Received: 20 January 2023 | Revised: 31 March 2023 | Accepted: 1 May 2023

DOI: 10.1002/bit.28428

ARTICLE

BIOTECHNOLOGY AND BIOENGINEERING WILEY

An adaptive soft-sensor for advanced real-time monitoring of an antibody-drug conjugation reaction

Robin Schiemer | Jan Tobias Weggen | Katrin Marianne Schmitt | Jürgen Hubbuch

Institute of Process Engineering in Life Sciences—Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Baden-Württemberg, Germany

Correspondence
Jürgen Hubbuch, Institute of Process Engineering in Life Sciences—Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology (KIT), Fritz-Haber-Weg 2, 76131 Karlsruhe, Baden-Württemberg, Germany.
Email: Juergen.Hubbuch@kit.edu

Funding information
Ministry of Science, Research and the Arts of the State of Baden-Württemberg within the initiative Ideenwettbewerb Biotechnologie

Abstract

In the production of antibody-drug conjugates (ADCs), the conjugation reaction is a central step defining the final product composition and, hence, directly affecting product safety and efficacy. To enable real-time monitoring, spectroscopic sensors in combination with multivariate regression models have gained popularity in recent years. The extended Kalman filter (EKF) can be used as so-called soft-sensor to fuse sensor predictions with long-horizon forecasts by process models. This enables the dynamic update of the current state and provides increased robustness against experimental noise or model errors. Due to the uncertainty associated with sensor and process models in biopharmaceutical applications, the deployment of such soft-sensors is challenging. In this study, we demonstrate the combination of an uncertainty-aware sensor model with a kinetic reaction model using an EKF to monitor a site-directed ADC conjugation reaction. As the sensor model, a Gaussian process regression model is presented to realize a time-variant determination of the sensor uncertainty. The EKF fuses the time-discrete predictions of the amount of conjugated drug from the sensor model with the time-continuous predictions from the kinetic model. While the ADC species are not distinguishable by on-line recorded UV/Vis spectra, the developed soft-sensor is able to dynamically update all relevant reaction species. It could be shown that the use of time-variant process and sensor noise computation approaches improved the performance of the EKF and achieved a reduction of the prediction error of up to 23% compared with the kinetic model. The developed framework proved to enhance robustness against noisy sensor measurements or wrong model initialization and was successfully transferred from batch to fed-batch mode. In future, this framework could be implemented for model-based process control and be adopted for other ADC conjugation reaction types.

KEYWORDS
antibody-drug conjugate, extended Kalman filter, Gaussian process regression, process analytical technology, reaction monitoring, soft-sensor

Robin Schiemer and Jan Tobias Weggen contributed equally to this study.

1 | INTRODUCTION

Antibody-drug conjugates (ADCs) are one of the fastest-growing classes of biopharmaceuticals using the target specificity of a monoclonal antibody (mAb) to exclusively deliver a cytotoxic drug to tumor cells (Thomas et al., 2016). Many different conjugation strategies exist consisting of multiple reaction steps while one of the greatest challenges remains in controlling the drug-to-antibody ratio (Joubert et al., 2020; Panowski et al., 2014). To ensure consistent final product quality, regulatory agencies have been promoting the implementation of advanced monitoring and control strategies according to the concepts of quality by design (QbD) and process analytical technology (PAT) during a decade (FDA, 2004; ICH, 2009). In line with the QbD framework, recent publications showed the successful application of model-based techniques in ADC process development, such as molecular dynamics simulations (Coumans et al., 2020), conjugation kinetic models (Andris et al., 2019), and chromatography models (Andris & Hubbuch, 2020; Keller & Wendeler, 2021). Mechanistic process models, however, often require simplification of complex phenomena which may lead to plant-model mismatches or poor predictability of the transient behavior of the process (Gargalo et al., 2021). With regard to PAT, spectroscopic sensors have shown to enable real-time measurement of certain critical quality attributes (Rathore et al., 2010). Consequently, these sensors have been deployed as PAT for various unit operations in downstream processing of mAbs (Großhans et al., 2018; Rolinger et al., 2020; Wei et al., 2022). In the field of ADCs, monitoring of the conjugation reaction using UV/Vis spectra (Andris et al., 2018) and quantification of aggregation using Raman spectroscopy (Zhang et al., 2019) are recently established PAT tools. Partial least-squares regression (PLSR) models have commonly been applied in chemometrics (Andersson, 2009; Wold et al., 2001) and, more recently, Gaussian process regression (GPR) models have been identified as a competitive alternative to PLSR (Chen et al., 2007; Cui & Fearn, 2017; Tulsyan et al., 2021). Furthermore, GPR models deliver an uncertainty estimate for each evaluated data point and therefore help perform risk-based dynamic adjustments during the process (Tulsyan et al., 2021).

The extracted information from sensors can be used to supply process models with direct feedback from the process in near real-time, which enables timely model updates and, hence, more accurate estimates of process- or product-relevant variables (Sinner et al., 2021). To fuse estimates from two noisy sources, various optimal state estimation algorithms have been developed and recently been applied to chemical (Haseltine & Rawlings, 2005; Schneider & Georgakis, 2013) and pharmaceutical (Elsheikh et al., 2021; Golabgir & Herwig, 2016; Simutis & Lübbert, 2017) processes. The extended Kalman filter (EKF), as one of the most prominently applied algorithms, has been used to combine nonlinear process models and sensor measurements for cell culture processes (Elsheikh et al., 2021; Narayanan et al., 2020) or for capture chromatography (Feidl et al., 2019). In these studies, the EKF computes a weighted average of both source models, but does not affect the long-horizon predictions of the process model. In many cases, however, only a few states are directly available from measurement probes. This problem can be circumvented by utilizing an EKF formulation that combines the measurement with mass or energy balances. Thus, an EKF can be utilized to update all states in the model which are not available for measurement. This approach is considered a model-based soft-sensor according to Luttmann et al. (2012) and was similarly demonstrated in recent studies for advanced monitoring of bioprocesses (Cabaneros Lopez et al., 2020; Ohadi et al., 2015; Spann et al., 2018; Tuveri et al., 2021).

This study aims to develop a robust and adaptive soft-sensor to monitor the site-directed conjugation reaction of a mAb with a surrogate payload. Therefore, an uncertainty-aware chemometric sensor and the system of equations from a kinetic model are combined using an EKF algorithm. This enables to deliver dynamically updated predictions of all reaction states not available for measurement. The experimental data set encompasses nine batch conjugations and two fed-batch conjugations. Both kinetic model and chemometric sensor model are calibrated on the batch data based on the methodology of previously published works (Andris et al., 2018, 2019). For the chemometric sensor, a novel GPR model is evaluated against benchmark alternatives and combined with different preprocessing strategies. Subsequently, an EKF is deployed as a soft-sensor fusing the information of the GPR model with the kinetic model. In detail, this approach recursively updates the unmeasurable ADC species predicted from the kinetic model by using the time-discrete GPR predictions of the conjugated drug concentration. The influence of time-variant uncertainties for the GPR and kinetic model are investigated based on the methods given in literature and the optimal configuration is selected. Finally, the robustness and accuracy of the EKF are assessed by comparison to the two underlying models in different scenarios. The described workflow suggests a unified calibration routine for the underlying models, enables the quantification of model uncertainties, and assesses the performance boost of the developed soft-sensor.

2 | MATERIAL AND METHODS

2.1 | Experiments

2.1.1 | Chemicals and buffers

The chemicals used in this study were obtained from Merck KGaA, unless otherwise stated. Standard working buffer was made with NaH_2PO_4, dihydrate (VWR International GmbH), and ultrapure water (PURELAB Ultra, ELGA LabWater). The buffer was titrated to pH 7.2 with 4 M NaOH and filtered through a 0.2 µm cellulose acetate membrane filter (Sartorius AG). For the mAb reduction and oxidation step tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and L-dehydroascorbic acid were used. To mimic a cytotoxic drug, the nontoxic surrogate payload *N*-(1-pyrenyl)maleimide (NPM) was used. For quenching of the conjugation reaction, unconjugated free drug

was quenched with *N*-acetyl cysteine (NAC). For the reversed-phase ultrahigh performance liquid chromatography (RP-UHPLC), acetonitrile (VWR International GmbH) and trifluoroacetic acid (Thermo Scientific) were used.

2.1.2 | mAb functionalization and conjugation procedure

The reactions were conducted using an engineered immunoglobulin G1 mAb with two inserted cysteines in the hinge region which was generously provided by AstraZeneca. All chemicals were dissolved in 50 mM sodium phosphate buffer, except NPM which was dissolved in dimethyl sulfoxide. In the beginning, the mAb was thawed, diluted, and three reaction steps were consecutively performed: a reduction reaction to uncap the engineered cysteines, a dialysis to remove excess TCEP and a reoxidation reaction of the reduced interchain disulfide bonds. This functionalization procedure was conducted under the same methodology and conditions as previously described in Andris et al. (2019). The subsequent conjugation reaction was performed in a stirred glass beaker (400 rpm) with a reaction volume of 10 mL. Batch conjugations were started by adding a certain volume of 0.68 mM NPM solution with a pipette to reach the desired molar drug excess. In nine runs, the mAb concentration and drug excess were varied, including some runs in duplicates. Two additional fed-batch runs were performed by adding the NPM solution with a syringe pump (Nemesys S, Cetoni GmbH) constantly over 5 and 10 min for run 10 and run 11, respectively. All conducted runs are summarized in Table 1.

2.1.3 | UV/Vis on-line monitoring and off-line RP-UHPLC

To monitor the conjugation reaction, the same setup as in Andris et al. (2018) was used. In detail, an on-line loop was installed where a peristaltic pump (Minipuls 3, Gilson) pumped the reaction mixture continuously through the loop tubing. The flow rate was adjusted to 1 mL min^{-1} by measuring the flow rate with a liquid flow meter SLS-1500 (Sensirion AG). To collect UV/Vis spectra every 0.2 s, a DAD-3000 RS diode array detector with a 0.4-mm path length flow cell (both Thermo Fisher Scientific) was installed in the loop. UV/Vis spectra were recorded between 250 and 390 nm at a resolution of 1 nm. Before starting the conjugation reaction, the entire loop was equilibrated with the oxidized mAb solution for up to 15 min and an autozero of the DAD was performed to collect the pure NPM spectra during the reaction.

Off-line samples were taken at certain time points and immediately quenched with an NAC solution. To determine the concentrations of the single ADC species for the time points, all samples were analyzed with RP-UHPLC. The same device, column, and analysis method as described in Andris et al. (2018) was applied. The RP-UHPLC data will further be treated as reference data for all models.

TABLE 1 Overview of conducted conjugation runs using different mAb concentrations and NPM excesses.

Experiment	mAb concentration (mg mL^{-1})	Molar drug excess (–)	Number of samples	Subset
Run 1	1	2×	20	Training
Run 2	1	4×	17	Test
Run 3	1.25	3×	21	Training
Run 4	1.5	2×	21	Test
Run 5	1.5	2×	20	Training
Run 6	1.5	3×	18	Test
Run 7	1.5	4×	21	Training
Run 8	1.75	2×	20	Training
Run 9	1.75	3×	20	Training
Run 10 (fed-batch)	1.5	3×	22	External
Run 11 (fed-batch)	1.5	3×	20	External

Note: Batch runs (1–9) were split into training and test sets for model calibration, EKF tuning, and performance evaluation. Fed-batch runs (10 and 11) were used later in the study to test the transferability of the designed framework.

Abbreviations: EKF, extended Kalman filter; mAb, monoclonal antibody; NPM, *N*-(1-pyrenyl)maleimide.

2.2 | Data organization

The full data set containing the reaction time courses from the RP-UHPLC and the spectra for each run was split into training and test subsets. The split was done based on a Kennard–Stone algorithm (Kennard & Stone, 1969) using the first-derivative mean spectra of each run to ensure that the choice of the calibration set is chosen objectively and spans the full experimental design space. The training subset was used for hyperparameter tuning of the chemometric models, calibration of the kinetic model parameters, and tuning of the EKF. The test subset was used for model performance evaluation and selection. The fed-batch runs were evaluated separately. The subsets used for chemometric model development consisted of 122 and 56 reference data points for training/cross-validation and testing, respectively. An overview of all conducted experimental runs is shown in Table 1.

2.3 | Chemometric model development

Initially, all absorbance spectra were averaged over an interval of 5 s and aligned with the RP-UPHLC reference data. Before regression modeling, the spectra were mean-centered and normalized by the maximum absorbance at 250 nm of the training subset. Several chemometric models were calibrated and compared to determine the best-performing option. In the following,

all mentions of cross-validation refer to a *leave-one-run-out* rotation.

PLSR models were calibrated in combination with a Savitzky–Golay derivative filter (SGF) or a variable selection algorithm. The number of latent variables, the window size, and the order of derivative for the SGF were tuned by employing a cross-validated grid search. The variable selection algorithm was adapted from Mehmood et al. (2012) using the regression coefficients of the PLSR as an importance metric and evaluated using cross-validation.

For GPR models, multiple kernel functions were screened and combined with an SGF or a variable selection algorithm analog to the one described above. Multiple combinations of a linear, a radial-basis function (RBF), and an automatic relevance determination (ARD) kernel were evaluated. The final implementation of the kernel is given in Supporting Information A.1. More theoretical information on GPR modeling can further be found in Supporting Information A.1. The kernel parameters were estimated by expectation maximization as implemented in Pedregosa et al. (2011). The implemented variable selection algorithm was adapted from Paananen et al. (2019). For the exact mathematical procedure the reader is referred to the cited literature. The computed variable sensitivities are used as an importance metric, ranked, and selected as performed for the PLSR variable selection algorithm.

All calibrated models were evaluated based on their root-mean-square error of cross-validation (*RMSECV*) and prediction (*RMSEP*). A detailed definition of the error metric used in this study can be found in Supporting Information A.4. All chemometric models presented in this study were implemented with recent versions of Python 3.8, NumPy ≥1.20 (Harris et al., 2020), pandas ≥1.3 (The Pandas Development Team, 2020), scikit-learn ≥1.1 (Pedregosa et al., 2011), and related packages.

2.4 | Kinetic model calibration

A previously developed kinetic model for the site-directed conjugation reaction (see Andris et al., 2019) was applied in this study to be adopted in the EKF framework. A detailed explanation of the model and the full ordinary differential equation (ODE) system can be found in Supporting Information A.2. The kinetic rates k_1, k_2, and k_3 were estimated by nonlinear least-squares optimization using the *lsqnonlin* solver in MATLAB 2019b (The Mathworks Inc.). The initial values were set to 0.8mMs^{-1}, 1.4mMs^{-1}, and 0.01s^{-1} for k_1, k_2, and k_3, respectively. Assuming that measurement errors in each sample are independent and normally distributed with zero mean and variance σ^2, the parameter covariance matrix $\mathbf{C}_k$ is computed according to Equation (1):

$$\mathbf{C}_k = cov(\mathbf{k}) = \sigma^2(\mathbf{J}^T\,\mathbf{J})^{-1}, \tag{1}$$

where $\mathbf{J}$ is the Jacobian with regard to the kinetic rates $\mathbf{k}$ and σ^2 is the estimated error variance from the nonlinear estimation. The error variance σ^2 is approximated by $RMSEP^2$ as suggested in literature (Feidl et al., 2019; Valappil & Georgakis, 2000).

2.5 | Extended Kalman filter

2.5.1 | Terminology and concept

As the herein applied models are calibrated using different target variables, the terminology is shortly introduced: The un-, mono-, and biconjugated ADC species will be referred to as mAb, Conj1, and Conj2 for the remainder of this study. It is worth mentioning that "mAb" corresponds to the sum of the three subspecies with different amounts of active cysteines: mAb_{2c}, mAb_{1c}, and mAb_{0c}, and "Conj1" to the sum of the two subspecies $Conj1_{1c}$ and $Conj1_{0c}$ of the ODE system. While the kinetic model continuously predicts all species, the chemometric model converts the absorbance spectra to the amount of conjugated drug at discrete time points t_k. The same values can also be computed from the single species predictions by $c_{drug,Conj} = c_{Conj1} + 2c_{Conj2}$, where c_{Conj1} and c_{Conj2} are the concentration of mono- and biconjugated mAbs, respectively. To enable the sensor model to update the individual ADC species in the ODE system, a special measurement function was chosen (see Section 2.5.2). Furthermore, both model outputs are thought to be noisy as they are subject to errors caused by model assumptions, plant-model mismatch, and experimental handling. An EKF framework was designed to combine the noisy model outputs and generate a more accurate estimate than either of the two sources. Since the two used models are not producing estimates for the same set of species, the EKF acts as a soft-sensor updating several unmeasurable species by an indirect measurement of the sum of conjugated drugs. For better understanding, the overall concept and data organization is summarized in Figure 1.

2.5.2 | Mathematical formulation

The theory of the EKF framework is well established and can be found, for example, in Simon (2006). In principle, all species are represented in a state vector $\mathbf{x}$. The seven ODEs of the kinetic model act as state transition function f propagating the state vector $\mathbf{x}$ through time according to Equation (2) given deterministic inputs $u(t)$ and the model parameters $\boldsymbol{\theta}$:

$$\frac{\partial \mathbf{x}}{\partial t} = f(\mathbf{x}(t), u(t), \boldsymbol{\theta}) + \mathbf{q}(t) \tag{2}$$

with the state vector $\mathbf{x}$ for all seven reacting species in the ODE system defined as

$$\mathbf{x}(t) = [c_{mAb_{2c}}\ c_{mAb_{1c}}\ c_{mAb_{0c}}\ c_{Conj1_{1c}}\ c_{Conj1_{0c}}\ c_{Conj2}\ c_{NPM}]^T, \tag{3}$$

where c_{mAb} denotes the concentration of unconjugated mAbs with the respective number of available cysteines (c) and the conjugates with the respective number of conjugated NPM molecules. The free NPM concentration is given as c_{NPM}. Measurements $y(t_k)$ available at discrete time points t_k are related to states vector using the selected measurement function h which is given in Equation (4). The process

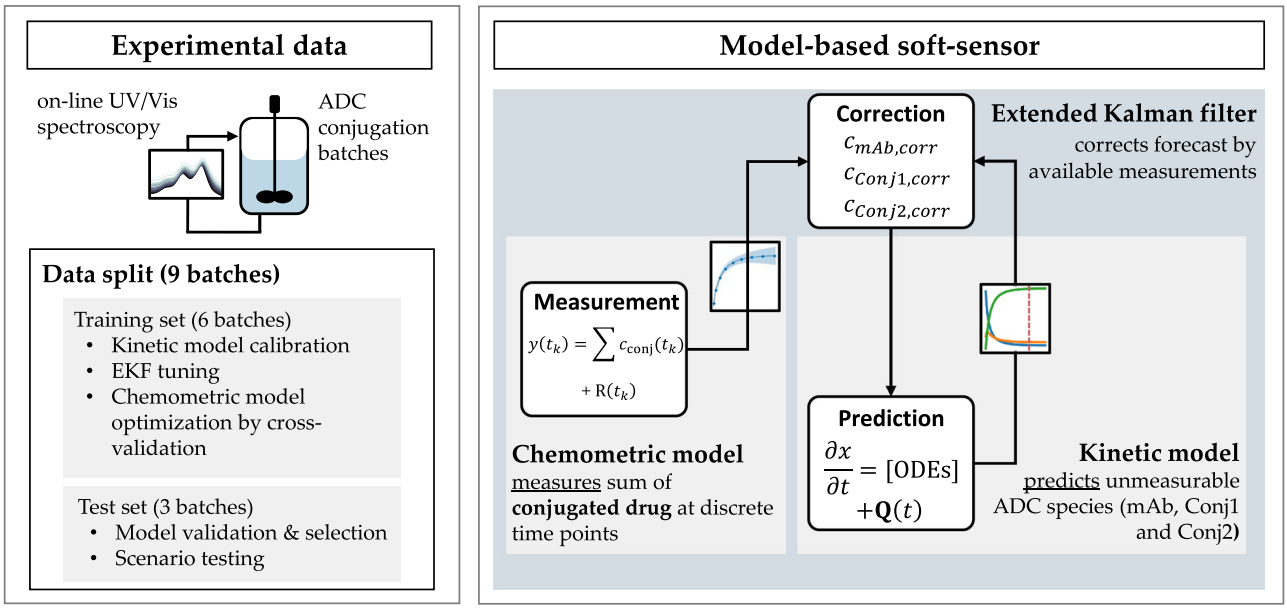

FIGURE 1 Schematic overview of the purposes of the data sets and the methodology of the designed model-based soft-sensor. ADC, antibody-drug conjugate; Conj1, monoconjugated; Conj2, biconjugated; EKF, extended Kalman filter; mAb, monoclonal antibody; ODE, ordinary differential equation; UV, ultraviolet; Vis, visible.

noise $\mathbf{q}(t)$ and the measurement noise $r(t)$ are assumed to be additive, zero-mean Gaussian noise:

$$\begin{aligned} y(t_k) &= h(\mathbf{x}(t_k)) + r(t_k) \\ &= [0\ 0\ 0\ 1\ 1\ 2\ 0] \cdot \mathbf{x}(t_k) + r(t_k). \end{aligned} \tag{4}$$

When a new measurement $y(t_k)$ is available, the measurement function effectively distributes the measured value to the fourth, fifth, and sixth reaction states, that is, the species with available binding sites. This measurement function was chosen as the available chemometric sensor model solely accounts for the amount of conjugated drug which is equivalent to the sum of the respective reaction states (cf. Section 2.3). A detailed mathematical formulation of the EKF working principle and its elementary functions can be found in Supporting Information A.3. The EKF was implemented using the *ExtendedKalmanFilter* class in MATLAB 2019b.

2.5.3 | Process noise covariance

The process noise $\mathbf{q}(t)$ is parameterized by the covariance matrix $\mathbf{Q}$ which provides a measure of uncertainty and cross-correlation for the states $\mathbf{x}$. For nonlinear process models, as used here, the implementation of a time-variant process noise covariance matrix $\mathbf{Q}(t)$ is crucial to capture process dynamics and allow the EKF to flexibly adjust the weighting between the models. In literature Valappil and Georgakis (2000), two methods have been suggested to implement a time-variant process noise covariance: A linearized approach that computes a time-dependent process noise covariance matrix $\mathbf{Q}_{\mathrm{lin}}(t)$ based on gradient evaluations with regard to the model parameters. This method uses the Jacobian $\mathbf{J}_{\theta,\mathrm{nom}}$ with regard to the nominal model parameters $\boldsymbol{\theta}_{\mathrm{nom}}$, namely, the kinetic rates $\mathbf{k}$, as well as their covariance matrix $\mathbf{C}_{\boldsymbol{\theta}}$ and computes $\mathbf{Q}(t)$ according to Equation (5):

$$\mathbf{Q}_{\mathrm{lin}}(t) = \mathbf{J}_{\boldsymbol{\theta},\mathrm{nom}}\ \mathbf{C}_{\boldsymbol{\theta}}\ \mathbf{J}^{T}_{\boldsymbol{\theta},\mathrm{nom}}. \tag{5}$$

Alternatively, the Monte Carlo (MC) approach uses linear estimates of the parameter covariance matrix $\mathbf{C}_{\boldsymbol{\theta}}$ to span a normal distribution for all model parameters θ_i with mean μ and variance σ as $\mathcal{N}_{\theta_i}(\mu = \theta_{i,\mathrm{nom}}, \sigma^2 = \mathbf{C}_{\theta_i,\theta_i})$. A definite number of n random samples are drawn from the distribution and the sampled parameter values are used to propagate the states $\mathbf{x}$ to the next time step t_{k+1}. The time-dependent process noise covariance $\mathbf{Q}_{\mathrm{MC}}$ is then computed as the covariance matrix of the mean-centered residuals $\tilde{\mathbf{w}}$ according to Equation (6):

$$\mathbf{Q}_{\mathrm{MC}}(t) = cov(\tilde{\mathbf{w}}) = cov(w_n - \overline{\mathbf{w}}) \tag{6}$$

with $\overline{\mathbf{w}}$ being the mean of the residuals and the residuals w_n being calculated by

$$w_n = f(t_{k+1}, \boldsymbol{\theta}_n) - f(t_{k+1}, \boldsymbol{\theta}_{\mathrm{nom}}), \tag{7}$$

where $\boldsymbol{\theta}_{\mathrm{nom}}$ and $\boldsymbol{\theta}_n$ denote the nominal and the MC samples of the model parameters, respectively.

2.5.4 | Measurement noise covariance

The measurement noise covariance matrix $\mathbf{R}$ is the sensor-equivalent of $\mathbf{Q}$. As the chemometric sensor in this study only accounts for the

amount of conjugated drug, R is a scalar. In contrast to the literature, for example, Feidl et al. (2019) and Narayanan et al. (2020), a time-variant measurement noise covariance $R(t_k)$ may be expressed using the predictive variance of the GPR evaluated at all discrete time points t_k. Further information on the computation of the predictive variance is given in Supporting Information A.1.

2.5.5 | EKF tuning

When updating the reaction states using a newly available measurement, the EKF combines the information stored in $\mathbf{Q}$ and R to produce an improved estimate of all states. Since the absolute values of $\mathbf{Q}$ and R depend on the estimation method, the process noise covariance is scaled to $\mathbf{Q}^* = k_Q\mathbf{Q}$ by a scaling factor k_Q specific for each method. To ensure an equal weighting of all ADC species i, the factor k_Q is optimized with regard to the normalized error *nRMSE* averaged over all species. The $nRMSE_i$ for each species is calculated according to Equation (8):

$$nRMSE_i = \frac{RMSE_i}{\bar{c}_i} \tag{8}$$

where $\bar{c}_i$ is the respective mean concentration of each species in each run and $RMSE_i$ being calculated for each species individually as defined in Supporting Information A.4. The total *nRMSE* is obtained by averaging over all runs and species.

2.6 | Study design

2.6.1 | Model-based soft-sensor discrimination

The best-performing chemometric model was selected based on the methodology described in Section 2.3 and a kinetic model was calibrated using the procedures given in Section 2.4. On the basis of the selected models, the influence of the different computation approaches for $\mathbf{Q}$ and R was compared. The performance was evaluated based on the prediction error regarding the unmeasurable species in the training and test subsets using the GPR predictions for the conjugated drug concentration. For each approach, the scaling factor k_Q was tuned independently.

2.6.2 | Challenge scenarios and transfer to fed-batch

The best-performing option was selected and further challenged in three scenarios. In scenario 1, the kinetic model was initialized with randomly sampled mAb and NPM concentrations with a standard deviation of 2% of the nominal mAb/NPM concentration. In scenario 2, the kinetic model was initialized with a randomly sampled cysteine distribution within the experimentally determined standard deviation (see Section 2.4). In scenario 3, the sensor was fed with noisy absorbance spectra. The spectral data for the test subset was modified by additive Gaussian white noise with a standard deviation of 0.02% of the respective absorption value. Additionally, a random, normally distributed offset was introduced with a standard deviation of 5% of the absorbance at 250 nm. In all scenarios, the random sampling and subsequent model evaluation were conducted for 100 repetitions.

Finally, the investigated models were transferred to fed-batch mode. The kinetic model was extended with a feeding term to account for the continuous addition of free drug solution. The GPR model was supplied with the on-line absorbance spectra which were treated as described in Section 2.3.

3 | RESULTS

3.1 | Chemometric model development

Multiple combinations of PLSR and GPR models with different preprocessing strategies were screened in this study and evaluated with regard to their predictive performance on the training and test set. A summary of the results is presented in Table 2. The PLSR model in combination with an SGF achieved the lowest *RMSECV*, while the GPR in combination with variable selection (GPR–VS) achieved the lowest *RMSEP*. Overall, the GPR–VS model performed best with regard to both error metrics and improved accuracy by 18% compared with the PLSR model in combination with variable selection (PLSR–VS). Hence for further evaluation, only the GPR–VS model is considered. The resulting selected variables for the GPR–VS model are shown in Supporting Information A.5.2.

The predictive mean and variance for the (cross-validated) training and test set are shown in Figure 2. The predictions for the GPR–VS model show good alignment with the reference data for the training and test set with R^2 of 0.968 and 0.973, respectively. For the cross-validation, the predictions for runs 3, 7, and 9 deviate from the reference data towards the end of the reaction. In run 3, an offset of the predictions from the reference data can be observed starting at roughly 500 s. For runs 7 and 9, the predictions reach a maximum of 17.5 μμM and decrease towards the end of the reaction to 16.0 μM. This tendency is not present in the reference data. The

TABLE 2 Overview of predictive performance of selected chemometric models.

Model	*RMSECV* (μM)	*RMSEP* (μM)
PLSR–SGF	0.31	0.84
PLSR–VS	0.54	0.51
GPR–SGF	0.61	0.82
GPR–VS	0.43	0.43

Note: Both PLSR and GPR models were combined with SGF and a model-specific VS method.

Abbreviations: GPR, Gaussian process regression; PLSR, partial least-squares regression; RMSECV, root-mean-square error of cross-validation; RMSEP, root-mean-square error of prediction; SGF, Savitzky–Golay derivative filter; VS, variable selection.

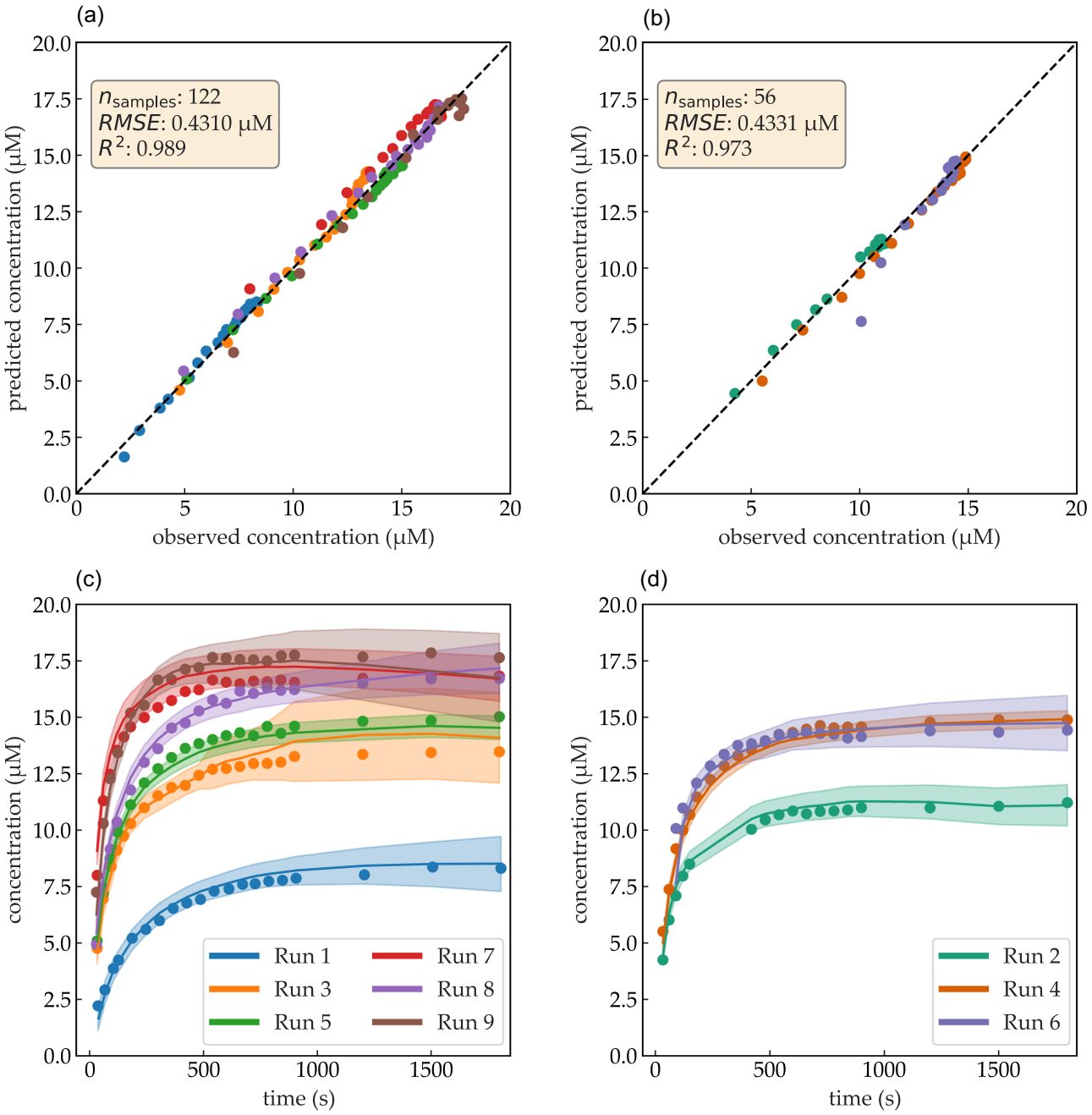

FIGURE 2 Predictions for the GPR–VS model. (a, b) Parity plots for the cross-validation and test subset. The tie line is depicted as dashed line and represents ideal model predictions. The data points are colored by the run numbers. (c, d) Predicted conjugated drug concentration over time. Reference data are illustrated as circles and colored by runs. The GPR–VS predictive mean is shown as solid lines, while the 95% confidence interval is indicated by the colored areas. The 95% confidence interval was approximated by 1.96 standard deviations. Note that the GPR–VS model was only evaluated at the reference data points and the lines are linearly interpolated to guide the eye. GPR, Gaussian process regression; RMSE, root-mean-square error; VS, variable selection.

95% confidence intervals were approximated by ±1.96 predictive standard deviations σ^* and show a growing uncertainty with time for all runs in all three subsets. For the cross-validation, the confidence intervals of runs 3, 7, and 9 are considerably wider than for the remaining runs indicating an increased uncertainty of the GPR–VS model. For the test subset, the predictions for run 4 deviate only 1 µM from the mean at a 95% confidence as opposed to 2 µM for run 6.

3.2 | Soft-sensor development

3.2.1 | Kinetic model calibration

Initially, the end-point compositions of runs with a drug excess of 3× or higher were used to determine the percentages of pre-inactivated cysteines as proposed by Andris et al. (2019).

This resulted in 85.98%, 10.14%, and 3.93% of mAbs with two, one, and none available cysteines, respectively. On the basis of this distribution and the training data, the kinetic rates k_1, k_2, and k_3 were estimated to 0.784 $mM^{-1}s^{-1}$, 1.625 $mM^{-1}s^{-1}$, and 0.0012 s^{-1}. The parameter covariance matrix was calculated by Equation (1) for the computation of the process noise covariance for the EKF. The calibrated model resulted in R^2 values for mAb, Conj1, and Conj2 of 94%, 95%, and 98% in the training data, and of 93%, 94%, and 96% for the test data.

3.2.2 | Tuning and selection of covariance matrices

In this section, the performance of the EKF with different approaches for the calculation of process noise **Q** and measurement noise *R* are compared. Each combination was tuned individually by the optimization of the scaling factor k_Q. The effect of the scaling factor on the *nRMSE* is exemplarily shown for one **Q**/*R* combination in Figure 3. Values in the range of $k_Q = 10^9 \ldots 10^{11}$ result in proper tuning so that the EKF beneficially considers both models to minimize the *RMSE* for all species. The optimal k_Q varies slightly in each combination of **Q** and *R* as shown in Table 3 together with the accomplished *RMSE* reduction by the EKF with regard to the kinetic model predictions. It could be demonstrated that the linearized method does not lead to an overall improvement, while the MC approach achieves an overall *RMSE* reduction in both subsets. Only the prediction error for Conj1 in the training set slightly increases. In summary, this approach yields an averaged *RMSE* reduction of 8.7% and 17.2% for the training and test set, respectively. The time-variant computation of *R* could further improve the performance by up to 4% for all species compared with the constant setting. In conclusion, the EKF using the $\mathbf{Q}_{MC}$ approach in combination with the *R* variable setting presents the most valuable combination and is used for the remainder of this study.

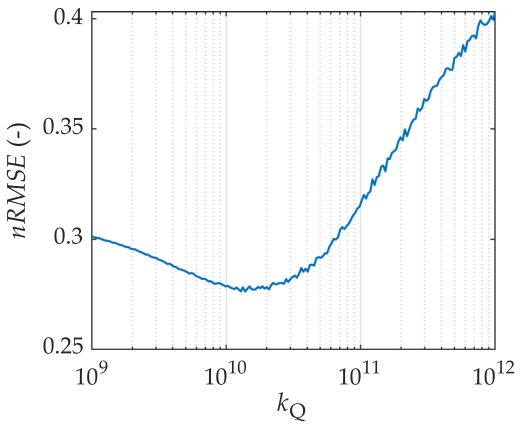

FIGURE 3 EKF prediction error regarding the *nRMSE* during k_Q tuning. EKF, extended Kalman filter; nRMSE, normalized root-mean-square error.

TABLE 3 Comparison of the three studied combinations of **Q** and *R* regarding their accomplished *RMSE* reduction compared with the kinetic model in training and test set.

Q	R	Tuned k_Q	Species	*RMSE* reduction (%) Training	Test
Linearized	Variable	9.48	mAb	2.19	2.11
			Conj1	−1.98	3.78
			Conj2	0.19	0.87
MC	Variable	10.11	mAb	17.37	11.95
			Conj1	−6.67	23.30
			Conj2	21.83	11.07
MC	Constant	10.25	mAb	17.78	8.04
			Conj1	−8.86	21.04
			Conj2	23.24	7.06

Abbreviations: Conj1, monoconjugated; Conj2, biconjugated; mAb, monoclonal antibody; MC, Monte Carlo; RMSE, root-mean-square error.

3.2.3 | EKF soft-sensor performance

The behavior and the soft-sensing capability of the selected EKF combination are exemplarily demonstrated for run 3. Figure 4a,b presents the predicted conjugated drug concentration as measurable states from the GPR, kinetic model, and EKF, as well as the predicted ADC species as the unmeasurable states from the kinetic model and the EKF, respectively. The reference data are overlayed in both figures. During the initial phase until approximately 500 s, it can be observed that the kinetic model predicts the conjugation reaction rate to be faster compared with the GPR sensor, which closely follows the reference data. For the unmeasurable states, the same trend is apparent for Conj1 and Conj2. After approximately 200 s, the EKF starts to adjust the trajectory by following the GPR estimates. Towards the end, the kinetic predictions converge to the reference data, while the GPR sensor overestimates the conjugation progress after roughly 600 s. In the same interval, the EKF switches and clearly follows the kinetic model to be aligned with the reference data. This trend can be similarly followed in Figure 4c,d, where the normalized time-variant process and sensor noise covariance, and the resulting Kalman gain are displayed. In the beginning, the measurement noise is comparably lower than the process noise regarding all ADC species. This leads the Kalman gain, and consequently the EKF, to rely more on the GPR predictions with values close to 1. Over the course of the reaction, the process noise is continuously decreasing as the reaction is slowing down and the GPR noise increases conversely relative to the measurement noise. This causes the Kalman gain to shift towards 0 resulting in the EKF relying more strongly on the kinetic model towards the end of the reaction. A similar adjustment dynamic could be observed for all runs in the training and test set and shows the advantage of using time-variant process and measurement noises. The fact that the EKF is converging towards the kinetic model at the end of the reaction, comes along

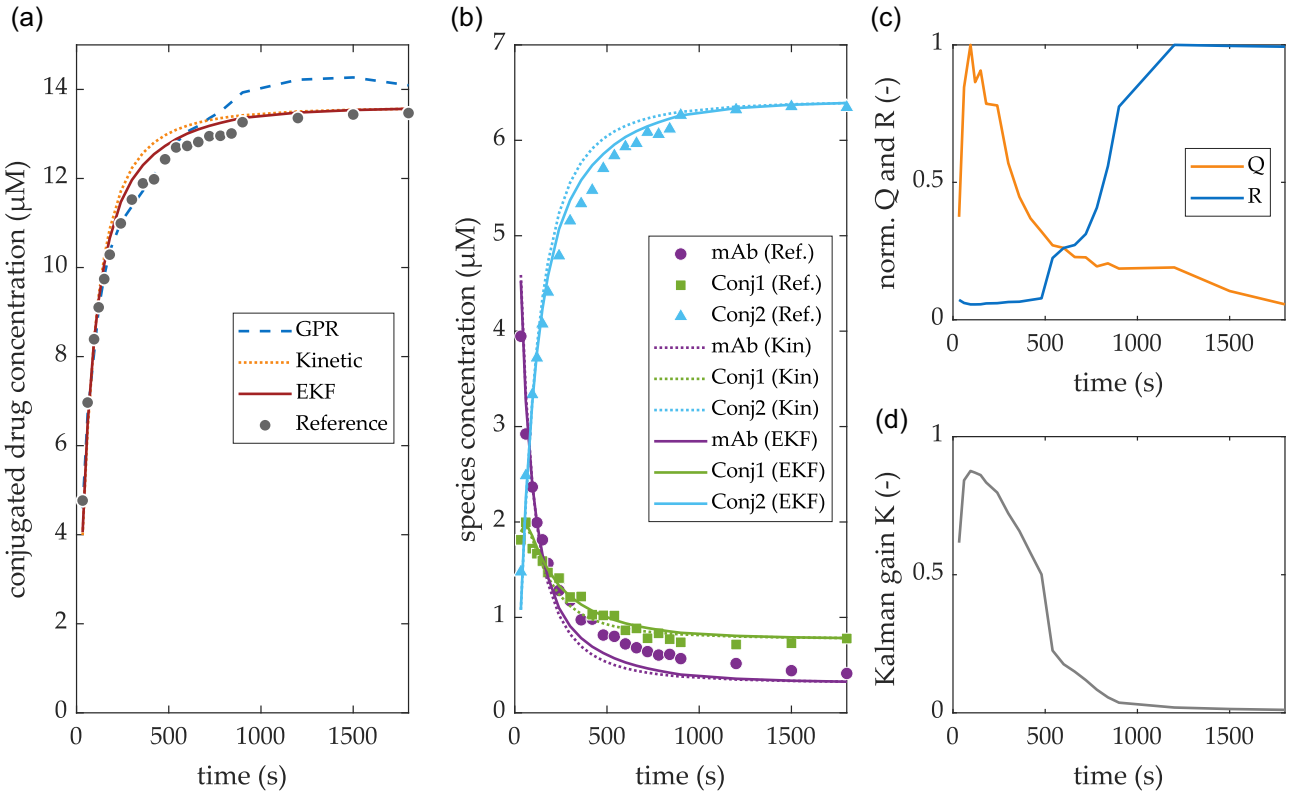

FIGURE 4 Model predictions and calculated process and sensor noise covariance for run 3. (a) Comparison of the conjugated drug concentration as the measured state for the GPR sensor, the kinetic model, and the EKF. The reference data from the RP-UHPLC are displayed as circles. (b) Comparison of the predictions for the three ADC species (mAb, Conj1, and Conj2) as the unmeasurable states for the kinetic model and the EKF. (c) Time course of the normalized process and sensor noise covariance $\mathbf{Q}_{MC}$ and R_{var} and (d) of the Kalman gain K. ADC, antibody-drug conjugate; Conj1, monoconjugated; Conj2, biconjugated; EKF, extended Kalman filter; GPR, Gaussian process regression; mAb, monoclonal antibody; MC, Monte Carlo.

with a decreasing state covariance which may be considered an internal EKF uncertainty. In Supporting Information A.6, the state covariances for all species are shown for the test subset. For runs, where the Kalman gain is close to 0 at the end of the reaction, the variance in the EKF estimates is small. Whereas, in Run 4 where the Kalman gain plateaus at approximately 0.75 due to the high confidence of the GPR sensor, a larger variance is observable. This variance is in agreement with additional data from replicate runs with the same starting conditions which were not used otherwise within this study. Furthermore, the model and EKF predictions for all training and test runs are depicted in Supporting Information A.6.

3.2.4 | Soft-sensor challenge scenarios

Three challenging scenarios were created to further investigate the robustness and accuracy of the designed EKF in comparison to the sole kinetic and sensor model. The resulting statistical distributions of the *RMSEP* after 100 repetitions are displayed as box plots in Figure 5.

The upper and lower rows present the comparison of the *RMSEP* distributions with regard to the ADC species and to the conjugated drug, respectively. In the first column, the effect of noisy initial concentrations of mAb and NPM (scenario 1) is depicted. By fusing the information from the unmodified GPR sensor with the faulty initialized kinetic model, the EKF achieves a reduction of the mean prediction error by 10.7%, 21.5%, and 18.2% compared with the single species predictions, mAb, Conj1, and Conj2, from the kinetic model. However, the interquartile ranges of the displayed boxes are overlapping and their span is similar for all species indicating that the noisy concentrations are projected to the EKF estimates. Regarding the conjugated drug, the EKF reduces the error by 18.0% compared with the kinetic model while not reaching the baseline of the GPR error entirely. In the second column, the effect of deviation in the activated cysteines (scenario 2) is shown. All achieved EKF reductions are similar to the results in scenario 1 with overall more dense distributions of the noisy data in both the kinetic model and EKF estimates. The third column demonstrates the effect of noise on the sensor data (scenario 3). With regard to the *RMSEP* of the single species, a reduction in the range of 12%–25% compared with the kinetic model is still accomplished by the EKF. In this case, the EKF effectively reduces the error in the conjugated drug by 13% compared with the noisy GPR as measured by the mean of both distributions.

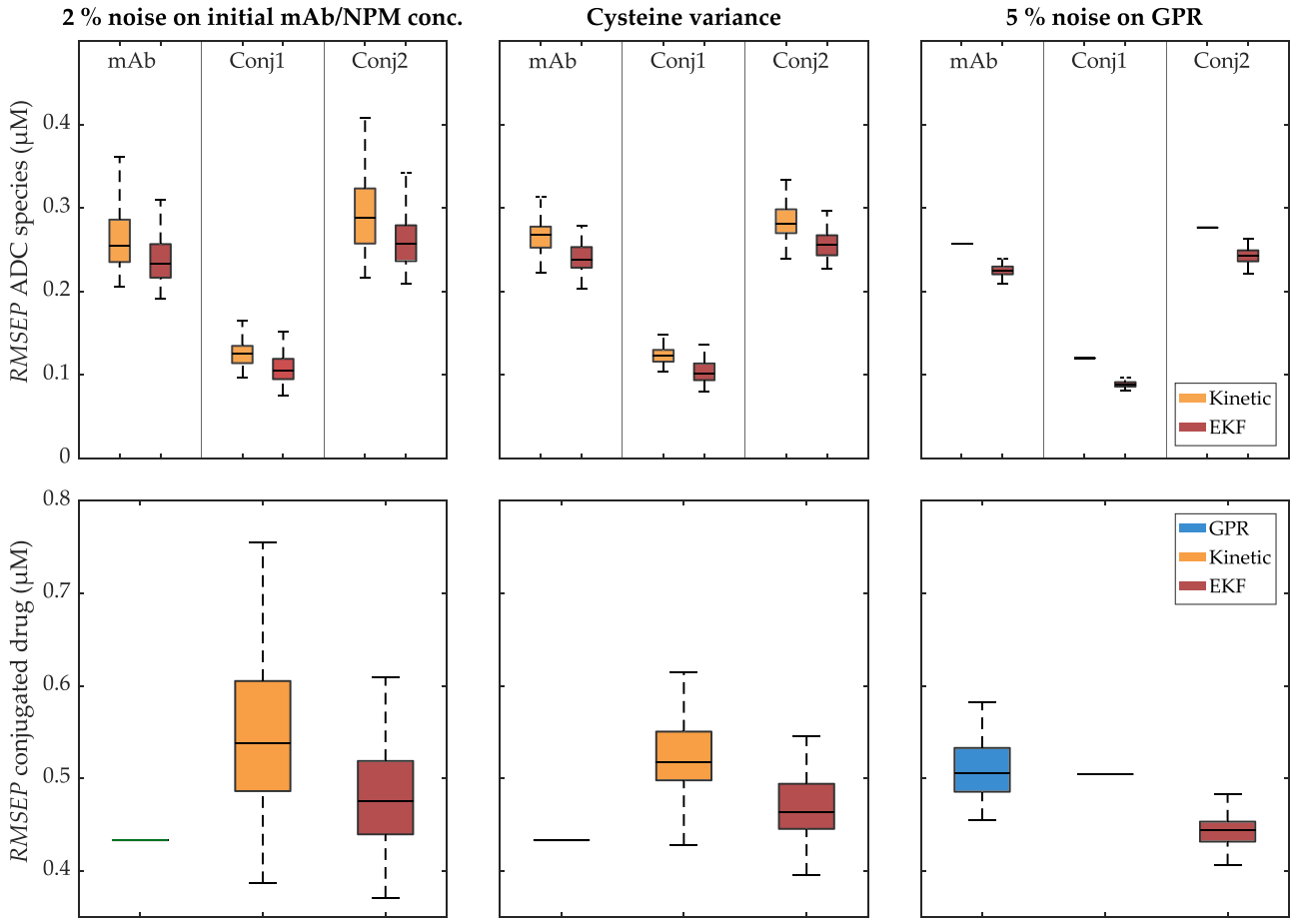

FIGURE 5 Comparison of the *RMSEP* values of the GPR, the kinetic model, and the EKF obtained from 100 simulations for each scenario. The upper row shows the *RMSEP* values regarding all ADC species and the lower row the *RMSEP* values regarding the conjugated drug concentration. ADC, antibody-drug conjugate; Conj1, monoconjugated; Conj2, biconjugated; EKF, extended Kalman filter; GPR, Gaussian process regression; mAb, monoclonal antibody; NPM, *N*-(1-pyrenyl)maleimide; RMSE, root-mean-square error.

3.2.5 | Transfer to fed-batch mode

Finally, the EKF was applied to fed-batch runs to test the ability of the EKF to perform under different reaction dynamics. In these runs, the relevant wavelengths exhibit a different behavior compared with the experiments operated in batch mode due to a timely overlay of spectral effects from the conjugation reaction and the NPM feed. A comparison of the spectral data from fed-batch and batch mode can be found in Supporting Information A.5.1. To compensate for this changing effect in the absorbance spectra, the GPR was recalibrated with fed-batch run 10 being part of the training subset. Figure 6 illustrates the model predictions for the fed-batch run 11. The modified kinetic model closely aligns with the reference data until a reaction time of 750 s. Towards the end of the reaction, the modified kinetic model undererstimates the conjugated drug concentration, while the GPR sensor follows the reference data. Combining the information of both models, the EKF shows minor improvements during the whole reaction and all species while converging to the predictions of the modified kinetic model towards the end of the reaction. Thus, the EKF errors for all species are improved compared with the kinetic model by 13.2%, 0.2%, and 16.4%. For the conjugated drug, the EKF follows the kinetic model after approximately 750 s and hence results in a relative error increase by 37.9% compared with the GPR model.

4 | DISCUSSION

4.1 | Chemometric model development

4.1.1 | Preprocessing and hyperparameter optimization

The developed GPR model presented in this study uses a composite kernel function and a variable selection algorithm based on the Kullback–Leibler divergence. The model showed elevated predictive performance compared with commonly applied PLSR models and

other combinations of kernel functions and preprocessing methods for the GPR (see Table 2).

SGF did not improve the predictive capability for either of the two studied regression models. In Andris et al. (2018), a baseline correction and subsequent first-derivative filter were found to achieve the best performance, while the herein-used GPR model does not rely on any further preprocessing method before variable selection. The herein-implemented variable selection methods notably improved the predictive accuracy for both the PLSR and the GPR model. This suggests that the absorbance spectra contain uninformative regions which can effectively be removed by the employed algorithms. Particularly wavelengths above 350 nm were found to have a negative impact on the sensor model performances due to the presence of spectral artifacts probably caused by the surrogate drug molecule (cf. Supporting Information A5).

The implemented kernel function which is a combination of the linear and the RBF kernel has previously been shown to be a viable combination for chemometric purposes in Chen et al. (2007). In other cases, the RBF kernel alone was used for GPR models using near-infrared and Raman spectroscopy data (Chen et al., 2007; Cui & Fearn, 2017; Tulsyan et al., 2021). In general, the authors consider the usage of covariance kernels beneficial to model the nonlinear observed spectral shift and baseline drift of the absorbance spectra compared with the linear subspace projections the PLSR models are based upon. It has previously been shown, that peak shifts which may be resembled by closely overlapping absorption maxima of multiple components have a strong influence on the regression vector in PLS modeling which may be a reason for the inferior performance compared with kernel-based models (Kjeldahl & Bro, 2010; Witjes et al., 2000).

4.1.2 | Quantification of spectra-based uncertainty

When the GPR model is evaluated at an unseen spectrum, the kernel function computes the covariance with all available training instances and uses that information to quantify the prediction estimate and the associated uncertainty. The uncertainty estimate is hence influenced by the distribution of available data points along the process as well as the spectral differences among the samples. As emphasized in Figure 2, the trajectories of runs 4 and 6 closely resemble one another. However, the associated uncertainty estimate is considerably higher for run 6. This observation is in agreement with the experimental data as run 4 is a replicate of run 5 and has been conducted on the same day. Run 6, instead, has a different initial condition than all training runs suggesting higher uncertainty due to unseen spectral variation. Similarly, runs 3 and 9 supposedly exhibit spectral features causing the covariance function to produce a wider confidence interval than the other runs. In fact, the variance increases towards the end of the reaction when the conjugation reaction is considerably slowing down. This may be influenced by an irregular chemical degradation of the NPM molecule which has been observed in experiments with pure NPM solutions (cf. Supporting Information A.5.1). However, this effect may not occur for other molecules such as real payloads due to incomparable spectral properties (Ouyang, 2013).

It should be noted that GPR models are a comparably novel method in chemometrics and the physical validity of the computed confidence intervals has not been extensively studied in the literature. As the confidence intervals are computed with regard to both the experimental variance of the reference and the spectral data, the predictive uncertainty of the GPR cannot be directly

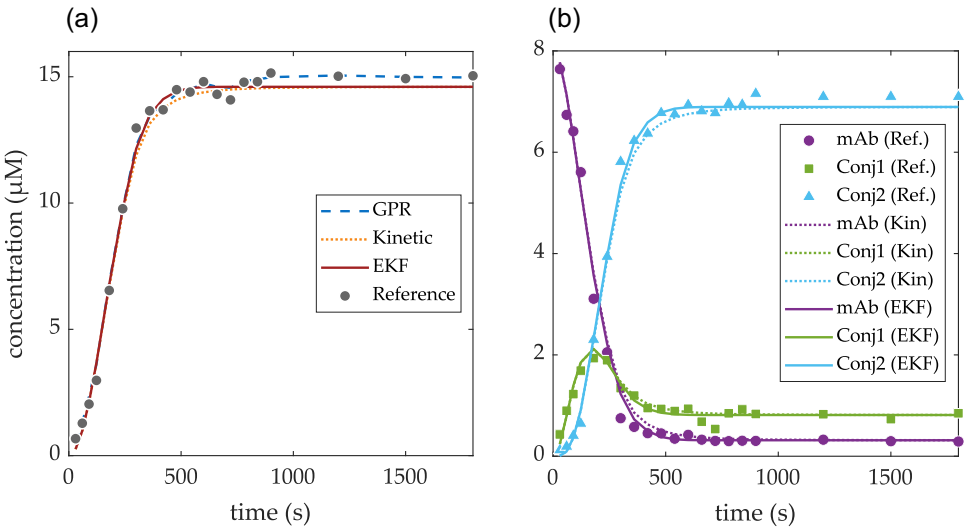

FIGURE 6 Model predictions for fed-batch run 11 for the concentration of conjugated drug (a) and the ADC reaction species (b). To compensate for a changing effect in the absorbance spectra compared with the experiments operated in batch mode, the GPR sensor model was recalibrated with run 10 being part of the training subset. ADC, antibody-drug conjugate; Conj1, monoconjugated; Conj2, biconjugated; EKF, extended Kalman filter; GPR, Gaussian process regression; mAb, monoclonal antibody.

compared with the variance of the reference data alone. Nonetheless, the implemented method provides further insights into the quality of the spectral data and the relation with the training instances and exceeds conventional models in terms of predictive accuracy. The GPR is therefore considered sufficiently accurate to be used as a sensor within the soft-sensor framework. In future, separate studies specifically designed for the validation of predictive uncertainty estimates should be conducted.

4.2 | Kinetic model calibration

The identified kinetic reaction rates are in agreement with the ones found in Andris et al. (2019) with k_1 and k_2 being 1.6% and 9.2% below and above the values in the referenced study. The difference might originate from minor differences in the determined cysteine distribution or the overall experimental or analytical procedure. The distribution of pre-inactivated cysteines is a necessary model parameter to account for the actual amount of available bindings sites on the oxidized mAb which in practice are lower than two. This phenomenon is already discussed in detail in Cao et al. (2019).

4.3 | Soft-sensor development

The selection and appropriate tuning of the process and sensor noise covariance play a crucial role in the development of an optimal state estimator, such as an EKF (Elsheikh et al., 2021). It has been observed that the EKF mainly especially selects the Conj1 state to compensate for the deviation when k_Q is set too high. This unfavorable adjustment is due to the underlying physical constraints from the kinetic model ODE system and could be effectively minimized by using a normalized *RMSE* for tuning.

Valappil and Georgakis (2000) stated that the applicability of the computation method for $\mathbf{Q}(\mathbf{t})$ has to be systemically tested for each individual case. In our study, the MC method outperforms the linearized method in terms of prediction error (see Table 3). Remarkably, the EKF using the MC method improves the errors for all ADC species in the test set by at least 11%. It appears that the sampling-based MC method is more adequate and robust in the case of the nonlinear ADC reaction. However, it should be noted that the parameter distributions used for the MC method are solely approximations. For increasingly nonlinear models, Gaussian distributions may no longer be assumed for all model parameters. Another disadvantage of the MC method may be the diminishing process variance and hence an overconfident EKF state estimate towards the steady state of the reaction. If applicable, Bayesian inference methods could be used to assess more accurate estimations of the parameter distributions and provide a means for more conservative variance estimation.

The sensor noise covariance R is commonly set to a constant scalar and derived from the calibration error of the sensor model (Elsheikh et al., 2021; Feidl et al., 2019; Narayanan et al., 2020). The usage of the GPR model firstly allows the implementation of a time-varying sensor noise covariance. As the uncertainty derived from the GPR increases with time, the sensor noise resembles this trend and the EKF counteracts to this by relying more on the kinetic model towards the end of the reaction. In comparison to a time-invariant R, the variable approach achieves small improvements, while the setting of $\mathbf{Q}$ majorly contributes to the presented error reduction. However, it should be emphasized that in this case the GPR provides an estimate of the cumulative conjugated drug and does not affect all states individually. It would be hence interesting to evaluate the performance of the variable sensor noise in cases where direct measurements of all states are available. Moreover, another drawback of the implemented EKF is the steady-state constraints, that is, the experimentally derived cysteine distribution assumed for all batches, which make the EKF rather stiff towards the end of the reaction. This is a known problem of the EKF (Tulsyan et al., 2021) when dealing with only a few states being available for measurement as well as for nonlinear models. In theory, this problem could be circumvented with other more flexible state estimation algorithms, such as particle filters as implemented in Golabgir and Herwig (2016) and Stelzer et al. (2017) or unconstrained Kalman filter algorithms as used in Kolås et al. (2009) and Simutis and Lübbert (2017) which were beyond the scope of this study.

The three challenging scenarios prove that even for a faulty EKF initialization and a noisy sensor model, the designed soft-sensor consistently leads to more accurate predictions compared with the sole chemometric or kinetic models. In cases of large offsets of the kinetic model, the EKF could be averted to rely on the sensor model instead, which is demonstrated in challenge scenarios 1 and 2. In case of noisy sensor inputs, the EKF produces less noisy predictions, also regarding the unmeasured species. As discussed before, the predictions of both the process and the sensor model are dependent on the variations in the absorbance spectra or initial conditions. The EKF was shown to be able to compensate foreseeable deviations and hence demonstrated improved robustness which is essential for the successful deployment of the soft-sensor.

Finally, the adaptability of the soft-sensor was investigated by transfer to the fed-batch operation mode. Under fed-batch conditions, the EKF also performed adequately while an extended retuning was not required. Despite the necessary addition of the fed-batch run 10 to the training data set for the GPR model (see Section 3.2.5), the configuration of the model remained the same. It is therefore recommended to perform multiple fed-batch experiments with varying initial conditions if the chemometric model is intended to be used for this operation mode. In the studied case, the kinetic model aligns perfectly with the reference data due to accurate initial values and precise feeding. Regarding the transferability to other setups or ADC reactions, a retuning of the EKF may be necessary as soon as one of the two underlying models' performances changes, as the EKF relies on weighting of the two models. However, a thorough investigation of this was out of the scope of this study. The scenarios and the fed-batch transfer showed that the EKF is less sensitive to deviations and could be implemented for model-based process

control, for example, for end-point determination of the conjugation reaction.

5 | CONCLUSION

In the ADC conjugation reaction, a cytotoxic payload is conjugated to a previously functionalized mAb with different levels of available binding sites. Current multivariate measurement approaches only enable the on-line measurement of the amount of conjugated drug as a whole but not the ADC species along the conjugation pathway. In this study, a robust and adaptive model-based soft-sensor for the advanced monitoring of a site-directed ADC conjugation reaction was presented combining a chemometric sensor based on a GPR model with a kinetic reaction model. The devised GPR model exhibited superior accuracy compared with commonly applied PLSR models and provided uncertainty estimates for each measurement. The associated uncertainty is directly influenced by the variation in the spectral data and could provide a means for risk-based monitoring approaches. The deployed soft-sensor EKF fuses the information from the GPR model with the timely forecasts of the kinetic model. The EKF is able to dynamically adjust the predictions of all unmeasurable ADC species throughout the entire reaction using time-variant computation approaches for the process and sensor noise covariances. The designed EKF improved the predictions on average by 8.7% and 17.2% for the training and test set, respectively. The EKF was further shown to be capable to compensate for experimental noise effects, wrong model initialization, and poorly determined kinetic model parameters compared with the unmodified kinetic or sensor model. In conclusion, the combination of a chemometric sensor and a kinetic model using an EKF-based soft-sensor is highly beneficial for accurate process monitoring and enables a high degree of flexibility with regard to process deviations. This ultimately helps facilitate robust process control or early end-point determination. The presented methodology further ensures a unified calibration routine for all underlying components and provides knowledge of sensor and model uncertainties.

AUTHOR CONTRIBUTIONS

Jürgen Hubbuch initiated and supervised the work. Robin Schiemer and Jan Tobias Weggen both evolved the concepts and methods presented in this manuscript, supervised experimental work, analyzed and interpreted the data, drafted the figures, and wrote the final manuscript. Katrin Marianne Schmitt performed the experiments, prepared the data for analysis, and prototyped the soft-sensor framework providing substantial knowledge for the further development of the methods. Robin Schiemer, Jan Tobias Weggen, Katrin Marianne Schmitt, and Jürgen Hubbuch read and approved the final manuscript.

ACKNOWLEDGMENTS

This project received funding from the Ministry of Science, Research and the Arts of the state of Baden-Württemberg within the initiative *Ideenwettbewerb Biotechnologie*. The authors would like to thank Dr. Michaela Wendeler for her great support and the valuable scientific discussion, and AstraZeneca for providing the mAb used in this study. We further would like to thank Christina H. Wegner for careful proofreading of the manuscript. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Robin Schiemer http://orcid.org/0000-0002-0083-7316
Jan Tobias Weggen http://orcid.org/0000-0002-9049-6868
Jürgen Hubbuch http://orcid.org/0000-0003-0839-561X

REFERENCES

Andersson, M. (2009). A comparison of nine PLS1 algorithms. *Journal of Chemometrics: A Journal of the Chemometrics Society*, *23*(10), 518–529. https://doi.org/10.1002/cem.1248

Andris, S., & Hubbuch, J. (2020). Modeling of hydrophobic interaction chromatography for the separation of antibody-drug conjugates and its application towards quality by design. *Journal of Biotechnology*, *317*(April), 48–58. https://doi.org/10.1016/j.jbiotec.2020.04.018

Andris, S., Rüdt, M., Rogalla, J., Wendeler, M., & Hubbuch, J. (2018). Monitoring of antibody-drug conjugation reactions with UV/Vis spectroscopy. *Journal of Biotechnology*, *288*(October), 15–22. https://doi.org/10.1016/j.jbiotec.2018.10.003

Andris, S., Seidel, J., & Hubbuch, J. (2019). Kinetic reaction modeling for antibody-drug conjugate process development. *Journal of Biotechnology*, *306*(September), 71–80. https://doi.org/10.1016/j.jbiotec.2019.09.013

Cabaneros Lopez, P., Udugama, I. A., Thomsen, S. T., Roslander, C., Junicke, H., Mauricio-Iglesias, M., & Gernaey, K. V. (2020). Towards a digital twin: A hybrid data-driven and mechanistic digital shadow to forecast the evolution of lignocellulosic fermentation. *Biofuels, Bioproducts and Biorefining*, *14*(5), 1046–1060. https://doi.org/10.1002/bbb.2108

Cao, M., De Mel, N., Jiao, Y., Howard, J., Parthemore, C., Korman, S., Thompson, C., Wendeler, M., & Liu, D. (2019). Site-specific antibody-drug conjugate heterogeneity characterization and heterogeneity root cause analysis. *mAbs*, *11*(6), 1064–1076. https://doi.org/10.1080/19420862.2019.1624127

Chen, T., Morris, J., & Martin, E. (2007). Gaussian process regression for multivariate spectroscopic calibration. *Chemometrics and Intelligent Laboratory Systems*, *87*(1), 59–71. https://doi.org/10.1016/j.chemolab.2006.09.004

Chen, T., Ou, X., & Martin, E. (2007). Calibration of spectroscopic sensors with Gaussian process and variable selection. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, *40*, 137–142. https://doi.org/10.3182/20070606-3-MX-2915.00141

Coumans, R. G., Ariaans, G. J., Spijker, H. J., Renart Verkerk, P., Beusker, P. H., Kokke, B. P., Schouten, J., Blomenröhr, M., Van Der Lee, M. M., Groothuis, P. G., Ubink, R., Dokter, W. H., & Timmers, C. M. (2020). A platform for the generation of site-specific antibody-drug conjugates that allows for selective reduction of engineered cysteines. *Bioconjugate Chemistry*, *31*(9), 2136–2146. https://doi.org/10.1021/acs.bioconjchem.0c00337

Cui, C., & Fearn, T. (2017). Comparison of partial least squares regression, least squares support vector machines, and Gaussian process regression for a near infrared calibration. *Journal of Near Infrared Spectroscopy*, *25*(1), 5–14. https://doi.org/10.1177/0967033516678515

Elsheikh, M., Hille, R., Tatulea-Codrean, A., & Krämer, S. (2021). A comparative review of multi-rate moving horizon estimation schemes for bioprocess applications. *Computers & Chemical Engineering*, *146*, 107219. https://doi.org/10.1016/j.compchemeng.2020.107219

FDA. (2004). *Guidance for industry PAT—A framework for innovative pharmaceutical development, manufacturing, and quality assurance* (Tech. Rep. No. September). http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070305.pdf

Feidl, F., Garbellini, S., Luna, M. F., Vogg, S., Souquet, J., Broly, H., Morbidelli, M., & Butté, A. (2019). Combining mechanistic modeling and Raman spectroscopy for monitoring antibody chromatographic purification. *Processes*, *7*(10), 683. https://doi.org/10.3390/pr7100683

Gargalo, C. L., de las Heras, S. C., Jones, M. N., Udugama, I., Mansouri, S. S., Krühne, U., & Gernaey, K. V. (2021). Towards the development of digital twins for the bio-manufacturing industry. In C. Herwig, R. Pörtner, & J. Möller (Eds.), *Digital twins: Tools and concepts for smart biomanufacturing* (pp. 1–34). Springer International Publishing. https://doi.org/10.1007/10_2020_142

Golabgir, A., & Herwig, C. (2016). Combining mechanistic modeling and Raman spectroscopy for real-time monitoring of fed-batch penicillin production. *Chemie-Ingenieur-Technik*, *88*(6), 764–776. https://doi.org/10.1002/cite.201500101

Großhans, S., Rüdt, M., Sanden, A., Brestrich, N., Morgenstern, J., Heissler, S., & Hubbuch, J. (2018). In-line Fourier-transform infrared spectroscopy as a versatile process analytical technology for preparative protein chromatography. *Journal of Chromatography A*, *1547*, 37–44. https://doi.org/10.1016/j.chroma.2018.03.005

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., delRío, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Haseltine, E. L., & Rawlings, J. B. (2005). Critical evaluation of extended Kalman filtering and moving-horizon estimation. *Industrial and Engineering Chemistry Research*, *44*(8), 2451–2460. https://doi.org/10.1021/ie034308l

ICH. (2009). *Q8(R2) Pharmaceutical development* (Tech. Rep. No. June). http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002872.pdf

Joubert, N., Beck, A., Dumontet, C., & Denevault-Sabourin, C. (2020). Antibody-drug conjugates: The last decade. *Pharmaceuticals*, *13*(9), 245. https://doi.org/10.3390/ph13090245

Keller, W. R., & Wendeler, M. (2021). Using multimodal chromatography for post-conjugation antibody-drug conjugate purification: A methodology from high throughput screening to in-silico process development. *Journal of Chromatography A*, *1653*, 462378. https://doi.org/10.1016/j.chroma.2021.462378

Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics*, *11*(1), 137–148. https://doi.org/10.2307/1266770

Kjeldahl, K., & Bro, R. (2010). Some common misunderstandings in chemometrics. *Journal of Chemometrics*, *24*(7–8), 558–564. https://doi.org/10.1002/cem.1346

Kolås, S., Foss, B. A., & Schei, T. S. (2009). Constrained nonlinear state estimation based on the UKF approach. *Computers and Chemical Engineering*, *33*(8), 1386–1401. https://doi.org/10.1016/j.compchemeng.2009.01.012

Luttmann, R., Bracewell, D. G., Cornelissen, G., Gernaey, K. V., Glassey, J., Hass, V. C., Kaiser, C., Preusse, C., Striedner, G., & Mandenius, C.-F. (2012). Soft sensors in bioprocessing: A status report and recommendations. *Biotechnology Journal*, *7*(8), 1040–1048. https://doi.org/10.1002/biot.201100506

Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, *118*, 62–69. https://doi.org/10.1016/j.chemolab.2012.07.010

Narayanan, H., Behle, L., Luna, M. F., Sokolov, M., Guillén-Gosálbez, G., Morbidelli, M., & Butté, A. (2020). Hybrid-EKF: Hybrid model coupled with extended Kalman filter for real-time monitoring and control of mammalian cell culture. *Biotechnology and Bioengineering*, *117*(9), 2703–2714. https://doi.org/10.1002/bit.27437

Ohadi, K., Legge, R. L., & Budman, H. M. (2015). Development of a soft-sensor based on multi-wavelength fluorescence spectroscopy and a dynamic metabolic model for monitoring mammalian cell cultures. *Biotechnology and Bioengineering*, *112*(1), 197–208. https://doi.org/10.1002/bit.25339

Ouyang, J. (2013). Drug-to-antibody ratio (DAR) and drug load distribution by hydrophobic interaction chromatography and reversed phase high-performance liquid chromatography. In L. Ducry (Ed.), *Antibody-drug conjugates*. Methods in Molecular Biology, (Vol. 1045, pp. 275–283). Humana Press. https://doi.org/10.1007/978-1-62703-541-5_17

Paananen, T., Piironen, J., Andersen, M. R., & Vehtari, A. (2019). Variable selection for Gaussian processes via sensitivity analysis of the posterior predictive distribution. *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS)*, *PLMR*(89), 1743–1752. http://arxiv.org/abs/1712.08048

Panowski, S., Bhakta, S., Raab, H., Polakis, P., & Junutula, J. R. (2014). Site-specific antibody drug conjugates for cancer therapy. *mAbs*, *6*(1), 34–45. https://doi.org/10.4161/mabs.27022

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Rathore, A. S., Bhambure, R., & Ghare, V. (2010). Process analytical technology (PAT) for biopharmaceutical products. *Analytical and Bioanalytical Chemistry*, *398*(1), 137–154. https://doi.org/10.1007/s00216-010-3781-x

Rolinger, L., Rüdt, M., Diehm, J., Chow-Hubbertz, J., Heitmann, M., Schleper, S., & Hubbuch, J. (2020). Multi-attribute PAT for UF/DF of proteins—Monitoring concentration, particle sizes, and buffer exchange. *Analytical and Bioanalytical Chemistry*, *412*, 2123–2136. https://doi.org/10.1007/s00216-019-02318-8

Schneider, R., & Georgakis, C. (2013). How to NOT make the extended Kalman filter fail. *Industrial and Engineering Chemistry Research*, *52*(9), 3354–3362. https://doi.org/10.1021/ie300415d

Simon, D. (2006). Nonlinear Kalman filtering. In *Optimal state estimation* (pp. 393–431). John Wiley & Sons Ltd. https://doi.org/10.1002/0470045345.ch13

Simutis, R., & Lübbert, A. (2017). Hybrid approach to state estimation for bioprocess control. *Bioengineering*, *4*(1), 21. https://doi.org/10.3390/bioengineering4010021

Sinner, P., Daume, S., Herwig, C., & Kager, J. (2021). Usage of digital twins along a typical process development cycle. In C. Herwig, R. Pörtner, & J. Möller (Eds.), *Digital twins: Tools and concepts for smart biomanufacturing* (pp. 71–96). Springer International Publishing. https://doi.org/10.1007/10_2020_149

Spann, R., Roca, C., Kold, D., Eliasson Lantz, A., Gernaey, K. V., & Sin, G. (2018). A probabilistic model-based soft sensor to monitor lactic acid bacteria fermentations. *Biochemical Engineering Journal*, *135*, 49–60. https://doi.org/10.1016/j.bej.2018.03.016

Stelzer, I. V., Kager, J., & Herwig, C. (2017). Comparison of particle filter and extended Kalman filter algorithms for monitoring of bioprocesses. In *Proceedings of the 27th European Symposium on Computer Aided Process Engineering—ESCAPE 27* (Vol. 40, pp. 1483–1488). Elsevier Masson SAS. https://doi.org/10.1016/B978-0-444-63965-3.50249-X

The Pandas Development Team. (2020). *Pandas-dev/pandas: Pandas*. Zenodo. https://doi.org/10.5281/zenodo.3509134

Thomas, A., Teicher, B. A., & Hassan, R. (2016). Antibody-drug conjugates for cancer therapy. *The Lancet Oncology*, *17*(6), e254–e262. https://doi.org/10.1016/S1470-2045(16)30030-4

Tulsyan, A., Khodabandehlou, H., Wang, T., Schorner, G., Coufal, M., & Undey, C. (2021). Spectroscopic models for real-time monitoring of cell culture processes using spatiotemporal just-in-time Gaussian processes. *AIChE Journal*, *67*(5), e17210. https://doi.org/10.1002/aic.17210

Tuveri, A., Pérez-García, F., Lira-Parada, P. A., Imsland, L., & Bar, N. (2021). Sensor fusion based on extended and unscented Kalman filter for bioprocess monitoring. *Journal of Process Control*, *106*, 195–207. https://doi.org/10.1016/j.jprocont.2021.09.005

Valappil, J., & Georgakis, C. (2000). Systematic estimation of state noise statistics for extended Kalman filters. *AIChE Journal*, *46*(2), 292–308. https://doi.org/10.1002/aic.690460209

Wei, B., Woon, N., Dai, L., Fish, R., Tai, M., Handagama, W., Yin, A., Sun, J., Maier, A., McDaniel, D., Kadaub, E., Yang, J., Saggu, M., Woys, A., Pester, O., Lambert, D., Pell, A., Hao, Z., Magill, G., … Chen, Y. (2022). Multi-attribute Raman spectroscopy (MARS) for monitoring product quality attributes in formulated monoclonal antibody therapeutics. *mAbs*, *14*(1), 2007564. https://doi.org/10.1080/19420862.2021.2007564

Witjes, H., Van Den Brink, M., Melssen, W. J., & Buydens, L. M. (2000). Automatic correction of peak shifts in Raman spectra before PLS regression. *Chemometrics and Intelligent Laboratory Systems*, *52*(1), 105–116. https://doi.org/10.1016/S0169-7439(00)00085-X

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Zhang, C., Springall, J. S., Wang, X., & Barman, I. (2019). Rapid, quantitative determination of aggregation and particle formation for antibody drug conjugate therapeutics with label-free Raman spectroscopy. *Analytica Chimica Acta*, *1081*, 138–145. https://doi.org/10.1016/j.aca.2019.07.007

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Schiemer, R., Weggen, J. T., Schmitt, K. M., & Hubbuch, J. (2023). An adaptive soft-sensor for advanced real-time monitoring of an antibody-drug conjugation reaction. *Biotechnology and Bioengineering*, *120*, 1914–1928. https://doi.org/10.1002/bit.28428