A Bayesian Solution to the Inverse Problem of Electrocardiography

Changxin Lai
Department of Biomedical Engineering
Johns Hopkins University
Baltimore, MD 21218
clai29@jhmi.edu

Abstract

The solution to the inverse problem of electrocardiography enables a novel imaging technology to noninvasively assess the heart function. Statistical methods have been used to tackle the ill-posedness of the inverse problem. In this study, a Gaussian process model is incorporated into the classical Bayesian solution framework to combine spatial information and reflect realistic activation patterns on the heart tissue. I demonstrate that the solution using the proposed method outperforms the most commonly used empirical method in robustness and fidelity.

1 Background

Heart rhythm disorders are among the leading causes of deaths worldwide. The Electrocardiography has remained a centerpiece non-invasive diagnostic tool for heart function assessment since its invention. An electrocardiogram (ECG) is a recording of voltage with respect to time that captures the heart's electrophysiological activities from different angles using electrodes placed over skin. The commonly used 12-lead ECG, which is derived from 9 electrodes placed on the torso, although provides a cheap an quick assessment of the heart function, it lacks the capacity to directly assess spatial electrical activity at the level of the heart muscle at high spatial resolution.

Electrocardiographic imaging (ECGI) is a novel noninvasive imaging modality for studying cardiac electrophysiology, which has much higher spatial resolution than the 12-lead ECG and assisted the personalized management and treatment for arrythmias. This technique combines measured body-surface potential maps and patient specific anatomy geometries from CT/MRI to reconstruct heart surface potential maps by solving a inverse problem.

ECGI is achieved by solving what is known as the inverse problem of electrocardiography. In very broad terms, the inverse problem is defined as the determination of the electrical function of the heart from a number of remote recordings of potentials on body surface. In contrast, the forward problem is defined as the determination of body surface potentials from cardiac activities. The body surface potentials are the results from cardiac electrical activities, so the forward problem has a unique solution by assuming a static electrical field scenario and solving the Greens' second identity. The inverse problem, on the contrary, is an ill-posed problem and cannot be solved uniquely. To overcome the ill-posedness and obtain a reasonable solution is the biggest challenge in the inverse problem.

Preprint. Under review.

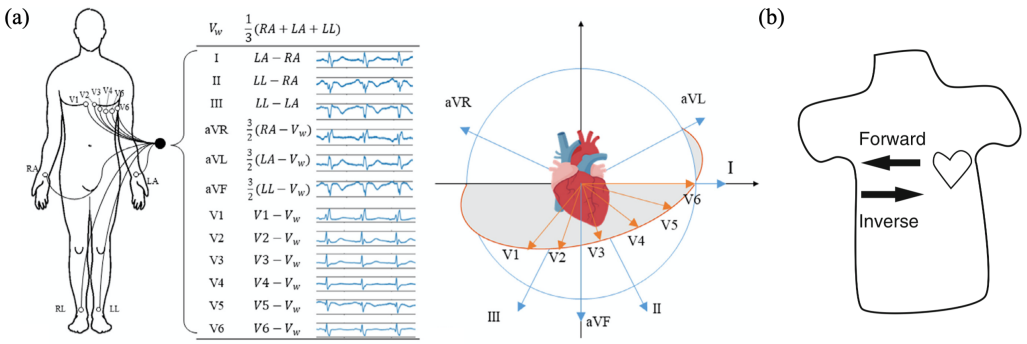

Figure 1: (a) The 12-lead ECG; (b) the forward and inverse problem of electrocardiography.

2 Literature review

2.1 Problem formulation

In the topic of electrocardiography, the heart-torso biophysical system at one time instant is modeled with a linear equation

$$\mathbf{y}(t) = \mathbf{A}\mathbf{x}(t) + \epsilon(t) \tag{1}$$

where $\mathbf{y}(t) \in \mathbb{R}^{N_y}$ is a vector containing the potentials at N_y different locations on the heart surface at the time instant t; vector $\mathbf{x}(t) \in \mathbb{R}^{N_x}$ contains the potentials at N_x different locations on the body surface; $\mathbf{A} \in \mathbb{R}^{N_y \times N_x}$ is the transfer matrix which maps the heart surface potentials to the body surface; $\epsilon(t) \in \mathbb{R}^{N_y}$ represents the measurement noise.

Under this linear model, the forward problem focuses on analyzing the geometry boundaries of the torso and heart and then obtaining a numeric solution to the Greens' second identity. The solution to the forward problem gives the transfer matrix $\mathbf{A}$. Then, with the matrix $\mathbf{A}$, one can easily map the heart surface potentials to the body surface by multiplication.

The inverse problem, which is the focus in this study, is then formulated as, given the transfer matrix $\mathbf{A}$ and the body surface potential measurements $\mathbf{y}$, to obtain the heart surface potentials $\mathbf{x}$. Due to the severe ill-conditioning of the forward operator $\mathbf{A}$, the inverse problem is ill-posed.

Equation (1) models the system at one time instant. When dealing with a period of time, one can concatenate the potential vectors together, and assume the the transfer matrix $\mathbf{A}$ to be time-invariant, then the system becomes

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{E} \tag{2}$$

where $\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), ..., \mathbf{y}(T)] \in \mathbb{R}^{N_y \times T}$ is the body surface potential matrix, $\mathbf{X} = [\mathbf{x}(1), \mathbf{x}(2), ..., \mathbf{x}(T)] \in \mathbb{R}^{N_x \times T}$ is the heart surface potential matrix, and $\mathbf{E} = [\epsilon(1), \epsilon(2), ..., \epsilon(T)] \in \mathbb{R}^{N_y \times T}$ is the noise.

2.2 Solutions to the inverse problem

The problem of determining $\mathbf{x}$ is an example of a rather general type of "inverse problem" that arises frequently in science and engineering. In inverse electrocardiography, as in other such inverse problems, common solutions fall into two categories. One is a deterministic framework, generally referred to as "regularization", in which an objective function to be minimized is composed of a combination of the norm of the residual error and some norm of a constraint function (or multiple constraint functions). The other category of formulation is a statistical framework, in which the solution is treated as random with an appropriate probability model, and a probabilistic error measure is minimized to find a likely solution.

Regularization The most classical deterministic regularization approach is Tikhonov regularization. One or more l_2 norm spatial penalties acting on a candidate solution are defined, and then a weighted

sum of these functions plus the residual norm is minimized. The penalty functions generally used constrain the magnitude of the inverse solution or its high spatial frequency content (often formulated via a first or second order spatial derivative). The resulting formulation can be summarized as:

$$\widehat{\mathbf{x}_\lambda} = argmin_{\mathbf{x}}\{\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda^2\|L\mathbf{x}\|_2^2\} \tag{3}$$

where L represents a "regularization operator" and λ is the regularization parameter controlling the the balance between the data fit and the amount of regularization. Then the Tikhonov solution simplifies to

$$\widehat{\mathbf{x}_\lambda} = (\mathbf{A}^T\mathbf{A} + \lambda^2 L^T L)^{-1}\mathbf{A}^T\mathbf{y} \tag{4}$$

Another approach of regularization aims to approximate the original forward operator $\mathbf{A}$ with an operator $\mathbf{A}_r$ which is similar to $\mathbf{A}$ in some sense, but much better conditioned. The most common version of this approach is the Truncated Singular Value Decomposition (TSVD).

Statistical modeling In statistical approaches, $\mathbf{x}$ and $\mathbf{n}$ are considered as random vectors with given probability models. The optimum Bayesian solution is the posterior mean of $\mathbf{x}$. Normally, Gaussian models are assumed for both $\mathbf{x}$ and $\mathbf{n}$. Then the Bayesian solution becomes identical to the maximum of the posterior probability distribution of $\mathbf{x}$, given the body surface measurements $\mathbf{y}$, and the solution is called the Maximum a Posteriori (MAP) solution. The simplest and most widely used model assumes that each time instant is a vector drawn from a stationary distribution; in the case of a Gaussian model, this means that a constant mean and a fixed covariance matrix determine the model. The constant mean is almost always taken to be zero, because we care more about the morphologies of the reconstructed signals, rather than the absolute values.

3 Methods

3.1 Data modeling

In this study, I initially followed the classical statistical modeling method for the electrocardiography data. The body surface potential vector $\mathbf{x}$ and the measurement noise ϵ is modelled as random vectors following multi-variate normal distribution.

$$\mathbf{x} \sim MN(\mu_{x,0}, \Sigma_{x,0}) \tag{5}$$

$$\epsilon \sim MN(0, \Sigma_{\epsilon,0}) \tag{6}$$

Then, the posterior probability density of $\mathbf{x}$ conditional on the measurements $\mathbf{y}$ and all the variables $\theta = \{\mu_{x,0}, \Sigma_{x,0}, \Sigma_{\epsilon,0}\}$ is

$$\begin{aligned} p(\mathbf{x}|\mathbf{y},\theta) \sim & p(\mathbf{y}|\mathbf{x},\theta)p(\mathbf{x}|\theta) \qquad (7)\\ =&(2\pi)^{-\frac{N_y}{2}} det(\Sigma_{\epsilon,0})^{-\frac{1}{2}} exp[-\frac{1}{2}(\mathbf{y}-\mathbf{A}\mathbf{x})^T\Sigma_{\epsilon,0}^{-1}(\mathbf{y}-\mathbf{A}\mathbf{x})]\\ &(2\pi)^{-\frac{N_x}{2}} det(\Sigma_{x,0})^{-\frac{1}{2}} exp[-\frac{1}{2}(\mathbf{x}-\mu_{x,0})^T\Sigma_{x,0}^{-1}(\mathbf{x}-\mu_{x,0})]\\ \sim& exp[-\frac{1}{2}(\mathbf{y}-\mathbf{A}\mathbf{x})^T\Sigma_{\epsilon,0}^{-1}(\mathbf{y}-\mathbf{A}\mathbf{x}) - \frac{1}{2}(\mathbf{x}-\mu_{x,0})^T\Sigma_{x,0}^{-1}(\mathbf{x}-\mu_{x,0})]\\ \sim& exp[-\frac{1}{2}(\mathbf{x}^T\mathbf{A}^T\Sigma_{\epsilon,0}^{-1}\mathbf{A}\mathbf{x} - 2\mathbf{x}^T\mathbf{A}^T\Sigma_{\epsilon,0}^{-1}\mathbf{y} + \mathbf{x}^T\Sigma_{x,0}^{-1}\mathbf{x} - 2\mathbf{x}^T\Sigma_{x,0}^{-1}\mu_{x,0})]\\ \sim& exp\left\{-\frac{1}{2}[\mathbf{x} - (\mathbf{A}^T\Sigma_{\epsilon,0}^{-1}\mathbf{A} + \Sigma_{x,0}^{-1})^{-1}(\mathbf{A}^T\Sigma_{\epsilon,0}^{-1}\mathbf{y} + \Sigma_{x,0}^{-1}\mu_{x,0})]^T(\mathbf{A}^T\Sigma_{\epsilon,0}^{-1}\mathbf{A} + \Sigma_{x,0}^{-1})[\mathbf{x} - ...]\right\} \qquad (8) \end{aligned}$$

Equation (8) suggests that the posterior distribution of $\mathbf{x}$ follows a multi-variate normal distribution

$$\mathbf{x}|(\mathbf{y},\theta) \sim MN(\mu_{x,n}, \Sigma_{x,n}) \tag{9}$$

where

$$\mu_{x,n} = (\mathbf{A}^T\Sigma_{\epsilon,0}^{-1}\mathbf{A} + \Sigma_{x,0}^{-1})^{-1}(\mathbf{A}^T\Sigma_{\epsilon,0}^{-1}\mathbf{y} + \Sigma_{x,0}^{-1}\mu_{x,0}) \tag{10}$$

$$\Sigma_{x,n} = (\mathbf{A}^T\Sigma_{\epsilon,0}^{-1}\mathbf{A} + \Sigma_{x,0}^{-1})^{-1} \tag{11}$$

If $\mathbf{x}$ is assumed to have a prior mean of 0 (i.e. $\mu_{x,0} = 0$), ϵ is assumed to be white noise (i.e. $\Sigma_{\epsilon,0} = \sigma_{n,0}^2 I$), then the posterior mean and variance simplifies to

$$\mu_{x,n} = (\mathbf{A}^T\mathbf{A} + \sigma_{\epsilon,0}^2\Sigma_{x,0}^{-1})^{-1}\mathbf{A}^T\mathbf{y} \tag{12}$$

$$\Sigma_{x,n} = (\mathbf{A}^T\mathbf{A} + \sigma_{\epsilon,0}^2\Sigma_{x,0}^{-1})^{-1}\sigma_{\epsilon,0}^2 \tag{13}$$

The inverse problem now boils down to setting prior variances for $\mathbf{x}$ and ϵ, and then calculating the posterior parameters using the equations above. Assume we have some training data of coupled $(\mathbf{x}_{training}, \mathbf{y}_{training})$, and the forward operator $\mathbf{A}$ is known (by solving the forward problem), then the problem becomes to estimate better $\Sigma_{x,0}$ and $\sigma_{\epsilon,0}^2$ from the training data. Instead of just calculating empirical covariance matrices from the training data, I further modeled them in the following way.

Assume the prior belief of $\sigma_{\epsilon,0}^2$ before learning from the training data is a random variable following inverse-Gamma distribution

$$\begin{aligned}\sigma_{\epsilon,0}^2 &\sim Inverse - Gamma(\alpha_0, \beta_0) \\ \alpha_0 &= 1, \beta_0 = 1\end{aligned} \tag{14}$$

Then the posterior probability density of $\sigma_{\epsilon,0}^2$ after learning from the training data is

$$\begin{aligned}p(\sigma_{\epsilon,0}^2|\mathbf{x}_{training}, \mathbf{y}_{training}) =& p(\mathbf{y}_{training}|\mathbf{x}_{training}, \sigma_{\epsilon,0}^2)p(\sigma_{\epsilon,0}^2) \\ \sim& [(\sigma_{\epsilon,0}^2)^{N_y}]^{-\frac{1}{2}} exp[-\frac{1}{2\sigma_{\epsilon,0}^2}(\mathbf{y}_{training} - \mathbf{A}\mathbf{x}_{training})^T(\mathbf{y}_{training} - \mathbf{A}\mathbf{x}_{training})] \\ & (\frac{1}{\sigma_{\epsilon,0}^2})^{\alpha_0 - 1} exp[-\frac{\beta_0}{\sigma_{\epsilon,0}^2}] \\ \sim& (\frac{1}{\sigma_{\epsilon,0}^2})^{\alpha_0 - 1 + \frac{N_y}{2}} \\ & exp[-\frac{\beta_0 + (\mathbf{y}_{training} - \mathbf{A}\mathbf{x}_{training})^T(\mathbf{y}_{training} - \mathbf{A}\mathbf{x}_{training})/2}{\sigma_{\epsilon,0}^2}]\end{aligned} \tag{15}$$

Equation (15) suggests that the posterior distribution of $\sigma_{\epsilon,0}^2$ after learning from the training data is also inverse-Gamma

$$\begin{aligned}\sigma_{\epsilon,0}^2|(\mathbf{x}_{training}, \mathbf{y}_{training}) &\sim Inverse - Gamma(\alpha_n, \beta_n) \\ \alpha_n &= \alpha_0 + \frac{N_y}{2} \\ \beta_n &= \beta_0 + (\mathbf{y}_{training} - \mathbf{A}\mathbf{x}_{training})^T(\mathbf{y}_{training} - \mathbf{A}\mathbf{x}_{training})/2\end{aligned} \tag{16}$$

To model the covariance of $\mathbf{x}$, I want to incorporate spatial information to reflect the realistic activation patterns on the heart tissue. The change of heart surface potentials is the result from cell activation, and the activation of one cell is triggered by the activation of its neighborhood. Therefore, cells with close distance should have similar activation patterns, and their potential signals should have higher correlations. To model this property in the covariance of $\mathbf{x}$, I further assume $\mathbf{x}$ follows a Gaussian Process as a response of spatial locations $u = (u_1, u_2, u_3)$.

$$\mathbf{x} \sim GP(0, K(u, u')) \tag{17}$$

$$K(u, u') = \sigma_x^2 \prod_{k=1}^{3} exp\left\{-\theta_k|u_k - u_k'|_2^2\right\} \tag{18}$$

The three dimensions are parametrized separately to allow for the anisotropy on the heart tissue (fiber orientations). The Gaussian Process model is fit using the training data $(\mathbf{x}_{training}, \mathbf{y}_{training})$ to estimate the parameters $(\sigma_x^2, \theta_1, \theta_2, \theta_3)$. With the Gaussian Process parameters, we can obtain a prior estimate of $\Sigma_{x,0} = K(u, u')$, then apply it to estimate unknown $\mathbf{x}$.

3.2 Experiment design

The data used in this study is from an atrial fibrillation case, containing a coupled 5-second recording of heart surface potentials $\mathbf{X}$ and body surface potentials $\mathbf{Y}$ with a sampling frequency of 500Hz (i.e. in total $T = 2499$ time instants). The geometries of the torso and the heart are known, and the forward transfer matrix $\mathbf{A}$ is hereby calculated using boundary element method. The heart surface mapping has $N_x = 2039$ nodes and the body surface mapping has $N_y = 659$ nodes.

The data is divided into a 1-second training set ($t = 1 \sim 500$) and a 4-second testing set ($t = 501 \sim 2499$). The prior noise variance $\sigma^2_{\epsilon,0}$ is estimated on the training set using equation (16). The Gaussian process parameters $(\sigma_x^2, \theta_1, \theta_2, \theta_3)$ is estimated using an R package `GPfit`. In `GPfit`, θ_k is further parametrized as $\theta_k = 10^{\beta_k}$ and it uses a multi-start L-BFGS-B algorithm for optimizing the deviance, which is more robust and faster. Despite these efforts, the computational complexity to fit the whole training set is still enormous. Therefore, only a small set of the training data is used in the GP parameter estimation. 20% of spatial locations are randomly sampled, and a time instant is sampled periodically every 0.1 sec. The prior estimate of $\Sigma_{x,0} = K(u, u')$ is then obtained using the estimated GP parameters.

When estimating the heart surface potentials $\mathbf{x}$, I used both MAP estimation and Markov chain Monte Carlo (MCMC). The MAP solution is

$$\widehat{\sigma^2_\epsilon}_{MAP} = \frac{\beta_n}{\alpha_n - 1} \tag{19}$$

$$\widehat{\mathbf{x}}_{MAP} = (\mathbf{A}^T\mathbf{A} + \widehat{\sigma^2_\epsilon}_{MAP} K(u, u')^{-1})^{-1}\mathbf{A}^T\mathbf{y} \tag{20}$$

The full conditional posterior distributions for σ^2_ϵ and $\mathbf{x}$ used in Gibbs sampling are

$$\sigma^2_\epsilon|(\mathbf{x}, \mathbf{y}) \sim Inverse - Gamma(\alpha_n + \frac{N_y}{2}, \beta_n + \frac{(\mathbf{y} - \mathbf{A}\mathbf{x})^T(\mathbf{y} - \mathbf{A}\mathbf{x})}{2}) \tag{21}$$

$$\mathbf{x}|(\sigma^2_\epsilon, \mathbf{y}) \sim MN(\mathbf{A}^T\mathbf{A} + \sigma^2_\epsilon K(u, u')^{-1})^{-1}\mathbf{A}^T\mathbf{y}, (\mathbf{A}^T\mathbf{A} + \sigma^2_\epsilon K(u, u')^{-1}\sigma^2_\epsilon) \tag{22}$$

However, due to the limitation of computational power, I cannot finish the MCMC computation in time. Therefore, the study only covers the results from MAP estimation. The code for MCMC is available in the attachment.

As a validation of the proposed modeling method (especially the modeling of the covariance matrix of $\mathbf{x}$, the proposed results are compared against results using empirical covariance matrix $\Sigma_{x,0}$ calculated on the training data.

4 Results

On the training data, the noise variance is estimated using equation (16). The parameters are

$$\alpha_n = 164751, \beta_n = 745.66$$

The parameters for the Gaussian process are estimated using `GPfit`

$$\sigma_x^2 = 492.89$$
$$\beta_1 = 2.07, \beta_2 = 2.07, \beta_3 = 2.14$$

Using equation (18), the covariance matrix $\Sigma_{x,0}$ is estimated.

Correlation coefficients (CC) between the reconstructed signal and the true signal are calculated at each node, as a measurement of signal reconstruction quality for each location. The distribution of the 2039 correlation coefficients are plotted in Figure 2. The reconstruction using Gaussian process variance has a mean CC of 0.55, and the reconstruction using empirical variance has a mean CC of 0.50. Figure 2 also suggests that the reconstruction using Gaussian process variance outperforms the reconstruction using empirical variance.

The reconstructed signals at one location are plotted in Figure 3. From the plots we can see that the reconstruction using Gaussian process variance has less fluctuations and is more stable than using

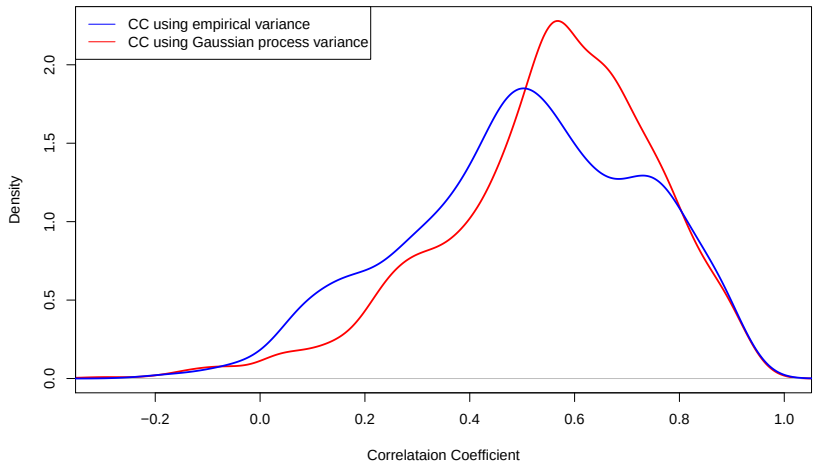

Figure 2: The distributions of correlation coefficients at different locations

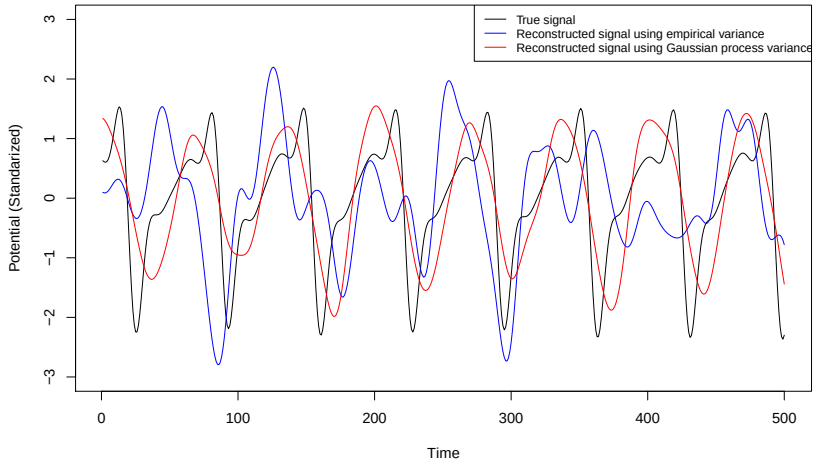

Figure 3: Reconstructed signals at one location on the heart surface

empirical variance. The locations of the signal spikes are of particular interest in this study, in that they represent the activation time of the tissue at this region. In this sense, Gaussian process variance enables a higher fidelity of reconstructed activation, because its spikes followed the true spikes much better than the empirical variance model does.

5 Conclusions

In this study, I investigated the statistical modeling method to solving the inverse problem of electrocardiography and improved the classical Bayesian solution by Gaussian process modeling of the heart surface potential maps. The Gaussian process modeling enables the model to incorporate spatial information and mimic realistic activation patterns. Comparison against results from empirical variance suggests that the Gaussian process-modeled variance outperformed the empirical variance with regard to CC, activation patterns and robustness.

References

[1] Figuera, C., Suárez-Gutiérrez, V., Hernández-Romero, I., Rodrigo, M., Liberos, A., Atienza, F., ... & Alonso-Atienza, F. (2016). Regularization techniques for ECG imaging during atrial fibrillation: a computational study. Frontiers in physiology, 7, 466.

[2] Macfarlane, P. W., Van Oosterom, A., Pahlm, O., Kligfield, P., Janse, M., & Camm, J. (Eds.). (2010). Comprehensive electrocardiology. Springer Science & Business Media.

[3] Oosterom, A. V. (1997). Incorporation of the spatial covariance in the inverse problem. Biomedizinische Technik, 42, 33-36.