Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 55-20 (2022) 487–492

Advanced monitoring of viral amplification process by soft sensing

Laurent Dewasme * **Guillaume Jeanne** ** **Lydia Saint Cristau** **
Céline Barraud ** **Alain Vande Wouwer** *

* *Systems, Estimation, Control and Optimization (SECO), University of Mons (UMONS), 7000 Mons, Belgium.*
** *Sanofi Pasteur, 69007 Lyon, France.*

Abstract: In this study, a software sensor monitoring a viral amplification process is developed and validated. First, a dynamic model structure is proposed, describing Vero cell growth as well as the impact of viral infection, in accordance with the considered industrial application. A parameter identification procedure is set up based on a nonlinear least-square optimization criterion using several data sets provided by Sanofi Pasteur (Lyon, France). Second, an extended Kalman filter is designed considering a specific measurement configuration including a Raman probe sensing biomass, glucose, lactate and glutamine concentrations, and the estimation of exogenous variables such as the cell growth rate and viral amplification parameters. The obtained results validate the possibility to consider the EKF software sensor as a useful tool to monitor and report on viral amplification dynamics.

Keywords: state estimation, parameter estimation, software sensor, viral amplification, Vero cells

1. INTRODUCTION

Describing and predicting virus amplification dynamics on host cells is critical to understand and optimize culture-based vaccine production processes, which are key economic challenges for pharmaceutical industries. In the framework of the 4.0 industry development and the increasing success of artificial intelligence, digital twins based on dynamic models constitute a promising tool to meet these expectations.

Several dynamic models of viral amplification are proposed in the literature for different animal cell strains such as MDCK, Vero or hybridoma cells (Möhler et al., 2005; Schulze-Horsel et al., 2009; Müller et al., 2013; Ursache et al., 2015), namely focusing on Influenza or poliovirus. In (Abbate et al., 2016), the viral amplification process dynamics is described as a function of the whole living biomass, either uninfected or infected, instead of segregating the different populations as it is common in the previous cited works. Considering that infected biomass evolves faster than other variables, the model is indeed reduced through a slow-fast assumption. This work is extended and elaborated in Abbate et al. (2019), where the global identifiability analysis of the proposed model is achieved using the DAISY Toolbox (Bellu et al., 2007) as well as a local identifiability analysis based on the Fisher Information Matrix (FIM). The resulting mechanistic models have good predictive capability and offer the possibility to develop advanced on-line monitoring tools, provided a system property called observability, which states that unmeasured variables can be estimated by an observer or software sensor, in finite time, if the model meets some structural conditions involving the available probe configuration. Application of software sensors to bioprocesses is a widely studied topic (Dochain, 2003; Bogaerts and Vande Wouwer, 2003; Goffaux and Vande Wouwer, 2005; Dewasme et al., 2009; Ali et al., 2015) and recent practical applications have assessed the potential monitoring improvements in the context of animal cell cultures (Amribt et al., 2014b,a; Dewasme et al., 2015; Dewasme and Vande Wouwer, 2020). To the best of the authors knowledge, software sensor design for monitoring viral amplification has not been reported yet. The main motivation of this work is therefore to assess the potential of software sensors in this specific context and to provide insight into the monitoring improvements.

This work, included in a vaccine development project framework, therefore aims at optimizing budding viral amplification cultures in several steps going from process modeling to advanced monitoring and control. This paper reports on the first phases of the project which are dedicated to model parameter identification and observer design, achieved with the support of an available, yet confidential, chemometric model delivering on-line measurements of a state variable subset.

This work is organized as follows. Section 2 presents the process of interest and the operating mode while a candidate dynamic model is proposed in section 3. An original parameter identification procedure is described in section 4 as well as model validation. The model is used to design a software sensor under the form of an extended Kalman filter in section 5 to estimate state variables and critical kinetic parameters. This software sensor is validated in section 6 under the assumption of the presence of a Raman probe and conclusions are drawn in section 7.

2. VIRAL AMPLIFICATION PROCESS

The current study considers a budding viral amplification process achieved in 15 mL bioreactors, starting with cell seeding at day 0 followed by a $2-$day batch culture. At day 3, a medium change is operated as well as cell infection with a specific multiplicity of infection representing the ratio $MOI = \frac{Vir}{X}$ (which is not divulged for the sake of confidentiality) where Vir and

⋆ e-mail: laurent.dewasme@umons.ac.be,celine.barraud@sanofi.com

10.1016/j.ifacol.2022.09.142

X respectively stand for the viral load and cell concentration. At day 5 (day 2 post-infection), one last medium change is operated. Off-line measurement samples are taken at days 0 and 3 during the first batch phase and every day following infection. A summary of these operating conditions is shown in Table 1.

Table 1. Viral amplification procedure - the index *pi* stands for post-infection.

Day	Operation	Sampling
d_0	Biomass inoculation	Yes
d_1	No operation	No
d_2	No operation	No
$d_3 = d_{0,pi}$	Medium change and infection	Yes (Pre and post)
$d_{1,pi}$	No operation	Yes
$d_{2,pi}$	Medium change	Yes (Pre and post)
$d_{3,pi}$	No operation	Yes

The provided database concerns the follow-up of 4 experiments where the following metabolite concentrations are measured, e.g., the viable biomass (10^6 $cells/mL$), glucose (g/L), glutamine (g/L), lactate (g/L), ammonium (mM), glutamate (g/L) and infection titer concentrations ($log(10^6\ Vir/mL)$).

3. DYNAMIC MODELING

The starting point of the current study is the modeling procedure reported in (Abbate et al., 2019), which is based on the following assumptions:

- The total biomass rules the substrate consumption dynamics as well as by-product production rates and biomass decay;
- The growth is dynamically driven only by the uninfected biomass;
- The biomass decay rate is constant all along the culture;
- Viral amplification is dynamically driven by the infected biomass;
- The infected biomass is assumed to be in quasi steady-state with respect to other states by time-scale separation (the conversion of uninfected to infected biomass is almost instantaneous from the slowly growing variable reference).

These assumptions lead to an ordinary differential equation model describing a batch process (Abbate et al., 2019):

$$\frac{d\xi}{dt} = K\,\varphi \tag{1}$$

where ξ is the state vector, containing all metabolite concentrations, i.e., biomass X, glucose G, glutamine Gn, lactate L, ammonium N, glutamate Glu and infection titer IT. K is the pseudo-stoichiometric matrix which reads:

$$K = \begin{pmatrix} \delta(t) & 0 & 0 \\ 0 & -Y_{glc} & 0 \\ 0 & -Y_{gln} & 0 \\ 0 & Y_{lac} & 0 \\ 0 & Y_{NH_3} & 0 \\ 0 & Y_{glu} & 0 \\ 0 & 0 & Y_{vir} \end{pmatrix} \tag{2}$$

where $\delta = \frac{t}{t+t_L}$ and $t_L = 1$ day represents the growth latency observed each time a medium renewal is achieved (i.e., medium renewal can be considered, regarding the cell acclimation, as the start of a new culture), $Y(.)$ are the stoichiometric coefficients and the reaction rate vector φ reads:

$$\varphi = \begin{pmatrix} \mu_{growth}(X - X_i) \\ \mu_{growth} X \\ X_i \end{pmatrix} \tag{3}$$

where:

$$\mu_{growth} = \frac{-1}{1 + exp\,(-S\,(X - X_{max}))} + 1 \tag{4}$$

S is a parameter related to the surface of the carrier beads, assumed to be known and set to 4 in the current study, and X_{max} is the maximum capacity concentration of biomass clustered on the beads. The infected biomass X_i is calculated as:

$$X_i = \frac{IT}{IT + K_{Vir}} X \tag{5}$$

where K_{Vir} stands as the half-saturation constant of the infection rate. Equation (5) results from a slow-fast dynamic assumption of the infection which is assumed to be much faster than the biomass growth (see (Abbate et al., 2019)).

4. PARAMETER IDENTIFICATION

4.1 Identification procedure

A weighted nonlinear least-squares criterion minimizing the distances between model predictions and experimental data is chosen:

$$J = \sum_{s=1}^{n_{exp}} \sum_{i=1}^{n_{samp}} \left(\xi(i,s) - \xi_{data}(i,s)\right) Q^{-1} \left(\xi(i,s) - \xi_{data}(i,s)\right)^T \tag{6}$$

where $\xi_{data}(i,s)$ is the state measurement at time i (going from 1 to n_{samp}) in experiment s (going from 1 to n_{exp}) and Q is the covariance matrix weighting the cost function J.

The optimization problem can therefore be enunciated as $min_\theta J$ such that (1) holds, where θ represents the parameter vector containing the stoichiometric and kinetic parameters but also the ordinary differential equation initial conditions $\xi(1,s)$.

4.2 Model validation

The validation process consists in dividing the data sets in two sets, selecting 3 experiments described in section 2 to identify the model parameters and achieve a direct validation (assessing the model fitting). The remaining fourth data set is used to cross-validate the model (assessing its predictive capacity in new situations).

The chosen optimization procedure is implemented on the MATLAB platform and proceeds as follows:

- The *fmincon* solver is called, which applies an interior-point algorithm with constraints on the parameter values (allowing to reduce the search space). These constraints are however chosen quite large (10^{-6} to 10^6) but could be tightened if required (in case of multiple local minima). The *fmincon* solver is called in a loop using each new optimization result as initial guess of the next solver call until a specific threshold is met, stopping the loop when no residual variation of more than 0.1% with respect to the previous optimization result is detected;

- The latter results are considered as initial guess of a new optimization calling the *lsqnonlin* solver, using a Levenberg-Marquardt algorithm and providing the Jacobian matrix of parameter sensitivities that can be used to build the lower bound of the parameter estimation error covariance from the inverse of the Fisher Information Matrix (FIM).

A multi-start strategy is also applied, generating multiple random initial guesses of the parameters and covering the search space to avoid falling into local minima. Model performances are evaluated taking into account the cost function J residual (fitting quality) and practical identifiability (parameter estimate confidence intervals at 95% inferred from the diagonal of the inverted FIM).

The identification results are shown in Table 2 and the direct and cross-validations respectively appear in Figures 1 and 2. Obviously, the fitting is satisfactory, with a mean residual of 0.055 in direct validation and 0.076 in cross-validation. The parameter values also seem to be accurately identified since the relative confidence intervals at 95 % are, for the majority, below 15 % excepted the yield coefficient of virus production which is more uncertain, mainly due to the rather small number of infectious titer data.

Table 2. Parameter identification results - Values and relative confidence intervals (C.I.)

Parameter	Value	C.I. (%)
X_{max}	0.843	7.988
Y_{glc}	3.564	13.429
Y_{gln}	0.439	12.904
Y_{lac}	3.049	11.808
Y_{glu}	16.290	12.359
Y_{NH_3}	2.706	11.687
Y_{Vir}	305.359	51.021
K_{Vir}	125.067	4.796

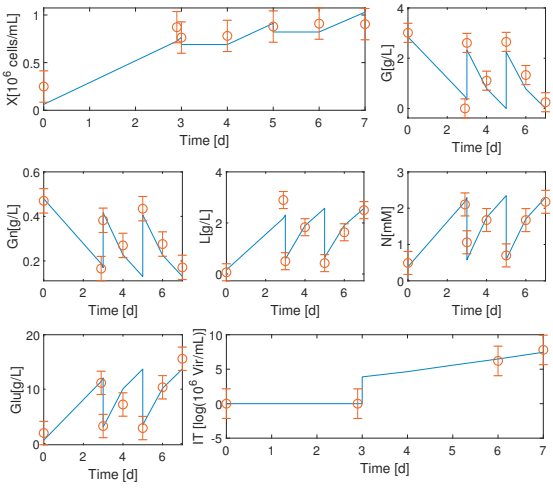

Fig. 1. Direct validation of the identified model. Continuous line: model. Bubbles: experimental data with 95 % confidence intervals.

The results of this modeling and identification procedure of the viral amplification process are therefore comparable to the work of Abbate et al. (2019), which is comforting with regard

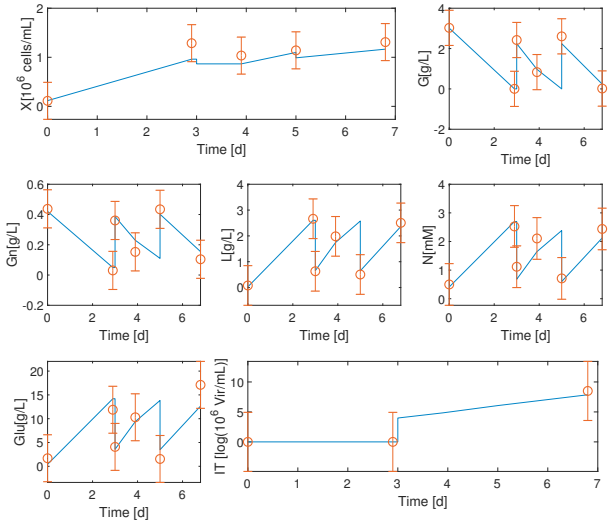

Fig. 2. Cross-validation of the identified model. Continuous line: model. Bubbles: experimental data with 95 % confidence intervals.

to the development of a software sensor, as proposed in the next section.

5. SOFTWARE SENSOR DESIGN

The definition of observability is first recalled:

Definition System (1) is observable if

$$\forall t_0, \exists t_1 < \infty | y(t;t_0,\xi(0),u(t)) = y(t;t_0,\xi'(0),u(t)), \\ \forall u(t), t_0 < t < t_1 \Rightarrow \xi(0) = \xi'(0) \tag{7}$$

where $y = h(\xi)$ is the system measurable output assumed as a nonlinear function of the states ξ. Observability of nonlinear systems is however very difficult to analyze. A weaker but very interesting concept is detectability.

Definition System (1) is detectable if considering a copy of this system inducing new states ξ', for any known output $y = h(\xi)$, the error $\varepsilon = \xi - \xi'$ tends asymptotically to zero, therefore inducing the distinguishability of the state trajectories (the interested reader can refer to (Moreno et al., 2014) for a methodology based on this concept).

5.1 Measurement mapping

Observability and detectability therefore depend on the on-line measurement mapping. Calibrated turbidimetric or capacitance probes allow monitoring on-line the biomass concentration, making the latter the most likely measurable candidate at the start of the following analysis.

Considering system (1), an error system can be proposed:

$$\frac{d\varepsilon}{dt} = K\left(\varphi(\xi) - \varphi(\xi')\right) \tag{8}$$

where ξ' is a copy of the state vector ξ.

Considering viable biomass as measurable, μ_{growth} is only a function of $X = X'$, $\varepsilon_X = 0$, inducing:

$$0 = \frac{d\varepsilon_X}{dt} = \delta X \mu_{growth} K_{Vir} \frac{IT' - IT}{(IT + K_{Vir})(IT' + K_{Vir})} \quad (9)$$

and $IT = IT'$, making system (1) partially detectable. Moreover, the vector $\frac{d\varepsilon_\xi}{dt}$ is zero for all ξ species which means that any initial error on the remaining state variables is never increasing and system (1) presents stable error dynamics.

In this study, not only biomass but also glucose, glutamine and lactate are measured on-line. This configuration is made possible by chemometric modeling which allows calibrating a Raman probe and correlating with machine learning techniques the spectra with off-line measurements of specific state variables. In turn, this allows reconstructing their on-line evolutions at a pace set by the Raman probe sampling.

Only ammonium and glutamate could possibly not converge to their true trajectory. If these components do not accumulate excessively, thus avoiding toxic levels for the cells, their monitoring is not essential.

It must also be noticed that detectability is a structural property which depends on the choice of particular kinetic structures.

5.2 Extended Kalman Filter with exogenous variables

The Kalman filter (KF) is an optimal estimator minimizing the estimation error variance. It is a popular approach for process monitoring due to its capacity to take measurement noise into account (under the assumption of Gaussian white noises) and its successful application in many reported studies (Wilson et al., 1998; Hitzmann et al., 2000; Arndt and Hitzmann, 2004; Arndt et al., 2005; Dewasme et al., 2013; Amribt et al., 2014b). However, since the KF is defined in a linear framework, its application to models such as (1) requires a first-order linearization around the estimated trajectory and the resulting estimator is called extended Kalman filter (EKF). Moreover, since most of the bioprocesses are usually modeled by continuous-time ordinary differential equation systems while on-line probes deliver data at discrete times, continuous-discrete EKF is the most appropriate version. Estimations are achieved in two steps: a prediction using the available model and a correction when the measurements are available:

Prediction between t_k and t_{k+1}:

$$\frac{d\xi(t)}{dt} = K\varphi(\xi(t),t) - D\xi(t) + D\xi^{in}; \quad \xi(t_k) = \xi(t_k^+), \quad t_k^+ \leq t < t_{k+1}^- \quad (10a)$$

$$\frac{dC(t)}{dt} = A(\xi(t))C(t) + C(t)A(\xi(t))^T + R_\eta; \quad C(t_k) = C(t_k^+), \quad (10b)$$

Correction at time t_{k+1}:

$$\Omega(\xi(t_{k+1})) = C(t_{k+1}^-)L^T \left[LC(t_{k+1}^-)L^T + R_\varepsilon(t_{k+1})\right]^{-1} \quad (11a)$$

$$\xi(t_{k+1}^+) = \xi(t_{k+1}^-) + \Omega(\xi(t_{k+1}))\left(y(t_{k+1}) - L\xi(t_{k+1}^-)\right) \quad (11b)$$

$$C(t_{k+1}^+) = C(t_{k+1}^-) - \Omega(\xi(t_{k+1}))LC(t_{k+1}^-) \quad (11c)$$

In these expressions, L is the measurement matrix, Ω the correction gain, C the covariance matrix of the state estimation errors, R_ε and R_η the covariance matrices of respectively the measurement and model errors, and t_{k+1}^- and t_{k+1}^+ the time instants relative to the a priori and a posteriori estimations between which the measurement is assumed to be delivered.

When some model parameters are uncertain, the EKF also offers the possibility to estimate these parameters under the assumption of a slow variation which is described in an exogenous model. Since the purpose of this work is to monitor viral amplification, it may be of interest to estimate the cell growth rate μ_{growth}, and uncertain parameters such as the virus production yield coefficient Y_{Vir} as well as the kinetic constant K_{Vir} which characterizes the quantity of infected biomass.

6. MONITORING OF VIRAL AMPLIFICATION: RESULTS

This section makes use of experimental data collected in several experiments where the Raman probe was not in operation. In order to test the performance of the EKF using a Raman probe, it is therefore necessary to emulate the signal information that will be provided by such a probe. This is achieved by generating synthetic data between actual off-line data points using the model identified with the same data. A fast sampling is assumed, e.g., $T_s = 0.1$ day, which is even slower than expected in real conditions (about half an hour). The emulated Raman probe reproduces the evolutions of the biomass, glucose, glutamine and lactate concentration measurements. The validation of the EKF can therefore be considered as a worse case study, in the sense that faster sampling will improve on the current results. In order to reproduce realistic conditions, white Gaussian noise is added to the model outputs, with standard deviations given by the chemometric model (which, for the sake of confidentiality, is not divulged), which are $\sigma_X = 0.113\ g/L$, $\sigma_G = 0.130\ g/L$, $\sigma_{Gn} = 0.026\ g/L$ and $\sigma_L = 0.230\ g/L$ respectively for biomass, glucose, glutamine and lactate. The EKF covariance matrices are therefore designed as follows:

$$R_\varepsilon = \begin{pmatrix} 0.113^2 & 0 & 0 & 0 \\ 0 & 0.130^2 & 0 & 0 \\ 0 & 0 & 0.026^2 & 0 \\ 0 & 0 & 0 & 0.230^2 \end{pmatrix} \quad (12)$$

$$R_\eta = I_{n_a\ x\ n_a} \quad (13)$$

where I stands for the identity matrix and $n_a = 10$ is the number of state variables of the augmented model considering the exogenous model for the parameter evolution:

$$\frac{d\mu_{growth}}{dt} = 0 \quad (14a)$$

$$\frac{dY_{Vir}}{dt} = 0 \quad (14b)$$

$$\frac{dK_{Vir}}{dt} = 0 \quad (14c)$$

R_η reflects the confidence in the model accuracy as compared to the measurement noise. In the present study, more confidence is given to the measurements as the diagonal of R_ε is at least 100 times smaller than the corresponding terms in R_η. C is initialized to consider initial estimates normally distributed around the first data sample with a relative standard deviation of 10 %.

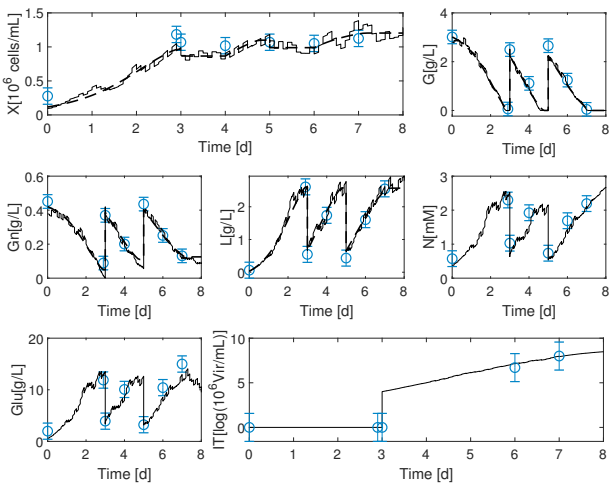

Fig. 3. Validation of the augmented EKF for the estimation of viral amplification state variables - Bubbles: experimental data with 95 % confidence intervals - Dashed line: model output - Continuous line: EKF output.

Figure 3 shows the EKF application in the context of the first experiment that was used to identify the model parameters, run over 8 days. The estimates are qualitatively satisfactory since there is no important deviation from the off-line data. Regarding the medium renewal at days 3 and 5, a 10 % relative error is applied to the expected dilution factor, which is likely to vary in true experimental conditions. This also allows challenging the EKF which has to restart converging to the true state trajectories between days 3 and 5 as well as days 5 and 8.

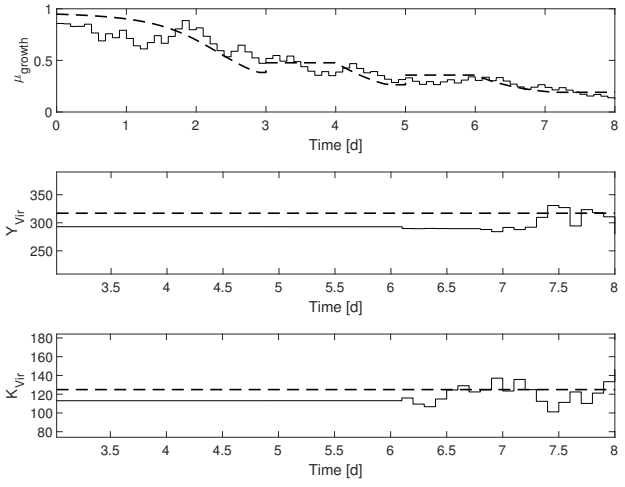

Fig. 4. Validation of the augmented EKF for the estimation of viral amplification key rate/parameters - Dashed line: model output/parameter value - Continuous line: EKF output.

More interestingly, Figure 4 shows that the selected rate and parameters are well estimated by the augmented EKF since the latter keeps track of the cell growth rate μ_{growth} before and after infection despite the discontinuous variations of the model prediction, due to the medium renewals, and provides, even if noisy, good estimates of the viral amplification parameters. The convergence to the values of Y_{Vir} and K_{Vir} starts at day 4 (post infection) and requires several days to approach the parameter nominal values. The application of the EKF to a second experimental dataset confirms these observations as shown in figures 5 and 6.

These different dynamic behaviors either related to the initial state variables or to the augmented ones should be taken into account in future experimental investigations which could consider the augmented EKF as a key software tool to trigger viral amplification.

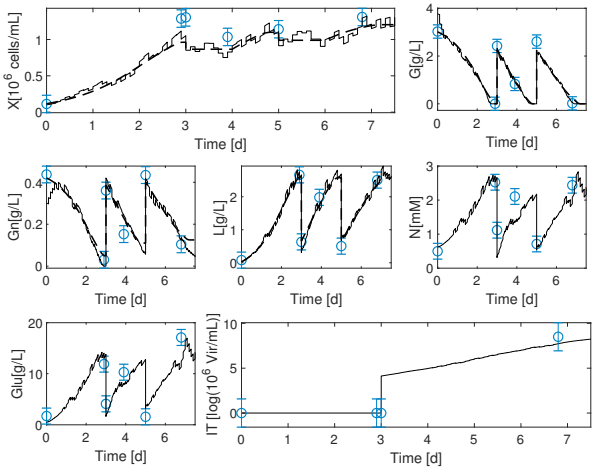

Fig. 5. Validation of the augmented EKF for the estimation of viral amplification state variables - Bubbles: experimental data with 95 % confidence intervals - Dashed line: model output - Continuous line: EKF output.

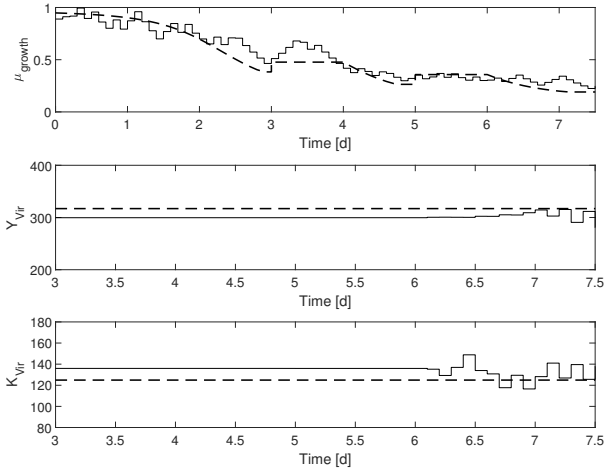

Fig. 6. Validation of the augmented EKF for the estimation of viral amplification key rate/parameters - Dashed line: model output/parameter value - Continuous line: EKF output.

7. CONCLUSION

In this paper, monitoring of a viral amplification process is developed based on a dynamic model whose parameters are

identified using a multi-step procedure and industrial data sets. The software sensor is an extended Kalman filter (EKF) combined to a chemometric model of a Raman probe delivering online biomass, glucose, glutamine and lactate concentration measurements. The EKF blends the information from the dynamic model and the chemometric model in order to reconstruct the trajectories of the unmeasured state variables, i.e., the infection titer (viral load) as well as a few key kinetic parameters. The results offer promising prospects, despite partial detectability of the system (ammonium and glutamate are not guaranteed to converge to the actual values), and future research entails the implementation of the EKF with the actual Raman probe to monitor the cell growth rate and the parameters related to the infected biomass and the infection yield.

ACKNOWLEDGEMENTS

This work was funded by Sanofi Pasteur. Conceptualization, investigation, software and original draft preparation: Laurent Dewasme. Resources and data curation: Lydia Saint Cristau, Guillaume Jeanne and Céline Barraud. Methodology and writing review-editing: Alain Vande Wouwer. Lydia Saint Cristau, Guillaume Jeanne and Céline Barraud are Sanofi employees and may hold shares in the company. All other authors declare no competing interests and have read and agreed to the published version of the manuscript.

REFERENCES

Abbate, T., Dewasme, L., and Vande Wouwer, A. (2016). Dynamic macroscopic model of dengue viral amplification in vero cell cultures. *IFAC-PapersOnLine*, 49(26), 159–164.

Abbate, T., Dewasme, L., and Vande Wouwer, A. (2019). Variable selection and parameter estimation of viral amplification in vero cell cultures dedicated to the production of a dengue vaccine. *Biotechnology Progress*, 35(1), e2687.

Ali, J., Hoang, N., Hussain, M., and Dochain, D. (2015). Review and classification of recent observers applied in chemical process systems. *Computers and Chemical Engineering*, 76, 27–41.

Amribt, Z., Dewasme, L., Vande Wouwer, A., and Bogaerts, P. (2014a). Optimization and robustness analysis of hybridoma cell fed-batch cultures using the overflow metabolism model. *Bioprocess Biosyst Eng*, 37, 1637–1652.

Amribt, Z., Dewasme, L., Vande Wouwer, A., and Bogaerts, P. (2014b). Parameter identification for state estimation: Design of an extended kalman filter for hybridoma cell fed-batch cultures. In *19th IFAC World Congress (IFAC'14)*.

Arndt, M. and Hitzmann, B. (2004). Kalman filter based glucose control at small set points during fed-batch cultivation of saccharomyces cerevisiae. *Biotechnology Progress*, 20, 377–383.

Arndt, M., Kleist, S., Miksch, G., Friehs, K., Flaschel, E., Trierweiler, J., and Hitzmann, B. (2005). A feedforward-feedback substrate controller based on a kalman filter for a fed-batch cultivation of escherichia coli producing phytase. *Computers and Chemical Engineering*, 29, 1113–1120.

Bellu, G., Saccomani, M., Audoly, S., and DAngiò, L. (2007). Daisy: a new software tool to test global identifiability of biological and physiological systems. *Comput Methods Prog Biomed.*, 88(1), 52–61.

Bogaerts, P. and Vande Wouwer, A. (2003). Software sensors for bioprocesses. *ISA Transactions*, 42, 547–558.

Dewasme, L., Bogaerts, P., and Vande Wouwer, A. (2009). *Monitoring of bioprocesses: mechanistic and data-driven approaches*, 57–97. Studies in Computational Intelligence, (Computational Intelligent Techniques for Bioprocess Modelling, Supervision and Control, Maria do Carmo Nicoletti, Lakhmi C. Jain, eds.). Springer Verlag.

Dewasme, L., Fernandes, S., Amribt, Z., Santos, L., Bogaerts, P., and Vande Wouwer, A. (2015). State estimation and predictive control of fed-batch cultures of hybridoma cells. *Journal of Process Control*, 30, 50–57.

Dewasme, L., Goffaux, G., Hantson, A.L., and Vande Wouwer, A. (2013). Experimental validation of an extended kalman filter estimating acetate concentration in *E. coli* cultures. *Journal of Process Control*, 23, 148–157.

Dewasme, L. and Vande Wouwer, A. (2020). Experimental validation of a full-horizon interval observer applied to hybridoma cell cultures. *International Journal of Control*, 93(11), 2719–2728.

Dochain, D. (2003). State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of Process Control*, 13(8), 801–818.

Goffaux, G. and Vande Wouwer, A. (2005). Bioprocess state estimation: some classical and less classical approaches. *Control and observer design for nonlinear finite and infinite dimensional systems*, 111–128.

Hitzmann, B., Broxtermann, O., Cha, Y.L., Sobieh, O., Stärk, E., and Scheper, T. (2000). The control of glucose concentration during yeast fed-batch cultivation using a fast measurement complemented by an extended kalman filter. *Bioprocess Engineering*, 23, 337–341.

Möhler, L., Flockerzi, D., Sann, H., and Reichl, U. (2005). Mathematical model of influenza a virus production in large-scale microcarrier culture. *Biotechnol Bioeng.*, 90(1), 46–58.

Moreno, J.A., Rocha-Cozatl, E., and Vande Wouwer, A. (2014). A dynamical interpretation of strong observability and detectability concepts for nonlinear systems with unknown inputs: application to biochemical processes. *Bioprocess Biosyst. Eng.*, 37, 37–49.

Müller, T., Dürr, R., Isken, B., Schulze-Horsel, J., Reichl, U., and Kienle, A. (2013). Distributed modeling of human influenza a virus-host cell interactions during vaccine production. *Biotechnol Bioeng.*, 110(8), 2252–2266.

Schulze-Horsel, J., Schulze, M., Agalaridis, G., Genzel, Y., and Reichl, U. (2009). Infection dynamics and virus-induced apoptosis in cell culture-based influenza vaccine production-flow cytometry and mathematical modeling. *Vaccine*, 27(20), 2712–2722.

Ursache, R., Thomassen, Y., van Eikenhorst, G., Verheijen, P., and Bakker, W. (2015). Mathematical model of adherent vero cell growth and poliovirus production in animal component free medium. *Bioprocess Biosyst Eng.*, 38(3), 543–555.

Wilson, D., Agarwal, M., and Rippin, D. (1998). Experiences implementing the extended kalman filter on an industrial batch reactor. *Computers & chemical engineering*, 22(11), 1653–1672.