

Confidence intervals on fit parameters derived from optical reflectance spectroscopy measurements

Arjen Amelink
Dominic J. Robinson
Henricus J. C. M. Sterenberg

Erasmus Medical Centre Rotterdam
Center for Optical Diagnostics and Therapy
Department of Radiation Oncology
3000 CA Rotterdam, The Netherlands

Abstract. We validate a simple method for determining the confidence intervals on fitted parameters derived from modeling optical reflectance spectroscopy measurements using synthetic datasets. The method estimates the parameter confidence intervals as the square roots of the diagonal elements of the covariance matrix, obtained by multiplying the inverse of the second derivative matrix of χ^2 with respect to its free parameters by χ^2/ν , with ν the number of degrees of freedom. We show that this method yields correct confidence intervals as long as the model used to describe the data is correct. Imperfections in the fitting model introduces a bias in the fitted parameters that greatly exceeds the estimated confidence intervals. We investigate the use of various methods to identify and subsequently minimize the bias in the fitted parameters associated with incorrect modeling. © 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2982523]

Keywords: confidence intervals; fitting; optical spectroscopy; statistical error; systematic error.

Paper 08123RR received Apr. 14, 2008; revised manuscript received Jul. 3, 2008; accepted for publication Jul. 3, 2008; published online Sep. 22, 2008.

1 Introduction

Fitting *in vivo* optical spectroscopic data with mathematical models that describe the underlying physical “principles” is common practice in biomedical optics. Examples of such analyses include the fitting of multicomponent fluorescence spectra in tissue (that can include tissue autofluorescence as well as exogenous fluorophores), the fitting of multicomponent Raman spectra to determine the tissue biochemistry, and the fitting of multicomponent extinction spectra of tissue chromophores using reflectance spectroscopic measurements. *In vivo* fluorescence, Raman, and reflectance spectroscopic measurements are all affected by the tissue optical properties. The effect of the optical properties on the measured spectra must first be accounted for using physical light-tissue interaction models to obtain the “intrinsic” fluorescence, Raman, and extinction spectra. These undistorted spectra can then be analyzed with linear component analysis techniques such as singular value decomposition or other least squares minimization techniques. However, if the physical models used in the “pre-processing step” are inaccurate, an error in the fitted component values will be introduced. Furthermore, even if the algorithms used to extract the intrinsic optical spectra from the measured spectra are completely accurate, a lack of knowledge of either the *in vivo* spectral shape or even the physical presence of some of the components may introduce another source of error in the analysis of *in vivo* optical spectroscopic measurements. Finally, even when the intrinsic spectra are determined correctly and full knowledge about the presence and spectral shapes of all the components is available, another

source of error (uncertainty) in the analysis of optical spectroscopic measurements is introduced by a limited quality of the data (signal-to-noise ratio) that can prevent accurate determination of some or all of the components in the component analysis. The latter type of error is called a *statistical* error, while the first two types of errors are considered as *systematic* errors. These errors will introduce an uncertainty in the fitted parameter values extracted from the component analysis. Currently, it is not common practice in the field of biomedical optics to estimate the accuracy of the fitted parameter values from single spectroscopic measurements. However, we believe that it is useful to at least determine the statistical error in each of the fitted parameter values. For example, in the case of averaging multiple *in vivo* measurements, calculating the weighted average of the parameters using the statistical errors as weight factors would be more appropriate than simply calculating the unweighted average and ignoring the statistical uncertainties associated with the quality of the fits of the different spectra. Moreover, calculating the statistical errors in the fit parameters facilitates objective assessment of the quality of spectra and fits with the possibility to reject poor quality spectra or fit values from an *in vivo* dataset.

In principle, it is possible to calculate the parameter uncertainty by repeating the measurements and analyzing the standard deviation of parameters extracted from each of these multiple measurements. However, this is not practical for *in vivo* applications such as medical diagnosis, in which typically only one or a few measurements can be acquired. Scepanovic et al.¹ recently published an error analysis for single Raman spectroscopic measurements. They presented an analytical formula for estimating uncertainty expressed as a function of measurement noise, signal strength, and spectral over-

Address all correspondence to: Arjen Amelink, Center for Optical Diagnostics and Therapy, Department of Radiation Oncology, Erasmus MC, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands; Tel: +31 10 4632104; Fax: +31 10 4632141; E-mail: a.amelink@erasmusmc.nl.

lap. Their analysis characterizes uncertainty for linear fitting functions such as Raman and fluorescence spectroscopy, but has not been tested for nonlinear fitting functions encountered in reflectance measurements. In this work we validate a general method for determining the statistical error or confidence interval (CI) on fitted parameters derived from modeling single optical reflectance spectroscopic measurements using synthetic datasets. The method estimates the parameter CIs as the square roots of the diagonal elements of the covariance matrix, which is obtained by multiplying the inverse of the second derivative matrix of χ^2 with respect to its free parameters by χ^2/ν , with ν being the number of degrees of freedom. While the method is in principle valid for all spectroscopic techniques (fluorescence, Raman, and reflectance), we demonstrate the validity of the approach using reflectance spectroscopy as an example. Reflectance spectroscopy is sensitive to the absorption and scattering properties of tissue. The absorption coefficient is proportional to the concentration of the tissue chromophores, and the scattering coefficient reflects the size and density of scattering structures in the tissue. Several methods have been employed to extract physically relevant information (such as blood volume fraction and blood saturation) from reflectance spectra, e.g., using an analytical approximation of the transport equation such as the diffusion equation,²⁻⁶ Monte-Carlo-based models,^{7,8} or empirical methods.⁹⁻¹¹ Irrespective of the approach, a model function minimization or fitting routine must always be used to find the optimal values for the free parameters of the models. In general, a least squares (χ^2) optimization routine is used where the (weighted) square of the difference between the data points and the fitted curve is minimized. Most commercial software packages that employ least squares fitting routines automatically generate the covariance (also called variance-covariance or error) matrix. In principle, the diagonal elements of this matrix represent the variances of the parameters if 1. the fitting function is linear in the parameters and 2. the measurement uncertainties of the individual data points are known exactly. However, neither of these points is fulfilled for a single reflectance spectroscopy measurement. Nevertheless, we demonstrate that our simple method to estimate the statistical error of the fit parameters from single reflectance spectroscopic measurements yields correct CIs as long as the model used to describe the data is correct. Furthermore, we analyze the effect of imperfections in the fitting model (systematic errors) and show that even minor imperfections in the model may introduce a bias in the fitted parameters that greatly exceeds the estimated statistical CIs. Finally, we investigate the use of methods to identify and subsequently minimize the bias in the fitted parameters associated with incorrect modeling.

2 Theory

In this work we use the empirical model used to analyze a particular type of reflectance spectroscopy that we have been investigating over the past 5 years, differential path-length spectroscopy (DPS),¹¹⁻¹⁴ as an example. The model to which the differential path-length spectra are fitted can be written as:

$$R(\lambda) = \left[a_1 \left(\frac{\lambda}{\lambda_0} \right)^{a_2} + a_3 \left(\frac{\lambda}{\lambda_0} \right)^{-4} \right] \exp[-0.4\mu_a^{\text{total}}(\lambda)]. \quad (1)$$

The scattering function (in square brackets) is modeled by a combination of Mie and Rayleigh scattering, given by power law functions with amplitudes a_1 and a_3 , and wavelength dependencies $(\lambda/\lambda_0)^{a_2}$ and $(\lambda/\lambda_0)^{-4}$, respectively. Here λ_0 is a normalization wavelength, which we usually set to 800 nm. The differential reflectance signal is attenuated due to the presence of absorbers following Lambert-Beer's law with a path length equal to the fiber diameter,^{15,16} which in this exercise is assumed to be 0.4 mm. The absorption coefficient $\mu_a^{\text{total}}(\lambda)$ is the sum of the absorption coefficients of the chromophores present in the interrogation volume, which in this example are assumed to be betacarotene and blood:

$$\mu_a^{\text{total}}(\lambda) = \left\{ a_4 \mu_a^{\text{Bcar}}(\lambda) + a_5 [a_6 \mu_a^{\text{HbO}_2}(\lambda) + (1 - a_6) \mu_a^{\text{Hb}}(\lambda)] \times \left(\frac{1 - \exp\{-a_7 [a_6 \mu_a^{\text{HbO}_2}(\lambda) + (1 - a_6) \mu_a^{\text{Hb}}(\lambda)]\}}{a_7 [a_6 \mu_a^{\text{HbO}_2}(\lambda) + (1 - a_6) \mu_a^{\text{Hb}}(\lambda)]} \right) \right\}. \quad (2)$$

Parameter a_4 represents the concentration of betacarotene, a_5 is the blood volume fraction, a_6 is the microvascular blood oxygenation, and a_7 is the average vessel diameter. Input spectrum $\mu_a^{\text{Bcar}}(\lambda)$ is the specific absorption coefficient of betacarotene,¹⁷ $\mu_a^{\text{HbO}_2}(\lambda)$ is the absorption coefficient of fully oxygenated whole blood, and $\mu_a^{\text{Hb}}(\lambda)$ is the absorption coefficient of fully deoxygenated whole blood.¹⁸

With our experimental setup,¹¹⁻¹⁶ the data acquired in each reflectance spectrum consist of $N=2048$ pixel values $R_{\text{data}}(j)$ corresponding to the differential reflectance at wavelengths ranging from 350 to 1000 nm. Since the transmission efficiency of the setup is different for different wavelengths, the signal-to-noise ratio (SNR) is wavelength dependent. Clearly it is desirable to put less weight on pixels with a low SNR in the fitting routine. To accomplish this, the spectra are first smoothed by binning p pixel values and using the $N/p=n$ average values $R_{\text{data}}(i)$ and corresponding standard deviations $s(i)$ in the fitting routine. In our simulations, we have typically used a bin width of 15 pixels, which corresponds to an optical resolution of 6 nm, but we have also tested the effect of smaller and larger bin widths (10 and 30 pixels, respectively). The data are then fitted to Eq. (1) using a least squares Levenberg-Marquardt fitting routine using the standard deviations of the binned data points as weight factors. Chi-square (the quantity that is minimized) is given by

$$\chi^2 = \sum_{i=1}^{i=n} \left[\frac{R_{\text{fit}}(i) - R_{\text{data}}(i)}{s(i)} \right]^2, \quad (3)$$

where $R_{\text{fit}}(i)$ are the fitted reflectance values according to Eq. (1), $R_{\text{data}}(i)$ are the corresponding measured bin values, and $s(i)$ are the standard deviations of the bins. Note that in this approach, it is critical to bin the input spectra $\mu_a^{\text{Bcar}}(\lambda)$, $\mu_a^{\text{HbO}_2}(\lambda)$, and $\mu_a^{\text{Hb}}(\lambda)$ in exactly the same way as the measurement spectra. Furthermore, it is important to keep the bin

Table 1 Parameter values used for the synthetic datasets.

Parameter	Parameter description	Values
α_1 (-)	Mie scattering amplitude	1
α_2 (-)	Mie scattering slope	-1
α_3 (-)	Rayleigh scattering amplitude	0; 0.01; 0.1
α_4 (μM)	Betacarotene concentration	0; 5; 30
α_5 (%)	Blood volume fraction	0; 0.2; 1.0; 2.0; 5.0; 10; 20
α_6 (%)	Blood saturation	0; 25; 50; 75; 100
α_7 (μm)	Average vessel diameter	10

width small enough so that the approximation

$$R_{\text{data}}(i) = \frac{1}{p} \sum_j^{j+p} R_{\text{data}}(j) \propto \frac{1}{p} \sum_j^{j+p} \exp[-0.4\mu_a^{\text{total}}(j)] \\ \approx \exp\left[-0.4 \frac{1}{p} \sum_j^{j+p} \mu_a^{\text{total}}(j)\right],$$

holds. Particularly for wavelength regions where the absorption coefficient varies rapidly, e.g., around the Soret absorption band of hemoglobin, this approximation will become invalid for large bin widths.

The minimum value of χ^2 corresponds to values for the seven fit parameters (α_1 to α_7) that best describe the data. In principle, the CI of each of these parameters is given by the square root of the diagonal elements of the covariance matrix. The covariance matrix is the inverse of the second derivative matrix of χ^2 with respect to its free parameters and is exact, if and only if; 1. the fitting function is linear in its parameters, and 2. the weight factors used in the minimization routine are exact.^{19,20} In the case of our nonlinear model Eq. (1) with estimated weight factors $s(i)$, we estimate the covariance matrix by multiplying the inverse of the second derivative matrix of χ^2 with respect to its free parameters by χ^2/ν , with ν being the number of degrees of freedom.²⁰ The CI of each of the parameters is given by the square root of the diagonal elements of the estimated covariance matrix under the assumption that the model [Eq. (1)] is correct. We validate that the estimated covariance matrix is reasonably accurate for our nonlinear fitting problem in Eq. (1).

The calculated CI represents the statistical error only and does not fully account for systematic errors due to inaccuracies in the physical model used to fit the data. For example, if an additional absorber was present in the sampling volume but not in the model fit function, or if the wavelength dependence of the scattering function is not modeled correctly, the fitted parameters will be biased with respect to their true values. We investigate the magnitude of the bias in the fitted parameters for the case of 1. a missing absorber in the model and 2. an error in the modeled scattering function, and inves-

Table 2 Average number of times (out of 100) each of the fitted parameter values is within 2 CIs of their true values when the correct model is used, for bin widths of 10, 15, and 30 pixels. Standard deviations are indicated in parentheses. Noise amplitude=0.01.

	10	15	30
α_1	94 (4)	96 (4)	96 (7)
α_2	91 (4)	94 (4)	94 (5)
α_3	96 (3)	97 (4)	95 (8)
α_4	94 (3)	95 (5)	84 (25)
α_5	94 (3)	95 (3)	95 (7)
α_6	94 (3)	95 (5)	85 (22)
α_7	93 (4)	95 (3)	94 (9)

tigate the use of methods to identify and subsequently minimize the bias in the fitted parameters associated with incorrect modeling.

3 Simulations

We have generated 315 synthetic datasets, each consisting of 2048 pixel reflectance values $R_{\text{syn}}(j)$ using Eqs. (1) and (2) with the parameters as described in Table 1.

For each combination of parameter values, 100 different fits are performed by generating noise on the synthetic spectrum. The noise was generated by adding a random number between -1 and 1, multiplied by a specified noise amplitude to the pixels: $R_{\text{data}}(j) = R_{\text{syn}}(j) + \text{noiseamp} \cdot \text{random}[-1, 1]$. Note that the noise generated in this way is independent of wavelength. Since the data $R_{\text{syn}}(j)$ are wavelength dependent, the relative noise amplitude is in fact wavelength dependent, which is typically the case for optical spectroscopic measurements. For our experimental setup, typical signal-to-noise ratios at $\lambda_0=800$ nm are better than 100, and we have used noiseamp=0.01 for most of the simulations, but also checked the validity of our approach for larger noise amplitudes (0.1 and 0.5). By definition, if the calculated confidence intervals are correct, the fitted parameter values should be within 2 CIs in 95 out of the 100 fits.

4 Results

4.1 Fitting with the Correct Model

We first analyze whether the CIs calculated by taking the square root of the diagonal elements of the estimated covariance matrix are correct, as a function of bin width and noise amplitude. The fits are performed using the correct model described in Eqs. (1) and (2) with seven free parameters. Furthermore, we calculate the magnitude of the CIs as a function of blood volume fraction for this case.

4.1.1 Influence of bin width

Table 2 shows the number of times each of the fitted parameter values is within 2 CIs of their true values, out of the 100 fits and averaged for the 315 synthetic datasets (standard de-

Table 3 Average number of times (out of 100) each of the fitted parameter values is within 2 CIs of their true values when the correct model is used, for noise amplitudes of 0.01, 0.1, and 0.5. Bin width = 15 pixels.

	0.01	0.1	0.5
a_1	96 (4)	93 (3)	93 (3)
a_2	94 (4)	93 (3)	93 (3)
a_3	97 (4)	93 (3)	93 (3)
a_4	95 (5)	93 (3)	92 (3)
a_5	95 (3)	93 (4)	91 (6)
a_6	95 (5)	94 (2)	94 (3)
a_7	95 (3)	92 (3)	92 (3)

viations are indicated in parentheses) for different bin widths. If this number is close to 95, the CI was estimated correctly. If this number is much smaller, the CI was underestimated. The noise amplitude was set to 0.01.

Table 2 shows that on average, close to 95% of the fitted parameter values is within 2 CIs, except for the largest bin width for parameters a_4 (betacarotene concentration) and a_6 (blood saturation). For this large bin width, the optical resolution of the binned data (12 nm) becomes larger than the full

width at half maximum (FWHM) of the narrowest betacarotene feature (8 nm), causing an underestimation in the CI of the betacarotene concentration a_4 . Since betacarotene absorbs in the same wavelength region as oxyhemoglobin and deoxyhemoglobin, the accuracy of the blood saturation a_6 is affected as well. From Table 2 it is clear that the best results were obtained for a bin width of 15, which is used in the remainder of this work.

4.1.2 Influence of noise amplitude

Table 3 shows the number of times each of the fitted parameter values is within 2 CIs of their true values, out of the 100 fits and averaged for the 315 synthetic datasets (standard deviations are indicated in parentheses) for different noise amplitudes and a bin width of 15 pixels. On average, close to 95% of the fitted parameter values is within 2 CIs, even for very high noise amplitudes. Table 4 shows the average CI for each of the parameters as a function of blood volume fraction a_5 for the three different noise amplitudes. The CI increases with increasing noise amplitude, as expected. Thus for large noise amplitudes (0.1 and 0.5, corresponding to a signal-to-noise ratio at $\lambda=800$ nm of 10 and 2, respectively), the fitted parameters are also correctly within 2 CIs in not significantly less than 95% of the simulations, because the CIs are dramatically increased compared to the low noise amplitude (0.01, or signal to noise of 100). Note that for the low noise amplitude, which is typical for our *in vivo* data,¹¹⁻¹⁴ the CI of the fitted blood saturation a_6 is smaller than 1% for blood volume fractions $a_5 > 1\%$. Even for a blood volume fraction as low as

Table 4 Average CI for each of the parameters as a function of blood volume fraction a_5 for the three different noise amplitudes.

	Noise amplitudes	Noise amplitudes						
		0%	0.2%	1%	2%	5%	10%	20%
a_1 [-]	0.01	0.3	0.0003	0.0003	0.0003	0.0004	0.0006	0.001
	0.10	0.5	0.002	0.002	0.002	0.002	0.003	0.01
	0.50	2.5	2.4	0.01	0.01	0.01	0.02	0.05
a_2 [-]	0.01	0.002	0.001	0.001	0.001	0.002	0.002	0.004
	0.10	0.01	0.01	0.01	0.01	0.01	0.02	0.03
	0.50	0.07	0.07	0.06	0.06	0.07	0.09	0.2
a_3 [-]	0.01	0.0005	0.0002	0.0002	0.0002	0.0003	0.0005	0.001
	0.10	0.002	0.001	0.001	0.001	0.002	0.003	0.01
	0.50	0.08	0.02	0.006	0.006	0.008	0.01	0.05
a_4 [μM]	0.01	0.15	0.1	0.1	0.2	0.3	0.4	0.9
	0.10	0.9	0.9	1.0	1.1	1.5	2.3	6.0
	0.50	4.5	4.6	4.9	5.3	7.0	11	30
a_5 [%]	0.01	10^4	0.008	0.009	0.01	0.01	0.02	0.04
	0.10	10^4	0.07	0.08	0.08	0.1	0.14	0.3
	0.50	10^5	10^4	0.4	0.4	0.5	0.7	1.4
a_6 [%]	0.01	10^6	5.0	1.0	0.6	0.4	0.4	0.4
	0.10	10^8	34	7	4	2.4	2.0	2.1
	0.50	10^7	10^7	37	20	11	10	10
a_7 [μm]	0.01	10^6	0.8	0.2	0.1	0.05	0.04	0.08
	0.10	10^7	6	1.3	0.7	0.4	0.3	0.7
	0.50	10^7	10^3	7	4	2	1.7	3.4

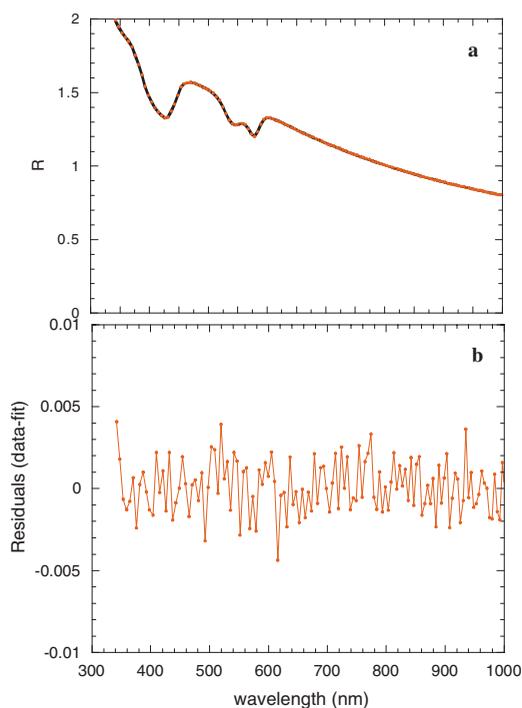


Fig. 1 (a) Typical fit and (b) the residuals for the synthetic spectrum with $a_3=0.01$, $a_4=5 \mu\text{M}$, $a_5=2\%$, and $a_6=75\%$ for a noise amplitude of 0.01 and a bin width of 15 pixels.

0.2%, the CI of a_6 is only 5% when the correct model is used in the fit. These numbers indicate that our instrument can extract the *in vivo* microvascular saturation with high accuracy, under the assumption that the model used to fit the data is correct.

Figures 1(a) and 1(b) show a typical fit and the residuals (\equiv data-fit), respectively, for the synthetic spectrum with $a_3=0.01$, $a_4=5 \mu\text{M}$, $a_5=2\%$, and $a_6=75\%$ for a noise amplitude of 0.01 and a bin width of 15. The fitted values for the parameters and the CIs were $a_1=1.0003 \pm 0.0002$, $a_2=-1.000 \pm 0.001$, $a_3=0.0099 \pm 0.0002$, $a_4=5.0 \pm 0.2 \mu\text{M}$, $a_5=2.00 \pm 0.01\%$, $a_6=74.9 \pm 0.7\%$, and $a_7=10.04 \pm 0.05 \mu\text{m}$.

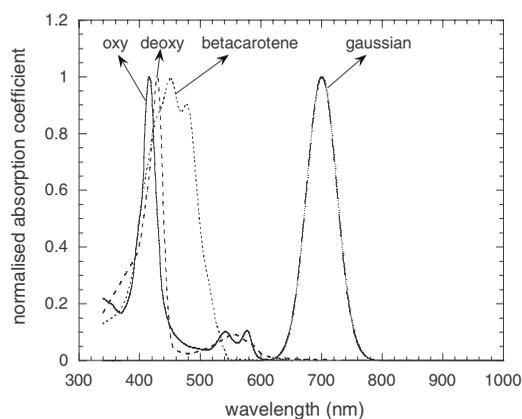


Fig. 2 Normalized absorption coefficients of the four absorbers used in the simulations.

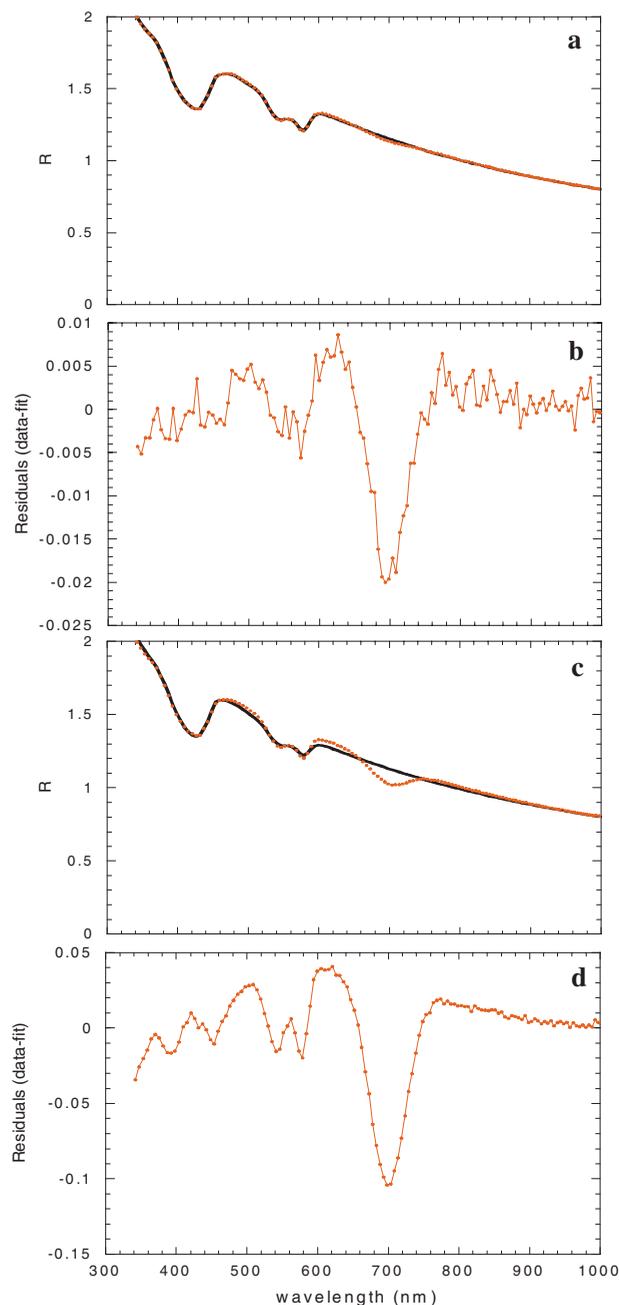


Fig. 3 (a) and (c) typical fit and (b) and (d) the residuals for the synthetic spectrum with $a_3=0.01$, $a_4=5$ and 30 , $a_5=2\%$, and $a_6=75\%$, fitted without the Gaussian absorber in the fitting model.

All fit values are unbiased and within 2 CIs of the true values. The residual spectrum shows only noise and is featureless.

In conclusion, when the correct model is used to fit the data, the fitted parameter values are unbiased, and the CIs are estimated correctly using the diagonal elements of estimated covariance matrix obtained by multiplying the inverse of the second derivative matrix of χ^2 with respect to its free parameters by χ^2/ν , with ν being the number of degrees of freedom.

4.2 Fitting with the Wrong Model: Missing Absorber

In this section, we investigate the magnitude of the bias in the fitted parameters introduced when the fitting model is missing

Table 5 Average number of times (out of 100) each of the fitted parameter values is within 2 CIs of their true values for different concentrations of missing Gaussian absorber.

	0	5	30
a_1	96 (4)	6 (21)	3 (16)
a_2	94 (3)	16 (36)	28 (44)
a_3	97 (4)	86 (31)	78 (40)
a_4			
a_5	95 (3)	2 (7)	3 (15)
a_6	95 (5)	80 (36)	71 (44)
a_7	95 (3)	1 (3)	1 (3)

an absorber. Two different situations are investigated: the absorption band of the missing absorber has 1. little or 2. significant spectral overlap with the other absorbers in the component fit. In Sec. 4.2.1 the missing absorber is an artificial Gaussian centered at 700 nm with a width of 25 nm. Such an absorber has only little overlap with the main absorption bands of oxyhemoglobin and deoxyhemoglobin. In Sec. 4.2.2 the missing absorber is betacarotene, which does overlap with the absorption bands of hemoglobin. Figure 2 shows the normalized absorption coefficients of the four absorbers used in the simulations. The noise amplitude is set to 0.01 and the bin width is 15 in this section.

4.2.1 Nonoverlapping Absorber: Gaussian Centered at 700 nm

An additional 210 synthetic datasets were generated using the Gaussian absorber instead of betacarotene in Eq. (2). Fitting was performed with only six free parameters by excluding $a_4 \cdot \mu_a^{\text{Bcar}}(\lambda)$ from Eq. (2). Figures 3(a) and 3(b) show a typical fit and the residuals, respectively, for the synthetic spectrum with $a_3=0.01$, $a_4=5$, $a_5=2\%$, and $a_6=75\%$. This “concentration” ($a_4=5$) of Gaussian absorber results in a hardly visible maximum signal decrease of 2.5% at $\lambda=700$ nm. The fitted values for the parameters and the CIs were $a_1=0.9977 \pm 0.0007$, $a_2=-0.992 \pm 0.004$, $a_3=0.0098 \pm 0.0005$, $a_5=1.88 \pm 0.03\%$, $a_6=74 \pm 2\%$, and $a_7=9.6 \pm 0.2 \mu\text{m}$.

Figures 3(c) and 3(d) show a typical fit and the residuals, respectively, for the synthetic spectrum with $a_3=0.01$, $a_4=30$, $a_5=2\%$, and $a_6=75\%$. This “concentration” ($a_4=30$) of Gaussian absorber results in a clearly visible maximum signal decrease of 14% at $\lambda=700$ nm. The fitted values for the parameters and the CIs were $a_1=0.982 \pm 0.004$, $a_2=-0.93 \pm 0.02$, $a_3=0.011 \pm 0.003$, $a_5=1.3 \pm 0.2\%$, $a_6=62 \pm 11\%$, and $a_7=7.4 \pm 0.8 \mu\text{m}$. Even though the statistical CIs of these fits are much larger than the statistical CIs of the parameters using the correct model (Sec. 4.1.2), only the Rayleigh amplitude (a_3) and saturation (a_6) fit values are within 2 CIs of the true values. The residual spectra very clearly show the features of hemoglobin absorption as well as the missing Gaussian absorber.

Table 6 Average bias and CI for each of the parameters for different concentrations of missing Gaussian absorber.

	Gauss absorber concentration	Average bias	Average CI
a_1 [-]	5	0.005	0.002
	30	0.03	0.008
a_2 [-]	5	0.02	0.007
	30	0.13	0.04
a_3 [-]	5	0.0016	0.0015
	30	0.009	0.007
a_4			
a_5 [%]	5	0.2	0.06
	30	1.0	0.3
a_6 [%]	5	1.3	1.5
	30	9	9
a_7 [μm]	5	0.9	0.3
	30	4.6	1.5

Table 5 shows the number of times each of the fitted parameter values is within 2 CIs of their true values, out of the 100 fits and averaged for the 105 synthetic datasets for different concentrations of Gaussian absorber.

Table 6 shows the average bias (\equiv absolute value of absolute difference between true and fitted parameter values) and CI for each of the parameters for different concentrations of Gaussian absorber. Blood volume fractions $a_5 < 1\%$ were excluded from the calculation of the averages due to the extremely large bias and CI of the blood parameters (a_5 , a_6 , and a_7) for small blood volume fractions (see for example Table 4). It is observed that, except for parameters Rayleigh amplitude (a_3) and saturation (a_6), the average bias in the parameters is much larger than twice the average statistical CI of the parameters.

4.2.2 Overlapping Absorber: Betacarotene

Similar to Sec. 4.2.1, fitting was performed with only six free parameters by excluding $a_4 \mu_a^{\text{Bcar}}(\lambda)$ from Eq. (2). Figures 4(a) and 4(b) show a typical fit and the residuals, respectively, for the synthetic spectrum with $a_3=0.01$, $a_4=5 \mu\text{M}$, $a_5=2\%$, and $a_6=75\%$. This concentration ($a_4=5 \mu\text{M}$) of betacarotene results in a hardly visible maximum signal decrease of 2.5% at $\lambda=450$ nm. The fitted values for the parameters and the CIs were $a_1=0.9987 \pm 0.0006$, $a_2=-0.984 \pm 0.003$, $a_3=0.0103 \pm 0.0004$, $a_5=2.02 \pm 0.03\%$, $a_6=91 \pm 1\%$, and $a_7=9.9 \pm 0.1 \mu\text{m}$. Even though the statistical CIs of this fit are larger than the statistical CIs of the parameters using the correct model (Sec. 4.1.2), only half of the fit values (a_3 , a_5 , and a_7) are within 2 CIs of the true values. The fitted saturation (a_6) is very biased with respect to the true value (91 versus 75%), while the CI is increased only marginally compared to the case where betacarotene was included in the fit. The residual spectrum clearly shows the features of hemoglobin absorption as well as betacarotene absorption.

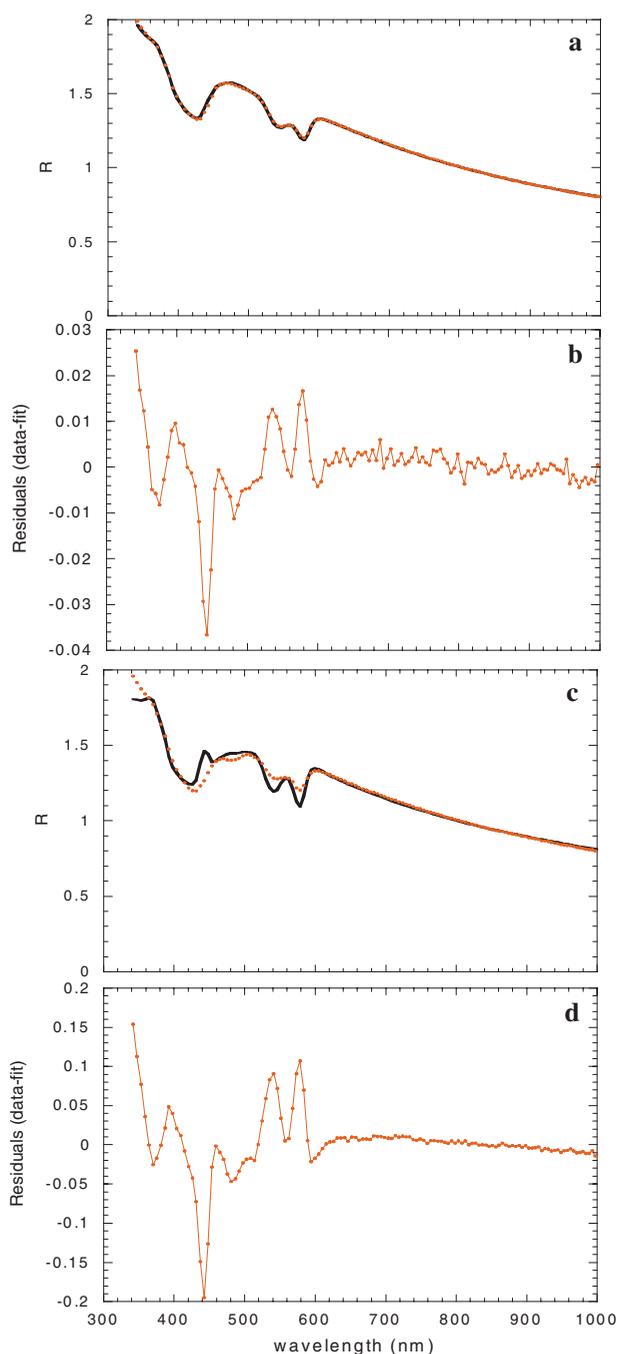


Fig. 4 (a) and (c) typical fit and (b) and (d) the residuals for the synthetic spectrum with $a_3=0.01$, $a_4=5$ and $30 \mu\text{M}$, $a_5=2\%$, and $a_6=75\%$, fitted without betacarotene in the fitting model.

Figures 4(c) and 4(d) show a typical fit and the residuals, respectively, for the synthetic spectrum with $a_3=0.01$, $a_4=30 \mu\text{M}$, $a_5=2\%$, and $a_6=75\%$. This concentration ($a_4=30 \mu\text{M}$) of betacarotene results in a clearly visible maximum signal decrease of 14% at $\lambda=450 \text{ nm}$. The fitted values for the parameters and the CIs were $a_1=0.995 \pm 0.003$, $a_2=-0.91 \pm 0.02$, $a_3=0.012 \pm 0.002$, $a_5=2.4 \pm 0.1\%$, $a_6=155 \pm 7\%$, and $a_7=10.7 \pm 0.6 \mu\text{m}$. Even though the statistical CIs of this fit are much larger than the statistical CIs of the parameters using the correct model (Sec. 4.1.2), only half of

Table 7 Average number of times (out of 100) each of the fitted parameter values is within 2 CIs of their true values for different concentrations of missing betacarotene.

	0	5	30
a_1	96 (4)	24 (37)	23 (40)
a_2	94 (3)	48 (44)	42 (44)
a_3	97 (4)	43 (42)	38 (44)
a_4			
a_5	95 (3)	58 (39)	46 (42)
a_6	95 (5)	10 (29)	22 (40)
a_7	95 (3)	45 (42)	39 (41)

the fit values (a_1 , a_3 , and a_7) are within 2 CIs of the true values. The fitted saturation (a_6) is very biased with respect to the true value (155 versus 75%), while the CI is increased only marginally compared to the case where betacarotene was included in the fit. Note that the fitted blood volume fraction a_5 has a smaller bias than expected from the figure; this is due to the fitted negative amount of deoxyhemoglobin (which also resulted in the physically impossible 155% saturation). The residual spectrum very clearly shows the features of hemoglobin absorption as well as betacarotene absorption.

Table 7 shows the number of times each of the fitted parameter values is within 2 CIs of their true values, out of the 100 fits and averaged for the 105 synthetic datasets for different concentrations of betacarotene.

Table 8 shows the average bias and CI for each of the parameters for different concentrations of betacarotene, for blood volume fractions $a_5 > 1\%$.

It is observed that the average bias in the parameters is larger than twice the average statistical CI of the parameters, especially for the saturation a_6 .

4.3 Fitting with the Wrong Model: Incorrect Scattering Model

In this section, we investigate the magnitude of the bias in the fitted parameters that is introduced when a wrong scattering function is used in the fitting routine. This is established by fitting with only six free parameters by excluding the Rayleigh scattering term $a_3(\lambda/\lambda_0)^{-4}$ in Eq. (1) from the fit function. Figures 5(a) and 5(b) show a typical fit and the residuals, respectively, for the synthetic spectrum with $a_3=0.01$, $a_4=5 \mu\text{M}$, $a_5=2\%$, and $a_6=75\%$. The fitted values for the parameters and the CIs were $a_1=1.0092 \pm 0.0009$, $a_2=-1.052 \pm 0.006$, $a_4=6.0 \pm 0.9 \mu\text{M}$, $a_5=1.75 \pm 0.05\%$, $a_6=73 \pm 4\%$, and $a_7=9.6 \pm 0.3 \mu\text{m}$.

Figures 5(c) and 5(d) show a typical fit and the residuals, respectively, for the synthetic spectrum with $a_3=0.1$, $a_4=5 \mu\text{M}$, $a_5=2\%$, and $a_6=75\%$. The fitted values for the parameters and the CIs were $a_1=1.105 \pm 0.003$, $a_2=-1.35 \pm 0.02$, $a_4=-10 \pm 3 \mu\text{M}$, $a_5=1.2 \pm 0.2\%$, $a_6=65 \pm 21\%$, and $a_7=10 \pm 2 \mu\text{m}$. Even though the statistical CIs of these fits are much larger than the statistical CIs of the parameters

Table 8 Average bias and CI for each of the parameters for different concentrations of missing betacarotene.

	Betacarotene concentration	Average bias	Average CI
a_1 [-]	5	0.003	0.001
	30	0.02	0.004
a_2 [-]	5	0.012	0.004
	30	0.07	0.02
a_3 [-]	5	0.003	0.0007
	30	0.02	0.003
a_4			
a_5 [%]	5	0.06	0.03
	30	0.38	0.14
a_6 [%]	5	8	1
	30	44	6
a_7 [μm]	5	0.5	0.2
	30	2.1	1.2

using the correct model (Sec. 4.1.2), only half the fit values are within 2 CIs of the true values. The blood volume fraction is particularly underestimated with respect to the true value. The residual spectrum very clearly shows the features of hemoglobin absorption as well as a strong peak at the start of the spectrum associated with an incorrect scattering function.

Table 9 shows the number of times each of the fitted parameter values is within 2 CIs of their true values, out of the 100 fits and averaged for 70 synthetic datasets (only the extreme betacarotene concentrations 0 and 30 μM were considered) for different Rayleigh amplitudes.

Table 10 shows the average bias and CI for each of the parameters for different Rayleigh amplitudes. Again, blood volume fractions $a_5 < 1\%$ were excluded from the calculation of the averages due to the extremely large bias and CI of the blood parameters (a_5 , a_6 , and a_7) for small blood volume fractions. For the largest Rayleigh amplitude, blood volume fractions $a_5 < 2\%$ were excluded from the calculation of the averages due to the large bias and CI of the blood parameters (a_5 , a_6 , and a_7) for these blood volume fractions as well.

It is observed that the scattering parameters (a_1 and a_2) as well as the absorber concentrations (a_4 and a_5) are very biased compared to the calculated statistical CIs, whereas the blood saturation and vessel diameter (a_6 and a_7 , respectively) are not affected as much by imperfections in the scattering function.

4.4 Identification and Minimization of Bias

In the previous sections we have shown that when the fitting model is correct, the fitted parameter values are unbiased and the CIs can be estimated correctly using the reduced chi-square normalized covariance matrix. When the fitting model is incorrect, however, the fitted parameters are biased and the calculated statistical CIs cannot correctly account for this bias. This is expected, since the systematic errors associated with using the wrong model are not fully accounted for by our

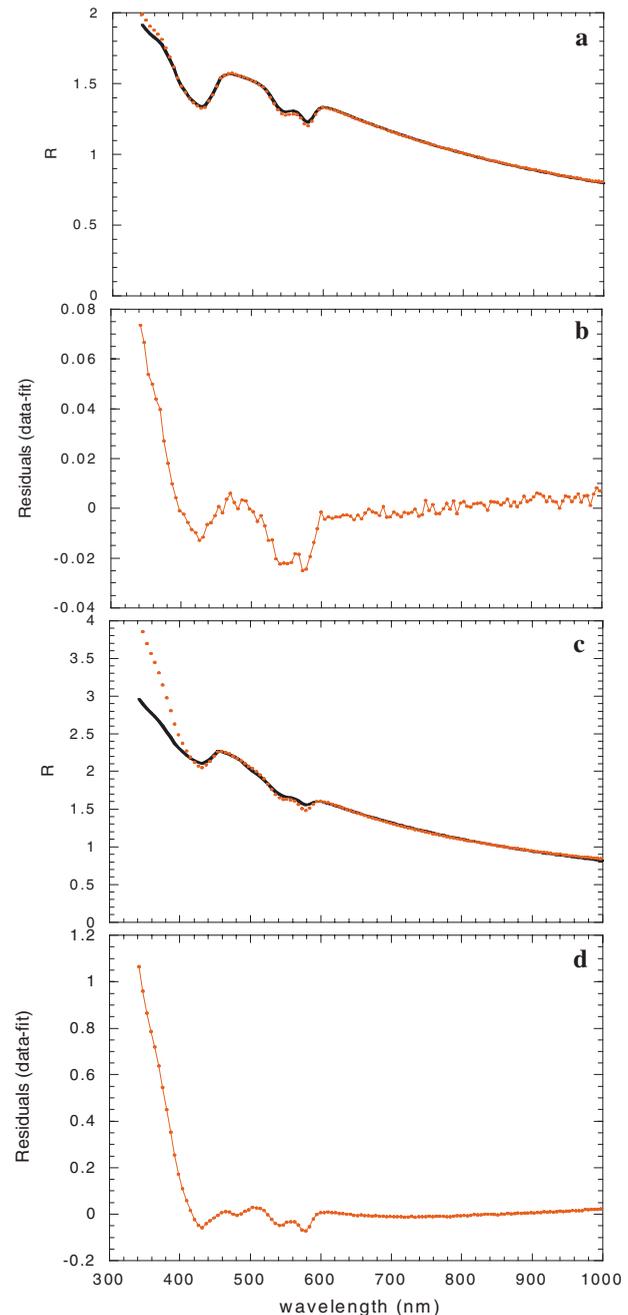


Fig. 5 (a) and (c) typical fit and (b) and (d) the residuals for the synthetic spectrum with $a_3=0.01$ and 0.1 , $a_4=5 \mu\text{M}$, $a_5=2\%$, and $a_6=75\%$, fitted without Rayleigh scattering in the fitting model.

purely statistical analysis. It is practically impossible to account for systematic errors using statistical methods only, and hence it should be realized that the calculated CIs represent the lower limits of the true CIs of the parameters. In this section we present a method to identify and minimize systematic errors in the fitting routine.

Identification of systematic errors is easily performed based on the shape of the residual spectrum. When the fitting model is correct, the residual spectrum is featureless and shows random noise only [Fig. 1(b)]. When the fitting model is incorrect, the residual spectrum will always show features

Table 9 Average number of times (out of 100) each of the fitted parameter values is within 2 CIs of their true values for different Rayleigh amplitudes with the Rayleigh term omitted from the fit function.

	0	0.01	0.1
a_1	91 (3)	0 (0)	0 (0)
a_2	91 (3)	0 (0)	0 (0)
a_3			
a_4	94 (3)	62 (43)	9 (27)
a_5	94 (3)	11 (25)	1 (4)
a_6	95 (3)	86 (25)	80 (34)
a_7	94 (3)	62 (42)	69 (43)

of the chromophore absorption coefficients, and in case the wrong scattering function is used, an additional (slow) background variation over the whole wavelength range can be observed [Figs. 3(b), 3(d), 5(b), and 5(d)]. The major difference in the residual spectra of fits with missing absorbers [Figs. 3(b), 3(d), 4(b), and 4(d)] and fits with an incorrect scattering function [Figs. 5(b) and 5(d)] is that in the first case, the residuals have a (strong) negative peak near the location of the absorption maximum of the missing absorber, while in the second case all peaks in the residuals are near the absorption maxima of the chromophores used in the fit. In Sec. 4.4.1, we discuss an approach to minimize the systematic errors associated with missing absorbers, and in Sec. 4.4.2, we discuss a method to minimize the systematic errors associated with an incorrect scattering function.

4.4.1 Minimization of Bias: Missing Absorber

A missing absorber in the model can be identified by a strong negative peak in the residuals (if defined as data minus fit) at a location away from absorption maxima of the chromophores used in the fit. After identification of a missing absorber, one can reduce the systematic error by refitting the spectrum, but now including an additional absorber that peaks near the wavelength of the negative peak in the residuals. Since a Gaussian function is a good first approximation for the shape of an absorption band, we propose to use a Gaussian function for this unidentified absorber, following Eq. (4):

$$\mu_a^{\text{unknown}}(\lambda) = a_8 \exp \left[-0.5 \left(\frac{\lambda - \lambda_{\text{peak}} - a_9}{a_{10}} \right)^2 \right]. \quad (4)$$

Here λ_{peak} is the location of the negative peak in the residual spectrum and a_8 , a_9 , and a_{10} are additional free parameters representing the “concentration,” the peak shift, and the width of the unknown absorber, respectively. The peak shift parameter a_9 accounts for the fact that the location of the negative peak in the residual spectrum is not necessarily the exact location of the absorption maximum of the absorber, but will be close to it. Therefore, this parameter should be constrained to small values to force the location of the absorption maximum of the Gaussian to be near the location of the negative peak in

Table 10 Average bias and CI for each of the parameters for different Rayleigh amplitudes.

	Rayleigh amplitude	Average bias	Average CI
a_1 [-]	0.01	0.0098	0.0006
	0.1	0.103	0.002
a_2 [-]	0.01	0.042	0.004
	0.1	0.30	0.02
a_3 [-]			
a_4 [μM]	0.01	2	1
	0.1	19	4
a_5 [%]	0.01	0.23	0.05
	0.1	1.4	0.2
a_6 [%]	0.01	3	3
	0.1	10	9
a_7 [μm]	0.01	0.7	0.4
	0.1	1.8	1.3

the residual spectrum. We have used $a_9 < 10$ nm in our analysis, but did not investigate in detail how the results depend on this constraint.

When this strategy is employed for the case of the missing absorbers in the fitting routine (Sec. 4.2), dramatic improvement in the fit quality is achieved. Table 11 shows the number of times each of the fitted parameter values is within 2 CIs of their true values, out of the 100 fits and averaged for the 105 synthetic datasets for different concentrations of Gaussian absorber. It is observed that on average, nearly 95% of the fitted parameter values is within 2 CIs of their true values, which indicates that the parameters and their CIs are estimated correctly.

The location of the peak λ_{peak} was determined by the minimum of the residual spectra, and the absorption maximum of the unknown absorber is $\lambda_{\text{peak}} - a_9$. For the small Gaussian absorber concentration $a_4 = 5$, the average fitted values of the missing absorber were $\lambda_{\text{peak}} - a_9 = 701$ nm and width $a_{10} = 24$ nm. For the larger Gaussian absorber concentration $a_4 = 30$, the average fitted values of the missing absorber were $\lambda_{\text{peak}} - a_9 = 699$ nm and width $a_{10} = 27$ nm. Thus in both cases the absorption maxima and widths of the unknown absorber reproduce the true absorption spectrum very accurately.

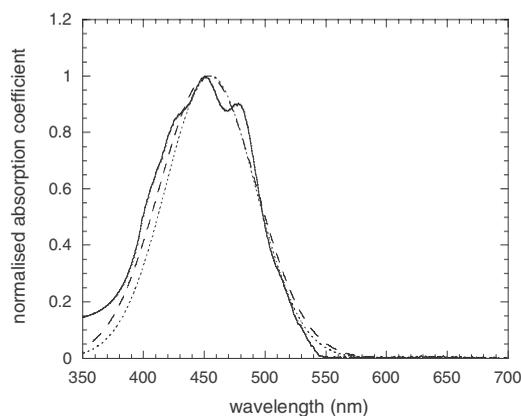
Table 12 shows the number of times each of the fitted parameter values is within 2 CIs of their true values, out of the 100 fits and averaged for the 105 synthetic datasets for different concentrations of betacarotene.

The location of the peak λ_{peak} was determined by the minimum of the residual spectra, and the absorption maximum of the unknown absorber is $\lambda_{\text{peak}} - a_9$. For the small betacarotene concentration $a_4 = 5 \mu\text{M}$, the average fitted values of the missing absorber were $\lambda_{\text{peak}} - a_9 = 453$ nm and width $a_{10} = 40$ nm. For the larger betacarotene concentration $a_4 = 30 \mu\text{M}$, the average fitted values of the missing absorber were $\lambda_{\text{peak}} - a_9 = 455$ nm and width $a_{10} = 37$ nm. These values

Table 11 Average number of times (out of 100) each of the fitted parameter values is within 2 CIs of their true values for different concentrations of Gaussian absorber with an additional absorber in the fit function.

	0	5	30
a_1	96 (4)	93 (4)	92 (6)
a_2	94 (3)	94 (6)	94 (7)
a_3	97 (4)	95 (8)	94 (10)
a_4			
a_5	95 (3)	91 (8)	89 (12)
a_6	95 (5)	91 (10)	89 (11)
a_7	95 (3)	93 (5)	93 (6)

were determined from the fits with blood volume fractions $a_5 < 10\%$ only; for larger blood volume fractions, the minimum of the residual spectra was collocated with the blood absorption peaks. As a result, no missing absorber could be identified for blood volume fractions $a_5 \geq 10\%$. Figure 6 shows the normalized betacarotene absorption spectrum and the average extracted absorption spectra of the unknown Gaussian absorber determined from the fits. Good agreement in both the peak position and the width is found between the true betacarotene absorption coefficient and the absorption coefficients of the unknown Gaussians determined from the fits. Despite these promising results, Table 12 shows that on average, much less than 95% of the fitted parameter values is within 2 CIs of their true values. This indicates that there are still systematic errors and that the resulting fit values are biased. The disappointing results are also due to the fact that the method fails to identify a missing absorber for large blood volume fractions. Table 13 shows the average bias (\equiv absolute value of the absolute difference between true and

**Fig. 6** Normalized betacarotene absorption spectrum and the average extracted absorption spectra of the unknown Gaussian absorber determined from the fits. Solid line: true betacarotene spectrum. Dotted line: Gaussian approximation from fits with large betacarotene concentration. Dashed line: Gaussian approximation from fits with small betacarotene concentration.**Table 12** Average number of times (out of 100) each of the fitted parameter values is within 2 CIs of their true values for different concentrations of betacarotene with an additional absorber in the fit function.

	0	5	30
a_1	96 (4)	72 (16)	39 (35)
a_2	94 (3)	53 (22)	11 (20)
a_3	97 (4)	38 (24)	9 (19)
a_4			
a_5	95 (3)	42 (21)	17 (23)
a_6	95 (5)	60 (33)	55 (33)
a_7	95 (3)	44 (26)	25 (32)

fitted parameter values) and CI for each of the parameters for different concentrations of betacarotene, with an unknown Gaussian absorber in the fit that has a maximum absorption at a wavelength no more than 10 nm away from the minimum of the residuals, as explained before. Blood volume fractions $a_5 < 1\%$ were excluded from the calculation of the averages, as discussed previously. It is observed that the average bias in the parameters is mostly larger than twice the average statistical CI of the parameters, but much smaller than before (Table 8) when no additional Gaussian absorber was included in the fit. In particular, the bias on the saturation is decreased by a large amount.

4.4.2 Minimization of Bias: Wrong Scattering Model

When a wrong scattering model is used, the residuals show features of the chromophore absorption coefficients on top of an additional background variation over the whole wavelength range. After identification of a wrong scattering model, one can reduce the systematic error by refitting the spectrum, but now including an additional scattering term. We propose to use a second-order polynomial function as an additional scattering term, since it provides a good first-order approximation to any nonperiodic continuous function, such as usually encountered in tissue scattering:

$$R(\lambda) = \left[a_1 \left(\frac{\lambda}{\lambda_0} \right)^{a_2} + a_3 \left(\frac{\lambda}{\lambda_0} \right)^2 + a_8 \left(\frac{\lambda}{\lambda_0} \right) + a_9 \right] \times \exp[-0.4\mu_a^{\text{total}}(\lambda)]. \quad (5)$$

Thus the scattering function is now modeled by Mie scattering plus a second-order polynomial. Whereas the Mie scattering model was based on physical principles where the parameters a_1 and a_2 are associated with the density and size distribution of scatterers,²¹ respectively, the new scattering model is primarily designed to describe the scattering function mathematically without associating physical meaning to the parameters. In other words, the scattering function is modified with the aim to reduce the bias in the concentration estimates of the chromophores, at the cost of modeling with a physically uninterpretable scattering function. Table 14 shows the

Table 13 Average bias and CI for each of the parameters for different concentrations of betacarotene, with an unknown Gaussian absorber in the fit, which has a maximum absorption at a wavelength no more than 10 nm away from the minimum of the residuals.

	Betacarotene concentration	Average bias	Average CI
α_1 [-]	5	0.002	0.008
	30	0.005	0.001
α_2 [-]	5	0.004	0.004
	30	0.011	0.005
α_3 [-]	5	0.001	0.001
	30	0.006	0.001
α_4			
α_5 [%]	5	0.04	0.02
	30	0.24	0.05
α_6 [%]	5	1.5	0.6
	30	3.5	1.5
α_7 [μm]	5	0.3	0.1
	30	1.6	0.3

number of times the fitted microvascular parameter values are within 2 CIs of their true values, out of the 100 fits and averaged for 35 synthetic datasets (betacarotene was excluded in this subanalysis for convenience) for different Rayleigh amplitudes. It is observed that on average, the in the case of a small Rayleigh component, nearly 95% of the fitted parameter values is within 2 CIs of their true values, which indicates that the parameters and their CIs are estimated correctly. For the larger Rayleigh amplitude, the accuracy of the results decreases but is still much better than before (Table 9) without the additional second-order polynomial scattering term. As for the Mie and Rayleigh scattering parameters, it should not be expected that they are accurately reproduced, since the mathematical scattering function used in the fit is very different from the scattering function of the synthetic datasets. Finally, Fig. 7(a) shows the synthetic scattering function

$$S_{\text{syn}}(\lambda) = 1 \left(\frac{\lambda}{\lambda_0} \right)^{-1} + 0.1 \left(\frac{\lambda}{\lambda_0} \right)^{-4},$$

and a fit of this function using the scattering function of Eq. (5). Figure 7(b) shows the residuals. Clearly the scattering function of Eq. (5) is a good but not perfect approximation for $S_{\text{syn}}(\lambda)$, leading to the results of Table 14.

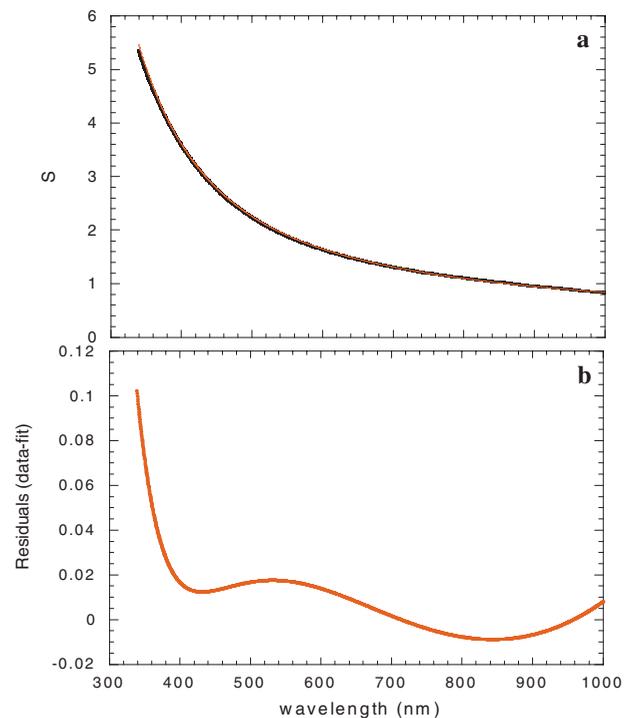
5 Discussion and Conclusion

We have validated a simple method for determining the confidence intervals on fitted parameters derived from modeling optical reflectance spectroscopic measurements using synthetic datasets. The method estimates the parameter confidence intervals as the square roots of the diagonal elements of the covariance matrix, which is obtained by multiplying the inverse of the second derivative matrix of χ^2 with respect to its free parameters by χ^2/v , with v being the degrees of free-

Table 14 Average number of times (out of 100) the fitted microvascular parameter values are within 2 CIs of their true values for different Rayleigh amplitudes fitted with an additional second order polynomial scattering function.

	0	0.01	0.1
α_5	96 (3)	92 (5)	81 (20)
α_6	95 (3)	92 (3)	80 (10)
α_7	95 (3)	89 (6)	75 (24)

dom. Most commercial fitting software packages, employing least squares analysis, automatically generate the covariance matrix. CIs are easily extracted after multiplication of the diagonal elements with χ^2/v to account for the uncertainty in the estimation of the statistical fluctuations and corresponding standard deviations $s(i)$ of the data points. We have shown in Sec. 4.1 that this method yields correct CIs for our nonlinear fitting function, as long as the model used to describe the data is correct. The statistical CIs do not depend on the number of pixels used to bin the data, as long as the optical resolution associated with this number does not exceed the optical resolution of the biological features. Furthermore, the CI of the fitted parameters correctly increases with decreasing signal-to-noise ratio, while in all cases the fitted parameter values are unbiased with respect to their true values. The CI calculated in this way represents the statistical error on the fit parameters and is a lower boundary of the true fit parameter uncertainty. We have also investigated the sensitivity of our parameters to systematic errors associated with incorrect models in results

**Fig. 7** (a) The synthetic scattering function $S_{\text{syn}}(\lambda)$ and a fit of this function using the scattering function of Eq. (5), and (b) the residuals.

Secs. 4.2 and 4.3. Even when there are only small amounts of missing absorber (Sec. 4.2), such that the fit does not deviate from the data by more than 2 to 3% over the entire wavelength range [Figs. 3(a), 3(b), 4(a), and 4(b)], the fitted parameters may be biased by large amounts compared to the calculated statistical CIs. Figures 4(a) and 4(b) are particularly instructive: the quality of the fit looks excellent, the residuals are smaller than 2.5% over the entire wavelength range, and yet the bias in the saturation is more than 15%. Thus imperfections in the fitting model introduce a bias in the fitted parameters that greatly exceeds the estimated statistical CIs. Since multicomponent spectral fitting of *in vivo* optical spectroscopic data is particularly sensitive to these problems, we have also introduced methods to identify and subsequently minimize the bias in the fitted parameters associated with incorrect modeling (Sec. 4.4). Identification of systematic errors is based on the shape of the residual spectrum. When the fitting model is correct, the residual spectrum is featureless and shows random noise only [Fig. 1(b)]. When the fitting model is incorrect, the residual spectrum always shows features of the included chromophore absorption coefficients, and in case the wrong scattering function is used, an additional background variation over the whole wavelength range can be observed [Figs. 3(b), 3(d), 5(b), and 5(d)]. In the case of a missing absorber, a strong negative peak in the residuals at a location away from absorption maxima of the chromophores used in the fit is observed. The systematic error associated with a missing absorber was reduced by refitting the spectrum, but now including an additional absorber that peaks near the wavelength of the minimum (maximum negative) in the residuals. Dramatic improvement in the fit quality was achieved when using a Gaussian absorber with three free parameters: amplitude, peak shift with respect to minimum of residuals, and peak width. In the case where the missing absorber truly was a Gaussian, the fits were perfect, and nearly 95% of the fitted parameter values was within 2 CIs of their true values, which indicates that the fitted parameters and their CIs were estimated correctly. Furthermore, the absorption maxima and widths of the unknown Gaussian absorber reproduced the true Gaussian absorption spectrum very accurately. In the case where the missing absorber was betacarotene, which has a non-Gaussian absorption spectrum overlapping the absorption spectra of oxy- and deoxyhemoglobin, the results were not as good. For large blood volume fractions, the minimum of the residual spectra was collocated with the blood absorption peaks, and no missing absorber could be identified for large ($a_5 \geq 10\%$) blood volume fractions. For smaller blood volume fractions, the extracted absorption spectra of the unknown Gaussian absorber determined from the fits showed good agreement in both the peak position and the width with the true betacarotene absorption coefficient. Despite this good agreement, much less than 95% of the fitted parameter values was within 2 CIs of their true values for both high and low betacarotene concentrations. This indicates that there are still systematic errors and that the resulting fit values are biased. However, the bias in the parameters was shown to be much smaller than without an additional Gaussian absorber included in the fit; the bias on the saturation was decreased by a large amount (from 8 to 1.5% for small betacarotene concentration, and from 44 to 3.5% for large betacarotene concentration).

Palmer et al.²² have previously used a similar approach in a clinical dataset of malignant and nonmalignant breast tissue. They included a Gaussian absorber with fixed width and peak position in their fits based on the shape of their previously obtained residuals in a large clinical dataset, under the assumption that the spectral shape of the missing absorber is the same across the whole set of tissue samples measured. Interestingly, they found that although the quality of the fits was improved with the addition of the Gaussian absorber, the conclusions regarding the extracted absorber and scattering parameters from the fits were not significantly affected by its addition, and so all subsequent analysis was carried out without the inclusion of this Gaussian absorber. Based on our present research, this result is surprising. Possibly this statement was made with respect to the average of the fitted parameters over the entire dataset. In that case, the biases in the fitted parameters for the individual spectra may be canceled on averaging. Furthermore, the biological variation in the parameters for multiple samples is probably larger than the bias associated with the systematic error introduced by the incomplete fitting model, as we describe next.

The systematic error associated with a wrong scattering model was reduced by refitting the spectrum, but now including a second-order polynomial as an additional scattering term. The new scattering model is primarily designed to describe the scattering function mathematically without associating physical meaning to the scattering parameters, with the aim to reduce the bias in the concentration estimates of the chromophores. The implementation of this polynomial scattering function greatly reduced the bias in the microvascular parameters, although the approach has its limits for large deviations in the model scattering function from the true scattering function. The mathematical description of the "true" scattering function in case of biological (*in vivo*) tissue actually depends on the measurement geometry and on the tissue itself. In the case of diffuse measurements, the power law behavior of scattering seems to be an adequate description of the scattering function. In the case of nondiffuse measurements, more complicated functions (including periodic elements²³⁻²⁵) may be necessary to fully describe the scattering function.

Other potential sources of systematic error that are not investigated in detail in this work are: 1. incorrect wavelength calibration of the spectrometer, and 2. incorrect or unknown specific absorption coefficients of the chromophores. Regarding the first point, we observed that a 1-nm offset in the spectrometer calibration typically resulted in a 3% bias in the fitted saturation a_6 (data not shown). Since it should always be possible to calibrate the wavelength axis of a spectrometer to within 1 nm, we consider this source of error to be minor. Regarding the second point, the specific absorption coefficients of chromophores are usually measured when dissolved in water, ethanol, hexane, or chloroform. Apart from water, these environments are dramatically different from the *in vivo* environment, and it is not easy to predict how the environment changes the specific absorption coefficient of the chromophore. Spectral shifting as well as spectral broadening may occur, which would also result in biased fit parameters, the amount of bias depending on the amounts of shifting and broadening. Therefore, it is imperative to use specific chromophore absorption coefficients in the fitting routine that

were measured in an environment that comes as close to the *in vivo* environment as possible. This is particularly important for component fits of *in vivo* fluorescence and Raman spectra, whose shapes can be altered even more dramatically than extinction spectra due changes in the microenvironment. Alternatively, any features in the residual spectra obtained from fitting *in vivo* spectra using *ex vivo* component spectra could be used to modify these component spectra and determine the true *in vivo* fluorescence, Raman, or extinction spectra. Moreover, environmental changes (such as pH variations) that are reflected in altered spectral shapes of the components will lead to features in the residual spectra. This provides an opportunity to characterize and monitor the microenvironment by analyzing the shape of the residual spectra.

As can be seen in Table 4, for a noise amplitude typical for our *in vivo* data, the CI of the fitted parameters is quite small (e.g., the CI for blood saturation a_6 is smaller than 1% for blood volume fractions $a_5 > 1\%$). This CI represents the statistical error associated with an individual spectrum. It is critical that this is interpreted differently from a standard deviation calculated from averaging multiple measurements. The latter standard deviation represents the biological variation in the parameters, which for most applications will be larger than the statistical uncertainty in the parameters related to the quality of the data. In fact, in our experience the biological variation in the parameters for multiple measurements is even larger than the bias associated with the systematic errors discussed in results Secs. 4.2 and 4.3, and for most applications it will be sufficient to specify these biological standard deviations only. However, completely disregarding the statistical CIs can also cause problems. For example, if the blood volume fraction is very low, the statistical uncertainty of the blood related fit values becomes very large (Table 4). Therefore, in the case of averaging multiple measurements, we feel that calculating the weighted average of the parameters, with the statistical CIs as weight factors, is more appropriate than calculating the unweighted average and ignoring the statistical CIs. Furthermore, calculating the statistical errors in the fit parameters facilitates objective assessment of the quality of spectra, and fits with the possibility to reject poor quality spectra or fit values from an *in vivo* dataset, thereby avoiding unnecessary pollution of the dataset by poor quality spectra.

Table 8 shows that a missing absorber that overlaps with the spectra of oxy- and deoxyhemoglobin can cause a large bias in the extracted saturation. This is particularly true for small concentrations of blood, where the calculated statistical uncertainty of the saturation is much smaller than the bias in the saturation introduced by a missing spectrally overlapping absorber (data not shown). Therefore, to be safe, we have in previous publications restricted the calculations of average blood saturation to spectra with blood volume fractions larger than 1%.^{11–14} It is important to note that the saturation CI was calculated for our specific DPS measurement geometry with 400- μm fibers, featuring a path length of 0.4 mm. Since the saturation CI will depend on signal attenuation rather than on blood volume fraction, the saturation CI (and potential bias) will be smaller for similar blood volume fractions in the case of DPS with larger fiber diameters, or any other measurement geometry with a longer path length.

In this work, the parameters were not constrained, i.e., all free parameters were allowed to range from minus infinity to

plus infinity. However, in reality the parameters are physically constrained to $a_1 \geq 0$, $0 \leq a_2 \leq -4$, $a_3 \geq 0$, $a_4 \geq 0 \mu\text{M}$, $a_5 \geq 0\%$, $0\% \leq a_6 \leq 100\%$, and $a_7 \geq 5 \mu\text{m}$ (the diameter of a red blood cell). We did not implement these boundary conditions here to fully explore the sources and magnitudes of the statistical and systematic errors. When fitting an *in vivo* spectrum, however, it would be appropriate to constrain the parameters by implementing these physical boundary conditions to exclude nonphysical parameter values. However, in that case, care should be taken that the final value of any parameter is not at one of these artificially imposed limits. Boundary conditions on the individual parameters may also be artificially imposed by transformations of the variables. For example, by making the transformation $a'_6 \rightarrow 1/2 + [\arctan(a_6)]/\pi$, the transformed parameter a'_6 is restricted to $0 < a'_6 < 1$ for $-\infty < a_6 < \infty$. In this way the fit is performed without boundary conditions on the fit parameters while still imposing physical boundary conditions on the transformed parameters. If the transformation of any fit parameter a is given by $a' \rightarrow f(a)$, the statistical error of the transformed parameter is given by $\sigma' = \sigma[df(a)/da]$.²⁰ For our example, $a'_6 \rightarrow 1/2 + [\arctan(a_6)]/\pi$ we find $\sigma'_{a'_6} = \sigma_{a_6} / \pi(1+a_6^2)$, where σ_{a_6} is the square root of the sixth diagonal element of the reduced chi-square normalized covariance matrix obtained from the fit.

In conclusion, we show that we can correctly estimate the statistical error on parameters extracted from fits to reflectance spectroscopic data. We demonstrate the use of various methods to identify and subsequently minimize the bias in the fitted parameters associated with systematic errors. Detailed analysis of the residual spectrum is potentially useful not only for minimization of bias, but also for monitoring changes in the *in vivo* microenvironment.

Acknowledgment

This research has been supported by the Dutch Technology Foundation STW, applied science division of NWO, and the Technology Program of the Ministry of Economic Affairs.

References

- O. R. Stepanovic, K. L. Bechtel, A. S. Haka, W. C. Shih, T. W. Koo, A. J. Berger, and M. S. Feld, "Determination of uncertainty in parameters extracted from single spectroscopic measurements," *J. Biomed. Opt.* **12**, 064012 (2007).
- T. J. Farrell, M. S. Patterson, and B. C. Wilson, "A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the non-invasive determination of tissue optical properties," *Med. Phys.* **19**, 879–888 (1992).
- R. M. P. Doornbos, R. Lang, M. C. Aalders, F. W. Cross, and H. J. C. M. Sterenborg, "The determination of *in vivo* human tissue optical properties and absolute chromophore concentrations using spatially resolved steady-state diffuse reflectance spectroscopy," *Phys. Med. Biol.* **44**, 967–981 (1999).
- G. Zonios, L. T. Perelman, V. Backman, R. Manoharan, M. Fitzmaurice, J. Van-Dam, and M. S. Feld, "Diffuse reflectance spectroscopy of human adenomatous colon polyps *in vivo*," *Appl. Opt.* **38**, 6628–6637 (1999).
- N. Ghosh, S. K. Mohanty, S. K. Majumder, and P. K. Gupta, "Measurement of optical transport properties of normal and malignant human breast tissue," *Appl. Opt.* **40**, 176–184 (2001).
- J. C. Finlay and T. H. Foster, "Hemoglobin oxygen saturations in phantoms and *in vivo* from measurements of steady-state diffuse reflectance at a single, short source-detector separation," *Med. Phys.* **31**, 1949–1959 (2004).
- P. Thueller, I. Charvet, F. Bevilacqua, M. St. Ghislain, G. Ory, P.

- Marquet, P. Meda, B. Vermeulen, and C. Depeursinge, "In vivo endoscopic tissue diagnostics based on spectroscopic absorption, scattering, and phase function properties," *J. Biomed. Opt.* **8**(4), 495–503 (2003).
8. G. M. Palmer and N. Ramanujam, "Monte Carlo-based inverse model for calculating tissue optical properties. Part I: theory and validation on synthetic phantoms," *Appl. Opt.* **45**, 1062–1071 (2006).
 9. T. J. Pfefer, L. S. Matchette, C. L. Bennett, J. A. Gall, J. N. Wilke, A. J. Durkin, and M. N. Ediger, "Reflectance-based determination of optical properties in highly attenuating tissue," *J. Biomed. Opt.* **8**(2), 206–215 (2003).
 10. R. Reif, O. A' Amar, and I. J. Bigio, "Analytical model of light reflectance for extraction of the optical properties in small volumes of turbid media," *Appl. Opt.* **46**, 7317–7328 (2007).
 11. A. Amelink, M. P. Bard, S. A. Burgers, and H. J. Sterenborg, "In vivo measurement of the local optical properties of tissue by use of differential path-length spectroscopy," *Opt. Lett.* **29**, 1087–1089 (2004).
 12. M. P. Bard, A. Amelink, V. Noordhoek Hegt, W. J. Graveland, H. J. Sterenborg, H. C. Hoogsteden, and J. G. Aerts, "Measurement of hypoxia-related parameters in bronchial mucosa by use of optical spectroscopy," *Am. J. Respir. Crit. Care Med.* **171**, 1178–1184 (2005).
 13. R. L. van Veen, A. Amelink, M. Menke-Pluymers, C. van der Pol, and H. J. Sterenborg, "Optical biopsy of breast tissue using differential path-length spectroscopy," *Phys. Med. Biol.* **50**, 2573–2581 (2005).
 14. A. Amelink, O. P. Kaspers, H. J. Sterenborg, J. E. van der Wal, J. L. Roodenburg, and M. J. Witjes, "Non-invasive measurement of the morphology and physiology of oral mucosa by use of optical spectroscopy," *Oral Oncol.* **44**, 65–71 (2008).
 15. A. Amelink and H. J. C. M. Sterenborg, "Measurement of the local optical properties of turbid media using differential pathlength spectroscopy," *Appl. Opt.* **43**, 3048–3054 (2004).
 16. O. P. Kaspers, H. J. C. M. Sterenborg, and A. Amelink, "Controlling the optical path length in turbid media using differential path-length spectroscopy: fiber diameter dependence," *Appl. Opt.* **47**, 365–371 (2008).
 17. S. W. van de Poll, "Raman spectroscopy of atherosclerosis," Ph.D. Thesis Univ. of Leiden, p. 123, ISBN 90-9016109-0 (2003).
 18. S. Prahl, see <http://omlc.ogi.edu/spectra/hemoglobin> (1999).
 19. F. James, *Minuit—function minimization and error analysis*, Reference Manual Version 94.1, Cern (1998).
 20. P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York (1969).
 21. J. R. Mourant, T. Fuselier, J. Boyer, T. M. Johnson, and I. J. Bigio, "Prediction and measurements of scattering and absorption over broad wavelength ranges in tissue phantoms," *Appl. Opt.* **36**, 949–957 (1997).
 22. G. M. Palmer, C. Zhu, T. M. Breslin, F. Xu, K. W. Gilchrist, and N. Ramanujam, "Monte Carlo-based inverse model for calculating tissue optical properties. Part II: Application to breast cancer diagnosis," *Appl. Opt.* **45**, 1072–1078 (2006).
 23. L. T. Perelman, V. Backman, M. Wallace, G. Zonios, R. Manoharan, A. Nusrat, S. Shields, M. Seiler, C. Lima, T. Hamano, I. Itzkan, J. Van Dam, J. M. Crawford, and M. S. Feld, "Observation of periodic fine structure in reflectance from biological tissue: a new technique for measuring nuclear size distribution," *Phys. Rev. Lett.* **80**, 627–630 (1998).
 24. M. B. Wallace, L. T. Perelman, V. Backman, J. M. Crawford, M. Fitzmaurice, M. Seiler, K. Badizadegan, S. J. Shields, I. Itzkan, R. R. Dasari, J. Van Dam, and M. S. Feld, "Endoscopic detection of dysplasia in patients with barrett's esophagus using light-scattering spectroscopy," *Gastroenterology* **119**, 677–682 (2000).
 25. V. Backman, R. Gurjar, K. Badizadegan, I. Itzkan, R. R. Dasari, L. T. Perelman, and M. S. Feld, "Polarized light scattering spectroscopy for quantitative measurement of epithelial cellular structures *in situ*," *IEEE J. Sel. Top. Quantum Electron.* **5**, 1019–1026 (1999).