Chapter 16

Data Assimilation Research Team

16.1 Members

Takemasa Miyoshi (Team Leader)
Koji Terasaki (Research Scientist)
Shigenori Otsuka (Research Scientist)
Shunji Kotsuki (Research Scientist)
Keiichi Kondo (Postdoctoral Researcher)
Guo-Yuan Lien (Postdoctoral Researcher)
Takumi Honda (Postdoctoral Researcher)
Atsushi Okazaki (Postdoctoral Researcher)
Toshiki Teramura (Postdoctoral Researcher)
Kohei Tatakama (Postdoctoral Researcher)
James Taylor (Postdoctoral Researcher)
Yasumitsu Maejima (Research Associate)
Hazuki Arakida (Technical Staff)
Taeka Awazu (Technical Staff)
Hideyuki Sakamoto (Technical Staff)
Marimo Ohhigashi (Technical Staff)
Shinichiro Shima (Visiting Scientist)
Juan J. Ruiz (Visiting Scientist)
Shu-Chih Yang (Visiting Scientist)
Stephen G. Penny (Visiting Scientist)
Kozo Okamoto (Visiting Scientist)
Michiko Otsuka (Visiting Scientist)
Yohei Sawada (Visiting Scientist)
Shohei Takino (Visiting Scientist)
16.2 Research Activities

Data Assimilation Research Team (DA Team) was launched in October 2012 and is composed of 23 research and technical staff including 8 visiting members as of March 2018. Data assimilation is a cross-disciplinary science to synergize computer simulations and real-world data, using statistical methods and applied mathematics. As computers become more powerful and enable more precise simulations, it will become more important to compare the simulations with actual observations. DA Team performs cutting-edge research and development on advanced data assimilation methods and their wide applications, aiming to integrate computer simulations and real-world data in the wisest way. Particularly, DA Team tackles challenging problems of developing efficient and accurate data assimilation systems for “big simulations” with real-world “big data” from various sources including advanced sensors. The specific foci include 1) theoretical and algorithmic developments for efficient and accurate data assimilation, 2) data assimilation methods and applications by taking advantage of the world-leading K computer and “big data” from new advanced sensors, and 3) exploratory new applications of data assimilation in wider simulation fields. These advanced data assimilation studies will enhance simulation capabilities and lead to a better use of the K computer.

In FY2017, we continued on the ongoing data assimilation research in the following aspects: 1) theoretical research on challenging problems, 2) leading research on meteorological applications, 3) optimization of computational algorithms, and 4) exploratory research on wider applications. We also explored close collaborations with several research teams within the AICS Research Division. We have made substantial progress on the following research items:

Theoretical research

- Non-Gaussian PDF in DA was investigated using the Lorenz-63 3 variable model.
- The observation error correlation was investigated with the Lorenz96 model. Reconditioning of observation error covariance matrix was introduced.
- A particle filter was applied to the cellular automaton of 3 state sheep model.
- Impact of assimilation order of the serial EnSRF was investigated (1 paper published).
Leading research on meteorological applications

- The development work of the SCALE-LETKF data assimilation system, which consists of SCALE (Scalable Computing for Advanced Library and Environment) and LETKF (Local Ensemble Transform Kalman Filter), was continued. New functions including the random additive noises, online-nested domain LETKF, and improved observation departure diagnostics were implemented. The computational speed was also continued to be improved.
- The study on the assimilation of Phased Array Weather Radar (PAWR) data with the SCALE-LETKF was continued. A new convective rainfall case on July 16, 2017 was conducted, with which the impacts of assimilation intervals, random additive noises, and a deterministic analysis member on the forecast skill were studied.
- A system to assimilate observations by two PAWRs simultaneously has been investigated.
- Weather radar observations from an Argentinian operational radar were assimilated by using the SCALE-LETKF system.
- Himawari-8 satellite all-sky infrared radiance observations were assimilated with SCALE-LETKF. Case studies on Typhoon Soudelor (2015) and the Kanto-Tohoku rainfall event were conducted (2 papers published).
- Himawari-8 satellite all-sky infrared radiances were assimilated in a typhoon-induced heavy precipitation event in Japan.
- A series of experiments on the Kanto-Tohoku heavy rain event in September 2015 was conducted with the SCALE-LETKF. Various observations including Himawari-8 infrared images and dense surface observations provided by NTT DoCoMo were assimilated. River discharge simulations using the SCALE-LETKF precipitation forecasts were conducted.
- Experiments with NHM-LETKF were conducted to investigate an impact of assimilating Himawari-8 observations on detecting a local severe storm at an early stage of its development.
- An ocean mixed layer model was implemented into SCALE-LETKF to investigate the role of flow-dependent sea surface temperature perturbations in the atmospheric DA system.
- The 3D precipitation nowcasting system with PAWR was operated in real time. Forecasts were disseminated via a smartphone application in collaboration with MTI Ltd. (press release on 4 July 2017).
- The 30-second-update 100-m-mesh DA experiments on a sudden local rainfall event on September 11, 2014 were performed with NHM-LETKF (1 paper published).
- A DA system for Advanced Microwave Sounding Unit (AMSU)-A radiance data was developed with Nonhydrostatic Icosahedral Atmospheric Model (NICAM)-LETKF (1 paper published).
- A new adaptive covariance inflation method was developed and applied to the NICAM-LETKF system (1 paper published).
- An ensemble-based model parameter estimation was investigated with NICAM-LETKF.
- A high resolution NICAM-LETKF was developed to assimilate dual frequency precipitation radar (DPR) of global precipitation measurement (GPM) core satellite.
- The heavy ice precipitation flag of the GPM/DPR was compared with the 3.5-km-resolution NICAM simulations.
- Including the observation error correlation in DA was investigated with the NICAM-LETKF system. We found that reconditioning the observation error covariance matrix stabilizes the data assimilation and improves the analyses.
- The NICAM-LETKF system was upgraded to assimilate microwave humidity sounder data.
- An impact of each observation in the NICAM-LETKF was investigated with the Ensemble Forecast Sensitivity to Observation (EFSO) method.
- Global precipitation nowcasting with the LETKF has been operated in real time using JAXA’s Global Satellite Mapping of Precipitation (GSMaP) (GSMaP RIKEN Nowcast (RNC)). Forecast data in the past two years were analyzed.
- A precipitation forecasting system which merges forecasts by NWP and nowcast has been developed.
- A DA method for dense precipitation radar data was investigated with SCALE-LETKF.
- Non-Gaussian statistics in global atmospheric dynamics were investigated with a 10240-member ensemble Kalman filter using an intermediate AGCM.
- An object-based verification method of precipitation pattern was investigated using pattern recognition techniques.
- An LETKF system to assimilate pattern features of precipitation areas was implemented with an intermediate AGCM.
• An atmosphere-river coupled DA system was developed with NHM-LETKF and a river discharge model. An OSSE assimilating river discharge observation showed a positive impact on the atmospheric variables (1 paper published).
• A prototype of dam operation optimization system was developed with the machine learning in collaboration with Tokyo Electric Power Company Holdings, Incorporated.
• A land-atmosphere-coupled DA system was developed.
• A new methodology to accelerate development of a new observation system was proposed using the ensemble forecast sensitivity to observations (EFSO) with a global numerical weather prediction system (1 paper published).

Computational optimization
• The computational performance of the “Big Data Assimilation” problem with the SCALE-LETKF was improved.
• We started incorporating the Data Transfer Framework (DTF) developed by System Software Research Team into the SCALE-LETKF system to enable the real-time “Big Data Assimilation” experiments.
• A tree-based geographical search algorithm in LETKF was implemented in collaboration with Prof. Hideyuki Kawashima’s group in the University of Tsukuba. The algorithm was implemented with the SCALE-LETKF and SPEEDY-LETKF.

Wider applications
• A particle filter was applied to a dynamical vegetation model known as the SEIB-DGVM (Spatially-Explicit, Individual-Based Dynamic Global Vegetation Model). Uncertainties in the state variables and the parameters were greatly reduced by assimilating satellite-based Leaf Area Index (1 paper published).
• The SEIB-DGVM DA system was applied to a wider region in the northeastern Eurasia to estimate spatially-varying vegetation parameters.
• The SEIB-DGVM DA system was applied to a deciduous broad-leaved forest in Japan to optimize leaf area index and the related variables.
• Feasibility of applying the SEIB-DGVM DA system to vegetation management was investigated toward the sustainable development goals (SDGs).
• A DA method for time-averaged data was investigated with SPEEDY-LETKF for paleoclimate reconstruction.
• A feasibility of predicting coastal ocean environments was investigated using a regional ocean model (ROMS).

Several achievements are selected and highlighted in the next section.

16.3 Research Results and Achievements

16.3.1 Himawari-8 satellite “Big Data Assimilation” for typhoon and heavy-rainfall prediction

Weather prediction models attempt to predict future weather by running simulations based on current conditions taken from various sources of data. However, the inherently complex nature of the systems, coupled with the lack of precision and timeliness of the data, makes it difficult to conduct accurate predictions, especially with weather systems such as sudden precipitation. As a means to improve models, we are using powerful supercomputers to run simulations based on more frequently updated and accurate data. We decided to work with data from Himawari-8, a geostationary satellite that began operating in 2015. Its instruments can scan the entire area it covers every ten minutes in both visible and infrared light, at a resolution of up to 500 meters, and the data is provided to meteorological agencies. Infrared measurements are useful for indirectly gauging rainfall, as they make it possible to see where clouds are located and at what altitude.

For one study, we looked at the behavior of Typhoon Soudelor (known in the Philippines as Hanna), a category 5 storm that wreaked damage in the Pacific region in late July and early August 2015. In a second study, we investigated the use of the improved data on predictions of heavy rainfall that occurred in the Kanto region of Japan in September 2015. These articles were published in Monthly Weather Review and Journal of Geophysical Research: Atmospheres. For the study on Typhoon Soudelor, we adopted a recently
16.3. RESEARCH RESULTS AND ACHIEVEMENTS

Figure 16.1: Himawari-8 band 13 (10.4 μm) brightness temperature analyses (K) for DA experiments (left) without and (middle) with Himawari-8, and (right) corresponding Himawari-8 observation.

Figure 16.2: Analyses (thick) and forecasts (thin) of minimum sea level pressure of Typhoon Soudelor. Red and black curves correspond to the experiments with and without Himawari-8 DA, respectively. Blue curve shows the JMA best track analysis.
CHAPTER 16. DATA ASSIMILATION RESEARCH TEAM

Figure 16.3: Horizontal maps of 12-hour forecast precipitation (mm h\(^{-1}\), previous 1-hour accumulation) for the experiments (left) without and (middle) with Himawari-8 DA, and (c) corresponding JMA radar estimate.

Figure 16.4: River discharge forecasts driven by the rainfall inputs from the experiments with and without Himawari-8 DA. Black curves show the forecasts without Himawari-8 DA, initiated at 0900 JST (solid) and 1500 JST (dashed). Colored curves show the forecasts with Himawari-8 DA, where warmer colors indicate a later initial time corresponding the colors shown at the top between 0900 JST and 1500 JST. Gray curve corresponds to the observed river discharge.
16.3. RESEARCH RESULTS AND ACHIEVEMENTS

Figure 16.5: Time-mean background root mean square differences (RMSDs) and ensemble spreads of temperature (K) at 500 hPa relative to the ERA Interim reanalysis for the global (GL), Northern Hemisphere (NH), tropics (TR), and Southern Hemisphere (SH) domains, averaged over a month in August 2014. Grey, yellow, and green bars show the RMSDs of adaptive-MULT, adaptive-RTPS, and adaptive-RTPP, respectively. Red dots indicate the ensemble spreads. Adopted from Kotsuki et al. (2017, QJRMS).

developed weather model called SCALE-LETKF—running an ensemble of 50 simulations—and incorporated infrared measurements from the satellite every ten minutes, comparing the performance of the model against the actual data from the 2015 tropical storm. We found that compared to models not using the assimilated data, the new simulation more accurately forecast the rapid development of the storm (Figs. 16.1 and 16.2). We tried assimilating data at a slower speed, updating the model every 30 minutes rather than ten minutes, and the model did not perform as well, indicating that the frequency of the assimilation is an important element of the improvement.

To perform the research on disastrous precipitation, we examined data from heavy rainfall that occurred in the Kanto region in 2015. Compared to models without data assimilation from the Himawari-8 satellite, the simulations more accurately predicted the heavy, concentrated rain that took place, and came closer to predicting the situation where an overflowing river led to severe flooding (Fig. 16.3). With every-10-minute Himawari-8 DA, we can refresh precipitation and river discharge forecasts every 10 minutes, i.e., 36 times in 6 hours (Fig. 16.4). The every-10-minute refresh can provide warning information at an earlier time; having a longer lead time by even 10 minutes may save lives. We plan to apply this new method to other weather events to make sure that the results are truly robust. This research result was highlighted by RINEK Press Release on 18 January 2018 (http://www.riken.jp/en/pr/press/2018/20180118_1/).

16.3.2 Improving the error covariance inflation approach in NICAM-LETKF

Covariance relaxation is a widely-used inflation technique, which plays an essential role in the ensemble Kalman filter because the ensemble-based error variance is usually underestimated mainly due to limited ensemble size and model imperfections. To avoid computationally-expensive manual tuning of the relaxation parameter, this study pioneers to propose adaptive covariance relaxation (ACR) approaches based on Desroziers’ innovation statistics (2005, QJRMS). Two ACR methods are implemented: relaxation to prior spread based on Ying and Zhang (2015, QJRMS), and relaxation to prior perturbations. We conduct a series of experiments in the real-world global atmosphere with both conventional observations and satellite radiances for the first time.

The results demonstrate that the proposed ACR approaches provide nearly optimal relaxation parameter values and improve the analyses and forecasts compared to a baseline control experiment with an adaptive multiplicative inflation method (Fig. 16.5). The adaptive relaxation methods are turned out to be robust to changes in the observing networks and observation error settings. We mathematically show that the innovation statistics for the analysis error covariance (a-minus-b o-minus-a statistics are more robust than those for the
background error covariance (\(o - minus-b\) or \(a - minus-a\) statistics) if the observation and background error variances are imperfect.

16.3.3 Phased-Array Weather Radar 3D Nowcast

Today, nowcasting—a term that refers to short-term weather forecasts made in real-time—is generally done using parabolic radar antennas, which take five to ten minutes to scan about 15 layers of the entire sky. Typically, it is done by looking at a single layer of the sky, detecting the rain there, and then extrapolating from weather conditions where the rain will be falling at a later time. However, though nowcasting requires much less computing power than weather simulations, it is not able to accurately predict rainfall from rapidly developing thunderclouds, where there are rapid vertical movements in the rain patterns. Recently, however, a novel type of system, called a phased-array radar, was installed on the Suita Campus of Osaka University. The radar can be targeted very quickly—it can scan the entire sky in ten to thirty seconds, looking at approximately 100 angles with a range of 60 kilometers. The radar can be precisely targeted by manipulating the beams emitted by a number of devices, allowing a flat radar to scan the whole sky very rapidly.

In an effort to improve the forecasting capability of the radar, we developed an algorithm that takes the enormous amount of observational data from the radar, updated every 30 seconds, and makes rapid forecasts based on the 3D rain data. This allows the extremely frequent and accurate forecasts to be made. The forecasts, which cover the Kansai (Osaka, Kyoto and Kobe) area of Japan, are available online at: https://weather.riken.jp/ (Figs. 16.6 and 16.7). As part of the effort to make the system practical, we are collaborating with an app designer that offers weather forecasting apps to smartphone users (see http://www.aics.riken.jp/en/topics/170324.html). We also look forward to collaborating with other phased-array radar facilities to provide forecasts for a more widespread area. This research result was highlighted by RINEK Press Release on 4 July 2017 (http://www.riken.jp/en/pr/press/2017/20170704_1/).

16.4 Schedule and Future Plan

In FY2017, DA team had three additional full-time research staff. We have been working on various aspects of DA including theoretical problems, meteorological applications, and wider applications. We continued working on the prototype system of “Big Data Assimilation” (BDA). In FY2017, we obtained promising results for the Himawari-8 satellite radiance assimilation. In addition, the computational speed of the phased-array weather radar assimilation was also improved. Although it is still not sufficient to run the BDA system in real time on
16.4. SCHEDULE AND FUTURE PLAN

Beyond the direct future of the BDA effort, we can extend the idea of BDA to a broader perspective: integration of “Big Data” and “Big Simulation.” New sensors provide orders of magnitude more data, and simulations become more precise. Collaborative work with computer scientists will be essential to utilize the complex high-performance computer system effectively. In addition, dense sensor data tend to have correlated errors, and the proper treatment is necessary to fully utilize the “Big Data.” The current DA methods commonly assume no correlation in the observation errors. Based on our previous theoretical research on the observation-error correlations, we are developing methods to consider the observation error correlations in realistic NWP applications. This is also relevant to satellite data assimilation. The current NWP systems use only a small fraction of available data partly because of the observation error correlations.

Treating the model errors and non-Gaussian probability distribution has been grand challenges in DA. Big Ensemble Data Assimilation with the largest-ever 10,240 samples was a milestone providing fundamental data to investigate the non-Gaussian probability distribution. We have developed expertise and exclusive dataset to tackle these challenges.

For further enhancing our research activities, collaborations with other AICS (R-CCS) teams, RIKEN-wide partners (iTHEMS, RIKEN Engineering Network), MOU partners (University of Maryland, University of Reading), domestic and international research partners (NICI, Tokyo Metropolitan University, JMA, Meteorological Satellite Center, Meteorological Research Institute, Kyoto University, Argentinean National Meteorological Service, University of Buenos Aires, Barcelona Supercomputing Center, LMU Munich, University of Tokyo, Taiwan National Central University, Pennsylvania State University, JAXA, JAMSTEC, Tokyo Institute of Technology, University of Tsukuba, Ritsumeikan University, Institut Mines-Télécom-Atlantique), and industry partners (Tokyo Electric Power Company Holdings, Incorporated, Meisei Electric Co., Ltd., MTI Ltd.) will be the key to success. We will continue expanding our collaborative activities further.

In FY2017, we continued the “DA innovation hub” project to expand our activity to wider fields and to attract more people in experimental, computational, and theoretical sciences. We organized the RIKEN DA Workshop, RIKEN DA Camp, RIKEN International School on Data Assimilation (RISDA2018), and the RIKEN Uncertainty Quantification Workshop. In addition, we awarded the DA fund to five RIKEN researchers in various fields (biology, medical science, brain science, and engineering) to seek for new DA applications. In FY2018, we will continue seeking for new collaborations and seeds of new DA applications.
16.5 Publications

16.5.1 Awards


16.5.2 Articles


16.5. PUBLICATIONS


16.5.3 Invited Talks


176  CHAPTER 16. DATA ASSIMILATION RESEARCH TEAM


[17] Lien, G.-Y., Toward an operational high-resolution regional LETKF data assimilation system for a small area: challenges and promises, Taiwan Typhoon and Flood Research Institute, Taipei, Taiwan, 9/22/2017.


[21] Lien, G.-Y., T. Miyoshi, and T. Honda, Roles and issues of a high-resolution regional ensemble data assimilation system for a small area, Central Weather Bureau, Taipei, Taiwan, 12/19/2017.


16.5.4 Oral Presentations


近藤圭一, 三好建正, アンサンブルデータ同化における鉛直誤差相関の調査, 日本気象学会春季大会, 東京, 5/25/2017.


澤田洋平, 岡本幸三, 国井勝, 三好建正, ひまわり 8 号輝度温度データの同化による局地的大雨の再現性向上の試み, 日本気象学会春季大会, 東京, 5/28/2017.


CHAPTER 16. DATA ASSIMILATION RESEARCH TEAM


[40] 前島公光, G.-Y. Lien, 三好建正, 高密度地上観測のデータ同化による降水量予測改善のプロセス, 日本気象学会 2017 年度秋季大会, 北海道大学, 10/31/2017.

[41] 澤田洋平, 岡本幸三, 国井勝, 三好建正, Running-In-Place 法を用いたひまわり 8 号全天輝度温度データ同化による局地的大雨の予測精度向上の試み, 日本気象学会春季大会, 札幌, 10/31/2017.

16.5. PUBLICATIONS


[51] Lien, G.-Y., T. Miyoshi, and T. Honda, Roles and issues of a high-resolution regional ensemble data assimilation system for a small area, Research Center for Environmental Changes, Academia Sinica, Taipei, Taiwan, 12/18/2017.

[52] 三好建正, 「ビッグデータ同化」の技術革新の創出によるゲリラ豪雨予測の実証, AIP ネットワークラボ主催 NSF 連携ビッグデータ合同領域会議, 東京, 12/19/2017 (英).


[59] 小槻峻司, 黒澤賢太, 三好建正: 全球大気データ同化システムNICAM-LETKF を使った EFSO観測インパクト推定. 第8回データ同化ワークショップ, 明治大学, 中野区, 東京, 1/19/2018.

[60] 前島康光, G.-Y. Lien, 三好建正, 平成27年関東東北豪雨事例における精密地上観測データ同化のインパクト, 第8回データ同化ワークショップ, 明治大学中野キャンパス, 1/19/2018.

CHAPTER 16. DATA ASSIMILATION RESEARCH TEAM


16.5.5 Poster Presentations


16.5. PUBLICATIONS


[33] Terasaki, K. and T. Miyoshi, Accounting for the observation error correlation in data assimilation. The 8th AICS International Symposium, RIKEN AICS, Kobe, Japan, 2/7/2018.


[38] Ohhigashi, M., S. Kotsuki, T. Miyoshi, and S. Takino: Dam Operation Optimization by Machine Learning, RIKEN International Workshop on Uncertainty Quantification (UQWS), RIKEN AICS, Kobe, Japan, 2/19/2018.


[40] Awazu, T., S. Otsuka, and T. Miyoshi, Data assimilation using shape features of rainfall areas, RIKEN International Workshop on Uncertainty Quantification (UQWS), RIKEN AICS, Kobe, Japan, 2/19/2018.

[41] Takatama, K., T. Miyoshi, J. Kikuchi, and K. Ito: Towards a coastal ocean prediction, RIKEN International Workshop on Uncertainty Quantification (UQWS), RIKEN AICS, Kobe, Japan, 2/19/2018.

[42] Sakamoto, H., T. Teramura, and T. Miyoshi, Non-Gaussian Measure in Gaussian Filtering Problem, RIKEN International Workshop on Uncertainty Quantification (UQWS), RIKEN AICS, Kobe, Japan, 2/19/2018.

[43] Lien, G.-Y. and T. Miyoshi, Issues regarding maintaining ensemble spreads, balance, and high-resolution information in rapid-update-cycle radar data assimilation with the LETKF, 6th International Symposium on Data Assimilation, Munich, Germany, 3/5/2018.


