European Journal of Pharmaceutics and Biopharmaceutics 119 (2017) 343–352

Contents lists available at ScienceDirect

European Journal of Pharmaceutics and Biopharmaceutics

journal homepage: www.elsevier.com/locate/ejpb

Research paper

Global regression model for moisture content determination using near-infrared spectroscopy

Matthieu Clavaud [a,b,*], Yves Roggo [a], Klara Dégardin [a], Pierre-Yves Sacré [b], Philippe Hubert [b], Eric Ziemons [b]

[a] F. Hoffmann-La Roche Ltd., Wurmisweg, CH-4303 Kaiseraugst, Switzerland
[b] University of Liege (ULg), CIRM, Department of Pharmacy, Laboratory of Analytical Chemistry, Quartier Hôpital, Avenue Hippocrate 15, B36, B-4000 Liege, Belgium

ARTICLE INFO

Article history:
Received 27 March 2017
Revised 2 June 2017
Accepted in revised form 15 July 2017
Available online 17 July 2017

Keywords:
Freeze-dried medicine
Moisture content
Near-infrared spectroscopy
Support vector regression
Global model

ABSTRACT

Near-infrared (NIR) global quantitative models were evaluated for the moisture content (MC) determination of three different freeze-dried drug products. The quantitative models were based on 3822 spectra measured on two identical spectrometers to include variability. The MC, measured with the reference Karl Fischer (KF) method, were ranged from 0.05% to 4.96%. Linear and non-linear regression models using Partial Least Square (PLS), Decision Tree (DT), Bayesian Ridge Regression (Bayes-RR), K-Nearest Neighbors (KNN), and Support Vector Regression (SVR) algorithms were created and evaluated. Among them, the SVR model was retained for a global application. The Standard Error of Calibration (SEC) and the Standard Error of Prediction (SEP) were respectively 0.12% and 0.15%. This model was then evaluated in terms of total error and risk-based assessment, linearity, and accuracy. It was observed that MC can be fastly and simultaneously determined in freeze-dried pharmaceutical products thanks to a global NIR model created with different medicines. This innovative approach allows to speed up the validation time and the in-lab release analyses.

1. Introduction

The freeze-drying process is widely used in the manufacturing of medicines. This operation consists in reducing the water content by applying sublimation and desorption [1]. Lab analyses are required to quantify the remaining water in the freeze-dried products. While Karl Fischer (KF) titration is the most employed method [2], studies demonstrated the efficiency of Near-infrared (NIR) spectroscopy to determine the remaining part of water in the freeze-dried products, as well as in powders, tablets or capsules [3–6]. The choice of NIR for water determination is first justified because it is a safe, fast, and non-destructive method. Secondly, water has an important signal in the NIR spectral range at around 6900 cm^{-1} and 5150 cm^{-1}, which is suitable for quantitative applications [7]. The NIR studies published so far are mainly product specific and creating a NIR model for one product is time-consuming. Thus, a solution would be to create a dataset composed of spectra of different products for one and unique quantitative application.

The majority of the published applications for MC determination are focused on single models [5,8] and few of them on universal models [9,10]. Nevertheless, some multivariate global models have been investigated for fast quantitative determination of compounds.

The first type of global models published deals with the quantification of compounds in the pharmaceutical or agricultural fields manufactured at different sites. In the agricultural field, universal NIR models were investigated for the quantification of compounds in pesticides, in milk (e.g. fat and protein content), or in feedstock (e.g. banana, coffee) [11–13]. In the pharmaceutical field, models were presented for active pharmaceutical ingredient (API) quantification (e.g. in capsules, tablets), and for vitamins quantification [14–18]. In all the studies, the possibility of replacing the case-by-case strategy conventionally used by a NIR universal model was demonstrated. A strategy has also proposed to update a universal quantitative model when products with different excipients and manufacturing processes are involved [19], as well as strategy to select a validation set [20].

The second type of global models that were published is the instrument-to-instrument transfer. Applying a bias adjustments [21] or adding noise in order to deteriorate data [22] are two examples of computations to show that model developed with multiple instruments can be transferred to other ones.

* Corresponding author at: F. Hoffmann-La Roche Ltd., Wurmisweg, CH-4303 Kaiseraugst, Switzerland.
E-mail addresses: matthieu.clavaud@roche.com (M. Clavaud), y_roggo@hotmail.fr (Y. Roggo), klara.degardin@roche.com (K. Dégardin), pysacre@ulg.ac.be (P.-Y. Sacré), ph.hubert@ulg.ac.be (P. Hubert), eziemons@ulg.ac.be (E. Ziemons).

http://dx.doi.org/10.1016/j.ejpb.2017.07.007
0939-6411/

In this paper, three different freeze-dried medicines were chosen to investigate a global regression model for MC determination. The main novelty of this work lies in the investigation of linear and non-linear chemometrics regression methods for a global application, namely the Partial least squares (PLS), the decision tree (DT), the Bayesian ridge regression (Bayes-RR), the K-Nearest Neighbors (KNN), and the Support vector regression (SVR). A dataset composed of 3822 spectra acquired on two identical NIR spectrometers was used to create and validate a model common to the three drug products. Each regression method was evaluated, and the best one based on SVR algorithm was retained.

2. Material and methods

2.1. Samples and validation concept

2.1.1. Selection of samples

2.1.1.1. Calibration and validation sets. Three different freeze-dried medicines, filled into transparent glass vials, have been measured. The products were annotated A, B, and C for confidentiality reasons. For the product B, two dosages B_1 and B_2 were analyzed that were composed of the same ingredients but presented a different active pharmaceutical ingredient (API) concentration. All the products differed in terms of type and content of API, diameter size of the glass vial, and excipients amount (Table 1).

At least 5 batches and 100 samples per product were selected to include batch-to-batch, product-to-product, and sample-to-sample variability in the dataset (Table 1).

In addition, a special attention was paid to the quality of the cake. Only sample passing the visual appearance test were used. The vials were visually inspected to retain the cakes presenting the best appearance in order to avoid meltbacks [1] (collapse resulting from the presence of liquid during the primary drying), tilted cake (non-correct position of the cake in the glass vial), and cake breaks or cracks (due to transport malfunction). All of them could have an impact on the NIR measurement. The NIR spectra will be impacted and consequently the prediction results will be incorrect.

2.1.1.2. External validation set. One freeze-dried medicine, named D, also filled into transparent glass vials, has been tested during the external validation. The product was different to the ones used to create the models (Table 1).

2.1.2. Sample preparation

In normal circumstances, it is not required to prepare the samples prior to NIR measurements. Nevertheless, since each of the measured product contains on average 1.00% of residual moisture on the routine, a preparation step was necessary in order to create a wider MC range.

On the whole, 140 vials were kept intact (unchanged) and stored at 5 °C, storage temperature of the studied products.

Then, 466 vials were humidified using the strategy protocol described in a previous work [23]. The vials were humidified from 5 to 105 min for the product A, from 5 to 240 min for the product B, from 5 to 180 min for product C, and from 15 to 150 min for product D, with 5 min interval. After this period of time, the samples were closed and stored at 5 °C. The humidification times were selected after feasibility studies which showed those were the necessary times to reach a wide MC range.

In parallel, 59 vials were dried using also the strategy protocol described in a previous work [23]. The samples of product A were kept inside desiccators from 2 to 22 days, with 2 days interval. In parallel, the samples of product B were kept inside desiccators during: 2, 6, 10, 14, 18 and 22 days. Then, when the indicated time was reached, the samples were closed and stored in a second desiccator containing silica gel. The drying times were selected after feasibility studies which showed those were the necessary times to reach a wide MC range. The samples of product C and D were not dried due to their to initial low MC.

2.1.3. Calibration and validation concept

For the global model evaluation, the whole dataset composed of products A, B, and C, was split into two independent calibration and validation sets. Each set contained samples from all the products. The calibration set was composed of two thirds of the dataset and the validation set of the remaining third. Therefore, 419 vials were used for the calibration and 218 vials for the validation (Table 1).

For the individual model evaluation, each product dataset (products A, B, and C) was also split into independent calibration and validation sets, with two thirds of the samples in the calibration set and one third in the validation set (Table 1).

The global and individual models were created using the calibration set and validated with the validation set. An external validation set composed of 28 vials from the product D was also used to evaluate the prediction effects of the best global model on an independent product.

2.2. Analytical methods

2.2.1. Near-infrared equipment and spectra recording

Two identical Antaris Fourier Transform spectrometers (Thermo Fisher Scientific®, Madison, WI, USA) were used to perform the measurements. The instruments were equipped with an InGaAs detector, a Michelson interferometer, an halogen NIR source, and an integrating sphere module.

Table 1
Main characteristics of the products used to calibrate and validate the multivariate models.

	Model calibration and validation				External validation
Products	A	B_1	B_2	C	D
Galenic form	Freeze-dried				
Type of product	Biopharmaceutical (protein based)			Chemically manufactured	
API type	Monoclonal antibody	Antibody drug conjugate		Antiviral	Antiviral
API content (mg/vial)	150	100	160	100	500
Vial format (mL)	15	15	20	3	15
Vial diameter size (cm)	2.5	2.5	3	1.5	2.5
Total amount of product in a vial (mg)	50	433	700	50	60
Number of batches	5	10	7	15	5
Number of vials	100	194	238	105	28
(calibration/validation)	*(67/33)*	*(124/70)*	*(158/80)*	*(70/35)*	*(0/28)*
Humid	58	126	166	95	21
Dry	30	14	15	n/a	n/a
Unchanged	12	54	57	10	7