I&EC research
Industrial & Engineering Chemistry Research

Article

pubs.acs.org/IECR

Fault Detection and Isolation in a Spiral-Wound Reverse Osmosis (RO) Desalination Plant

Xavier Pascual,[†] Han Gu,[‡] Alex Bartman,[‡] Aihua Zhu,[‡] Anditya Rahardianto,[‡] Jaume Giralt,[†] Robert Rallo,[§] Panagiotis D. Christofides,[‡] and Yoram Cohen*[,‡]

[†]Departament d'Enginyeria Química, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Catalunya, Spain

[‡]Department of Chemical and Biomolecular Engineering, Water Technology Research Center, University of California, 5531 Boelter Hall, Los Angeles, California 90095-1592, United States

[§]Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Catalunya, Spain

ABSTRACT: Sensor fault detection and isolation (SFDI) approaches, based on support vector regression (SVR) plant sensor models and self-organizing-map (SOM) analysis, were investigated for application to reverse osmosis (RO) desalination plant operation. SFDI-SVR and SFDI-SOM were assessed using operational data from a small spiral-wound RO pilot plant and synthetic faulty data generated as perturbations relative to normal plant operational data. SFDI-SVR was achieved without false negative (FN) detections for sensor deviations of $\gtrsim$|10%| and FN detections of, at the most, $\lesssim$|5%|, and for sensor deviations of $\gtrsim$|4%| at sensor fault detection (FD) thresholds of up to ∼|4%|. False positive (FP) detections were almost invariant, with respect to sensor FD, being $\lesssim$|5%| for sensor deviations of $\gtrsim$|5%|. Corrections of faulty sensor readings were within SVR model accuracy (AARE < 1%) for SFDI-SVR and $\lesssim$|5%| for SFDI-SOM. Although SFDI-SOM has lower detection accuracy, it requires a single overall plant model (or SOM), while providing pictorial representation of plant operation and depiction of faulty operational trajectories.

1. INTRODUCTION

Water desalination by reverse osmosis (RO) membrane technology has been increasingly deployed for potable water production from seawater and water reuse applications including municipal wastewater and agricultural drainage water. Most RO plants are generally designed to operate at relatively steady-state conditions with traditional control strategies to attain the target permeate productivity and quality. Given the complexity of RO plants, plant process models, which consider specific plant characteristics and equipment, are needed to describe plant operation in order to optimize water production and design robust process control strategies.[1−4]

Plant process models require reliable sensor measurements as deviations of sensor readings (e.g., due to sensor failure or drift) can result in drift of plant operational variables beyond acceptable limits.[5] In this regard, fault tolerant control (FTC)[6] strategies need to be provided with effective fault detection and isolation (FDI) methods to identify faults in critical plant sensors or actuators that could degrade control system performance. FDI methods can be designed to detect system faults, and to identify their root cause by isolating system components whose operation lies outside the nominal range.[7] Indeed, model-based FDI systems have been successfully applied in different fields, and they have been integrated with systems of vehicle control, power, manufacturing, as well as in robotics and process control systems.[6,8−13]

In recent years, model-based FDI methods have been applied in RO water desalination.[9,14−16] For example, actuator FDI integrated with a fault-tolerant-control (FTC) strategy was reported for a single-membrane-unit RO desalination process without pre-treatment or post-treatment. The approach[15] relied on model-based feedback control laws making use of the fundamental RO transport equations and was successfully applied to operation with varying levels of feed salinity fluctuations. A supervisory switching law was derived to guarantee closed-loop stability by determining the activation time of fall-back control configurations in the presence of faults in the primary control configuration. The above approach was able to detect and isolate actuator faults in the system's adjustable retentate and pump bypass valves, as well as to recover the desired system operational regime by switching to the appropriate control strategy (i.e., using redundant actuators different than the actuators used in the primary control configuration). The above approach was feasible due, in part, to the deployment of redundant controls that provided alternatives to compensate for faulty elements.[17,18]

Sensor fault detection and isolation (SFDI), also known as gross error detection or sensor validation,[5] is a specific case of the general fault diagnosis and handling problem. Once a faulty sensor is detected and isolated, data reconciliation or rectification is necessary in order to estimate the sensor true reading values for the faulty data trace. The above approach was demonstrated for RO desalination plants for detection of sensor along with other faults in the plant[14,16] as the basis for model-based fault tolerant RO plant control.[14] Faults (e.g., in system hydraulics, pumps, sensors, and actuators) were induced in a computerized model of the RO plant (treatment capacity

Received: October 25, 2013
Revised: January 19, 2014
Accepted: January 22, 2014
Published: January 22, 2014

   dx.doi.org/10.1021/ie403603x | *Ind. Eng. Chem. Res.* 2014, 53, 3257−3271

of 500 L/h of feed with a conductivity value of 800 μS/cm, producing 250 L/h of permeate with a conductivity of 7 μS/cm). FTC was demonstrated[14] for two different scenarios: the first involved reduction of the retentate valve speed, and the second involved the application of a constant negative offset to the permeate flow rate sensor. It should be noted that, in the latter case, the permeate flow rate sensor was considered faulty when balance closure errors exceeded a specific tolerance. Such an approach can be beneficial when there is a priori knowledge (i.e., isolation) of the faulty sensor. However, in actual plant operation, detection and isolation are crucial, since one needs to identify (via appropriate procedures) and confirm which sensors are providing normal or abnormal (including missing) readings.

Statistical techniques such as principal component analysis (PCA), based on the analysis of process history, are commonly used for fault detection and monitoring. For example, monitoring and fault detection was carried out applying a PCA-based scheme and an unfolded PCA (U-PCA) using an RO desalination plant simulation (~1.7 m^3/h capacity with sand and cartridge filters for RO feed pretreatment) that operated with cyclical cleaning phases.[16] The approach was tested by simulating plant operation using an object-oriented and dynamic simulation tool (EcosimPro[19]). Three different U-PCA models were developed (for the two filters and for the RO membranes). A PCA model was first established using nominal data, where the square prediction error (Q) was used for all plant variables (including pressures, flow rates, and concentrations) to monitor and detect faulty data. Three types of faults (generated as 0%−60% deviation from the nominal sensor readings) were considered in the plant that included offset in the sand filter pressure sensor, as well as various types of membrane blockages and integrity losses. Faults detection was established on the basis of Q values exceeding a specific threshold, and whereby the Trimmed Score Regression method (TSR)[20] was used for correction of faulty data. The number of false positives observed with the U-PCA method, relative to classical PCA, was reduced from 11.6%, 12.9%, and 10.5% to 10.2%, 11.3%, and 4.0% for the sand filter, cartridge filter, and membrane models, respectively. Faults were detected (i.e., 100% detection) without delay (i.e., instantaneous detection), since only abrupt faults were considered. It is important to note that abrupt faults are only a subset of the faults that can occur in the operation of RO plants. For example, instrument behavior may deviate gradually from its nominal operation (for example, during fouling, scaling, and sensor drifting). In such cases, additional strategies that take into account historical plant operational data are essential to confirm fault identification. However, detection of the above would clearly involve a time delay, because of the nature of such faults.

The majority of previous studies have focused on FDI of equipment and actuators in RO plant operation. However, FDI of RO plant sensors is equally important in order to ensure optimal and safe plant operation. FDI of plant sensors requires accurate plant and sensor models, which are typically unavailable or impractical to develop for RO plants/processes that involve complex processes (i.e., involving coupled fluid flow, mass transfer, and energy transport) whose nature (and coupling) often cannot be predicted a priori (e.g., due to fouling and mineral scaling, and gradual deterioration of plant components including sensors). The above challenge can be overcome, for the purpose of FDI, using data-driven plant process models developed using machine learning (ML) techniques that have been shown to provide highly accurate description of RO plant operation.[21−24] The advantage of ML models of plant operation is that they can be used for plant process data validation and reconciliation and, thus, data correction and imputation to compensate for abnormal or missing (respectively) sensor readings.[21,25,26]

In the present work, two approaches for RO plant sensor fault detection and isolation (FDI), as well as sensor data imputation (SenDI) were assessed based on the application of self-organizing maps (SOM) and support-vector regression (SVR). The application of SVR and SOM machine learning methods to sensor FDI (termed here as SFDI) are fundamentally different in their approach, the latter being a classification-based method that provides a visual portray of overall plant and sensor behavior, while the former is based on providing a quantitative data-driven model for relating output to input variables. The use of SOM for FDI has been proposed for various applications including, for example, vehicle cooling systems,[27] aircraft engines,[28] induction motors and electrical machines,[29,30] power transformers[31,32] and anesthesia systems.[33] Although SOM have been used to analyze the operation of RO systems,[21,34] SOM analysis has not been previously proposed as the basis for FDI in RO plants. Accordingly, in the present work, FDI and SenDI approaches are developed, using the SOM and SVR machine-learning methods, and demonstrated based on operational data from a small laboratory spiral-wound RO pilot plant.

2. EXPERIMENTAL PROCEDURE

2.1. Feed Solution and Materials. Aqueous salt feed solutions were prepared using analytical-grade sodium chloride (Fisher Scientific, ACS grade, Pittsburgh, PA) in deionized (DI) water. The spiral-wound RO membranes utilized in pilot RO system (Dow Filmtec XLE-2540, The Dow Chemical Company, Midland, MI) had elements with an outer diameter of 2.5 in. (0.0635 m) and length of 40 in. (1.02 m), with a per element surface area of 2.6 m^2. A single membrane element had a manufacturer-reported permeate flow rate of 3.2 m^3/day, and salt rejection of 99%, determined at a pressure of 6.9 bar for a 500 mg/L NaCl feed solution.

2.2. RO System. Data for model development were generated using the UCLA spiral-wound Mini-Mobile-Modular (M3) pilot RO desalination system shown schematically in Figure 1.[3,35,36] The M3 system was designed for permeate water

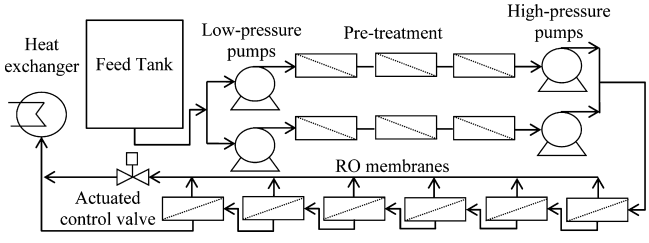

Figure 1. Configuration of the spiral-wound RO pilot plant.

production capacity up to 1.2 m^3/h (7560 gallons/day) for brackish water (5000 mg/L TDS) operating at 75% recovery and up to 0.64 m^3/h (4058 gallons/day) for seawater desalination (at recovery of 40%), using up to 18 spiral-wound elements in various configurations. In the present study, a configuration of 6 elements (one per pressure vessel) in series was utilized with the system operating in a total recycle mode

with permeate and concentrate streams returned to the feed tank. Briefly, the M3 RO plant consisted of a 450-L feed tank with two low-pressure feed pumps (Model JM3460-SRM, Sea Recovery, Carson, CA) pumping the RO feed through a series of cartridge microfilters (5, 0.45, and 0.2 μm; 08P GIANT, pleated 177 polypropylene filter cartridges, Keystone Filter, Hatfield, PA). The RO plant was operated such that filtered feed was fed to the RO membranes via two-high pressure pumps (Danfoss Model CM 3559, 3HP, 3450 rpm, Baldor Reliance Motor, Sea Recovery Corp. Carson, CA) operating in parallel and controlled by variable-frequency drives (VFDs) (Model FM50, TECO Fluxmaster, Round Rock, TX). The retentate flow rate and pressure in the RO unit were set by a model-based controller[35] that adjusted both an electrically actuated needle valve (valve V-1) (model VA8 V-7-0-10, ETI Systems, Carlsbad, CA) on the retentate stream of the M3 RO system and the pump VFD. In order to maintain the RO feed temperature, a heat exchanger (Model BP 410-030 refrigerant heat exchanger, ITT Industries) was installed on the retentate side of the RO system. Permeate and retentate streams were monitored via in-line conductivity sensors, conductivity/resistivity sensor electronics (Signet 2839 to 2842 and Signet 2850, George Fischer Signet, Inc. El Monte, CA) and pH sensor (DryLoc pH electrodes 2775, George Fischer Signet, Inc., El Monte, CA). The M3 plant was equipped with a centralized data acquisition system that received all sensor outputs (0−5 V, 0−10 V, 4−20 mA) and converted the signals to process variable values.

2.3. Experimental Procedure. RO desalting experiments covered the range of operating conditions that were feasible by the operability limit of the system for the specific feed salinities. The experiments covered feed pressure and feed flow rate ranges of 7.9−24.8 bar (or 115−360 psi), and 0.26−0.68 m^3/h (or 1−3 gpm), respectively. The RO plant control system was programmed to autonomously step through a range of feed flow rates and transmembrane pressures, whereby each experiment was carried out until the attainment of steady state. Feed flow rate and pressure were varied by the system controller which provided the necessary adjustments of the high-pressure pumps VFD and actuated valve settings (Figure 1). The data were logged into both the system embedded computer and a remote network database. Data could be logged at a frequency range as high as 1 kHz, although, for the present study, a frequency of a 1 Hz was deemed sufficient.

3. MODEL DEVELOPMENT

3.1. Data Preprocessing. A total number of 81 different operational states, over the entire operational domain of the M3 RO plant (see section 2.2), were generated by inducing pressure and flow rate step changes of up to 15% and 25%, respectively, with operational parameters recorded at a frequency of 1 Hz. The recorded raw data included natural plant fluctuations (i.e., due to operation of pumps and valves) as well as noise from normal operation of sensors, actuators, and system pumps. Data were recorded from a total of eight different sensors (Table 1) and the average standard deviation (ASTD) for each measured sensor variable was determined from:

$$\mathrm{ASTD} = \frac{1}{N}\sum_{i=1}^{N}\frac{\mathrm{STD}_i}{y_{i,\mathrm{ave}}} \times 100 \tag{1}$$

Table 1. Range of Steady-State Plant Operating Parameters Covered in the Study

sensor number	variable	range	steady-state STD[a]
1	feed flow rate	0.26−0.68 m^3/h	1.21%
2	feed conductivity	9842−10 766 μS	0.49%
3	feed pressure	7.90−24.82 bar	0.38%
4	permeate flow rate	0.06−0.28 m^3/h	0.94%
5	permeate conductivity	629−1823 μS	0.55%
6	retentate flow rate	0.11−0.56 m^3/h	1.28%
7	retentate conductivity	11 539−20 411 μS	0.33%
8	retentate pressure	3.59−10.75 bar	0.98%

[a]The steady-state STD represents the average percent standard deviation of the readings for a given sensor, at steady-state operation, relative to its average steady-state value, based on all the steady-state traces.

where STD is the standard deviation for the given variable within a steady state (for the set of operating conditions) trace i, $y_{i,\mathrm{ave}}$ the average variable value for the given trace, and N the number of the experimental steady-state traces.

Data-driven models (i.e., SVR and SOM, see sections 3.2 and 3.4, respectively) of the state-of-the-plant (i.e., steady state operation) were developed using steady-state data traces of 30 s, for the different steady-state operational states, at a sampling frequency of 1 Hz. The steady-state period was established as that for which the measured process variables did not vary with time by more than 3%, with respect to the time-averaged values. The above short (30 s) traces at a sampling frequency of 1 Hz were sufficient for model development; longer operational data traces and higher sampling frequencies did not provide significant model improvements but did increase the computational burden.

The SVR models were developed with sensor data (Table 1) that were normalized (for each sensor) in the range of [0, 1], using min-max normalization:

$$y_i^{(\mathrm{n})} = \frac{y_i - \min(y)}{\max(y) - \min(y)} \tag{2}$$

where $y_i^{(\mathrm{n})}$ is the normalized value of the experimental data y_i, and $\min(y)$ and $\max(y)$ are the minimum and maximum variable values in the dataset, respectively. Subsequently, for the purpose of SVR model building, the data were divided into two sets: one for model training and the other for model testing. The training dataset was selected to cover the entire plant operational domain (i.e., applicability domain[37]), by assuring that the maximum and minimum values of all variables were included in the training set. Data for model testing were selected randomly from the remaining data (i.e., not including the training set) but within the applicability domain to avoid extrapolation during predictions. In all cases, complete steady-state sequences were selected for the training and test sets.

Data for SOM generation were normalized to unit variance, in order to facilitate clustering of the different variables at the same scale (i.e., same mean and variance), such that

$$y_i^{(\mathrm{n})} = \frac{y_i - \mu}{\sigma} \tag{3}$$

where $y_i^{(\mathrm{n})}$ is the normalized value of the experimental variable (y_i), and μ and σ are the mean and standard deviation of the

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

variable values in the dataset, respectively. It is noted that, in contrast to SVR models which are based on supervised learning, SOM is a neural network method based on unsupervised learning,[38] which utilizes the entire dataset.

In order to test the ability of both the SVR and SOM approaches to identify abnormal sensor behavior, anomalous sensor readings were generated by introducing perturbations relative to the real plant data (i.e., deviations). Synthetic faulty sensor behavior was generated by multiplying a trace of real plant sensor data by a perturbation vector of the same length as the sensor data trace, while the remaining sensor readings were kept at their real values (i.e., only one sensor was perturbed at a time). The generation of perturbed traces was repeated for each of the eight sensors, such that eight sets of faulty data (i.e., one for each of the eight sensors) were generated with each containing only one faulty sensor data. Sensor reading (SR) deviations that ranged from −50% to +50%, relative to the real sensor readings, in 10% intervals, were then generated. SR deviations below ±4% were not assessed, since the observed variability of sensor reading during steady-state operation was up to approximately ±3%. Two different types of abnormal traces were generated: (a) a constant sensor drift (i.e., all values of the perturbation vector were the same for a specific deviation), and (b) an initial progressive sensor drift that stabilizes at a specific SR deviation. For the second type of abnormal sensor behavior, the first third of the sensor readings in the perturbation vector were set to either increase or decrease linearly until a preset percentage SR deviation (from the expected plant reading under normal plant operation) was attained, after which each successive sensor reading was taken to deviate by the same multiplier. It is important to recognize that, within the framework of fault detection and isolation, when a given sensor reading is outside its nominal range of operation (i.e., above or below the range of operability as specified by the manufacturer), this sensor would be immediately identified as being faulty. However, sensor readings that are within the nominal range could still be abnormal and thus one must perform detailed fault detection and isolation as per the SVR and SOM approaches described in sections 3.2 and 3.3, respectively.

3.2. Support Vector Regression (SVR) Models. Steady-state models (see sections 3.3) were developed with the support vector regression (SVR) algorithm,[37,39] using model training and test datasets (see section 3.1). Briefly, given a vector x of the RO process variables (e.g., feed pressure, feed flow rate, feed conductivity, retentate pressure, retentate flow rate, retentate conductivity, and permeate flow rate), the goal of SVR is to find a function $f(x)$ that has, at most, a deviation ε, with respect to the actual values of the target RO process variable y (e.g., permeate conductivity) and, at the same time, is as flat as possible (see Figure 2). SVR can be formulated as a convex optimization problem where a set of coefficients w for the regression model are computed in such a way that the flatness and accuracy of $f(x)$ are maximized. Since it is not always possible to keep the error within the margin ε for all the available data points, a pair of slack variables ξ and ξ^* must be introduced within the SVR formulation to cope with otherwise infeasible optimization constraints (Figure 2). In most cases, the optimization problem can be solved more readily by projecting x (i.e., the vector of the input RO process variables) onto a higher-dimensional space, where linear regression models can be developed for the target RO process variable of interest. The functions used to perform the above linear to

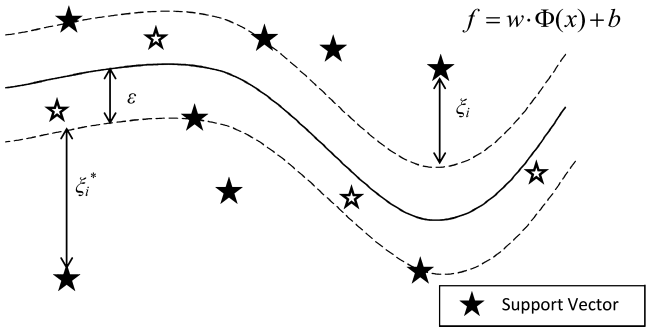

Figure 2. Support vector regression (SVR) structural parameters. Regression function (f) supported on the most representative vectors of information (support vectors), w the normal vector to the hyperplane generated, the kernel function (Φ), the vector of biases (b), slack variables (ξ), and radius of the insensitive tube (ε).

nonlinear mapping are known as kernel functions. Finally, the nonlinear models that relate input and target variables are obtained by mapping the data back to the original (i.e., nonlinear) space of the RO process variables. Only a subset of the training data points, representing the overall data behavior (i.e., support vectors), were used in the SVR formulation to generate the regression model (see Figure 2).

In the current work, the radial basis function (RBF) was selected as the kernel function, since it is suitable for systems of highly nonlinear behavior.[37] The key parameter characterizing this kernel is the width of the Gaussian (σ_g); it determines the area of influence of the support vectors over the data space and, here, its optimal value was determined via a grid search. The SVR-based models were developed in MATLAB using the LS-SVMlab1.7 package.[40,41] This SVR implementation utilizes a regularization parameter γ, which controls the tradeoff between the flatness (or smoothness) of the models and their accuracy, where the optimal parameter value was also determined via a grid search. The optimal values of γ and σ_g, obtained via grid optimization, for each of the models are summarized in Table 2.

3.3. Spiral-Wound RO Plant Sensor Fault Detection and Isolation (SFDI) Scheme Based on Support Vector Regression (SVR) Models. *3.3.1. State-of-the-Plant SVR Models.* The SFDI-SVR approach for detection of abnormal sensor behavior (Figure 3) was based on assessing the deviation (i.e., quantified as residuals) between actual sensor readings and those expected for normal plant response, as predicted from the SVR plant models. The use of SVR RO plant model was chosen, given their previously demosntrated accurate performance for describing spiral-wound RO plant operation.[21] In the SFDI-SVR approach, an individual state-of-the-plant (STP) SVR model was first developed for seven of the plant process sensors, excluding the feed conductivity ($cond_{feed}$) (see Table 1). Although the total number of sensors in the plant was eight, the feed conductivity was considered as a variable independent of all other process variables, which was not subject to process fluctuations. However, feed conductivity sensor readings were included as inputs to develop the SVR models for the remaining seven sensors. Accordingly, SVR models were developed for each of the seven sensors (excluding the feed conductivity sensor):

$$y_s^* = f(y_j) \qquad j = 1, \ldots, m+1;\ j \neq s \tag{4}$$

where y_s^* is the predicted reading for sensor s, j designates the specific plant sensor, and m represents the seven sensors (i.e.,

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

Table 2. Performance of SVR-Based Models for Predicting Sensor Values of RO Plant Operating under Steady-State Conditions

predicted variable	regularization parameter, γ	width of the Gaussian, σ_g	average absolute error, AAE	average absolute relative error, AARE (%)	linear correlation coefficient, r^2
feed flow rate (m^3/h)	10 000	20	1.1×10^{-4}	2.3×10^{-2}	1
feed pressure (bar)	200	20	9.4×10^{-2}	0.65	0.994
retentate conductivity (μS)	1000	20	42.50	0.32	0.997
retentate flow rate (m^3/h)	10 000	20	1.2×10^{-4}	3.2×10^{-2}	1
retentate pressure (bar)	8400	9	5.0×10^{-2}	0.82	0.997
permeate conductivity (μS)	200	20	8.23	0.70	0.999
permeate flow rate (m^3/h)	10 000	14	9.7×10^{-5}	3.2×10^{-3}	1

total number of sensors excluding the feed conductivity sensor). For example, the SVR model for the permeate conductivity sensor (sensor 5 in Table 1) was based on input data from the remaining sensors listed in Table 1 (i.e., 1−4 and 6−8, where $j \neq s$). For SVR model development, the steady-state data traces (for the range of operating conditions listed in Table 1) were divided into a training set (62% of the dataset, 50 traces) and a test set (38% of the overall dataset, 31 traces).

Performance of the different SVR sensor models (with the optimal γ and σ_g values listed in Table 2) was quantified using the linear r^2 correlation coefficient (i.e., between the predicted, y_i^*, and experimental, y_i, variable values), and the average absolute (AAE) and percent average absolute relative (AARE) errors given as

$$\mathrm{AAE}_s = \frac{1}{n}\sum_{i=1}^{n} |y_{i,s}^* - y_{i,s}| \tag{5}$$

$$\mathrm{AARE}_s = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_{i,s}^* - y_{i,s}}{y_{i,s}}\right| \times 100 \tag{6}$$

where n is the total number of data points for the dataset (training or test) and s is the sensor under consideration.

3.3.2. Fault Detection and Isolation with SFDI-SVR. It is important to note that fault detection and isolation for the conductivity sensor ($\mathrm{cond_{feed}}$) sensor was based on the maximum deviation of sensor reading, from its expected value for the specific feed solution, over the course of each experiment. Therefore, feed conductivity sensor readings were assessed as the first step in the process of fault detection and isolation (see Figure 3). A fault in the $\mathrm{cond_{feed}}$ sensor was inferred when its value differed more than 3.5% from the expected initially determined (preset). Once $\mathrm{cond_{feed}}$ was confirmed to be correct, the fault detection and isolation strategy for the remaining sensors was carried out. Otherwise, when the conductivity sensor ($\mathrm{cond_{feed}}$) was found to be faulty, it was corrected, given knowledge of the preset feed conductivity (see section 3.3.3) and the remaining sensors were checked for additional faults. Sensor fault detection and isolation (i.e., SFDI) strategy for the remaining seven sensors followed the basic structure of the so-called observer-based techniques.[7] Accordingly, the expected readings from each of the sensors are predicted by their respective SVR model and compared to their actual (experimental) readings, as quantified by the percent absolute relative difference (PARD):

$$\mathrm{PARD}_j = \frac{|y_j^* - y_j|}{y_j} \times 100 \tag{7}$$

where y_j^* and y_j are the predicted and experimental values of the sensor under consideration j. Sensor behavior is considered normal when PARD_j is below a prescribed PARD_j threshold (T_j) and is otherwise considered abnormal. In the present approach, the sensor thresholds represent the maximum absolute sensor deviation, with respect to normal plant operation, which is deemed as acceptable (or allowable). Since the identity of the faulty sensor is not known a priori, each of the m plant sensors (i.e., total number of sensors excluding the feed conductivity sensor) are checked sequentially assuming that each one (in turn) could be a faulty sensor. Accordingly, in a sequential order, the expected reading, y_s^*, is predicted for each given sensor (s) of the plant's $m-1$ sensors, using the sensor's SVR model. It is noted that the SVR model for a given sensor s is developed based on the operational (experimental) readings of the other $m-1$ sensors plus the feed conductivity sensor which are used as the SVR model input for the given sensor. For each sensor s, the PARD_j^s for each of the remaining $m-1$ sensors is then calculated; if the PARD_j^s of all these $m-1$ sensors are below the threshold T_j, then sensor s is flagged as a faulty candidate (FC) sensor. The above isolation sequence is repeated for all the sensors. However, it is stressed that SFDI becomes challenging when sensor deviations are small. Specifically, when sensor deviations are close to the PARD threshold, multiple faulty sensor candidates could be erroneously flagged as being potentially faulty. Also, if multiple sensors are corrected this could lead to erroneously identifying the remainder of the sensors as normal (i.e., accomplishing $\mathrm{PARD}_j^s < T_j$, and thus identifying only one fault). In order to avoid the above dilemma, for each FC sensor, the total percent absolute relative difference (TPARD) (eq 8) is evaluated as the sum of the PARD of the remaining sensors (i.e., the PARD for the FC sensor is not included in TPARD, since its measured value is flagged as being potentially faulty). It is noted that the FC sensor reading needs to be predicted in order to use its value to predict the remaining $m-1$ sensors by using the SVR model (eq 4). Accordingly,

$$\mathrm{TPARD}_{\mathrm{FC}} = \sum_{\substack{j=1 \\ j\neq \mathrm{FC}}}^{m} \mathrm{PARD}_j \tag{8}$$

where j designates the specific sensor (i.e., sensor number), m designates the seven plant sensors being assessed (i.e., total number of sensors excluding the conductivity sensor) and FC designate the sensor being assessed. The FC sensor of the minimum TPARD is identified as providing abnormal readings (i.e., faulty), while the remaining sensors are considered to be operating normally. The use of TPARD as the criterion for identifying the faulty sensor provides a relative measure of comparison of the deviation of predicted compared to experimental sensor readings. Accordingly, the FC sensor with the smaller TPARD is more likely to be faulty (or providing abnormal readings).

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

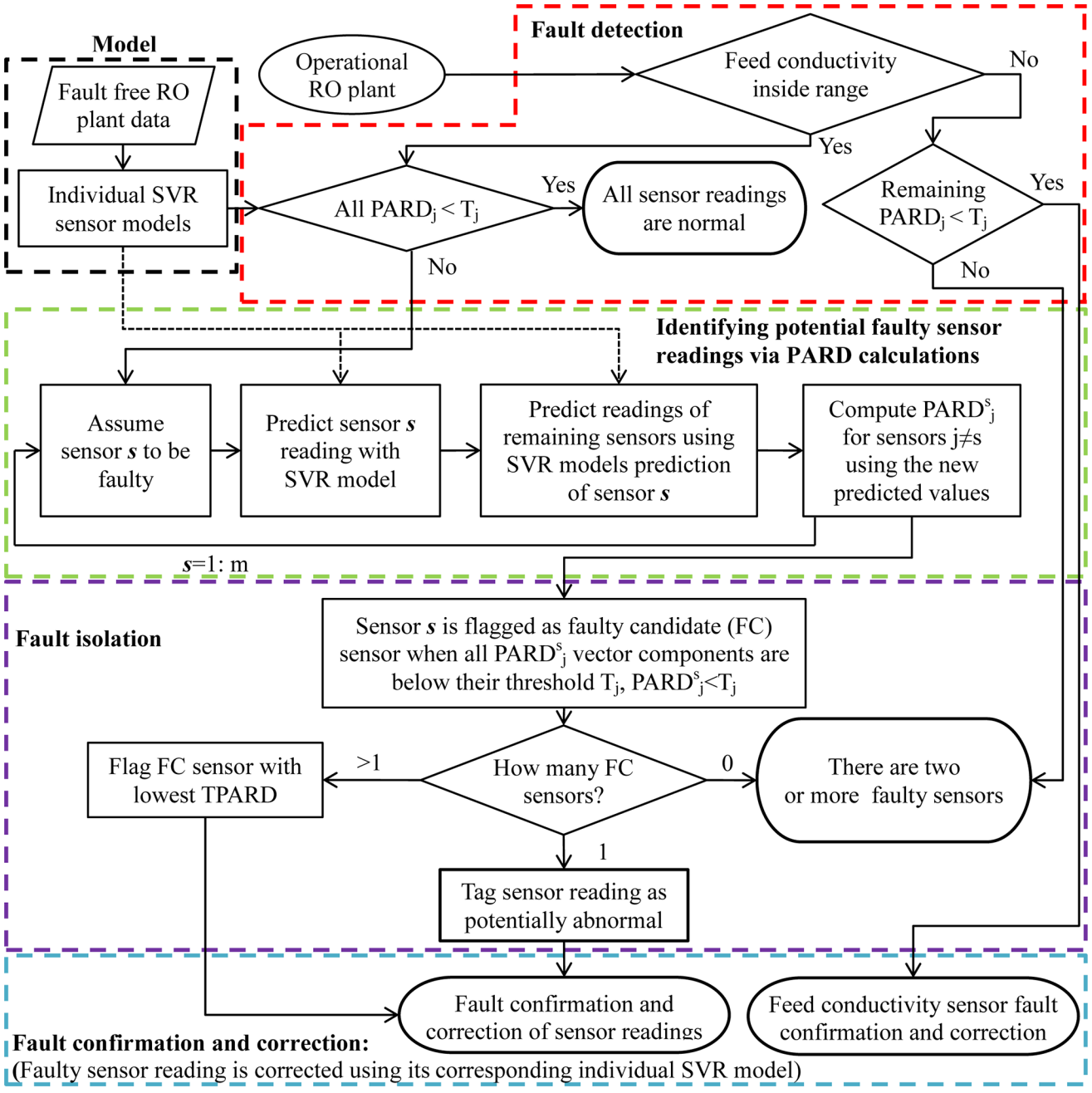

Figure 3. SVR model-based algorithm for detection and isolation of abnormal sensor operation. The parameters *m*, PARD, TPARD, and T refer to the number of plant sensors (not including the feed conductivity sensor), the percent absolute relative difference, total percent absolute relative difference between sensor reading and expected values, and the sensor threshold, respectively. $PARD^s_j$ contains the PARD of all sensors, except for the FC sensor under consideration (i.e., $j \neq s$) and the feed conductivity sensor ($cond_{feed}$). Sensor *s* is identified as an FC sensor when each $PARD^s_j$ of the remaining sensors is below its corresponding threshold T_j. If only one FC sensor is identified, then it is directly flagged as being faulty. Note: If FC sensors are not encountered, this could suggest that more than one sensor may be faulty. If more than a single FC sensor is suspected, then the FC sensor with predictions closer to the experimental values (minimum TPARD) is selected as faulty.

3.3.3. Fault Confirmation with SFDI-SVR. In order to confirm the presence of abnormal sensor readings, sensor faults are checked over the entire diagnostic period (Figure 3). A minimum number of abnormal measurements is needed before flagging a sensor as a fault candidate. This is required in order to avoid errant fault detection due to sensor measurement spikes or fluctuations, which often occur in plant operation. In the present RO plant test case, a minimum of one-third of the total trace measurements was set as being required in order to confirm a sensor as being flagged as faulty. Subsequently, the "faulty" sensor readings were reconstructed based on predictions from its sensor SVR model using as input the readings from the other $(m - 1)$ sensors plus the feed conductivity sensor. When the feed conductivity sensor was identified as a potentially faulty sensor, the sensor reading was then replaced by the known present feed conductivity.

Although there is a generally low likelihood of the simultaneous occurrence of multiple faulty sensors, the present approach can be extended to such cases. However, fault isolation would then require SVR sensor models based on, in

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

addition to the feed conductivity sensor ($cond_{feed}$), $m - p$ sensor readings for each modeled sensor, where p is the number of potential faulty sensors to be identified. Such an approach may be feasible provided that the number of faulty sensors is significantly less than the total number of sensors.[6] Notwithstanding, the present approach of dealing with incidences of a single faulty sensor occurrence should be of great practical applicability, particularly since it can also provide a warning regarding the potential occurrence of multiple faulty sensors.

3.4. Spiral-Wound RO Plant Sensor Fault Detection and Isolation (SFDI) Scheme Based on Self-Organizing Maps (SOM) Analysis. *3.4.1. SOM and RO Operational Domain.* The Self-Organizing Map (SOM) clustering algorithm is used as an alternative approach for developing SFDI schemes (depicted in Figure 5, which is presented later in this work). The main advantage of SOM is that it can be used to characterize RO plant operational states in a manner that does not require an individual model for each plant sensor (see section 3.3). SOM also enables projection and visualization of multidimensional RO plant data into a two-dimensional (2D) lattice (map) of SOM units arranged in a hexagonal geometry (i.e., each SOM unit has six neighboring units).[38,42] Each of the input data vector dimensions (i.e., eight plant sensors of the present plant) is represented by a component plane (Figure 4).

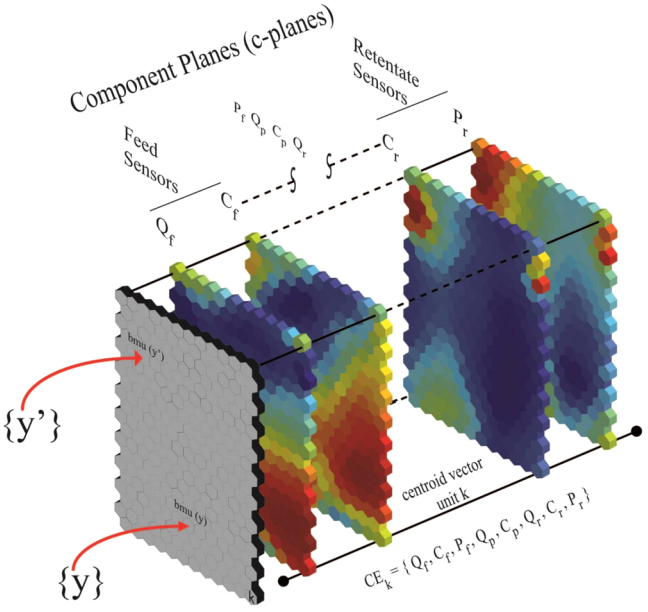

Figure 4. Each of the component planes (generated from its corresponding RO plant sensor data) are presented by a slice from the SOM. For the sake of clarity, the feed flow rate and conductivity, retentate pressure and retentate conductivity component planes are shown. Note that the overlaying SOM units contain information regarding a given plant operational state (e.g., SOM units along the solid line ended with circles, through the component planes, are for a similar plant operational domain).

The aggregation of all the component planes constitutes the complete SOM whose centroids contain information regarding all the data vector components (i.e., dimensions). In this approach, the identification of specific plant operation regimes is based on similarity calculations (i.e., Euclidean distance) between vectors formed by process variables (i.e., sensor readings) and real vectors that are represented by the centroid in each SOM unit. The SOM unit, which is most similar to a given plant operational data vector is known as the "best matching unit" (or bmu). In the current application, the vector $\mathbf{y} = [Q_f, C_f, P_f, Q_p, C_p, Q_r, C_r, P_r]$ represents each state of operation of the RO plant. Similarly, the centroid of a given unit i is a real vector $\mathbf{CE_i} = [\mu_1, \mu_2, ..., \mu_m]^T \in \Re^m$, where m is the dimension of the input data (i.e., m represents the number of plant sensors (8 for the present plant)). During SOM development, unit centroids are iteratively adapted to preserve, over the SOM projection, the topological relationships (i.e., ordering and distances) of the original high-dimensional RO plant operational space. As a consequence, similar SOM units are located close to each other resulting in clusters of similar operational states. In the current study, the SOM was developed using the MATLAB SOM Toolbox package.[42]

A SOM for the domain of normal plant operation is first developed based on fault-free data (Figure 4). Accordingly, the constructed SOM consists of SOM units encoding information from the various plant variables over the complete range of feasible operating conditions. Thus, the data vectors are classified using the SOM developed from fault-free plant data. The distance (i.e., similarity) of a given data vector relative to the centroid of its bmu (in the SOM built using the fault-free data) provides an indication of the consistency and reliability of the sensor reading relative to fault-free operational data. In other words, correct sensor readings are expected to be similar to readings obtained from fault-free plant data for the same operational state.

3.4.2. Fault Detection with SOM. In order to detect a sensor with faulty readings, an appropriate fault detection threshold was defined. As a first step in the process, for each unit in the SOM, it is necessary to identify the data vector furthest away from the centroid (i.e., representing the average of all normal operational data vectors of the SOM unit) in terms of the Euclidean distance. The threshold was then defined as the absolute differences between the data vector components (i.e., readings of each sensor) and those of the centroid vector components, thereby yielding a vector containing a fault detection SOM (FDS) threshold (T_s) for each of the sensors. The FDS served to delimit the operational regime boundary captured by that SOM unit. It can be assumed that data vectors located beyond the FDS threshold have a significant deviation relative to the normal operation regime captured by the unit. Based on this criterion, the absolute difference (**AD**) vector between the components of a RO plant data vector $\mathbf{y}$ and the components of the centroid $\mathbf{CE}^{\mathbf{bmu}(y)}$ of its corresponding SOM bmu is first calculated:

$$AD_s^y = |CE_s^{\mathbf{bmu}(y)} - y_s| \tag{9}$$

where y_s is the reading of sensor s and $CE_s^{\mathbf{bmu}(y)}$ is the corresponding component of the bmu centroid. The data vector $\mathbf{y}$ is flagged as being faulty when any of its $AD_s^y > T_s^{\mathbf{bmu}(y)}$, where $T_s^{\mathbf{bmu}(y)}$ is the FDS threshold corresponding to sensor s for the given SOM bmu. Similarly, the data vector $\mathbf{y}$ is considered as not faulty when the AD_s^y values for all eight plant sensors are smaller than their respective FDS thresholds ($T_s^{\mathbf{bmu}(y)}$).

3.4.3. Fault Isolation with SOM. Similar to the SVR-SFDI approach, in the SOM approach to fault isolation (i.e., identifying the faulty sensor once faulty data vector was identified), the feed conductivity sensor requires special attention, since feed salinity is typically fixed by the feed

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

source composition. Moreover, in the present work, feed salinity (or conductivity) was fixed over the course of each experiment. Therefore, in the present approach, readings of the feed conductivity sensor were first checked to evaluate if those were faulty. A new data vector of m components was created excluding the feed conductivity sensor, and it was classified inside the SOM to find its bmu and calculate its **AD** vector. A fault in the feed conductivity sensor was identified when all the **AD** component values were below their corresponding FDS threshold ($T_s^{\mathbf{bmu}(y)}$). Otherwise, the feed conductivity sensor was considered fault-free. In the process of fault isolation, a new set of m reduced data vectors ($\mathbf{y}(m-1)$ where m is the number of plant sensors (excluding the feed conductivity sensor), each containing information from only $m-1$ sensors (i.e., one sensor at a time is discarded in each of the new reduced data vectors along with the feed conductivity sensor), are generated from the original data vector. Therefore, with eight sensors in the present plant, seven new reduced data vectors are created, each containing 6 components. The bmu values corresponding to these new reduced data vectors are then identified, together with their corresponding **AD** vectors (eq 9). The fault isolation algorithm (Figure 5) then considers the following three different fault scenarios:

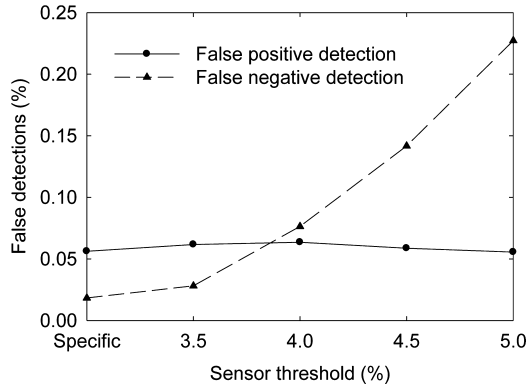

Figure 5. Detection and isolation algorithm for abnormal sensor operation based on self-organizing maps (SOM). The parameters m, AD, TAD and T refer to number of plant sensors (not including the feed conductivity sensor), absolute difference, total absolute difference, and sensor FDS threshold, respectively. Note: The $\mathbf{AD}^y_s$ contains the **AD** of all sensors except the FC sensor under consideration (i.e., $s \neq z$) or except for the feed conductivity sensor ($\text{cond}_{\text{feed}}$) (i.e., $s \neq \text{cond}_{\text{feed}}$). Sensor s is identified as FC sensor when each $\mathbf{AD}^y_s$ of the remaining sensors is below its corresponding FDS threshold (T_j). If only one FC sensor is identified, then it is directly flagged as being faulty. Note: If FC sensors are not encountered, this could suggest that more than one sensor may be faulty. If more than a single FC sensor is suspected, then the FC sensor with predictions closer to the experimental values (minimum TAD) is selected as faulty.

(a) *Only one of the seven reduced data vectors has all of its six **AD** vector components below their corresponding FDS thresholds.* In this case, the faulty sensor can be isolated and identified as the missing component in the above reduced data vector. The remaining reduced data vectors contain information from the faulty sensor and have at least one of their **AD** components above the FDS threshold;

(b) *The case where none of the seven reduced data vectors have all their **AD** components values below their FDS thresholds.* In the above situation, the set of SOM units most similar to the bmu (i.e., units located in the vicinity of the bmu) are also checked to ensure that there are no other neighboring SOM units (i.e., corresponding to similar operational states) with **AD** components below the FDS thresholds. In the present work, for all reduced data vectors, the maximum number of consecutive bmu values that must be checked, to ensure that there are no other proximal units that may also have **AD** components below the FDS threshold, was set to nine (i.e., the six SOM units adjacent to the hexagonal bmu plus three additional SOM units with the shortest proximal distance from the bmu). The above conservative heuristic strategy (to reduce computational time) was employed in order to check if there are proximal SOM units next to the bmu that also satisfy the FDS threshold). When the first bmu that satisfies the FDS threshold (i.e., with **AD** values below the FDS threshold) is found, the confirmation and correction of the faulty sensor is then carried out as described in section 3.4.4. However, when none of the first nine bmu values satisfy the FDS threshold, fault isolation cannot be carried out, since none of the reduced data vectors correspond to a normal plant operational state, which is an indication that there are likely to be two or more faulty sensors. Similar to the SFDI-SVR approach (see section 3.3), a new set of reduced data vectors, containing information for only $m-k$ sensors, needs to be used in order to isolate k faulty sensors; and

(c) *The situation in which there is more than a single reduced data vector that has all its **AD** components below the FDS threshold, suggesting two or more faulty candidate (FC) sensors.* For each FC sensor, the total absolute difference TAD^{FC} (eq 10) is computed as the sum of AD components, excluding the FC sensor.

$$\text{TAD}^{\text{FC}} = \sum_{\substack{s=1 \\ s\neq \text{FC}}}^{m} \text{AD}^y_s \tag{10}$$

In the above scenario, the reduced data vector (i.e., in which the FC sensor has been removed) with the lowest TAD^{FC} is identified as the data vector more likely to be fault-free, and therefore its discarded sensor component FC is inferred as faulty.

3.4.4. Fault Confirmation and Correction. Similar to the SFDI-SVR (see section 3.3), in the SFDI-SOM one also needs to ensure that the analysis is not biased by typical plant fluctuations that may be interpreted as produced by faulty sensor readings. Therefore, a minimum percentage of the diagnostic trace period is required to confirm abnormal sensor behavior and thus the approach was evaluated up to one-third of the trace period (at 1 kHz data acquisition frequency; see section 3.1). Once a sensor fault is confirmed, the faulty sensor reading is corrected by replacing it with the corresponding component value in its bmu centroid.

4. RESULTS AND DISCUSSION

4.1. Performance of SFDI-SVR. State-of-the-plant (STP) models (see section 3.3) were first developed in order to provide predictions for each one of the sensors of the plant. Performances of STP models applied to the test dataset are provided in Table 2. Models predicted feed, retentate, and permeate flow rates with average absolute relative error

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

Table 3. Performance of the SFDI-SVR for Induced Sensor Deviations in the Range of ±4% to ±50% for both Global and Individual Sensor Thresholds[a]

	Individual Threshold		Global Sensor Threshold							
			3.5%		4%		4.5%		5%	
Dev. (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)
−50	0	0	0	0	0	0	0	0	0	0
−40	0	0	0	0	0	0	0	0	0	0
−30	0	0	0	0	0	0	0	0	0	0
−20	0	0	0	0	0	0	0	0	0	0
−10	0	0	0	0.40	0	0.40	0	0.40	0	0.40
−9	0	0.81	0	0.81	0	0.81	0	0.81	0	0.81
−8	0	0.81	0	0.81	0	0.81	0	0.81	0	0.81
−7	0	1.61	0	2.02	0	2.02	0	1.61	0	1.61
−6	0.4	2.42	0.4	3.23	0.4	3.63	0.4	3.63	0.4	3.23
−5	0.4	3.63	0.81	5.24	1.21	6.05	2.02	5.65	6.45	4.84
−4	1.61	7.26	4.03	6.85	8.47	7.26	34.27	5.65	48.39	5.24
+4	5.24	6.85	8.06	6.85	31.85	6.45	45.16	6.05	52.82	6.05
+5	2.42	4.44	2.82	4.84	5.24	4.84	8.06	4.84	34.27	4.84
+6	0.81	3.23	1.21	4.03	1.61	3.63	2.02	3.23	4.44	3.23
+7	0.4	2.42	0.4	2.42	0.81	2.42	0.81	2.42	1.21	2.02
+8	0.4	2.02	0.4	2.02	0.4	2.02	0.4	2.02	1.21	2.02
+9	0.4	0.81	0.4	0.40	0.4	0.81	0.4	0.81	0.4	0.81
+10	0	0.81	0	0.81	0	0.81	0	0.81	0.4	0.81
+20	0	0	0	0	0	0	0	0	0	0
+30	0	0	0	0	0	0	0	0	0	0
+40	0	0	0	0	0	0	0	0	0	0
+50	0	0	0	0	0	0	0	0	0	0

[a]FN and FP respectively designate the percentages of false negatives and false positive of the total number of existent.

(AARE) values of 2.3×10^{-2}%, 3.2×10^{-2}%, and 2.3×10^{-2}%, respectively. Similarly, feed, retentate, and permeate conductivities were predicted with AARE levels of 0.33%, 0.32%, and 0.70%. Feed and retentate pressures were predicted with AARE levels of 0.65% and 0.82%, respectively, with greater accuracy of the flow rate compared to the pressure and conductivity sensor models. It is noted that the linear correlation coefficient (r^2) (for the predicted versus measured values) was, in all cases, ≥0.994. As discussed in section 3.3.1, a model for the feed conductivity sensor was not implemented since the feed conductivity is an independent variable not affected by the remainder of the process variables. In fact, SVR model for this sensor (in terms of the readings from the other seven sensors) yielded a model with a correlation coefficient of only 0.76. This was not surprising, since the average deviation of the feed sensor, relative to the steady-state average, was only ∼0.49% while the stated manufacturer sensor performance was rated within an error of ±2%.

The SFDI-SVR approach was evaluated for different sensor thresholds (see Table 3). Steady-state operation was considered established (see section 3.1) when the measured process variables did not vary with time by more than 3%, with respect to the time-averaged values. Accordingly, the effect of the global sensor threshold was evaluated for sensor deviations over a range of 3.5% - 5% (i.e., slightly above the steady-state deviation of normal plant operation). In addition, the SFDI-SVR approach was assessed with individual sensor thresholds for each sensor; these were based on each sensor's maximum deviation observed (relative to the average) during steady-state plant operation. Sensor thresholds for the retentate and permeate conductivity and feed pressure were set at 2%, feed flow rate at 2.5% and 3.5% for the retentate and permeate flow rate, retentate pressure and feed conductivity sensors (see section 3.3.2). Using the above, performance of the SFDI-SVR approach for sensor deviations (i.e., faults) between ±4% to ±50% and for sensor thresholds from 3.5% to 5%, as well as for individual sensor thresholds is shown in detail in Table 3.

SFID-SVR performance (Table 3) was assessed in terms of false negative (FN) and false positive (FP) detections being respectively defined as faulty sensor readings not identified as being faulty or fault-free sensor readings incorrectly identified as faulty. For both global and individual sensor fault detection thresholds, FN detections were zero (i.e., 100% detection of all faulty readings) for the above range of sensor thresholds (3.5%−5%) for up to sensor deviations of ±10%. FP detections occurred for −10% ≤ sensor deviation ≤10%, with FP detection of 0.81% and 0.4% for sensor deviations of +10% and −10%, respectively, when using the global sensor thresholds. There were no FP detections for the above range of sensor deviations when using the individual sensor thresholds. When the SFDI-SVR algorithm (Figure 3) was challenged with lower sensor deviations (i.e., ±4% to ±9%), the percentage of FPs and FNs increased as shown in Figure 6. As expected, FN deviations increased with decreasing sensor deviations and increasing global sensor thresholds. On the other hand, FP detections remained essentially invariant with respect to global sensor thresholds but increased with decreasing sensor deviations. Overall, individual sensor thresholds demonstrated lower FN detections, relative to the use of global sensor thresholds, while FP detections were similar for the two types of sensor thresholds.

In deploying the SFID-SVR, it is desirable to minimize the number of FN detections from the viewpoint of plant safety and maintenance of process and equipment integrity. On the

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

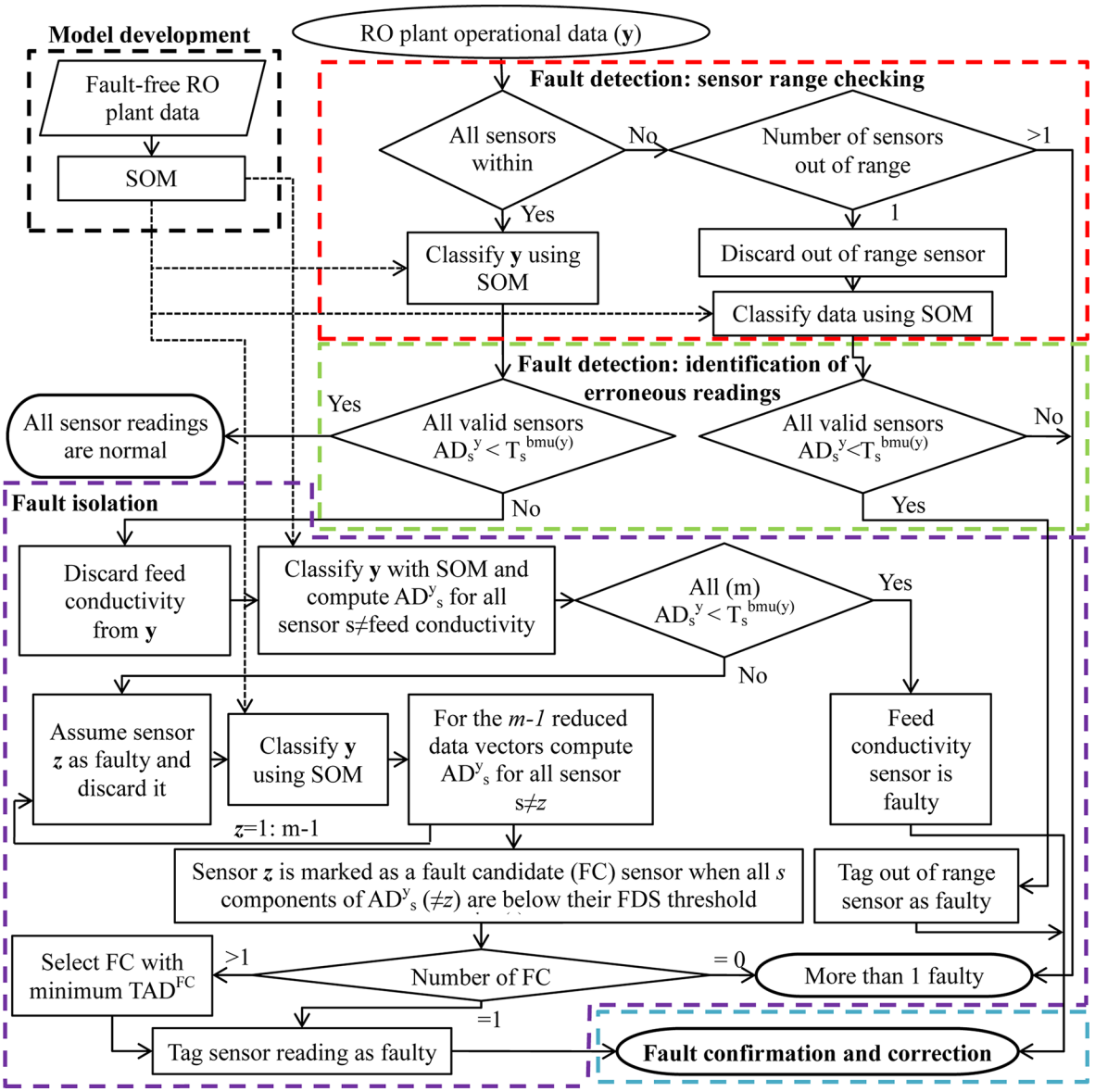

Figure 6. Total false positive and false negative detections for the plant eight sensors over the range of sensor deviations of ±4% to ±50%. The results are shown for both global and individual (first data points as specified in the sensor threshold% axis) sensor thresholds.

other hand, the occurrence of FP detections in a plant would trigger false alarms; these, however, can be minimized through a secondary process of using traditional decision tree analysis. It is noted that, in the current study, FP detections were encountered only for small sensor deviations. Clearly, small sensor deviations that are identified as either FP or FN detections are of lesser concern, relative to those associated with large sensor deviations, since the former are less likely to have significant impact on plant operation. As is evident from Figure 6, the number of FN detections increased (from 0.028% to 0.23%) for higher sensor thresholds, while the number of FP detections decreased marginally (from 0.062% to 0.056%) over the range of global sensor thresholds of 3.5%−5%. Overall, it can be concluded that for the present RO pilot system, a greater allowance of sensor readings fluctuations, over a steady-state operational period, can result in a higher rate of FN detections, while the number of FP detections is only slightly impacted. Therefore, the above observation suggests that individual sensor thresholds may be preferred when sensor performance accuracy data are available, while in the absence of the above global sensor thresholds can be utilized.

The SFDI-SVR method was also tested with increasing and decreasing sensor drifts (see section 3.1) to determine the minimum sensor deviations at which fault detection can be made (i.e., detection limit), with the detection performance summarized in Table 4. Accordingly, fault detection was evaluated for each one of the m plant sensors (not including the conductivity sensor). The minimum deviation (MD) for a given sensor that enabled fault detection, during a sensor drift (as described in section 3.1), are provided in Table 4 representing the average (for each sensor) for 372 data traces of sensor drifts (i.e., for periods of faulty sensor operation). To check for consistency in the minimum sensor deviation at which fault detection can be made, this detection limit was compared for sensors measuring the same variable type. As shown in Table 4, the fault detection limits for the feed and retentate flow rates were similar (3.45% and 3.66%, respectively), but higher (5.56%) for the permeate flow rate.

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

Table 4. Minimum Sensor Deviation That Enabled First Fault Detection[a]

	avg. det. limit[b] (%)	accuracy of sensor SVR model, AARE[c] (%)
feed flow rate	3.45	0.02
retentate flow rate	3.66	0.03
permeate flow rate	5.56	0.09
retentate conductivity	3.73	0.32
permeate conductivity	3.73	0.70
feed pressure	5.45	0.65
retentate pressure	5.56	0.82

[a]Based on individual sensor thresholds. [b]avg. det. limit = minimum sensor deviations at which fault detection can be made (based on 372 data traces of sensor drift tests; see section 3.1). [c]AARE = average absolute relative error.

The fault detection limits for the feed and retentate pressures were also similar (5.45% and 5.56%, respectively), and essentially identical for the retentate and permeate conductivity meters. As expected, the order of the fault detection limits correlated with the accuracy of the sensor SVR models (Table 2), whereby, for similar type sensor measurements, the SVR model of higher accuracy enabled a lower fault detection limit.

Once a faulty sensor was identified, faulty data were corrected using the state-of-the-plant sensor SVR models (see section 3.3.1). Thus, data correction was within the same accuracy as that of the sensor SVR models (Table 2). As an illustration, the feed pressure sensor corrections, for two selected abnormal traces of sensor drifts (induced by increasing perturbations of up to ±5% relative to normal sensor operation), are shown in Figure 7. It is important to recognize

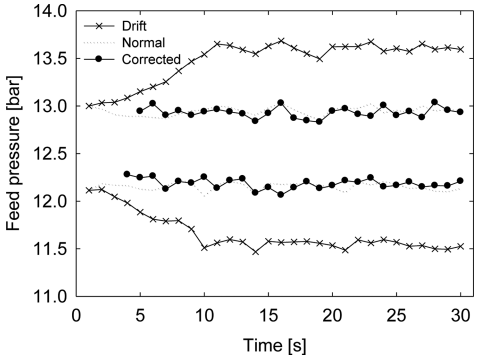

Figure 7. Corrections of feed pressure sensor for sensor readings with sensor drifts up to ±5%, relative to expected normal operation values and using the individual sensor thresholds.

that sensor drifts can be mistaken for natural fluctuations of sensor readings when sensor deviations are small. However, the SFDI-SVR approach enabled detection of small sensor deviations (3.45%−5.56%) with a reasonable accuracy (see Tables 2 and 5).

4.2. Performance of SFDI-SOM. In the first step of the SFDI-SOM approach, the fault-free RO plant data were utilized to construct a SOM that characterizes the normal plant operational states. This resulted in a 240-unit (20 rows and 12 columns) SOM with quantization and topographic errors of 0.397 and 0.214, respectively[42] for which a hexagonal grid was selected for the 2D map representation. Although the shape of

Table 5. Performance of SOM-Based Abnormal Behavior Detection and Isolation Tool for Induced Deviations from −50% to +50%

	Percent of Faulty Sensor Data Required for Fault Confirmation[a]					
	13%		20%		33%	
deviation (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)
+50	0	0	0	0	0	0
+40	0	0	0	0	0	0
+30	0.62	0.15	0.62	0	0.93	0
+20	1.7	0.31	1.7	0.15	1.85	0
+10	6.64	1.54	7.41	0.15	8.02	0
−10	3.24	4.17	3.55	1.54	4.17	0
−20	0.46	1.85	0.46	0	0.46	0
−30	0.15	1.85	0.15	0	0.15	0
−40	0	0.46	0	0	0	0
−50	0	0.46	0	0	0	0

[a]Percent of faulty data readings (false negatives (FN) and false positives (FP)) of the total data trace required for fault confirmation.

the complete SOM is toroidal, in the current work, for the sake of clarity, the SOM is presented as a series of 2D component plane maps (see Figures 4 and 8). In order to confirm detection of abnormal sensor behavior (see section 3.4), a minimum percentage of faulty readings (MPFR), of the total sensor readings in a prescribed period, was required to be identified as faulty. Accordingly, the performance of the SFDI-SOM method was evaluated, for MPFR (of the total data trace; see section 3.1) in the range of 13%−33% and sensor deviations (from normal operation) in the range of ±10% to ±50% (see Table 5). In contrast to the SFDI-SVR approach, SFDI-SOM had a lower level of detection provision, whereby deviations below ±10% resulted in a significant number of FN detections; thus, the SFDI-SOM approach was only evaluated for sensor deviations of ±10% or above. While the above is a limitation, the SFDI-SOM approach has other advantages, as discussed later in this section.

The number of undetected sensor faults (i.e., FN readings) was greater with increasing requirement of higher percentage of faulty data (of the total trace data) for fault confirmation, while the number of false positive (FP) detections decreased. For example, for sensor deviations of +20%, FN detections increased from 1.7% to 1.85% as the percentage of trace readings declared as faulty increased from 13% to 33%, while FP detections decreased from 0.31% to none for the same range. As expected, the number of FN and FP detections increased with smaller sensor deviations (from normal behavior), as it was more challenging for SFDI-SOM to identify such faults. For example, for sensor deviations of +10%, the percent FN and FP detections were as high as 6.64% and 1.54%, respectively (Table 5), based on the MPFR requirement of at least 13% for fault confirmation. As is evident from Table 5, one can tune the acceptance criterion of percentage of false negative relative to false positive detections by tuning the MPFR requirement for fault confirmation. Here, we note again (as in section 4.1), that the occurrence of FN detections is more problematic than FP detections; hence, one should establish fault confirmation criteria based on the acceptable balance between FN and FP detections.

Once the faulty sensors were detected and isolated, data reconciliation was carried out by replacing the faulty values with the values associated with the centroids of the bmu values (see

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

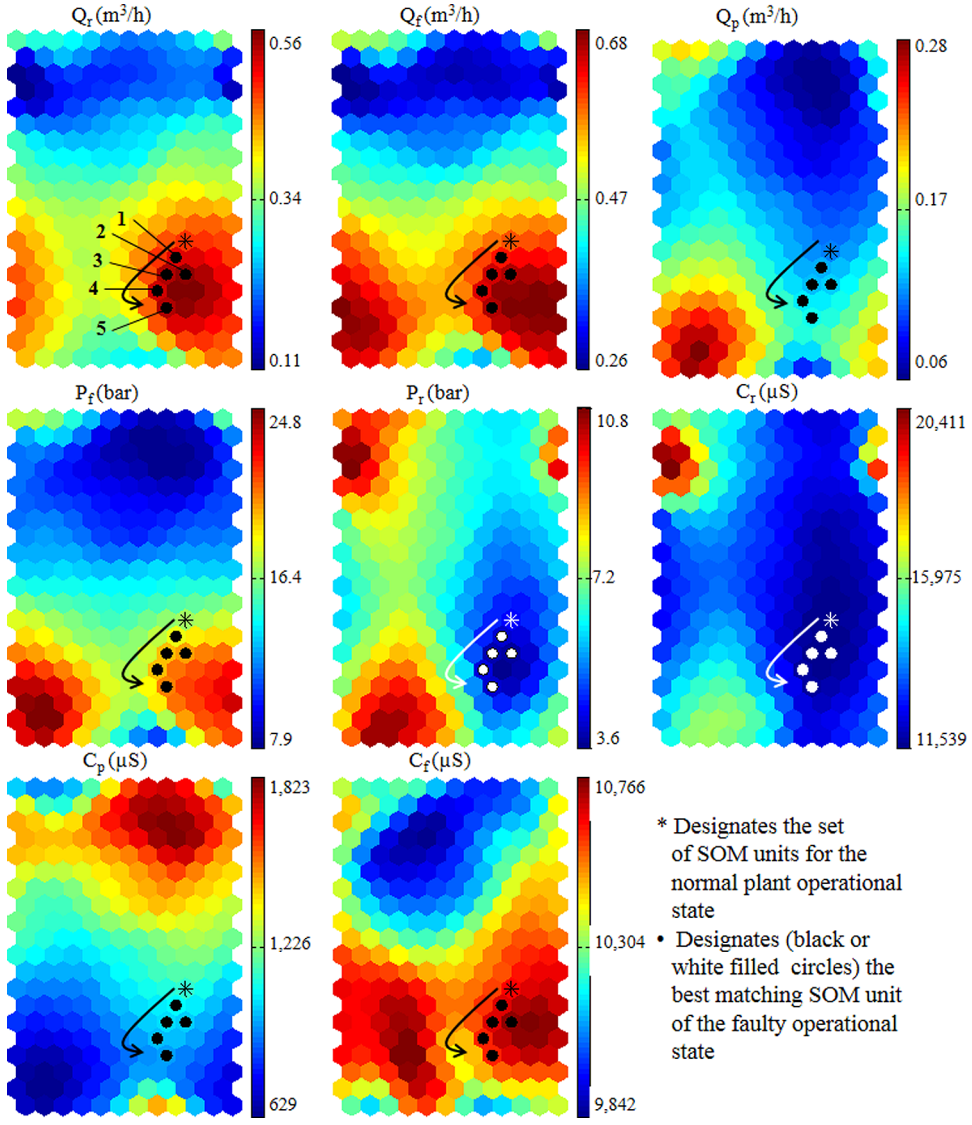

Figure 8. Plant behavior trajectory for a data trace with a drift in the feed rate sensor reading. The deviation from the nominal value increases linearly with time up to a deviation of +50%. Note: The trajectory of the best matching unit (bmu) for the faulty data vector with time is marked as 1−5 (at successive 5-s intervals).

section 3.4.4) corresponding to the reduced data vectors and discarding the faulty sensor data. The differences between the corrected values and the expected values (known from the experimental data) were calculated by means of the AARE values, being, in all cases, smaller than 5%. The accuracy of data correction (of faulty sensors) via the SFDI-SOM approach was in the AARE range of 0.19%-4.22% for the eight plant sensors (Table 6). The above level of accuracy for faulty data correction is somewhat lower than that which was achieved (0.02%−0.82% AARE) for the SFDI-SVR approach. However, the SFDI-SOM is simpler to use, because it does not require individual sensor models but only a single SOM construction, based on the fault-free plant operational data.

An attractive element for using SOM, which can make its use beneficial as either an add-on to SFDI-SVR or SFDI-SOM is in that SOM can provide a visual representation of sensor deviations from normal plant operation. For example, based on historical plant data used to construct the plant SOM and the desired operational domain, one can establish visually the location of a given plant operation state over the SOM. When the plant operates under normal conditions, the bmu values representing plant data are located in a contiguous SOM area corresponding to normal plant operation. In contrast, when sensor readings are faulty, significant drifts of the bmu are observed. As an illustration, Figure 8 shows one example of the effect of an increasing perturbation up to +50% for the feed

Table 6. Accuracy of Corrections Using the SFDI-SOM Tool

corrected variable	AAE[a]	AARE[b] (%)
feed conductivity (μS)	20.31	0.19
feed flow rate (m³/h)	0.01	2.95
feed pressure (bar)	0.42	2.75
retentate conductivity (μS)	283	1.96
retentate flow rate (m³/h)	0.01	4.22
retentate pressure (bar)	0.26	3.89
permeate conductivity (μS)	40.81	3.37
permeate flow rate (m³/h)	0.02	3.96

[a]AAE = average absolute error. [b]AARE = average absolute relative error.

dx.doi.org/10.1021/ie403603x | Ind. Eng. Chem. Res. 2014, 53, 3257−3271

flow rate sensor. Each subfigure represents the so-called component planes of the SOM, which are generated by plotting each individual component of the centroid of the SOM units. Accordingly, each component plane depicts the distribution of the sensor readings for each of the eight plant sensors (i.e., retentate, feed and permeate flow rates, feed and permeate pressures, and retentate, permeate and feed conductivities) over the SOM grid. Note that each component plane corresponds to a slice of the SOM and therefore the overlaying SOM units are all for the same plant operational state (see Figure 4). In the example of Figure 8, normal plant operation associated with SOM unit marked with an asterisk (*) and the deviated operational state resulting from information provided by the faulty feed flow rate is depicted by black or white circles. Figure 8 shows the bmu of the faulty data moving away (i.e., trajectory) from the location corresponding to the normal operation bmu (i.e., bmu of the vector with fault-free operational parameters). The drift in sensor readings is apparent in the example of Figure 8, where the bmu associated with the faulty data vector drifts, along the shown trajectory of increasing time (at 5-s intervals), further away from the bmu associated with the expected normal plant operation (marked with an asterisk). The above depiction of the state of the plant and sensor deviation can provide a rapid and visual aid for plant operators to monitor plant operational status and to detect abnormal plant behavior. The above visual approach does not identify the specific faulty sensor, but can provide an early warning, which could then proceed with the process of faulty sensor identification and isolation (by either SFDI-SVR or SFDI-SOM), as described in sections 3.3.2 and 3.4.3.

CONCLUSIONS

Sensor fault detection and isolation (SFDI), as well as sensor data corrections approaches for spiral-wound reverse-osmosis (RO) plant operation were assessed based on the use of both support-vector regression (SVR) plant sensor models and self-organizing map (SOM) representation of plant operation. SFDI-SVR was based on data-driven models developed for individual RO plant sensors using the support vector regression model building approach. SFDI-SVR enabled fault detection without false negative detections for sensor deviations in the range of ±10% or greater when sensor thresholds (global or individual) were just below 5%. False positive (FP) detections were higher than false negative (FN) detections by up to a factor of 2 and greater in some cases, particularly for increasing global sensor detection threshold and small sensor deviations (e.g., approaching the expected normal sensor fluctuations under normal steady-state conditions). Overall, individual sensor thresholds provided lower false negative detection rate for all sensor deviations, but false positive detections appeared to be nearly invariant with respect to sensor threshold in the range of 3.5%−5%. Corrections of faulty sensor readings were within the accuracy level of the SVR sensor models (based on evaluation of sensor performance for normal plant operation). The SFDI-SOM's performance was inferior to SFDI-SVR, demonstrating no FN detections only for sensor deviations in the range of ±40% or higher, with increased FN detections of 3.24%−6.64% for sensor deviations in the range of ±10% with SOM minimum percentage of faulty readings (MPFR) of 13% (i.e., percent of sensor reading in a given monitoring trace being identified as faulty) for fault confirmation and somewhat higher (4.17%−8.02%) for MPFR of 33%. Corrections of the abnormal sensor readings with SFDI-SOM were with average absolute relative error (AARE) values of <5%. Although SFDI-SOM was of lower accuracy than SFDI-SVR, its application is simpler, since it does not require a model to be developed for each of the plant sensors. Moreover, SOM provides a pictorial representation of plant operation and trajectory of faulty sensor behavior that could be beneficial to plant operators. Overall, both approaches were robust and able to handle noisy plant data. Clearly the choice of using either one or integration of both of the above approaches for RO plants would have to be made on the basis of acceptability of FN, relative to FP fault detections and desirability for visual representation of plant operation that is feasible via SOM.

AUTHOR INFORMATION

Corresponding Author

*Tel.: 310-825-8766. E-mail: yoram@ucla.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by funding from the Catalan Government (No. 2009 SGR 1529), the Spanish Ministry of Education (No. CTQ2009-14627), the California Department of Water Resources, U.S. Bureau of Reclamation, the United States Environmental Protection Agency, the National Water Research Institute, the Office of Naval Research, and the UCLA Water Technology Research Center.

NOMENCLATURE

AAE = average absolute error
AARE = average absolute relative error
AD = absolute difference
ANN = artificial neural network
b = bias vector
bmu = best matching unit
C_f = feed conductivity
C_p = permeate conductivity
C_r = retentate conductivity
CE = centroid
$cond_{feed}$ = feed conductivity sensor reading
FDI = fault detection and isolation
FDS = fault detection SOM
FN = false negative
FP = false positive
FTC = fault tolerant control
$f(x)$ = function obtained by SVR
m = number of sensors not including feed conductivity
n = number of experimental values
P_f = feed pressure
P_r = retentate pressure
PARD = percent absolute relative difference
PCA = principal component analysis
Q_f = feed flow rate
Q_p = permeate flow rate
Q_r = retentate flow rate
RO = reverse osmosis
r^2 = correlation coefficient
s = sensor under consideration
SFDI = sensor fault detection and isolation
SOM = self-organizing-map
STDV = standard deviation

SVM = support vector machines
SVR = support vector regression
T = threshold
TAD = total absolute difference
TDS = total dissolved solids
TPARD = total percent absolute relative difference
V_i = average variable value for the given steady state
VFD = variable frequency drive
$\mathbf{w}$ = normal vector to hyperplane
$\mathbf{x}$ = vector of input RO parameters in the SVR model
$\mathbf{y}$ = vector of input parameters in the SOM
y_i = experimental value for a given variable i
y_i^* = predicted value for a given variable i
$y_i^{(n)}$ = normalized value for a given variable i

Greek Symbols

ε = tolerance of the SVR models
γ = regularization parameter
ξ = upper slack variable
ξ^* = lower slack variable
σ = variance/width of the Gaussian
Φ = kernel function

REFERENCES

(1) Jamal, K.; Khan, M. A.; Kamil, M. Mathematical modeling of reverse osmosis systems. *Desalination* **2004**, *160* (1), 29−42.
(2) Abbas, A. Model predictive control of a reverse osmosis desalination unit. *Desalination* **2006**, *194* (1−3), 268−280.
(3) Bartman, A. R.; Christofides, P. D.; Cohen, Y. Nonlinear Model-Based Control of an Experimental Reverse-Osmosis Water Desalination System. *Ind. Eng. Chem. Res.* **2009**, *48* (13), 6126−6136.
(4) Bartman, A. R.; McFall, C. W.; Christofides, P. D.; Cohen, Y. Model-predictive control of feed flow reversal in a reverse osmosis desalination process. *J. Process Control* **2009**, *19* (3), 433−442.
(5) Venkatasubramanian, V.; Rengaswamy, R.; Yin, K.; Kavuri, S. N. A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. *Comput. Chem. Eng.* **2003**, *27* (3), 293−311.
(6) Mhaskar, P.; Liu, J.; Christofides, P. D. *Fault-Tolerant Process Control: Methods and Applications*; Springer: London, 2013.
(7) Patton, R. J.; Chen, J. Observer-based fault detection and isolation: Robustness and applications. *Control Eng. Pract.* **1997**, *5* (5), 671−682.
(8) Isermann, R. Model-based fault-detection and diagnosis−Status and applications. *Ann. Rev. Control* **2005**, *29* (1), 71−85.
(9) McFall, C. W.; Christofides, P. D.; Cohen, Y.; Davis, J. F. Fault-tolerant control of a reverse osmosis desalination process. In *Proceedings of the 8th IFAC Symposium on Dynamics and Control of Process Systems*, 2007; pp 163−168.
(10) Eski, I.; Erkaya, S.; Savas, S.; Yildirim, S. Fault detection on robot manipulators using artificial neural networks. *Robotics Comput.-Integrated Manuf.* **2011**, *27* (1), 115−123.
(11) Chilin, D.; Liu, J.; Davis, J. F.; Christofides, P. D. Data-based monitoring and reconfiguration of a distributed model predictive control system. *Int. J. Robust Nonlinear Control* **2012**, *22* (1), 68−88.
(12) Jack, L.; Nandi, A. Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms. *Mech. Syst. Signal Process.* **2002**, *16* (2), 373−390.
(13) Samanta, B.; Al-Balushi, K.; Al-Araimi, S. Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. *Eng. Appl. Artif. Intell.* **2003**, *16* (7), 657−665.
(14) Gambier, A.; Miksch, T.; Badreddin, E. A reverse osmosis laboratory plant for experimenting with fault-tolerant control. In *American Control Conference, 2009 (ACC'09)*; IEEE: Piscataway, NJ, 2009; pp 3775−3780.
(15) McFall, C. W.; Bartman, A.; Christofides, P. D.; Cohen, Y. Control and Monitoring of a High Recovery Reverse Osmosis Desalination Process. *Ind. Eng. Chem. Res.* **2008**, *47* (17), 6698−6710.
(16) Garcia-Alvarez, D.; Fuente, M. J. A UPCA-based monitoring and fault detection approach for reverse osmosis desalination plants. *Desalin. Water Treat.* **2013**, 1−15.
(17) Yang, G.-h.; Zhang, S.-Y.; Lam, J.; Wang, J. Reliable control using redundant controllers. *Autom. Control, IEEE Trans. Autom. Control* **1998**, *43* (11), 1588−1593.
(18) Šiljak, D. D. Reliable control using multiple control systems. *Int. J. Control* **1980**, *31* (2), 303−329.
(19) *EcosimPro*; EA International: Madrid, Spain.
(20) Arteaga, F.; Ferrer, A. Framework for regression-based missing data imputation methods in on-line MSPC. *J. Chemom.* **2005**, *19* (8), 439−447.
(21) Pascual, X.; Gu, H.; Bartman, A. R.; Zhu, A.; Rahardianto, A.; Giralt, J.; Rallo, R.; Christofides, P. D.; Cohen, Y. Data-driven models of steady state and transient operations of spiral-wound RO plant. *Desalination* **2013**, *316*, 154−161.
(22) Khayet, M.; Cojocaru, C.; Essalhi, M. Artificial neural network modeling and response surface methodology of desalination by reverse osmosis. *J. Membr. Sci.* **2011**, *368* (1−2), 202−214.
(23) Libotean, D.; Giralt, J.; Giralt, F.; Rallo, R.; Wolfe, T.; Cohen, Y. Neural network approach for modeling the performance of reverse osmosis membrane desalting. *J. Membr. Sci.* **2009**, *326* (2), 408−419.
(24) Abbas, A.; Al-Bastaki, N. Modeling of an RO water desalination unit using neural networks. *Chem. Eng. J.* **2005**, *114* (1−3), 139−143.
(25) Narasimhan, S.; Jordache, C. Industrial Applications of Data Reconciliation and Gross Error Detection Technologies. In *Data Reconciliation and Gross Error Detection*; Gulf Professional Publishing: Houston, TX, 1999; Chapter 11, pp 327−372.
(26) Batur, C.; Ling, Z.; Chien-Chung, C. In Support vector machines for fault detection. In *Proceedings of the 41st IEEE Conference on Decision and Control*, Dec. 10−13, 2002; Vol. *2*, pp 1355−1356.
(27) Svensson, M.; Byttner, S.; Rognvaldsson, T. In Self-organizing maps for automatic fault detection in a vehicle cooling system. In *Intelligent Systems, 2008 (IS '08), 4th International IEEE Conference*, Sept. 6−8, 2008; pp 24-8−24-12.
(28) Cottrell, M.; Gaubert, P.; Eloy, C.; François, D.; Hallaux, G.; Lacaille, J.; Verleysen, M. Fault Prediction in Aircraft Engines Using Self-Organizing Maps. In *Advances in Self-Organizing Maps*, Príncipe, J., Miikkulainen, R., Eds.; Springer: Berlin, Heidelberg, Germany: 2009; Vol. *5629*, pp 37−44.
(29) Bossio, J. M.; De Angelo, C. H.; Bossio, G. R.; Garcia, G. O. In Fault diagnosis on induction motors using Self-Organizing Maps. In *9th IEEE/IAS International Conference on Industry Applications (INDUSCON)*, Nov. 8−10, 2010; pp 1−6.
(30) Fei, Z.; Tielin, S.; Tao, H. In Fault diagnosis of motor bearing using self-organizing maps. In *Proceedings of the Eighth International Conference on Electrical Machines and Systems, 2005 (ICEMS 2005)*, Sept. 29, 2005; Vol. *3*, pp 2411−2414.
(31) Thang, K. F.; Aggarwal, R. K.; McGrail, A. J.; Esp, D. G. Analysis of power transformer dissolved gas data using the self-organizing map. *IEEE Trans. Power Delivery* **2003**, *18* (4), 1241−1248.
(32) Pengju, K.; Birtwhistle, D. Condition assessment of power transformer onload tap changers using wavelet analysis and self-organizing map: field evaluation. *IEEE Trans. Power Delivery* **2003**, *18* (1), 78−84.
(33) Vapola, M.; Simula, O.; Kohonen, T.; Meriläinen, P., Representation and Identification of Fault Conditions of an Anaesthesia System by Means of the Self-Organizing Map. In *ICANN '94*; Marinaro, M., Morasso, P., Eds.; Springer: London, 1994; pp 350−353.
(34) Libotean, D.; Giralt, J.; Rallo, R.; Cohen, Y.; Giralt, F.; Ridgway, H. F.; Rodriguez, G.; Phipps, D. Organic compounds passage through RO membranes. *J. Membr. Sci.* **2008**, *313* (1−2), 23−43.
(35) Bartman, A. R.; Zhu, A.; Christofides, P. D.; Cohen, Y. Minimizing energy consumption in reverse osmosis membrane desalination using optimization-based control. *J. Process Control* **2010**, *20* (10), 1261−1269.

(36) Gu, H.; Bartman, A. R.; Uchymiak, M.; Christofides, P. D.; Cohen, Y. Self-adaptive feed flow reversal operation of reverse osmosis desalination. *Desalination* **2013**, *308*, 63−72.
(37) Bishop, C. M., *Pattern Recognition and Machine Learning*; Springer Science and Business Media, LLC: New York, 2006.
(38) Kohonen, T. *Self-Organizing Maps*, Vol. *30*; Springer: Berlin, Heidelberg, Germany, 2001.
(39) Vogt, M.; Kecman, V. Active-set methods for support vector machines. In *Support Vector Machines: Theory and Applications*, Wang, L., Ed.; Springer: Berlin, 2005; Vol. *177*, pp 133−158.
(40) Vapnik, V. N. *The Nature of Statistical Learning Theory*; Springer−Verlag: New York, 1995; p 188.
(41) De Brabanter, K.; Karsmakers, P.; Ojeda, F.; Alzate, C.; De Brabanter, J.; Pelckmans, K.; De Moor, B.; Vandewalle, J.; Suykens, A. K. *LS-SVMlab Toolbox User's Guide*; K.U. Leuven: Leuven, Belgium, 2011.
(42) Alhoniemi, E.; Himberg, J.; Parhankangas, J.; Vesanto, J. *SOM Toolbox, 2.0*; 2000.