doi:10.3788/gzxb20174601.0112006

近红外漫反射光谱法快速检测苯磺酸氨氯地平片辅料含量

韩君[1],孙长海[1],陈爱明[2],方洪壮[1]
(1 佳木斯大学 药学院,黑龙江 佳木斯 154007)
(2 大连达硕信息技术有限公司,辽宁 大连 116023)

摘 要:将近红外光谱技术和化学计量学相结合快速检测苯磺酸氨氯地平片辅料含量.通过随机青蛙法、变量投影重要性和竞争自适应重加权采样筛选特征波长变量点,采用 9 种光谱预处理方法对原始光谱进行处理后,分别建立了偏最小二乘法模型和支持向量回归分析模型,并将这两种模型进行了比较.应用优选模型对样品进行了测试,结果表明:对于所涉及的样本,在最优特征波长变量选择上,随机青蛙法效果较好;在模型预测结果上,与支持向量回归分析模型相比,5 个指标的偏最小二乘法定量模型的决定系数,预测均方根误差评价参数效果较好,相对分析误差值均大于 3.0.样品测试值与实测值标准误差均小于 1.30,配对 t 检验表明,在 $a=0.05$ 显著性水平上,两者无显著性差异.因此,可采用近红外漫反射光谱法用于苯磺酸氨氯地平片辅料含量的快速检测,该方法重复性、中间精密度、线性、精确性良好,且可为其他药用辅料含量快速检测提供借鉴.

关键词:近红外光谱;辅料;支持向量机回归分析;随机青蛙法;苯磺酸氨氯地平片
中图分类号:O433; R917 **文献标识码**:A **文章编号**:1004-4213(2017)01-0112006-10

Near Infrared Diffuse Reflectance Spectroscopy for Rapid Detection of the Excipients′ Contents in Amlodipine Besylate Tablets

HAN Jun[1], SUN Chang-hai[1], CHEN Ai-ming[2], FANG Hong-zhuang[1]
(1 *School of Pharmacy, Jiamusi University, Jiamusi, Heilongjiang* 154007, *China*)
(2 *Dalian ChemDataSolution Information Technology Co. Ltd, Dalian, Liaoning* 116023, *China*)

Abstract: The excipients contents in the Amlodipine Besylate Tablet were rapidly detected combing of the near infrared spectroscopy and Chemometrics. Characteristic wavelength variation points were screened by methods of Random Frog, Variable Importance Projection and Competition Self-Adaptive Reweighted Sampling. After processing the original spectrum by 9 kinds of spectrum pre-processing methods, the Partial Least Squares (PLS) model and Support Vector Regression analysis (SVR) model were established respectively and compared to each other. And then the optimized model was applied to test the samples. The results show that the effect of Random Frog is better for the selection of optimal characteristic wavelength variables in the samples involved; For the model predictions, the effect of PLS quantitative model is better for the evaluation parameters in the determination coefficient and RMSEP of the five indexes, when compared with that of the SVR model, and the Relative Percent Difference (RPD) values are all more than 3.0. The standard error of the tested values and measured values for samples are

Foundation item: The Postgraduate Student Scientific Innovation Project in Jiamusi University (No. LM2015_082).
First author: HAN Jun (1988—), male, M. S. degree candidate, mainly focuses on NIR and chemometrics. Email:hjzy630@163.com
Supervisor(Contact author): FANG Hong-zhuang (1956—), male, professor, mainly focuses on calculation drug analysis. Email:fhz-chjms@sohu.com
Received: Jul. 22, 2016; **Accepted**: Oct. 10, 2016

both less than 1.30; Paired t-test shows that there is no significant difference at the significance level of $a=0.05$. So near infrared diffuse reflectance spectroscopy can be used to rapidly detect the excipients′ contents in the Amlodipine Besylate Tablets, with a good repeatability, an intermediate precision, a linearity, accuracy, can provide a good reference for the rapid detection of other pharmaceutical excipients′contents.

Key words: Near-infrared; Excipients; Support vector regression analysis; Random Frog; Amlodipine Besylate Tablets

OCIS Codes: 120.4630; 070.0070; 170.1580; 200.4560; 200.3050

0 Introduction

The quality differences between foreign original drugs and native generic drugs are always being noticed. And one reason for this quality differences is due to the differences in the composition and contents of pharmaceutical excipients. Although the quality of pharmaceutical formulations is related to the production equipment, quality management control and the production process, the quality, the category proportions of different excipients and the contents of the selected pharmaceutical excipients have directly effect on the efficacy and quality of medicinal formulations[1]. When one excipient or many types of excipient coexist with the active pharmaceutical ingredient, the analysis for the excipients in solid formulations is extremely difficult to implement. The chemical method, a determination method for excipients contents, is often difficult to use; the chromatography, which is suitable for the analysis of mixture, is also restricted due to the chemical properties of excipients, failing to reach the requirement for separation or the need of unconventional detector. Furthermore, due to the solubility characteristics of excipients in solid formulation, it is often difficult to make a suitable solution for analysis detection. With the existence of the above factors, a fast, efficient and accurate method for the analysis and detection of excipients in solid formulation are more urgently needed. While the near infrared spectroscopy has the advantages of detecting at a high speed, a good test reproducibility, and being suitable for simultaneous analysis on complex multi-components [2-4], *etc*.

When the whole segment data is used to model and analyze, a lot of redundant information in spectrum will make the workload of modeling calculation bigger and the time spent on it longer. In addition, due to the strong correlation between variables, not all the wavelengths can provide useful information. Optimizing wavelength variables, on the one hand, can eliminate the irrelevant variables, and at the time to simplify the model, can also improve the prediction accuracy and robustness of the model; on the other hand, can apply it to the implantable minimally invasive spectrometer based on the obtained characteristic wavelength, thus to effectively reduce the costs. Yan-Ru Zhao et al[5] select important wavelengths from the whole wavelengths using Random Frog algorithm (Random Frog), and establish the quantitative prediction model for mulberry fruit soluble solids. Qianyi Luo et al[6] select the key variables that with the most information through Competition Self-Adaptive Reweighted Sampling (CARS), and establish a rapid quantitative analysis method for icariin. Goodarzi M et al[7] combine the Interval Partial Least Squares (iPLS), Variable Importance Projection (VIP), Uninformative Variable Elimination (UVE), Moving Window Partial Least Squares (MWPLS) and random Interval Frog method (Interval Random Frog), and use them to screen the most informative combination, which is applied in a spectrum implantable minimally invasive sensor devices for the detection of glucose content in serum.

Support Vector Regression (SVR) is a new non-linear correction method developed in recent years, and shows many advantages that traditional algorithms do not have in the issues relevant to the small sample numbers, nonlinear, high-dimensional data space and other issues. Currently, SVR has been widely used in the structure-activity relationships, pattern recognition and spectral quantitative analysis[8-11]. This experiment takes amlodipine besylate as a sample, which is in the breed catalog of generics conformance assessment, establishes near infrared quantitative models for its corresponding excipients, and explores near infrared simultaneous quantitative method for excipients in complex analysis system of solid pharmaceutical formulations. The comparison of a variety of mathematical methods in this study shows that the effect of Random Frog is better for selection of optimal characteristic wavelength

variables in the samples involved; For the model prediction, PLS prediction model is superior to SVR, and is simple and fast, thus can be used for the rapid detection of excipients' contents in amlodipine besylate tablets.

1 Experiment section

1.1 Instrument and test parameters

Antaris Ⅱ Fourier transform near infrared spectrometer (Thermo Fisher, United States), ChemDataSolution 1.1.0 (Dalian ChemDataSolution Information Technology Co. Ltd), takes a built-in instrument background as a reference, selects the spectral range from 10 000 cm^{-1}～4 000 cm^{-1}, and with a scanning number of 64 times, a resolution of 8 cm^{-1}. The collected near infrared spectroscopy for amlodipine besylate tablets is shown in Fig. 1. Curves show sample of amlodipine besylate, microcrystalline cellulose, pregelatinized starch, Sodium carboxymethyl starch, povidone and magnesium stearate respectively.

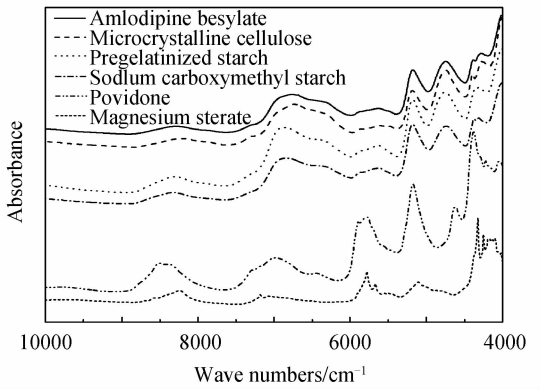

Fig. 1 Near-infrared diffuse reflectance spectrum charts of amlodipine besylates tablets

1.2 Preparation for samples

In accordance with the formla of Amlodipine besylate and counted by mass portion, there should be 100～400 portions of microcrystalline cellulose, 3～150 portions of pre-gelatinized starch, 100～200 portions of sodium carboxymethyl starch, 5～25 portions of povidone and 1～5 portions of magnesium stearate. The active ingredient and related pharmaceutical excipients of amlodipine besylate should be prepared into 100 portions of sample containing the main drug of 5 mg per portion, with the same weight of 1 g. OPUS 7.0 shall be adopted to conduct modeling design for the 100 samples, and the mortar is used to finely grind and mix evenly them, and pack in 10 mL penicillin bottles. In the data collection process, the temperature shall be maintained at about 25℃ and humidity at about 60% RH. Sample component content statistics are shown in Table 1.

Table 1 Statistical table of sample component content

Component content	English abbreviations	Minimum value/%	Maximum value/%	Average value/%	Standard deviation/%
Microcrystalline cellulose	MCC	29.09	68.83	48.99	9.52
Pregelatinized starch	PS	0.56	28.12	13.92	8.13
Sodium carboxymethyl starch	CMS-Na	19.30	37.75	28.66	5.30
Povidone	PVP	0.95	4.61	2.97	1.09
Magnesium stearate	MS	0.23	0.95	0.57	0.20

Kennard-Stone method is used to divide the sample set, and 65 samples are divided as a calibration set, while the remaining 35 as the validation set. It is used for model establishment and evaluation, which are shown in Fig. 2. Prepare the samples with the low, medium, high concentration accessory materials MCC (29.09%, 47.72% and 68.56%), PS (0.56%, 13.70% and 28.12%), CMS-Na (19.30%, 28.64% and 37.51%) ,PVP (0.95%, 3.10%and 4.60%) and MS (0.27%, 0.54% and 0.95%). Use them in the repeated and intermediate precision tests. And prepare 6 samples for the model test.

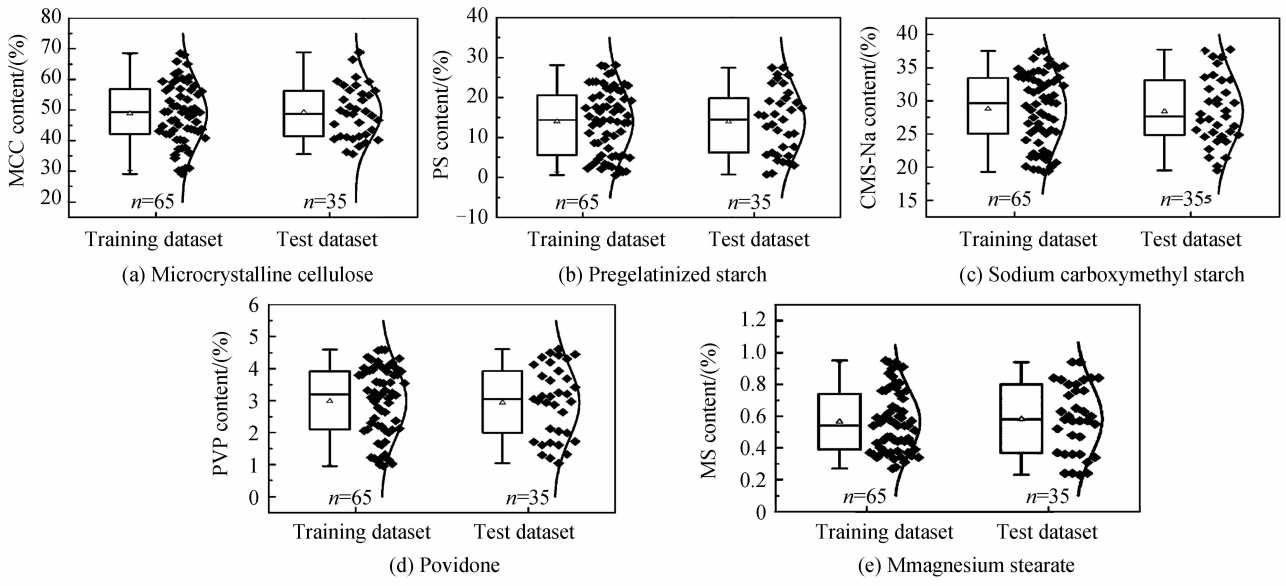

Fig. 2 The box plot and distribution of the contents of five pharmaceutical excipients training set and test set

1.3 Model evaluation

Root Mean Square Error of Calibration (RMSEC) and determination coefficient (R^2) are used to compare the predictive ability of calibration model, the Root Mean Square Error of Cross Validation (RMSECV) is used to select the number of latent variables, the Root Mean Squared Error of Prediction (RMSEP) is used to evaluate predictive ability of different PLS models. The homogeneity test of variance (reliability 95%) is used to compare the forecast error of the built model. In addition, the paired t test (reliability 95%) is used to evaluate the accuracy. Variance analysis (reliability 95%) is used to compare the experiment result and realize the accuracy assessment. The assessment of calibration models is related to the repeatability, intermediate precision, accuracy, linear relationship and other related parameters.

2 Results and discussion

2.1 Spectrum pre-processing methods

In the rapid detection of near infrared diffuse reflectance spectroscopy, since the near infrared spectroscopy measured by instrument not only contained the information of the sample itself, but also is affected by the testing environment of the equipment, personnel operation and the influence of the sample itself, so noise interference and some irrelevant information are appeared in the detection of near infrared spectroscopy. If the signal generated by original spectrum is used to establish quantitative analysis model, it will inevitably result in decreased model accuracy and precision. Therefore, adopting an appropriate combination and operation order of spectral pre-processing method for the near infrared spectroscopy to eliminate the effect of noise and some irrelevant information on spectrum, can improve the stability and predictive ability of the calibration model.

The common pretreatment methods for NIR include smoothing, derivative, Multiplicative Scatter Correction (MSC), Standard Normal Variate Transformation (SNV) and Orthogonal Signal Correction (OSC) *etc*. They can correct the different noise signal in the near infrared spectrum respectively, and derivative spectrum can effectively eliminate the interference of baseline and other background, distinguish overlapping peaks and improve resolution and sensitivity, in which, Gap method can improve its ability of eliminating the interference signal, which is the special case of Gap-Seg method to derivate; smoothing is a common method for noise reduction, in which, the core of Penalized Least Squares Smoothing (PLSS)[12] is that the optimization objective function of smoothing contains both the original spectrum and the least squares items of spectrum after noise removal, as well as the penalty term of least square (the first derivative of spectrum after noise removal), in which, the former is used to control the fitting error, while the latter is used to define the smooth degree of spectrum after noise removal; MSC is mainly used to eliminate the scattering influence induced by the particle′s uneven distribution and the particle′s size; SNV

is mainly used to eliminate the influence of the solid particle's size and surface scattering and change of optical path on the near-infrared diffuse reflectance spectroscopy; OSC[13-14] removes the information that is irrelevant to the tested component in spectral array by orthogonal projecting first, and then conducts multivariate calibration operation, lastly, it is anticipated to achieve the goal of simplifying the model and improving the predictive ability of the model. The quantitative analysis model of povidone is taken as an example, and the statistical parameters of quantitative analysis models established by different spectral pre-processing methods are shown in Table 2.

Table 2 The results comparison of Povidone models in different variable selection methods

Pre-Processing Methods	LVs	R_C^2	RMSEC	R_p^2	RMSEP
MSC+PLSS+SG	13	0.947 9	0.23	0.891 6	0.36
OSC+MedianFilter+Gap	9	0.998 0	0.049	0.997 3	0.058
SNV+GaussianFilter+Gap-Seg	15	0.970 5	0.19	0.889 9	0.36

The determination coefficient (R^2) and the RMSEP of models are taken as the index to find the best pretreatment method. R^2 is the determination coefficient of validation set, and the closer of its value to 1, indicating the more accurate prediction for the model; RMSEP is the root mean square error of prediction, the smaller the value, the higher predicted accuracy for models. It can be seen that the correlation coefficient of the model is closer to 1 after it is processed by the OSC + Median Filter+ Gap, and the RMSEC and RMSEP are both small, indicating that the model's prediction ability has been improved. Therefore, OSC + Median Filter + Gap are selected to preprocess the original spectra, which are shown in Fig. 3.

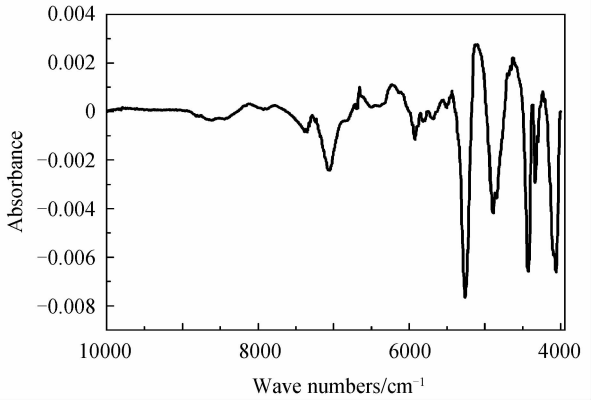

Fig. 3 NIR spectra after pretreatment by OSC+Median Filter+Gap

2.2 Variable selection

In order to investigate the optimal selection method of wavelength variable, Random Frog, CARS, VIP are used to optimize the wavelength variable for the whole spectral data, that is generated by pre-processing the spectra, in which, Random Frog[15-17] adopts the improved reverse jump Markov chain for model sampling, obtaining a plurality of models. The appearance probability of each variable in the model is counted, analyzed, and taken as the indicators and basis for the selection of important variables; CARS[18-19] firstly chooses the optimal subset of variables based on the so-called principle of "survival of the fittest", and then randomly divides the samples to build models, calculates the interaction test error of the variables subsets, and lastly takes the variables subsets that with the smallest mean value as the output; the core idea of VIP[20] is that the model score can reflect the explanation ability for the response value y through variable, if the score corresponded to a variable has strong explanatory power, and also has significant contribution to the construction of the model, then it will indicate that the variable is very important and needs to be selected.

As can be seen from Fig. 1, within the entire near-infrared spectral range, the near infrared spectroscopy characteristic peak positions of samples and pharmaceutical excipients of amlodipine besylate can not be identified by comparison visually, and therefore chemical metrological must be adopt for analysis. The spectral analysis of Random Frog preferred features variable of 5 kinds of excipients and pharmaceutical excipients for Amlodipine Besylate Tablets is shown in Table 3.

Table 3 Functional group assignment in NIR selected wavenumbers

Excipients	Selected wavenumbers/cm^{-1}	Assignments
MCC	6 897	2vO-H
	5 618	2vC-H
	5 495	O-H str. +C-O str.
	4 283～4 386	C-H str. +CH_2 def.
	4 019	C-H str. +CH_2 str.
PS	6 897	2vO-H
	4 762	O-H def. +C-O str.
	4283～4 386	C-H str. +CH_2 def.
CMS-Na	5 063～5 128	RCO_2R
	6 897	2vO-H
	4 000	C-H str.
PVP	8 000～8 150	3vC-H str.
	6 840～7 145	3vC-H
	6 840～7 450	2vC-H str. +C-H def.
	6 740	2vCONHR
	5 470～5 720	2vC - H
	4 730～4 750	C-H def. + C=O str.
MS	7 282	2vC-H str. + C-H def.
	4 945	C-H str. + C=O str.
	4 880	C-H str. + C=O str.

Symbol in the table:

"2v", "3v": respectively, is the first overtone and second overtone;

"str.", "def.": respectively, is the stretch vibration and the deformation vibration.

Table 4 The results comparison of Povidone models in different pretreatment methods

Method	Variable number	LVs	R_C^2	RMSEC	R_p^2	RMSEP
Random Frog-PLS	38	4	0.998 3	0.046	0.997 8	0.051
CARS-PLS	21	8	0.997 5	0.055	0.996 6	0.063
VIP-PLS	249	9	0.996 0	0.068	0.994 1	0.085

The PLS model that is established through the three screening methods for wavelengths variables point is shown in Table 4, with the quantitative model of povidone as an example, RMSEC and RMSEP in the model established by Random Frog are both small, indicating that the predictive ability of the model has been improved.

Fig. 4 is the screening process of the wavelength variables of the random frog method. Fig. 5 is the wavelength selection result of the variable importance in the projection. The final selected wavelength variables are 38 and 249.

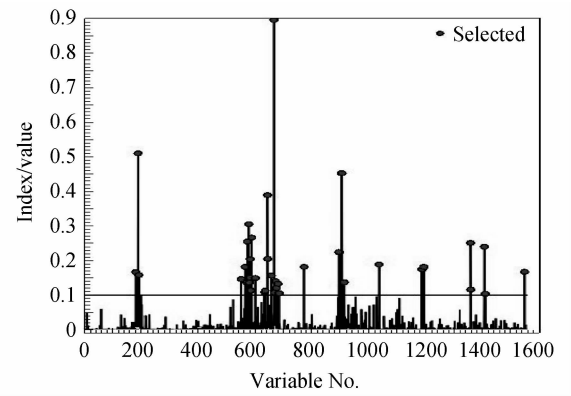

Fig. 4 Selection results of characteristic wavelength by Random Frog

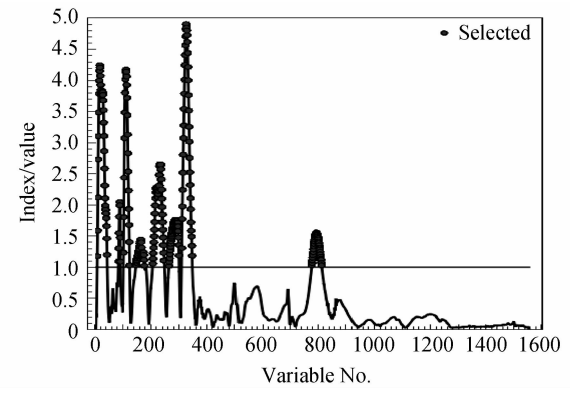

Fig. 5 Selection results of characteristic wavelength of variable importance projection

2.3 The comparison for modeling method

Support Vector Machine (SVM) method was firstly proposed for pattern recognition, and with the introduction of ε-insensitive function, SVM has been extended to nonlinear regression and function approximation, and has shown good learning performance, which is called supporting vector regression correction method[21-23]. The establishment of the optimal quantitative model in SVR needs to optimize the model parameters, and the key is to intensively find the optimal parameter setting in samples. In this study, the type of kernel is RBF kernel, which is used to build SVR model, and the parameters that are needed to be optimized include: kernel parameter gamma and penalty coefficient C. After they are optimized by the grid search, the usage of SVR method for modeling and the forecasting results are shown in Table 5. Paired t-test is used to detect the significance of differences between the predicted values and the measured values, Paired t-test results (shown in Table 5) show that, $|t| < t(0.05,34)$ (look up the T Distribution Table $t_{0.05,34}=2.032$), $P>0.05$, the results tested by two methods have no significant difference.

Table 5 SVR method modeling and prediction results

Excipients	C	Gamma	R_C^2	RMSEC	R_p^2	RMSEP	t-Distribution	P Values
MCC	100	0.01	0.998 6	1.03	0.984 4	1.06	1.527	0.136
PS	100	0.01	0.999 5	0.19	0.998 4	0.32	0.307	0.761
CMS-Na	10	0.01	0.991 0	0.51	0.987 0	0.58	0.923	0.363
PVP	100	0.01	0.996 3	0.066	0.996 3	0.067	1.537	0.138
MS	1	0.01	0.917 8	0.055	0.875 2	0.076	0.464	0.523

PLS predictive model is established with the optimal characteristic wavelength variable selected by Random Frog and the contents of excipients in amlodipine besylate tablets, and the results of modeling and forecasting are shown in Table 6. Paired t-test is used to detect the significance of differences between the predicted values and the measured values, Paired t-test results (shown in Table 6) show that, $|t| < t(0.05,34)$ (look up the T Distribution Table $t_{0.05,34}=2.032$), $P>0.05$, the results tested by two methods have no significant difference.

Table 6 PLS method on five kinds of excipients modeling and prediction results

Excipients	LVs	R_C^2	RMSEC	R_p^2	RMSEP	t-Distribution	P Values
MCC	7	0.997 2	0.50	0.996 5	0.53	0.043	0.966
PS	4	0.999 2	0.22	0.998 9	0.27	0.751	0.458
CMS-Na	5	0.994 0	0.39	0.993 0	0.45	0.277	0.784
PVP	4	0.998 3	0.046	0.997 8	0.051	0.252	0.803
MS	6	0.997 7	0.010	0.997 0	0.011	1.465	0.152

In order to analyze the reliability and applicability of different models, PLS and SVR models are analyzed and compared in predicting the contents of excipients in amlodipine besylate tablets, in which, the optimal model is determined based on the size of R^2, RMSEC and RMSEP, and the relative percent deviation of validation set (RPD) is used for further evaluation of model effect; if RPD≥3.0, then it indicates that the calibration effect is good and the established calibration model can be used for actual detection; if 2.5 <RPD<3.0, it indicates that the quantitative analysis for this ingredient by near infrared spectroscopy is feasible, but the prediction accuracy is required to improve; if RPD<2.5, it indicates that quantitative analysis for the ingredient by near-infrared spectroscopy is difficult[24-25]. Although we can't observe there are great differences between RMSEP value in the PLS model and RMSEP value in the SVR model (F test, confidence interval 95%), Amlodipine nesylate tablets excipients quantitative model parameters R^2, RMSEC, RMSEP which are established by PLS are obviously better than those of the SVR model, and RPD value is greater than 3.0. It can be seen that the applicability and stability of PLS quantitative model is slightly better, and the PLS prediction model can also be more reliable and stable in the practical application.

2.4 The verification for near infrared quantitative analysis model

2.4.1 *Repeatability*

Use the samples with the low, medium, high concentration accessory materials, and repetitively

measure the near infrared spectrum 6 times, and the excipients′ content of each analytical concentrations are calculated and obtained through the established quantitative analysis model of amlodipine besylate tablets respectively, to investigate the repeatability of the method. The measured relative standard deviation values of the repeatability of various excipients under three concentration levels are below 0.38%, 1.45%, 0.78%, 1.20%, 1.64%respectively, showing that model has good repeatability.

2.4.2 *Intermediate precision*

Intermediate precision shows the laboratory internal changes, such as different equipment, different dates and different analysts. In this thesis, the factors include three samples, two analysts, different concentrations and different three days. The first analyst tests the sample in the morning, and the second analyst tests the sample in the afternoon. In this case, standard deviation of five excipients is below 0.36%, 0.49%, 0.35%, 0.02%, 0.01%. There is no significant difference between the results which are obtained by different analysts on different dates (variance analysis, reliability 95%). Finally, the assessed relative standard deviation is under 1.23%, 1.72%, 0.95%, 0.80%, 1.61% on different dates.

2.4.3 *Accuracy*

Accuracy is evaluated by the statistical parameters RMSEC and RMSEP, The averagevalue of the deviation of the validation set and the test set has been calculated, and the accuracy of the model is better, and the results are shown in Table 7.

Table 7 Sample accuracy results summary

Excipients	RMSEC	RMSEP	RMSEP of test set	Average value of the deviation of the validation set	Average value of the deviation of the test set
MCC	0.50	0.53	1.42	0.37	1.10
PS	0.22	0.27	1.32	0.15	1.16
CMS-Na	0.39	0.45	0.99	0.28	0.81
PVP	0.046	0.051	0.43	0.034	0.23
MS	0.010	0.011	0.06	0.008	0.04

2.4.4 *Linear relationship*

The near infrared method is a secondary analysis method, its linear mainly studies the relationships between prediction results and the first methods [26] and the results are shown in Table 8. Showing the NIR analysis method of validation set and the linear equation of known value, the P value corresponding to regression coefficients obtained by the test results is less than 0.05, which is significant and has statistical significance. That means the regression equation is established.

Table 8 Sample linear results summary

Excipients	Linear range	Linear equation	Slope confidence interval	Residual sum of squares
MCC	35.61%～68.83%	$Y=0.986X+0.693$	[0.967;1.004]	6.900
PS	0.72%～27.50%	$Y=0.991X+0.155$	[0.983;0.998]	1.100
CMS-Na	19.48%～37.75%	$Y=0.964X+1.006$	[0.940;0.988]	4.211
PVP	1.04%～4.61%	$Y=0.985X+0.043$	[0.971;0.998]	0.065
MS	0.23%～0.94%	$Y=0.979X+0.015$	[0.964;0.995]	0.003

2.5 The confirmation of model validity

To be precise, after the model is established, it can not be used for the measurement analysis directly. In order to check whether it can predict the result correctly, it is necessary to confirm the validity of the model before it is in actual use. The predication of a good model should be well consistent with the measurement result of the standard method, when it is evaluated by the validation samples. Besides, it should avoid not to be influenced by the factors such as the change of instrument and temperature and the background interference, but only be sensitive to the change of physical and chemical properties of the sample. A group (usually 5 to 10, of course, the more the better) of qualified samples has been found to be analyzed and tested with the established model. The property parameters of the qualified samples have been known, but they do not participate in the model building (these samples are known as validation set).

The near infrared model of amlodipine besylate tablets has been used to analyze the six samples which do not participate in the modeling. It can be found that the standard error of the test value and measured

value of five accessories are 1.30, 1.21, 0.90, 0.39 and 0.06 respectively. Then the paired *t*-test has been conducted to the test value and the measured value of near infrared model, and the *t*-distribution value is 1.142, 0.787, 1.868, 1.469 and 1.903 respectively, and the significance level is 0.305, 0.787, 0.121, 0.202 and 0.115 respectively. The results show that at the significant level of $a = 0.05$, the two have no significance difference.

3 Conclusions

Through the performance comparison of the quantitative model (established by three characteristic wavelength variable screening methods: Random Frog, CARS, VIP) for the excipients in amlodipine besylate tablet, it can be seen that the optimized method that is most suitable for the quantitative analysis of excipients in amlodipine tablets is Random Frog, which can reduce the calculation amount of modeling, shorten the time on building correction models, and improve the model prediction accuracy. The quantitative models established with characteristic wavelength variable selection method are much sensitive, so they could change with the wavelength variable point, making the mean square root error of model prediction too large, therefore, the method is more suitable for the uniform state test samples. Only when the screening methods for characteristic wavelength variable point is combined with the sample state and the target values of the model to be established, it can play a crucial effect in modeling.

With the comparison of the modeling methods of two different models, it can be found that the quantitative model (established with PLS) parameters R^2, RMSEC, RMSEP of excipients in amlodipine besylate tablets are significantly better than those of SVR model, and RPD values are all more than 3.0. The model, with high accuracy, could meet the needs of practical application, and could solve the promotion issue for model predictive of excipients' contents in amlodipine besylate tablets. This study also indicates that SVR model has more advantages in the quantitation for excipients, such as: the selection process for model parameters is easy to control, and only few parameters need to be optimized, etc., furthermore, it is expected to become a commonly used quantitative analysis and correcting method for the near infrared spectral pharmaceutical excipients.

The test results of unknown samples that are tested with quantitative analysis model of pharmaceutical excipients established with PLS showed that, the standard error of the tested values and measured values of unknown samples are both less than 1.30, and the results of paired t-test showed that there was no significant difference between the tested values and measured values of unknown samples at the significant level of $a=0.05$. So the near infrared diffuse reflectance spectroscopy can be used for the rapid detection of the excipients' contents in amlodipine besylate tablets, with the advantages of being operated simply and rapidly, green, accurate and reliable tested results, and good repeatability, stability, linearity, accuracy, which may provide a reference for the rapid detection of other pharmaceutical excipients' contents. In practical application, by increasing the sample volume of calibration set and prediction set to further optimize, validate and perfect the model, its applicability and reliability can be continuously improved, it can meet the needs of actual production better, and has important guiding significance for on-line detection; furthermore, it can further expand the application of near infrared spectroscopy, and it's expected to give power to the consistency evaluation of generic drugs.

References

[1] VIKRANT S, BHUPINDER S S. The regulation of pharmaceutical excipients[J]. *Journal of Excipients and Food Chemicals*, 2013, **4**(3): 95-106.

[2] QUYANG Q, ZHAO J, CHEN Q. Measurement of non-sugar solids content in chinese rice wine using near infrared spectroscopy combined with an efficient characteristic variables selection algorithm[J]. *Molecular and Biomolecular Spectroscopy*, 2015, **151**:280-285.

[3] ZHOU S, WANG Z, LU L, *et al*. Rapid quantification of stabilizing agents in single-base propellants using near infrared spectroscopy[J]. *Infrared Physics & Technology*, 2016, **77**: 1-7.

[4] DOU Y, QU N, WANG B, *et al*. Simultaneous determination of two active components in compound aspirin tablets using principal component artificial neural networks (PC-ANNs) on NIR spectroscopy[J]. *European Journal of Pharmaceutical Sciences*, 2007, **32**(3): 193-199.

[5] ZHAO Y R, YU K Q, HE Y. Hyperspectral imaging coupled with random frog and calibration models for assessment of

total soluble solids in mulberries[J]. *Journal of Analytical Methods in Chemistry*, 2015(2): 343782.

[6] LUO Q, YUN Y, FAN W, *et al*. Application of near infrared spectroscopy for the rapid determination of epimedin A, B, C and icariin in Epimedium[J]. *Rsc Advances*, 2014, **5**(7): 5046-5052.

[7] GOODARZI M, SAEYS W. Selection of the most informative near infrared spectroscopy wavebands for continuous glucose monitoring in human serum[J]. *Talanta*, 2016, **146**: 155-165.

[8] YAO X J, PANAYE A, DOUCET J P, *et al*. Comparative study of QSAR/QSPR correlations using support vector machines, radial basis function neural networks, and multiple linear regression[J]. *Journal of Chemical Information and Computer Sciences*, 2004, **44**(4): 1257-1266.

[9] THISSEN U, VAN BRAKEL R, DE WEIJER A P, *et al*. Using support vector machines for time series prediction[J]. *Chemometrics and Intelligent Laboratory Systems*, 2003, **69**(1): 35-49.

[10] PIERNA J A, BAETEN V, RENIER A M, *et al*. Combination of support vector machines (SVM) and near-infrared (NIR) imaging spectroscopy for the detection of meat and bone meal (MBM) in compound feeds[J]. *Journal of Chemometrics*, 2004, **18**(7-8): 341-349.

[11] AMENDOLIA S R, COSSU G, GANADU M L, *et al*. A comparative study of k-nearest neighbour, support vector machine and multi-layer perceptron for thalassemia screening[J]. *Chemometrics and Intelligent Laboratory Systems*, 2003, **69**(1): 13-20.

[12] KRÄMER N, BOULESTEIX A L, TUTZ G. Penalized partial least squares with applications to B-spline transformations and functional data[J]. *Chemometrics and Intelligent Laboratory Systems*, 2008, **94**(1): 60-69.

[13] GHASEMI J, NIAZI A. Spectrophotometric simultaneous determination of nitroaniline isomers by orthogonal signal correction-partial least squares[J]. *Talanta*, 2005, **65**(5): 1168-1173.

[14] NIAZI A, YAZDANIPOUR A. Spectrophotometric simultaneous determination of nitrophenol isomers by orthogonal signal correction and partial least squares[J]. *Journal of Hazardous Materials*, 2007, **146**(1): 421-427.

[15] YUN Y H, LI H D, WOOD L R E, *et al*. An efficient method of wavelength interval selection based on random frog for multivariate spectral calibration[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2013, **111**: 31-36.

[16] ALEXANDRIDIS A, PATRINOS P, SARIMVEIS H, *et al*. A two-stage evolutionary algorithm for variable selection in the development of RBF neural network models[J]. *Chemometrics and Intelligent Laboratory Systems*, 2005, **75**(2): 149-162.

[17] LI H D, XU Q S, LIANG Y Z. Random frog: an efficient reversible jump Markov chain Monte Carlo-like approach for variable selection with applications to gene selection and disease classification[J]. *Analytica Chimica Acta*, 2012, **740**: 20-26.

[18] LI H, LIANG Y, XU Q, *et al*. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration[J]. *Analytica Chimica Acta*, 2009, **648**(1): 77-84.

[19] ZHENG K, LI Q, WANG J, *et al*. Stability competitive adaptive reweighted sampling (SCARS) and its applications to multivariate calibration of NIR spectra[J]. *Chemometrics and Intelligent Laboratory Systems*, 2012, **112**: 48-54.

[20] CHONG I G, JUN C H. Performance of some variable selection methods when multicollinearity is present[J]. *Chemometrics and Intelligent Laboratory Systems*, 2005, **78**(1): 103-112.

[21] CHAPELLE O, HAFFNER P, VAPNIK V N. Support vector machines for histogram-based image classification[J]. *IEEE Transactions on Neural Networks*, 1999, **10**(5): 1055-1064.

[22] DRUCKER H, WU D, VAPNIK V N. Support vector machines for spam categorization[J]. *IEEE Transactions on Neural Networks*, 1999, **10**(5): 1048-1054.

[23] BELOUSOV A I, VERZAKOV S A, VONFRESE J. Applicational aspects of support vector machines[J]. *Journal of Chemometrics*, 2002, **16**(8-10): 482-489.

[24] MALLEY D F, MCCLURE C, MARTIN P D, *et al*. Compositional analysis of cattle manure during composting using a field-portable near-infrared spectrometer[J]. *Communications in Soil Science and Plant Analysis*, 2005, **36**(4-6): 455-475.

[25] MALLEY D F, RÖNICKE H, FINDLAY D L, *et al*. Feasibility of using near-infrared reflectance spectroscopy for the analysis of C, N, P, and diatoms in lake sediments[J]. *Journal of Paleolimnology*, 1999, **21**(3): 295-306.

[26] RITCHIE G E, ROLLER R W, CIURCZAK E W, *et al*. Validation of a near-infrared transmission spectroscopic procedure: part B: application to alternate content uniformity and release assay methods for pharmaceutical solid dosage forms[J]. *Journal of Pharmaceutical and Biomedical Analysis*, 2002, **29**(1): 159-171.