A PAT Study of an Industrial Catalytic Hydrogenation of an Active Pharmaceutical Ingredient

Licinia O. Rodrigues¹, João A. Lopes², Joaquim P. Cardoso¹ and José C. Menezes²*

¹ CIPAN SA, Vala do Carregado, P-2601-906 Alenquer, Portugal
² Centre for Biological & Chemical Engineering, Technical University of Lisbon, Av. Rovisco Pais, P-1049-001 Lisbon, Portugal

Abstract
Process analytical technologies (PAT) designate the integrated use of industrial process analytical chemistry techniques with classical process systems engineering tools, for the analysis and control of manufacturing processes. In this work we report on the development and integrated use of an at-line monitoring technique with a kinetic model of the process, to establish a basic understanding of an industrial catalytic hydrogenation of an active pharmaceutical ingredient (API). A suitable process spectroscopy technique (NIR) is described to monitor the most relevant reaction constituents, in terms of process performance (product distribution, reagent conversion and catalyst selectivity and stability). A kinetic model is also proposed for the process, which is capable of describing the industrial process under diverse operating conditions. The success of the industrial PAT application described is indicative of this methodology general value for industrial process analysis, control and optimization without production disruption.

Keywords: process analytical technologies (PAT), near-infrared (NIR), multivariate modelling (PLS), pharmaceutical production

1. Introduction
Process analytical technologies (PAT) are systems for analysis & control of manufacturing processes based on timely measurements of critical quality parameters and performance attributes of raw materials and in-process products, to assure acceptable end-product quality at the completion of the process (i.e., quality by design) (FDA, 2003). PATs are a landmark in the acceptance of process systems engineering tools in modern pharmaceutical manufacturing and quality assurance of food and drug processes in general. PATs involve the application of process analytical chemistry, chemometrics and process control techniques (viz., intelligent use of process data with multivariate supervision and diagnosis strategies). PATs as such were born in XXI century for this century’s industries and are more than a simple sum of existing disciples as there is a significant gain in their integrated use. Successful PATs applications are the major driver for developments in some of its supporting sciences

* Author to whom correspondence should be addressed: bsel@ist.utl.pt
and not the other way around, as in many other cases in which science advances precedes technological applications. In terms of process monitoring the use of PATs represent a paradigm shift in the sense that sophisticated quality control moves from lab-based to process-based (i.e., in-process). Numerous publications on this subject have appeared in the last few years (2000-2003) as simple web search will reveal using with the acronym. In fact, the discipline is going through an exponential growth phase with dedicated sessions (e.g., ACHEMA, 2003) and even congresses. Here we report on the integrated use of an at-line process spectroscopy monitoring technique (Fourier transform near-infrared spectroscopy, FT-NIR), with process kinetic models, to establish a basic understanding of an industrial catalytic hydrogenation of an active pharmaceutical ingredient (API). Similar efforts on the use of process spectroscopy to study non-ideal API catalytic hydrogenations have been reported earlier (Dyson et al., 2000) but no industrial description of the combined use of monitoring and modelling in the PAT sense has been reported to the best of our knowledge.

2. Experimental

For confidentiality reasons process details must be kept to a minimum. The reaction mixture is a non-ideal multiphase multi-component system made up of a gas phase (H₂), an organic solvent phase containing the reaction reagent (A), product (B), intermediates and decomposition products (C and D), and a catalyst-on-carbon solid phase. Both A and B are pharmaceutical active principles (API) with antibiotic activity. The molecules A through D have a large common skeleton with minor chemical differences. The catalyst properties (activity, selectivity and stability) are strongly correlated with the number of utilizations (i.e., between regenerations) and the number of regenerations (i.e., catalyst lives). Thus, those two numbers will be used as state variables.

Samples were collected from the reactor for analysis at equal intervals throughout reaction time and filtered for catalyst removal. NIR and HPLC procedures were carried simultaneously. This method was repeated for 7 batches operating under different conditions such as temperature, initial concentrations and catalyst’s live and utilization. An ABB BOMEM MB-160 spectrophotometer, InGaAs detector, equipped with an AXIOM fibre optical interface coupled to a adjustable optical path length transflectance probe (Flex C22-Quartz probe from SOVIAS AG) was used. The instrument was controlled via GRAMS AI® 7.00 software. Transflectance probe with a 2mm path length was immersed into the samples for spectra collection. The spectrum for each sample was recorded with triplicate, over the wavenumber range 4000-12000 cm⁻¹. Spectra was pre-processed (Savitzky-Golay, 2nd derivative) and divided into calibration (40 samples from 6 batches) and validation sets (7 samples from one batch). All models tested were based on the PLS (partial least-squares) algorithm.

3. Results and Discussion

3.1 Process monitoring with FT-NIR

Several monitoring options based on different process spectroscopy solutions (viz., UV and mid-infrared with ATR, transmittance and reflectance near-infrared probes) were tested. Considering the reaction matrix and the similarity between compounds A to D, FT-NIR and a transflection probe were selected. (Rodrigues et al, 2003). After the FT-NIR spectra acquisition of samples from several industrial batches a simple exploratory data analysis by principal component analysis (PCA) was carried out for outlier detection in booth calibration and validation data sets. The scores plot in Fig. 1 contains
most of the variance in the pre-processed spectra and shows that there are no abnormal
batches as all samples cluster together inside the 95% confidence limit. Fig. 1 also
supports the choice of batch #2 for validation set since all samples from this run are
within the range spanned by the other batches samples.

Figure 1. Scores plot for the first and second principal component of each sample’s pre-
processed FT-NIR spectra. (●) Calibration set, (○) Validation set (batch #2), (---) 95%
confidence limit.

Table 1. Results for internal and external validation of calibration models

<table>
<thead>
<tr>
<th>Analyte</th>
<th>PLS Model</th>
<th>Prediction Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC’s</td>
<td>R²cv</td>
</tr>
<tr>
<td>Reagent (A)</td>
<td>6</td>
<td>0.998</td>
</tr>
<tr>
<td>Product (B)</td>
<td>6</td>
<td>0.987</td>
</tr>
<tr>
<td>By-product (C)</td>
<td>3</td>
<td>0.952</td>
</tr>
</tbody>
</table>

Figure 2. External validation of FT-NIR monitoring (batch #2).  a) Correlation plot of
predicted values (arbitrary units) with HPLC values;  b) Predicted NIR reaction
profile. (○) Reactant, (●) Product, (⊙) By-Product.

Wavelengths selection for NIR calibration model is one critical step in model
construction. It was found that variable selection by a genetic algorithm gave better
results than conventional wavelength selection techniques (Naes et al., 2002). PLS models were built for each analyte by leaving-one-batch-out for cross-validation. These models were then used to predict the concentrations profiles observed in the external validation batch. The predictions, as well as some statistical figures of merit for the external validation batch, are presented in Table 1 and Fig. 3. The results obtained are very satisfactory since a global correlation coefficient of 0.998 between NIR predictions and the HPLC was achieved, as well as a global mean relative error (MRE = 3.2%) similar to HPLC reference method error (3%).

### 3.2. Process modelling

As suggested by the observed concentration profiles the hydrogenation reaction can be described by a set of ordinary differential equations (equations (1) to (4)). The $K_i$'s are formal rate constants in the proposed reaction scheme and lump together temperature effects and catalyst properties. It was necessary to include a second by-product, D, which is not relevant in terms of reaction monitoring but whose explicit consideration is needed to close mass balances. The simple kinetic model proposed is adequate for the envisaged future use of the model in process optimisation and control studies.

$$\frac{d[A]}{dt} = -(K_1 + K_2 + K_3)[A]^\alpha$$  \hspace{1cm} (1)  
$$\frac{d[B]}{dt} = K_1[A]^\alpha$$  \hspace{1cm} (2)  
$$\frac{d[C]}{dt} = K_2[A]^\alpha$$  \hspace{1cm} (3)  
$$\frac{d[D]}{dt} = K_3[A]^\alpha$$  \hspace{1cm} (4)

A set of parameters $K_1$, $K_2$, $K_3$ and reaction order $\alpha$, were estimated for each hydrogenation batch, by fitting the proposed model to the concentration profiles, minimizing the sum of square errors (SSE) with a non-linear least squares technique. Table 2 shows the estimated values for each batch and summarizes the respective operating conditions used in each batch: temperature (T), catalyst life (L = 1, 2, 3), number of catalyst utilizations in each life (U = 1, 2, …, 6) and a scaling factor (Z) related to the ratio of catalyst to reagent in each hydrogenation. Not only the order of magnitude of each estimated parameter in Table 2 was the same for all seven hydrogenation runs, but their value was in general very similar in all batches. Even though that is already a strong indication of model structure adequacy, the model was modified to explicitly account for the operating conditions mentioned in Table 2 and enabling a unique set of parameters to describe all batches. Thus, parameter $K_i$ was considered to be a product of two contributions: $K_i = k_i \Phi_i$ with the first term representing a formal kinetic constant, $k_i$, for which an Arrhenius law can be used to account for temperature effects in $K_i$; while the second term enclose all other properties that indirectly define the catalyst state as $\Phi = f (L, U, Z)$. To set up an adequate function for $\Phi$, several meaningful equations in terms of process engineering knowledge were examined (Table 3). It is known that increasing the reaction temperature will get better conversion but will also increase catalyst deactivation and have a negative impact in subsequent catalyst utilizations. This information is contained in any of the examined formal models for $\Phi$ in Table 3. Model discrimination and parameter estimation of $a$ and $b$ constants of models in Table 3, as well as the Arrhenius law parameters for each reaction, were determined by standard non-linear least-squares. The best results were obtained with function (IV) for both the main product B and the secondary by-product D, while function (III) was best suited to describe by-product C. For each compound a
single set of parameters was obtained to describe its kinetics in all but one batch, as explained below.

Table 2. Predicted constants

<table>
<thead>
<tr>
<th>Batch</th>
<th>Operating Conditions</th>
<th>Estimated parameters</th>
<th>Fitting Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>L</td>
<td>U</td>
</tr>
<tr>
<td>1</td>
<td>40.5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>43.1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>49.7</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>48.9</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>41.9</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>42.9</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>44.9</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3. Tested functions for $\Phi$

<table>
<thead>
<tr>
<th></th>
<th>(I)</th>
<th>(II)</th>
<th>(III)</th>
<th>(IV)</th>
<th>(V)</th>
<th>(VI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi = \frac{Z}{aU^b}$</td>
<td>$\Phi = \frac{1}{aU^b}$</td>
<td>$\Phi = \frac{1}{aU^b}$</td>
<td>$\Phi = \frac{Z}{aU^b}$</td>
<td>$\Phi = \frac{1}{aL^b}$</td>
<td>$\Phi = \frac{Z}{aL^b}$</td>
<td></td>
</tr>
</tbody>
</table>

Validation of the revised kinetic models $K_i = k_i \Phi_i$, was carried out by cross-validation: i.e., taking one batch out, performing parameter estimation on the remaining batches and predicting $K_i$ in the left out batch with the already calibrated $K_i = k_i \Phi_i$ model. This procedure was repeated for product B and by-products C and D. In fact, the predicted $K_i$ values ($<K_i>$) for each batch are not significantly different from the values obtained from the experimental data for that particular batch ($K_i$). That similarity is a good indicator of the proposed kinetic expressions predictive ability. In fact, the mean relative errors between $K_i$ and $<K_i>$ are well below 10%.

3.3 PAT approach validation

To validate the combined monitoring / modelling approach proposed (PAT) a test was made with a batch not used neither in NIR calibration nor in model development. Figure 4 summarizes the comparisons between NIR and process model predictions. There is a very good agreement between the observed trends by both types of models and a generally good agreement between absolute values of both models. The relative accuracy of the two different types of predictions opens several possibilities as to the usefulness of combined at-line monitoring / on-line model-based process supervisory control schemes in fault detection and diagnosis. Based on the developed kinetic model several optimisation studies can be made with regard to different criteria and with varying detail as to process economics and constraints considered. To demonstrate that the predictions made with the proposed model are realistic enough we will accept the current sub-optimal practice in the real plant of a constant temperature profile throughout a batch, as optimal and derive the best temperature history for consecutive utilizations of the catalyst that maximizes conversion (this will yield batches of uneven duration). The obtained results are depict in Fig. 4 for six consecutive utilizations of a catalyst in its second life ($L=2$ and $U=1, \ldots, 6$). There is a striking agreement not only of the trends observed but also on the clustering of industrial batches around the determined optimal temperature set-points (data omitted from Fig. 4). In fact, in the
studied process the index to be optimised is essentially process conversion as the operating costs have a significantly much lower impact on the overall profitability and the formation of secondary products is preferred to the presence of unreacted reagent at the completion of batch.

Figure 3. Validation batch. NIR analysis (dots) vs. kinetic model predicted profiles (lines).

Figure 4. Simulated conversion profiles for different reactor temperatures in six consecutive utilizations of the catalyst in its second life (*) - optimal reaction temperature).

4. Conclusions

A NIR method was developed as an alternative to conventional HPLC method for monitoring a hydrogenation reaction with simultaneous determination of the three main components (reagent, main product and most relevant by-product). A kinetic model was established that is able to describe well the process within the operating range of the real industrial process. Presently we are deriving optimal operating points for each particular condition of the reaction catalyst, while future work is aimed at developing the outlined monitoring/modelling PAT application into a routine process supervision strategy to increase process consistency through better and faster diagnosis of process operating deviations from the optimal trajectories derived.

References

ACHEMA, 2003 - Pharmaceutical Production Technology (PAT I & II) Sessions, May, Frankfurt, Germany