Received: 26 February 2021 | Revised: 28 May 2021 | Accepted: 19 June 2021

DOI: 10.1002/bit.27870

ARTICLE

BIOTECHNOLOGY AND BIOENGINEERING WILEY

Technology transfer of a monitoring system to predict product concentration and purity of biopharmaceuticals in real-time during chromatographic separation

Anna Christler[1,2] | Theresa Scharl[1,3] | Dominik G. Sauer[1,4] | Johannes Köppl[5] | Anne Tscheließnig[6] | Cabir Toy[6] | Michael Melcher[1,3] | Alois Jungbauer[1,2] | Astrid Dürauer[1,2]

[1]Austrian Centre of Industrial Biotechnology (ACIB), Vienna, Austria

[2]Department of Biotechnology, Institute of Bioprocess Science and Engineering, University of Natural Resources and Life Sciences, Vienna, Austria

[3]Department of Landscape, Spatial and Infrastructure Sciences, Institute of Statistics, University of Natural Resources and Life Sciences, Vienna, Austria

[4]Biopharmaceutical Process & Product Development, Novartis, Schaftenau, Austria

[5]Technical Operations Large Molecules, Manufacturing Science & Technologies, Novartis, Kundl, Austria

[6]Boehringer Ingelheim Regional Center Vienna GmbH & Co KG, Vienna, Austria

Correspondence
Astrid Dürauer, Department of Biotechnology/ Institute of Bioprocess Science and Engineering, University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria.
Email: astrid.duerauer@boku.ac.at

Funding information
Österreichische Forschungsförderungsgesellschaft, Grant/Award Number: 848951

Abstract

Technological developments require the transfer to their location of application to make use of them. We describe the transfer of a real-time monitoring system for lab-scale preparative chromatography to two new sites where it will be used and developed further. Equivalent equipment was used. The capture of a biopharmaceutical model protein, human fibroblast growth factor 2 (FGF-2) was used to evaluate the system transfer. Predictive models for five quality attributes based on partial least squares regression were transferred. Six out of seven online sensors (UV/VIS, pH, conductivity, IR, RI, and MALS) showed comparable signals between the sites while one sensor (fluorescence) showed different signal profiles. A direct transfer of the models for real-time monitoring was not possible, mainly due to differences in sensor signals. Adaptation of the models was necessary. Then, among five prediction models, the prediction errors of the test run at the new sites were on average twice as high as at the training site (model-wise 0.9–5.7 times). Additionally, new prediction models for different products were trained at each new site. These allowed monitoring the critical quality attributes of two new biopharmaceutical products during their purification processes with mean relative deviations between 1% and 33%.

KEYWORDS
critical quality attributes, modeling, online monitoring, PLS, prediction, soft sensor

Anna Christler and Theresa Scharl equally contributing authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2021 The Authors. *Biotechnology and Bioengineering* Published by Wiley Periodicals LLC

1 | INTRODUCTION

According to the WHO, transfer of technology is defined as "a logical procedure that controls the transfer of any process together with its documentation and professional expertise between development and manufacture or between manufacture sites" (World Health Organization, 2011). In this study, we describe the transfer of real-time monitoring and pooling system for preparative chromatographic separation which we developed previously (Sauer, et al., 2019a; Walch et al., 2019). Usually, pooling decisions are made based on off-line or at-line analysis (Mendhe et al., 2015; Rathore, Wood, et al., 2008; Rathore, Yu, et al., 2008; Shekhawat & Rathore, 2019). Our monitoring system allows to predict critical quality attributes such as product concentration and content of process- and product-related impurities by statistical models in real-time. The main advantage of this approach is that it saves a lot of process time compared with the determination of quality attributes by offline wet lab analysis. Moreover, real-time monitoring enables prospective process control (Jiang et al., 2017). Predictive models were built on process data derived from a panel of online sensors and offline analyses of the corresponding quality attributes. As online sensors, conventional pH, conductivity, and UV absorbance measurements were connected in series with commercially available multiangle light scattering (MALS), refractive index (RI), and ATR-FTIR (IR) probes. A prototype fluorescence measurement device was also integrated which allowed to collect emission spectra alternatingly at six different excitation wavelengths. Overall, several thousands of predictors from the online sensors were available for model building. Structured additive regression (STAR) and partial least squares (PLS) regression were applied as modeling techniques (Sauer, et al., 2019a; Walch et al., 2019). In the present work, we used partial least squares (PLS) regression as a modeling technique. PLS regression has been used for similar purposes and with similar online sensors (Grote et al., 2014; Roychoudhury et al., 2006; Wasalathanthri, 2020). This chemometric technique reduces the dimensionality of the data set by projecting the original variables to latent structures. The method is particularly suited for such highly correlated variables generated by the online sensors. For model training, the effluent of the chromatography column was fractionated and analyzed for the respective quality attributes. The optimal number of fractions and replicates of chromatographic runs is a function of the precision of the off-line method and the requested quality of the prediction frequently assessed by the root mean squared error (RMSE) (Felföldi et al., 2020).

The transfer of the monitoring system between different sites is described in this study. A chromatographic capture step of recombinant human fibroblast growth factor 2 (FGF-2) based on ion exchange (Sauer, et al., 2019b) was used as an industry-relevant model process for system comparison. The elution phase was monitored. We used 12 training runs with 15 fractions each to reach a total number of 180 observations. This illustrates that many fractions must be analyzed to establish a model. Hence, a direct transfer of the method from one site to the other would save a lot of time and material. We hypothesize that the transfer of models from the training site to the new sites is possible because sensors of identical configuration of the same vendors where implemented into commercially available chromatographic workstations of the same type at all sites. Furthermore, the same separation process protocol and equivalent feed material were used for all experiments. We tested the hypothesis by evaluating the test errors at the new sites, that is, the difference between predictions and measured quality attributes. The product, FGF-2, was pooled based on offline analyses and based on model-predicted values and the respective models are being described and discussed. Case studies were performed at the new sites showing the functionality with newly generated models specific for the site.

2 | MATERIALS AND METHODS

2.1 | Materials

All chemicals were purchased from Merck unless stated otherwise. Basic human FGF-2 was expressed in *Escherichia coli* BL21 in soluble form. Cells harvest, disintegration, and clarification were carried out as described by Sauer et al. (2019b). Aliquots of the homogenates obtained from fermentation broth carried out under same conditions were used as feed material for the experiments at the three sites. For the case studies, biopharmaceutical proteins were produced by proprietary processes.

2.2 | Methods

2.2.1 | Chromatography

FGF-2 was purified by chromatographic purification on an Äkta Pure 25 (Cytiva) as described in Sauer, et al. (2019b). In brief, *E. coli* homogenate was 0.22 μm filtered and 118 ml (10 CV) of the clarified homogenate were loaded on a column packed with weak cation exchanger Carboxymethyl-Sepharose Fast Flow (Cytiva) with 11.8 ml CV (1 cm diameter, 15 cm bed height). The column was equilibrated before and washed after loading with 100 mM Na-phosphate, pH 7.0. FGF-2 was eluted by a linear gradient from 0 to 1 M NaCl in 100 mM Na-phosphate pH 7.0. During the elution phase, the effluent was collected in 1 ml fractions. Fifteen fractions were analyzed around the peak center. The column was sanitized after each run with 1.0 M NaOH for 1 h (5 CV). For the case studies, proprietary purification protocols were used. For model training, 8–9 replicate runs were performed at each of the new sites, 6–7 of them as training runs, and 2 runs as test set.

2.2.2 | Online sensors and database

Sensors were integrated in the column effluent stream in-line in the order of increasing flow cell void volume and/or increasing pressure sensitivity. Details are described in Sauer, et al. (2019a) and Walch et al. (2019). A mid-infrared spectrometer MATRIX-FM (Bruker) was used to measure ATR-FTIR spectra in the wavenumber range from 3500 to 750 cm^{-1} at a resolution of 2 cm^{-1}. Intrinsic protein

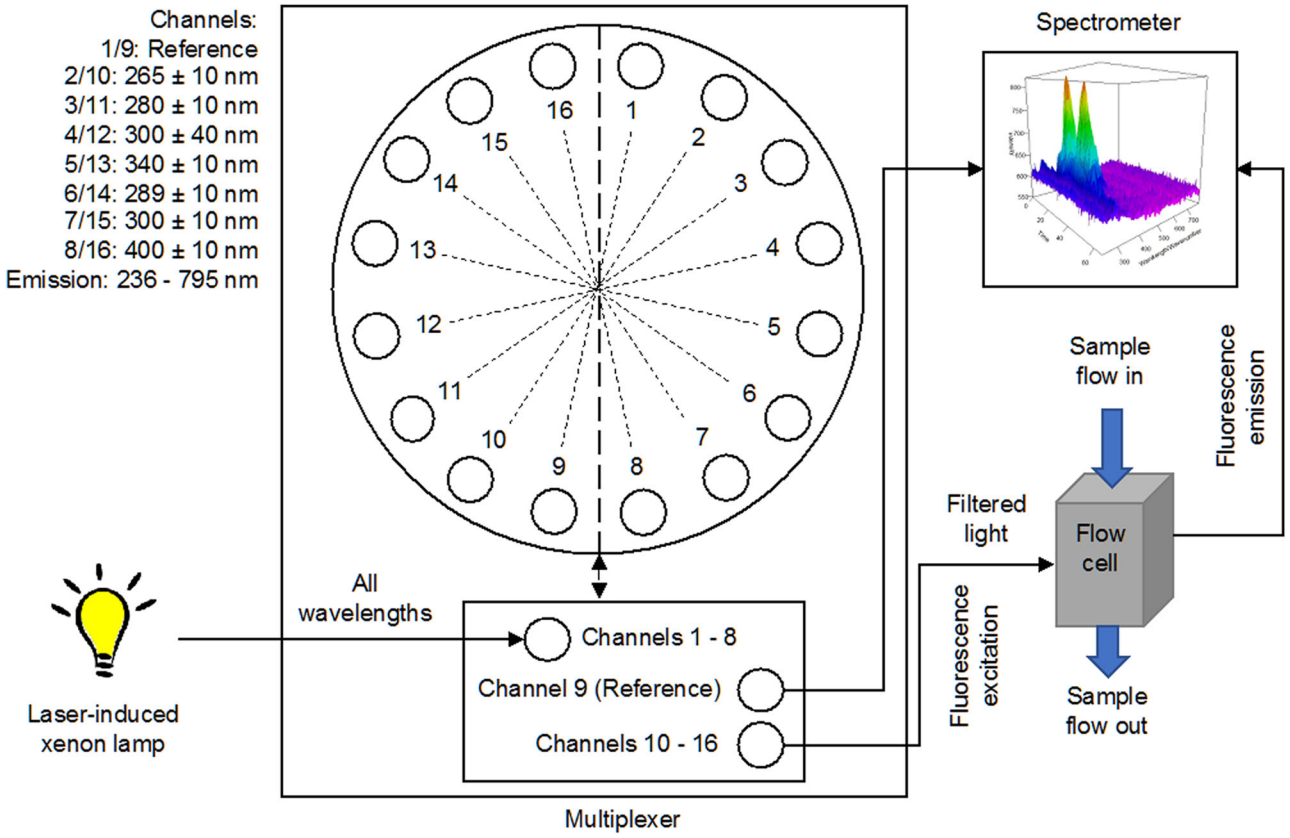

FIGURE 1 Schematic of the in-house assembled multi-wavelength fluorescence detector. Arrows represent optical fibers. Not drawn to scale

fluorescence was recorded at emission wavelengths between 236 and 795 nm at a resolution of 0.3 nm. Excitation was done at six different wavelengths and one (300 nm) with a small and a large filter width: 265 ± 10 nm, 280 ± 10 nm, 289 ± 10 nm, 300 ± 10 nm, 300 ± 40 nm, 340 ± 10 nm, and 400 ± 10 nm. The sensor was assembled in-house using a laser-induced xenon lamp (type EQ-99XFC LDLS, Energetiq), a fiber optic multiplexer (Avantes), a flow cell (FIAlab Instruments), and the spectrometer AvaSpec-ULS-TEC (Avantes) (Figure 1). Measurement time for all seven emission spectra including multiplexer switching time was 16 s. All other detectors were standalone commercial devices. An RI detector Optilab T-rEX (Wyatt) was used allowing differential RI detection in the range of −0.0047 to + 0.0047 RIU. The RI detector also traced a forward monitor for evaluation of system integrity and the LED intensity. Light scattering signals were recorded with the MALS detector miniDAWN TREOS (Wyatt) at angles of 43.6° (LS1), 90° (LS2), and 136.4° (LS3). Additionally, forward monitor intensity and temperature were recorded.

Volume delay between sensors was determined gravimetrically and sensor signals aligned accordingly before modeling and prediction. All buffers used in the process were aqueous based, therefore water was used as blank for all measurements. Blanks were measured with UV, IR, RI, and MALS before each run and signals adjusted. The pH probe was calibrated with a linear calibration between pH 4 and pH 7 before each run. The IR detector was cooled with liquid nitrogen at least 20 min before each run. Signals were recorded and stored by the control software XAMIris (evon).

2.2.3 | Offline analytics for quality parameters

All offline analyses were performed as described in Sauer, et al. (2019b). In essence, protein quantity was determined by reversed-phase (RP) HPLC using a TSKgel Super-Octyl column (2 μm bead diameter, 4.6 × 50/100 mm, 110 Å, Tosoh Bioscience). Monomer and high molecular weight impurity (HMWI) contents were determined by their relative peak areas after size-exclusion (SEC) UPLC using an ACQUITY UPLC BEH125 SEC column (1.7 μm bead diameter, 4.6 × 150 mm, Waters). Low molecular weight impurities were present in negligible amounts in the peak center.

Host cell proteins (HCP) were determined by anti-*E. coli*-HCP sandwich ELISA in 96-well plate format with antibodies from Cygnus. Values of 3–6 doubling dilutions per sample were averaged. Double-stranded DNA (dsDNA) was quantified by Quant-iT® Picogreen (Thermo Fisher Scientific) fluorescence dying in 96-well plate format. Values of 3–4 doubling dilutions per sample were averaged. Offline analyses for the case studies were performed by proprietary analytical methods.

2.3 | Statistical modeling

PLS regression was used to generate prediction models of quantity and purity. All data handling was performed within the statistical computing environment R (Team R. C., 2020) using R package pls (Mevik et al., 2019) as in Walch et al. (2019). This linear regression method is particularly suited for multicollinear variables. Models were trained on 8–12 replicate runs (120–180 data points) at the training site as there were some missing sensor data at the training site. Targets were the quality attributes quantity (measured in mg/ml), monomer content (in %), HMWI content (in %), and HCP and dsDNA contents (both in ng/ml). Based on expert knowledge (such as amide bands and fingerprint regions), the spectral data were reduced to certain regions of interest and different combinations of sensor signals were tested. Numerous subsets of in total 15,725 predictors (online variables) that were available and usable at all three sites were selected. Prediction models were generated on the training data using leave-one-run-out cross-validation (CV) for each of the five responses. The error measure used here was the RMSE, which is a measure of the average prediction error. It is given in the unit of the respective quality attribute and calculated by

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2},$$

where n is the number of samples, y_i are the measured values, and $\hat{y}_i$ the predicted values. The set of predictors yielding the lowest RMSE for a given response was selected for the prediction model of this response. In case of equivalence or almost equivalence between models for one target, models with fewer predictors were preferred for reasons of simplification and greater robustness against sensor fall-outs. Prediction quality was assessed by applying the models on test data sets which were not used for model training. The error measure is then called "test RMSE."

To assess the quality of the proposed models we also included so-called null models. In a null model the response of a certain run is simply predicted by the average values of all the training runs without the use of any predictors. If the proposed model outperforms the null model, the contained predictors are considered important and useful for the prediction of the response.

Another measure of the quality of prediction used in this study was the mean relative deviation (MRD, in %) defined by

$$\mathrm{MRD} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|.$$

Contrary to the RMSE, the MRD is a relative measure and can be regarded as independent of the range of the measured values. Therefore, it was used for the case studies where not all information could be disclosed.

2.4 | Method transfer to new sites

Online sensors at two new sites (A and B) were purchased from the vendors mentioned above. Offline analyses were performed at the training site to reduce differences between sites. Operators received hands-on training for the use of the monitoring system. A run checklist was transferred to ensure correct system handling. Three multi-day hands-on trainings were performed for data analysis and modeling.

Data of runs B1 and B2 were shifted forward by 2 and 1 min respectively, corresponding to equal milliliters, for easier visual comparison.

2.5 | Pooling

Fractions were manually selected for pooling so that the pools met defined quality criteria while maximizing the yield (collected protein over eluted protein). Negative predictions of dsDNA, HCP, and HMWI were set to 0 before fraction selection for calculation of the pool averages. Predictions of more than 100% Monomer were set to 100%. For pooling, HCP and dsDNA were calculated in parts per million (ppm) by dividing the values in ng/ml by the FGF-2 concentration in mg/ml. Runs were pooled independently based on measured quality attributes and based on model predictions. The pooled fractions can be identified in Online Supporting Information.

2.6 | Case studies

All offline analyses were performed at the new sites. PLS models for six critical quality attributes were trained on 6–7 training runs and tested on data of two independent test runs. In each run, 13 fractions were collected.

3 | RESULTS

Online monitoring systems for chromatographic purification were set-up at the two new sites. Automated sensor control was enabled by custom control software. The functionality of each system was first tested using a human serum albumin solution (data not shown). As a model process, we used the chromatographic capture process of FGF-2 from clarified *E. coli* homogenate. First, models were established for real-time prediction of product concentration and contents of monomer, HMWI, HCP, and dsDNA at the training site. These models were transferred to the new sites.

3.1 | Online sensor data of FGF-2 capture at three sites

Online sensor data of three FGF-2 capture runs at site A (runs A1–A3) and three runs at site B (runs B1–B3) were compared with the arithmetic mean of 12 runs performed at the training site (Figure 2). To include the variability at the training site, one, two, and three standard deviations (*SD*) of the signals obtained at the training

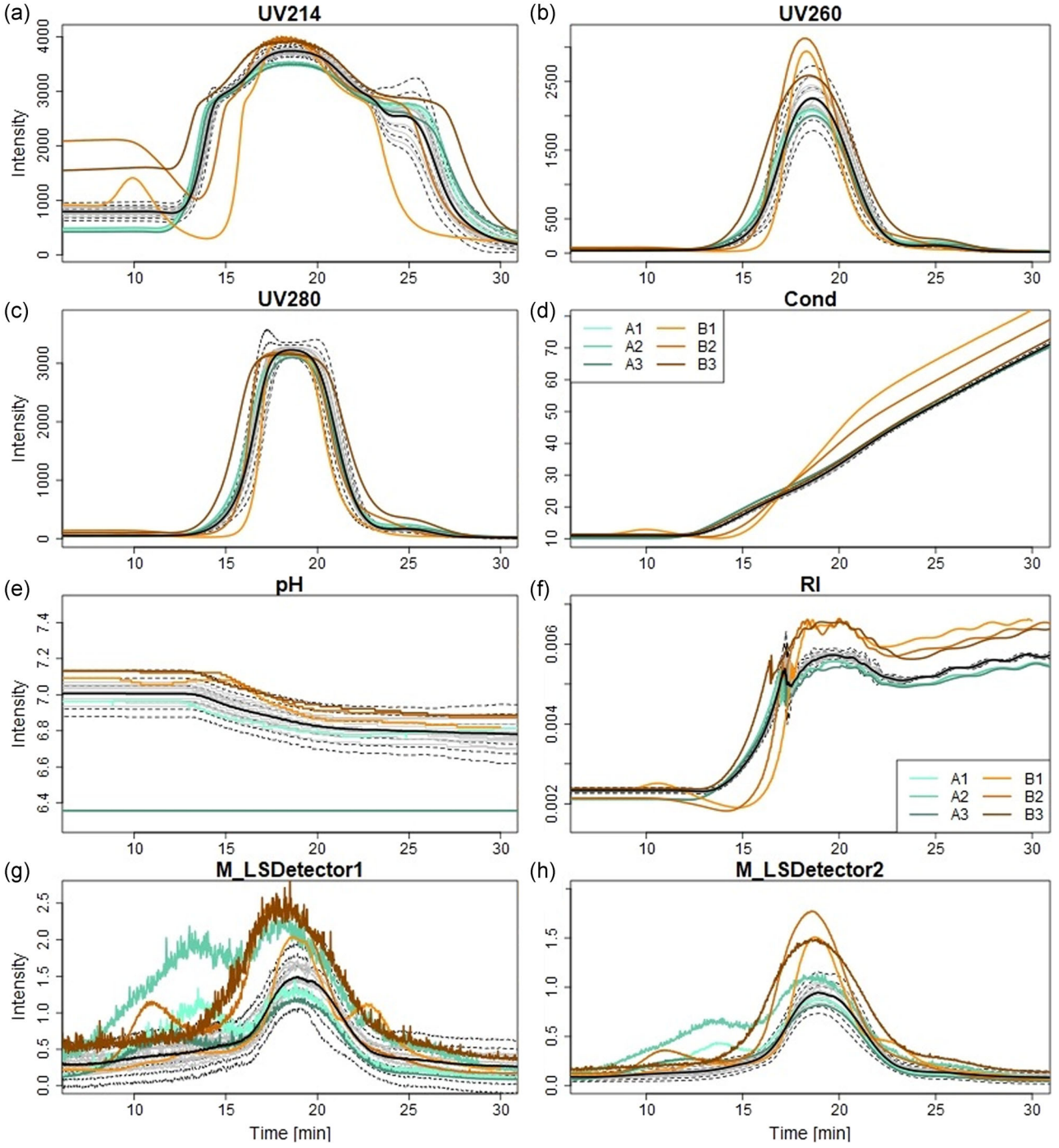

FIGURE 2 Online sensor data during the elution phase from three test runs at sites A and B compared with data from the training site (gray lines). Black lines: averages of 12 training runs. Dashed lines: 1, 2, and 3 standard deviations of training data. Sensors: UV (214, 260, and 280 nm, a–c), conductivity (d), pH (e), refractive index (f), light scattering (g) and (h)

site were included in the graph. The three runs and the respective sensor signals at site A were highly reproducible except for the MALS detector. At site B, the runs were not as reproducible. Elution was delayed in runs B1 and B2 as can be seen in the UV absorption, conductivity, and refractive index signals (Figures 2a–d and 2f). The reason for this delay was a later start and then steeper salt gradient used for protein elution due to insufficient priming of the tubes before the runs. As a result, peaks were narrower with higher peak protein concentrations in runs B1 and B2 (Figure 2b). All sensors except conductivity and pH showed higher maximum intensities compared with site A and the training site. Modifications of the process conditions of the final run B3 caused a broadening of the peak. UV absorption signals showed large differences at site B compared to the other two sites where the signals were comparable

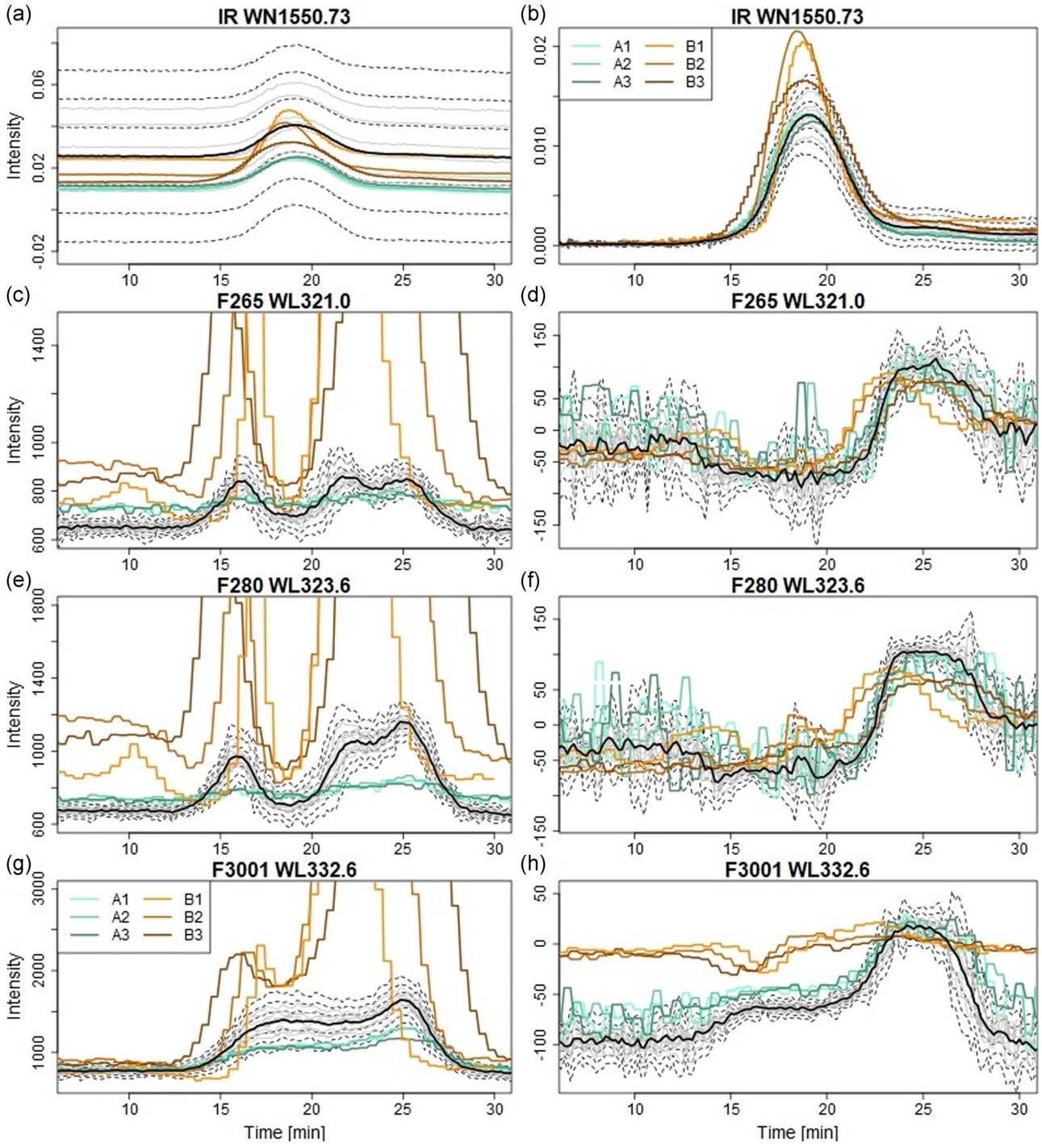

FIGURE 3 Exemplary wavenumbers and wavelengths of IR (a) and (b) and fluorescence data (c)–(h) before and after preprocessing (left and right column, respectively). Gray lines: model training data. Black lines: averages of 13 training runs. Dashed lines: 1, 2, and 3 standard deviations of training data

(Figure 2a–c). pH was not available for runs A2 and A3, but similar in all other runs at all sites (Figure 2d). More pronounced pre-peaks were observed in MALS signals at the new sites in the range of 5–15 min after start of the elution phase (Figure 2g,h). Noisy scattering signals at 43.6° angle (Figure 2g) indicated the requirement for cleaning of the flow cell at both new sites. Processing of cell homogenates led to (fast) fouling of the flow cell. Light scattering at 90° (Figure 2h) was less affected by fouling. Overall, the signals from the two new sites often differed significantly from the training site.

IR and fluorescence sensors recorded absorption and emission spectra, respectively, at each time point in a very high resolution. Thus, representative wavenumbers and wavelengths in the center of the peaks were visualized over the time of the elution phase. Figure 3 shows raw and preprocessed IR and fluorescence data.

Spectral data needed one or more of the following preprocessing techniques: smoothing, baseline correction, first or second derivative, and normalization to equal length or area. Preprocessing methods and their parameter values influenced the suitability for predictive modeling (not shown). They must be optimized iteratively by comparing a score such as the RMSE. Here, smoothing, taking the first derivative and normalization to equal length were selected for fluorescence. IR was preprocessed by subtraction of a reference spectrum recorded before elution in addition to smoothing and baseline correction by asymmetric least squares (Eilers & Boelens, 2005) implemented in the R package *baseline* (Liland et al., 2010).

Among all sensors, the fluorescence data showed the largest differences between sites. Intensities of peaks were different at all sites (Figures 3c, 3e, and 3g). For excitation at 265 nm and 280 nm, signal preprocessing by smoothing, derivatization (first derivative), and normalization to equal length led to a higher agreement of the spectra (Figure 3c–f). For all higher excitation wavelengths (289–340 nm), no suitable preprocessing method or a combination thereof could be found to compensate for the differences of the signals obtained at the different sites (e.g., Figure 3g,h). In addition, a decrease of signal intensities over time as a result of the aging of the lamp and optical fibers was observed (see Figure S1).

3.2 | Model predictions

Predictive models for all described quality attributes were applied to the online data obtained at the new sites. Due to the large differences in fluorescence signals which could not be compensated by preprocessing, the RMSEs of the models were large in the beginning. Fluorescence variables were only included for excitation at 265 and 280 nm. Models for monomer, HMWI, and HCP which contained fluorescence excitation variables above 280 nm, were retrained with data from the training site to be able to apply them to the new sites' data. Figure 4 shows the results of the offline analyses obtained for FGF-2 concentration and contents of monomer, HMWI, HCP, and dsDNA. Furthermore, corresponding model predictions and differences between measured and predicted values (errors) are shown for

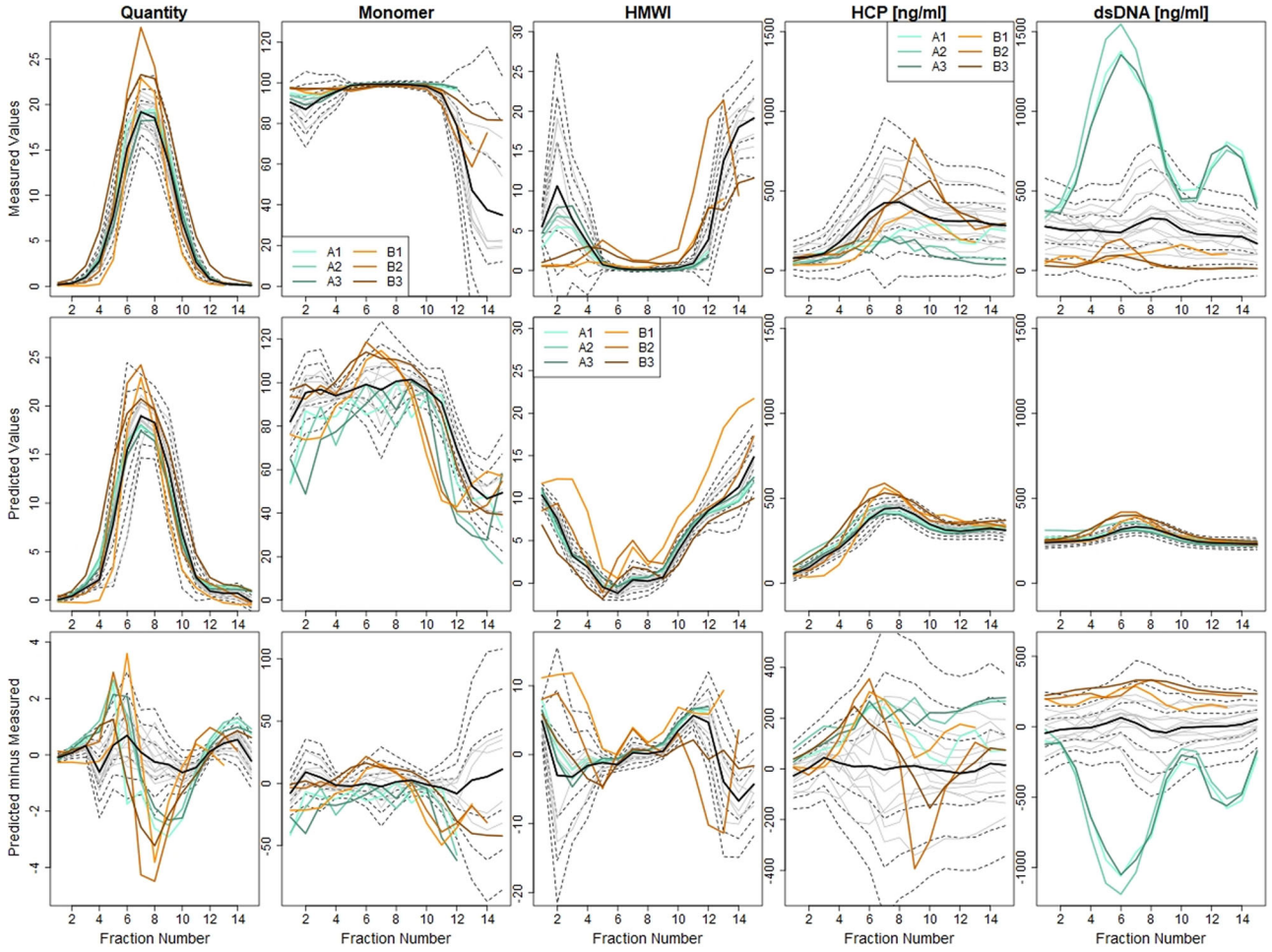

FIGURE 4 Fraction-wise offline measured quality attributes for FGF-2 concentration, monomer, HMWI, HCP, and dsDNA content, model-based predictions, and corresponding differences to offline measured values (prediction errors). Gray lines: model training data. Blackline: average of training data. Dashed lines: 1, 2, and 3 standard deviations of training data. dsDNA, double-stranded DNA; FGF-2, fibroblast growth factor 2; HCP, host cell protein; HMWI, high molecular weight impurities

TABLE 1 Predictors used in the models, RMSEs of the respective null models, and PLS models for test runs at three sites

Model	Predictors (number)	RMSE	Training site	Site A	Site B
Concentration (mg/ml)	UV, RI, conductivity (5)	Null model	7.0	7.0	8.8
		PLS model	0.8 (0.5–1.0)	1.2 (1.1–1.3)	1.1 (0.9–1.4)
Monomer (%)	UV, RI, Fluor265 (25)	Null model	21.9	11.1	24.9
		PLS model	8.5 (4.7–10.4)	17.7 (14.4–20.9)	16.7 (14.9–19.2)
HMWI (%)	UV, conductivity (4)	Null model	6.2	3.7	4.9
		PLS model	3.2 (2.0–4.7)	2.8 (2.7–3.0)	4.2 (1.8–6.1)
HCP (ng/ml)	UV, conductivity, MALS (8)	Null model	156	169	191
		PLS model	81 (37–125)	171 (113–209)	121 (98–141)
dsDNA (ng/ml)	MALS (4)	Null model	110	623	199
		PLS model	87 (39–139)	496 (488–501)	233 (186–271)

Note: The ranges of test RMSEs per run are given in brackets.

Abbreviations: dsDNA, double-stranded DNA; FGF-2, fibroblast growth factor 2; HCP, host cell protein; HMWI, high molecular weight impurities; MALS, multiangle light scattering; PLS, partial least squares; RI, refractive index; RMSEs, root mean squared errors.

TABLE 2 Quality attributes of the collected FGF-2 product pools at the three sites based on model-predicted values and based on offline measurements, respectively

	Training site (n = 6)		Site A (Run A2)		Site B (Run B3)	
	Measured	Predicted	Measured	Predicted	Measured	Predicted
Pool volume (ml)*	8 or 9	8 or 9	4	5	10	10
FGF-2 concentration (mg/ml)	11.8 ± 0.6	10.5 ± 0.4	15.2	13.9	12.5	11.7
Monomer (%)	96.0 ± 1.5	97.6 ± 1.6	99.5	91.1	97.1	93.7
HMWI (%)	0.8 ± 0.4	1.7 ± 0.2	0.1	1.4	2.5	2.0
HCP (ppm)	44 ± 6	56 ± 3	12	34	45	53
dsDNA (ppm)	34 ± 14	56 ± 5	56	26	4	46
Yield (%)	97.1 ± 0.3	96.3 ± 1.3	65.1	73.0	97.6	95.9

Note: Plus/minus values represent one standard deviation. For the new sites, data of one representative run is shown.

Abbreviations: dsDNA, double-stranded DNA; FGF-2, fibroblast growth factor 2; HCP, host cell protein; HMWI, high molecular weight impurities.

*For the exact fractions refer to Online Supporting Information.

each fraction. As before, data from new sites were plotted over the data obtained at the training site including their respective averages and 1, 2, and 3 standard deviations thereof. Predictors selected in each model are given in Table 1.

Errors of the prediction of concentration were distributed around zero in a slightly larger range as for the training data with errors up to 4.5 mg/ml. The monomer content was underpredicted at site A by −8.9% on average and by up to −20.5% in collected fractions. The monomer content was over- and underpredicted at site B by up to 31.2% overall and up to −20.7% in collected fractions. On average, the monomer content at site B was underpredicted by −1.5% in collected fractions. HMWI showed lower absolute values in the beginning of the peak at site B compared with the other sites which was not recognized by the model. Errors of the HMWI predictions were in the same range as at the training site (±12% HMWI). The model for HCP was not able to recognize the different profiles at the new sites, yet the errors were in a similar range as for the training runs. Measured dsDNA was higher at site A compared to the training site and lower at site B, due to the different fermentation batches. The model for dsDNA was not able to predict the different dsDNA profiles accurately. Average prediction errors for dsDNA were about two and five times as high as at the training site for site A and B, respectively (Table 1).

Model quality can also be evaluated with regard to the errors obtained by so-called null models which predict the target value by simply taking the average of the respective variable from all the

training runs. Any trained model has to outperform the null model. For the training site, data from six independent test runs from two new fermentation batches (3 runs per batch) was used as a reference (Sauer, et al., 2019a). These test runs represent a similar situation as was faced at the new sites since the data was not used for model training and the processed material was from different fermentation batches. Predictions and offline measured values can be found in Online Supporting Information.

The PLS model for FGF-2 concentration performed much better than the corresponding null model at all three sites (Table 1). The model for monomer content also performed better than the null model at two of three sites. At site A, the PLS model gave a higher RMSE than the null model. This was probably due to many missing observations at the tail of the peak where the error usually was the largest (compare Figure 4). At the training site, the models predicting HMWI, dsDNA, and HCP allowed better predictions compared with the null models. Average RSMEs of the null models were between 1.3 and 8.8 times larger than the RMSEs of the PLS models at the training site. At the new sites, the advantage of the models over the null models was moderate or none at all. In three cases (monomer at site A, HCP at site A, and dsDNA at site B) the model prediction errors were even higher than the null model RMSE. The performance of the transferred PLS models at the new sites was only satisfying for the FGF-2 concentration.

3.3 | Product pooling based on model predictions

The ability of the new systems to produce product that meets specified quality criteria was evaluated by pooling fractions of each of the test runs. Quality attributes of pools based on model-predicted values were compared to pools based on offline measured data. The following criteria had to be fulfilled by the pools: FGF-2 monomer more than or equal to 90%, HMWI less than or equal to 5%, HCP less than or equal to 60 ppm, DNA less than or equal to 60 ppm, and FGF-2 concentration in the fraction of at least 1.0 mg/ml. Measured and predicted HCP and dsDNA contents in ng/ml were converted into ppm by division through the measured and predicted FGF-2 concentrations, respectively. In an iterative process, fractions were included to maximize the share of collected product from the total eluted protein (=yield) while fulfilling all quality criteria. Table 2 shows the obtained pool quality attributes at the training site and for exemplary runs at the new sites. Run A2 was selected randomly since all runs at site A were very similar. Run B3 was selected as an example because in this run the process conditions best matched those of the other sites. Pool volumes were multiples of 1 ml fractions and thus integer numbers. At the training site, very similar pools were obtained using the predictions and the offline measurements. At site A, less fractions were pooled offline than online due to high measured dsDNA contents. At site B, the model predictions led to the collection of the same fractions as were pooled offline. The FGF-2 concentration in the peak center was underpredicted about equally at all sites: On average, the collected FGF-2 mass was predicted 7.0% lower than the measured mass. This was most probably due to the saturation of many sensors at the peak center, such as UV or RI detectors. Despite deviations in the monomer prediction at the training site of up to 17% in collected fractions, the average monomer content of the pools was similar to the reference analytics. At site A, the monomer content was the limiting variable for pooling based on predictions due to its strong underprediction which led to moderate yields of 73%–88%. The mass balance of %monomer and % HMWI was not closed either due to the presence of low molecular weight impurities (not shown) or due to inaccuracies of the predictions. Mass balances were not considered with these models.

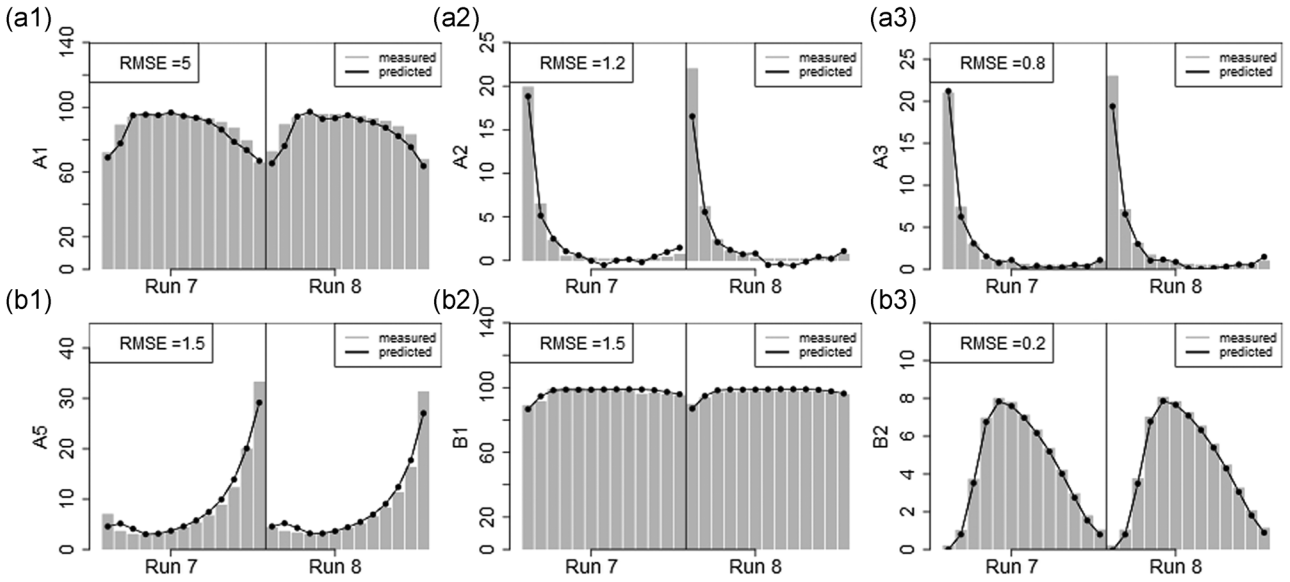

FIGURE 5 PLS predictions and offline measurements of six critical quality attributes of two test runs at one of the new sites. The test RMSE is given for each attribute. PLS, partial least squares; RMSEs, root mean squared errors

The high dsDNA content of the material at site A (compare Figure 4) was responsible for the low yields of the pools based on offline measurements of 65%–82%.

3.4 | Case studies at the new sites

Case studies at the new sites were performed to test the abilities of the systems to predict critical quality attributes of biopharmaceutical proteins during chromatographic separation. One of them is exemplarily shown in Figure 5. Among the modeled product attributes were typical process-related and product-related quality attributes, respectively. Mean relative deviations (MRDs) among the six models were between 1% and 33%.

Among the critical quality attributes were charge variants which represent a form of micro-heterogeneity. The composition of charge variants is a major quality attribute that needs to be controlled in the manufacturing of biopharmaceuticals (Hintersteiner, Lingg, Janzek, et al., 2016; Hintersteiner, Lingg, Zhang, et al., 2016). Protein charge variants have very similar structural and spectroscopic properties and present a challenge for online monitoring and prediction. The developed models were able to predict the different critical quality attributes with satisfying accuracy (not shown).

4 | DISCUSSION

The aim of the technology transfer was to ensure that the systems are fit for their purpose at the new sites. The performances of the online monitoring systems transferred to the new sites were compared with the training site by two different measures: the RMSEs of the test runs quantified the average errors overall fractions and thus the overall quality of the PLS models. The pooling example illustrated the ability of the models to produce qualitative products.

Even though equivalent load material was processed following the identical protocol and using online sensors from the same suppliers at all sites, significant differences between the sensor signals were observed for the system established at the two new sites. Site-specific influences such as operator effects, ambient temperature, humidity, or chemicals might have influenced the processes. Fouling is another common problem when working with biological solutions. Usually, signals are being background corrected by resetting them just before the measurement however at the cost of decreased sensitivity with an increasing amount of fouling. Fouling and sensor aging must be continuously monitored and controlled to ensure that the monitoring system is in a functional state. The IR detector came with a built-in performance qualification test which was done before each run. The test fails, for example, when the device is not cooled enough, when the lamp intensity is too low or humidity too high which would indicate leakage. IR spectra were strongly influenced by ambient conditions, for example, temperature. IR absorption measurements are usually background corrected by a blank spectrum recorded just before the measurement. This was not sufficient here due to the long duration of the loading process of about 2 h. Spectra of different runs showed very different appearances. Subtraction of a spectrum just before the elution was necessary to use the sensor data for modeling. For RI and MALS detectors, Forward monitor intensity of more than 90% was used for quality control. With these measures, the commercial sensors were robust and delivered comparable results across the sites. The reason for the different fluorescence signals was probably that this sensor was not an optimized commercial setup but an in-house assembled prototype (see Figure 1). The flow cell was free-standing and not encased as in the other sensors. The optical fibers must be manually arrested and the exact position impacted the measurement through the transmission of light. Mechanical switching of channels by the multiplexer must be precise to transmit all light. This setup was chosen to allow the scanning of six excitation wavelengths in parallel to gather as much information as possible. However, this was at the expense of the robustness of this fluorescence sensor. In a manufacturing environment, a simpler and more robust sensor would be needed. Fluorescence signals at site A were closer to the training site than signals at site B. The technology was transferred to site A about 2 years after set-up at the training site was completed and to site B about 1 year after that, that is, 3 years after set-up at the training site. Signal preprocessing could not level out the differences sufficiently. Lamp aging is generally common for such lamps and can be predicted but here the effects of fiber aging and lamp aging overlayed.

Accurate representative process data is a major requirement for the generation of reliable prediction models. Therefore, sufficient time and resources must be invested in data generation. However, data generation was work-intensive (compare Christler et al., 2020). The appropriate number of fractions and number of runs depends on the coefficient of variation of the reference analytics (Felfödi et al., 2020). If more fractions are analyzed, less runs can be sufficient because a certain number of total data points is necessary. More observations usually increase the reliability of predictions. More than the planned runs were necessary at the training site because at least one sensor showed fouling, aging, or had a defect. Moreover, the complex system required experience to handle it correctly. Offline and online data of the runs at site A were generally more similar to the training site than at site B. This shows that technology transfer concerns not only the transferred technology but also the technology handling process. System handling is expected to become much easier for an optimized commercial sensor battery.

Model training was conducted within a few days, once the workflow was established and knowledge was obtained on data processing. The main parameters for model optimization were data preprocessing techniques and predictor selection. Simpler models with fewer predictors were selected in this investigation whenever the RMSEs were similar between models containing a different number of predictors. Simpler models bear a lower risk of dysfunction of the monitoring system if single sensors become dysfunctional or lose their connection to the database during a run. Missing predictors can lead to totally missing predictions because all predictors of a model are needed to calculate the target. Missing predictors

could be imputed by other highly correlated predictors using for example tree-based modeling techniques (Kuhn & Johnson, 2013), however at the cost of increased model complexity.

Overall, the UV absorbance sensor gave the most useful predictors, conductivity the second most useful, then RI and MALS, and last fluorescence. IR was not used in a model. However, this is no general statement. In our previous work, fluorescence and infrared sensors yielded very useful predictors, especially for process-derived impurities such as HCP and DNA, but also for monomer and HWMI (Sauer, et al., 2019a; Walch et al., 2019). This could be a reason for the poorer performance of these models here, since fluorescence from excitation above 280 nm could not be used. The fact that a sensor is not included in a model does not necessarily mean that it is useless in predicting a response. In the case of several prediction models with a similar performance the simpler models were selected due to higher robustness. RSMEs of the five models on the test runs of the new sites were on average twice as high (2.0 times) as the test errors at the training site (0.9–5.7 times). Higher errors at the new sites were expected since some of the new data were not represented in the training data set, such as the high dsDNA values at site A or the high HCP values at site B. In such cases, predictions may not be accurate (Kuhn & Johnson, 2013). Variations in product and impurity contents are common in biological processes. The training and test data set should include at least as much variation as is expected later in the application, better a bit more.

Product pooling based on model predictions was possible at all three sites with yields between 73% and 99%. Nevertheless, the differences between offline measurements and model predictions showed that the accuracy of the transferred models was not sufficient for monitoring in the manufacture of biopharmaceuticals. A properly functioning fluorescence sensor is expected to significantly improve the models' performance. Moreover, the test situation was rather complex due to the very different impurity contents between the sites.

The case studies showed that on-site model training allowed to predict six critical quality attributes with good accuracy. A mean relative deviation of 33% may be borderline for manufacturing but there is still room for optimization, for example by using different modeling techniques. Nonlinear techniques such as STAR used in our previous work (Sauer, et al., 2019a) generally bear an advantage in cases where predictors and responses are nonlinearly related. However, nonlinear modeling techniques often need more computation power and time for training. An advantage of PLS is that it can be trained very fast. For proof of principle, we used the simpler method here.

A prerequisite for real-time process monitoring and control is a central powerful data collection point. Large data was generated: for one run about 450 MB of process monitoring data. Sufficient computing capacity was necessary to allow the estimation of several quality parameters within a few seconds. Overall, time for quality estimation could be reduced by days because no offline analytics was necessary anymore. Once the functionality of the monitoring system was shown, it can also be used for process development (e.g., Chemalil, 2020). Sensor data can be used directly, for example, the fluorescence pattern to distinguish protein of interest from HCP. Model predictions can be used to compare process variants as long as the conditions are within the trained design space of the models.

In summary, we were able to transfer statistical models for real-time prediction of five critical quality attributes of a recombinant human FGF-2 process from the training site to two other sites. Up to 5.7 times higher test errors at the new sites compared with the training site were observed with an average of 2.0 times among all quality attributes. The accuracy of the transferred models would not be high enough for real-time process monitoring and product pooling. Model re-training would be needed for application in manufacturing. However, these were not optimized models but used to assess system similarities and general functionality under real process conditions.

5 | CONCLUSION

For a successful transfer of a statistical model for real-time prediction we conclude that the quality of prediction at new sites depend on (a) how close the process parameters can be matched with the training site and (b) how robust and reliably the sensors work at the different sites. The biggest source of errors in our work was the different sensitivity of these highly sensitive sensors. Differences between sensor signals at the three sites could partially be compensated by preprocessing methods and this is considered as advantage of the statistical models. For process monitoring, model re-training at each site was necessary. The case studies showed that on-site model training allowed to predict six critical quality attributes with good accuracy. We conclude that it will be necessary to at least optimize transferred statistical models at new sites. Sensor robustness and thus data reliability are key elements of a monitoring system.

ACKNOWLEDGMENTS

Anna Christler, Theresa Scharl, Dominik G, Sauer, and Michael Melcher have been supported in part by the Federal Ministry for Digital and Economic Affairs (BMWD), the Federal Ministry for Transport, Innovation, and Technology (BMVIT), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol, Government of Lower Austria and ZIT–Technology Agency of the City of Vienna through the COMET funding program managed by the Austrian Research Promotion Agency FFG. The funding agencies had no influence on the conduct of this study. The authors thank Manuel Garbe, Anna-Carina Frank, Edit Felföldi and Magdalena Mosor for technical support.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author, (A. D.), upon reasonable request.

ORCID

Alois Jungbauer http://orcid.org/0000-0001-8182-7728

Astrid Dürauer http://orcid.org/0000-0002-6007-7697

REFERENCES

Chemalil, L. (2020). Online/at-line measurement, analysis and control of product titer & critical product quality attributes (CQAs) during process development. *Biotechnology and Bioengineering*. *117*, 3757-3765. https://doi.org/10.1002/bit.27531

Christler, A., Felföldi, E., Mosor, M., Sauer, D., Walch, N., Dürauer, A., & Jungbauer, A. (2020). Semi-automation of process analytics reduces operator effect. *Bioprocess and Biosystems Engineering*, *43*(5), 753–764. https://doi.org/10.1007/s00449-019-02254-y

Eilers, P. H., & Boelens, H. F. (2005). Baseline correction with asymmetric least squares smoothing. *Leiden University Medical Centre Report*, *1*(1), 5.

Felfödi, E., Scharl, T., Melcher, M., Dürauer, A., Wright, K., & Jungbauer, A. (2020). Osmolality is a predictor for model-based real time monitoring of concentration in protein chromatography. *Journal of Chemical Technology and Biotechnology*, *95*(4), 1146–1152. https://doi.org/10.1002/jctb.6299

Grote, B., Zense, T., & Hitzmann, B. (2014). 2D-fluorescence and multivariate data analysis for monitoring of sourdough fermentation process. *Food Control*, *38*(1), 8–18. https://doi.org/10.1016/j.foodcont.2013.09.039

Hintersteiner, B., Lingg, N., Janzek, E., Mutschlechner, O., Loibner, H., & Jungbauer, A. (2016). Microheterogeneity of therapeutic monoclonal antibodies is governed by changes in the surface charge of the protein. *Biotechnology Journal*, *11*(12), 1617–1627. https://doi.org/10.1002/biot.201600504

Hintersteiner, B., Lingg, N., Zhang, P., Woen, S., Hoi, K. M., Stranner, S., Wiederkum, S., Mutschlechner, O., Schuster, M., Loibner, H., & Jungbauer, A. (2016). Charge heterogeneity: Basic antibody charge variants with increased binding to Fc receptors. *mAbs*, *8*(8), 1548–1560. https://doi.org/10.1080/19420862.2016.1225642

Jiang, M., Severson, K. A., Love, J. C., Madden, H., Swann, P., Zang, L., & Braatz, R. D. (2017). Opportunities and challenges of real-time release testing in biopharmaceutical manufacturing. *Biotechnology and Bioengineering*, *114*(11), 2445–2456. https://doi.org/10.1002/bit.26383

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Liland, K. H., Almoy, T., & Mevik, B.-H. (2010). Optimal choice of baseline correction for multivariate calibration of spectra. *Applied Spectroscopy*, *64*(9), 1007–1016.

Mendhe, R., Thukkaram, M., Patil, N., & Rathore, A. S. (2015). Comparison of PAT based approaches for making real-time pooling decisions for process chromatography: Use of feed forward control. *Journal of Chemical Technology and Biotechnology*, *90*(2), 341–348. https://doi.org/10.1002/jctb.4448

Mevik, B.-H., Wehrens, R., & Liland, K. H. (2019). pls: Partial least squares and principal component regression. *R package version*, *2* 7-2.

Rathore, A. S., Wood, R., Sharma, A., & Dermawan, S. (2008). Case study and application of process analytical technology (PAT) towards bioprocessing: II. Use of ultra-performance liquid chromatography (UPLC) for making real-time pooling decisions for process chromatography. *Biotechnology and Bioengineering*, *101*(6), 1366–1374. https://doi.org/10.1002/bit.21982

Rathore, A. S., Yu, M., Yeboah, S., & Sharma, A. (2008). Case study and application of process analytical technology (PAT) towards bioprocessing: Use of on-line high-performance liquid chromatography (HPLC) for making real-time pooling decisions for process chromatography. *Biotechnology and Bioengineering*, *100*(2), 306–316. https://doi.org/10.1002/bit.21759

Roychoudhury, P., Harvey, L. M., & McNeil, B. (2006). At-line monitoring of ammonium, glucose, methyl oleate and biomass in a complex antibiotic fermentation process using attenuated total reflectance-mid-infrared (ATR-MIR) spectroscopy. *Analytica Chimica Acta*, *561*(1–2), 218–224. https://doi.org/10.1016/j.aca.2006.01.037

Sauer, D. G., Melcher, M., Mosor, M., Walch, N., Berkemeyer, M., Scharl-Hirsch, T., Leisch, F., Jungbauer, A., & Dürauer, A. (2019a). Real-time monitoring and model-based prediction of purity and quantity during a chromatographic capture of fibroblast growth factor 2. *Biotechnology and Bioengineering*, *116*(8), 1999–2009. https://doi.org/10.1002/bit.26984

Sauer, D. G., Mosor, M., Frank, A.-C., Weiß, F., Christler, A., Walch, N., Jungbauer, A., & Dürauer, A. (2019b). A two-step process for capture and purification of human basic fibroblast growth factor from E. coli homogenate: Yield versus endotoxin clearance. *Protein Expression and Purification*, *153*, 153–182. https://doi.org/10.1016/j.pep.2018.08.009

Shekhawat, L. K., & Rathore, A. S. (2019). Mechanistic modeling based process analytical technology implementation for pooling in hydrophobic interaction chromatography. *Biotechnology Progress*, *35*(2), e2758. https://doi.org/10.1002/btpr.2758

Team, R. C. (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.r-project.org

Walch, N., Scharl, T., Felföldi, E., Sauer, D. G., Melcher, M., Leisch, F., Dürauer, A., & Jungbauer, A. (2019). Prediction of the quantity and purity of an antibody capture process in real time. *Biotechnology Journal*, *14*(7), 1–10. https://doi.org/10.1002/biot.201800521

Wasalathanthri, D. (2020). Real-time monitoring of quality attributes by in-line fourier transform infrared spectroscopic sensors at ultrafiltration and diafiltration of bioprocess. *Biotechnology and Bioengineering*, *117*, 3766–3774. https://doi.org/10.1002/bit.27532

World Health Organization. (2011). *WHO guidelines on transfer of technology in pharmaceutical manufacturing* (pp. 285–309). WHO. https://www.who.int/medicines/areas/quality_safety/quality_assurance/TransferTechnologyPharmaceuticalManufacturingTRS961Annex7.pdf

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Christler, A., Scharl, T., Sauer, D. G., Köppl, J., Tscheließnig, A., Toy, C., Melcher, M., Jungbauer, A., & Dürauer, A. (2021). Technology transfer of a monitoring system to predict product concentration and purity of biopharmaceuticals in real-time during chromatographic separation. *Biotechnology and Bioengineering*, 118, 3941–3952. https://doi.org/10.1002/bit.27870