosion process. In [3] it has been shown that the allocated values of oil, gas and water flow rates can cause large inaccuracies and uncertainties in the calculation of the actual valve flow coefficient. In this work, an ensemble of Kernel Regression (KR) models has been devised to correct these values based on the relations among all parameters and aggregated to the outcome of the physical model to increase its accuracy. KR is a distance-based regression algorithm [4] [5]; an ensemble of multiple KR models is used to avoid the need of selecting the optimal model and to increase the robustness and reduce the uncertainty of the estimate [6]. Diversity is injected in the ensemble by differentiating the training procedure for each KR model. The aggregation of the KR model outcomes is obtained based on the models performance on historical data closed to the test data under reconstruction. This approach can be found in literature under the name of local fusion (Barald, 2010).

The paper is framed as follows. The traditional procedure for the construction of a health indicator assessing the choke valve erosion state is presented in Section 2; in Section 3, an ensemble of KR models is proposed to improve the accuracy of the allocated flow rates; Section 4 shows the results of the application of the method to the choke dataset; finally, conclusion and potential perspectives for future work are drawn in Section 5.

2. Choke valve erosion assessment

In oil and gas industries, choke valves are normally located on top of each well and are used to balance the pressure on several wells into a common manifold to control flow rates and protect the equipment from unusual pressure fluctuations.

In Figure 1, a choke valve is sketched. The throttle mechanism consists of two circular disks, each with a pair of circular openings to create variable flow areas. One of the disks is fixed in the valve body, whereas the other is rotated either by manual operation or by actuator, to vary or close the opening. For large pressure drops, the well streams which contain gas, liquid and sand particles can reach 400-500 m/s and produce heavy metal loss mainly due to solids, liquid droplets, cavitation and combined mechanisms of erosion-corrosion, resulting in choke lifetimes of less than a year. Erosion management is vital to avoid failures that may result in loss of containment, production being held back, and increased maintenance costs. Moreover, several chokes are located subsea, where the replacement cost is high. Then, the need has increased for reliable models to estimate erosion and lifetime of choke valves, in order to allow implementing effective maintenance strategies [8] [9] [10].

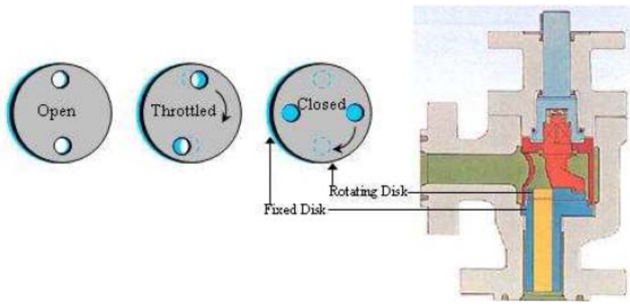

Figure 1. Typical choke valve of rotating disk type (http://www.vonkchokes.nl/).

A common indicator of the valve flow capacity is the flow coefficient C_V, which is related to the effective flow cross-section of the valve.

For a specific valve opening, erosion produces a gradual increase of the valve area available for the flow transit, thus determining an increase of C_V (eq. 1). For this reason, knowing the value of the flow coefficient is fundamental for assessing the health state of the choke. During operation, C_V is not directly measured but computed for a two-phase flow as [11]:

$$C_V = \frac{\dot{m}}{27.3 \cdot F_p \sqrt{\Delta P \left(\frac{f_g}{\rho_g \cdot J^2} + \frac{f_w}{\rho_w} + \frac{f_o}{\rho_o} \right)}} \quad (1)$$

where $\dot{m} = \dot{m}_o + \dot{m}_w + \dot{m}_g$ is the total mass flow rate of the oil-water-gas mixture, $f_{o,w,g} = \dot{m}_{o,w,g} / \dot{m}$ is the fraction of the oil, water and gas mass flow rates, respectively, $\rho_{o,w,g}$ are the corresponding densities, J is the gas expansion factor, $F_p(\theta)$ is the piping geometry factor accounting for the geometry of the valve/pipe reducer assembly and ΔP is the pressure drop through the choke. Eq. (1) and the values of $\rho_{o,w,g}$, J, $F_p(\theta)$ and N_6 are derived from fluid dynamics; parameters ΔP, θ, $\dot{m}_o$, $\dot{m}_w$ and $\dot{m}_g$ are measured or allocated during operation.

2.1 Choke valve dataset

For a correct assessment of the choke erosion state, it is fundamental to obtain frequent and reliable measurements or estimates of the parameters ΔP, θ, $\dot{m}_o$, $\dot{m}_w$ and $\dot{m}_g$ used to compute the flow coefficient C_V. Nevertheless, only the pressure drop ΔP and the valve opening θ are measured during standard daily inspections (SI), whereas measures of water, oil and gas flows rates are taken downstream of the choke only during well tests (WT) with a multiphase flow separator. On a daily basis, the values of $\dot{m}_o$, $\dot{m}_w$ and $\dot{m}_g$ are allocated for a single well by a software based on the measured total production from a number of wells and on physical parameters (pressures and temperatures) related to the specific well. In this work, we consider data regarding valve degradation collected on a daily basis from five different wells. Table 1 and 2 outlines the available information: the daily allocated values of $\dot{m}_o$, $\dot{m}_w$ and $\dot{m}_g$, the daily measured value of ΔP and θ and the $\dot{m}_o$, $\dot{m}_w$ and $\dot{m}_g$ real values measured during well tests. Since degraded valves are replaced, data collected for a single well refer to different chokes.

Table 1. Available information

	Standard Inspections (SI)	**Well Test Inspections (WT)**
ΔP and θ	Measured	Measured
$\dot{m}_o$, $\dot{m}_w$ and $\dot{m}_g$	Allocated	Measured

Table 2. Number of SI and WT patterns for each choke.

Well	N_{SI}	N_{WT}	***Nval***
1	1854	87	68
2	2143	96	59
3	657	39	20
4	1859	96	54
5	1678	71	36

3. Improving the quality of the allocated parameters

Since the allocated values of $\dot{m}_o$, $\dot{m}_w$, and $\dot{m}_g$ are noisy and unreliable [3], an on-line procedure for improving the accuracy of the estimates of those parameters values is here proposed. The procedure is based on empirical models which learn from a training set the relationships between the parameters, and provides as output an estimate $\hat{\mathbf{x}}$ of the input parameters $\mathbf{x}$. Different regression techniques such as those based on the use of principal component analysis [12], artificial neural networks [13] [14], support vector machines [15], evolving clustering methods [16] have been applied to this purpose. In this work, Kernel Regression models [4] [5] have been chosen.

3.1 Kernel regression models

Nonparametric Kernel Regression (KR) is used to build a model for improving the quality of the allocated values of oil, water and gas mass flow rates. Compared with parametric methods, which are defined by sets of parameters and predefined functional relationships, nonparametric methods have the advantage that they do not require any assumption about the mathematical structure of the regression model [4].

KR models provide estimates by developing local models in the neighborhoods of the test patterns they are fed with. Estimates are obtained as weighted averages of the training patterns, with weights decreasing as the distance between the test and the training pattern increases. In this view, training patterns closer to the test pattern are conjectured to be more similar to it, thus giving the most relevant contribution to its estimate.

Let $\mathbf{X}^{trn}=\{\mathbf{x}_k\}$, $k=1,\dots, N_{trn}$ be the training set used for the estimate of the test pattern $\mathbf{x}_{tst}$. To develop the KR model, parameters are divided into a predictor group (PG) and a response group (RG) (with the two groups possibly overlapping). For the estimate of $\mathbf{x}_{tst}$, the KR algorithm assigns to each training pattern $\mathbf{x}_k$ a weight $w_k=K[d_{PG}(\mathbf{x}_{tst},\mathbf{x}_k)]$, where K is the kernel function which produces the weight for a given distance $d_{PG}(\mathbf{x}_{tst},\mathbf{x}_k)$, between the training and the test patterns, computed considering only the parameters of the predictor group. The estimate $\hat{\mathbf{x}}_{tst}^{RG}$ of the RG parameters of the test patterns is obtained as a weighted average of the RG parameters of the training patterns:

$$\hat{\mathbf{x}}_{tst}^{RG} = \frac{\sum_{k=1}^{N_{trn}} w_k \cdot \mathbf{x}_k^{RG}}{\sum_{k=1}^{N_{trn}} w_k} \tag{2}$$

The kernel function K must be such that training patterns with small distances from the test pattern are assigned large weights and vice versa. Among the several functions which satisfy this criterion, the Gaussian kernel is commonly used [17]:

$$K(d_{PG}) = \frac{1}{\sqrt{2\pi h^2}} \exp\left(-\frac{{d_{PG}}^2}{2h^2} \right) \tag{3}$$

where the parameter h defines the kernel bandwidth and is used to control how close training patterns must be to the test pattern to be assigned a large weight. In order to compute d_{PG}, the PG parameters are normalized to mean equal to 0 and standard deviation equal to 1.

In the present case study, the choice of training dataset and predictor parameters is critical. In this respect, four different models can be devised by differentiating the training set as listed in Table 2.

Table 3. Model training procedures.

Model	Training set	Predictor parameters
1	Well test data X^{WT}	Measured $\mathbf{x}_k^{PG} = [\Delta P, \theta]$
2	Standard inspections data X^{SI}	Measured $\mathbf{x}_k^{PG} = [\Delta P, \theta]$
3	Well test data X^{WT}	Measured & allocated $\mathbf{x}_k^{PG} = [\Delta P, \theta, \dot{m}_o, \dot{m}_w, \dot{m}_g]$
4	Standard inspections data X^{SI}	Measured & allocated $\mathbf{x}_k^{PG} = [\Delta P, \theta, \dot{m}_o, \dot{m}_w, \dot{m}_g]$

In any case, for estimating the test pattern $\mathbf{x}_{tst}$ collected for a specific choke, only patterns concerning the same choke are considered. This is done because, although valves are similar, the well behavior, and thus the relation among the observed parameters, tends to vary with time and from one well to another. Thus, data collected from other chokes do not provide useful information about the behavior of the choke under study.

In practice, when estimating the test pattern $\mathbf{x}_k$ for a specific choke c, only the patterns $\mathbf{x}_j$, j=1,…,k-1 previously collected during the life of the c-th choke are used as training patterns.

In order to verify the performance of the process parameter estimation models, we have considered a test set formed by the N_{WT} patterns of $\mathbf{X}^{\mathrm{SI}}$ collected the same day of a well test. Since an accurate measure of the process parameters under estimation is available for these patterns, they can be used to assess the performance of the estimation models.

The response group is formed by the unreliable parameters that need to be estimated $\mathbf{x}_k^{RG} = [\dot{m}_o, \dot{m}_w, \dot{m}_g]$.

3.2 Ensemble approach

Since the performance of the models depends on the characteristics of the parameter to be estimated and the intensity of the noise, it is difficult to identify a single best model [3].

Using an ensemble of models allows overcoming this dilemma. Indeed, the general idea underlying ensembles is to create many models and combine their outputs in order to achieve a performance better than that provided by each individual model in the ensemble [6]. Models' prediction diversity plays a fundamental role when ensemble approaches are devised. In fact, individual models committing diverse errors can be opportunely combined in such a way that the error of the aggregated prediction is smaller than the error of any of the individual models.

Different techniques for the aggregation of the outcomes of individual models have been proposed in literature, the most common being statistics methods like the simple mean, the median and the trimmed mean [12] [18]. Other aggregation techniques, which allow improving the ensemble performance, consider weighted averages of the model outcomes with weights proportional to the performance of the individual models. In this respect, both global approaches (in which the performance is computed on all the available patterns) and local approaches (which measure the performance only on the patterns closed to the test pattern) have been proposed [19]. Since in the choke valve case study a complete input-output set of patterns is not available, model weighting cannot be based on a measure of the performance of the individual model. For this reason, a new strategy is here proposed based on the use of the Analytic Hierarchy Process (AHP) [7].

3.3 Outcome aggregation with Analytic Hierarchy Process

AHP is used to assign performance weights to the models of the ensemble. The procedure allows ranking different models outcomes using relative performance measurements, without resorting to an absolute measurement of the model performance. AHP is a multi-criteria decision method that uses hierarchic structures to represent a decision problem and provides ranking of different choices [7]. It consists of two main steps: 1) structuring a hierarchy; 2) assigning priorities to the elements of each hierarchy level by comparative judgments of the elements based on a pre-defined scale.

In this application, the hierarchy structure sketched in Figure 2 is used. The four models on level 3 are compared with respect to the two criteria Z_1 and Z_2 of the level 2 towards the goal (level 1) of obtaining high model accuracy.

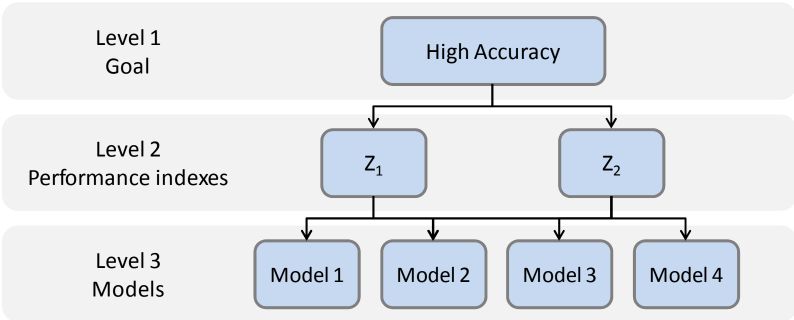

Figure 2. Model weighting hierarchy structure.

The basic tool for assigning priorities to the elements of a level of the hierarchy are matrices of pairwise comparisons based on the criteria defined at the previous level. For the hierarchy of Figure 2, two matrices of comparisons $\mathbf{A}_{Z_1}$ and $\mathbf{A}_{Z_2}$ have to be defined, each one containing elements a_{ij} representing the relative importance of model i when compared to model j based, respectively, on criteria Z_1 and Z_2.

Once a matrix of comparisons $\mathbf{A}_{Z_l}$ is defined, the vector of priorities $\boldsymbol{\pi}_{Z_l}$ of the models in level 3 of the hierarchy with respect to criterion Z_l is given by the eigenvector associated to the maximum eigenvalue of matrix $\mathbf{A}_{Z_l}$. The priority vectors obtained for each criterion are weighted with the priority assigned to the corresponding criterion and averaged to obtain the overall priority vector $\boldsymbol{\pi}=[\pi_1, \pi_2, \pi_3, \pi_4]$ assigning the priority π_m to model m.

In the proposed aggregation method, the priorities assigned to each model are used as weights to aggregate the models' outcomes through a weighted average:

$$\hat{\mathbf{x}}_{tst}^{RG} = \frac{\sum_{m=1}^{4} \pi_m \hat{\mathbf{x}}_{tst}^{RG,m}}{\sum_{m=1}^{4} \pi_m} \tag{4}$$

where $\hat{\mathbf{x}}_{tst}^{RG,m}$ is the estimate provided by model m of the response group parameters in $\mathbf{x}_{tst}^{RG}$.

In this application, the first criterion Z_1 chosen to evaluate the relative importance $a_{ij}(\mathbf{x}_{tst})$ of model i with respect to model j in the reconstruction of a test pattern $\mathbf{x}_{tst}$ is the relative similarity of the two models outcomes $\hat{\mathbf{x}}_{tst}^{RG,i}$ and $\hat{\mathbf{x}}_{tst}^{RG,j}$ to the remaining models outcome $\hat{\mathbf{x}}_{tst}^{RG,m}$, $m \neq i, j$. Assuming that the model outcomes of the models left out of the pair-wise comparison are distributed around the correct value, this criterion assigns larger weights to the model (i or j) whose outcome is more similar to that of the models left out.

The similarity of two patterns $\hat{\mathbf{x}}_{tst}^{RG,i}$ and $\hat{\mathbf{x}}_{tst}^{RG,m}$ has been estimated by the inverse of their Euclidean distance $d(\hat{\mathbf{x}}_{tst}^{RG,i}, \hat{\mathbf{x}}_{tst}^{RG,m})$; the relative importance $a_{ij}^{m}(\mathbf{x}_{tst})$ of a model i with respect to model j when model m is taken as reference is defined by:

$$a_{ij}^{m}(\mathbf{x}_{tst}) = \frac{d(\hat{\mathbf{x}}_{tst}^{RG,j}, \hat{\mathbf{x}}_{tst}^{RG,m})}{d(\hat{\mathbf{x}}_{tst}^{RG,i}, \hat{\mathbf{x}}_{tst}^{RG,m})} \tag{5}$$

and the entry a_{ij} of the comparison matrix $\mathbf{A}_{Z_1}$ is given by the product of the relative importance values $a_{ij}^{m}(\mathbf{x}_{tst})$ m=1,…,4, m≠i, j:

$$a_{ij} = \prod_{m \neq i,j} a_{ij}^{m}(\mathbf{x}_{tst}) \tag{6}$$

According to the AHP method, the quality of a matrix of comparison can be evaluated considering its consistency. Matrix $\mathbf{A}_{Z_1}$ is consistent if the following equation is satisfied for any i, j and k [7]:

$$a_{ij} a_{jk} = \frac{\pi_i}{\pi_j} \frac{\pi_j}{\pi_k} = \frac{\pi_i}{\pi_k} = a_{ik} \tag{7}$$

In our case, substituting eqs. (5) and (6) in eq. (7) gives:

$$a_{ij} a_{jk} = \prod_{m \neq i,j} \frac{d^{jm}}{d^{im}} \prod_{m \neq j,k} \frac{d^{km}}{d^{jm}} = \left[\frac{d^{jk}}{d^{ik}} \prod_{m \neq i,j,k} \frac{d^{jm}}{d^{im}} \right] \left[\frac{d^{ki}}{d^{ji}} \prod_{m \neq j,k,i} \frac{d^{km}}{d^{jm}} \right] = \frac{d^{jk}}{d^{ji}} \prod_{m \neq i,j,k} \frac{d^{km}}{d^{im}} = \prod_{h \neq i,k} \frac{d^{km}}{d^{im}} = a_{ik}$$

where $d^{ij} = d(\hat{\mathbf{x}}_{tst}^{RG,i}, \hat{\mathbf{x}}_{tst}^{RG,j})$ and, by definition, $d^{ij} = d^{ji}$. This shows that, in the proposed approach, matrix $\mathbf{A}_{Z_1}$ is consistent.

A second criterion Z_2 for evaluating the performance of a model takes into account the RMSE in reconstructing the reliable parameters ΔP and θ, i.e. the root mean square difference between the reconstructed and measured values. This second criterion takes into account the fact that robust and reliable models should be able to correctly reconstruct the reliable parameters despite the noise on the mass flow rates.

Since all model performances are evaluated with respect to the same reference, i.e., the reliable measurements of ΔP and θ, the pair-wise comparison is not needed, and the vector of priorities $\boldsymbol{\pi}_{Z_2}$ is computed by taking for each model h=1,…,4, the inverse of its RMSE, i.e., $\pi_{Z_2}^{m} = 1/RMSE^{m}$.

Finally, the two criterions Z_1 and Z_2 of level 2 of the hierarchy are given the same importance and thus the priority vector $\boldsymbol{\pi}$ is given by:

$$\boldsymbol{\pi} = [0.5 \quad 0.5] \begin{bmatrix} \boldsymbol{\pi}_{Z_1} \\ \boldsymbol{\pi}_{Z_2} \end{bmatrix} \tag{8}$$

4. Results

The KR models in Table 2 are applied to the choke valve case study to obtain estimates of the mass flow rates $\dot{m}_o$, $\dot{m}_w$ and $\dot{m}_g$. The performance of the models is evaluated by considering the mean square error (MSE) between the estimates of the mass flow rates $\dot{m}_o$, $\dot{m}_w$ and $\dot{m}_g$ and the corresponding well test measurements.

Table 4 compares the MSE of the four models and that which is obtained by considering the allocations as mass flow rate estimates. Notice that only model 2 allows achieving estimates more accurate than the allocations. Furthermore, using all five parameters as predictor parameters produces, in general, best results.

Table 4. Comparison of the performance of the SI allocations with the KR models estimates.

	SI allocations	Model 1	Model 2	Model 3	Model 4
$\dot{m}_o$	0.1205	0.0925	0.0794	0.1086	0.1267
$\dot{m}_w$	0.1547	0.1967	0.1445	0.2658	0.2060
$\dot{m}_a$	0.1993	0.2875	0.2068	0.3087	0.2096
Average	**0.1582**	**0.1923**	**0.1436**	**0.2277**	**0.1808**

Table 5 compares the performance of the allocations with respect to those obtained by three different AHP ensembles: a first ensemble aggregating all four KR models in Table 4, a second ensemble considering only the three best performing models (i.e., models 1, 2 and 4) and a third one considering both the allocations and the three best performing models.

Table 5. Comparison of the performance of the SI allocations with those of three different ensemble estimates

	Allocations	Ensemble 1	Ensemble 2	Ensemble 3
$\dot{m}_o$	0.1205	0.0672	0.0646	0.0647
$\dot{m}_w$	0.1547	0.1666	0.1302	0.1271
$\dot{m}_a$	0.1993	0.2103	0.1915	0.1862
Average	0.1582	0.1480	0.1288	0.1260

.

Results show that all the three AHP-based ensembles outperform the allocations; moreover, ensembles 2 and 3 outperform all single models.

In Figure 3, the MSE of the allocations and of Ensemble 3 are compared for every choke valve.

The obtained results show that the ensemble estimates do not always outperform the allocation. This can be due to two main reasons:

1. for some choke valve the allocations are highly accurate, and thus they cannot be improved through the KR estimates;
2. in some cases characterized by abrupt changes in the operating conditions of the choke (e.g., large variations of the choke opening), the pattern used for the model training may not cover the range of parameter values of interest for the test pattern.

To this purpose, future research work should be devoted to the a priori identification of those cases in which the KR ensemble will not outperform the allocations. In particular, with respect to the situation at point 1) one can have an a-priori indiction of the allocations and the KR ensemble accuracy by considering the corresponding MSE obtained in the estimates of the previous well test performed on the same choke, whereas with respect to the situation at point 2) an estimate of the confidence of the ensemble outcome should be provided by analyzing the position of the test pattern with respect to the distribution of the training patterns.

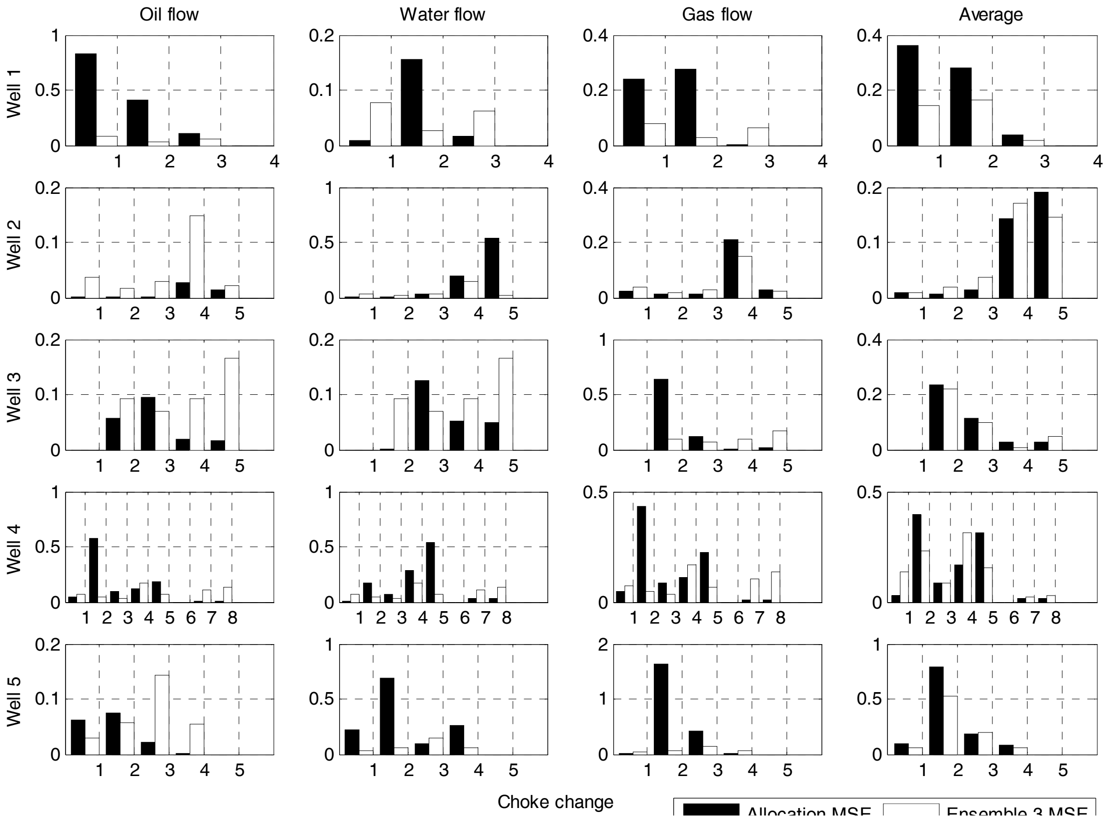

Figure 3. Comparison of the allocation and estimation performance for each choke.

5. Conclusions

In this paper, we have tackled the problem of improving the quality of the estimates of some process parameters used in offshore oil platforms for assessing the health state of choke valves in which undergoes erosion. To this purpose, we have proposed a method which estimates the allocated process parameters based on available measurements of other parameters, which are conjectured to be reliable, and on few reliable measurements of the process parameters collected during a number of well tests performed throughout the valve life.

The method is based on an ensemble of Kernel Regression models trained using different procedures in order to inject diversity into the models ensemble. To aggregate the outcomes of the individual models, an original technique based on the Analytic Hierarchy Process (AHP) method has been used. The results obtained on a number of similar eroding choke valves have confirmed the improved performances of the ensemble with respect to any of the single KR models, allowing significant improvement of the oil, water and gas mass flow rates estimates.

It has been shown, however, that in some cases the allocated value of the uncertain parameters are more accurate than those estimated by the proposed KR models ensemble; two main situations which causes this have been identified and a possible strategy for the improvement of the flow rate estimates, based on the identification of these situations, has been briefly sketched. The development of a method for the implementation of this strategy will be the object of future research work.

Althoght the methodology has been developed in the oil and gas context, a general application of the proposed approach is envisioned in situations in which unreliable parameters' measurements can be improved by resorting to a set of reliable parameters. The goal of the research is the application of such methodology to obtain reliable component health state indicators and prognostic models for remaining useful life estimation also in the nuclear energy field where diagnostics and prognostics are critical issues.

6. Acknowledgments

The authors wish to thank Erling Lunde and Morten Løes at Statoil ASA for proving us with the operational choke valve data and the IO Center for Integrated Operations in the Petroleum Industry (www.ntnu.no/iocenter) for partly funding this research project.

7. References

[1] B.H. Nystad, G. Gola, J.E. Hulsund, D. Roverso. Technical Condition Assessment and Remaining Useful Life Estimation of Choke Valves subject to Erosion. Proc. PHM Society 2010 Ann. Conf., October 11-13, 2010, Portland, OR.

[2] K. Haugen, O. Kvernvold, A. Ronold, R. Sandberg. Sand Erosion of Wear Resistant Materials: Erosion in Choke Valves. Wear, Vol. 186-187, Part 1, pp. 179-188, 1995.

[3] P. Baraldi, E. Zio, F. Mangili, G. Gola, B.H. Nystad. Ensemble of kernel regression models for assessing the health state of choke valves in offshore oil platforms. Submitted for pubblication to SPE journal, 2011.

[4] E.A. Nadaraya. On Estimating Regression. Theory of Probability and Its Applications, Vol. 10, pp. 186–190, 1964.

[5] C.G. Atkeson, A.W. Moore, S. Schaal. Locally Weighted Learning. Artificial Intelligence Review, Vol. 11, pp. 11–73, 1997.

[6] M.P. Perrone, L.N. Cooper. When networks disagree: ensemble methods for hybrid neural networks, National Science Fundation, USA, 1992.

[7] T.L. Saaty. The analytic Hierarchy Process, Planning, Priority Setting, Resource Allocation. McGraw-Hill, New York, 1980.

[8] K. Wold, S. Hopkins, T. Jakobsen, S.E. Lilleland, R.S Roxar, Ø Brandal. New Generation Software Integrates Intrusive and Non-intrusive Systems for Corrosion and Sand/erosion Monitoring. SPE International Conference on Oilfield Corrosion, 24-25 May 24-25, 2010, Aberdeen, UK.

[9] L. Ngkleberg, and T. Sontvedt. Erosion in choke valves-oil and gas industry applications. Wear, Vol. 186-187, Part 2, pp. 401-412, 1995.

[10] P.M. Birchenough, D. Cornally, S.G.B. Dawson, P. McCarthy, S. Susden. Assessment of Choke Valve Erosion in a High-Pressure, High-Temperature Gas Condensate Well Using TLA. SPE European Petroleum Conference, October 25-27, 1994, London, UK.

[11] Metso Automation. Flow Control Manual. 4th edition, 2005.

[12] P. Baraldi, E. Zio, G. Gola, D. Roverso, M. Hoffmann. Aggregation of randomized model ensemble outcomes for reconstructing nuclear signals from faulty sensors, Proc. ESREL Conf., Sept 7-10, 2009, Prague, CZ.

[13] P.F. Fantoni and A. Mazzola. Multiple-Failure Signal Validation in Nuclear Power Plants using Artificial Neural Networks, Nuclear technology, Vol. 113, Issue 3, pp. 368-374, 1996.

[14] M. Marseguerra, E. Zio, F. and Marcucci. Continuous Monitoring and Calibration of UTSG Process Sensors by Autoassociative Artificial Neural Network. Nuclear Technology, Vol. 154, Issue 2, pp. 224-236, 2006.

[15] B.Y. Sun, D.S. Huang, H.T. Fang. Lidar Signal Denoising Using Least-Squares Support Vector Machine. IEEE Signal Processing Letters, Vol. 12, Issue 2, pp.101-104, 2005.

[16] R. Chevalier, D. Provost, R. Seraoui. Assessment of Statistical and Classification Models For Monitoring EDF's Assets, Sixth American Nuclear Society Int. Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies, Knoxville, USA, 2009.

[17] M.P. Wand, W.R. Schucany. Gaussian-based kernels for curve estimation and window width selection. Canadian Journal of Statistics, Vol. 18, pp. 197–204, 1990.

[18] R. Polikar. Bootstrap-inspired techniques in computational intelligence, IEEE Signal Processing Magazine, Vol. 59, pp. 59–72, 2007.

[19] P. Baraldi, A. Cammi, F. Mangili, E. Zio. Local Fusion of an Ensemble of Models for the Reconstruction of Faulty Signals, IEEE Trans. on Nucl. Sci., Vol.57, Issue 2, pp. 793-806, 2010.