Data-driven geoscience: concept and applications

T. KUWATANI[1]*, K. NAGATA[2], M. OKADA[2] AND T. KOMAI[3]

[1]Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka 237-0061, Japan (correspondence: kuwatani@jamstec.go.jp)
[2]University of Tokyo, Kashiwa 277-8561, Japan
[3]Tohoku University, Sendai 980-8579, Japan

High-dimensional data sets in earth sciences generally show very complex behavior and often have large uncertainty. In order to approach the dynamic behavior of the earth, it is significant to select a small number of essential parameters which can explain the target phenomenon from such high-dimensional data.

Recently, we have launched a big scientific project entitled as “Initiative for high-dimensional data-driven science through deepening sparse modelling” financially aided by the Ministry of Education, Culture, Sports, Science and Technology in Japan. The main object of the project is to develop data-driven methodologies for understanding the world of nature by tight fusion of information science and natural sciences. The project consists of several tens of research teams; their target fields are very wide including geosciences, astronomy, biology, medicine and brain science.

We consider that ‘Bayesian estimation’ and ‘sparse modeling’ are key technologies of data-driven analyses. Bayesian estimation is a probabilistic inversion scheme by incorporating forward models and prior information through the Bayes’ theorem. Sparse modeling is a mathematical framework which can effectively extract a small number of essential explanatory variables from high-dimensional data sets. It is based on the fundamental principle of sparseness: most of useful information is embedded in the low-dimensional subspace for high-dimensional observation data in the various fields of natural science.

In this contribution, we will introduce the concept of data-driven analyses with some applications to petrology [1], geochemistry [2, 3] and geophysics [4, 5, 6].

[1] Kuwatani et al. (2012) *Contrib. Min. Petrol.* **163**, 547-562.
[2] Kuwatani et al. (2014) *Sci. Rep.* **4**, 7077-1-6.
[3] Nakamura et al. (2016) *Chemosphere* **144**, 1241-1248.
[4] Kuwatani et al. (2014) *Earth Planet. Space* **66**, 5-1-9.
[5] Kuwatani et al. (2014) *Phys. Rev. E* **90**, 042137-1-7.
[6] Nakata et al. (2016) *Earth Planet. Space* **68**, 20-1-10.