

# Analytical Methods

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the RSC Publishing peer review process and has been accepted for publication.

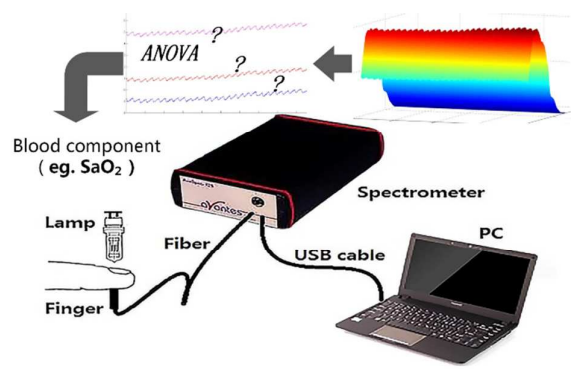
*Accepted Manuscripts* are published online shortly after acceptance, which is prior to technical editing, formatting and proof reading. This free service from RSC Publishing allows authors to make their results available to the community, in citable form, before publication of the edited article. This *Accepted Manuscript* will be replaced by the edited and formatted *Advance Article* as soon as this is available.

To cite this manuscript please use its permanent Digital Object Identifier (DOI®), which is identical for all formats of publication.

More information about *Accepted Manuscripts* can be found in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics contained in the manuscript submitted by the author(s) which may alter content, and that the standard [Terms & Conditions](#) and the [ethical guidelines](#) that apply to the journal are still applicable. In no event shall the RSC be held responsible for any errors or omissions in these *Accepted Manuscript* manuscripts or any consequences arising from the use of any information contained in them.

A novel method that uses ANOVA was proposed to the wavelength selection for oximetry, which can be put into other spectral analysis.



# Wavelength Selection Method based on Test Analysis of Variance: Application to Oximetry

Wei Li <sup>a</sup>, Ling Lin <sup>a\*</sup>, Gang Li <sup>b</sup>

<sup>a</sup> State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin 300072

<sup>b</sup> Tianjin Key Laboratory of Biomedical Detecting Techniques & Instruments, Tianjin University, Tianjin 300072

**Abstract**—In order to improve the precision and the reliability of the spectral measurement of blood oxygen saturation, the method of test analysis of variance was employed. We selected the preferred wavelength combination by the analysis of the coefficient distribution of oximetry from different wavelength combinations. Calculated by different combinations of wavelengths (660nm and 940nm, 660nm and 805nm, 805nm and 940nm, which are common used.) through the clinical data under different oxygen saturation, we firstly established the single factor test analysis of variance model of the oxygen saturation coefficient. Then the relative preferably wavelength combination can be selected out by comparative analysis of different combinations of wavelengths from the photoelectric volume pulse to provide a reliable intermediate data for further modeling. The experiment results showed that the wavelength combination of 660nm and 805nm responded more significantly to the changes in blood oxygen saturation. The introduced noise and method error are relatively smaller of this combination than other wavelength combination, which could improve the measurement accuracy of oximetry. The study applied the test variance analysis to the selection of wavelength combination in the blood oxygen result measurement, which result is significant. The study provided a new idea for the blood oxygen measurements and other spectroscopy quantitative analysis.

**Keywords**—The analysis of variance; Oximetry; Wavelength selection; Photoplethysmography (PPG); NIR Spectroscopy; Biomedical optics

## INTRODUCTION

In the near-infrared spectroscopy noninvasive measurement of blood components, only blood oxygen monitoring have been widely used in clinic. The value of human blood oxygen saturation has a direct relationship with cardiopulmonary function, the respiratory and circulatory system. In anesthesia, medical care and in the research on the process during exercise and sleep, monitoring blood oxygen has a very important role <sup>1-3</sup>. Scholars have done a lot of studies on oxygen detection <sup>4-8</sup>, and a variety of commercial oximeter are sold on the market <sup>9</sup>. The low radiation attenuation in the human tissues in the spectral interval [600–1100nm], the so-called therapeutic window, makes the light of the window to the most potential tool for the blood component noninvasive measurement <sup>10</sup>. However, according to the fact that the blood volume changes in the cyclical fluctuation followed by the heartbeat, different researchers extracted different wavelength combinations from Photoplethysmography (PPG) for oximetry. Silva <sup>11</sup> et al. presented a transmittance pulse oximetry system based on near-infrared (NIR) laser diodes (750 and 850 nm) for monitoring a wide range of oxygen saturation levels of arterial blood hemoglobin. Lopez-Silva. <sup>12</sup> et al. have been working onto

the application of near infrared range wavelengths to replace those of the classical pulse oximeter with the development of laser diodes based sensors, processing algorithms, a calibration procedure. Meanwhile, the relevant researchers have discussed a series of methods which is used to select the wavelengths of the visible-near infrared spectral for the measurements of various components in blood. They have analyzed the wavelength range which can be more effective for spectral measurements. Zourabian <sup>13</sup> et al chose the most optimal wavelengths (the first in the range of 670–720 nm and the second in the range of 825–925 nm) and discuss the dependence of their measurements on the fetal position. Mannheimer <sup>14</sup> et al found that sensors fabricated with 735 and 890 nm emitters had read more accurately at low saturation under a variety of conditions than sensors made with conventionally used 660 and 900 nm band emitters from numerical modeling. Camille <sup>15</sup> et al. have taken the standard deviation as the indicator, based on the two-dimensional correlation spectroscopy at the range of 650nm~1000nm, they found that the obtained standard deviation are the smallest at all the noise levels in a variety of oxygen measurement, when the wavelength combination the 660nm and 735nm, 660nm and 640nm selected. Uyuklu <sup>16</sup> et al. have used standard deviation and the variance as an indicator, then observed that the reversible aggregation of red blood cells is significant around 800nm of the transmission spectrum at 500~900nm. The above analysis method for wavelength selecting is much more complex, and some researchers selected the wavelengths without a description. Furthermore, no wavelength combination selected from actual clinical oximetry to be compared and analyze underlying reasons. Thus, it is of a very important significance to find a suitable method to wavelength selection of oxygen measurement for further study.

ANOVA <sup>17-21</sup> (Analysis of Variance) is a method to study the correlation between a number of factors and an indicator, i.e. the significance level of the impact of factors on the indicator. Single-factor test is a test that only one factor changes, the other factors remain unchanged. In the tobacco industry <sup>19</sup>, metal smelting <sup>20</sup>, psychological behavior <sup>21</sup> and many other fields of scientific research and experimental data processing, test analysis of variance has been widely used. However, in the near-infrared measurement of blood components, analysis of variance test is rarely used for the analysis and processing of measurement data. A variety of spectral processing methods inevitably produce the theoretical error, and it is impossible to completely eliminate spectral measurement noise. Therefore, these errors can be combined as the total random noise <sup>22</sup>, we can apply the single factor analysis of variance model to determine whether the changes in spectral data are really caused by component concentration.

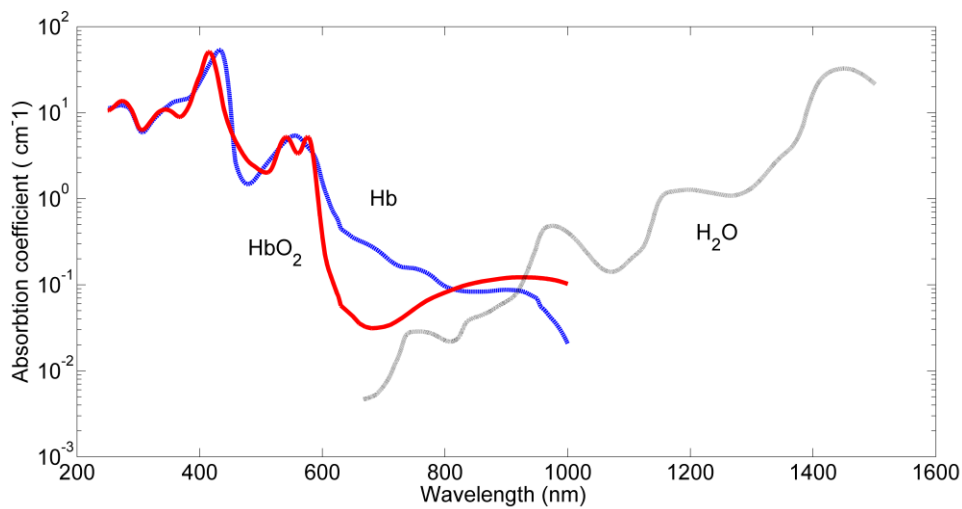
In the measurement of blood oxygen saturation, two specific wavelengths has been selected to calculate the intermediate value, oxygen saturation coefficient, then the oxygen saturation value can be obtained by empirical formulas or fitting methods. In this work, we apply the statistical method of test analysis of variance to comparatively analyze the significance of oxygen saturation coefficients, calculated by different wavelength combinations (660nm and 940nm, 660nm and 805nm, 805nm and 940nm). Furthermore, we should analyze the difference of the intermediate values under the specified wavelength combination of different oxygen saturations. If this difference becomes larger than random noise, then this difference cannot be explained by random disturbances, but only caused by the changes of oxygen saturations, so it is valid to extract this wavelength combination for further analysis model. Conversely, if this difference equally matches random noise, then it is considered that the wavelength combination cannot reflect the information of oxygen saturation. Also the significance strength lever of the difference can be judged by the indicators of

analysis of variance.

This idea has a potential to be applied to other spectral analysis. The effective information which adequately represents the measured can be extracted by the analysis of variance. The total measured interference information can be also eliminated.

## METHODOLOGY

In visible-near infrared band, 300~1500nm, the absorption coefficient curves<sup>23</sup> of reduced hemoglobin (Hb), oxyhemoglobin (HbO<sub>2</sub>) and water (H<sub>2</sub>O) are shown in Figure 1. For oxygen saturation measurement, Lambert-Beer's law is commonly used based on the method of light transmission and the difference between the light absorption coefficient of Hb and HbO<sub>2</sub> is employed. In the commercial pulse oximeter, oxygen pulse components are considered only caused by the arterial pulse, not the other part. The corresponding pulse wave signal can be obtained by filtering and separating from alternating current (AC) and direct current (DC) components of photoplethysmography (PPG). Usually finger and ear lobe are selected to measure the blood oxygen.



**FIGURE 1 Absorption spectra of hemoglobin and water at 300~1500nm**

When a specific wavelength incident light  $I_0$  irradiating to the finger, the transmitted light intensity can be written as

$$I = I_0 e^{-\varepsilon_0 C_0 L_0} e^{-\varepsilon_{HbO_2} C_{HbO_2} L} e^{-\varepsilon_{Hb} C_{Hb} L} \quad (1)$$

The transmitted optical signal is consist of the two parts. One is the attenuated signal after absorbed by the skin, muscles and tissues. This transmitted light intensity, the non-pulsating component, is called *DC* component, denoted  $I_{DC}$ . Its absorption coefficient, the concentration of light-absorbing material and the light path length are expressed as  $\varepsilon_0$ ,  $C_0$  and  $L_0$ , respectively. The other is the reduction light intensity absorbed by Hb and HbO<sub>2</sub> in blood. The transmitted light intensity is marked as  $I_{AC}$ . The absorption coefficient and the concentration of oxygenated hemoglobin (HbO<sub>2</sub>) are expressed as  $\varepsilon_{HbO_2}$  and  $C_{HbO_2}$ . The absorption coefficient and the concentration of reduced hemoglobin (Hb) are expressed as  $\varepsilon_{Hb}$  and  $C_{Hb}$ , and the optical path length is denoted by  $L$ . Since the vasodilation is caused by arterial pulse, the optical length  $L$  increases  $\Delta L$  and the transmitted light intensity  $I_{DC}$  changes to

$$I_{DC} - I_{AC} = I_{DC} e^{-(\varepsilon_{HbO_2} C_{HbO_2} + \varepsilon_{Hb} C_{Hb}) \cdot \Delta L} \quad (2)$$

$$D = \ln[(I_{DC} - I_{AC})/I_{DC}] = -(\varepsilon_{HbO_2} C_{HbO_2} + \varepsilon_{Hb} C_{Hb}) \cdot \Delta L \quad (3)$$

D is defined as a variable for calculating conveniently. Although the change of optical path length of the incident light transmitting in the human tissue is difficult to determine, path length change  $\Delta L$  of different wavelength light is basically the same. So two beams of different wavelengths ( $\lambda_1$  and  $\lambda_2$ ) were selected simultaneously to transact with human tissues and the following formula can be obtained.

$$\frac{D_{\lambda_1}}{D_{\lambda_2}} = \frac{\ln[(I_{DC}^{\lambda_1} - I_{AC}^{\lambda_1})/I_{DC}^{\lambda_1}]}{\ln[(I_{DC}^{\lambda_2} - I_{AC}^{\lambda_2})/I_{DC}^{\lambda_2}]} = \frac{\varepsilon_{HbO_2}^{\lambda_1} C_{HbO_2} + \varepsilon_{Hb}^{\lambda_1} C_{Hb}}{\varepsilon_{HbO_2}^{\lambda_2} C_{HbO_2} + \varepsilon_{Hb}^{\lambda_2} C_{Hb}} \quad (4)$$

When  $\lambda = 805nm$ ,  $\varepsilon_{HbO_2} = \varepsilon_{Hb}$ . The oxygen saturation,  $SaO_2$ , can be calculated by

$$SaO_2 = \frac{C_{HbO_2}}{C_{HbO_2} + C_{Hb}} = \frac{\varepsilon_{Hb}^{\lambda_2}}{\varepsilon_{Hb}^{\lambda_2} - \varepsilon_{HbO_2}^{\lambda_2}} \cdot \frac{D_{\lambda_1}}{D_{\lambda_2}} - \frac{\varepsilon_{Hb}^{\lambda_1}}{\varepsilon_{Hb}^{\lambda_2} - \varepsilon_{HbO_2}^{\lambda_2}} \quad (5)$$

In the transmitted light, AC component representing the DC component ratio is usually much less than 1%, so according to  $\lim_{x \rightarrow 0} \frac{\ln(1-x)}{x} = -1$ ,  $\frac{I_{AC}}{I_{DC}}$  can be used to approximate the alternative

$\ln[(I_{DC} - I_{AC})/I_{DC}]$ . Two constant coefficients A and B are introduced in the actual application,

$$A = \frac{\varepsilon_{Hb}^{\lambda_2}}{\varepsilon_{Hb}^{\lambda_2} - \varepsilon_{HbO_2}^{\lambda_2}}, \quad B = \frac{\varepsilon_{Hb}^{\lambda_1}}{\varepsilon_{Hb}^{\lambda_2} - \varepsilon_{HbO_2}^{\lambda_2}} \quad (6)$$

$$SaO_2 = A \cdot \frac{I_{AC}^{\lambda_1}/I_{DC}^{\lambda_1}}{I_{AC}^{\lambda_2}/I_{DC}^{\lambda_2}} - B = A \cdot R - B \quad (7)$$

Where R is called as oxygen saturation coefficient,

$$R = \frac{I_{AC}^{\lambda_1}/I_{DC}^{\lambda_1}}{I_{AC}^{\lambda_2}/I_{DC}^{\lambda_2}} \quad (8)$$

In the actual calculation, the oxygen saturation calculation is modified <sup>24-26</sup> by

$$SaO_2 = A_1 \cdot R^2 + B_1 \cdot R + C_1 \quad (9)$$

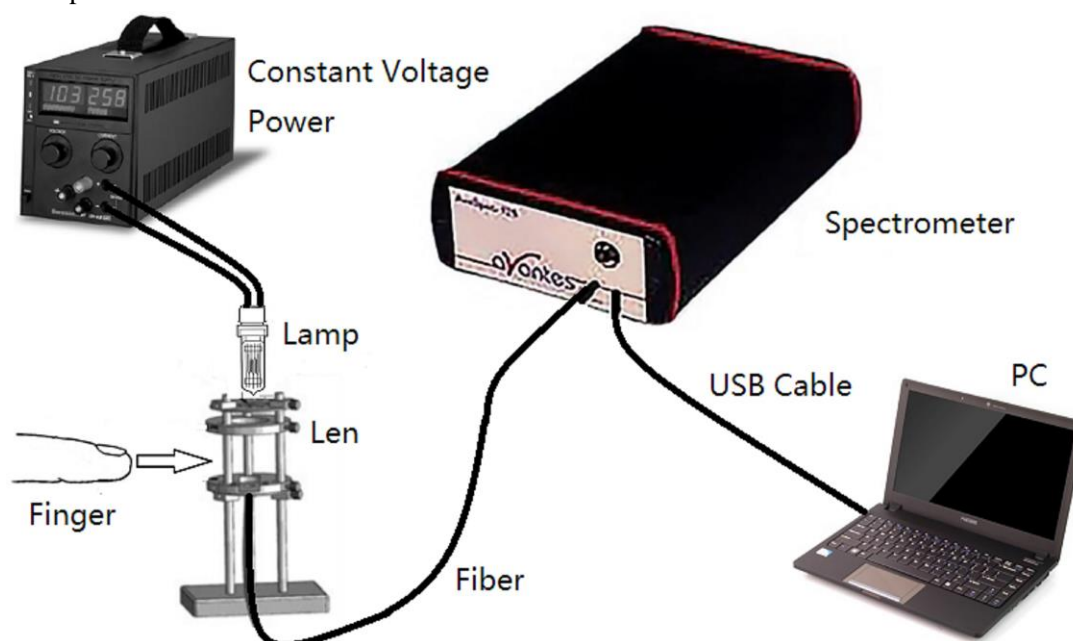
Therefore, after determining the constant values,  $A_1$ ,  $B_1$  and  $C_1$ , the oxygen saturation values can be calculated by R values measured by two-wavelength absorption and Eq.9.

## EXPERIMENT

### Experiment Setup

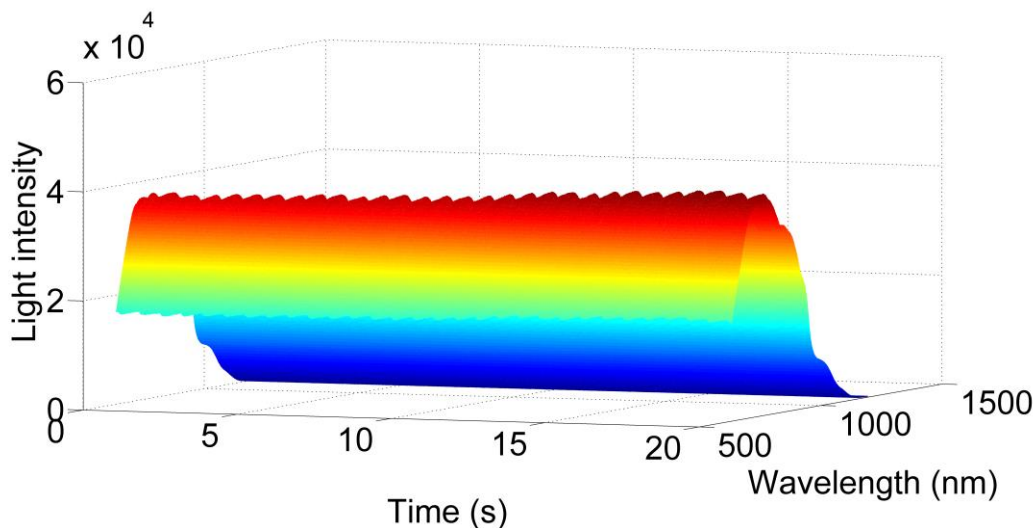
The spectral data for blood oxygen saturation were measured by the light transmission through finger in vivo. Figure 2 shows the schematic diagram of the measurement system. The measurement

system comprises four main parts- a light source, the optical fiber, the spectrometer, and a PC. The broadband source was a 50 W tungsten-halogen lamp (Philips). The broadband source lamp was supplied by a direct-current power source. The final spectra has been processed through the pre-elimination of background noise. The convex lens with a focal length  $F=63$  mm was used to concentrate light. The fiber bundle as a “-” line consisted of several fibers with numerical aperture (NA) of 0.22, which was connected to the spectrometer (AvaSpec-Hs1024×58TEC-USB2, Avantes). The spectrometer has a built-in cooling type back-illuminated CCD detector. The wavelength of the spectrometer ranged from 200 to 1160nm and the wavelength resolution can be variable from 1.5nm to 20nm. A USB was employed to connect the spectrometer and the computer. After the pending completion of data collection, data were transferred to a PC for saving. We chose the wavelength range 600.29 to 1147.50 nm in the study. During the data collection process, the spectrometer integration time per data point was set to 10ms, so the sample frequency was 100 Hz. Each person has been measured 2000 times.



**FIGURE 2 The schematic diagram of the measurement system**

We have performed a serial experiments from July 23 to August 24 in 2012 at the intensive care unit (ICU) of PLA 304 Hospital. From 7:00 to 9:00am every day, ICU patients were being measured. The instrument was allowed to warm up for at least 15 minutes before the measurement. Patients were requested to keep quietly and keep their one finger at the entrance of the fiber on the platform shown in Figure 2 during the measurement. We assisted them to remain the contact pressure stable and avoid body noise caused by the swing. During the measurement time, the patient's oxygen saturation was considered to maintain a constant. Figure 3 showed the photoelectric volume pulse wave (PPG) of No.1 patient in the test. At the same time, the patients' oxygen saturation were measured by a commercial pulse oximeter. When data acquisition completed, the tested object's arterial blood sample was immediately collected by the hospital professionals and the oxygen content has been analyzed through the Blood Gas Analyzer (Critical Care Xpress, Nova, USA).



**FIGURE 3** The 20s PPG of No.1 patient. In the timeline, each pulse wave can be seen corresponding to the wavelength, the transmission spectrum at different times can be seen from the wavelength axis, the depth of color in the figure represents the intensity of the transmitted light intensity. During this time the patient’ s oxygen saturation is considered to be not changeable.

These experiments were carried out in accordance with the guidelines issued by the Ethical Committee of Tianjin University. The experiments were also performed according to the relevant laws and institutional guidelines in order to not introduce any damages to the volunteers and all involved volunteers and their doctors were agreed to participate in this experiment in advance.

**Spectral data preprocessing**

According to the principle of oximetry, we have taken the extraction method of frequency domain analysis <sup>23, 27-29</sup> in this experiment. Based on the similarities of pulse waveforms and the linearity of the Fourier transform, frequency-domain analysis is proposed to solve the difficulties of extracting a single pulse peak directly at the low sample rate in the actual acquisition. The frequency spectrum has been obtained from each wavelength  $\lambda$  of each example spectral data by discrete Fourier transform (DFT). Then we extracted the frequency of 0Hz, the DC component,  $I_{DC}^\lambda$ , and extracted the first concentrated energy points of the spectrum (the frequency is 0.5~2Hz corresponding to the pulse period), the AC component,  $I_{AC}^\lambda$ . The DC and AC components were used to calculate the oxygen saturation coefficient,  $R$ .

At the same time, we have collected oxygen saturation (rounded to integer) of each patient through blood gas analysis. The standard value could be enrolled in carrying on the preliminary analysis. We have taken the data of 57 patients into use in all. Table 1 shows the distribution of persons at different oxygen saturations from 90% to 100%.

**TABLE 1** The distribution of persons at different oxygen saturations

Oxygen saturation (%)	90	95	96	97	98	99
Numbers (n <sub>i</sub> )	1	2	3	5	10	36

**Variance analysis model by selecting 660 nm and 805 nm wavelength combination**

Many researchers have selected the wavelengths of 660nm and 805nm as a combination to

measure the oxygen saturation<sup>30-32</sup>. As is shown in Figure 1, at 650nm, the light absorption coefficient difference between the Hb and HbO<sub>2</sub> is the maximum, so the detection sensitivity is the highest there. At 805nm, the light absorption coefficients of HbO<sub>2</sub> and Hb have the maximum gradient varying with the wavelength. From 805nm to 1000nm, the absorption coefficient difference of both is not particularly large. According to the data conversion of these two wavelengths, the value of HbO<sub>2</sub> and Hb can be easily obtained<sup>30, 32</sup>.

Thus we first extracted the spectral data near 660nm and 805nm from the clinical tests. The actual spectrometer data acquisition were at the wavelength of 660.04nm and 805.29nm. After extracting  $I_{DC}$  and  $I_{AC}$  of each patient by DFT from spectral data, we have calculated 57 oxygen saturation coefficients, according to Eq.8. Based on the theory of test of analysis of variance, the 57 oxygen saturation coefficients were listed in Table 2.

**TABLE 2 The R-values when 660nm and 805nm selected**

oxygen saturation (%)	C <sub>1</sub> 90	C <sub>2</sub> 95	C <sub>3</sub> 96	C <sub>4</sub> 97	C <sub>5</sub> 98	C <sub>6</sub> 99
Y <sub>1</sub>	0.785956	0.615788	0.571652	0.498135	0.482101	0.455877
Y <sub>2</sub>		0.585246	0.555875	0.532186	0.461732	0.488569
Y <sub>3</sub>			0.560315	0.520278	0.488512	0.440352
Y <sub>4</sub>				0.519985	0.469901	0.450758
Y <sub>5</sub>				0.538634	0.494435	0.471654
Y <sub>6</sub>					0.502347	0.407414
Y <sub>7</sub>					0.523014	0.421324
Y <sub>8</sub>					0.472834	0.431245
Y <sub>9</sub>					0.431347	0.453435
Y <sub>10</sub>					0.452375	0.407275
Y <sub>11</sub>						0.454925
.....						.....
Y <sub>39</sub>						0.418437

In Table 2, each column is response to a different oxygen saturation value. Based on the oxygen saturation measuring principle, R is a single theoretical value corresponding to each oxygen saturation value. In fact, because of the experimental measurement error, individual differences, the wavelength spectrum selection and extraction method, and the method of calculating the R-value, actual R-value has the deviation with the theoretical value. According to the basic idea of analysis of variance, we can compare the deviation and the impact of oxygen saturation to determine whether the R value still plays a significant role under such a series of errors in the spectral data and its extraction method when oxygen saturation changes.

The method of ANOVA has the following hypotheses:

- (a) at different oxygen saturation  $C_i$ , the spectral data are normally distributed,

$$Y_{ij} \sim N(\mu_i, \sigma^2), i = 1, 2, \dots, 6, \text{ where } \mu_i, \sigma^2 \text{ are unknown parameters.}$$

- (b) Under the  $C_i$  oxygen saturation, the obtained  $Y_{ij}, j = 1, 2, \dots, n_i$ , is a simple random sample. The samples of each of the oxygen saturation are independent. There is  $n_i$  samples under each blood oxygen saturation.

- (c) The method of ANOVA test the hypotheses that,  $H_0: \mu_1 = \mu_2 = \dots = \mu_6$ .

However, as ANOVA is not robust to unequal variances with unequal sample sizes— as in the

present example. The homogeneity (equivalent) of variance should be tested based on the Levene's Test or the modified Bartlett test before using ANOVA<sup>18, 33, 34</sup>. Only the significance value corresponding to the Levene statistic is greater than 0.05, the hypothesis that the group variances are equal could not be rejected. In general, variances of different groups from the experiment data of repetitive measurements are commonly treated to be approximately equal. The results of test of homogeneity of variances were listed in Table 3 calculated by SPSS software.

$Y_{ij}$  is the R-value calculated from the spectral intensity at the  $j$ -th measurement under the oxygen saturation  $C_i$ , as is shown in Table 1.  $\bar{Y}_{i\cdot}$  is the mean of R-values under the oxygen saturation of  $C_i$ , so the grand mean of all R-values can be calculated by the table data.

$$\bar{Y} = \frac{1}{6} \sum_{i=1}^6 \bar{Y}_{i\cdot} = 0.565361 \quad (10)$$

Take the calculated values as a group under the same oxygen saturation and the squared deviations of this group is

$$V_i = \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i\cdot})^2 \quad (11)$$

The  $V_i$  reflects the size of the random error. Since the noise levels of the tests are the same, i.e. the experimental variances are the same. The sum of each  $V_i$  can be used to represent the test error, which is denoted as

$$S_E = \sum_{i=1}^6 \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i\cdot})^2 = V_1 + \dots + V_6 = 0.030463 \quad (12)$$

Eq.12 is called as the sum of squares within the group samples, or the sum of squares of the error. The difference of R-values under the same oxygen saturation among the different spectral measurement data can be reflected by their own means  $\bar{Y}_{i\cdot}$  and the total average is  $\bar{Y}$ , which is related with the number of tests.

$$S_A = \sum_{i=1}^6 n_i (\bar{Y}_{i\cdot} - \bar{Y})^2 = 0.19825 \quad (13)$$

$S_A$  is called as the sum of squares between the group samples.

The basic idea of variance analysis is to find ratio of the sum of squares between groups and within the group, the ratio is denoted as  $F$ . The greater  $F$  is, the more significant the factor is. On the contrary, the smaller  $F$  shows the factor has no significant effect. In order to eliminate the disturbance of the number of data and the number of groups, the  $F$  value should be also divided by degrees of freedom. Here, the  $S_A$  should be divided by 5 (the number of groups minus 1), the  $S_E$  is divided by 51 (total number of trials minus the number of groups), we can get

$$F = \frac{S_A/5}{S_E/51} = 66.38102 \quad (14)$$

Thus the analysis of variance table can be listed, as is shown in Table.4.

### Variance analysis model by selecting 660 nm and 940 nm wavelength combination

To achieve the desired measurement sensitivity, the biggest different points of the absorption coefficients of the HbO<sub>2</sub> and Hb are selected at two wavelengths as the oxygen changes can be more

sensitive to be reflected. In the vicinity of 660nm and 940nm, light absorption coefficients of HbO<sub>2</sub> and Hb are of the maximum difference, and the absolute value of it were the opposite, as shown in Figure 1. In recent years, LEDs, with their center wavelengths at 660nm and 940nm, are usually selected as a light source in commercial oximetry instruments.

We also have extracted the spectral data near 660nm and 940nm from the clinical test. The actual spectrometer data acquisition were at the wavelength of 660.04nm and 940.27nm. According to Eq.8, 57 oxygen saturation coefficients have been calculated and  $I_{DC}$  and  $I_{AC}$  of each patient by DFT from spectral data have been extracted. Based on Levene's Test and the theory of test of analysis of variance, the results were listed in Table 3 and 4.

### Variance analysis model by selecting 805 nm and 940 nm wavelength combination

From the above analysis, it is known that 640nm, 940nm and 805nm are the important points for the oximetry, which represent the light absorption coefficient maximum difference points of Hb and HbO<sub>2</sub> and represent the maximum gradient point of light absorption coefficient varies with the wavelength. To a further comparative study, we have selected the wavelength combination of 805nm and 940nm to measure the oxygen saturation.

The data around the center wavelength of 805nm and 940nm were extracted from the clinical trial data. The actual spectrometer data acquisition were at the wavelength of 805.29nm and 940.27nm. According to Eq.8, 57 oxygen saturation coefficients have been calculated and  $I_{DC}$  and  $I_{AC}$  of each patient by DFT from spectral data have been extracted. The results of test of homogeneity of variances were listed in Table 3 and the analysis of variance table was listed in Table 4.

**Table 3 Test of Homogeneity of Variances (R-vlaues)**

wavelength combine	Levene Statistic	Degree of Freedom 1#	Degree of Freedom 2#	Sig.
660nm and 805nm	1.299	4	51	0.283
660nm and 940nm	1.478	4	51	0.223
805nm and 940nm	1.790	4	51	0.145

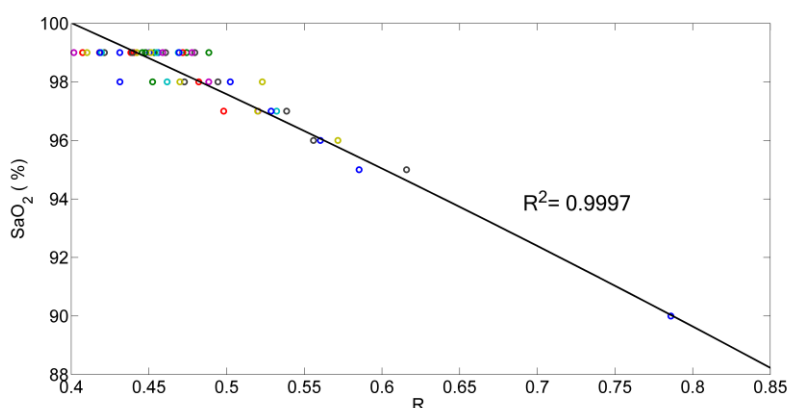
**TABLE 4 The test analysis of variance table when different wavelength combines selected**

wavelength combine	Source	Sum of Square	Degree of freedom	Mean Square	F-criterion	p
$\lambda_1=660\text{nm}$ and $\lambda_2=940\text{nm}$	Between Samples	0.136942	5	0.027388	$F_1$	$p_1$
	Within Samples	0.099485	51	0.001951	14.040458	1.23823E-08
	Total	0.236427	56			
$\lambda_1=660\text{nm}$ and $\lambda_3=805\text{nm}$	Between Samples	0.198250	5	0.039650	$F_2$	$p_2$
	Within Samples	0.030463	51	0.000597	66.381023	4.00104E-21
	Total	0.228712	56			
$\lambda_3=805\text{nm}$ and $\lambda_2=940\text{nm}$	Between Samples	0.025349	5	0.005070	$F_3$	$p_3$
	Within Samples	0.341881	51	0.006704	0.756283	0.585460
	Total	0.367230	56			

## DISCUSSION

According to the Levene statistic, as all the Sig. were all greater than 0.05 of different wavelength combinations in Table 3 (at the confidence level of 0.95), the hypothesis that the group variances were equal could not be rejected. So these data sets of the oxygen coefficient could be put into the analysis of variance. As can be seen from Table 4, when R-value obtained by the wavelength combination of  $\lambda_1=660\text{nm}$  and  $\lambda_2=940\text{nm}$ , at the confidence level of 0.99,  $F_1$  is greater than the F distribution quantile  $F_{0.99}(5, 51)=3.363276$  and  $p_1$  is less than the 0.01, the very small p-value indicates that differences between column means are highly significant, so the hypothesis  $H_0$  has been rejected. The R-value calculated by the wavelength selection of 660nm and 940nm can reflect relatively significantly the oxygen saturation changes at 99% confidence level, which can be applied to further analysis and modeling. Through the analysis of variance on the above two wavelength combinations,  $\lambda_1=660\text{nm}$  and  $\lambda_2=940\text{nm}$ ,  $\lambda_1=660\text{nm}$  and  $\lambda_3=805\text{nm}$ , the measuring principle has been verified and calculation method has been proved to be feasible as the oxygen saturation can be significantly reflected by the oxygen saturation coefficient calculated from the spectral data. While when the wavelength combination of  $\lambda_3=805\text{nm}$  and  $\lambda_2=940\text{nm}$  selected to be analyzed, the calculated oxygen saturation coefficient could not significantly reflect the different oxygen saturations, which has indicated that this group of combination could not effectively be used to measure oxygen saturation. From the principle of test analysis of variance, the larger the F-value of an extracting method is, i.e. the smaller p-value is, the more significantly to apply this extraction method to composition measurements and which can be more effective in distinguishing the component concentration differences. In the first two wavelength combinations,  $F_1 < F_2$ ,  $p_1 > p_2$ . Therefore, the comparative study of the test analysis of variance has shown that, under the same condition of the measurement principle and the calculation method, the wavelength combination of  $\lambda_1=660\text{nm}$  and  $\lambda_3=805\text{nm}$ , compared with the wavelength combination of  $\lambda_1=660\text{nm}$  and  $\lambda_2=940\text{nm}$ , can more significantly reflect the change in oxygen saturation. That is to say, the former carried more information of the oxygen saturation and its overall synthesized noise from the measurement method and the calculate error was smaller, so the wavelength combination of  $\lambda_1=660\text{nm}$  and  $\lambda_3=805\text{nm}$  is proved to be more efficient for oxygen saturation measurements.

In order to verify the select effectiveness of the wavelength combination,  $\lambda_1=660\text{nm}$  and  $\lambda_3=805\text{nm}$ , we compared differences between the oxygen saturation by the wavelength combination after ANOVA and accurate oxygen saturation by the blood gas analysis. According to Eq.9 and R-value in Table 2, the oxygen saturations were fitted. The calibration curve of oxygen saturation is shown in Figure 4.

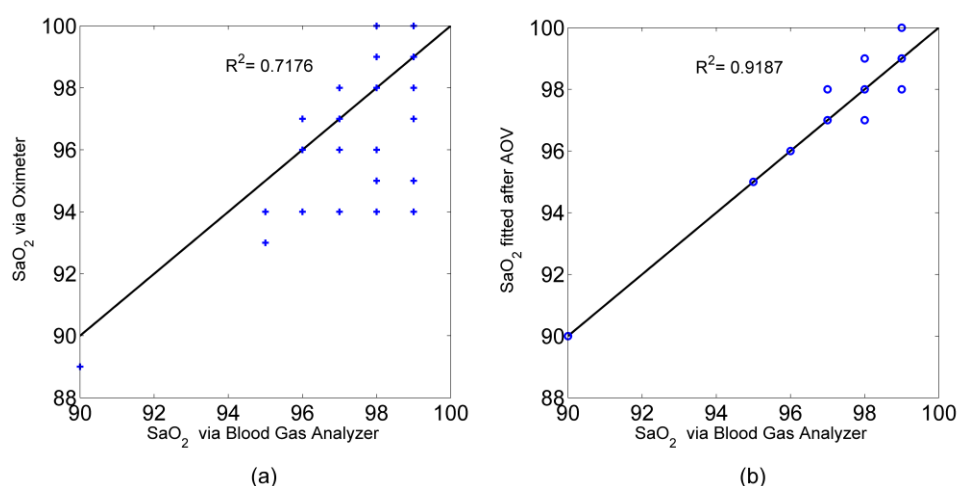


**Figure 4 The calibration curve of oxygen saturation.  $R^2$  is the correlation coefficient between the fitted curve and the measured values. In axis X,  $R$  is the oxygen saturation coefficient.**

When the wavelength combination of  $\lambda_1 = 660\text{nm}$  and  $\lambda_3 = 805\text{nm}$  selected, in the range of 88% -100%, oxygen saturation is calculated as

$$SaO_2 \% = -5.264 \cdot R^2 - 19.61 \cdot R + 108.7 \quad (15)$$

Linear correlation analysis was performed between oximeter measurements and the results of blood gas analysis, as is shown in Figure 5(a). According to Eq.15, oxygen saturations of 57 cases were calculated with  $R$ -value in Table 2. The calculated oxygen saturations were rounded off. Linear correlation analysis between the oxygen saturations after ANOVA and the results of blood gas analysis is shown in Figure 5(b). Compared to the actual application of the oximeter measurements, the calculated oxygen saturation value is closer to the true value. The coefficient  $R$ -value of the selected wavelength combination represented the change of oxygen saturation more significantly than the oximeter.



**Figure 5 Comparison with the results of blood gas analyzer**

**(a) Results by the oximeter (b) Results by the combination after ANOVA**

Furthermore, the absorption coefficients of the main components of blood in the visible-near infrared band were analyzed to make a preliminary explanation of above results. The absorption curve of reduced hemoglobin (Hb), oxygenated hemoglobin (HbO<sub>2</sub>) and water (H<sub>2</sub>O) at 300 ~ 1500nm is shown in Figure 1. Water absorption is the minimum at the 660nm and water has an absorption valley at 805nm where the absorption curves of Hb and HbO<sub>2</sub> are intersecting. However, near 940nm, the absorption coefficient difference of Hb, HbO<sub>2</sub> is less. Blood consists of four main components: plasma, red blood cells, white blood cells, and platelets. Plasma Accounts for 55-60% of total blood, while in plasma, water accounts for about 90-92%. Blood cells have a proportion about 40-45% in total blood volume, and in which water has a larger proportion than hemoglobin. The oximetry can eliminate influence of other components theoretically, but in the actual measurement process, the influence of other ingredients cannot be completely eliminated, especially the effects of water inevitably introduces the noise. Therefore it can be considered that, when  $\lambda_3=805\text{nm}$  or  $\lambda_2=940\text{nm}$  involved to model, the large degree of information of water results in the relatively poor significance. So when the wavelength combination of  $\lambda_3=805\text{nm}$  and  $\lambda_2=940\text{nm}$  selected to the analysis of variance, the oxygen saturation coefficient  $R$  cannot significantly explain

the changes in oxygen saturation. When the wavelength combination of  $\lambda_1=660\text{nm}$  and  $\lambda_2=940\text{nm}$  selected for analysis, compared to the wavelength combination of  $\lambda_3=805\text{nm}$  and  $\lambda_2=940\text{nm}$  wavelength combinations, the oxygen saturation coefficient R-value can just relatively significantly reflect the changes in oxygen saturation as water absorption is relatively small. Instead, when the wavelength combination of  $\lambda_1=660\text{nm}$  and  $\lambda_3=805\text{nm}$  selected, because of the minimal water absorption, the oxygen saturation coefficient thus can better significantly reflect the changes in oxygen saturation, which has verified the results of the analysis of variance. Through the above analysis, in the future study, we can select the large point of the water absorption curve for the first calculation step, and then calculate the error caused by the absorption of water at other wavelengths (such as 805nm, 940nm), finally make the error compensation to increase the accuracy of oxygen measurement.

## CONCLUSION

A new method based on test to the wavelength selection for oximetry has been proposed. Through the analysis and calculation of the wavelength combination, F-value and p-value, the indicators of test analysis of variance can be applied to determine the significance between the wavelength combination and oxygen saturation. In the experiment, after establishing the analysis of variance model of three wavelength combinations, 660nm and 940nm, 660nm and 805nm, 805nm and 940nm, we found that the wavelengths combination of 660nm and 805nm can be more effective for oximetry. In the analysis of variance model, the calculated oxygen saturation coefficient  $R$  contains a combination influence brought by the wavelength selection, calculation methods, measurement noise, etc. In comparative experiments, the differences between the results of the analysis of variance mainly caused by the different selections of wavelength combination. After analyzing the absorption coefficients of Hb, HbO<sub>2</sub> and water at 300~1500nm, we have explained the phenomenon why the measurements were relatively poor when the spectral data of 805nm, 940nm involved into calculation. Therefore, test analysis of variance can be used as a method to judge the superior wavelength selection for oximetry. This work has given a new idea to the characteristic wavelength selection, noise separation and the validation for the calculation method in the NIR noninvasive blood component measurement. Also the study has provided a way to further improve the accuracy of spectral data analysis.

## ACKNOWLEDGEMENTS

This work was supported in part by the State Key Laboratory of Precision Measurement Technology and Instruments (Tianjin University) under the National Natural Science Foundation of China (No.30973964) and the Tianjin Application Basis & Front Technology Study Programs (No.11JCZDJC17100).

## RERERENCES

- 1 P. K. Narotam, J. F. Morrison and N. Nathoo, *J. Neurosur.*, 2009, 111, 672.
- 2 J. P. Janssens, J. C. Borel and P épin J L, *Thorax*, 2011, 66, 438.
- 3 F. Chung, P. Liao and Elsaid H, *Anesth. Anal.*, 2012,114, 993.
- 4 M. K. Diab, W. M. Weber, and A.Al-Ali, *Method and apparatus for demodulating signals in a pulse oximetry system*, U.S. Patent 8,185,180. 2012-5-22.

- 5 M. Cannesson, and P. Talke, *F1000 medicine reports*, 2009, 1.
- 6 M. Raghuram, K. V. Madhav and E. H. Krishna, *Medical Measurements and Applications Proceedings (MeMeA)*, 2012 IEEE International Symposium on. IEEE, 2012, 1.
- 7 J. A. Sukor, S. J. Redmond and N. H. Lovell, *Physiol. Meas.*, 2011, 32, 369.
- 8 W. Yongqing W, Z. Zhipeng, W. Hongrui, L. Xiuling and S. Rongxia, *Biomedical Engineering and Informatics (BMEI)*, 2010 3rd International Conference on. IEEE, 4: 2010, 1376.
- 9 T. Pedersen, A. M. Møller and K. Hovhannisyan, *Cochrane Database Syst Rev*, 2009, 4.
- 10 S. J. Barker, J. Curry, D. Redford and S. Morgan. *Anesthesiology*, 2007, 105, 892.
- 11 S. M. L. Silva, M. L. D. Castilla and J. P. S. Martin, *J. Biomed. Opt.*, 2003, 8, 525.
- 12 S. M. Lopez-Silva, M. L. Dotor, J. P. Silveira, R. Giannetti, D. Golmayo, P. Martin and L. Herrera *Transmittance photoplethysmography and pulse oximetry with near infrared laser diodes*. In Instrumentation and Measurement Technology Conference, IMTC 04. Proceedings of the 21st IEEE, May, 2004.
- 13 A. Zourabian, A. Siegel, B. Chance, N. Ramanujam, M. Rode and D. A. Boas, *J. Biomed. Opt.*, 2000, 5, 391.
- 14 P. D. Mannheim, J. R. Cascini, M. E. Fein and S. L. Nierlich, (1997). *IEEE T.Bio-med.Eng.*, 1997, 44, 148.
- 15 C. Vazquez-Jacaud, G. Paez and M. Strojnik, *Annals. Biomed. Eng.*, 2011, 39, 1994.
- 16 M. Uyuklu, M. Canpolat and H. J. Meiselman, *J. Biomed. Opt.*, 2011, 16, 117006.
- 17 J. R. Crawford, P. H. Garthwaite and D. C. Howell, *J. Int. Neuropsych. Soc.*, 2003, 9, 989.
- 18 D. S. Moore, G. P. McCabe and W. M. Duckworth, *The practice of business statistics: using data for decisions*, Macmillan, 2003.
- 19 A. R. Vansickel, A. Shihadeh and T. Eissenberg, *Tob. Control*, 2012, 21, 377.
- 20 T. Stafilov, R. Šajin and Z. Pančevski, *J. Hazard. Mater.*, 2010, 175, 896.
- 21 S. Liu, M. J. Rovine and P. C. M. Molenaar, *Psychol. Meth.*, 2012, 17, 15.
- 22 G. Li, Z. Li, X. Li, L. Lin, B. Zhang and W. Wang, *Spectrosc. Spectral Anal.*, 2013, 26, 33, 1456.
- 23 S. Fantini, A. M. Franceschini, *Chapter 7 in Handbook of Optical Biomedical Diagnostics*, V. V. Tuchin, Ed., pp. 405–453, SPIE Press, Bellingham, WA, 2002.
- 24 G. LIU, W. GUO, P. ZHU, and M. BAI, *Laser & Infrared.*, 2009, 39, 169.
- 25 J. D. Passafaro, *U.S. Patent No. 4,651,741*. Washington, DC: U.S. Patent and Trademark Office. 1987.
- 26 G. Zonios, U. Shankar, and V. K. Iyer, *IEEE T.Bio-med.Eng.*, 2004, 51, 818.
- 27 G. Li, Q. Li and L. Lin, *Spectrosc. Spectral Anal.*, 2006, 26, 2177.
- 28 S. Fantini, M. A. Franceschini and J. S. Maier, *Opt. Eng.*, 1995, 34, 32.
- 29 M. Zhou, G. Li and L. Lin, *J. Biomed. Opt.*, 2013, 18, 057003.
- 30 Q. Yao, X. Liu, Y. Shi and H. Wang, *Opt. Precision Eng.* 2008, 16, 375.
- 31 F. Neudel, S. Takatani, H. Reul and G. Rau, *Med. Eng. Phys.*, 2002, 24, 301.
- 32 A. J. Spielman, G. Zhang, C. M. Yang, P. D'Ambrosio, S. Serizawa, M. Nagata, and R. R. Alfano, *Brain Res.*, 2000, 866, 313.
- 33 T. K. Keyes and M. S. Levy, *J. Edu. Behav. Stat.*, 1997, 22, 227.
- 34 T. S. Lim and W. Y. Loh, *Comput. Stat. Data An.*, 1996, 22, 287.