Routine infrared phosphorus determination in ex-farm milk providing better insight in the phosphorus cycle on dairy farms

M. Bastian and H. van den Bijgaart

Qlip B.V., P.O. Box 119, 7200 AC Zutphen, The Netherlands
Corresponding Authors: bastian.mikael@gmail.com; bijgaart@qlip.nl

Abstract

EU limitations incited the Dutch government in 2017 to implement a national regulation to reduce phosphorus losses from dairy husbandry. The demand for a better insight in the phosphorus cycle on Dutch dairy farms led Qlip to develop a FTIR calibration model for phosphorus measurements in raw milk.

The calibration model was developed with a training set of 210 milk samples and tested on 80 milk samples. The model allows for a precise estimation of P content (RMSEP = 2-3 mg/100 g milk). Model performance is stable across the year, similar between herd bulk milk and individual cows' milk, and robust to specific breeds, e.g. Jersey cows.

The application was successfully implemented in routine Dutch herd bulk milk testing in early 2019. With this tool, farmers can now monitor the phosphorus balance in their dairy cattle and better fine-tune the supply of phosphorus through the ration. Furthermore, these detailed milk data can serve as a basis for farm-specific reporting of phosphorus output through ex-farm milk supplies.

If wished for, the application can be extended with models on a number of other minerals.

Keywords: Phosphorus excretion, milk, infra red spectrometry.

Introduction

Phosphorus (P) is an important mineral in milk and for the dairy sector. P deficiency in cows may cause health disorders such as demineralization of the skeleton, growth problems, lameness, infertility and a decrease in milk yield (Brooks, Cook, Mansell, and Walker, 1984; Gerloff and Swensen, 1996). To avoid these negative effects, farmers have long preferred *over*feeding cows with P (Klop *et al.*, 2013). However, this costly strategy also has environmental side effects: while intake of P (P_{intake}) above the physiological needs does not further increase milk yield nor P concentration in milk (P_{milk}, Wang *et al.*, 2014), it does increase the P concentration in manure (P_{manure}) in the form of phosphates (PO_4^{3-}, P_4O_{10}). Phosphates in manure contribute to the eutrophication of soils and waters.

To face environmental challenges, the EU published her Nitrates Directive in 1991. It aimed to protect water quality across Europe by preventing nitrates from agricultural sources leaching into ground and surface waters and by promoting the use of good farming practices. In 2006, Dutch farmers with their highly productive grassland were granted a derogation (as were Germany, Denmark, UK, Ireland, Flanders and areas in Italy). Farms with at least 80% grassland were allowed to spread up to 230 to 250 kg N (instead of 170 kg N) with manure from grazing animals per hectare per year. As a consequence, national quotas on production of nitrates, but also on phosphates, were imposed.

In 2016, after the lifting of the EU milk quota system, it became apparent that the Dutch farmers did increase the number of cattle and milk production more rapidly than anticipated, and that they would exceed the phosphate quota. This led the Dutch government to regulate phosphate production via a phosphate reduction plan in 2017. This included subsidies for farmers to quit farming, reduction of the phosphate content in concentrates and a generic reduction in number of cattle. Phosphate production rights that had been awarded based on the number of cattle in 2015 underwent an overall deduction of about 8%. As a consequence, a better insight and an understanding of the phosphorus cycle on dairy farms became of economic importance in the Dutch dairy sector from 2017 on.

The phosphorus balance in dairy cows can be expressed with the following formula:

$$P_{manure} = P_{intake} - P_{milk}$$

In the calculations for regulatory purposes P_{milk} was taken as a constant of 97 mg / 100 g (of milk). Yet, other constants had been proposed before: 90 mg / 100 g (NRC, 2000; Valk, Sebek, and Beynen, 2002), and 100 mg / 100 g (Commissie Onderzoek Minerale Voeding, 2005). This degree of discrepancy for P_{milk} undermines the calculus for P_{manure}. Scientific research over the last decade (Alvarez-Fuentes *et al.*, 2016; Klop *et al.*, 2014; Soyeurt *et al.*, 2009) has shown that P_{milk} varies between cows from 56 to 149 mg / 100 g, with an average of 103 mg / 100 g and a standard deviation of 11 mg / 100 g (Alvarez-Fuentes *et al.*, 2016). This meant that P content in milk varies as much as protein content and more than lactose contents (Qlip internal data). Since the variation of fat, protein and lactose contents is routinely monitored at both individual cows and dairy farm levels, a method that can routinely measure P_{milk} was deemed to be useful to provide farmers a better insight in the phosphorus cycle on their farms and better means tofine-tune the supply through the ration.

A clear example of inter-cow and inter-herd differences are herds that are partly or totally constituted of Jersey cows. Compared to Holstein-Friesian cows, which constitute more than 90% of the Dutch dairy cattle, Jersey cows produce milk that is richer in both fat (5.0% vs. 3.6%) and protein contents (3.6% vs. 3.0%, Reinart and Nesbitt, 1956 ; Qlip data for Jersey cows in The Netherlands over the year 2018: fat: 6.0% vs. 4.4%, protein: 4.2% vs. 3.6%). Since P content is correlated with both fat and protein content, it would thus be expected that P content is higher in Jersey milk than in Holstein-Friesian milk.

The current reference method to measure P_{milk} is ICP-MS. ICP-MS requires complex laboratory procedures to measure the mineral content of a milk sample. In comparison, Fourier Transformed Infra-Red (FTIR) spectroscopy is part of the routine analysis of raw milk. Amongst other, FTIR is applied in routine to determine fat, protein, lactose and urea contents at both individual cow and herd levels. A model to predict P_{milk} based on the FTIR spectrum would allow a cost-effective estimation of P content of raw milk samples. Where FTIR spectra have been stored, the estimation can also be made for past raw milk samples.

Qlip wished to develop a FTIR-based model to predict P content in raw milk based on the work of Soyeurt *et al.*, 2009. The present paper summarises the development and validation of the final calibration model. Subsequently, the implementation in routine and large scale prediction of P content are discussed.

Development and validation of the calibration model

In-house development of a FTIR calibration model

In total, the P content of 290 raw milk samples was measured with the reference method ICP-MS (ISO 21424|IDF 243). Milk samples were collected covering the whole range of compositional variation, on multiple instruments and across various periods in the year 2018. The model was a PLS regression based on the FTIR spectra (925.92 to 5011.54 cm^{-1}) of 210 training samples (P reference: mean = 105, SD = 15, range = 64 to 179 mg / 100 g): 105 herd bulk milk samples measured with four Foss MilkoScan™ FT+ instruments and 105 individual cow milk samples measured with ten Foss MilkoScan™ FT6000 instruments (N = 100) and one Foss MilkoScan™ 7 RM instrument (N = 5), thereby having applied beforehand spectrum standardization with all instruments in accordance with manufacturer's instructions.

One way of selecting samples was by chance (e.g.: taking samples out of routine processing). These were 10 herd bulk tank milk samples and 10 individual cow milk samples in February, then in April, June, Augustus and October 2018 (total random samples: 50 herd bulk tank milk and 50 individual cow milk). In each set of 10 samples, 6 were randomly set apart for external validation, constituting a first external validation set (N = 60 samples = 6 x 2 milk types x 5 periods). The remaining 4 samples of each set of 10 samples were used in the calibration set. A second external validation set was composed of 8 Jersey herd bulk milk samples and 8 individual Jersey cow milk samples. All were collected in April 2018, regardless of milk composition (i.e. close to a selection "by chance").

Another way of selecting samples was based on milk composition or milk origin. 174 non-random samples were selected during 2018: N = 70 in January, N = 37 in February, N = 11 "extreme" samples and N = 16 Jersey milk samples (4 of which outliers) in April, N = 40 in Augustus). "Extreme" samples were selected on expected extreme P content based on the FTIR spectrum. Jersey milk samples were selected based on knowing that some dairy farms only housed Jersey cows. The rest of the non-random samples were selected on variations of fat, protein, lactose, urea and/or expected P content. These comprised a total of 85 herd bulk milk samples and 89 individual cow milk samples.

The final "whole year" calibration model was trained on 210 samples:

- 170 non-randomly selected samples (174 samples above, minus the 4 outliers).
- 40 randomly selected samples (4 remaining of each set of 10 random samples).

For simplification, the results on external validation with the two external sets (random routine N = 60 and Jersey milk N = 16 + 4 outliers eliminated from the training) were first pooled in a total set of 80 samples. Note that the "training outlier" samples were known to have had mild to severe fat distribution problems (due to the high fat content of Jersey milk) and / or high acidity. For these reasons these were not included in the training set, but they were still included in the validation set to provide an estimate of the robustness of the predictions. However, of the 80 samples two Jersey milk samples did present really abnormal values (pH = 4.8 and 6.0, urea = 62 and 46 mg/100 g, protein = 7.4% and 6.8%) and were thus excluded from the validation set.

Reference values for P content ranged from 64 to 192 mg / 100 g: 192 mg / 100 g was for one of the two samples excluded. P content thus ranged from 64 to 179 mg / 100 g in the training set (N = 210, mean = 105, SD = 15) and from 78 to 135 mg / 100 g in the validation set (N = 78).

Validation of the calibration model at Qlip

The overall performance of the calibration model on the 78 samples (60 random, 18 Jersey) in external validation was: MAE (Mean Absolute Error) = 2.1 mg / 100 g, RMSE (Root Mean Squared Error) = 2.6 mg / 100 g and R^2 = .94. In comparison, using the same training set and the same validation set but using protein content as predictor (with or without lactose, fat, or urea as co-predictors) led to a performance of MAE = 4.5, RMSE = 5.9, R^2 = .70 at best. The added value of using the spectrum was therefore that the prediction error could be reduced by a factor two as compared to predicting P content from protein content with or without co-predictors.

Figure 1 shows the external validation plot for herd bulk milk (left panel) and for individual cow milk (right panel) separately, and allows the identification of Jersey samples (triangles). There was no significant difference in performance as measured by RMSE between individual cow milk versus herd bulk milk. Similarly, although the set of Jersey samples was small, no significant / obvious degradation of performance could be found when comparing the random set (N = 60, RMSE = 2.2 and 2.9 mg / 100 g respectively for herd bulk milk individual cow milk) to the Jersey set (N = 18, RMSE = 2.7 and 2.8 mg / 100 g).In general, absolute errors were randomly distributed when the two outliers were excluded: no obvious linear or quadratic pattern involving predicted P content, reference P content, fat, protein, lactose or urea contents or pH could be found. This is important since this indicates that the model appears to be robust within the range covered in this validation set (reference P content: 78 – 134 mg / 100 g, predicted P content: 81 – 130 mg / 100 g, fat: 2.4 – 7.0%, protein: 2.6 – 4.7%, lactose: 4.1 – 5.2%, urea: 9 – 37 mg / 100 g, pH: 6.4 – 6.9). Yet, if anything, absolute errors might have been slightly larger with extreme protein values, hence providing an explanation as to why the absolute error with individual cow milk might be slightly larger than the absolute error with herd bulk milk – the range is larger. This is a possible explanation that is to

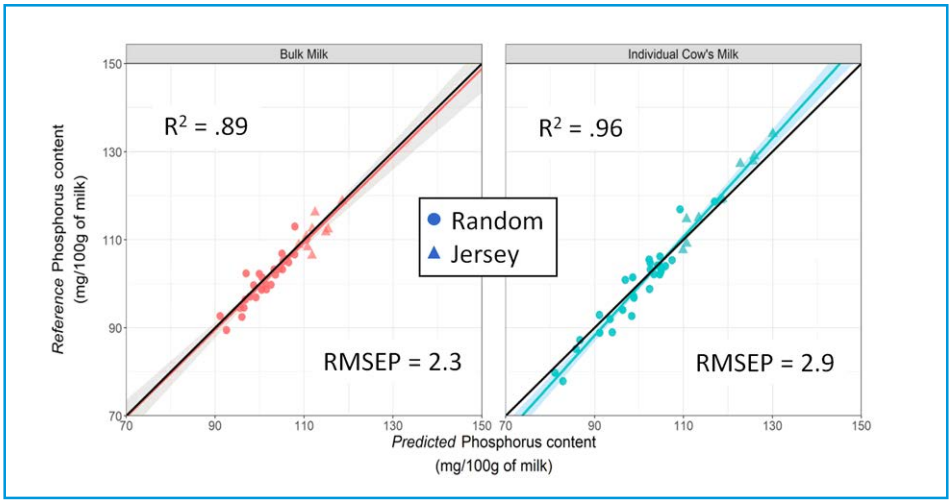

Figure 1. External validation plots for herd bulk milk and individual cow milk. RMSEP in mg / 100 g.

be confirmed with more data. Finally, no significant sinusoidal pattern could be found in the distribution of absolute error during the course of a year, suggesting no obvious effect of season on model performance, and hence a robustness to various seasons.

Validation of the calibration model by another laboratory

In the context of a collaboration with another laboratory that also measures P content in milk, 95 herd bulk milk samples with vast variation in P content (predicted P content from 80 to 122, mean 102, SD = 11) were selected in February 2019 and sent for reference analysis at this laboratory. This occurred about four months after the last results had been included in the calibration set. After correction for a known and understood bias (removing the same bias value to all samples), the agreement between Qlip's FTIR model and the extern laboratory reference method ICP-MS was MAE (Mean Absolute Error) = 2.3 mg / 100 g, RMSE (Root Mean Squared Error) = 3.0 mg / 100 g and R^2 = .96.

Comparison to previous research

In sum, the FTIR model for P content developed at Qlip has a precision between 2.5 and 3 mg / 100 g. The authors are not aware of previous research reporting P content models for bulk milk, but pioneering models for predicting P content in individual cow milk have been published (e.g. Soyeurt *et al.*, 2009) with a RMSE in cross-validation around 5.0 mg/100 g. The differences between previous research and the current work are numerous, and include:

- differences regarding the reference method (ICP-MS following an ISO norm here vs. ICP-AES on frozen-defrosted milk samples by Soyeurt *et al.* 2009);
- the use of multiple infra-red instruments and a very standardized execution of infra-red measurements in a routine laboratory (vs. 1 unique MilkoScan™ FT 6000);
- differences in sample selection and nature of the samples: our final model used herd bulk milk samples as well as individual cow milk samples. This comprised about 210 samples in the training set with a sample selection protocol focused on variability of the chemical composition of milk and of its P content (vs. focused on spectral variability and Ca content by Soyeurt *et al.* 2009).

Some statistics for herd bulk milk 2017-2018

Using the data stored at Qlip, we could derive P content predictions at a large scale, predictions that covered about 4 million herd bulk milk samples.

Overall distribution of milk samples

Herd bulk milk samples from routine were distributed around 102 mg / 100 g, ranging from 90 to 115 mg / 100 g. There was a clear effect of breed, since samples coming from dairy farms raising Jersey cows only had a distribution with higher levels of P content ranging from 100 to 130 mg / 100 g with a median of 115 mg / 100 g.

Seasonal effect

P content presented a seasonal variation that mirrored closely that of fat and protein content: in both 2017 and 2018 maximum values attained 104 mg / 100 g on average in November and December and minimum values attained were 98 mg / 100 g in June, July and Augustus. No significant / obvious difference was found between the years 2017 and 2018.

Relation to protein

P in milk is for the larger part present as phosphate (Walstra and Jenness, 1984). 10% of these phosphate groups are soluble esters and part of phospholipids that are elementary constituents of the fat globule membrane. About 20% of the phosphate groups are also organic but esterified to the protein molecules, notably the various forms of caseins. For those 20% there is therefore a direct relationship between protein and P content: more protein means more P content. The remaining part of P in milk is inorganic and bound to other minerals such as calcium – as calcium phosphate. However, these inorganic phosphate groups are for a considerable part contained in the casein micelles. The relation is here indirect, in that more protein comes with more phosphate. In sum, the majority of P content is directly or indirectly related to protein content. Yet, predicting P content from ex-farm milk at a large scale, we found that this correlation tends to vary during the course of the year, the correlation between P content and protein content being stronger in the winter (at the level of dairy farms in one given day), than in the summer (R^2 = .62 vs. R^2 = .41). The reasons for this difference are still to be explored. We speculate that dairy farm management has a stronger impact in the winter, when all cows are inside, than in the summer, when most cows are grazing. Another reason may be breed since the variability between cows is larger in winter than in summer.

The ability to predict P content at a large scale opens up possibilities to further research this question and other observations at minimal costs.

Take home message

From a trigger, available knowledge and decisive acting, Qlip was able to rapidly implement a new application in her routine FTIR testing portfolio at the beginning of 2019. The Dutch situation regarding phosphate regulation for dairy farms created the need for a better understanding of their phosphorus cycle. Reducing uncertainty about the amount of P in milk was considered helpful in 1) promoting awareness and bringing insight in the P cycle on dairy farms, 2) providing means to improve P utilization on dairy farms and 3) exploring the underpinning farm-specific registration of phosphate production.

Careful sample selection and execution of both the infra-red and the reference method allowed to develop a precise FTIR calibration model with robust performance (RMSE = 2.5 – 3 mg P / 100 g milk). Combining that with milk spectral data stored at Qlip provided insight about the past situation of phosphorus content in milk, and may well help identifying trends regarding the phosphorus cycle of dairy farms. Research institutes that would like to better understand differences in P content between farms, the effect of feed (grass, P supplementation) on P-content of milk, or even the nutritional/ technological / functional properties of P rich milk for human consumption have now, thanks to a routine implementation of the calibration model, a tool to quickly identify farms with interesting milk composition regarding P content, and this independently from protein or fat content.

List of references

Alvarez-Fuentes, G., J.A.D.R.N. Appuhamy and E. Kebreab, 2016. Prediction of phosphorus output in manure and milk by lactating dairy cows. J. Dairy Sci., 99(1), 771–782. *https://doi.org/10.3168/jds.2015-10092*.

Brooks, H., T.G. Cook, G.P. Mansell and G.A. Walker, 1984. Phosphorus deficiency in a dairy herd. New Zealand Vet. Journal, 32(10), 174–176. https://doi.org/10.1080/00480169.1984.35113.

Commissie Onderzoek Minerale Voeding, 2005. Handleiding Mineralenvoorziening Rundvee, Schapen, Geiten.

Gerloff, B.J. and E.P. Swensen, 1996. Acute recumbency and marginal phosphorus deficiency in dairy cattle. J. Am. Vet. Medical Ass., 208(5), 716–719.

International Organization for Standardization, 2018. ISO 21424/IDF 243:2018 - Milk, milk products, infant formula and adult nutritionals — Determination of minerals and trace elements — Inductively coupled plasma mass spectrometry (ICP-MS) method.

Klop, G., J.L. Ellis, A. Bannink, E. Kebreab, J. France and J. Dijkstra, 2013. Meta-analysis of factors that affect the utilization efficiency of phosphorus in lactating dairy cows. J. Dairy Sci., 96(6), 3936–3949. *https://doi.org/10.3168/jds.2012-6336*.

Klop, G., J.L. Ellis, M.C. Blok, G.G. Brandsma, A. Bannink and J. Dijkstra, 2014. Variation in phosphorus content of milk from dairy cattle as affected by differences in milk composition. J. Agr. Sci., 152(05), 860–869. *https://doi.org/10.1017/S0021859614000082*.

NRC, 2000. Nutrient Requirements of Dairy Cattle: Seventh Revised Edition, 2001. *https://doi.org/10.17226/9825*.

Soyeurt, H., D. Bruwier, J.-M. Romnee, N. Gengler, C. Bertozzi, D. Veselko and P. Dardenne, 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. J. Dairy Sci., 92(6), 2444–2454. *https://doi.org/10.3168/jds.2008-1734*.

Valk, H., L.B.J. Sebek, and A.C. Beynen, 2002. Influence of phosphorus intake on excretion and blood plasma and saliva concentrations of phosphorus in dairy cows. J. Dairy Sci., 85(10), 2642–2649. *https://doi.org/10.3168/jds.S0022-0302(02)74349-X*.

Walstra, P. and R. Jenness, 1984. Salts. In: Dairy chemistry and physics, Wiley and Sons, New York, USA, 42-57.

Wang, C., Z. Liu, D. Wang, J. Liu, H. Liu and Z. Wu, 2014. Effect of dietary phosphorus content on milk production and phosphorus excretion in dairy cows. J. Animal Sci. Biotechn., 5(1), 23. *https://doi.org/10.1186/2049-1891-5-23*.