

## Data-driven approaches to frontier of earth and planetary sciences

KUWATANI, Tatsu<sup>1\*</sup> ; NAGATA, Kenji<sup>2</sup> ; OKADA, Masato<sup>2</sup> ; KOMAI, Takeshi<sup>1</sup>

<sup>1</sup>Graduate School of Environmental Studies, Tohoku University, <sup>2</sup>Graduate school of Frontier Sciences, University of Tokyo

It is important to extract essential processes and structures from observed data sets in order to understand the dynamic behavior of the earth and planetary systems. Recently, many powerful methodologies have been proposed to extract useful information from high-dimensional data sets in information sciences. Since the summer of 2013, we have launched a big scientific project entitled as 'Initiative for high-dimensional data-driven science through deepening sparse modelling' supported by the Ministry of Education, Culture, Sports, Science and Technology in Japan (<http://sparse-modeling.jp/index.e.html>). The main purpose of this project is to develop the innovative mathematical methodology for understanding the world of nature by tight fusion of information science and natural science. The project includes a wide variety of natural sciences such as biology, medicine, brain science, earth and planetary sciences and astronomy. Two main key technologies are important to develop data-driven sciences: One is 'Bayesian estimation', which is a probabilistic methodology which can estimate cause from effect by reversing law of causality, and the other is 'Sparse modeling', which is a mathematical framework which can effectively extract a small number of essential explanatory variables from high-dimensional data sets. In this presentation, based on these two key technologies, we will overview data-driven analytical technologies with some examples (Kuwatani et al., 2014a *Earth, Planets and Space*; 2014b *Physical Review E*; 2014c *Scientific Reports*).

## Estimation of bulk composition for meteoritic samples using handheld XRF

NIIHARA, Takafumi<sup>1\*</sup> ; MIYAMOTO, Hideaki<sup>1</sup> ; HONG, Peng K.<sup>1</sup> ; OGUMA, Midori<sup>1</sup> ; DOHM, James M.<sup>1</sup>

<sup>1</sup>UMUT

Over 50,000 meteorites are recovered in hot and cold desert, and most of these meteorites are thought to originate from asteroids. The Hayabusa mission confirmed that the samples from asteroid Itokawa (S-type) have compositions similar to ordinary chondrites, therefore meteorites and asteroids have direct correlation, and thus the importance of meteorites increases to understand small bodies that could record early solar system histories. Direct comparison of meteorites and asteroids may produce important information of the surface materials of asteroids. Surface chemical composition can be obtained by onboard instruments such as APXS and GRS. On the other hand, bulk chemical composition for most meteorites are missing (only 10% have been analyzed) due to a limited mass of each sample. Therefore, most meteoritic samples cannot be analyzed using conventional techniques (e.g. INAA and XRF). To solve this issue, we are now developing an analytical procedure using a handheld XRF (Olympus Delta) to estimate bulk chemical composition without any sample preparation, although the accuracy of measurement is relatively lower than conventional techniques. The tube voltages of 40 kV and 10 kV are used for heavy elements (heavier than Ti) and light elements, respectively. The fundamental parameter method is adopted for calculation of elemental abundances and gives us semiquantitative values. One big problem of this method is analyzable elements are limited; light elements (lighter than Mg) cannot be detected with our instrument. In this presentation, we will report preliminary results of the precision and accuracy of measurement tested through geochemical standard (JB-3, JA-3, and JG-1a, supplied from geological survey of Japan). We currently calibrated for 6 major elements (Si, Ti, Al, Fe, Mn, and Ca) and can analyze within the variation of <5 % (GSJ values +/- 1 wt. %) in powder condition (soil cup with Mylar filter). Our purpose of this study is to expeditiously analyze the composition of a slab (or block) of the meteorites without any kind of sample preparation. We will continue to calibrate using slabs of diverse rock types stored in the Umut.

Keywords: meteorite, asteroid, bulk composition

## Microscopic model for nucleation pathways of polymorphic minerals from solutions

OKAMOTO, Atsushi<sup>1\*</sup>; KUWATANI, Tatsu<sup>1</sup>; OMORI, Toshiaki<sup>3</sup>; HUKUSHIMA, Koji<sup>2</sup>

<sup>1</sup>Tohoku University, <sup>2</sup>University of Tokyo, <sup>3</sup>Kobe University

Rock-forming processes in the Earth (e.g., reaction, mass transport, deformation) are usually too slow for us, and thus it is difficult to obtain the direct time-series data of the system. Therefore, it is important to develop the methodology how we extract the information on dynamics of rock-formation from “frozen pictures” (textures). In this study, as an example, we focus on the nucleation of metastable phases from solutions, because the formation of metastable minerals is essential on the progress of some rock-forming reactions (hydrothermal and metamorphic reactions), and often provides unique textures.

A phenomena, that a metastable mineral firstly crystallizes from liquid and translates into more stable phases, has been known as the Ostwald rule, and such phenomena has been reported in crystallization of various materials, including proteins, clay minerals, iron oxides, carbonate (Pouget et al., 2009) and silica (Morse and Casey, 1988). Okamoto et al. (2010) found characteristic textures of silica that cristobalite is formed as a cover of spherical amorphous silica particles, and that quartz is formed within a deposit of cristobalite. The phenomena following the Ostwald rule would result from competitive processes of bulk chemical potential and interface energies for solid-solid and solid-liquid; however, the detailed dynamics is not clear.

Inspired by the occurrences of silica polymorphs, we present a microscopic model (two-dimensional Potts lattice gas model; c.f., Sanders, 2007) for understanding nucleation mechanism of polymorphic phases from solutions. We construct a system which is composed of a solvent (s4) and three polymorphic solid phases (s1, s2, s3). The configuration energy is expressed as combination of interaction term and bulk chemical potential. One of the most characteristic features of this model is that both state and position of the solid phases are updated by Metropolis dynamics. We evaluated the equilibrium conditions by exchange Monte Carlo method (Hukushima and Nemot, 1996), as well as temporal evolution of the system. At temperature higher than critical temperature ( $kbT/J = 1.1$ ), nucleation of the solid phase does not occur. At lower temperature, a large nucleus of the least stable phase is initially formed and then it translates into more stable phases following the Ostwald rule. The resulting textures are similar to those observed in natural system (i.e. silica precipitates from aqueous solutions), indicating that our simplified model captures an essential process of the Ostwald rule. The free energy surface of the system, which has several local minimums, changes depending on temperature and cluster size, which controls the nucleation pathways. In usual, the interfacial energies among the minerals are much more difficult to be estimated than bulk chemical potential (i.e., solubility); therefore, the characteristic textures among the polymorphic minerals could be useful for evaluating the magnitudes of the interfacial energies relative to the chemical potential values, by comparing the snapshots of the configuration in the microscopic model.

### References

- Hukushima, K., and Nemoto, K., 1996. *J. Phys. Soc. Japan*, 65, 1604-1608.  
Morse J. W. and Casey W. H., 1988. *Am. J. Sci.* 288, 537-560.  
Okamoto, A., Saishu, H., Hirano, N., and Tsuchiya, N., 2010. *Geochimica et Cosmochimica Acta*, 74, 3692-3706.  
Pouget, E.M., Bomans, P.H.H., Goos, J.A.C., et al., 2009, *Science*, 323, 1455-1458.  
Sanders D. P., 2007. *Phys Rev B*, 75, 132101.

Keywords: rock texture, Ostwald step rule, nucleation, Monte Carlo Simulation, Potts model, interfacial energy

## Data-driven imaging of seismic wave field in the Tokyo metropolitan area based on lasso

MIZUSAKO, Sadanobu<sup>1</sup> ; NAGAO, Hiromichi<sup>1\*</sup> ; HIROSE, Kei<sup>2</sup> ; KANO, Masayuki<sup>1</sup> ; HORI, Muneo<sup>1</sup> ; SAKAI, Shin'ichi<sup>1</sup> ; NAKAGAWA, Shigeki<sup>1</sup> ; HONDA, Ryou<sup>3</sup> ; KIMURA, Hisanori<sup>4</sup> ; HIRATA, Naoshi<sup>1</sup>

<sup>1</sup>Earthquake Research Institute, The University of Tokyo, <sup>2</sup>Graduate School of Engineering Science, Osaka University, <sup>3</sup>Hot Springs Research Institute of Kanagawa Prefecture, <sup>4</sup>National Research Institute for Earth Science and Disaster Prevention

Rapid prediction of damage due to large earthquakes on constructions such as buildings and bridges through a numerical simulation that computes seismic responses of all constructions provides important information for making a decision of priority order relating to rescue and rehabilitation activities without waiting for on-site reports. Such a simulation requires ground motion input to each construction, which usually distributes much denser than seismometers. An estimation of the ground motion should rely on data-driven modeling from observational data because of insufficient information related to hypocenters or underground structure models needed in a numerical simulation of seismic wave propagation.

We propose a new methodology based on lasso (least absolute shrinkage and selection operator) for the purpose of data-driven imaging of seismic wave field in an urban area from seismograms obtained by a dense array. The target of this study is the Tokyo metropolitan area, in which MeSO-net (Metropolitan Seismic Observation network) consisting of 296 accelerometers has been established. We assume that a Taylor's expansion model is capable of estimating the seismic wave field at an arbitrary point from seismograms at nearby observatories. This model is reduced to a multivariate linear regression of which the matrix of predictors is a set of products of positions of the observatories, the outcome matrix consists of the observational data, and the coefficient matrix to be estimated consists of the partial derivative coefficients. The coefficient matrix is estimated by group lasso, which minimizes an evaluation function consisting of a loss function given as the squared discrepancy between the observation data and the regression model, and a regularization term based on the L2 norm for selection of grouped variables. The reason why we implement group lasso rather than ordinary lasso is that it can avoid the dependency of subjective coordinate settings in time and space.

We have applied the proposed method to synthetic seismograms obtained from an analytic solution of seismic wave field assuming that an earthquake of M7 class occurs in a horizontally-layered underground. The resulting image reproduced better than those obtained by other methods such as the ordinary least square, ridge regression, and lasso. We have also applied these methods to the actual MeSO-net data when the 2011 Earthquake off the Pacific coast of Tohoku occurred, and confirmed that group lasso reproduced the actual seismic wave field better than other methods.

Keywords: urban mega earthquake, MeSO-net, sparse modeling, lasso, imaging, regularization

## AIC model selection in microboudin palaeo-piezometer

MATUMURA, Taroujiro<sup>1\*</sup> ; MASUDA, Toshiaki<sup>2</sup>

<sup>1</sup>Graduate school of science and Technology, Shizuoka University, <sup>2</sup>Institute of geosciences, Shizuoka University

Microboudin structure is a deformation structure of rocks at microscopic scale. It is formed as the fracture of mineral grains and their subsequent separation within the matrix minerals, which accompanies with infiltration of matrix minerals into the inter-boudin gaps by flowing. In particular case in which microboudin structure consists of columnar minerals such as tourmaline, amphibole, and piemontite embedded within quartz or calcite matrix minerals, it can be utilized as a palaeo-piezometer if a set of data (proportion of boudinaged columnar mineral grains with respect to the their aspect ratio) is supplied. Estimation of the magnitude of palaeo-stress in the rock is generally a difficult subject, although many geologists have been recognized its importance. The microboudin palaeo-piezometer can estimate the magnitude of stress with increasing strain of rocks under metamorphic conditions. However, flow of rocks is generally treated by the viscous model with a Newtonian material. In this presentation, we measure the relative quality of both elastic and viscous models by using the Akaike Information Criterion (AIC) for selecting model. The set of data for the palaeo-piezometer is given by measurements of total 6171 tourmaline grains within 9 quartzoze metamorphic rocks, which are collected from the greenstone belt around a granite complex in East Pilbara, Western Australia. It was clarified that AIC shows that the elastic model is preferable to the viscous one.

Keywords: Microboudin structure, palaeo-piezometer, AIC, Model selection

## Application of principal component analysis to seafloor basalt and its quantitative interpretation

UNO, Masaaki<sup>1\*</sup> ; KUWATANI, Tatsu<sup>1</sup> ; UEKI, Kenta<sup>2</sup>

<sup>1</sup>Graduate School of Environmental Studies, Tohoku University, <sup>2</sup>Japan Agency for Marine-Earth Science and Technology, Department of Solid Earth Geochemistry

Bulk chemical compositional data of rocks consist of tens of elements, and analysis of high dimensional data is needed for interpretation of such data sets. In the traditional studies, a certain pairs of elements were empirically chosen to represent specific processes such as degree of melting and effect of crystal differentiation. To elucidate hidden processes without using empirical knowledge, principal component analysis (PCA) has been utilized for the interpretation of the compositional variations of sediments and igneous rocks (Kuwatani et al., 2014; Ueki, 2014). For example, application of PCA to compositions of arc basalt has revealed that principal components differ among volcano, and distinguishes processes such as mixing and crystal differentiation (Ueki, 2014). To further develop the application of PCA to the interpretations of bulk rock compositions, a quantitative evaluation of contribution rates and factor loadings are needed. In this study, to establish a quantitative interpretation of factor loadings, bulk compositions of seafloor basalt are chosen for PCA, and the results were quantitatively compared with theoretical fractionation processes.

Sample data sets were from seafloor basalt collected from Cocos plate by IODP expedition 334. The samples were collected from the depth of 99-164 meter below seafloor (mbsf). As these samples were collected from narrow region, it is expected that the processes controlling the compositions of these basalt are relatively simple. Empirical factors such as Nb/Zr and Cr/Zr ratios suggest that the compositional variations of the samples originate from different degree of melting and crystal differentiations (Uno et al., in prep.).

Principal component analysis was applied to the compositions of 43 basalt samples for 10 major elements (Si, Ti, Al, Fe, Mg, Mn, Ca, Na, K, P), and 33 minor elements (Li, Sc, V, Cr, Co, Ni, Cu, Sr, Rb, Y, Zr, Nb, Cs, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, Hf, Ta, Pb, Th, U).

The first three principal components account for 80% of the data variance: the contribution factor of PC1 is 52%. The factor loading is high for Middle REE (MREE) and low for incompatible elements (i.e., elements that partition preferentially to silicate melt compared to rock) such as Rb and Ba. The PC2 has contribution factor of 19%, and factor loadings are high for Ni, Co and heavy REE (HREE), whereas those are low for incompatible elements such as Rb, Ba and Th. As the factor loadings of PC2 change continuously from high values for compatible elements to low values for incompatible elements, the PC2 suggest fractionation of melt. The PC3 contributes 8% of the data variance, and factor loadings are high for Cs, Li, U, and loss on ignition (i.e., bulk water content). As solubility for aqueous fluids are high for these elements, the PC3 suggests chemical alteration by aqueous fluids.

The depth profiles of PC1-3 are consistent with the empirical factor of Nb/Zr, Cr/Zr ratios and the amount of altered minerals, respectively. These results suggest that the bulk compositions of seafloor basalt of the Cocos plate are controlled by crystal differentiation, degree of melting and hydrothermal alteration.

For the quantitative interpretations of the PCA, the factor loadings of each PCA were compared with the various kinds of partition coefficients. For PC2, the patterns of factor loadings can be quantitatively explained by the partition coefficients of basaltic melt/peridotite. On the other hand, PC1 cannot be explained by simple melt/mineral partitioning, but can be explained quantitatively by olivine/pyroxene partitioning. These quantitative relations between factor loadings and partition coefficients show the effectiveness of PCA for elucidating fractionation processes.

In this presentation, the cause of quantitative relation between the factor loadings and fractionation processes will be discussed.

Keywords: principal component analysis, seafloor basalt, elemental differentiation, Costa Rica

## Distribution mechanism of heavy metals revealed by correspondence analysis in the Natori River, Send

NAKAMURA, Kengo<sup>1\*</sup> ; KUWATANI, Tatsu<sup>1</sup> ; KAWABE, Yoshishige<sup>2</sup> ; KOMAI, Takeshi<sup>1</sup>

<sup>1</sup>Tohoku University Graduate School of Environmental Studies, <sup>2</sup>National Institute of Advanced Industrial Science and Technology

It is important to understand the distribution mechanism of heavy metal elements in environmental systems, because the toxicities of various metals and metalloids significantly affect health risks. The heavy metals concentrations are controlled by a wide of variety environmental factors. This study aims to clarify main influence factors on the heavy metal distribution in the sediment, soil, river in the Natori River, Sendai City.

We firstly produced geochemical maps for heavy metal concentrations (As, Pb and Cr) of river sediment, neighboring soil (bulk component, soluble component in 1N HCl and in pure water) and river water. The heavy metals concentrations in river water increase from upper stream to down stream, whereas heavy metal concentrations in river sediments and soil vary little from upper stream to downstream. There are no anomalous distributions of heavy metals in the studied area, although many anthropogenic materials are found in sediment, soil and river water.

The correspondence analysis was used to investigate relationships of heavy-metal concentrations between sediment, soil (bulk component, soluble component in 1N HCl and in pure water) and river water. The results indicate that the heavy metals (As, Pb and Cr) in sediment, soil and river water are dominantly controlled by natural sedimentary processes, such as denudation and sedimentation. The proposed method could be further applied for other pollutants in various environment systems.

Keywords: geochemical map, heavy metals, factor analysis, risk assessment

## Spatial modeling of metal contents in a kuroko-type deposit with application to estimating ore-solutions paths

UCHIKURA, Risa<sup>1</sup> ; KUBO, Taiki<sup>1</sup> ; LIU, Chunxue<sup>2</sup> ; KASHIWAYA, Kouki<sup>1</sup> ; KOIKE, Katsuaki<sup>1\*</sup> ; YAMADA, Ryoichi<sup>3</sup>

<sup>1</sup>Graduate School of Eng., Kyoto Univ., <sup>2</sup>School of Urban and Environment, Yunnan Univ. of Finance and Economics, <sup>3</sup>Graduate School of Sci., Tohoku Univ.

Demand for metal resources has largely increased because of constructing sustainable society and innovative technologies. This trend will continue in the future, which requires accurate characterization of ore grade in a metal deposit and deep understanding of ore generation mechanism for exploring new deposits. Based on that background, this study is aimed to (i) develop a method for highly precise spatial modeling of metal contents in a metal deposit and (ii) clarify a physical law that formed a deposit. Matsumine mine, the largest kuroko deposit in the Hokuroku district, Akita Pref., northern Japan, is selected as a case study site. Kuroko is a Japanese term that was used originally in the mining industry for massive, compact black-ore. It mainly contains sphalerite, galena, and pyrite. Kuroko deposits are massive sulfide deposits associated with felsic to intermediate submarine volcanic activity (e.g. Yamada and Yoshida, 2013).

Geological columns at 77 sites and the metal contents of Cu, Zn, and Pb (chief metals of kuroko) at 1457 measurement points were used for the spatial analyses over a region of 500 m×1000 m in the horizontal direction and 300 m along the vertical direction from 0 to -300 m a.s.l. 3D geological model was produced by classifying the columns into ten main rock types and calculating their appearance probabilities at each grid point using 3D optimization principle method (Koike et al., 1998). Geo-statistics, known as the best linear unbiased estimator, was adopted for 3D modeling of metal contents. The result of variography clarified that, common to the three metals, the spatial correlation structures of the metal contents was well approximated by the spherical model. Correlation length (range) along the horizontal direction is longer by a factor of 2 to 3 than that of the vertical direction, which implies the horizontal extents of ore solutions. Ordinary kriging (OK) was selected as a spatial modeling method, because its estimation accuracy was higher than a multivariate kriging.

Through an integration of 3D geological and metal-content models, high content zones were revealed to be overlapped with the patchy kuroko zones and extend horizontally as connecting the zones. Secondary sedimentation of kuroko ores by small-scale volcanic activity and subsequent submarine landslide is one possible cause of this extent. Assuming that the transport of ore solutions and the deposition of metals are approximated by an advective-diffusion spread phenomenon, the advective velocities and the diffusion coefficients were calculated at each grid point by a combination of the metal content model and an advection-diffusion equation. One noteworthy feature detected is that the advective velocities are mainly directed upward and the grid points with this direction are continued in the distributions of silicious ore and rhyolite. This may imply the main paths of ore solutions. The biased signs of vertical diffusion coefficients, 2 to 1 is considered to originate from the general trend of metal contents, which is higher in the shallow parts than the deep parts. Our next step is to improve the metal content estimation by combining a physical law of the deposit formation, which may be expressed by the above advection-diffusion equation, and kriging.

**Acknowledgments:** The authors wish to express their gratitude to Dowa Metals & Mining Co., Ltd. and Hanaoka Eco-System Co., Ltd. for providing the precious drilling investigation materials and cooperating to organize them.

### References

Koike, K., Shiraishi, Y., Verdeja, E. and Fujimura, K (1998) Three-dimensional interpolation and lithofacies analysis of granular composition data for earthquake-engineering characterization of shallow soil, *Mathematical Geology*, v. 30, no. 6, p. 733-759.

Yamada, R. and Yoshida, T. (2013) Kuroko deposits and related back-arc volcanism in the Hokuroku district, *Jour. Geol. Soc. Japan*, v. 119, Supplement, p. 168-179.

Keywords: Kuroko, Geostatistics, kriging, spatial modeling, advection-diffusion equation, Matsumine mine



## Bayesian analysis of nonlinear dynamics in heterogeneous reactions – toward understanding of rock-forming processes –

OMORI, Toshiaki<sup>1\*</sup> ; KUWATANI, Tatsu<sup>3</sup> ; OKAMOTO, Atsushi<sup>3</sup> ; HUKUSHIMA, Koji<sup>2</sup>

<sup>1</sup>Kobe University, <sup>2</sup>The University of Tokyo, <sup>3</sup>Tohoku University

Heterogeneous reactions, which occur at interface between two phases, play important roles in rock-forming processes. In heterogeneous reactions, dynamics are intrinsically nonlinear, because surface area governs actual reaction rates and changes temporally. In general, even static surface area, not to mention the temporal change of surface area, is very difficult to be measured. Therefore, it has still been difficult to determine kinetic parameters of heterogeneous reactions even by well-designed laboratory experiments, because we can only use nonlinear and incomplete data sets.

In this study, we develop novel methodologies which estimate kinetic parameters from nonlinear dynamics of heterogeneous reactions using a Bayesian probabilistic approach. With an increasing demand of data-driven analysis, the Bayesian estimation has been widely applied to various fields in the natural sciences including physics, medicine, brain science, and earth sciences (e.g. Watanabe et al., 2009; Naruse et al., 2009; Shouno and Okada, 2010; Omori et al., 2013; Kuwatani et al., 2014; Omori, 2014). In this study, we first formulate probabilistic models of nonlinear dynamics of heterogeneous systems based on state space model. Sequential Monte Carlo Method and EM algorithms are employed for the proposed state-space model in order to simultaneously estimate time course of heterogeneous systems and kinetic parameters underlying nonlinear dynamics. The proposed method provides a very suitable framework for inversion analysis of nonlinear dynamics from incomplete datasets.

In this study, we focus a simple heterogeneous reaction, in which the solid reactant changes to solid product via intermediate product dissolved in fluid phase. It can be regarded as one of the most fundamental and substantial heterogeneous reactions for solid-fluid interactions. Using our proposed method, the reaction rate constants of dissolution and precipitation as well as the temporal changes of solid reactants and products were successfully estimated only by the observable temporal change of the dissolved intermediate product. The proposed method can be potentially applied to actual laboratory experiments of heterogeneous kinetics in various fields of earth and planetary sciences.

Keywords: data analysis in geoscience, data-driven approach, Bayesian statistics, nonlinear dynamics, heterogeneous reaction, rock-forming processes

## Deconvolution analysis of reflectance spectra of synthetic clinopyroxenes using the exchange Monte Carlo method

HONG, Peng<sup>1\*</sup>; MIYAMOTO, Hideaki<sup>1</sup>; NAGATA, Kenji<sup>2</sup>; SUGITA, Seiji<sup>3</sup>; NIIHARA, Takafumi<sup>1</sup>; JAMES, Dohm M.<sup>1</sup>; HENMI, Ryodo<sup>1</sup>; OKADA, Masato<sup>2</sup>

<sup>1</sup>Univ. Museum, Univ. of Tokyo, <sup>2</sup>Complexity Sci. & Eng., Univ. of Tokyo, <sup>3</sup>Earth & Planetary Sci., Univ. of Tokyo

Deconvolution analysis of reflectance spectra has been a useful method to infer mineral composition and crystal structure. Clinopyroxene (CPX) is one of the most important mineral groups due to both its rich abundance on solid bodies in the solar system and distinguished absorption features. The band centers of visible to near-infrared (Vis/NIR) spectra of CPX are known to vary due to total iron and calcium content. Klima et al. (2011) characterized the Vis/NIR spectra of synthetic CPX based on a modified Gaussian Model (MGM). The numerical algorithm of the widely used MGM (Sunshine et al., 2011), however, utilizes the steepest descent method, which has a local minima problem. With bad initial parameters, the steepest descent method converges into a local minimum, thus the analyzer needs to manually adjust initial parameters and calculate the model repeatedly to obtain the desired solution. In order to avoid the local minimum problem, we analyzed Vis/NIR spectra of CPX using Bayesian spectral deconvolution with the exchange Monte Carlo method (Nagata et al., 2012). This method is an improved algorithm of the Markov chain Monte Carlo method, aimed to both avoid local minima traps and remove the arbitrariness originated from initial parameters. We used the Vis/NIR spectra of CPX, measured by Klima et al. (2011) at 5 nm intervals over the wavelength range of 0.3-2.6  $\mu\text{m}$ . We chose only sieved powder samples whose grain sizes are smaller than 45  $\mu\text{m}$  with individual grains about 15-25  $\mu\text{m}$  in size. We collected 31 CPX spectra with wide ranging Ca, Mg, Fe compositions (8-52%, 0-52% and 3-90%, respectively). We found that: (1) 1  $\mu\text{m}$  band shifts regularly to longer wavelengths with an increase of Ca content; (2) 1  $\mu\text{m}$  band shifts to shorter wavelengths with an increase of Fe content, although the variance is larger than that of orthopyroxene; (3) for Ca contents <30%, 2  $\mu\text{m}$  band shifts to longer wavelength as a function of Ca, while for Ca contents >30%, the center of 2  $\mu\text{m}$  band remains almost constant at  $\sim 2.4 \mu\text{m}$ ; and (4) the center of 2  $\mu\text{m}$  band of augites (Ca <20%) does not depend on Fe content significantly, while the center of 2  $\mu\text{m}$  band of pigeonites (Ca >20%) shifts to longer wavelengths as a function of Fe content. These results are consistent with Klima et al. (2011), suggesting that the exchange Monte Carlo method can yield the same results obtained by conventional MGM analysis. The successful application of the exchange Monte Carlo method to wide range of CPX would pave the way for further deconvolution analysis of reflectance spectra of mineral mixtures, such as olivine, orthopyroxene, and clinopyroxene.

## Study on direction finding method based on sparsity

OTA, Mamoru<sup>1\*</sup> ; KASAHARA, Yoshiya<sup>1</sup> ; GOTO, Yoshitaka<sup>1</sup>

<sup>1</sup>Kanazawa University

Investigating characteristics of plasma waves observed by scientific satellites in the Earth's magnetosphere is the effective key to understand not only generation mechanisms of the waves but also a plasma environment which influences its generation and propagation conditions. In particular, direction finding of the waves is important for understanding the propagation characteristics of VLF waves. In order to find the directions, the wave distribution function (WDF) method was proposed. This method can estimate direction of arrival for multiple waves included in electromagnetic waves observed by the satellite, and the estimated direction is represented as wave energy density distribution (WDF). For the estimation of WDFs, this method uses dispersion relation for the plasma surrounding the satellite. However, this estimation is an ill-posed problem and cannot determine a unique solution. In order to obtain a unique solution, therefore, we need to give some prior information. In previous work, some models were proposed as the prior information, and its performance was evaluated by using computer-generated data.

In the present study, we propose a new method using a sparsity assumption of the WDF. This assumption expresses that the effective direction of arrival to explain the data is a smaller number, therefore, it can be expected that the reconstructed WDF by the new method is more precise than conventional methods. In conventional evaluations, the spectral matrix which is used as the input data of the estimation was assumed precisely calculable. However, in the case of the actual observation data, there is a problem about how to calculate precise spectral matrix. We examined the problem that arises in the case that the WDF method is applied for the actual observation data.

Keywords: Waves in plasma, Ill-posed problem, Wave distribution function, Sparsity

## Data-driven spatial modeling of frictional features at plate subduction zones

KANO, Masayuki<sup>1\*</sup> ; SUZUKI, Akihiro<sup>2</sup> ; NAGAO, Hiromichi<sup>1</sup> ; KOMAKI, Fumiyasu<sup>2</sup>

<sup>1</sup>Earthquake Research Institute, The University of Tokyo, <sup>2</sup>Graduate School of Information Science and Technology, The University of Tokyo

Frictional properties at a plate boundary are considered to control the time evolution of fault slips, so that clarification of their spatial distribution is one of the major issues to predict the states in the Earth's crust.

Recently, various aspects of the whole earthquake generation such as intervals of occurrence times, interseismic tectonic loading, afterslips, and episodic slow slips, were qualitatively reproduced, empirically giving the frictional parameters in the rate and state friction law [e.g., Kato and Yoshida (2011), Hori and Miyazaki (2011)]. For a more realistic simulation, the frictional parameters should be quantitatively determined based on observational data and theoretical prior information. Data assimilation (DA) is a computational technique based on the Bayesian statistics to integrate numerical simulation models and observational data [Higuchi et al. (2011)], which is widely used in geoscience including the solid earth science [e.g., Nagao et al. (2013)]. DA has also been applied to clarify the frictional features at plate boundaries, which are considered to control postseismic phenomena, estimating the frictional parameters in afterslip regions [e.g., Fukuda et al. (2009), Mitsui et al. (2010), Kano et al. (2013), Kano (2014)]. These previous studies assumed that the frictional parameters were unrealistically uniform in the entire fault region or subjectively divided the afterslip region into several areas in each of which the frictional feature is uniform in order to avoid substantial computational cost due to too much high-resolution spatial grids never to be determined by the limited observations on the Earth's surface. Therefore, it is meaningful to develop a method to divide the region appropriately in an automatic and objective way. We propose a data-driven procedure consisting of the  $k$ -means-based clustering method to obtain candidate division patterns in the afterslip region and the Akaike's Information Criterion to determine the optimum model among the candidates. We have confirmed that the model obtained by an application to synthetic data is almost the true one. We will report results when the proposed method is applied to the case of the afterslip region of the 2003 Tokachi-oki earthquake. This method will help to extract the large-scale frictional features and make relevant simulations more effective, objective and realistic.

## Principal Component Analysis based determination of chemical differentiation processes of volcanic rocks

UEKI, Kenta<sup>1\*</sup>

<sup>1</sup>Japan Agency for Marine-Earth Science and Technology

Bulk chemical compositions of igneous rocks represent sums of a series of chemical processes during generation and migration of magmas. Various processes from the mantle to the crust in various pressures, temperature and chemical composition modify magma composition (e.g., Annen et al., 2006). Multivariable analysis can be used to investigate such reactions involving multi-element and multi process.

Differentiation in terms of major element composition is controlled by non-linear thermodynamic relation (e.g., Ueki and Iwamori, 2014). As such, major element composition can be a proxy for the pressure and temperature of the source region of magma. However, major element composition is often overwritten by shallower processes because it exhibits a relatively restricted degree of freedom in terms of composition. On the other hand, trace element shows a large degree of freedom. In addition, the concentration of a specific trace element can be a tracer of a reaction in which a specific phase is involved, because the partition coefficient of a specific element in phases show wide range of variation. Moreover, partition of trace elements between phases exhibits a weak non-ideality, and can be modeled with a simple equation. Consequently, trace element concentrations can be used as a proxy of a specific phase and chemical mass reaction process during the formation of the rock (e.g., Depaolo, 1981; Pearce et al., 2005).

In this study, principal component analysis is used to analyze compositional variations of volcanic rocks in Northeastern Japan Arc. 14 trace elements of 262 samples including basalt to rhyolite, sampled from 17 different volcanoes in a volcanic cluster of northeastern Japan, called the Sengan region, are processed with principal component analysis.

The result of the principal component analysis shows that the only 3 processes are enough to cover the ~90% of the geochemical variation of the magmas of the Sengan region. It is estimated that the three principal components represent magma mixing, crystallization-fractionation of mafic minerals at deep crustal depth and crystallization-fractionation of plagioclase at shallower crustal depth, respectively. This result shows that intermediate-felsic magmas ( $\text{SiO}_2 > 60$  wt. %) can only be derived through magma mixing, not by crystal fractionation. No mantle signature is observed; indicating that differentiation processes in the arc crust, are the primary controlling factor to derive compositional variation in terms of trace element. The first principal component (magma mixing) shows good correlations with petrological observation (core composition of plagioclase phenocryst), and geophysical observations (distribution of seismic low velocity zone in the crust; Nakajima et al., 2001 and geothermal gradient; Tamanyu, 1994), indicating consistent description between the trace element principal components, major element composition, and petrological information such as mineral composition can be derived from the analysis.

Keywords: volcanic rock, multivariable analysis, magma, arc, Northeastern Japan

## Bayesian inversion with sparsity constraint for spatial distribution of afterslip

NAKATA, Ryoko<sup>1\*</sup>; KUWATANI, Tatsu<sup>2</sup>; OKADA, Masato<sup>3</sup>; HORI, Takane<sup>1</sup>

<sup>1</sup>JAMSTEC, <sup>2</sup>Tohoku University, <sup>3</sup>University of Tokyo

Heterogeneous distribution on the plate interface is considered to be the manifestation of heterogeneity in plate coupling, which may slip as earthquake sources in the future. However, it seems to be difficult to resolve clearly such heterogeneous distribution for M7 class earthquakes in offshore region because of stations only on land and narrow expected heterogeneous area.

We demonstrated the reproducibility of spatial distribution of afterslip following a M<sup>7</sup> earthquake through numerical experiments, which estimate slip distribution on the plate interface from displacements observed on the free surface (Nakata et al., 2014). We calculated synthetic displacement data from the result of numerical simulation conducted for the afterslip following a M 6.8 earthquake, for existing global navigation satellite system stations on land (GEONET) and planned ocean floor pressure gauge network stations (DONET). The spatial distribution of fault slip is estimated using a Kalman filter-based inversion. The slip distribution estimated by using ocean floor stations demonstrated the heterogeneity of plate coupling within the coseismic area of the M 6.8 earthquake with a radius of 10 km. The estimated slip amount in the coseismic area is nearly half of the peak one around it, although no slip is the true answer. This discrepancy is caused by the smoothness constraint in the inversion.

To improve the reproducibility of the slip distribution, it is necessary to introduce different type of constraints. We have incorporated three constraints into the evaluation function as follows: First is continuous and smooth constraint in non-zero slip area. Second is the sparsity constraint that the number of subfaults with non-zero afterslip is significantly smaller than that of residual area. Third constraint is discontinuity between locked area (slip = 0) area and afterslip area.

By developing a method of Kuwatani et al. (2014), we proposed an evaluation function combined with Markov random field (e.g., Kindermann & Snell, 1980; Geman & Geman, 1984) model and sparse modeling (e.g., Tibshirani, 1996). Obtained afterslip distribution that minimizes the evaluation function could be resolved heterogeneity due to the M<sup>7</sup> seismic source sharper than that using previous methods.

## Cluster analyses on bulk elemental compositions of meteorites

MIYAMOTO, Hideaki<sup>1\*</sup> ; NIIHARA, Takafumi<sup>1</sup> ; KURITANI, Takeshi<sup>2</sup> ; HONG, Peng<sup>1</sup> ; SUGITA, Seiji<sup>1</sup> ; JAMES, Dohm<sup>1</sup>

<sup>1</sup>University of Tokyo, <sup>2</sup>Hokkaido University

Meteorites are now determined to be precious samples from extraterrestrial bodies and thus identifications of the corresponding meteorite types of the targeted bodies of asteroid missions are among the major goals of asteroid missions. Besides such scientific importance, rapid identification of the surface materials during the reconnaissance phase is also critically important even for maneuvering a spacecraft. Modern spacecraft carry cameras/spectrometers in the visible to infrared wavelengths, which are powerful tools in identifying surface materials. However, irradiation by cosmic and solar wind ions as well as bombardment by interplanetary dust particles modify the surface of airless bodies through processes known as space weathering. Impact events also mix materials at the surface of the body. These processes may flatten or change the absorption characteristics of reflectance spectra. In this sense, elemental compositions, which can be obtained by X and gamma-ray spectrometers, may be useful for the above purpose. However, it has not been investigated extensively how well we can classify these planetary materials based on elemental composition alone. In this study, we perform principal component and cluster analyses on 12 major and minor elements of the bulk compositions of 500 meteorites reported in the NIPR database. Our unique approach, which includes using hierarchical cluster analysis, indicates that meteorites can be classified into about 10 groups purely by their bulk elemental compositions. We suggest that Si, Fe, Mg, Ca, and Na are the optimal set of elements, as this set has been used successfully to classify meteorites of the NIPR database with more than 94% accuracy. Principal components analysis indicates that elemental compositions of meteorites form 8 clusters in the 3 dimensional space of the components. The three major principal components (PC1, PC2, and PC3) are interpreted as degree of differentiations of the source body (i.e., primitive vs. differentiated), degree of thermal effects, and degree of chemical fractionation, respectively. Though the exact ranges of elements of each cluster suffer from the systematic intra-laboratory error, realized through comparing our results with those of another elemental composition database, our new method shows promise in the classification of the surface materials of a small body into a known group of meteorites, having a significant bearing in future reconnaissance.

Keywords: meteorite, bulk composition, cluster analysis, asteroid, XRF

## Application of advanced statistical analysis in geoscience and its evolution

KOMAI, Takeshi<sup>1\*</sup> ; KUWATANI, Tatsu<sup>1</sup>

<sup>1</sup>Tohoku University

An innovative methodology using sparse modelling as Bayes' theorem for discrimination between tsunami deposits and marine stratum is significantly important, because of its identification of sedimentation origin and evidences of past tsunami event frequency. The authors started an interdisciplinary research area for innovative modelling in the field of geoscience. One of the topics is the geo-statistical sparse modelling of tsunami deposits and soils using metal elements. Sparse modelling is based on fundamental principle of sparseness; most important information is embedded in the low-dimensional subspace for high-dimension data.

Keywords: geoscience, statistical analysis, sparse modeling