Perfusion Quantification using Gaussian Process Deconvolution

I.K. Andersen[1,2], A. Szymkowiak[1], C. Rasmussen[3], L.G. Hanson[2], J.R. Marstrand[2], H.B.W. Larsson[4], L.K. Hansen[1]

[1]Informatics and Mathematical Modelling, Technical University of Denmark,
[2]Danish Research Centre for Magnetic Resonance, Hvidovre Hospital, Denmark
[3]Gatsby Computational Neuroscience Unit, University College, London, UK
[4]MR Centre, Medical Section, University Hospital (Sst. Olavs Hospital, HF), Trondheim, Norway

Running head: Perfusion Quantification using GP Deconvolution

Correspondence to:

Irene Klærke Andersen
Danish Research Centre for Magnetic Resonance
Hvidovre Hospital, Afd. 340B
Kettegård Allé 30
DK-2650 Hvidovre
Denmark

Phone: +45 3632 2978
Faxno: +45 3647 0302
Email: ireneka@magnet.drcmr.dk

Abstract

The quantification of perfusion using dynamic susceptibility contrast MR imaging requires deconvolution to obtain the residual impulse-response function (IRF). Here, a method using a Gaussian process for deconvolution, GPD, is proposed. The fact that the IRF is smooth is incorporated as a constraint in the method. The GPD method, which automatically estimates the noise level in each voxel, has the advantage that model parameters are optimized automatically. The GPD is compared to singular value decomposition (SVD) using a common threshold for the singular values and to SVD using a threshold optimized according to the noise level in each voxel. The comparison is carried out using artificial data as well as using data from healthy volunteers. It is shown that GPD is comparable to SVD with variable optimized threshold when determining the maximum of the IRF, which is directly related to the perfusion. GPD provides a better estimate of the entire IRF. As the signal to noise ratio increases or the time resolution of the measurements increases, GPD is shown to be superior to SVD. This is also found for large distribution volumes.

key words: perfusion quantification, Gaussian process deconvolution, singular value decomposition, dynamic susceptibility contrast.

Introduction

Perfusion, the capillary blood flow, may be measured using dynamic susceptibility contrast MR imaging (DSC-MRI) (1, 2). The first perfusion measurements using DSC-MRI were performed by Villringer et al. (3) and Boxerman et al. (4). Until recently, many studies used only semiquantitative approaches (5, 6), but inter-subject comparisons and patient follow up require quantitative approaches. To obtain this, the tissue concentration of contrast agent should be deconvolved with the concentration in the supplying vessels in order to obtain the response to an ideal input. This response is termed the residual impulse response function (IRF), the maximum of which is directly related to the perfusion. Quantification using deconvolution by means of gamma-variate functions was proposed by Rempp et al. (7). The IRF may also be modelled by an exponential (8), but in cases where the exponential assumption is not valid, the IRF may be better estimated by deconvolution using a method where no specific shape of the IRF is assumed such as singular value decomposition (SVD), as shown by Østergaard et al. (9). This method involves a choice of a threshold for signal variance. A low threshold leads to a noisy estimation of the IRF, whereas a high threshold introduces systematic errors to the perfusion estimate. Østergaard et al. have developed software for SVD deconvolution that has been used by other groups (10–12). This approach uses a common threshold in all voxels set at an optimized value, found by Monte Carlo simulations (9), of 20% of the maximum singular value (13). However, Liu et al. (14) showed that the optimal threshold depends on the signal to noise ratio (SNR). They suggest a voxel-wise threshold determined by the contrast SNR, which leads to significant improvements in the perfusion determination. The threshold was calculated from artificial data using exponential IRFs and a mean transit time (MTT) of blood through tissue of three seconds. If the IRF is not an exponential or the MTT changes, the optimal threshold may change. Vonken et al. (15) estimated the IRF in a voxel-wise manner by optimizing the likelihood of the data using an expectation maximization algorithm. The voxel-wise optimization clearly improved perfusion estimation, however difficulties in determining the optimal number of iterations were reported.

The Gaussian process for deconvolution, GPD, presented here, estimates the IRF as the mean value of an optimized joint Gaussian distribution. The method is also likelihood based but extended to include the smoothness of the IRF, which is incorporated as *a priori* information. This is achieved by initializing the joint Gaussian distribution by a special distribution called a Gaussian process (GP), which is a joint Gaussian distribution with an infinite input space (16). The correlation length of the IRF is a parameter of the joint Gaussian distribution and is constructed initially from the delay between the sampling times only. In this way, the smoothness of the IRF is incorporated as a regularizing term. This is necessary for improving the IRF estimation when the SNR is limited. The GPD automatically estimates the noise level as well as all other parameters of the joint Gaussian distribution. Furthermore, the method supplies the standard deviation of the estimated IRF in each voxel. Whilst GP for regression has been used in geostatistics, where it is known as Kriging (17), the statistical framework has only been developed (16) recently. In this work, GP is used for deconvolution.

Theory

The MR signal intensity, $S(t)$, during passage of a bolus of contrast agent is related to the tissue concentration of the contrast agent, $C_{\mathrm{t}}(t)$, as:

$$S(t) = S_0 \exp(-kC_{\mathrm{t}}(t)T_{\mathrm{E}}) + \epsilon(t), \qquad [1]$$

where S_0 is the prebolus signal; k is a constant; T_{E} is the echo time; and ϵ is noise.

The tissue concentration as a function of time, $C_{\mathrm{t}}(t)$, depends on the concentration of contrast agent in the feeding vessels, $C_{\mathrm{a}}(t)$, the regional cerebral blood volume, $rCBV$, and the perfusion, f. The relationship can be expressed as

$$rCBV = \frac{\int_0^\infty C_{\mathrm{t}}(\tau)d\tau}{\int_0^\infty C_{\mathrm{a}}(\tau)d\tau}, \qquad [2]$$

where, under the usual assumption of negligible inter-voxel diffusion of contrast agent,

$$C_{\mathrm{t}}(t) = \kappa f C_{\mathrm{a}}(t) \otimes \tilde{h}(t), \qquad [3]$$

where $\otimes$ denotes convolution, κ is the ratio of 1-hematocrit in capillaries to 1-hematocrit in arteries (taken to be 1.2 in correspondance with (9)), and $\tilde{h}$ is the IRF. Since the MR signal reflects the residual amount of contrast agent, the tissue is characterized by a residual IRF that describes the distribution of particles following different paths through the tissue. The IRF is smooth, since the contrast agent mixes with blood. If the mixing is instantaneous for the whole blood compartment, then the IRF is an exponential function. Even if the mixing is less instantaneous, the contrast agent is subject to bulk motion, and consequently the IRF resembles a Fermi function[1]. The IRF equals the tissue signal when the input curve is a delta function and is by definition scaled to a maximum of one at time zero. The function $h(t) = f \cdot \tilde{h}(t)$ is the kernel of the convolution. The maximum of the kernel is the perfusion. The area under the kernel gives the rCBV as seen by combining Eqs. [2] and [3]:

$$rCBV = \int_0^\infty f \cdot \tilde{h}(\tau)d\tau = f \cdot MTT, \qquad [4]$$

[1] Fermi function: $[1 + \exp(-ab)]/(1 + \exp[a(x-b)])$

where MTT is the mean transit time through the tissue.

If the subject has a deficient blood brain barrier, the above equations are still valid except that the distribution volume of the contrast agent, described by the variable $rCBV$, is no longer restricted to the cerebral vascular volume. In this situation, the IRF is no longer mono-exponential, but may be biexponential or even more complex. However, it is still a smooth function with a maximum value of unity, and hence the maximum of the kernel is still the perfusion.

For computational convenience, Eq. [3] can be rewritten as a matrix product

$$\mathbf{c}_{\mathrm{t}} = \mathbf{C}_{\mathrm{a}}\mathbf{h} + \epsilon, \qquad [5]$$

where $\mathbf{c}_{\mathrm{t}}$ is the vector of sampled tissue concentrations $[C_{\mathrm{t}}(t^{(1)}), .., C_{\mathrm{t}}(t^{(N)})]^T$, and $\mathbf{h}$ is given by $[h(0), .., h(t_N)]^T$, where N is the number of points on the curves and T denotes the transposed vector. $\mathbf{C}_{\mathrm{a}}$ is the matrix given as

$$\mathbf{C}_{\mathrm{a}} = \begin{bmatrix} C_{\mathrm{a}}(t^{(1)}) & 0 & \dots & 0 \\ C_{\mathrm{a}}(t^{(2)}) & C_{\mathrm{a}}(t^{(1)}) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ C_{\mathrm{a}}(t^{(N)}) & C_{\mathrm{a}}(t^{(N-1)}) & \dots & C_{\mathrm{a}}(t^{(1)}) \end{bmatrix}. \qquad [6]$$

If some rows of $\mathbf{C}_{\mathrm{a}}$ are close to being linear combinations of each other, the matrix is nearly singular and inversion hereof may not be performed directly. Two methods for deconvolution to recover $\mathbf{h}$ from Eq. [5] are presented below.

Singular Value Decomposition

One way to recover $\mathbf{h}$ is to minimize the squared error $E(\mathbf{h}) = \frac{1}{2}(\mathbf{c}_{\mathrm{t}} - \mathbf{C}_{\mathrm{a}}\mathbf{h})^T(\mathbf{c}_{\mathrm{t}} - \mathbf{C}_{\mathrm{a}}\mathbf{h})$. This leads to

$$\mathbf{h} = (\mathbf{C}_{\mathrm{a}}\mathbf{C}_{\mathrm{a}}^T)^{-1}\mathbf{C}_{\mathrm{a}}^T\mathbf{c}_t. \qquad [7]$$

If the product $\mathbf{C}_{\mathrm{a}}\mathbf{C}_{\mathrm{a}}^T$ is close to singular, the problem has more solutions. The optimal solution giving the least squared error is found using the singular value decomposition method (SVD) (18). SVD decomposes $\mathbf{C}_{\mathrm{a}}$ into two orthogonal matrices, $\mathbf{U}$ and $\mathbf{V}^T$ and a diagonal matrix, $\mathbf{W}$, with singular values ordered descending in the diagonal, $\mathbf{C}_{\mathrm{a}} = \mathbf{U}\ \mathbf{W}\ \mathbf{V}^T$. The solution is then given as:

$$\mathbf{h} = \mathbf{V}\ \widetilde{\mathbf{W}}^{-1}\ \mathbf{U}^T\mathbf{c}_{\mathrm{t}}, \qquad [8]$$

where $\widetilde{\mathbf{W}}^{-1}$ is the inverse of $\mathbf{W}$, with some elements of $\widetilde{\mathbf{W}}^{-1}$ set to zero if the singular value is below a certain threshold. The threshold is set depending on the data and the noise levels (19). This approach may appear to lead to a less accurate solution, but the discarded equations are corrupted by numerical instabilities and should be removed in order to obtain a robust solution. Choosing a threshold that excludes more than the nearly singular components will lead to regularization of the solution where components with high frequencies are excluded because they are thought to originate from biological and instrumental noise.

Gaussian Process Deconvolution

This approach assumes that each measured value of the kernel is distributed normally around the true value as indicated in figure 1. Since there is some correlation between points on the curve, the data points belong to one joint Gaussian distribution. When the underlying kernel is expected to be smooth, Gaussian process deconvolution (GPD) may be used to calculate a predictive distribution for any point in time. The estimate of the kernel at the time point is the mean of the predictive distribution.

In the perfusion experiment, $S(t)$, is the measured signal, from which the tissue curve $\mathbf{c}_{\mathrm{t}}$ is calculated (Eq. [1]). The signal noise associated with $S(t)$ (with standard deviation σ_{S}) is approximately Gaussian and since the noise level is small compared to the signal, so is the noise on $\mathbf{c}_{\mathrm{t}}$ (with standard deviation σ_{C}). The measured kernel, $\mathbf{h}_0$, may be calculated from the contrast concentrations, Eq. [7], and belongs to the joint Gaussian data distribution (see appendix A)

$$\mathbf{h}_0 \in \mathcal{N}(\mathbf{h}, \mathbf{A}), \qquad [9]$$

where $\mathbf{h}$ is the vector of means of the joint normal distribution, and $\mathbf{A}$ is the covariance matrix that describes the covariance or correlation between points on the kernel. It is given as $\mathbf{A} = \sigma_{\mathrm{C}}^2(\mathbf{C}_{\mathrm{a}}^T\mathbf{C}_{\mathrm{a}})^{-1}$.

The mean vector and covariance matrix that optimize the likelihood, Eq. [9], is the least squares solution. In perfusion imaging, each data point is only measured once and with noisy data, this estimate is poor. With GPD, the expected smoothness of the IRF is used to regularize the solution through a prior distribution.

Nothing is assumed about the values of $\mathbf{h}$ prior to measurements, however nearby points on the $h(t)$ curve are expected to correlate. The joint distribution prior to measurements may thus be described by a Gaussian Process that is a joint distribution with an infinite input space, in this case time(16). The distribution may be constructed so that the covariance matrix is a function of the sampling times only:

$$\mathbf{h}_{\text{prior}} \in \mathcal{N}(0, \mathbf{B}), \tag{10}$$

where the covariance matrix, as in (20), is chosen to be:

$$\mathbf{B}_{(p,q)} = v \exp\left(-\frac{1}{2} w (t^{(p)} - t^{(q)})^2\right), \tag{11}$$

where $t^{(p)}$ and $t^{(q)}$ are sampling times and w and v are scale parameters. If $p = q$ then $B_{(p,p)} = v$, which is the variance. As the distance between the sampling points increases, the covariance or correlation, $B_{(p,q)}$, of the points decreases. In this way, nearby points have a larger influence on an estimate than more distant points. The scale parameters, v, w, are also called hyper parameters, since they control $\mathbf{B}$, and hence the prior distribution. v is the bias or off-set parameter that determines the influence of the prior distribution. w controls the correlation length.

A new joint distribution of the data may be found by multiplying the distribution of data, Eq. [9], and the prior distribution, Eq. [10], and integrate the product over $\mathbf{h}$ (see appendix A). This leads to $\mathbf{h}_0 \in \mathcal{N}(0, \mathbf{D})$, where $\mathbf{D} = \mathbf{A} + \mathbf{B}$. This distribution is the likelihood for the parameters of $\mathbf{D}$ (i.e. v, w, and σ_{C}^2) given the measured kernel $\mathbf{h}_0$. The parameters are estimated by optimizing this likelihood (see *Methods*). The influence of the prior distribution is determined by the distance between the sampling times $t^{(p)}$ and $t^{(q)}$, whereas the influence of the data distribution is dependent on the noise level through σ_{C}^2.

From the likelihood, the predictive distribution at any point, t_{p}, may be calculated. The mean of this distribution is the estimate of h_{p} and the variance of the estimate is the diagonal elements of the covariance matrix. The distribution is derived in appendix B:

$$h_{\text{p}} \in \mathcal{N}(\mathbf{a}^T\mathbf{D}^{-1}\mathbf{h}_0, b - \mathbf{a}^T\mathbf{D}^{-1}\mathbf{a}), \qquad [12]$$

where $\mathbf{a}$ is the covariance between the new input and the old points, $\mathbf{a} = \mathbf{B}(t^{(1)}..t^{(N)}, t_{\text{p}})$, and b is the variance of the new input, $b = \mathbf{D}(t_{\text{p}}, t_{\text{p}})$. The mean and variance of the distribution are calculated by simple matrix multiplications. Note that the predictive value may be found for any input t_{p}, but that $\mathbf{D}$ only has to be inverted once.

In summary, a regularized expression for $\mathbf{h}$ is obtained, by only assuming that $\mathbf{h}$ is distributed normally with a correlation between points. The introduced hyper parameters are optimized automatically for each voxel (see *Methods*). This means that no threshold has to be set using GPD. Moreover, a covariance matrix of the estimate is an inherent result of the method (see Eq. [12]), where the diagonal elements are the variance of the estimated kernel in each point.

Methods

The MR experiments

DSC-MRI data were obtained from nine healthy volunteers using a 1.5 T Siemens Vision scanner and a gradient echo sequence with echo planar imaging (TR/TE 1000/66 ms, flip angle 60°, field of view 250 mm, matrix 128×128, slice thickness 5 mm, 128 acq.). Gd-

DTPA (0.1 mmol/kg body weight. Magnevist, Schering) was injected in an antecubital vein with an MR-compatible double syringe power injector (Medrad Inc., Spectris™ MR Injector) starting at the 15th measurement with an injection rate of 3 ml/s, followed by 20 ml saline at 3 ml/s. The study was approved by the Central Scientific-Ethical Committee of Denmark (c-199-08;KF 01-357/98) and informed consent was obtained from the volunteers.

Concentration curves were calculated using Eq. [1]. In healty subjects, the concentration in the supplying vessels, $C_{\mathrm{a}}(t)$, is well represented by the artery concentration. Thus $C_{\mathrm{a}}(t)$ was chosen in the slice from the voxel with the earliest and largest signal peak in the insular area to minimize the transport delay from the input voxel to the tissue voxel.

Simulations

Artificial kernels and data were created to investigate the performance of the deconvolution methods. Four kernels were chosen: an exponential function, a gamma-variate function, a box, and a triangle. The exponential kernel corresponds to instantaneous mixing of the contrast agent in the blood volume. The gamma function would result if the transit of the input function from the input voxel to the tissue voxel is non-negligible (21, 22). The box- and the triangle functions were constructed as in (9) in order to test the ability of the methods to extract non-exponential features.

All kernels were constructed with maxima from f=0.002 to 0.01 ml/g/s, where f=0.003 ml/g/s corresponds to the white matter (WM) level in healthy adults, and f=0.01 ml/g/s corresponds to the grey matter (GM) level. The area under the curves (Eq. [4]) was set to *rCBV*=3%, 5%, 8% and 60% corresponding to typical values for WM, GM, slightly elevated capillary blood volume, and a deficient blood brain barrier. The arterial input function, $C_{\mathrm{a}}(t)$, was calculated from the measured arterial signal (Eq. [1]) with the constant, k, selected to insure a signal drop of the artificial tissue curves similar to the observed mean signal drop in GM (14).

The input concentration curve was convolved with the artificial kernels to generate tissue output curves, $\mathbf{c}_{\mathrm{t}}$. The signal curves were generated from Eq. [1], with T_{E} and S_0 equal to those of an MR experiment, so that kernels with f=0.01 ml/g/s were given the mean S_0 value of GM (S_0 = 480 MR signal units (MRu)) and kernels with f=0.003 ml/g/s were given the mean S_0 value of WM (S_0 = 440 MRu). The S_0 for other perfusion values were calculated by linear inter- or extrapolation of these values.

The standard deviation of the signal due to noise before contrast agent administration was also measured and σ_{S}=15, 22, and 30 MRu were found to be in the lower range, the middle range, and the upper range. Gaussian noise with these standard deviations was added to the constructed signal curves. To test the performance of the methods close to ideal conditions, noise was also added with σ_{S}=4 MRu. The noisy tissue-concentration curves were then constructed using Eq. [1]. The measured input curve and the noisy output curves were deconvolved using each method to create estimates of the kernels.

SVD

Two different thresholds were used with the SVD: In the A-SVD method, singular values above 20% of the maximum were kept (13). In the B-SVD method, the threshold was calculated from the signal to noise contrast ratio, SNR_{C}, as proposed by Liu et al. (14). The SNR_{C} is given by

$$SNR_{\mathrm{C}} = SNR_{\mathrm{i}} \frac{S_{\mathrm{min}}}{S_0} \ln\left(\frac{S_0}{S_{\mathrm{min}}}\right), \qquad [13]$$

where S_{min} is the minimum value of the tissue curve and $SNR_{\mathrm{i}} = S_0/\sigma$. Liu et al. (14) performed Monte Carlo simulations, using exponential kernels with MTT=3 s, SNR_{i} between 10 and 100, and f from zero to five times the GM value, to estimate the optimal threshold corresponding to each SNR_{C}. This provided a set of empiric equations describing the relationship between SNR_{C} and the optimal threshold. These equations were used in the comparison.

GPD

In the Gaussian process, the hyper parameters, v, w, and σ_{C}^2 of Eq. [11] were optimized by minimizing the logarithm of the likelihood function, $\mathcal{L}$ given by Eq. [A.4]. This was achieved using the conjugate gradient method with the Polak-Ribiere minimization direction (18). The gradient of the log-likelihood is

$$\mathbf{g}_i = \frac{\partial \log(\mathcal{L})}{\partial \theta_i} = \frac{1}{2}\mathrm{trace}\left(\mathbf{D}^{-1}\frac{\partial \mathbf{D}}{\partial \theta_i}\right) - \frac{1}{2}\mathbf{h}_0^T\mathbf{D}^{-1}\frac{\partial \mathbf{D}}{\partial \theta_i}\mathbf{D}^{-1}\mathbf{h}_0, \qquad [14]$$

where $\partial\mathbf{D}/\partial\theta_i$ is the partial derivative of the covariance function with respect to each of the parameters: $\theta = [v, w, \sigma_{\mathrm{C}}^2]$. Before using the conjugate gradient algorithm, the hyper parameters were initialized: w was set to 0.1; v was set to the variance of $\mathbf{h}_0$; and σ_{C}^2 was set to the variance of the tissue curve, $\mathbf{c}_{\mathrm{t}}$, before the bolus. The algorithm was then used iteratively until a relative precision of 10^{-6} was achieved. The same minimum of the cost function was found regardless of the initialization for large SNR, but for very noisy data, local minima in the cost function occurred.

Determination of the gradient requires a calculation of $\mathbf{h}_0$, which again demands an inversion of the matrix product, $\mathbf{C}_{\mathrm{a}}^T\mathbf{C}_{\mathrm{a}}$. In order to ensure that $\mathbf{C}_{\mathrm{a}}^T\mathbf{C}_{\mathrm{a}}$ is positive definite, the number of columns of $\mathbf{C}_\mathrm{a}$ is reduced by the number of singular values.

Comparisons

The abilities of the GPD and SVD methods to reproduce the kernel maximum (perfusion) for the noise levels, perfusion values, and blood volumes described previously, were compared. For each setting, the estimation of the kernel was repeated 50 times with different random noise to give the average perfusion or RMSE and the 95% confidence interval, CI_{95}=mean $\pm SER$, where SER=1.96·standard deviation is called the standard error in the following. The shapes of the curves were investigated, both by visual inspection of individual curves, and by comparison of the root of the mean squared error (RMSE) between estimated kernels and the original kernel. The influence of an increased sampling rate was also investigated. This required a higher temporal resolution of the input concentration curve. It was thus modelled by the sum of five gamma variate functions. Gaussian noise at the level of a typical measured input function was added.

Results

Simulations

Examples of deconvolved exponential kernels with $rCBV$=5% and f=0.01 ml/g/s are shown in figure 2 for the a) low, b) middle and c) high noise levels found in the MR data. The true kernel is plotted with the thick black line. GPD (thin line) performed increasingly well as noise was reduced and provided no artificial undulations. The A-SVD estimate (dark gray) was independent of the noise level. The fixed threshold resulted in a severe underestimation of the maximum regardless of the noise level. B-SVD (light gray) reproduced maximum well for the low noise level, however the curve shape was poorly determined. When more noise was added, the estimated curve shape became smoother but the maximum was underestimated.

Figure 2 indicates the performance of the methods in one voxel, i.e. without averaging, whereas the mean of the maxima from 50 deconvolved kernels is shown as a function of the true maximum in figure 3 for each of the four kernel types. In all situations, $rCBV$ was at the GM level (5%). Figure 3a shows that GPD performed better at the low noise level (σ_S=15 MRu) than at the high noise level (figure 3b, (σ_S=30 MRu)). For the exponential kernel and for the triangle, GPD underestimated high perfusion. It performed very well, when the kernel was a gamma function or generally when perfusion was low. Figures 3c and 3d show that A-SVD estimated the same perfusion value regardless of the noise level. A-SVD also performed better for the gamma function than for the exponential and triangle kernels. Figures 3e and 3f show that B-SVD performed well for the exponential kernel at the low noise level, but overestimated the perfusion at this $rCBV$ for the other kernels. At low perfusion, the method overestimated perfusion. B-SVD was the method with the largest variations of the estimate at the low noise level.

The kernels estimated by all methods had the correct areas under the curves, and the same variations of them (results not shown).

The RMSEs of the estimated kernels are shown in figure 4 as function of the true maximum. The RMSEs for GPD were very small at low perfusion. When perfusion or noise increased, so did the RMSEs. When noise was increased, A-SVD provided higher RMSEs. The RMSE seemed to increase with perfusion, but there was an off-set RMSE even at very low perfusion. The B-SVD RMSEs were large and varied considerably.

The results described above were for $rCBV$ set at the GM level. The result of varying *rCBV* are shown in figure 5. At the average noise level, σ_S=22 MRu, the estimated maximum (mean of 20 repetitions) as a function of the true maximum is shown for the gamma kernel at four different *rCBV* values (3%, 5%, 8% and 60%). The figure reveals that as *rCBV* increased, the estimate performed by GPD improved, whereas the estimate of the other methods increased above the true value.

The results of increasing the sampling rate (*resol*) from 1 to 5 Hz are shown in table 1 for the exponential kernel at σ_S=22 MRu, $rCBV$=5%, and f=0.01 ml/g/s. Under these conditions, B-SVD reproduced the maximum of the kernel well when *resol*=1 Hz, while GPD performed rather poorly for the exponential kernel. However, table 1 shows that GPD performed increasingly well as resolution was increased. A-SVD was unaffected by the resolution increase, but the performance of B-SVD clearly worsened as resolution increased.

Volunteer Studies

The maximum of the kernel in each voxel estimated from the *in vivo* experiment is shown in figure 6a for the GPD method. In this perfusion image, GM, WM, basal ganglia, and CSF are distinguished easily. Figure 6b shows the mean of the standard error of the kernel for each voxel. As expected, the errors were high in CSF and in voxels along the edge of the brain.

In general, DSC-MRI perfusion measurements provide values that are too high. This is due to partial volume effects, different relaxivity values in tissue and vessels, and under-sampling of the input function. Whilst various solutions have been suggested to overcome some of these problems, they were not used in this study, since they are not yet well established and are not necessary for the comparison of deconvolution methods.

The maxima found *in vivo* by A-SVD, B-SVD and GPD were almost identical. However, the average GM and WM perfusion values were calculated for the nine subjects using all methods. The values estimated with A-SVD and B-SVD are plotted against the GPD estimates in figure 7. It is evident that B-SVD always gives a larger estimate than GPD. For high perfusion values, GPD gives larger estimates than A-SVD, whereas A-SVD gives larger estimates of low perfusion.

The areas under the estimated kernels were identical (images not shown).

Discussion

The GPD performs excellently for the artificial kernels with very low noise (data not shown). So does B-SVD, but A-SVD gives the same (wrong) estimate regardless of the noise level. The kernel steepness increases with perfusion for fixed rCBV, leading A-SVD to be more likely to underestimate high perfusion than low perfusion, due to the low-pass filtering. It overestimates low perfusion if the kernel is smooth at the maximum like the gamma function. This effect is seen in the MR data where A-SVD provides lower estimates than GPD in GM where perfusion is high, but slightly higher estimates than GPD in WM where perfusion is low.

The B-SVD threshold was initially optimized (14) from exponential kernels with constant MTT, meaning that kernels with high $rCBV$ had high f (Eq. [4]). Under these conditions B-SVD is significantly better than A-SVD in determining the maximum. However, due to severe fluctuations of the estimated kernel the whole curve is often better estimated using A-SVD (RMSE is smaller). The maximum of kernels more flat than the exponential is overestimated by B-SVD since the kernel estimates are fluctuating. Overestimation also occurs if rCBV increases for constant perfusion, since this also leads to more flat kernels. This, together with the tendency for B-SVD to overestimate low perfusion, leads B-SVD to give higher estimates of the perfusion than GPD as observed in *in vivo* experiments.

Increasing the sampling rate reduced the performance of B-SVD markedly. The reasons are two fold: The S_{min} estimate decreases with improved resolution, leading to a decrease in SNR_{c} (Eq. [13]), which results in a higher threshold and therefore an increased low-pass filtering.

Moreover, an increased sampling rate allows for higher frequencies to be measured. This increases the number of singular values, but reduces the values representing higher frequencies. The same percentage threshold thus includes fewer singular values which results in increased low-pass filtering. The A-SVD method is not affected by this, since the low frequencies represented by the singular values with a threshold above 20% are measured accurately at a sampling rate of 1 Hz. To increase the performance of B-SVD at high sampling rates, new equations for the threshold have to be computed.

The B-SVD method involves an estimation of the optimal SVD threshold, but it is not evident that a threshold that separates noise from the kernel exists. Setting a threshold removes high frequency variations that are thought to originate from biological and instrumental noise. However, the bolus peak also has high frequency components and is truncated. Moreover, the SNR used to determine the optimal threshold depends on the patient morphology and movement; factors that should not influence the perfusion estimate systematically.

GPD performs increasingly well as the noise level is reduced and the influence of the prior distribution is reduced. This is not necessarily the case for the other methods. In noisy data, when the prior distribution is required, the performance depends on the smoothness of the kernel. The smoothness increases with increased rCBV for fixed perfusion, which leads to improved performance of the GPD method. In patients with enlarged distribution volume, of the presented methods only GPD can provide a valid estimate.

The steepness of the kernel increases with perfusion for fixed rCBV, leading to reduced performance at high perfusion. This is one reason why GPD gave a lower estimate of maximum than B-SVD in the *in vivo* experiments. However, the data shows that at high perfusion, the whole kernel is better determined with GPD (RMSE is smaller). The derivative of the kernel is the distribution of transit times. It can be used to obtain metabolic information(23), which makes precise estimation of the whole kernel very important. At low perfusion, the GPD estimate was very good.

The smoothness of all kernels is improved with increased sampling rate. This suggests that acquisitions performed using high field scanners ($>$1.5T), where the SNR and/or the time resolution is increased, provide optimal conditions for GPD that benefit more than the SVD methods.

GPD performed better for the gamma function than for the more steep exponential function. The constructed gamma kernel would result from the convolution of an exponential IRF and a transport function(21, 22). Even though the desired IRF is not measured directly, the deconvolution should be done accurately in order to facilitate subsequent corrections for the transport.

The GPD method provides an estimate of the uncertainty of the calculated IRF per voxel automatically. An uncertainty estimate is not supplied by SVD.

It should be noted that GPD is computational demanding. The computation time for one brain slice (approx. 4000 voxels) was 2 s with A-SVD, 39 s with B-SVD, and 46 min with GPD (on a 800MHz Pentium III computer, 512 Mb RAM, with Matlab for Linux).

The computational time for GPD can be reduced significantly using parallel computing. Calculations in limited regions could be utilised to minimize computation time in acute stroke patients. More measurements before contrast administration could provide an accurate estimate of the noise level, so that σ_S^2 could be fixed during optimization halving the computation time.

The parameter related to the correlation length, w, was optimized to one distinct value in each voxel. The correlation length is, however, expected to vary along the kernel. Including this effect may lead to improvements of the GPD method, since the estimate will become less rigid at the kernel maximum. This is a better approach than to smooth the measured signals, since the IRF is known to be smooth for physiological reasons. Low pass filtering of the measured signals will also decrease the sensitivity to IRF variations. Other warranted constraints like the IRF being a decreasing function of time may also be incorporated in the prior distribution. The suggested improvements of the GPD prior distribution may result in further improvement, but even with the prior distribution and the time resolution used here, the method provided more accurate results.

Conclusions

The Gaussian process deconvolution, GPD, is a statistically sound way of estimating the residual impulse-response function (IRF) in each voxel. Since DRC-MRI is SNR-limited, it is crucial to make use of *a priori* knowledge when available. GPD has the smoothness of the IRF incorporated as a constraint. The GPD method estimated the maximum of the IRF, the perfusion, as accurately as the SVD methods. The GPD method provided a more accurate estimate of the complete IRF, which is highly important in studies of metabolism. When rCBV increased, only GPD managed to provide a valid perfusion estimate. Furthermore, GPD gives an estimate of the noise level of data, and GPD automatically provides a measure of the uncertainty on the IRF estimate. Using high performance scanners, the SNR and/or the time resolution is improved and GPD automatically benefits from these improvements. Therefore, GPD has important advantages over the optimized SVD, as a method for deconvolution of the residual impulse-response curve in dynamic susceptibility contrast perfusion measurements.

Acknowledgements

This work was supported by the Danish Research Councils through the projects "THOR Center for Neuroinformatics" and "SITE Signal and Image Processing for Telemedicine".

Appendix A

The tissue concentrations, $\mathbf{c}_\mathrm{t}$, are distributed normally with the likelihood function (18):

$$p(\mathbf{c}_\mathrm{t}|\mathbf{c}_\mathrm{a},\mathbf{h},\sigma_\mathrm{c}^2) = \left(\frac{1}{\sqrt{2\pi\sigma_\mathrm{c}^2}}\right)^N \exp\left[-\frac{1}{2\sigma_\mathrm{c}^2}(\mathbf{c}_\mathrm{t}-\mathbf{C}_\mathrm{a}\mathbf{h})^T(\mathbf{c}_\mathrm{t}-\mathbf{C}_\mathrm{a}\mathbf{h})\right], \qquad [\mathrm{A.1}]$$

where σ_c^2 is the noise variance, and N is the length of $\mathbf{c}_\mathrm{t}$. Multiplying the parentheses $(\mathbf{c}_\mathrm{t}-\mathbf{C}_\mathrm{a}\mathbf{h})^T(\mathbf{c}_\mathrm{t}-\mathbf{C}_\mathrm{a}\mathbf{h}) = \mathbf{c}_\mathrm{t}^T\mathbf{c}_\mathrm{t} - 2\mathbf{c}_\mathrm{t}^T\mathbf{C}_\mathrm{a}\mathbf{h} + \mathbf{h}^T\mathbf{C}_\mathrm{a}^T\mathbf{C}_\mathrm{a}\mathbf{h}$ and defining the matrix $\mathbf{A}$ by $\mathbf{A}^{-1} = \frac{1}{\sigma_\mathrm{c}^2}\mathbf{C}_\mathrm{a}^T\mathbf{C}_\mathrm{a}$, and the vector $\mathbf{h}_0$ by $\mathbf{h}_0^T\mathbf{A}^{-1} = \frac{1}{\sigma_\mathrm{c}^2}\mathbf{c}_\mathrm{t}^T\mathbf{C}_\mathrm{a}$ leads to

$$p(\mathbf{c}_\mathrm{t}|\mathbf{c}_\mathrm{a},\mathbf{h},\sigma_\mathrm{c}^2) = \left(\frac{1}{\sqrt{2\pi\sigma_\mathrm{c}^2}}\right)^N \exp\left[-\frac{1}{2}(\mathbf{h}_0^T\mathbf{A}^{-1}\mathbf{h}_0 - 2\mathbf{h}_0^T\mathbf{A}^{-1}\mathbf{h} + \mathbf{h}^T\mathbf{A}^{-1}\mathbf{h} + \frac{1}{\sigma_\mathrm{c}^2}\mathbf{c}_\mathrm{t}^T\mathbf{c}_\mathrm{t} - \mathbf{h}_0^T\mathbf{A}^{-1}\mathbf{h}_0)\right],$$

where the term $\mathbf{h}_0^T\mathbf{A}^{-1}\mathbf{h}_0$ is added and subtracted. The first three terms are recognized as a Gaussian quadratic form in $\mathbf{h}$ and the last two terms are independent of $\mathbf{h}$ and form the constant K:

$$p(\mathbf{c}_\mathrm{t}|\mathbf{c}_\mathrm{a},\mathbf{h},\sigma_\mathrm{c}^2) = K\exp\left[-\frac{1}{2}(\mathbf{h}_0-\mathbf{h})^T\mathbf{A}^{-1}(\mathbf{h}_0-\mathbf{h})\right].$$

Transformation of variables to a distribution of $\mathbf{h}_0$ leads to

$$p(\mathbf{h}_0|\mathbf{h},\mathbf{A}) = \frac{1}{\sqrt{|2\pi\mathbf{A}|}}\exp\left[-\frac{1}{2}(\mathbf{h}_0-\mathbf{h})^T\mathbf{A}^{-1}(\mathbf{h}_0-\mathbf{h})\right], \qquad [\mathrm{A.2}]$$

where $\mathbf{h}_0 = (\mathbf{C}_\mathrm{a}^T\mathbf{C}_\mathrm{a})^{-1}\mathbf{C}_\mathrm{a}^T\mathbf{c}_\mathrm{t}$ is the observed data, $\mathbf{A}^{-1} = 1/\sigma_\mathrm{c}^2\ \mathbf{C}_\mathrm{a}^T\mathbf{C}_\mathrm{a}$ is the covariance matrix, and $\mathbf{h}$ is the mean vector.

The equation for the Gaussian process prior distribution, with $\mathbf{B}$ given by Eq. [11], is

$$p(\mathbf{h}|\mathbf{B}) = \frac{1}{\sqrt{|2\pi\mathbf{B}|}} \exp\left(-\frac{1}{2}\mathbf{h}^T\mathbf{B}^{-1}\mathbf{h}\right). \qquad [A.3]$$

The joint Gaussian likelihood is calculated from the data distribution, Eq. [A.2], and the prior distribution, Eq. [A.3], by marginalization of $\mathbf{h}$ i.e. $p(\mathbf{h}_0|\mathbf{A},\mathbf{B}) = \int_{-\infty}^{\infty} p(\mathbf{h}_0|\mathbf{h},\mathbf{A})p(\mathbf{h}|\mathbf{B})d\mathbf{h}$. This leads to

$$p(\mathbf{h}_0|\mathbf{A},\mathbf{B}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{|2\pi\mathbf{A}|}} \frac{1}{\sqrt{|2\pi\mathbf{B}|}} \exp\left[-\left(\frac{1}{2}\mathbf{h}^T(\mathbf{A}^{-1}+\mathbf{B}^{-1})\mathbf{h} - \mathbf{h}_0^T\mathbf{A}^{-1}\mathbf{h}\right) - \frac{1}{2}\mathbf{h}_0^T\mathbf{A}^{-1}\mathbf{h}_0\right] d\mathbf{h}.$$

To proceed, the relation: $\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - q^T\mathbf{x} = \frac{1}{2}(\mathbf{x}-\mathbf{Q}^{-1}q)^T\mathbf{Q}(\mathbf{x}-\mathbf{Q}^{-1}q) - \frac{1}{2}q^T\mathbf{Q}^{-1}q$ is used, where $\mathbf{Q} = \mathbf{A}^{-1}+\mathbf{B}^{-1}$, $q = \mathbf{A}^{-1}\mathbf{h}_0$, and $\mathbf{x} = \mathbf{h}$. This leads to

$$\begin{aligned} & p(\mathbf{h}_0|\mathbf{A},\mathbf{B}) = \frac{1}{(2\pi)^N} \frac{1}{\sqrt{|\mathbf{A}||\mathbf{B}|}} \\ \cdot & \int_{-\infty}^{\infty} \exp[-\frac{1}{2}(\mathbf{h} - (\mathbf{A}^{-1}+\mathbf{B}^{-1})^{-1}(\mathbf{h}_0^T\mathbf{A}^{-1})^T)^T(\mathbf{A}^{-1}+\mathbf{B}^{-1})(\mathbf{h} - (\mathbf{A}^{-1}+\mathbf{B}^{-1})^{-1}(\mathbf{h}_0^T\mathbf{A}^{-1})^T) \\ + & \frac{1}{2}(\mathbf{h}_0^T\mathbf{A}^{-1})(\mathbf{A}^{-1}+\mathbf{B}^{-1})^{-1}(\mathbf{h}_0^T\mathbf{A}^{-1})^T - \frac{1}{2}\mathbf{h}_0^T\mathbf{A}^{-1}\mathbf{h}_0]d\mathbf{h}. \end{aligned}$$

The last two terms are independent of $\mathbf{h}$. Integration of the Gaussian distribution gives

$$\begin{aligned} p(\mathbf{h}_0|\mathbf{A},\mathbf{B}) &= \frac{\sqrt{|2\pi(\mathbf{A}^{-1}+\mathbf{B}^{-1})^{-1}|}}{(2\pi)^N\sqrt{|\mathbf{A}||\mathbf{B}|}} \exp\left[-\frac{1}{2}\left(-(\mathbf{h}_0^T\mathbf{A}^{-1})(\mathbf{A}^{-1}+\mathbf{B}^{-1})^{-1}(\mathbf{h}_0^T\mathbf{A}^{-1})^T + \mathbf{h}_0^T\mathbf{A}^{-1}\mathbf{h}_0\right)\right] \\ &= \frac{\sqrt{|(\mathbf{A}^{-1}+\mathbf{B}^{-1})^{-1}|}}{(2\pi)^{\frac{N}{2}}\sqrt{|\mathbf{A}||\mathbf{B}|}} \exp\left[-\frac{1}{2}(\mathbf{h}_0^T(\mathbf{A}^{-1} - \mathbf{A}^{-1}(\mathbf{A}^{-1}+\mathbf{B}^{-1})^{-1}\mathbf{A}^{-1})\mathbf{h}_0)\right] \end{aligned}$$

using matrix rules and defining $\mathbf{D}$ as $\mathbf{D} = \mathbf{A}+\mathbf{B}$ leads to the likelihood function:

$$p(\mathbf{h}_0|\mathbf{D}) = \frac{1}{\sqrt{|2\pi\mathbf{D}|}} \exp\left[-\frac{1}{2}\mathbf{h}_0^T\mathbf{D}^{-1}\mathbf{h}_0\right] = \mathcal{L}. \qquad [A.4]$$

Appendix B

Inserting the distributions into in Eq. [12] in the text leads to

$$p(h_{\mathrm{p}}|\mathbf{h}_0,\mathbf{D}) = \frac{\left|2\pi\begin{bmatrix}\mathbf{D} & \mathbf{a}^T\\ \mathbf{a} & b\end{bmatrix}\right|^{\frac{1}{2}}}{\sqrt{|2\pi\mathbf{D}|}}\exp\left(-\frac{1}{2}\left(\begin{bmatrix}\mathbf{h}_0\\ h_{\mathrm{p}}\end{bmatrix}^T\begin{bmatrix}\mathbf{D} & \mathbf{a}^T\\ \mathbf{a} & b\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{h}_0\\ h_{\mathrm{p}}\end{bmatrix} - \mathbf{h}_0^T\mathbf{D}^{-1}\mathbf{h}_0\right)\right).$$

If defining

$$\begin{bmatrix}\tilde{\mathbf{D}} & \tilde{\mathbf{a}}^T\\ \tilde{\mathbf{a}} & \tilde{b}\end{bmatrix} = \begin{bmatrix}\mathbf{D} & \mathbf{a}^T\\ \mathbf{a} & b\end{bmatrix}^{-1},$$

then the expression within the inner parentheses becomes $\mathbf{h}_0^T(\tilde{\mathbf{D}} - \tilde{\mathbf{D}}^{-1})\mathbf{h}_0 + 2\mathbf{h}_0^T\tilde{\mathbf{a}}^T h_{\mathrm{p}} + h_{\mathrm{p}}\tilde{b}h_{\mathrm{p}}$. This corresponds to the Gaussian form with covariance Σ and $\boldsymbol{\mu}$ if $\Sigma^{-1} = \tilde{b}$, -$2\boldsymbol{\mu}^T\Sigma^{-1} = 2\mathbf{h}_0^T\tilde{\mathbf{a}}^{\mathbf{T}}$, and $\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu} = \mathbf{h}_0^T(\tilde{\mathbf{D}} - \mathbf{D}^{-1})\mathbf{h}_0$. The matrix inversion lemma immediately gives the covariance matrix, since

$$\Sigma^{-1} = \tilde{b} = b - \mathbf{a}^T\mathbf{D}^{-1}\mathbf{a}. \qquad [A.5]$$

Using the lemma a few times more leads to $\mathbf{h}_0^T(\tilde{\mathbf{D}} - \mathbf{D}^{-1})\mathbf{h}_0 = \mathbf{h}_0^T(\mathbf{D}^{-1}\mathbf{a}\Sigma^{-1}\mathbf{a}^T\mathbf{D}^{-1})\mathbf{h}_0 = \boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}$ from which $\boldsymbol{\mu}$ is derived:

$$\boldsymbol{\mu} = \mathbf{a}^T\mathbf{D}^{-1}\mathbf{h}_0. \qquad [A.6]$$

The values of Eqs. [A.5] and [A.6] also fulfill the third equality: -$2\boldsymbol{\mu}^T\Sigma^{-1} = 2\mathbf{h}_0^T\tilde{\mathbf{a}}^{\mathbf{T}}$.

References

1. Calamante F, Thomas D, Pell G, Wiersma J, Turner R. Measuring cerebral blood flow using magnetic resonance imaging techniques. J Cereb Blood Flow Metab 1999; 19:701–735.

2. Barbier EL, Lamalle L, Décorps M. Methodology of brain perfusion imaging. J Magn Reson 2001;13:496–520.

3. Villringer A, Rosen BR, Belliveau JW, Ackerman JL, Lauffer RB, Buxton RB, Choa YS, Wedeen VJ, Brady TJ. Dynamic imaging with lanthanide chelates in normal brain: Contrast due to magnetic susceptibility effects. Magn Reson Med 1988;6:164–174.

4. Boxerman JL, Hamberg LM, Rosen BR, Weisskoff RM. MR contrast due to intravascular magnetic susceptibility pertubations. Magn Reson Med 1995;34:555–566.

5. Nighoghossian N, Berthezene Y, Meyer R, Cinotti L, Adeleine P, Philippon B, Froment J, Trouillas P. Assessment of cerebrovascular reactivity by dynamic susceptibility contrast-enhanced MR imaging. J Neurol Sci 1997;149:171–176.

6. Speck O, Chang L, DeSilva NM, Ernst T. Perfusion MRI of the human brain with dynamic susceptibility contrast: Gradient-echo versus spin-echo techniques. J Magn Reson Imaging 2000;12:381–387.

7. Rempp KA, Brix G, Wenz F, Becker CR, Guckel F, Lorenz WJ. Quantification of regional cerebral blood flow and volume with dynamic susceptibility contrast-enhanced MR imaging. Radiology 1994;193:637–641.

8. Larsson HBW, Stubgaard M, Søndergaard L, Henriksen O. In vivo quantification of the unidirectional influx constant for Gd-DTPA diffusion across the myocardial capillaries with MR imaging. J Magn Reson Imaging 1994;4:433–440.

9. Østergaard L, Weisskoff RM, Chesler DA, Gyldensted C, Rosen BR. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: mathematical approach and statistical analysis. Magn Reson Med 1996;36:715–725.

10. Chabriat H, Pappata S, Østergaard L, Clark C, Pachot-Clouard M, Vahedi K, Jobert A, Bihan DL, Bousser M. Cerebral hemodynamics in CADASIL before and after acetazolamide challenge assessed with MRI bolus tracking. Stroke 2000;31:1904–1912.

11. Murase K, Kikuchi K, Miki H, Shimizu T, Ikezoe J. Determination of arterial input function using fuzzy clustering for quantification of cerebral blood flow with dynamic susceptibility contrast-enhanced MR imaging. J Magn Reson Imaging 2001;13:797–806.

12. Li TQ, Zheng GC, Østergaard L, Hindmarsh T, Moseley ME. Quantification of cerebral blood flow by bolus tracking and artery spin tagging methods. Magn Reson Imaging 2000;18:503–512.

13. Wirestam R, Andersson L, Østergaard L, Bolling M, Aunola JP, Lindgren A, Geijer B, Holtås S, Ståhlberg F. Assessment of regional cerebral blood flow by dynamic susceptibility contrast MRI using different deconvolution techniques. Magn Reson Med 2000;43:691–700.

14. Liu HL, Pu Y, Liu Y, Nickerson L, Andrews T, Fox PT, Gao JH. Cerebral blood flow measurement by dynamic contrast MRI using singular value decomposition with an adaptive threshold. Magn Reson Med 1999;42:167–172.

15. Vonken EPA, Beekman FJ, Bakker CJG, Viergever MA. Maximum likelihood estimation of cerebral blood flow in dynamic susceptibility contrast MRI. Magn Reson Med 1999;41:343–350.

16. Williams CKI, Rasmussen CE. Gaussian processes for regression. In: Advances in Neural Information Processing Systems, 8. eds. 1996; pp. 514–520.

17. Isaaks, Srivastava. An introduction to applied geostatistics. Oxford University Press 1989.

18. Bishop CM. Neural network for pattern recognition. Clarendon Press, Oxford 1995.

19. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in C. Cambridge 1992.

20. Goutte C, Nielsen FÅ, Hansen LK. Modelling the haemodynamic response in fMRI using smooth FIR filters. IEEE Trans Med imaging 2000;19:1188–1201.

21. Calamante F, Gadian DG, Conelly A. Delay and displersion effects in dynamic susceptibility contrast MRI: simulations using singular value decomposition. Magn Reson Med 2000;44:466–473.

22. Østergaard L, Chesler D, Weisskoff R, Sorensen A, Rosen B. Modeling cerebral blood flow and flow heterogeneity from magnetic resonance residue data. J Cereb Blood Flow Metab 1999;19:690–699.

23. Østergaard L, Gjedde A, Wu O, Gyldensted C, Rosen B, Sorensen G. Microscopic flow heterogeneity: Its determination by PWI and role in cerebral oxygen metabolism. In: Proc. Intl. Soc. Mag. Reson. Med 2001; p. 1433.

Table 1: The average maximum of exponential kernels (10 repetitions) and standard errors for three different temporal resolutions for each of the three methods. The performance of GPD was increasingly improved, while that of A-SVD was unaffected by the resolution. The performance of B-SVD decreased markedly as resolution increased.

	resol=1 Hz	*resol*=2 Hz	*resol*=5 Hz
GPD	0.008±0.001	0.009±0.001	0.010±0.001
A-SVD	0.008±0.001	0.0074±0.0008	0.0074±0.0007
B-SVD	0.010±0.002	0.009±0.002	0.006±0.004

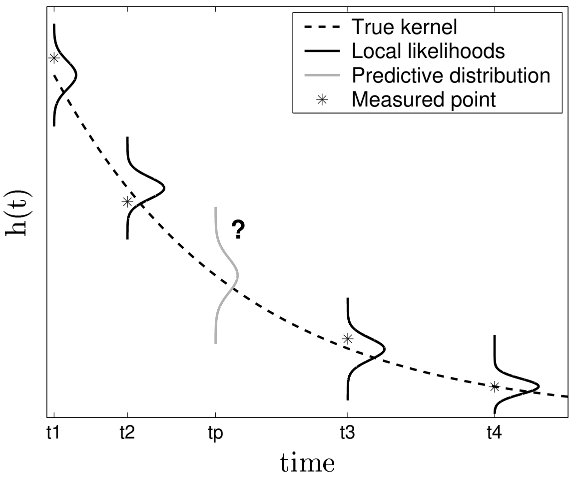

Figure 1: The kernel (dashed line) is measured at times t1, t2, t3 and t4. Each measured value belongs to a distribution. Gaussian process deconvolution estimates a new distribution at time t_{p} from the measured points and prior knowledge of smoothness. The estimate $h(t_{\mathrm{p}})$ is the mean of this distribution.

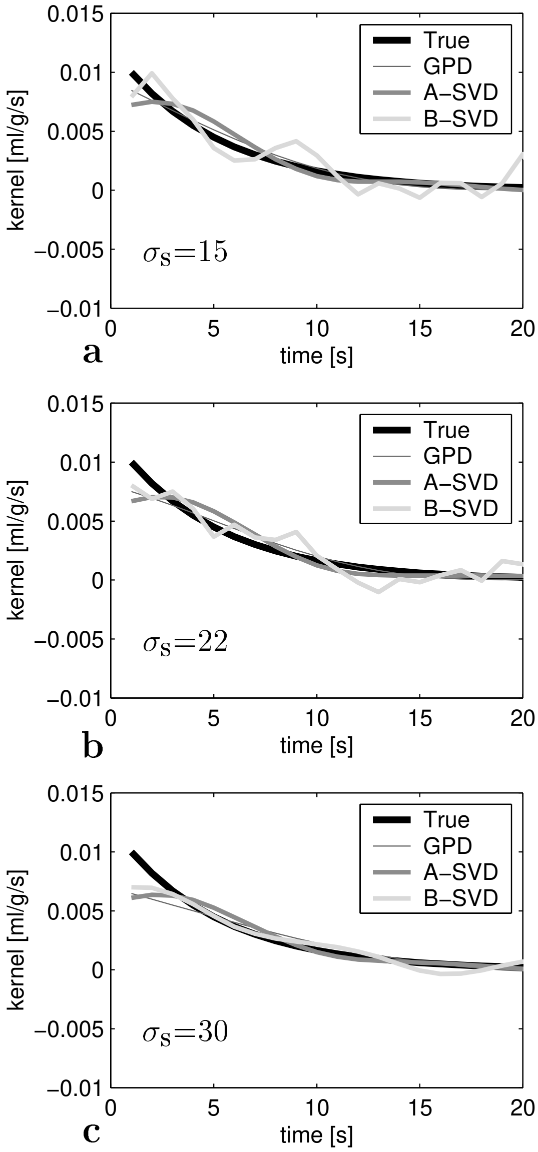

Figure 2: The true kernel and deconvolved exponential kernels with GM values of *rCBV*=5% and f=0.01 ml/g/s. B-SVD retained too much noise (plot a and b) or smoothed too much (plot c). The A-SVD method low-pass filtered the kernel at this perfusion. The GPD method performed increasingly well as the noise level was reduced and gave no artificial undulations.

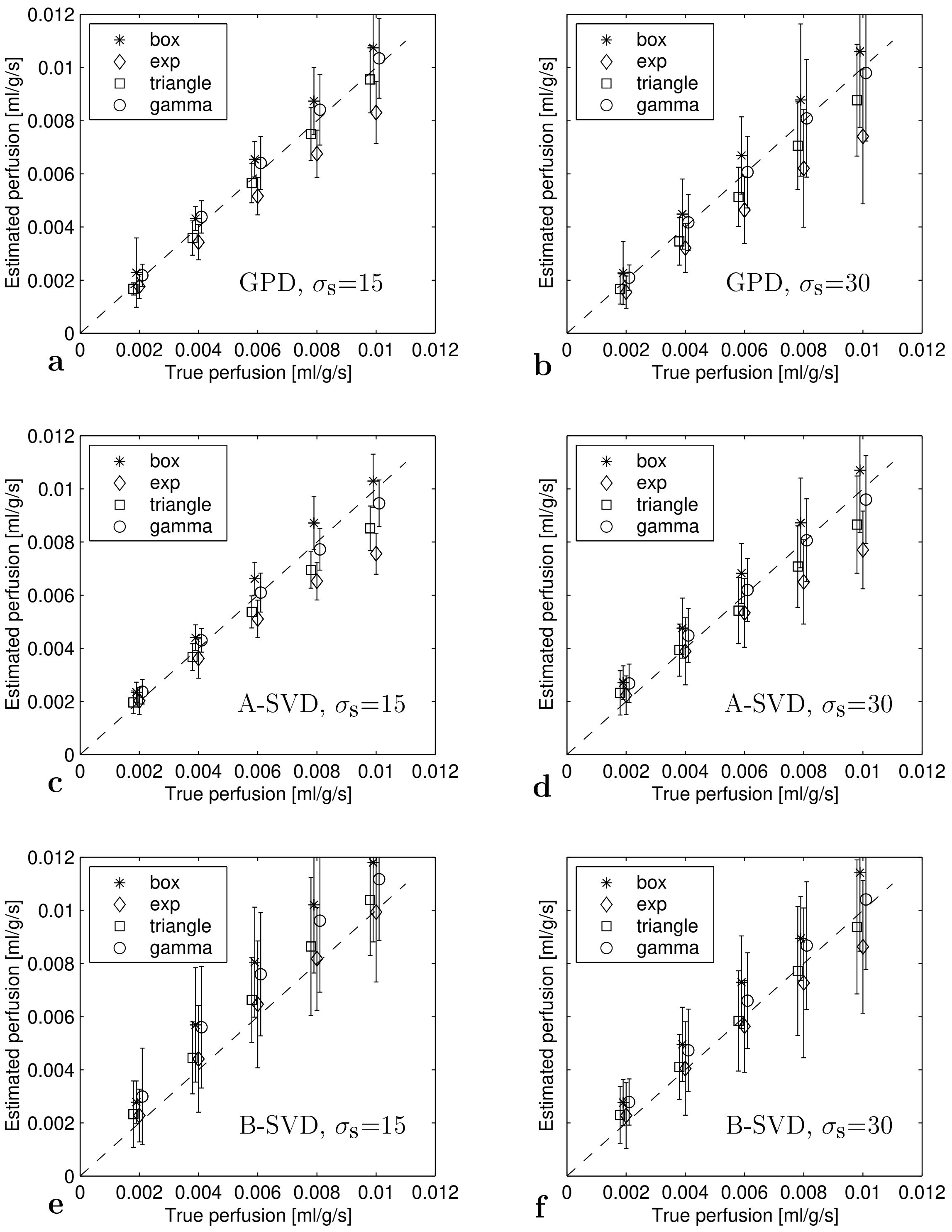

Figure 3: The estimated perfusion versus the true perfusion for different kernels. The dashed line is identity. Plots a) and b) show that GPD performed best for low perfusion and low noise. For the gamma-function, perfusion was reproduced, but for the exponential it was underestimated at high perfusion. Plots c) and d) show that A-SVD underestimated high perfusion and overestimated low perfusion, regardless of the noise level. Plots e) and f) show that B-SVD performed well for the exponential at the low noise level but overestimated for other kernels. Low perfusion values were overestimated.

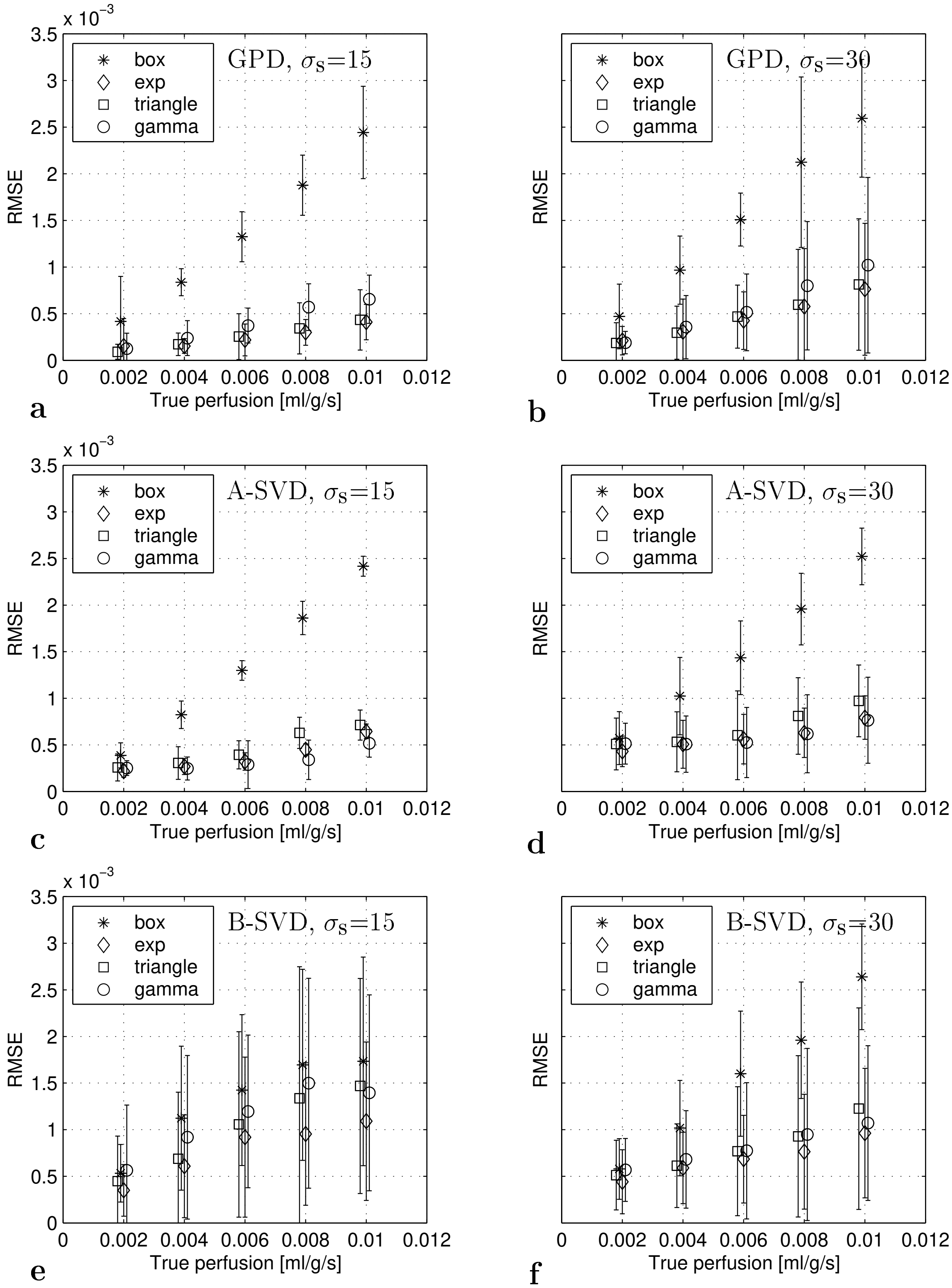

Figure 4: The RMSE versus the estimated perfusion for various kernels using three methods (rows) and two noise levels (columns). Plots a) and b) show that GPD estimated the curves well at low perfusion and low noise level. Plots c) and d) show that A-SVD had a certain off-set at low flows. Plots e) and f) show that B-SVD had the largest RMSE and the largest errors in this case of $rCBV$=5%.

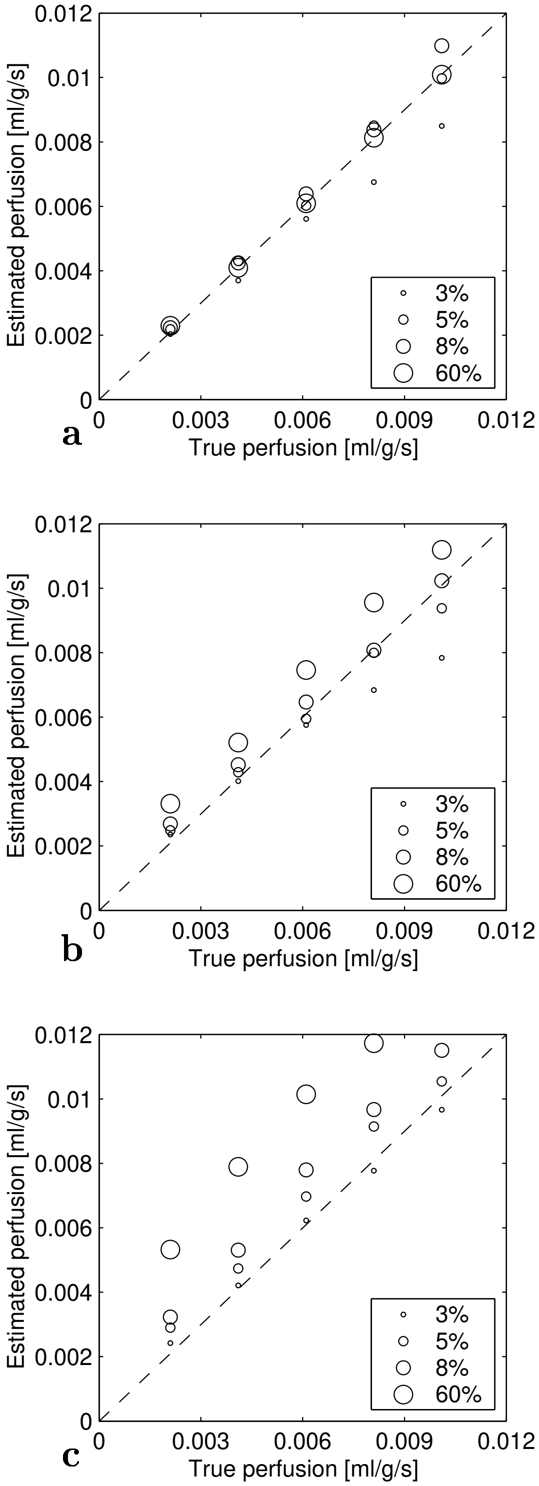

Figure 5: The estimated perfusion versus the true perfusion at $rCBV$ =3%, 5%, 8% and 60% for the gamma kernel at the medium noise level. The larger the marker the larger the $rCBV$. a) GPD. b) A-SVD. c) B-SVD. The dashed line is identity. As $rCBV$ increased, the estimate performed by GPD improved, whereas the estimates of the other methods increased above the true values.

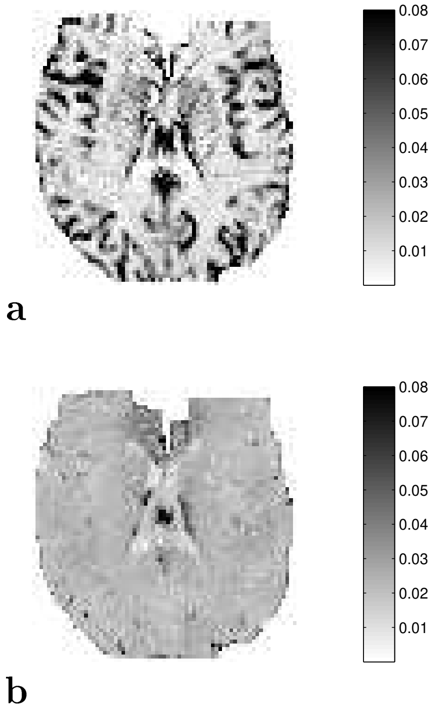

Figure 6: a) The maxima of the kernels, and b) the standard error estimated voxel-wise by GPD. The perfusion was higher in GM than in WM as expected. Also the basal ganglia are seen. The standard error was high in CSF and in edge voxels, indicating an expected poor estimate.

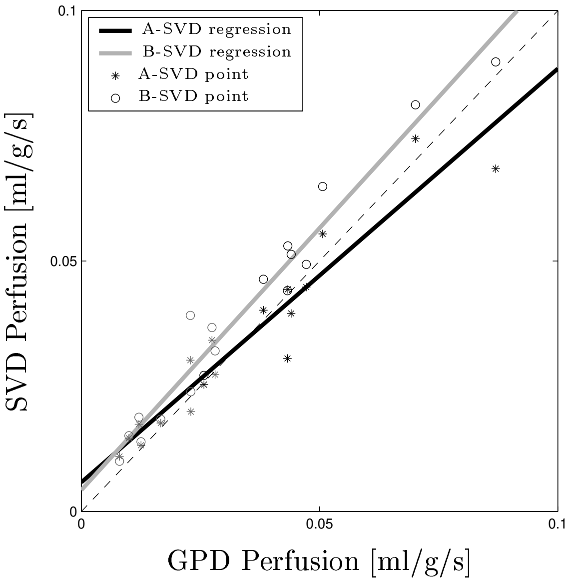

Figure 7: The GM and WM perfusion values obtained from nine subjects. The SVD estimates are plotted against the GPD estimates. The dashed line is identity. B-SVD always provided a higher estimate than GPD. A-SVD gave a higher estimate for low perfusion whereas GPD gave the higher estimate at high perfusion. These findings are in accord with the results obtain by simulations. The slopes of the regression lines are significantly different.

Table 1: The average maximum of exponential kernels (10 repetitions) and standard errors for three different temporal resolutions for each of the three methods. The performance of GPD was increasingly improved, while that of A-SVD was unaffected by the resolution. The performance of B-SVD decreased markedly as resolution increased.

Figure 1: The kernel (dashed line) is measured at times t1, t2, t3 and t4. Each measured value belongs to a distribution. Gaussian process deconvolution estimates a new distribution at time t_p from the measured points and prior knowledge of smoothness. The estimate $h(t_\mathrm{p})$ is the mean of this distribution.

Figure 2: The true kernel and deconvolved exponential kernels with GM values of $rCBV$=5% and f=0.01 ml/g/s. B-SVD retained too much noise (plot a and b) or smoothed too much (plot c). The A-SVD method low-pass filtered the kernel at this perfusion. The GPD method performed increasingly well as the noise level was reduced and gave no artificial undulations.

Figure 3: The estimated perfusion versus the true perfusion for different kernels. The dashed line is identity. Plots a) and b) show that GPD performed best for low perfusion and low noise. For the gamma-function, perfusion was reproduced, but for the exponential it was underestimated at high perfusion. Plots c) and d) show that A-SVD underestimated high perfusion and overestimated low perfusion, regardless of the noise level. Plots e) and f) show that B-SVD performed well for the exponential at the low noise level but overestimated for other kernels. Low perfusion values were overestimated.

Figure 4: The RMSE versus the estimated perfusion for various kernels using three methods (rows) and two noise levels (columns). Plots a) and b) show that GPD estimated the curves well at low perfusion and low noise level. Plots c) and d) show that A-SVD had a certain off-set at low flows. Plots e) and f) show that B-SVD had the largest RMSE and the largest errors in this case of $rCBV$=5%.

Figure 5: The estimated perfusion versus the true perfusion at $rCBV$ =3%, 5%, 8% and 60% for the gamma kernel at the medium noise level. The larger the marker the larger the $rCBV$. a) GPD. b) A-SVD. c) B-SVD. The dashed line is identity. As $rCBV$ increased, the estimate performed by GPD improved, whereas the estimates of the other methods increased above the true values.

Figure 6: a) The maxima of the kernels, and b) the standard error estimated voxel-wise by GPD. The perfusion was higher in GM than in WM as expected. Also the basal ganglia are seen. The standard error was high in CSF and in edge voxels, indicating an expected poor estimate.

Figure 7: The GM and WM perfusion values obtained from nine subjects. The SVD estimates are plotted against the GPD estimates. The dashed line is identity. B-SVD always provided a higher estimate than GPD. A-SVD gave a higher estimate for low perfusion whereas GPD gave the higher estimate at high perfusion. These findings are in accord with the results obtain by simulations. The slopes of the regression lines are significantly different.