

Data-driven approaches to frontier of earth and planetary sciences

KUWATANI, Tatsu^{1*} ; NAGATA, Kenji² ; OKADA, Masato² ; KOMAI, Takeshi¹

¹Graduate School of Environmental Studies, Tohoku University, ²Graduate 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*).