DEVELOPMENT AND TUNING OF A 3-D STOCHASTIC INVERSION METHODOLOGY TO THE EUROPEAN ARCTIC

Juerg Hauser[1], Kathleen M. Dyer[2], Michael E. Pasyanos[2], Hilmar Bungum[1], Jan Inge Faleide[3], and Stephen A. Clark[3]

NORSAR[1], Lawrence Livermore National Laboratory[2], and University of Oslo[3]

Sponsored by the National Nuclear Security Administration

Award Nos. DE-AC52-08NA28651[1,3] and LL08-BAA08-38-NDD03[2]
Proposal No. BAA08-38

ABSTRACT

The development of seismic models for the European Arctic, as for most regions, has traditionally been limited to proposing one model that provides the best possible fit to one or several datasets while observing some regularization constraints. This approach ignores a fundamental property of any inverse problem in geophysics, non-uniqueness, that is, if a model can be found to satisfy given datasets an infinite number of alternative models will exist that satisfy the datasets equally well. In this study we apply a Bayesian formulation to the problem of developing a geophysical model for the Barents Sea. The solution to the inverse problem presented here is no longer only the best-fitting model, but an ensemble of models that satisfy the data to the same degree: the posterior distribution. It is based on two sources of information: (1) the different datasets, notably surface-wave group velocities, regional body-wave travel times, gravity data, compiled 1D velocity models, and thickness relationships between sedimentary rocks and underlying crystalline rocks; and (2) prior information, which is independent from the data. In this work, the posterior distribution is generated using a Monte Carlo Markov Chain (MCMC) algorithm, which has successfully been applied to Yellow-Sea Korean Peninsula region. It samples models from the prior distribution, the set of plausible models based on a priori information, and tests them against the different datasets. While being computationally much more expensive, such a stochastic inversion provides a more complete picture of solution space. It gives an overview of the different structures that can explain the observed datasets while taking the uncertainties in the data into account. A notable improvement with respect to BARENTS3D and previous applications of the MCMC methodology is the model parameterization introduced in this work. We allow for linear changes of seismic parameters within the layers. This results in a parameterization that is superior to the concept of constant seismic parameters for each layer.

The European Arctic, encompassing oceanic crust, continental shelf regions, rift basins, and old cratonic crust, is a challenging setting for any imaging technique and therefore an ideal environment for demonstrating the practical advantages of an MCMC methodology. The MCMC technique under development in this study allowed us to develop a first stochastic model for the region, albeit at a low spatial resolution. The preliminary results show the posterior distribution to be in agreement with knowledge of the regional tectonic setting. We can for example resolve the changes in crustal thickness associated with the northern edge of the continental shelf. More important than absolute values for the crustal thickness are the uncertainties associated with them. The predicted uncertainties correlate well with the variation in the data coverage and data quality in the region. The uncertainties are significantly smaller in data rich regions (e.g., Northern Norway) than in regions with poor data coverage (e.g., Novaya Zemlya). This indicates that the technique behaves as expected; thus, we are properly tuning the methodology by allowing the Markov Chain adequate time to fully sample the model space. Plans for the next phase of the project include increasing the spatial resolution, optimizing the parameterization, and focusing on a more detailed analysis of the posterior distribution.

OBJECTIVES

The area of interest for this study is the European Arctic, in particular the Barents Sea and surrounding regions such as the Norwegian-Greenland Sea, the Southern Eurasian Basin, Novaya Zemlya, the Kara Sea, the East-European Lowlands, the Kola Peninsula and the Arctic plate boundary. The region is characterized by large variations in sediment thicknesses, crustal velocities, crustal thicknesses, and upper mantle velocities. The aim is to develop a stochastic geophysical model i.e., sample the posterior distribution, the ensemble of models that are all in agreement with our set of observations. The posterior distribution is generated from the prior distribution by probabilistically sampling the model space and preferentially accepting models that produce a good fit between the model predictions and the observed data. The observations used in this study are surface-wave group velocities, regional body-wave travel times, gravity data, compiled 1D velocity models, and thickness relationships between sedimentary rocks and underlying crystalline rocks. The MCMC methodology employed here is based on the technique used by Pasyanos et al. (2006) to derive a geophysical model for the Yellow Sea/Korean Peninsula region given surface-wave group velocities, body-wave travel times, receiver functions and gravity data.

Our initiative stands in contrast to recent approaches to modeling the crust and mantle lithosphere in the European Arctic, where the aim has been to find one best-fitting model and little to no attention was devoted to systematically analyzing the model uncertainties (e.g., Levshin et al., 2007; Ebbing et al., 2007; Ritzmann et al., 2007; Ritzmann and Faleide, 2008). A notable exception in this context is the global shear wave velocity model of the crust and upper mantle by Shapiro and Ritzwoller (2002). In the final stage of their inversion procedure, an MCMC technique is used to generate a posterior distribution and thereby estimate model uncertainties.

The stochastic model under development in this study is an ensemble of models that satisfy the datasets to similar degrees. Given a posterior distribution one can draw inferences about the true nature of the Earth and identify robust features in these models. We characterize the posterior distribution by computing the mean and standard deviation for each model parameter. It is however important to realize that these summary statistics do not provide the full picture of the posterior distribution. In the case of a multimodal posterior distribution computing a mean is a poor reflection of the ensemble. A more complete picture of the posterior could be obtained by cluster analysis. Vasco et al. (1993) uses this statistical technique to characterize the posterior and make inferences about properties that are shared by all the models forming the posterior. The importance of knowledge about the robustness of features in a geophysical model can not be over-emphasized, especially when interpreting such models in a geological context. Another important advantage of a stochastic model is that it allows estimating the uncertainties in model predictions due to uncertainties in the model parameters, e.g., uncertainties for travel times and travel-time corrections based on the uncertainties in the model parameters. Such information can be included in earthquake location procedures, thus resulting in much more realistic estimates of location uncertainties.

One of the aims of this project is to improve the MCMC methodology of Pasyanos et al. (2006) for application to a variety of regions of interest. What is an appropriate model parameterization in one region might not be capable of capturing the geological characteristics in another area. Constraining the structure at the nodes of a mesh based on a fixed degree interval between the nodes is an appropriate parameterization near the equator but not in the Barents Sea region. We therefore changed the parameterization to allow for an arbitrary distribution of nodes. Another important improvement is that seismic parameters can now vary linearly with depth in the layers used to describe the sediments, the crystalline crust, and the mantle. This is an enhancement over the constant velocity and density layers used in previous studies (e.g., Pasyanos et al., 2006; Ritzmann et al., 2007).

RESEARCH ACCOMPLISHED

A practical MCMC scheme needs to address four challenges: (1) Development of a mechanism to propose random perturbations to the state in model space that are in agreement with fundamental knowledge about the Earth's structure; (2) Limited computational resources, requiring efficient forward solvers and the smallest number of free parameters necessary to capture the structure; (3) Assignment of uncertainties to all observables; and (4) Application of statistical concepts to describe the posterior. In the first phase of the project we have made significant progress in overcoming these challenges for the European Arctic and have consequently developed a first stochastic model for the European Arctic, albeit at a low spatial resolution.

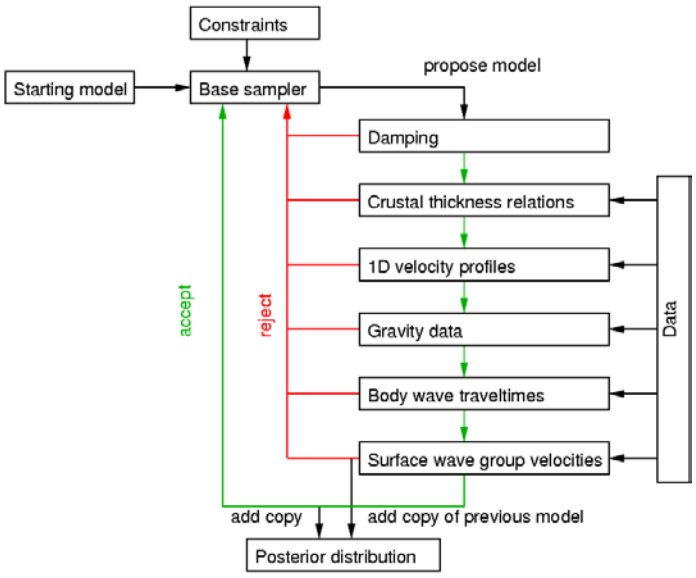

Figure 1. Cartoon showing the multistage MCMC inversion methodology used in this study.

We begin this paper with a short introduction of the MCMC methodology. After discussing the datasets and starting model, we will present our stochastic model and conclude with a discussion of areas of the methodology where we expect to make progress in the second half of the project.

Methodology and Approach

Different inversion algorithms can be classified by the tradeoff between exploration and exploitation (Sambridge and Mosegaard, 2002). Exploration means that one tries to improve the misfit between predictions and observations by choosing a new model entirely at random. Exploitation means that one tries to propose a better fitting model based on the information available from previous samples or sometimes just the current best-fit model. The steepest descent algorithm (Gill et al., 1981) is a good example of a method that only exploits information available from the current sample of model space, while a uniform search algorithm is an example of a method that tries to explore the model space completely. Exploitative algorithms will be more efficient, but are more likely to fall into a local minimum. Explorative algorithms on the other hand are less efficient at converging to a solution, but less likely to get trapped in a local minimum. MCMC and other Monte Carlo techniques try to be exploitative and explorative at the same time.

An MCMC algorithm recovers the posterior distribution using a cleverly constructed Markov Chain. In our work, a Markov Chain is a perturbative sequence of random changes to the starting model. Constructing the Markov Chain in such a way that it preferably contains models that satisfy data allows one to reconstruct the posterior, given sufficient time to move away from the starting model. Figure 1 illustrates the MCMC algorithm used in this study. The base sampler generates a proposed model m_j, a random perturbation of the current model m_i, also referred to as the current state in model space. The current model is the last model that has been added to the posterior or in the case of the first iteration of the algorithm the starting model. Given the model prediction $d_i(m)$ (i.e., the data predicted using the proposed model), the observed data d_i and the standard deviation (i.e., uncertainty) for each observation σ_i, one can determine the likelihood $L(m|d)$ for a model m using Equation 1.

$$L(m|d) = k \exp\left[\sum_i \left(\frac{\left|d_i(m) - d_{obs,i}\right|^n}{\sigma_i^{\,n}}\right)\right] \qquad (1)$$

The normalization constant k will be discussed in more detail later. The value of n determines the weight assigned to outliers. For n=2 the L^2 norm of the misfit vector is used, resulting in a good tradeoff between minimizing misfit weights (i.e., being insensitive to outliers) and maximizing them (i.e., being overly sensitive to outliers).

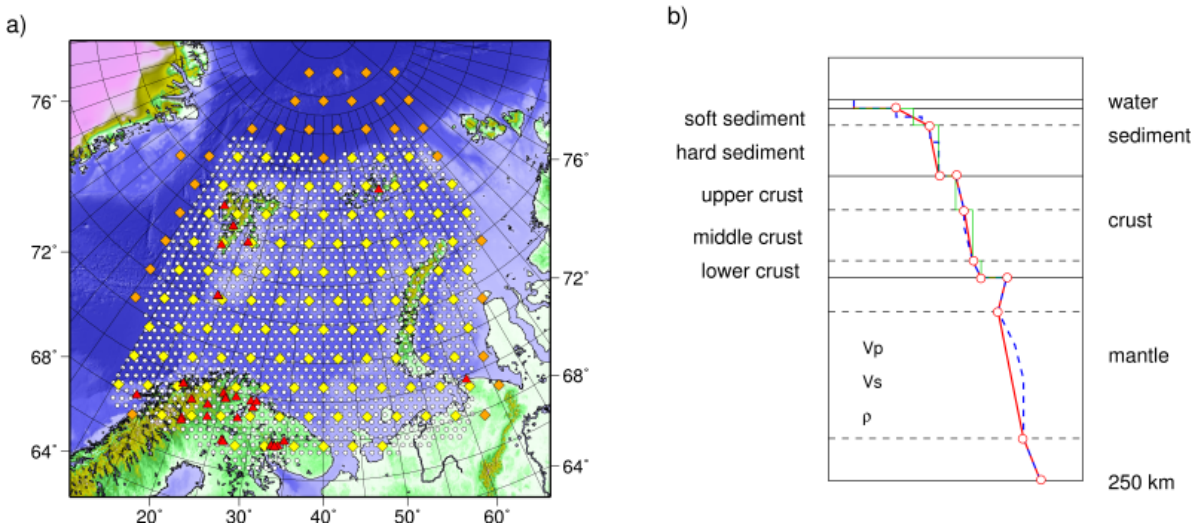

Figure 2. Model parameterization used in this study. a) Map of the region of interest. Locations of the 1D models in BARENTS50 are marked by white circles. The diamonds mark the positions of the 1D models to be constrained in this study. They are yellow if their initial value is based on BARENTS50 and BARMOD and orange if CRUST2.0 has been used to initialize them. b) 1D model parameterization. The dashed blue line represents a typical Vp versus depth curve for the region of interest. The red line is a piecewise interpolation of the red circles where the seismic parameters are defined in this study. The green line illustrates for the crust the limitations of a parameterization based on layers with constant seismic parameters, when trying to describe Vp as a function of depth.

Given the likelihood for a proposed model $L(m_j|d)$ and the current model $L(m_i|d)$, a decision has to be made whether or not the proposed model m_j should be accepted into the posterior distribution. If we would accept all models, we would simply recover the prior, the ensemble of models based on a priori information contained in the base sampler and starting model and not constraint by the data. For $L(m_j|d) \leq L(m_i|d)$ we accept the proposed model and add it to the posterior distribution. For $L(m_j|d) > L(m_i|d)$, the proposed model is accepted if the probability $L(m_j|d) / L(m_i|d)$ is larger than a random number drawn from a uniform distribution between 0 and 1. Otherwise, we reject the proposed model and add a copy of the current model to the posterior. The explorative component of the MCMC methodology, accepting models with worse fit to the data with a certain probability, allows one to move away from a local extremum. The resulting MCMC algorithm performs a guided search through model space. It tends to focus on regions in model space that better fit the prior information and the data, a process known as importance sampling.

In this work we develop a model that satisfies more than one dataset. Given a proposed model, a separate likelihood function is computed for each dataset (i.e., stage of the inversion) and a model is only added to the posterior if it has been accepted in all stages; otherwise, the previous model is added to the posterior (see Figure 1). The normalization constant k plays an important role in adjusting the influence the different datasets have on the posterior distribution. Ideally, when generating the posterior distribution, each stage should have the same influence, the acceptance rates in the individual stages should be comparable to each other, and no dataset should dominate the inversion result. This can be achieved by carefully tuning the normalization constant k.

Prior Information

Our starting model is based on BARENTS50 (Ritzmann et al., 2007), BARMOD (Levshin et al., 2007), and CRUST2.0 (Bassin et al., 2000). CRUST2.0 is used to initialize the model for nodes that fall outside the area covered by BARENTS50 and BARMOD (see Figure 2a). We parameterize the model as a set of 1D models, which define the main layers (water, sediments, crystalline crust, and mantle) that represent the geological structure. Figure 2b illustrates the parameters used to describe the structure at each node. This parameterization of structure associated with every 1D model represents a trade-off between trying to describe the structure as accurately as possible and limiting the number of parameters to avoid having to search a model space which is unnecessarily large. The parameterization results in 47 unknowns for each 1D model and a total of 6,721 unknowns. Note that

these unknowns are not completely independent of each other due to the constraints imposed on the prior distribution.

The success of an MCMC methodology relies partly on a mechanism for generating plausible models based on the prior information. In a Bayesian formulation this process is known as sampling the prior distribution. The prior distribution is given by the starting model and constraints for the individual parameters. The most basic constraints are the upper and lower limits and the standard deviation for the normal distribution from which we randomly draw new model parameters when perturbing the model. More sophisticated constraints employed in this study are ranges for acceptable Vp/Vs ratios, a positive velocity contrast imposed across the Moho and across the interface between the crystalline crust and sediments, and a limit on the maximum decrease of seismic parameters with increasing depth. Plausible ranges for seismic parameters in the area of interest and the relationship between them were derived from previous studies covering the region (e.g., Levshin et al., 2007; Ritzmann, et al., 2007; Breivik et al., 2002).

It is possible that the starting model represents a local minimum of the misfit function. The starting model is therefore jumbled before starting the inversion to ensure that the MCMC algorithm does not get trapped in a local minimum of the misfit function. This is commonly done by exchanging a subset (e.g., 10%) of the 1D models between randomly selected nodes. In this work, the 1D models have been randomly exchanged between all the nodes to begin the inversion very far away from the starting model and to thereby also demonstrate the explorative capabilities of our MCMC algorithm. Another important consideration is the step size in model space - the degree of perturbation applied to the model. If the current model exhibits a good fit to the data, we might be interested in taking smaller steps in model space. On the other hand, if the current model does not satisfy the data, we might be interested in taking larger steps in model space. When perturbing the model, we randomly choose a set of nodes where we will perturb the seismic parameters. We vary the degree of perturbation by randomly varying the size of the set of nodes. Any realistic geological structure is likely to exhibit a certain degree of spatial smoothness. The structure at an individual node should therefore not be completely independent from its neighbors. The base sampler takes this into account by applying spatial smoothing to each node that has been perturbed when proposing a new model. This is critical for proposing models which are more likely to be accepted in the different stages of the inversion.

Given the high quality of a priori knowledge about the structure in the target region already included in the starting model, one could decide to reject models which deviate too far from the starting model. We therefore added a stage to our inversion scheme whereby each proposed model can be compared with a reference model, e.g., the starting model (Figure 1). The same effect is achieved in iterative non-linear tomography through damping. It remains to be seen whether or not we need a damping stage to obtain a meaningful model for the target region. The low-resolution model presented in this paper was developed without using the damping stage. However, a damping stage might be necessary for obtaining meaningful results once the spatial resolution is increased.

Datasets

An important result of Ritzmann et al. (2007) is that the thickness relationships between sedimentary rocks and underlying crystalline rocks can be resolved to some degree within discrete geological provinces. Comparing layer thicknesses is computationally far less demanding than calculating group velocities. The first stage of our inversion is therefore the crustal thickness relations stage. Only when a model has passed this stage do we move to the computationally more expensive stages to determine whether or not it should be added to the posterior. Ordering the stages by the computation time helps to decrease the overall computation time.

Local crustal structure is explicitly constrained by seismic reflection and refraction lines. A velocity profile database, based on a subset of the sampled profiles used by Ritzmann et al. (2007), provides important constraints for the crustal part of the model. The 1D crustal velocity models are available as Vp as a function of depth for positions along the reflection and refraction lines. They were obtained by sampling velocity grids derived from seismic reflection or refraction lines. It would be beyond the scope of this work to revisit the methods used by Ritzmann et al. (2007) to derive the velocity profiles in detail. We parameterize the profile database in the same way our model is parameterized and compare each proposed model with the profiles in the database at predefined depth intervals. Uncertainties have been assigned to each profile based on the categories derived by Ritzmann et al. (2007). Figure 3b shows the coverage of the target region by the velocity profiles colored according to their uncertainty in Vp. Clearly the data quality in the Eastern Barents Sea (Figure 3c) is much lower than in the Western Barents Sea (Figure 3a). The crustal velocities shown in Figure 3c justify the relatively large uncertainty of 0.4 km/s

assigned to the whole seismic line. New deep seismic surveys have been collected and processed by the University of Oslo and should become available for this study in the coming months.

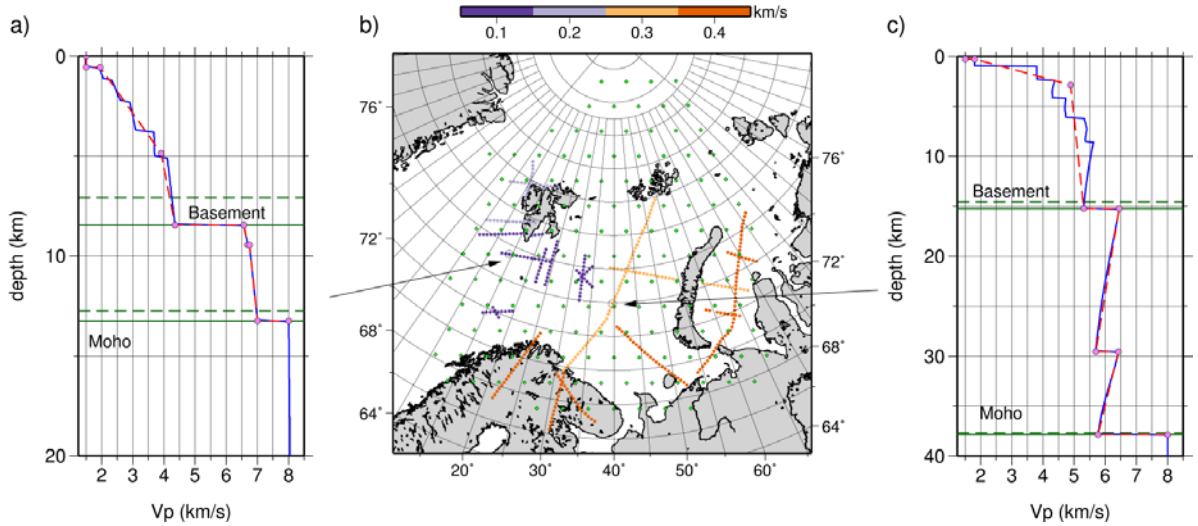

Figure 3. 1D velocity profile data. a) Example of a 1D velocity profile with a low uncertainty assigned to the data. The blue line shows the Vp derived by Ritzmann et al. (2007). The green dashed lines indicate the depths of the basement and Moho, interpolated from BARENTS50. The solid green lines mark the depth of the basement and Moho based on the sampled Vp. The violet points connected by the dashed red line are the control points for the best-fitting profile in the parameterization used in this study. b) Points where the seismic lines were sampled have been colored according to the uncertainty in Vp. The green diamonds mark the locations of the 1D models used to describe the structure in the region of interest. c) Example for a 1D velocity profile with a high assigned uncertainty.

Surface-wave group velocities (Levshin et al., 2007) and regional body-wave travel-time data provided by Lawrence Livermore National Laboratory are used to constrain the lithosphere and upper mantle. The Mid-Atlantic ridge and the Arctic plate boundary are the dominant sources for seismicity in the region. We only use events where the entire propagation path is within the boundaries of the model. Our region of interest therefore extends further north than BARENTS50. This means that we are now also imaging the changes in structure associated with the northern edge of the continental shelf. Figure 4 shows the path coverage of the target region by the different types of body and surface waves. Clearly events associated with the plate boundary to the north provide additional ray paths in the Barents Sea which provide valuable constraints. For the regional body-wave data and surface-wave group velocities, uncertainty information was available from previous studies using these data. All four datasets discussed so far provide much more constraints for the Western Barents Sea than for the eastern part, and do not constrain the density.

There are two sources of information that can be used to constrain density: (1) information about the relationship between Vs and density for different crustal lithologies (Breivik et al., 2002) and (2) a uniform coverage of the area with free air gravity data from the Arctic Gravity project (http://earth-info.nga.mil/GandG/wgs84/agp/). The surface wave tomography by Levshin et al. (2007) recovered a higher-velocity body in the upper mantle below the Eastern Barents Sea Basin, Novaya Zemlya, and the Kara Sea. Such a seismic velocity anomaly within the upper mantle is predominantly caused by temperature variations, rather than compositional effects (Goes et al. 2000), thus one would expect a density anomaly. However, gravity modeling (Ebbing et al., 2007; Ritzmann and Faleide, 2008) suggests a simple horizontally layered mantle density structure. This suggests, that the Eastern Barents Sea is underlain by a cold cratonic keel similar to the Siberian Craton (Ritzmann and Faleide, 2008), as density contrasts due to temperature variations below cratons are presumed to be entirely compensated by chemical depletion. This assumption has an important consequence for the base sampler: we do not impose any constraints on the relationship between Vs and density within the mantle.

Preliminary Results

One Markov chain consisting of 20,000 iterations of the engine was computed. The posterior distribution is given by the last 25% of the chain to ensure that the results have achieved the independent post burn in convergence. The number of iterations is significantly larger than in the previous study by Pasyanos et al. (2006). There are three factors contributing to the slower convergence in this work: (1) we have randomly swapped the profiles for all and not just 10% of the nodes in our model; (2) we take smaller steps in model space by only perturbing the model at randomly selected nodes and (3) we have two additional stages of model evaluation.

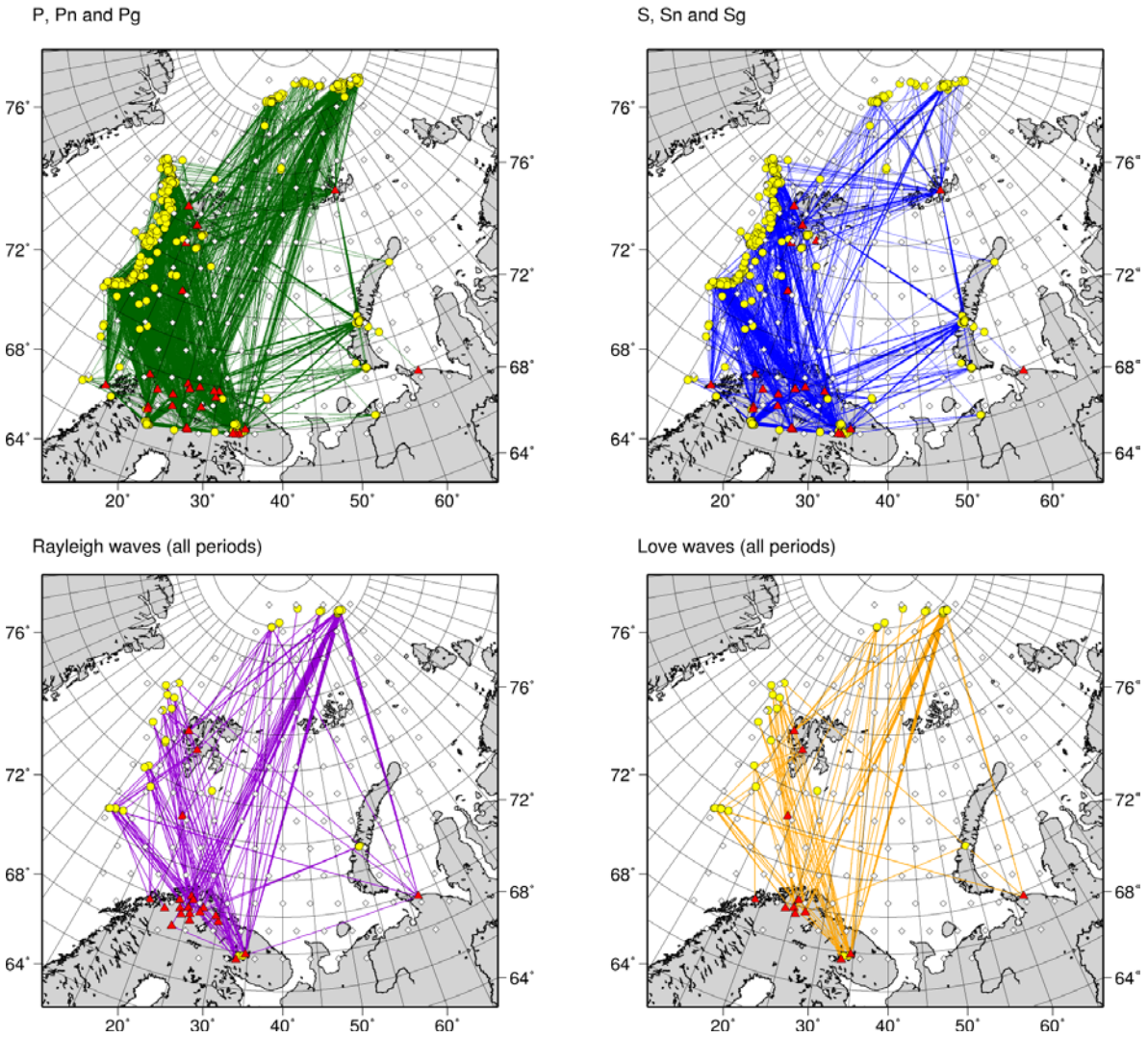

Figure 4. Path coverage of the target region for various seismic phases; white diamonds denote the location of the 1D models, earthquakes are shown as yellow circles and the stations are plotted as red triangles.

Figure 5 (a and b) shows the summary statistics for the thickness of crystalline crust. Clearly the method has recovered the dominant feature of crystalline crustal thickness in the region, the change in thickness associated with the continent-ocean transition. Figure 5a also illustrates the problem caused by the low spatial resolution of the model. Our preliminary model can only recover large scale features extending for more than 100km. The model developed by Ritzmann et al. (2007) proposes thicknesses for the crystalline crust under Novaya Zemlya of locally up to 50 km. We cannot recover such features at this stage because our 1D models are representative of an average structure in the vicinity of the node, not for local extrema of, for example, thickness of the crystalline crust.

More interesting than the mean value for the thickness of the crystalline crust is its standard deviation shown in Figure 5b. Overall the standard deviations (i.e., uncertainties) tend to be lower in Northern Norway when compared with the ones in the region around Novaya Zemlya. Nevertheless, they vary significantly between neighbouring nodes of the model. We recall that seismic reflection and refraction lines are the dominant constraints for the crustal structure. It therefore comes as no surprise that the amount of data available in the vicinity of a given node correlates to some degree with the standard deviation, in particular in the Western Barents Sea (see Figure 5b).

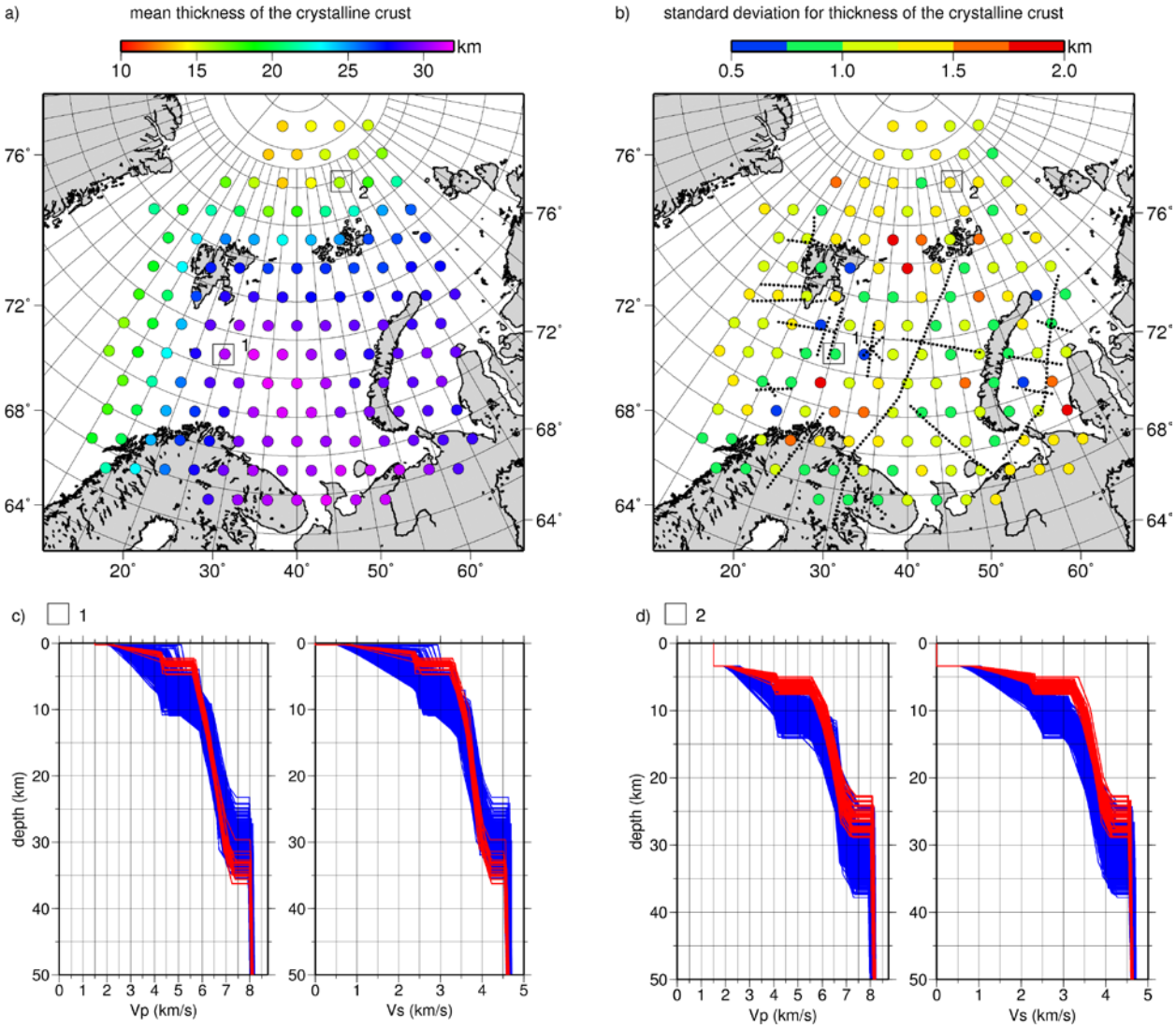

Figure 5. Stochastic model for the Barents Sea. Mean (a) and standard deviation (b) for thickness of the crystalline crust. The black dots in (b) show the locations of the 1D velocity profiles obtained from seismic reflection and refraction lines. For the two nodes labeled 1 and 2 the model profile distributions for Vp and Vs are shown (c and d). Blue lines show the prior distribution and red lines show the posterior distribution.

One way to see how the method is performing is to look at the prior and posterior distribution of individual nodes. Figure 5 (c and d) shows the prior and posterior distribution for two selected nodes, one located in a data-rich region (Figure 5c) and one in a data-poor region (Figure 5d). In both cases, the data have constrained the posterior distribution but, as we would expect, the variability of the models forming the posterior distribution is much higher for the node located in the data-poor region.

It is important to be aware of the fact that computing summary statistics for the posterior distribution can be deceptive. The posterior distribution might be multimodal especially in a joint inversion, where the constraints imposed on model parameters by the different datasets might not be in agreement with each other. Plotting the posterior distributions of profiles for individual nodes can help in determining whether or not the posterior distribution is multimodal. For the two nodes shown in Figure 5, computing the mean and standard deviation can clearly capture the main characteristics of the posterior distribution.

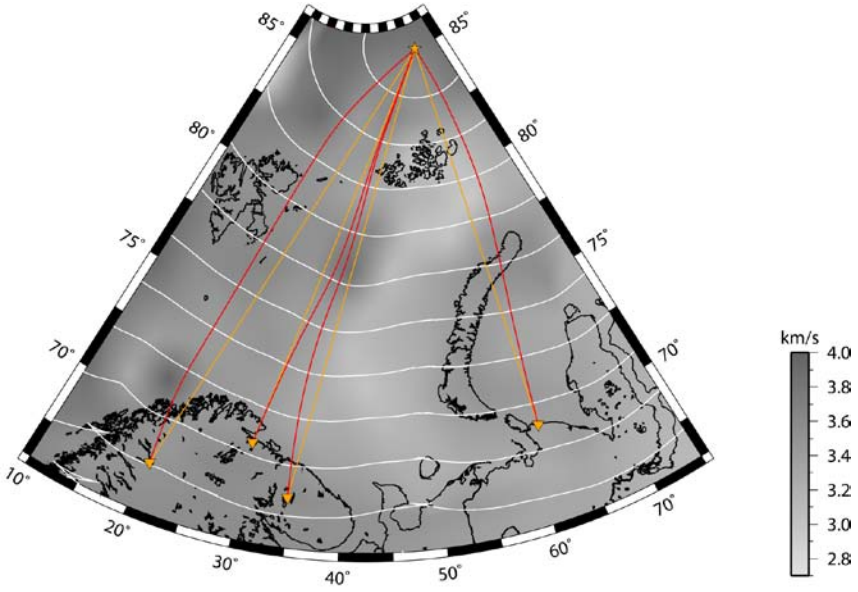

Figure 6. Phase velocity map for Rayleigh waves with a period of 16s based on the starting model. The orange star marks the source location and the orange triangles denote receivers. The great circle ray path is plotted as an orange line. Ray paths have been computed using ray theory and are plotted red. Note that the wave front plotted in white starts to triplicate due to the lateral changes in velocity.

CONCLUSIONS AND RECOMMENDATIONS

The preliminary results clearly show that an MCMC technique can be used to recover large-scale geological features in our region of interest. The main drawback of the stochastic model presented here is its low spatial resolution; the distance between the individual 1D models is about 180 km. The final model will need to have a spatial resolution comparable to the best-fitting models published for this region. We would prefer to use a variable spatial resolution higher in data rich regions and lower in data poor regions. Our implementation currently allows for an arbitrary distribution of the nodes. It is therefore possible to locally increase the spatial resolution by adding additional nodes.

Ideally, one would optimize the distribution of 1D models as part of the inversion process. The idea of a self-adapting parameterization is not new to seismic imaging and has for example been applied to cross-borehole tomography by Curtis and Snieder (1997). Part of optimizing the model parameterization is finding the optimum number of 1D models needed to constrain the structure. Trans-dimensional MCMC algorithms (e.g., Green, 1995; Geyer and Møller, 1994) can deal with cases where the number of model parameters is unknown. Whether or not these methods can be used in the construction of our model will depend to some degree on the influence they have on the computation time and convergence.

Computation time is an important consideration when using an MCMC technique. Predicting the different datasets for thousands of different models requires efficient forward solvers. In practice this means that one has to resort to using very simple forward solvers, such as 1D ray tracing for body waves using a path-average model. The dilemma between the need for an efficient forward solver and the need to be as accurate as possible can be illustrated using

the example of the surface-wave data. In a first approximation, surface waves follow the great circle paths. This allows us to compute group velocities by sampling the model along the great circle path. In the presence of lateral velocity variations, however, surface waves will no longer follow the great circle path. Rayleigh wave paths (at T = 16s), computed using ray theory, clearly deviate from the great circle path (Figure 6). The largest lateral variations in velocity are found in the upper crust. The shorter the period the more sensitive a surface wave is to structure near the surface and therefore the more the surface wave paths deviate from the great circle path. For longer periods (50s in the case of our starting model) the surface wave paths no longer deviate significantly from the great circle paths. The planned use of ray theory will be a significant improvement over the great circle approximation especially for the group velocity observations made for high frequencies. Nevertheless, one should keep in mind that ray theory is still an approximation; it neglects finite frequency effects which could be captured by solving the full wave equation. This is however currently not feasible given the computational requirements.

So far we only have computed summary statistics for the posterior distribution. This means we are not harvesting all the information contained in the posterior distribution. In the second half of the project, we will have to investigate new ways to analyze and visualize the posterior distribution. Cluster analysis which tries to group models based on their similarity appears to be a promising tool in this context.

ACKNOWLEDGEMENTS

We thank the Department of Geosciences at the University of Oslo for providing us with access to the Titan III high performance computing facilities.

REFERENCES

Bassin, C., G. Laske, and G. Masters (2000). The Current Limits of Resolution for Surface Wave Tomography in North America, *EOS Trans AGU* 81: F897.

Breivik, A. J., R. Mjelde, P. Grogan, H. Shimamura, Y. Murai, Y. Nishimura, and A Kuwano (2002). A possible Caledonide arm through the Barents Sea imaged by OBS data, *Tectonophysics* 355: 67–97.

Curtis, A., and Snieder, R. (1997). Reconditioning inverse problems using the genetic algorithm and revised parameterisation, *Geophysics* 62: 1524–1532.

Ebbing, J., C. Braitenberg, and S. Wienecke (2007). Insights into the lithospheric structure and the tectonic setting of the Barents Sea region by isostatic considerations, *Geophys. J. Int.* 171: 1390–1403, doi:10.1111/j.1365-246X.2007.03602.x.

Geyer G. and Møller J. (1994). Simulation procedures and likelihood inferences for spatial point processes, *Scandinavian Journal of Statistics* 21: (4), 369–373.

Gill, P. E., W. Murray, and Wright M. H. (1981). *Practical Optimization*, Academic San Diego, CA.

Goes, S., Govers, R. and Vacher, P. (2000). Shallow mantle temperature under Europe from P and S wave tomography, *J. Geophys. Res*. 105: (B5), 11153–11169.

Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika* 82: (4), 711–732; doi:10.1093/biomet/82.4.711.

Levshin, A., J. Schweitzer, C. Weidle, N. Shapiro, N. Maercklin, and M. Ritzwoller (2007). Surface wave tomography for the Barents Sea and Surrounding Regions, *Geophys. J. Int*. 170 : 441–459, doi:10.1111/j.1365-246X.2006.03285.x.

Pasyanos, M. E., G. A. Franz, and A. L. Ramirez, (2006). Reconciling a geophysical model to data using a Markov Chain Monte Carlo algorithm: An application to the Yellow Sea-Korean Peninsula region, *J. Geophys. Res.* 111: B03313, doi:10.1029/2005JB003851.

Ritzmann, O. and J. I. Faleide (2008). The crust and mantle lithosphere in the Barents Sea/Kara Sea region, *Tectonophysics*, TECTO 124236, doi: 10.1016/j.tecto.2008.06.018.

Ritzmann, O., N. Maercklin, J. I. Faleide, H. Bungum, W. D. Mooney, and S. T. Detweiler (2007). A 3D geophysical model for the crust in the greater Barents Sea region: Model construction and basement characterization, *Geophys. J. Int*. 170: 417–435, doi: 10.1111/j.1365-246X.2007.03337.x.

Sambridge, M. and K. Mosegaard (2002). Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.,* 40: (3), 1009, doi:10.1029/2000RG000089.

Shapiro, N. M. and M. H. Ritzwoller (2002). Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle, *Geophys. J.Int*. 151: 88–105.

Vasco, D. W., L. R. Johnson, and E. L. Majer (1993). Ensemble inference in geophysical inverse problems, *Geophys. J. Int*. 115: (3), 711–728.