

Feasibility Study of Veterinary Drug Residues in Honey by NIR Detection

Hongqian Chen, Zhenhua Tu, Zhaoshen Qing, Xiaobin Qiu, Chaoying Meng

► **To cite this version:**

Hongqian Chen, Zhenhua Tu, Zhaoshen Qing, Xiaobin Qiu, Chaoying Meng. Feasibility Study of Veterinary Drug Residues in Honey by NIR Detection. Daoliang Li; Yingyi Chen. 6th Computer and Computing Technologies in Agriculture (CCTA), Oct 2012, Zhangjiajie, China. Springer, IFIP Advances in Information and Communication Technology, AICT-392 (Part I), pp.150-156, 2013, Computer and Computing Technologies in Agriculture VI. <10.1007/978-3-642-36124-1_19>. <hal-01348093>

HAL Id: hal-01348093

<https://hal.inria.fr/hal-01348093>

Submitted on 22 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Feasibility Study of Veterinary Drug Residues in Honey by NIR Detection

Hongqian Chen¹, Zhenhua Tu², Zhaoshen Qing³, Xiaobin Qiu¹,
Chaoying Meng^{3*},

1. China Agriculture University Net Center, Beijing, 100083, China,

2. China Food Industry Promotion Center, Beijing, 100062, China

3. China Agriculture University, Beijing, 100062, China

Abstract. In order to verify the feasibility of the near-infrared to detect the veterinary drug in honey, studying the impact of different background correction methods for linear modeling and nonlinear modeling. Use PLSR to establish the linear model between near-infrared spectroscopy and to be measured, and LSSVR to establish the nonlinear model between near-infrared spectroscopy and to be measured. At the same time, we used Wavelet transform, Multiplicative Scatter Correction, Standard Normalized Variate, First derivative transform, Second derivative transform, Orthogonal Signal Correction and other background correction methods to improve the feasibility of the near-infrared detection.

Keywords: Near-infrared; Honey; Tetracycline.

1 Introduction

Honey is the nectar which is collected by bees from many plants, or is the sweet substance which is fully brewed in a special substance. In the process of feeding bees, beekeeper fed the bees with a large number of antibiotics such as tetracycline, chloromycetin and so on, in order to prevent the illness of the bees. This drug can be stored in honey's cells, tissues, organs, and with the food chain will threat human health. Especially in 2008, the exposure of the melamine incident highlights the problems in the field of food safety in China.

According to Hygienic Standard for Honey(GB/T5009.95): tetracycline antibiotics which are the major residues of veterinary medicine have been explicitly included in the detection range, and its limit of physicochemical index residue must be less than 0.05mg/kg. At present, as the requirements of European Union and The United States of America on pesticide and veterinary medicine residue, the quality of honeys which China exports should be improved. In January 2002, The

* This work is supported by National Natural Science Foundation of China(30901127) and Rapid detection of food in small and medium-sized enterprises of public service platform construction project(2011GH552096), National torch plan that environment of industrialization of construction project,e-mail: mcy@cau.edu.cn

European Union banned imports of Chinese products which originated in animals, because chloramphenicol residues of shrimp exceeded the standard. It also caused a chain reaction of Japan, Canada and Hong Kong. The "Green barriers" policy which EU adopted in the honey trade has not only seriously affected the honey export of China, but also influenced the whole industry^[2]. In China, the phenomenon that residues of tetracycline exceed the standard becomes more and more serious. In July 2004, Industrial and commercial bureau of Sichuan Province investigated the honey on the market: only 4 batches of honey qualified in the 30 batches, and the unqualified rate is 87%. The most prominent problem is the tetracycline antibiotic residues exceeding, and 10 batches of honey didn't qualify while the unqualified rate is 33%. In order to safeguard the interests of consumers, to ensure the quality of honey for import and export and find an effective detection method of tetracycline antibiotic residues exceeding is imminent. Efficient and practical detection technology contributes to the safety appraisal of honey, and it can ensure that China's export of honey meets all the international standards.

At present, there are several methods to detect the veterinary drug residue of honey, such as traditional chemical method, liquid chromatography, ultraviolet spectrophotometer, electrochemical analysis and so on. However, these methods are defective, as some of them are devastating while others are time-consuming and expensive. So, it needs a quick, lightweight, accurate and effective detection method in the research of tetracycline compositions in honey.

Some of previous studies have shown that Near-Infrared Spectroscopy can precisely detect the composition of honey. However, the content of tetracycline ingredients in honey is quite low, which leads that the information of near infrared region was disturbed by the background noise, so it is difficult to extract and utilize it. Meanwhile, Near infrared spectrum detection technology has been studied and applied in the detection of honey quality, identification of adulteration, plant source and place of origin. The near infrared spectrum detection technology has been used in the quality detection of honey products, and the adulteration detection, and the detection of plant source, and origin detection, but the detection of veterinary drug residues in honey has not been reported. In this study, the feasibility of detecting veterinary drug residues in honey by Near-Infrared Spectroscopy was investigated, in which variety of spectral preprocessing and non-linear quantitative model will be used.

2 Materials and methods

2.1 Materials

This study collected the main origin of honey, include Sichuan, Qingzang, Beijing and so on. The honey's variety include uniferous honey and multifloral honey, like *Robinia pseudoacacia* L, *Friobotryajaponica* (Thunb.) Lind, *Zizyphus jujuba* NM ver. *iaerius*(Bunge.)Rehrl, *Astragalus siviuis* L, *Vites negundo* van heterophylla (Franch.)Rehd, *Tilia amurensis* Rupr, *Dimocarpuslongan* Lour, *Brassica campestris* L,

Litchi chine-is Soon, Eurya, Citrus reticulata Blanco, Sophoraviciifelia Hanue. There are 153 honey samples, and stored at 4°C freezer.

2.2 Instruments and parameters

In this study, spectral acquisition used ISF/28N Fourier near-infrared spectrometer instrument which is product in BRUKER co. This instrument's range is 3600~12500 cm^{-1} , a minimum resolution is 1cm^{-1} . When collecting the spectrum of liquid, we can use transmission pool (optical path is 2mm or 20mm), quartz liquid transmission and reflection optical fiber (fixed optical path is 2mm) Annex.

2.3 The acquisition of near-infrared spectroscopy

The experimental spectra are collected in a temperature controlled laboratory (In this study, the temperature is 26°C). The instrument must be warm up for 30 minutes or more before test. At the same time, if honey is crystallized, the crystal honey samples must be heated to nine hours in 40 °C water bath and then place samples in 26°C environment temperature.

2.4 Detection of tetracycline content in honey

We use high performance liquid chromatography - ultraviolet detect method to detect the content of tetracycline in honey. Honey samples dissolved in 0.1mol/L Na₂EDTA-McIlvaine solution, after centrifuged, the supernatant use Oasis HLB Solid phase extraction column and the carboxylic acid anion exchange column to purify, and use Elution to determine the volume. Finally, we use high performance liquid chromatography with UV detection at 350nm, and use external standard method (peak area) - the standard curve for quantification, and the concentration of tetracycline is 0.005mg/kg^[6]. In 153 samples, we select 101 samples to detect, there are 41 samples have chemical value, and 60 samples don't have the chemical value.

3 Results and Analysis

3.1 PLSR model of tetracycline content in honey

There are 6 abnormal samples in the 41 samples, so we get 35 samples to calculate. After removing the abnormal samples, in order to get more stability of the model, the calibration set and predict set are in the ratio of 2:1, 3:1, 7:3, 4:1, 3:2, according to K-S law to divide. And then, all set establish PLSR model respectively, all indicators shows in Table 1.

Table 1. Statistic data of honey including calibration and prediction for 5 times.

| Honey quality indicators | Divide number | Sample set | Number of samples | Range | Average | Standard deviation |
|---------------------------|---------------|-----------------|-------------------|-----------|---------|--------------------|
| Tetracycline(μ g/kg) | 1 | Calibration set | 27 | 10.2-66.8 | 22.1333 | 14.1161 |
| | | Prediction set | 8 | 11.3-47.6 | 24.9125 | 13.2533 |
| | 2 | Calibration set | 24 | 10.2-66.8 | 21.7958 | 14.9162 |
| | | Prediction set | 11 | 11.3-47.6 | 24.8909 | 11.2625 |
| | 3 | Calibration set | 21 | 10.2-66.8 | 21.8571 | 15.4482 |
| | | Prediction set | 14 | 11.3-47.6 | 24.1357 | 11.2371 |
| | 4 | Calibration set | 25 | 10.2-66.8 | 22.1080 | 14.6854 |
| | | Prediction set | 10 | 11.3-47.6 | 24.42 | 11.7571 |
| | 5 | Calibration set | 28 | 10.2-66.8 | 21.8143 | 13.9547 |
| | | Prediction set | 7 | 11.3-47.6 | 26.5857 | 13.3714 |

The calibration set and the prediction set be calculate with Auto-scaling method, and then we use PLSR method for multivariate statistical analysis of the experimental data. Nonlinear iterative partial least squares (NIPALS) algorithm is used to get partial least square factors. The optimum factor number of the calibration model (# LV) is determined by two factors, LOOCV and prediction residual sum of squares (PRESS). The result of honey tetracycline Fourier transmission and reflection spectra PLSR model showed in Table 2.

Table 2. The PLSR models results of honey tetracycline content for 5 times.

| Honey quality indicators | Divide number | Calibration set | | | | | Prediction set | | |
|---------------------------|---------------|-----------------|------------------|---------------|----------------|----------------------|----------------|----------------|----------------------|
| | | n ^a | #LV ^b | r | SEC | RSD _c (%) | n ^c | SEP | RSD _p (%) |
| Tetracycline(μ g/kg) | 1 | 27 | 4 | 0.6287 | 10.9774 | 49.6% | 8 | 14.6035 | 58.62% |
| | 2 | 24 | 4 | 0.6170 | 11.7386 | 53.86% | 11 | 12.9852 | 52.12% |
| | 3 | 21 | 4 | 0.5946 | 12.4208 | 56.83% | 14 | 11.7928 | 48.86% |
| | 4 | 25 | 4 | 0.6165 | 11.5629 | 52.3% | 10 | 13.6492 | 55.89% |
| | 5 | 28 | 4 | 0.6271 | 10.8693 | 49.83% | 7 | 15.2219 | 57.26% |

Note: a: n is number of samples in the calibration set, b: # LV is the best number of factors in calibration set, c: n is the number samples in prediction set

The predict effect of PLSR modeling to establish honey tetracycline Fourier transmission and reflection spectra is not good, because the tetracycline content in honey is too low. The second division is much better, r is 0.6170, RSD is 52.12%. Figure 1 shows the correlation of the predictive value and the true value of the model built in the 2nd division of the tetracycline content in honey. From the figure, poor model prediction accuracy is difficult to directly meet the forecast requirements.

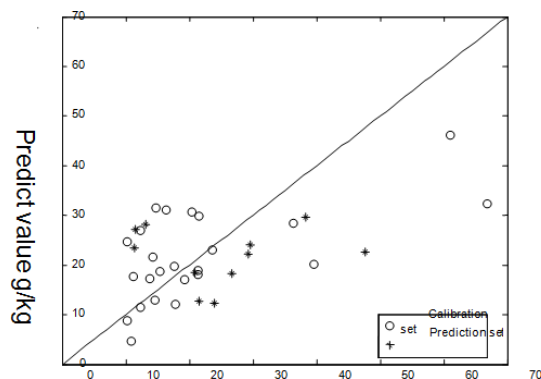


Fig. 1. Correlation between the measured and predicted values of honey tetracycline content for the second time: FT transfectance spectra ($4000\text{-}12500\text{cm}^{-1}$) .

3.2 Map of the signal optimization

Multiplicative scatter correction, standard variable transformation, the Fourier transform and wavelet transform are separately processed for the spectral matrix (X matrix), without taking into account the specific characteristics of the test the amount of information to be analyzed. Therefore, they are difficult to remove extraneous interference while without loss and to be measured on the spectral information. For this case, Wold, orthogonal signal correction (Orthogonal Signal Correction, OSC) method used to avoid the removal of important information on the forecast to be measured. Wold use Orthogonal Signal Correction method to resolve this problem. The OSC algorithm make the measurement information is also taken into account, and it remove the spectral information which is independent with variable matrix Y. The OSC algorithm has so many improved algorithms, one is a direct orthogonal signal correction (Direct Orthogonal the Signal Correction. DOSC), which is Westerhuis proposed, a direct orthogonal signal correction (DOSC) algorithm provides exact solution to orthogonal signal correction proposed by Wold. Just by a simple least squares method, DOSC can figure out those with the Y orthogonal, in the X matrix space and to the greatest extent describe the X matrix variation (variation) of the orthogonal factor.

We use DOSC to preprocess the near-infrared Fourier transmission and reflection spectra of honey tetracycline, and DOSC removed irrelevant information, which makes the difference among the spectra of each sample more obvious. Removing the irrelevant information with DOSC, we use PLSR model to establish the near-infrared spectrum and honey tetracycline.

Selecting parameters of DOSC and establishment the PLSR model, the number of factor selected from 1,2, and tolerance factor of DOSC from the 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, and PLSR best the main factor number is selected from 1-20. When DOSC tolerance factor is set too high, the result of model appeared over-fitting;

contrary, DOSC tolerance factor is set too low, the unwanted information filter not complete, and the precision of the model are also affected. When removing a factor of DOSC, and the tolerance factor is set to 0.35, the accuracy of the model is the highest, r increased from 0.617 to 0.6781, RSD_p (%) decreased from 52.12% to 45.65%.

3.3 Nonlinear calibration model of Tetracycline content in honey

The tetracycline content in honey is very low ($10^{-9}\sim 10^{-7}$), its non-linear effects and interference may also be larger. Therefore, we use LSSVM to establish the nonlinear model of tetracycline in honey, in order to explore the feasibility of quantitative detection. RBF kernel is used in this article, and LSSVR which used RBF kernel just need confirm two parameters, namely the regularization parameter γ and kernel bandwidth parameters. The parameters γ and σ^2 used two-step grid-search and remove one across-validation to determine.

Firstly, we follow the exponential growth to primaries parameter γ and σ^2 , and in the first choice, the range of parameter γ is from 1 to 109, and the range of parameter σ^2 is from 0.01 to 100. After first choice, the value of parameter γ is 1 and the value of parameter σ^2 is 81.92. Secondly, we use uniform growth to feature parameter γ and parameter σ^2 , in the first choice the initial range of the parameter γ is 0.1-10, and the initial range of σ^2 is 0.01~100. The growth step of parameter γ is 0.1, and the growth step of parameter σ^2 is 0.1. After second choice, the value of parameter γ is 1.2 and the value of parameter σ^2 is 9.8. After selecting the best parameter γ and the bandwidth parameters of kernel function, we establish the LSSVR model of Tetracycline content in honey, and the result of the model shown in Table 3. From the modeling results, the prediction of using LSSVR model is better than the PLSR model, RSD_p (%) decreased from 52.12% to 48.95%.

Table 3. The PLSR and LSSVR models constructed with different spectral preprocessing methods.

| Honey quality indicators | Model | Parameter | r | SEP | RSD_p (%) |
|---------------------------|--------------------------|----------------------------|--------|---------|-------------|
| Tetracycline(μ g/kg) | PLSR | LV=4 | 0.6170 | 12.9852 | 52.12% |
| | DOSC-PLSR | n=1, tol=0.35, LV=1 | 0.6781 | 11.3634 | 45.65% |
| | MSC-PLSR | LV=2 | 0.5736 | 12.4985 | 50.21% |
| | SNV—PLSR | LV=2 | 0.5734 | 12.4955 | 50.2% |
| | WT-PLSR | Bior3.3, J=1, LV=4 | 0.6169 | 12.9852 | 52.17% |
| | First derivative - PLSR | 2-9, LV=2 | 0.5327 | 12.3201 | 49.50% |
| | Second derivative - PLSR | 2-9, LV=2 | 0.5744 | 12.3930 | 49.79% |
| | LS-SVR | $\gamma=1.2, \sigma^2=9.8$ | 0.6618 | 12.1838 | 48.95% |

LV: PLSR, The number of main factors; n: DOSC, Number of factors; tol: DOS, The number of tolerance factor; J: Wavelet decomposition level; γ : LSSVR, The regularization parameter, σ^2 RBF, Bandwidth of kernel function parameters

4 Conclusion

DOSC remove the spectral information with Y orthogonal, and it make the linear relationship between spectra and to be measured is more obvious, which can improve the predictive ability of PLS Model and also simplify model. The value of RSD_p from 52.12% drop to 45.65% after DOSC treatment, so DOSC is a very effective pre-processing method in PLSR regression model.

We use LSSVR to build nonlinear model of tetracycline in honey, because our data base has less number of samples. The predict precision of LSSVR is better than PLSR, the value of RSD_p(%) decrease from 52.12% to 48.95%.

However, due to complex background, peak overlap, and the strong absorb effect of water, it is hard to accurately extract the effective information of tetracycline in honey with these methods, and to achieve rapid detect requirements.

References

1. Thomas E.V. A primer on multivariate calibration [J]. *Analytical Chemistry*, 1994, 66: 795A -804A
2. Daviesa M.C., Radovic B., Fearn T., et al. A preliminary study on the characterisation of honey by near infrared spectroscopy. *J. Near-Infrared Spectrosc*, 2002, 10(2): 121~135
3. Despaigne F., Massart D.L., Chabot P. Development of a robust calibration model for nonlinear in-line process data. *Analytical Chemistry*, 2000, 72: 1657~1665
4. Delphine J.R., Leardi R., Noord O.D., et al. Genetic Algorithms as a Tool for Wavelength Selection in Multivariate Calibration. *Analytical Biochemistry*, 1995, 67(23): 4295~4301
5. Serra B., Pesudo E.F., Pallarés G.J. Quantitative determination of free amino acids in honeybee collected pollen using gas chromatography and spectrophotometry. *Annales des Falsifications Expertise Chimique*, 1991, 897: 153~166