Hybrid Extended Kalman Filtering and Noise Statistics Optimization for Produce Wash State Estimation

Vahid Azimi[†*], Daniel Munther[‡], Seyed Abolfazl Fakoorian[†], Thang Tien Nguyen[†], Dan Simon[†]

[†]Department of Electrical Engineering and Computer Science
‡Department of Mathematics

The authors are with Cleveland State University, Cleveland, OH, USA
*Corresponding author, e-mail: v.azimi@csuohio.edu

Abstract

Food-borne diseases from fresh produce consistently cause serious public health issues. Although sanitization is used to enhance the safety of fresh produce, pathogen cross-contamination via process water continues to be associated with major disease outbreaks. Because of the complex and time-varying nature of the produce wash process, its effectiveness is limited. There is thus an urgent need for new approaches in produce washing to reduce the probability of outbreaks. As an important step in this direction, we design a hybrid extended Kalman filter (HEKF) and a particle swarm optimization (PSO)-based noise statistics optimization algorithm for a produce wash system. The HEKF uses discrete-time free chlorine (FC) measurements, and we use PSO to optimize the noise statistics of the process noise model. We apply the HEKF to estimate *Escherichia coli* O157:H7 contamination of shredded lettuce during an industrial wash, and we dynamically predict the chemical oxygen demand (COD) (RSME 8.24 mg/L), FC concentration (RMSE 0.09 mg/L), *E. coli* concentration (PC) in the wash water (RMSE 0.19 MPN/ml), and *E. coli* level (P) on the lettuce (RSME 0.04 MPN/g). In addition to providing intuitive validation, our sensitivity analysis demonstrates that the estimator has good robustness.

Keywords: Produce wash, Kalman filter, Particle swarm optimization, Sensitivity

I. Introduction

Fresh fruits and leafy green vegetables are a crucial part of a healthy diet. However, produce carries a risk of pathogen contamination, and eating contaminated fresh produce can result in food-borne illness and food poisoning [1], [2]. Therefore, washing fresh-cut produce with sanitizers such as chlorine is considered a crucial factor to improve public health [3].

Recent years have witnessed numerous studies in fresh-cut produce washing. For instance, the effect of free chlorine concentrations in wash water on E. coli O157:H7 reduction, survival, and transference during the washing of fresh-cut lettuce was investigated in [1]. A semi-commercial pilot-scale evaluation of a new process aid, T128, was reported in [2] for its impact on enhancing the antimicrobial efficacy of chlorinated wash water against pathogen survival and cross-contamination. Improvements in water disinfection and sanitation strategies, including a

shower pre-washing step and a final rinse of the produce, were suggested in [3]. The efficacy of five commercial sanitizer treatments against E. coli O157:H7 on iceberg lettuce, in wash water, and on equipment during simulated commercial production in a pilot-scale processing line was assessed in [4]. A mathematical model for pathogen cross-contamination dynamics during produce wash was constructed in [5].

This paper deals with a commercial double wash system with process and measurement noise for washing a mixture of lettuce and spinach [2], [5]. We model the system as a combination of the dynamics of water chemistry and pathogen transmission with four nonnegative states: COD in the water wash, FC concentration, PC in the water wash (which is denoted as X_W), and P on the lettuce (which is denoted as X_L). FC is added to the wash system according to the dosing strategy in [2] and we suppose that FC concentration is the only available measurement during the wash process.

The contribution of this research is the design of an HEKF for state estimation of the produce wash system, and a noise statistics optimization algorithm using PSO. State estimation has not been applied to produce wash prior to this research, but will soon be an essential component of wash systems for the following reasons: (1) state feedback controllers require accurate estimates of the system states; (2) real-time measurements contain errors due to inaccurate sensors; (3) sensors may be expensive and prone to failure; and (4) there are time delays between the sensor measurements and their availability for real-time control. State estimation algorithms can exploit real-time measurements to provide optimal estimates and predictions of the system states [6], [7]. The states of generic constrained nonlinear systems have been estimated using moving horizon estimation and Carleman linearization [8], [9]. The EKF continues to be the most popular state estimation technique for nonlinear systems [10]. The produce washing system is comprised of a continuous-time system and discrete-time measurements, so we design an HEKF for state estimation [6].

HEKF performance is heavily dependent on the noise model; the estimator performance may degrade or even diverge if we model the noise incorrectly. Various approaches have been used to identify unknown noise characteristics to achieve accurate estimation [11]. In [12] an adaptive method for online process and measurement noise identification during joint state and parameter estimation of nonlinear systems was presented. Identification of the noise covariance matrices using an autocovariance least-squares method for linear systems was implemented in [13].

In this research we identify system noise from experimental system measurements, which we then use in our system simulation. However, since the system and the estimator are nonlinear, we may want to use covariance matrices in the estimator that are different than the true covariances. Therefore, we use PSO to optimize the noise statistics to use in our estimator to obtain the best possible estimation performance. In summary, we first identify the system noise offline using experimental data captured during a washing process. We then optimize the noise statistics to use in the filter offline using PSO.

Our contribution can be summarized as follows: (1) we present methods to identify system noise covariances and optimize filter noise covariances offline, which means that online computational effort is minimized compared with online approaches; (2) we employ PSO as a swarm intelligence technique for noise covariance optimization, which potentially converges to the global optimum and avoids local optima; (3) we use only the innovations samples as inputs to our HEKF process noise statistics optimization approach, which are based on experimental measurement data (i.e., FC concentration).

Section II describes the chlorine / cross-contamination dynamic model. Section III presents the HEKF and the noise statistics optimization framework. Section IV presents the PSO algorithm for noise statistics optimization for the HEKF. Our approach uses the fact that in an optimal linear Kalman filter, the measurement innovations is zero-mean white noise with a known covariance; we can thus tune the noise model to achieve theoretically desired HEKF behavior. Section V presents simulation results, which show that the system provides good estimation of the states. PSO reduces the estimation cost function by 34% and drives the innovations as close as possible to zero-mean white noise with the desired covariance. Simulation results provide state estimation with an RMSE of 8.24 mg/L for COD, 0.09 mg/L for FC concentration, 0.19 MPN/ml for PC in the water wash, and 0.04 MPN/g for P on the lettuce. Section V also analyzes the sensitivity of the state estimates to modeling errors to investigate the robustness of the state estimator. The greatest sensitivity is the sensitivity of the estimated COD to the COD increase rate parameter, which is 73; that is, an error of one unit in the modeled COD increase rate results in a 73-fold increase in COD estimation error. The lowest sensitivity is the sensitivity of the estimated lettuce P to the natural decay rate of the FC, which is 0.26. Section VI presents concluding remarks and a discussion of future work.

II. Chlorine / Cross-Contamination Dynamic Model

We consider a dynamic model for a commercial double wash system as illustrated in Fig. 1 [2], [5]. The wash system comprises two separate tanks, where turbulence for washing the product is created by pumping air into both of the tanks. Shredded lettuce and inoculated baby spinach are on a conveyor belt that feeds the primary tank. Spinach leaves and lettuce shreds are mixed in the primary tank, where concentrated sodium hypochlorite (CSH) is added to achieve FC in the wash water.

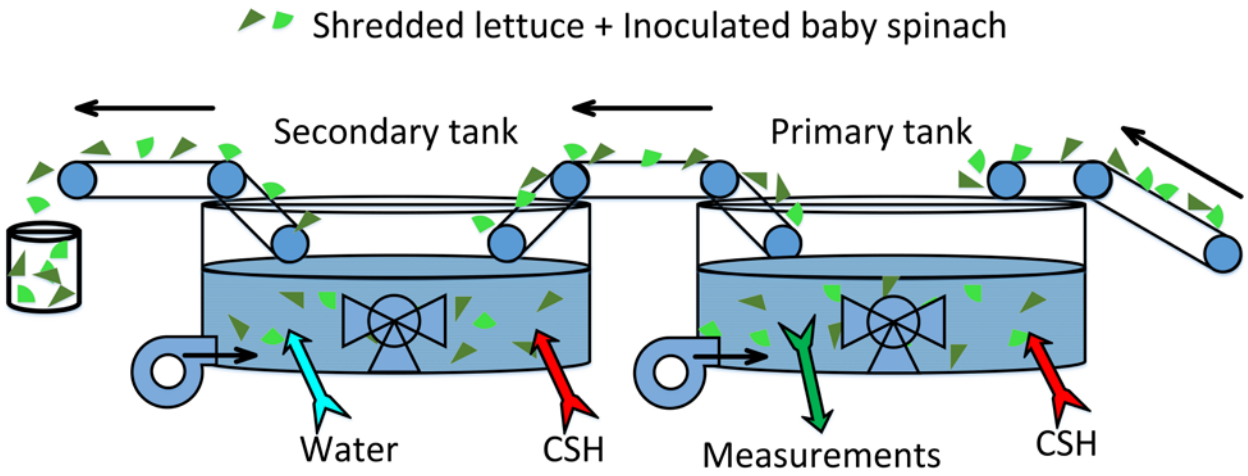

Fig. 1. Commercial double wash system

When fresh-cut produce enters the wash system, the abrupt increase in organic matter from exudates results in a significant decrease in hypochlorous acid and a linear increase in COD, so it is challenging to keep FC concentration relatively constant in the primary tank. Experiments have shown that FC concentration remains stable in the secondary tank [5]. To avoid increasing pathogen cross-contamination, sodium hypochlorite is added to the washing system every $\tau = 12$ minutes.

To compensate for the decrease in hypochlorous acid, we use an FC dosing method with a fixed time period τ. We use three doses with a fixed dosing width (duration) $\tau_0 = 2$ minutes, and the FC rate for each dose is denoted as r_{fc_i}, $i = 1, 2, 3$ [1], [2], [5]. Fig. 2 illustrates the FC dosing strategy over a washing time of 36 minutes.

FC concentration and pathogen can be measured every two minutes during washing using a chlorine photometer and polymerase chain reaction (PCR) respectively. COD is determined via reactor digestion method 10236 [2], [3]. A secondary tank is used to refill water that is lost in the primary wash tank to rewash the fresh-cut produce. The chlorine and cross-contamination dynamics can be modeled as follows [5]:

$$\dot{\underline{x}} = f + gu + w_s(t)$$

$$f(\underline{x}) = \begin{bmatrix} K_0 \\ -\gamma_c x_2 - \beta_c x_1 x_2 \\ \beta_{ws} - \beta_{lw}\left(\frac{L}{V}\right) x_3 - \alpha x_2 x_3 \\ \beta_{lw} x_3 - \alpha x_2 x_4 - C_l x_4 \end{bmatrix} \quad g = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$y_k = x_{2k} + v_{s_k} \tag{1}$$

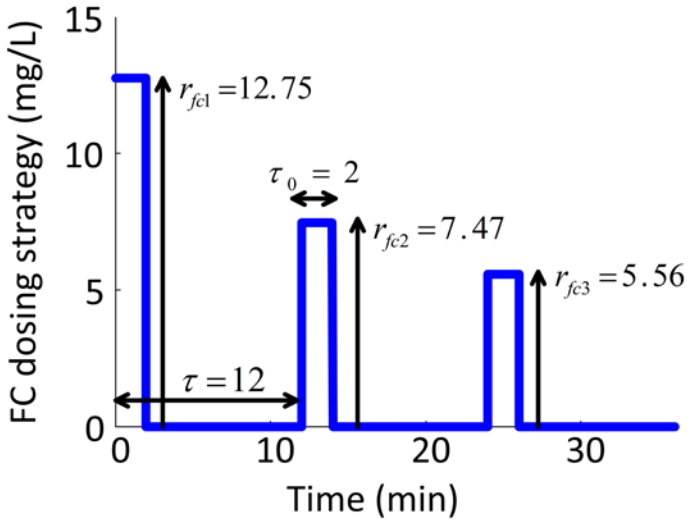

Fig. 2. FC dosing strategy. The magnitude of the FC doses $r_{fc_i}, i = 1, 2, 3$ are calculated in [5] with a least-squares fit to the model of Eq. (1). The dosing width τ_0 and the FC dosing period τ are from [2].

The first two dynamic system equations represent the water chemistry dynamics in the wash tank, and the last two dynamic system equations represent the pathogen contamination dynamics in the wash water and on produce in the tank. $\underline{x} = [x_1 \quad x_2 \quad x_3 \quad x_4]^T = [O \quad C \quad X_W \quad X_L]^T$ denotes the states of the system; O (mg/L) is COD in the water wash; C (mg/L) represents FC concentration; X_W (MPN/ml) is the PC in the water wash; X_L (MPN/g) represents the P on the lettuce; u (mg/l min) represents the FC injection rate to the wash tank (see Fig. 2); y_k represents the discrete-time measurement, which is FC concentration; $w_s(t)$ represents the four-element continuous-time white process noise vector of the system with $n \times n$ covariance Q_s; v_{s_k} represents the scalar discrete-time white measurement noise with $m \times m$ covariance R_{s_k}; and n and m are the number of states and measurements respectively.

In this paper, we consider $R_{s_k} = 1$ in the state estimator; this is because it is only the ratio of Q_s to R_{s_k} (not their absolute values) that determines estimation performance. Note that K_0 and β_c are positive constants and $1/C_l$ denotes the average dwell time for the lettuce in the wash tank. In the next section we present an HEKF to estimate the states of the system based on the measurements and the system model. We summarize the system parameters in Table 1. Now we use experimental data to derive the discrete-time process noise covariance Q_{s_d} which is equivalent to the continuous-time covariance Q_s. Equation (1) shows that the equivalent discrete-time process noise can be approximated as

$$w_{s_m} = \frac{x_{exp_{m+1}} - x_{exp_m} - (f(x_{exp_m}) + gu_m)\Delta t}{\Delta t} \tag{2}$$

where w_{s_m} represents a four-element discrete-time process noise vector for $m = 1, \dots, m_{max}$, where that m_{max} is the number of experimental data points; x_{exp_m} is the m-th experimental data point captured during the washing process (measurement $\Delta t = 2$ min); $f(x_{exp_m})$ is the system dynamics in (1) evaluated at x_{exp_m}; and u_m is the FC dosing strategy shown in Fig. 1 at sample index m. We estimate the covariance Q_{s_d} as

$$Q_{s_d} = \sum_{m=1}^{m_{max}} \frac{w_{s_m}^2}{m_{max}} \tag{3}$$

The continuous-time process noise covariance Q_s can then be obtained as follows [6, Chapter 8]:

$$Q_s = \frac{Q_{s_d}}{\Delta t} \tag{4}$$

We use 19 measured experimental samples of chlorine during the wash process, so $m_{max} = 19$ in Eq. (3). Using Eqs. (2)-(4), we calculate the 4×4 covariance matrix of the system's process noise Q_s from our experimental data as follows:

$$Q_s = \text{diag}(468 \quad 0.64 \quad 1.77 \quad 1.34)$$

Table 1. System parameters used in Eq. (1); see [5] for details

Parameter	Description	Value (Units)
K_0	COD increase rate	$32.3\ (\mathrm{mg/(L\ min)})$
γ_c	Natural decay rate of FC	$1.7 \times 10^{-3}\ (1/\mathrm{min})$
β_c	FC consumption rate	$5.38 \times 10^{-4}(\mathrm{L/mg\ min})$
β_{ws}	Increase rate of the pathogen in the water	$1.95\ (\mathrm{MPN/(ml\ min)})$
β_{lw}	Pathogen binding rate (water to produce)	$0.38\ (\mathrm{ml/(g\ min)})$
L	Amount of lettuce in the tank	$19526\ (\mathrm{g})$
V	Volume of the tank	$3.2 \times 10^6(\mathrm{ml})$
C_l	Reciprocal of average wash time	$2.3\ (1/\mathrm{min})$
α	Interaction rate of the pathogen via FC	$0.5\ (1/(\mathrm{mg\ min}))$

III. Hybrid Extended Kalman Filter And Noise Statistics Optimization

In this section we first design an HEKF to estimate the system states in Eq. (1), and then we present a PSO-based framework for noise statistics optimization so that the HEKF is approximately optimal. The HEKF is applied to the continuous-time dynamic system model in Eq. (1) and the discrete-time measurements (i.e., FC concentration). Consider a general continuous-time system with discrete time measurements [6]:

$$\begin{aligned} \dot{x} &= f(x, u, w_s, t) \\ y_k &= h_k(x_k, v_{s_k}) \\ w_s(t) &\sim (0, Q_s) \\ v_{s_k} &\sim (0, R_{s_k}) \end{aligned} \tag{5}$$

$f(.)$ and $h_k(.)$ represent the nonlinear system and measurement equations; and y_k, x, and u are the system measurement, state, and input, respectively. Recall that we computed continuous-time process noise covariance Q_s in Eq. (4). We initialize the filter as follows:

$$\begin{aligned} \hat{x}_0^+ &= \mathrm{E}[x_0] \\ P_0^+ &= \mathrm{E}[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T] \end{aligned} \tag{6}$$

where $\hat{x}_k^+$ and P_k^+ are the posterior estimate and its covariance, respectively, at time k; x_0 and $\hat{x}_0^+$ are the initial state and estimated state, respectively; $\mathrm{E}(.)$ denotes the expected value operation; and P_0^+ is the covariance of the initial estimate. We integrate the state estimate and its covariance from time $(k-1)^+$ to time k^- as follows [6]:

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}, u, 0, t) \\ \dot{P} &= AP + PA^T + LQ_fL^T \end{aligned} \tag{7}$$

where A and L denote the partial derivatives of $f(x, u, w_s, t)$ with respect to x and w_s respectively, both evaluated at $\hat{x}$; and Q_f denotes the continuous-time process noise covariance of the HEKF algorithm. Note that Q_f is, in general, different from Q_s because Q_s is unknown in practice. Although we identified Q_s Section II, we did so only for one particular experiment, and

only for simulation purposes. In general, Q_s is unknown, and so the HEKF needs to be implemented apart from any knowledge of Q_s.

We start integrating Eq. (7) with $\hat{x} = \hat{x}_{k-1}^{+}$ and $P = P_{k-1}^{+}$ at time k–1, and end the integration at the next measurement time k with $\hat{x} = \hat{x}_k^{-}$ and $P = P_k^{-}$, where $\hat{x}_k^{-}$ and P_k^{-} are the prior estimate and covariance respectively. Fig. 3 provides a conceptual diagram of the prior and posterior estimates and covariances [6]. We incorporate the measurement y_k into the state estimate and estimation covariance at time k as follows:

$$
\begin{aligned}
K_k &= P_k^{-} H_k^T \left(H_k P_k^{-} H_k^T + M_k R_{f_k} M_k^T \right)^{-1} \\
\hat{x}_k^{+} &= \hat{x}_k^{-} + K_k \left(y_k - h_k(\hat{x}_k^{-}, 0, t_k) \right) \\
P_k^{+} &= (I - K_k H_k) P_k^{-} (I - K_k H_k)^T + K_k M_k R_{f_k} M_k^T K_k^T
\end{aligned}
\tag{8}
$$

where H_k and M_k represent the partial derivative of $h_k(x_k, v_{sk})$ with respect to x_k and v_{sk} respectively, both evaluated at $\hat{x}_k^{-}$; K_k is the Kalman filter gain; y_k denotes the measurement at time k; and R_{f_k} represents the discrete-time measurement noise covariance.

Now that the HEKF algorithm has been outlined, we propose a method to optimize the noise statistics of the covariances that are used in the HEKF algorithm. Both measurement and process noise statistics are unknown, and the performance of the HEKF heavily depends on accurate noise models, so we derive a noise model optimization method based on PSO. In other words, we optimize the process noise covariance Q_f and the measurement noise covariance R_{f_k} that is used in the HEKF. We assume that the process and measurement noise are uncorrelated and have diagonal covariance matrices.

We have four states, and FC concentration is the only measurement ($n = 4$, $m = 1$), so the problem is to tune the four diagonal elements of Q_f and the single element of R_{f_k}. For this purpose we use the innovations, which are defined as

$$
r_k = y_k - h_k \hat{x}_k^{-} \tag{9}
$$

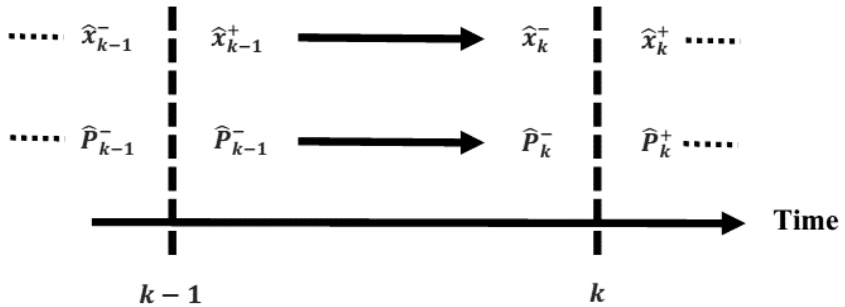

Fig. 3. Prior and posterior estimates and covariances

The innovations can be interpreted as the part of the measurement with new information which is used to update the state estimate in Eq. (8). We employ PSO to identify the diagonal elements of Q_f and R_{f_k} so that the proposed HEKF is approximately optimal, which means the innovations satisfy three different conditions [6]: (1) white noise, (2) zero mean, and (3) covariance of $H_k P_k^- H_k^{\mathrm{T}} + R_{s_k}$. Fig. 4 depicts the structure of the HEKF / noise statistics optimization system.

IV. Particle Swarm Optimization

In this section we use PSO to optimize the process noise statistics for the HEKF; that is, we tune the diagonal elements of the covariance matrices Q_f and R_{f_k} so that the innovations are approximately white noise and zero mean with a covariance of $H_k P_k^- H_k^{\mathrm{T}} + R_{s_k}$.

For the first condition, which is the whiteness of the innovations, we note that Gaussian white noise is a time sequence of uncorrelated random variables. We calculate the lower and upper bounds of a 99% statistical confidence interval for the whiteness of the innovations.

If the sample autocorrelation values are within the 99% confidence bounds, we conclude that the innovations are white. If the autocorrelation values exceed the 99% confidence bounds, the innovations are not white, which means that the diagonal elements of Q_f and R_{f_k} must be changed based on Fig. 4.

We use PSO to minimize the sample autocorrelation values, the mean of the innovations, and the difference between the sample covariance of the innovations and its theoretically desired covariance $H_k P_k^- H_k^{\mathrm{T}} + R_{s_k}$. We also incorporate the RMSE difference between the measured and estimated chlorine in our cost function as a heuristic adjustment to ensure reasonable estimation performance. We thus define four individual cost function components, one for each condition, and we combine them with weighting coefficients to obtain the complete cost function:

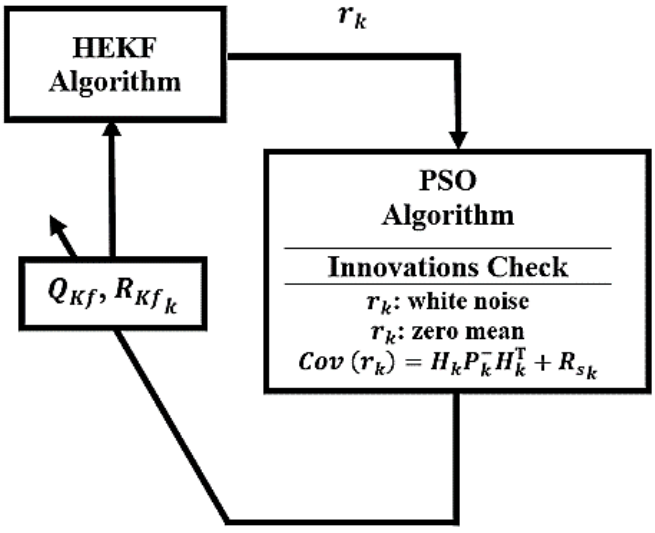

Fig. 4. Proposed structure for the HEKF and noise identification

$$Cost_1 = \frac{1}{g_{max}} \sum_{g=1}^{g_{max}} |R_g|$$

$$Cost_2 = \frac{1}{N} \sum_{k=1}^{N} |r_k|$$

$$Cost_3 = \frac{1}{N} \sum_{k=1}^{N} \left| r_k^2 - (H_k P_k^- H_k^{\mathrm{T}} + R_{s_k}) \right|$$

$$Cost_T = \sum_{e=1}^{3} \zeta_e Cost_e + \zeta_4 RMSE_y \tag{10}$$

where $R_g = \frac{1}{N-g} \frac{\sum_{i=1}^{N-g} (r_i - \bar{r})(r_{i+g} - \bar{r})}{\|r\|^2 \sigma_r^2}$ is the normalized autocorrelation of N innovation samples $\{r_1, \dots, r_N\}$ (N is the number of FC measurement samples); $\bar{r}$ and σ_r^2 denote the sample mean and variance of the innovations; $g = 1, \dots, g_{max}$ represents the lag index of the autocorrelation sample values, where $g_{max} < N$ is the maximum lag; $RMSE_y$ represents the root mean square error of the measured and estimated chlorine data; and ζ_e (e = 1, 2, 3, 4) are weighting coefficients that indicate the relative importance in $Cost_T$ of each individual cost function component.

PSO uses Eq. (10) to optimize the diagonal elements of Q_f and R_{f_k} to minimize $Cost_T$. PSO adjusts the diagonal elements of Q_f and R_{f_k} so that the sample autocorrelation values are within 99% confidence intervals so that the innovations look like a white noise process; this is the first condition. For the second condition, PSO minimizes the innovations mean through $Cost_2$. For the third condition, PSO minimizes the difference between the covariance of the innovations and the desired covariance through $Cost_3$. Finally, $Cost_T$ blends these cost functions and the RMSE of the chlorine fit.

Overview of PSO

PSO can be considered as a population-based optimization algorithm that is motivated by the social behavior of animals for achieving a specific goal [14], [15]. In this paper, we use a PSO algorithm to find the optimal candidate solution, also called an individual, which is comprised of the diagonal elements of the HEKF covariance matrices and minimizes the cost function $Cost_T$ in (10). We first generate a population (swarm) of individuals which are assigned random initial positions that represent candidate solutions. Their velocities represent the directions that they move through the search space. We update their velocities based on various factors, such as the difference between their best positions in the past and their current positions, and the difference between the best positions of their neighbors and their current positions. We then update the position of the individuals using their velocities. If the cost function in (10) for an individual (which corresponds to the diagonal elements of Q_f and R_{f_k}) is less than the cost function for its neighbor, the individual's best known position is updated. We define a maximum number of iterations as the stopping criteria. We initialize the population randomly and we initialize all velocity vectors to zero. We present the velocity update equation as follows [16]:

$$v_i(\kappa+1) = \Upsilon v_i(\kappa) + C_1\big(b_i(\kappa) - z_i(\kappa)\big) + C_2\big(s_i(\kappa) - z_i(\kappa)\big), \qquad i = 1, \dots, l \tag{11}$$

$$\Upsilon = K\,\Upsilon_d \;\;,\;\; C_1 = K\theta_1 \;\;,\;\; C_2 = K\theta_2 \tag{12}$$

Eq. (11) is a *D*-element vector operation; *D* is dimension of the optimization problem; $\kappa = 1,\dots, \kappa_{max}$ where κ_{max} is the maximum number of generations; $i=1,\dots,l$ where l is the number of particles (individuals); $z_i(\kappa)$, $v_i(\kappa)$, and $b_i(\kappa)$ are the position, velocity, and prior best position of the i^{th} particle at the κ^{th} generation; $s_i(\kappa)$ is the best position of the i^{th} neighborhood at the κ^{th} generation; $v_i(\kappa+1)$ is the velocity of the i^{th} particle at the $(\kappa+1)^{st}$ generation; θ_1 and θ_2 represent the cognition learning rate and the social learning rate respectively, which are random numbers distributed in $[0,\ \theta_{1,max}]$ and $[0,\ \theta_{2,max}]$ respectively, where $\theta_{1,max}$ and $\theta_{2,max}$; Υ is the inertia weight, which decreases from 0.9 in the first generation to 0.1 in the last one; Υ_d is the damping ratio of Υ; and C_1 and C_2 are velocity update coefficients. *K* is the constriction coefficient which we set as follows for stability of the algorithm [16]:

$$K < \frac{2}{\theta_{1,max} + \theta_{2,max} - 2 + \sqrt{\big(\theta_{1,max} + \theta_{2,max}\big)^2 - 4\big(\theta_{1,max} + \theta_{2,max}\big)}} \tag{13}$$

We apply velocity limiting to the i^{th} particle as follows:

$$v_{ij}(\kappa+1) \leftarrow \begin{cases} v_{ij}(\kappa+1) & , \ \big|v_{ij}(\kappa+1)\big| \le v_j^{max} \\ v_j^{max}\,\mathrm{sgn}(v_{ij}(\kappa+1)) & , \ \big|v_{ij}(\kappa+1)\big| > v_j^{max} \end{cases} \tag{14}$$

where $j=1,\dots,D$; and v_j^{max} is the maximum velocity, which we define as

$$v_j^{max} = \sigma(z_j^{max} - z_j^{min}) \tag{15}$$

where z_j^{max} and z_j^{min} are the minimum and maximum values of the j^{th} dimension of the search domain; and σ is a scale factor. After the velocity update of Eqs. (11)-(15), we formulate the position update and position limiting as follows:

$$z_i(\kappa+1) = v_i(\kappa+1) + z_i(\kappa) \tag{16}$$

$$z_{ij}(\kappa+1) \leftarrow \min(z_{ij}(\kappa+1),\ z_j^{max}) \tag{17}$$

$$z_{ij}(\kappa+1) \leftarrow \max(z_{ij}(\kappa+1), z_j^{min})$$

where Eq. (16) is a *D*-element vector operation, and $z_i(\kappa+1)$ is the position of the i^{th} particle at the $(\kappa+1)^{st}$ generation.

V. Simulation Results

We use PSO to identify the process and measurement noise statistics for HEKF so the innovations form a zero-mean white noise process with a covariance of $H_k P_k^- H_k^{\mathrm{T}} + R_{s_k}$. The initial state of the system [2] is given as $\underline{x}_0{}^T = [300 \quad 0 \quad 0 \quad 0]$, and the washing period is 36 min. The initial covariance of the estimation error $P_0 = 10\, I_{4\times 4}$, and the maximum number of lags $g_{max} = 5$.

Table 2 presents the parameters that we use in the PSO algorithm and the cost function of Eq. (10). The PSO parameters in Table 2 are chosen to provide fast convergence and achieve small value for $Cost_T$. Note that we could consider a larger number of iterations κ_{max} and a larger population size l to possibly obtain better performance, but better performance would come at the expense of more computational time. The initial covariance of the estimation error P_0 and the maximum number of lags g_{max} were chosen to achieve the lowest value for the total cost function in Eq. (10). The cost function coefficients ζ_e were chosen to make the individual cost function components $Cost_1$, $Cost_2$, $Cost_3$, and $RMSE_y$ of the same order of magnitude so the PSO algorithm would weight them all approximately equally.

Recall that we computed the continuous-time process noise covariance Q_s in Section II. The standard deviation of the simulation process noise is the square root of the covariance matrix $SD_{Qs} = \sqrt{Q_s} =$ diag$(21.65 \quad 0.79 \quad 1.33 \quad 1.15)$ and the standard deviation of the measurement noise is $SD_{R_{s_k}} = \sqrt{R_{s_k}} = 1$. We need to bound the PSO search domain of the standard deviation of the HEKF process and measurement noise. We selected the search domains to comfortably include the standard deviations of the simulation noise, which are the square roots of the system simulation noise covariance matrices Q_s and R_{s_k}. The search domains were chosen as

Table 2. PSO parameters

Parameter	Description	Value
D	Optimization problem dimension	5
κ_{max}	Number of iterations	20
l	Population size	20
$\theta_{1,max}$	Maximum rate for cognition learning	2.05
$\theta_{2,max}$	Maximum rate for social learning	2.05
γ_d	Damping ratio for inertia rate	0.9
σ	Scale factor	0.1
ζ_1	$Cost_1$ coefficient	100
ζ_2	$Cost_2$ coefficient	1
ζ_3	$Cost_3$ coefficient	1
ζ_4	$RMSE_y$ coefficient	1
g_{max}	Maximum number of lag	5
N	Number of innovations samples	3601

$sq_{11} \in [0, \quad 40]$, $sq_{22} \in [0, \quad 2]$, $sq_{33} \in [0, \quad 3]$, $sq_{44} \in [0, \quad 3]$
$SD_{R_{f_k}} \in [0.1, \quad 2]$

PSO decreases $Cost_T$ of Eq. (10) from 4.73 at the initial generation to 3.08 at the 20th generation, which means that the total cost improves by 34%. Fig. 5 shows the best cost values of the PSO population over 20 generations. The best solution found by PSO is given as $SD_{Qf} = \text{diag}(28.32 \quad 0.28 \quad 1.40 \quad 1.41)$ and $SD_{R_{s_k}} = 1.33$. The individual cost functions values from Eq. (7) for this solution are given as $Cost_1 = 0.008$, $Cost_2 = 0.79$, $Cost_3 = 1.38$, and $RMSE_y = 0.98$.

Fig. 6 compares the states of the produce wash system with the estimated HEKF states for two situations: zero initial state estimation error (i.e., $\underline{x}_0{}^T = [300 \quad 0 \quad 0 \quad 0]$), and non-zero initial state estimation error (i.e., $\underline{x}_0{}^T = [350 \quad 10 \quad 5 \quad 0.5]$). Fig. 6 shows that with zero initial state estimation error the HEKF can accurately estimate the states of the system with an RMSE value of 8.24 mg/L for COD, 0.09 mg/L for FC, 0.19 MPN/ml for X_W, and 0.04 MPN/g for X_L. In the case of non-zero initial state estimation error the HEKF estimates the states of the system with an RMSE value of 18.29 mg/L for COD, 0.12 mg/L for FC, 0.63 MPN/ml for X_W, and 0.1 MPN/g for X_L. We see that the HEKF demonstrates good robustness against initial state estimation error.

Fig. 7 shows the Kalman gains. We observe that the Kalman gain of the second state, i.e., FC concentration, converges to zero as time goes to infinity. This implies that we eventually gain sufficient knowledge from our measurements and additional ones do not provide us with any new information, so we ignore additional measurements. On the other hand, the Kalman gains of the first, third, and fourth states do not converge to zero, so those estimations are more responsive to measurements. Note that we heuristically limited the magnitude of the first Kalman gain to 0.5 to obtain robust state estimation for COD.

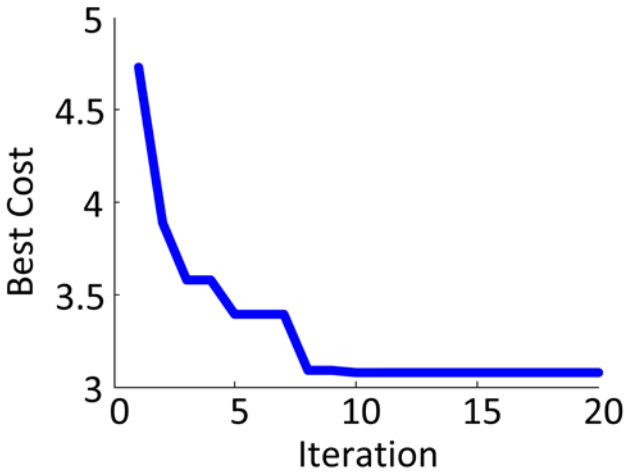

Fig. 5. PSO performance for identifying the process and measurement noise covariances

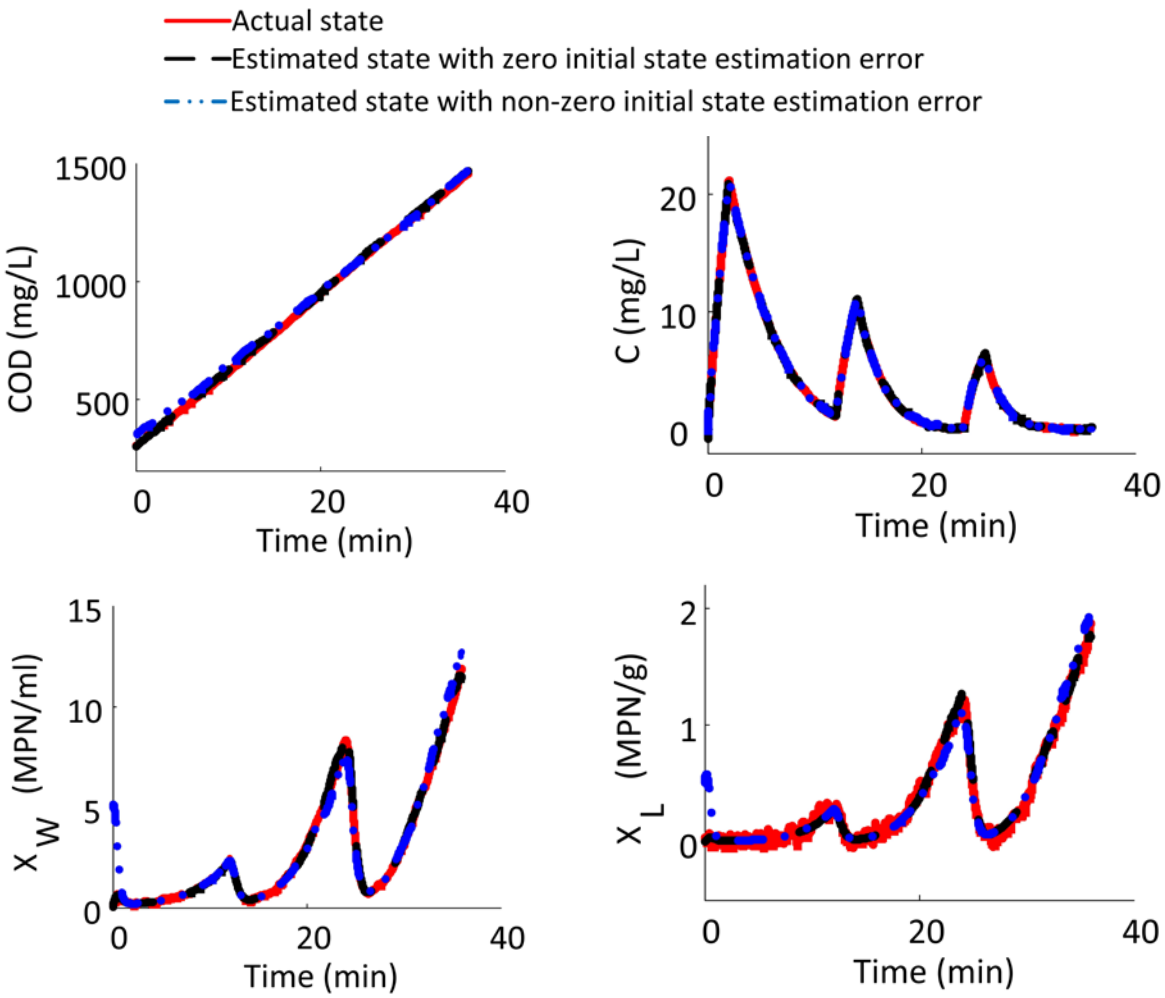

Fig. 6. State estimation performance: simulated states (red solid line), estimated states with zero initial estimation error (black dashed line), and estimated states with non-zero initial estimation error (blue dotted-dashed line)

Fig. 8 demonstrates the innovations of Eq. (9). Recall that the proposed PSO algorithm identifies the HEKF process and measurement noises so that the innovations are zero-mean white noise with a covariance of $H_k P_k^- H_k^{\mathrm{T}} + R_{s_k}$. The figure shows that the innovations appear to be approximately zero mean, and that they appear to be white, although the covariance cannot be easily identified from the figure.

Fig. 9 shows that the sample autocorrelation values of the innovations are bounded within the 99% confidence interval, except at zero lag. We conclude that the innovations are approximately white.

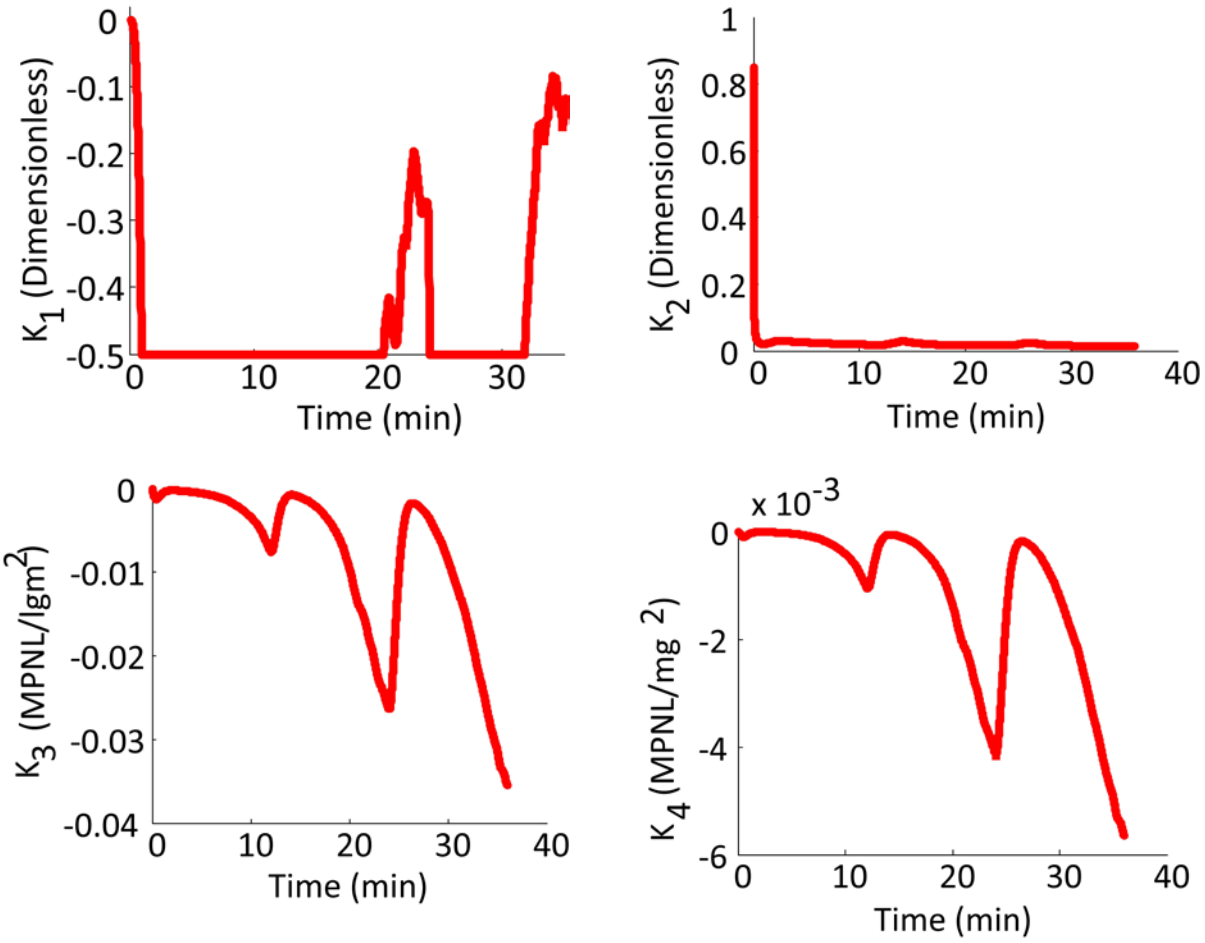

Fig. 7. Kalman gains

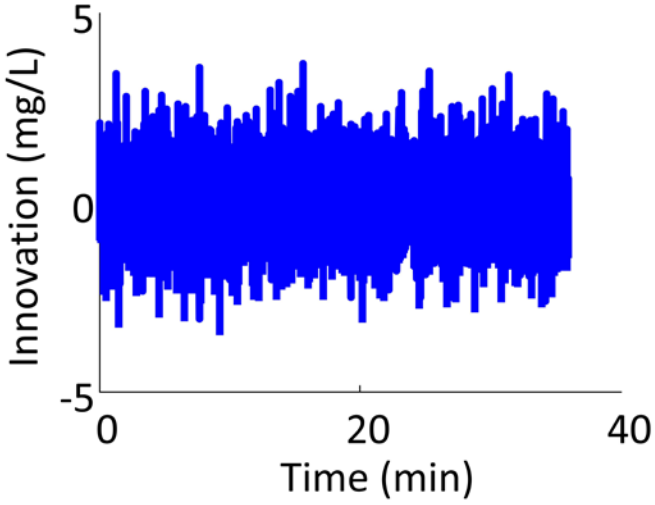

Fig. 8. Innovations

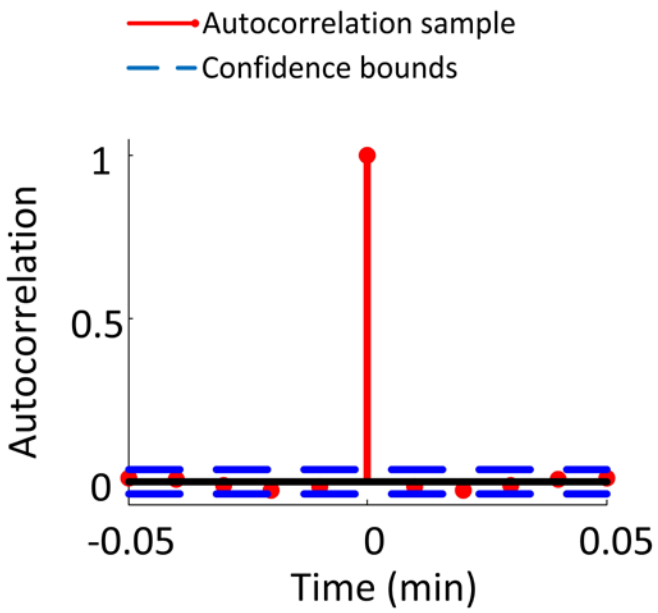

Fig. 9. White noise test of the innovations: autocorrelation of the innovations (red solid circle), and 99% confidence bounds (blue dashed lines)

Monte Carlo simulation provides an additional tool that can provide us with confidence in our estimates. Here we measure the sample standard deviation of the state estimation error (S_e) over multiple Monte Carlo simulations. We calculate the average standard deviation of the estimation error for the j^{th} state over N_{mc} Monte Carlo simulations as follows:

$$S_{e_j} = \sqrt{\frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} \left(x_{ji} - \hat{x}_{ji}\right)^2} \tag{18}$$

where x_{ji} and $\hat{x}_{ji}$ $(j = 1, \dots, 4)$ denote the state and estimated state at the i^{th} time step .

Fig. 10 shows the average standard deviation of the estimation error for each state over 100 Monte Carlo simulations ($N_{mc} = 100$), each with a 36 min simulation time. It is seen that the fluctuations of the average standard deviations of the state estimation errors are bounded within reasonable ranges: 5 mg/L for COD, 0.1 mg/L for FC, 0.1 MPN/ml for PC, and 0.025 MPN/g for P on the lettuce. The state estimates are clustered around the true states and the HEKF provides accurate estimates of the system states.

Sensitivity Analysis

Robustness is a basic requirement for a state estimator. The question that we address here is whether our HEKF and noise identification algorithm provide robustness to system parameter deviations (that is, modeling errors) and are thus reliable as a practical methodology for real-world plant wash systems. We perform a robustness test by varying all of the system parameters, one at a time, and calculating the sensitivity of each state estimation error relative to each

perturbed system model parameter. This will quantify the robustness of our proposed method to system parameter uncertainties. We propose the following equations to quantify the effect of system model parameters on state estimation errors:

$$\bar{S}_j^{P_l\delta} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(S_{e_j}^{P_l\delta}(i)\right)^2} \tag{19}$$

$$D_j^{P_l} = \frac{\left(\frac{\bar{S}_j^{P_l+\delta} + \bar{S}_j^{P_l-\delta}}{2}\right) - \bar{S}_j^{P_{l_0}}}{\bar{S}_j^{P_{l_0}}} \times 100 \tag{20}$$

$$SE_j^{P_l} = \frac{D_j^{P_l}}{\delta\%} \tag{21}$$

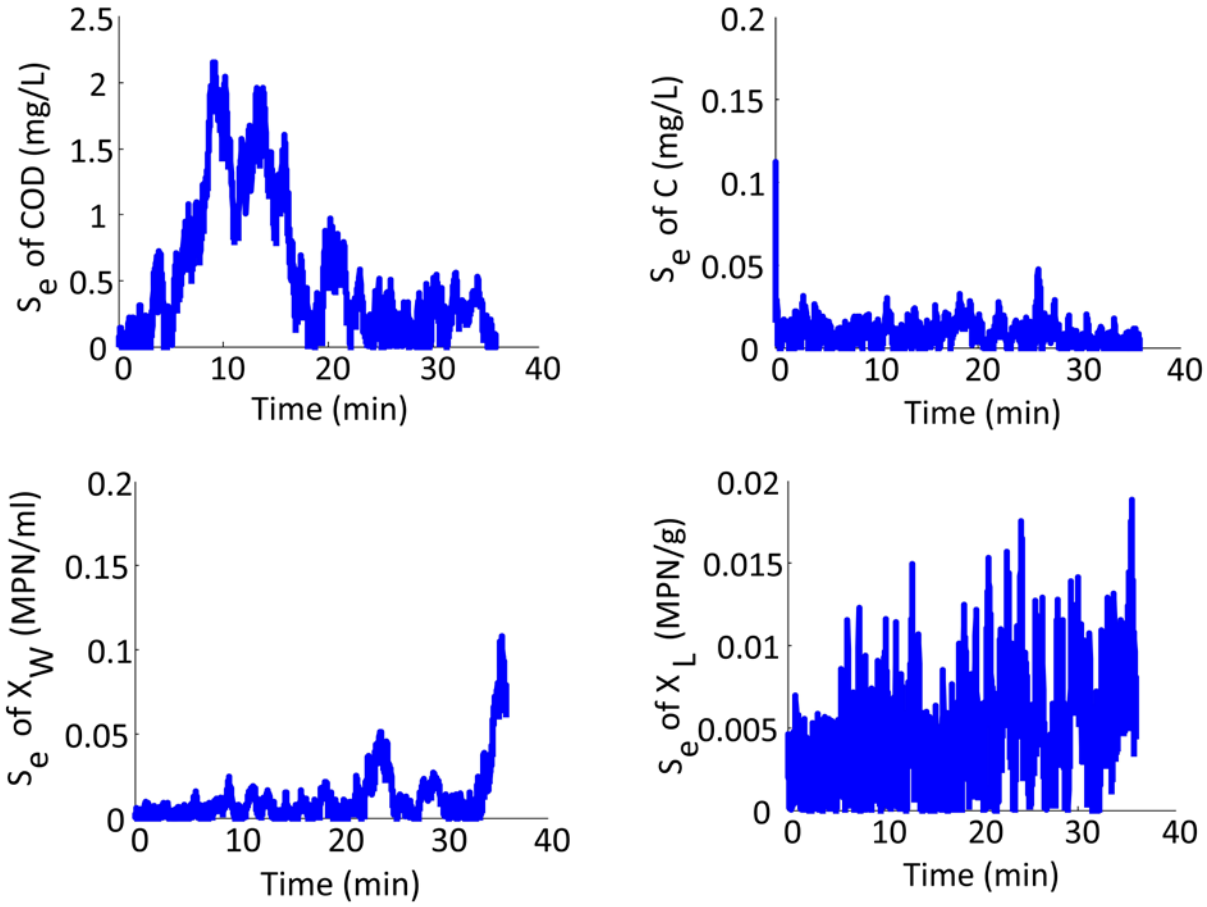

Fig. 10. Standard deviation of the estimation errors for the system states, averaged over 100 Monte Carlo simulations, each with a 36 min simulation time

$S_{e_j}^{P_l\delta}(i)$ represents the standard deviation of the estimation error of the j^{th} state at the i^{th} time step, averaged over N_{mc} Monte Carlo simulations, when the estimator's model of the l^{th} system parameter has an error of δ% relative to its true value while the other system parameters are modeled correctly as shown in Table 1. $\bar{S}_j^{P_l\delta}$ denotes the mean of $S_{e_j}^{P_l}$ over N time steps. $D_j^{P_l}$ indicates the RMSE percentage change of the estimation error of the j^{th} state due to a modeling error in the l^{th} system parameter. $SE_j^{P_l}$ represents the sensitivity of the estimation error of the j^{th} state relative to an error in the l^{th} system model parameter, where a positive sensitivity means that the estimation error of the j^{th} state increases, and a negative sensitivity means that the estimation error of the j^{th} state decreases.

Fig. 11 shows the sensitivity of the error of each state estimate to the error of each system model parameter, which were calculated with a value of $\delta = 10$ in Eqs. (19)-(21). The largest sensitivity is that of the COD (O) estimation error relative to the COD increase rate (K_0); the reason why the estimate of COD has such a large sensitivity to K_0 is that derivative of the COD state is exactly equal to K_0, as seen in Eq. (1). The smallest sensitivity is that of the P on the lettuce (X_L) estimation error relative to the decay rate of FC (γ_c), which is about 0; although γ_c appears in the derivative of the FC concentration in Eq. (1) and chlorine measurements are used to update the state estimates, the cross-contamination dynamics, and in turn X_L, are not directly dependent on γ_c.

As seen from Fig. 11, the estimate of COD is also highly sensitive to the FC consumption rate β_c. The reason for this sensitivity is as follows. Note that the first element of the covariance matrix Q_s (shown at the end of Section II), which corresponds to COD process noise, is very large relative to the other elements of Q_s; this indicates that there is a significant difference between the experimental data and the COD model dynamics of Eq. (1).

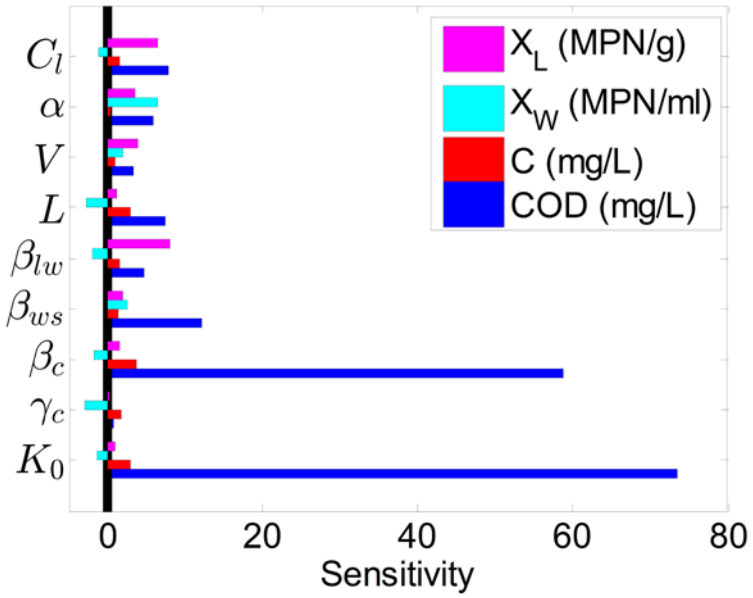

Fig. 11. Sensitivities of the four state estimation errors relative to modeling errors in the nine system parameters

Conversely, the second element of the covariance matrix Q_s is a very small number, indicating a close match between the experimental data and the FC model dynamics of Eq. (1). Combining these two observations, we see that when β_c is modeled incorrectly the Kalman filter adjusts its estimate of COD (x_1), but not FC (x_2), in order to maintain a close match between the experimental data and the $\beta_c x_1 x_2$ term in the second dynamic equation of Eq. (1).

Similarly, we see from Fig. 11 that COD estimation errors have a relatively high sensitivity to almost all of the model parameters. Again, this is because the first element of the covariance matrix Q_s, which corresponds to COD process noise, is very large relative to the other elements of Q_s. This, in turn, allows the Kalman filter to adjust its estimate of COD by a large amount in order to compensate for modeling errors.

Fig. 11 shows that the estimated FC concentration (C) is not highly sensitive to model errors, but its greatest sensitivity is to the modeled value of FC consumption rate (β_c). This is because β_c appears directly in the derivative of the FC state in Eq. (1).

The estimated PC in the water (X_W) is not highly sensitive to model errors, but its greatest sensitivity is to the interaction rate of the pathogen via FC (α). Again, this is because of their direct relationship in the system dynamics of Eq. (1). We also note that the estimate of X_W has a negative sensitivity to several modeling parameters. Negative sensitivity means that the estimate actually improves in the presence of modeling errors. This phenomenon is difficult to explain intuitively, but is due to the highly nonlinear nature of the system dynamics, a characteristic which often precludes the possibility of analytical or intuitive explanations of sensitivity relationships. The negative sensitivities indicate that we can estimate X_W accurately even without a good knowledge of system parameter models.

Finally, the estimated P on the lettuce (X_L) is not highly sensitive to model errors, but its greatest sensitivity is to the pathogen binding rate (β_{lw}). Again, this is because of their direct relationship in the system dynamics.

In summary, the sensitivity analysis shows that the proposed estimator has relatively good performance even when the true system parameters deviate from their modeled values.

VI. Conclusion And Future Work

We presented an HEKF and noise statistics optimization algorithm for a wash system. We designed the HEKF to estimate the states of a continuous-time wash system using discrete-time measurements of the FC. We then designed a noise statistics optimization algorithm based on PSO to identify the HEKF process and measurement noise covariances so that the innovations will be approximately zero mean and white with a specified covariance. Simulation results show that PSO identified the noise statistics sufficiently to decrease the cost function by 34%. RMS estimation errors were 8.24 mg/L for the COD, 0.09 mg/L for the FC concentration, 0.19 MPN/ml for PC in the water wash, and 0.04 MPN/g for the P on the lettuce in the tank. Our sensitivity analysis shows that, as expected, the highest sensitivity is that of the estimation error of the COD to modeling errors in the COD increase rate. In general, the estimation errors have a low sensitivity to modeling errors, showing acceptable robustness of the estimator.

For future work, we will perform a sensitivity and robustness analysis to make sure that our HEKF and noise statistics optimization algorithm are robust to system parameter deviations and are thus reliable as a practical methodology for real-world plant wash systems. In addition, we will design a state-based controller for the produce wash system to control FC concentration, the

pathogen concentration in the water wash, and the amount of pathogen on the lettuce during the washing process, while minimizing the FC input.

The simulation results in this paper can be reproduced with the Matlab source code that is available at http://embeddedlab.csuohio.edu/plant-wash-estimation.

References

[1] Y. Luo et al., "Determination of free chlorine concentrations needed to prevent Escherichia coli O157:H7 cross-contamination during fresh-cut produce wash," *Journal of Food Protection*, vol. 74, no. 3, pp. 352–358, 2011.

[2] Y. Luo et al., "A pilot plant scale evaluation of a new process aid for enhancing chlorine efficacy against pathogen survival and cross-contamination during produce wash," *International Journal of Food Microbiology*, vol. 158, no. 2, pp. 133–139, 2012.

[3] M. I. Gil, M. V. Selma, F. López-Gálvez, A. Allende, "Fresh-cut product sanitation and wash water disinfection: Problems and solutions," *International Journal of Food Microbiology*, vol. 134, no. 2, pp. 37–45, 2009.

[4] G. R. Davidson, A. L. Buchholz, E. R. Ryser, "Efficacy of commercial produce sanitizers against nontoxigenic Escherichia coli 0157:H7 during processing of iceberg lettuce in a pilot-scale leafy green processing line," *Journal of Food Protection*, vol. 76, no. 11, pp. 1838–1845, 2013.

[5] D. Munther et al., "A mathematical model for pathogen cross-contamination dynamics during produce wash," *Food Microbiology*, vol. 51, pp. 101–107, 2015.

[6] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*, John Wiley & Sons, 2006.

[7] S. A. Fakoorian, D. Simon, H. Richter and V. Azimi, "Ground reaction force estimation in prosthetic legs with an extended Kalman filter," *IEEE International Systems Conference*, Orlando, Florida, 2016.

[8] N. Hashemian, A. Armaou, "Simulation, model-reduction, and state estimation of a two-component coagulation process," AIChE Journal, DOI: 10.1002/aic.15146, 2016.

[9] N. Hashemian, A. Armaou, "Fast Moving Horizon Estimation of nonlinear processes via Carleman linearization," *American Control Conference*, Chicago, IL, 2015.

[10] J. Bellantoni, K. Dodge, "A square root formulation of the Kalman-Schmidt filter," *American Institute of Aeronautics and Astronautics Journal*, vol. 5, no. 7, pp. 1309–1314, 1967.

[11] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[12] T. Kontoroupi, A. W. Smyth, "Online noise identification for joint state and parameter estimation of nonlinear systems," *SCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, pp. B4015006-1/12, 2015.

[13] O. Kost, O. Straka, J. Duník, "Identification of state and measurement noise covariance matrices using nonlinear estimation framework," *Journal of Physics: Conference Series*, vol. 659, no. 1–12, pp. 1, 2015.

[14] J. Kennedy, R. C. Eberhart, "Particle swarm optimization," *IEEE International Conf. on Neural Networks*, pp. 1942–1948, 1995.

[15] R. C. Eberhart, J. Kennedy, "A new optimizer using particle swarm theory," *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43, 1995.

[16] D. Simon, *Evolutionary Optimization Algorithms*, John Wiley & Sons, 2013.