

A Content-Based Database System for Large Ensemble Future Climate Simulations

*Yujin Nakagawa¹, Shintaro Kawahara¹, Fumiaki Araki¹, Daisuke Matsuoka¹, Yoichi Ishikawa¹, Mikiko Fujita¹, Shiori Sugimoto¹, Yasuko Okada¹, Sho Kawazoe¹, Shingo Watanabe¹, Masayoshi Ishii², Ryo Mizuta², Akihiko Murata², Hiroaki Kawase²

1. Japan Agency for Marine-Earth Science and Technology, 2. Meteorological Research Institute

Analyses of large ensemble data of future climate are quite useful in order to produce probabilistic future projection of climate change in various interdisciplinary fields related to, e.g., hydrology, civil engineering, adaptation planning for global warming. Ensemble data of "+2K future climate simulations" are currently generated by "Social Implementation Program on Climate Change Adaptation Technology" (SI-CAT) as a part of a database for Policy Decision making for Future climate change (d4PDF; Mizuta et al. 2016) produced by Program for Risk Information on Climate Change. Those data consist of global warming simulations and regional downscaling simulations around Japan. Considering that those data volumes are too large, i.e. a few petabytes, to download to a local computer of users, a user-friendly system is required to search and download data which satisfy requests of the users. In order to provide functions to extract necessary data from +2K future climate simulations to the users, we developed "System for Efficient content-based retrieval to Analyze Large volume climate data (SEAL)". Existing web-based search systems for climate simulations are designed to find data files using metadata associated with a data file itself. On the other hand, the SEAL allows the users to extract necessary data using metadata associated with contents of a data file such as physical values. The SEAL mainly consists of a relational database, a data download function and web-based user interface. Among the three main features, a key role is provided by the relational database using PostgreSQL which is designed to be registered temporally and spatially compressed data. The data download function allows the users to download temporally and spatially extracted data based on search results obtained by the relational database. In addition, the web-based user interface allows the users to easily use the relational database without knowledge about PostgreSQL. We also prepare a guidebook for usage of the SEAL. The SEAL is currently in test operation on our local server for SI-CAT members. The SEAL and the guidebook will be released on Data Integration and Analysis System Program (DIAS) after end of fiscal year 2018. Techniques of the SEAL might be quite useful for simulation, experimental and observational data in other research fields. We report functions and some case studies of the SEAL.

Keywords: Climate Change, Relational Database