ISDA-Online
Friday, September 06, 2024 from 07 – 09h UTC

"Open Session on Data Assimilation"

Organizers: *James Taylor (RIKEN, Japan)*
Javier Amezcua (Tec. De Monterrey, Mexico; U. Reading, UK)
Lars Nerger (AWI, Germany)
Nora Schenk (DWD, Germany)
Tobias Necker (ECMWF, Germany)

Program: (All talks: 15' + 3' Q&A)

07:00 – 07:03   **Welcome**

07:04 – 07:22   **Artificial Neural Networks Approach to Predict and Forecasting Significant Wave Height Using Long Short-Term Memory (LSTM)**
Balla Maggero

07:23 – 07:41   **A physics-based method for calculating the displacement of background field**
Yicun Zhen, Valentin Resseguier, Bertrand Chapron

07:42 – 08:00   **Sequential Markov Chain Monte Carlo for Data Assimilation with Applications to Unknown Data Locations**
Hamza Ruzayqat, Alexandros Beskos, Dan Crisan, Nikolas Kantas, Ajay Jasra

08:01 – 08:19   **Ensemble Kalman filter based Land Data Assimilation for the Soil Diffusion based ISBA model**
Abhishek Lodh

08:20 – 08:38   **HGS-PDAF (version 1.0): a modular data assimilation framework for an integrated surface and subsurface hydrological model**
Qi Tang, Hugo Delottier, Wolfgang Kurtz, Lars Nerger, Oliver S. Schilling, Philip Brunner

08:39 – 08:57   **Latest Results of Including ZDR Column for Enhanced Radar Data Assimilation at German weather Service (DWD)**
Kobra Khosravian, Klaus Stephan, Jana Mendrok, Alberto De Lozar, Ulrich Blahak

08:57 – 09:00   **Closing / Outlook**

Please note:
- When you login to the session before 07:00 UTC, and everything is quiet, this is most likely because we muted the microphones.
- The times in UTC are approximate. In case of technical problems, we might have to change the order of the presentations.
- **Time Zones**: 07 – 09 UTC

08 – 10 am BST (London)	09 – 11 am CEST (Berlin)	
03 – 05 pm CST (Shanghai)	04 – 06 pm JST (Tokyo)	05 – 07 pm AEDT (Sydney)
12 – 02 am PDT (San Fran.)	01 – 03 am MDT (Denver)	03 – 05 am EDT (New York)

Artificial Neural Networks Approach to Predict and Forecasting Significant Wave Height Using Long Short-Term Memory (LSTM)

Balla Maggero[1]

[1]Institute for Meteorological Training & Research, Kenya

Artificial neural networks (ANN) algorithms are capable of producing data-driven decisions. Supervised learning is a type of machine learning that uses labelled data to train the algorithm, where both input and output data are given in this training phase. The algorithm learns the mapping function from the input to output variables, and after that, with a new set of data, the algorithm produces an output. In the context of ocean wave modelling, given the comprehensive amount of significant wave height time series documented and available, NN seems like a good alternative to improve predictions of this variable specially due to its 'forget gate' layer, which gives the net the capability to remember long-term propagation of data, or forget them, when is suitable. We investigate short-term NN to forecast significant wave height on several locations in Kenya's coast, for four different lead times. We perform forecast with verification times within a time period of one month or 744 time-steps and the training phase of the LSTM considers four different lead times. Training of the LSTM network is performed using both ERA5 reanalysis data, and real buoy measurements. With the former showing a better accuracy in the prediction period when compared to its own training database, but the latter shows higher accuracy with respect to buoy observations. As a contribution, we show an architecture for LSTM network that can be suitable for short-term forecasts of significant wave height, and potentially for other ocean waves variables as well, with a large decrease in computational time if compared to traditional physical models, and with accuracy, superior to the reanalysis, when the training is done base exclusively with observational data.

A physics-based method for calculating the displacement of background field

Yicun Zhen[1], Valentin Resseguier[2], Bertrand Chapron[3]

[1]Hohai University, China
[2]INRAE, France
[3]IFREMER, France

To estimate displacements of physical fields, a general framework is proposed. Considering that for each state variable, a tensor field can be associated, ways these displacements act on different state variables will differ according to the tensor field definitions. This perspective provides a differential-geometry-based reformulation of the generalized optical flow (OF) algorithm. Using the proposed framework, optimisation procedures can explicitly ensure the conservation of certain physical quantities (total mass, total vorticity, total kinetic energy, etc.). Existence and uniqueness of the solutions to the local optimisation problem are demonstrated, leading to a new nudging strategy using all-available observations to infer displacements of both observed and unobserved state variables. Using the proposed nudging method before EnKF, numerical results show that ensemble data assimilation better preserves the intrinsic structure of underline physical processes if the ensemble members are aligned with the observations.

Sequential Markov Chain Monte Carlo for Data Assimilation with Applications to Unknown Data Locations

Hamza Ruzayqat[1], Alexandros Beskos[2], Dan Crisan[3],
Nikolas Kantas[3], Ajay Jasra[4]

[1]King Abdullah University of Science and Technology, Saudi Arabia
[2]University College London, UK
[3]Imperial College London, UK
[4]The Chinese University of Hong Kong

We consider the problem of high-dimensional filtering/data assimilation of continuous and discrete state-space models at discrete times. This problem is particularly challenging as analytical solutions are usually
not available and many numerical approximation methods can have a cost that scales exponentially with the dimension of the hidden state. We utilize a sequential Markov chain Monte Carlo method to obtain samples from an approximation of the filtering distribution. For certain state-space models, this method is proven to converge to the true filter as the number of samples, N, tends to infinity. We benchmark our algorithms on linear Gaussian state-space models against competing ensemble methods and demonstrate a significant improvement in both execution speed and accuracy (the algorithm cost can range from O(Nd) to O(Nd[d+1]/2) based on the model noise covariance matrix structure, where d is the dimension of the hidden state. We then consider a state-space model with Lagrangian observations such that the spatial locations of these observations are unknown and driven by the partially observed hidden signal. This problem is exceptionally challenging as not only is high-dimensional, but the model for the signal yields longer-range time dependencies through the observation locations. Finally, the algorithm is tested on the high-dimensional rotating shallow water model with real data obtained from drifters in the ocean.

Ensemble Kalman filter based Land Data Assimilation for the Soil Diffusion based ISBA model

Abhishek Lodh[1]

[1]Swedish Meteorological and Hydrological Institute (SMHI), Sweden

The current research in seamless numerical weather prediction particularly in land surface and soil processes involves accurate estimation of soil moisture, soil temperature and snow, particularly over the European domain. Over the Arctic land region the land surface variables like snow depth, snow temperature exerts deeper impact on evaporation, latent and sensible heat flux and influences weather. So, it is important to provide accurate initial condition of land surface variables like snow, soil moisture and soil temperature to the land surface model across the different soil levels down to root zone. Both NWP and hydrology models requires proper root zone soil moisture data and microwave sensors provides only top few centimeters of soil moisture data. Even after getting soil moisture observations in real time, various sequential land DA techniques are being adopted by different modeling communities to properly initialize the soil moisture in the NWP models with a higher spread extending down to root zone. In the CERISE project at SMHI we are developing an Ensemble Kalman filter based land data assimilation to initialize soil variables from wide range of surface observations. Ensemble Kalman filter solves the limitations of saturated soil moisture in Extended Kalman filter, thus less requirements of non-linearity issues. Tests are ongoing for evaluating the spread in screen level observations and soil moisture (both surface and root zone) and to find the correlation between the screen level humidity and temperature to the soil moisture, soil temperature across the soil levels in the ISBA land surface model (CISBA = DIF; CSNOW = 3-L).

HGS-PDAF (version 1.0): a modular data assimilation framework for an integrated surface and subsurface hydrological model

Qi Tang[1,2], Hugo Delottier[1], Wolfgang Kurtz[3], Lars Nerger[4],
Oliver S. Schilling[1,2,5], Philip Brunner[1]

[1]Centre for Hydrogeology and Geothermics (CHYN), University of Neuchâtel, Switzerland
[2]Hydrogeology, Department of Environmental Sciences, University of Basel, Switzerland
[3]German Meteorological Service, Centre for Agrometeorological Research, Branch Office Weihenstephan, Germany
[4]Alfred-Wegener-Institut, Helmholtz Zentrum für Polar- und Meeresforschung, Germany
[5]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Switzerland

In this work we developed a modular ensemble-based data assimilation (DA) system which is developed for an integrated surface–subsurface hydrological model. The software environment for DA is the Parallel Data Assimilation Framework (PDAF), which provides various assimilation algorithms like the ensemble Kalman filters, non-linear filters, 3D-Var and combinations among them. The integrated surface–subsurface hydrological model is HydroGeoSphere (HGS), a physically based modelling software for the simulation of surface and variably saturated subsurface flow, as well as heat and mass transport. The coupling and capabilities of the modular DA system are described and demonstrated using an idealised model of a geologically heterogeneous alluvial river–aquifer system with drinking water production via riverbank filtration. To demonstrate its modularity and adaptability, both single and multivariate assimilations of hydraulic head and soil moisture observations are demonstrated in combination with individual and joint updating of multiple simulated states (i.e. hydraulic heads and water saturation) and model parameters (i.e. hydraulic conductivity). With the integrated model and this modular DA framework, we have essentially developed the hydrologically and DA-wise robust toolbox for developing the basic model for operational management of coupled surface water–groundwater resources.

Latest Results of Including ZDR Column for Enhanced Radar Data Assimilation at German weather Service (DWD)

Kobra Khosravian[1], Klaus Stephan[1], Jana Mendrok[1], Alberto De Lozar[1], Ulrich Blahak[1]

[1]Deutscher Wetterdienst, Germany

Radar data assimilation has been operationally used in the short-range ensemble numerical weather prediction (SRNWP) system (ICON-D2-KENDA LETKF system) at DWD since 2020 (radial wind from March 2020 and reflectivity from June 2020). It is in addition to the traditional Latent Heat Nudging (LHN) of 2D radar-derived precipitation rates.

Moreover, the study of radar data assimilation in NWP models and its effect particularly on short-term forecast has been intensified recently at DWD. In particular, the seamless Integrated ForecastiNg sYstem (SINFONY) project that leads a short-term forecasting system focusing on convective events from minutes up to 12 h shows clearly the benefit of radar data assimilation in improving the short-term forecast.

Furthermore, the integration of polarimetric radar parameters as a novel observational source into the data assimilation system at DWD has been recently considered. One of the most notable parameters is the so-called ZDR column, defined based on the differential radar reflectivity (ZDR), indicating a narrow vertical layer of positive ZDR above the 0°C level. This feature is closely linked to deep convective storms and can serve as an indicator of large raindrops or hail particles. We will present our latest results on incorporating the ZDR column as a new observational input in Observing System Simulation Experiments (OSSE) complementing radial wind and reflectivity data, based on the operational data assimilation framework used at DWD.