

# Developing Cloud Computing Architecture for Modeling (Model as a Service) using Data Assimilation Techniques

Dhanashree Kuthe  
Assistant Professor  
JNTUH  
Department of Computer Engineering

## ABSTRACT

Data assimilation refers to any use of observational information to improve a model. To solve any problem like weather forecasting, traffic management, water management, agricultural management, urban planning modeling of that particular problem is important. For developing the perfect model real time observational information is necessary. To get the correct solution and forecasting incorporating the observational data in the model will definitely improve the results and the perfect model has been build. Data assimilation techniques like statistical interpolation, Kalman Filter, 4d-Var, Ensemble Kalman filter, Optimal Interpolation, Nudging, Analysis Correction and Successive correction can be used to improve the model. But the question is how to get the real time data and improve the model, since to develop any model and to incorporate the huge amount of real time data into the model huge amount of computing resources is necessary. Cloud Computing provide the resources as required in agile way with its characteristics like elasticity, broad network access and resource pooling. The integration of cloud computing and data assimilation will help to build new applications to solve the above problems and get the instant access to those applications on the internet so any common man or any researcher can use it.

## General Terms

Data Assimilation and Machine Learning.

## Keywords

Model as a Service, Cloud Computing, Data Assimilation.

## 1. INTRODUCTION

Cloud Computing is a new paradigm to provide computing as a utility service for addressing different processing needs with on demand self services, resources pooling, elasticity, broad network access and pay per use. The utility of delivering computing capability adoptive a potential solution for the integration of numerical model and data assimilation. Data assimilation was meant for state estimation, but in the broadest sense, data assimilation refers to any use of observational information to improve a model.

Basically, there are four methods for “model updating”, as follows:

**Input:** corrects model input forcing errors or replaces model-based forcing with observations, thereby improving the model’s predictions;

**State:** corrects the state or storage of the model so that it comes closer to the observations (state estimation, data assimilation in the narrow sense);

**Parameter:** corrects or replaces model parameters with observational information (parameter estimation, calibration);

**Error correction:** correct the model predictions or state variables by an observed time-integrated error term in order to reduce systematic model bias (e.g. bias correction).

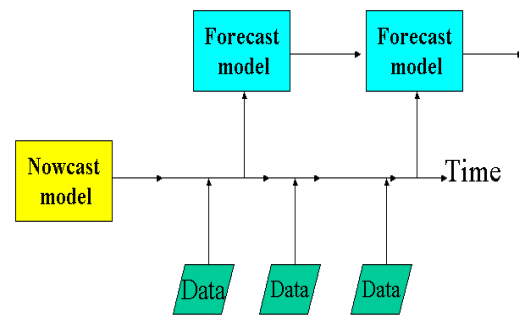


Fig 1: Data Assimilation Process

Online data acquisition, data assimilation and modeling have become more and more important in data science and data analysis. In this research I am going to use cloud computing for integrating field data acquisition and stochastic, modeling in a data assimilation and optimization framework as a service to solve any problem.

The focus of the proposed research is to develop a cloud computing architecture that will be used to design, formulate and improve a model using real time data that will be provided by the user of the system. For developing perfect a perfect model and algorithms huge real time data is necessary at particular time interval as shown in the above figure. It also needs huge computational power to processed the big real time data. The data assimilation techniques will provide the algorithms to improve the model and cloud computing will provide the computational power as it need.

## 2. LITERATURE SURVEY

Accurate initial conditions are needed in order to produce an accurate forecast. Numerical weather prediction model should be used to get the desired results but are not guaranteed for wind farms in isolated locations. In [1] authors, proposed that the wind forecast can be improved by assimilating anemometer wind speed observations from wind farm turbines into a numerical weather forecast system. A technique was developed to circumvent the requirement of simultaneously ingesting the wind speed and direction in a data assimilation system.

In [2] authors presented to map urban air quality in real time with low-cost sensors. The method merges low-cost sensor data with time-invariant model output.

Satellite-based rainfall products contain errors in real time processing. Soil moisture dynamics are impacted for several days by the accumulated amount of rainfall following within a particular event. In [3], authors collected data from the Soil Moisture Ocean Salinity (SMOS) satellite and it is used to

study the correct rainfall accumulation estimates provided by satellite-based real-time processing. An algorithm based on the SMOS measurements data assimilation is developed and tested in two land-surface models of different complexity.

Big Data and high speed and large processing power of computers has emerged with new opportunities for research, development, innovation. Cloud Computing has emerged as a new paradigm to provide computing as a utility and elastic computing service with pay per use model for addressing different processing needs. In [4] authors investigate how Cloud Computing can be utilized to address Big Data challenges to enable such transformation. The method is presented to leverage Cloud Computing for Big Data solutions. This framework can also be referred as a guidance to develop potential solutions for other big geospatial data challenges and initiatives, such as smart cities.

The study in [5] proposes a framework that uses data assimilation as a post processing technique to increase the accuracy of water depth prediction.

With the significant advancements in Information and Communications Technology, cloud based applications provide a novel approach to access applications which are not installed on the local computers. The integration of cloud computing and Internet of Things (IoT) indicated a bright future of the Internet. In [6], authors proposed a new architecture of cloud computing—Model as a Service (MaaS). The feasibility of the proposed architecture is proved by implementing a groundwater model on cloud as a case study. This study provides a flexible and effective approach for analyzing the uncertainties and time variant properties of the parameters and the proposed architecture of cloud computing provides a novel approach for the researchers and decision-makers to construct numerical models and follow-up researches.

Weather-related research requires high processing power for synthesizing vast amounts of data using GPU enable technologies and other solutions that are both economical and viable during and past the lifetime of the project. In [7] authors, illustrate the use of a private cloud object store developed by the Center for High Performance Computing.

The research of network service scalability is essentially the research of the capability of providing service, which is based on services and the relationship among the services. In [8] authors propose an analytical method of network service scalability, which is composed of a network description model and a service evaluation model. The method uses the network description model to describe services and the relationship among them.

In [9] authors build a service model for analysis of Mental health of people in England.

### 3. RESEARCH QUESTION

#### 3.1 Problem Statement

Please use a Developing Cloud Computing Architecture for Modeling (Model as a Service) using Data Assimilation Techniques.

#### 3.2 Sub Problems

- Designing cloud computing architecture for assimilating the real time data.
- Developing cloud-based real-time modeling and data assimilation framework.

- Connecting to a cloud-based data acquisition and monitoring module.
- Testing each data assimilation algorithms on the developed cloud computing architecture.
- Developing a easy to use application for Model as a Service to use by any normal user.

## 4. METHOD

### 4.1 Problem to be Solved

Representation of soil moisture conditions for hydrologic forecasting.

### 4.2 Methodology and Steps

1. Developed a private cloud using the resources.
2. Developed the algorithm for fast processing of data and modeling using the private cloud resources.
3. Developing the basic model using the cloud resources for the representation of soil moisture using the previous and current data.
4. Efficient Deployment of the sensors based on IoT (Internet of Things) to get the real time data.
5. Use the data assimilation techniques to improve the basic model for soil moisture hydrologic forecasting. We are using Ensemble Kalman Filter and 4d-Var algorithm for improving the model and forecasting.
6. Developing the web application for the use of our system over the internet as Model as a Service.

### 4.3 Data Analysis

Soil moisture sensors are used to measure the volumetric water content of the soil.

The data of the soil moisture is in following form.

*Soil sampling, Volume, Soil sample initial weight, Dried sample weight, Mass of water, Mass of dry soil.*

### 4.4 Importance

The present work has both computer engineering and mathematical importance. From the mathematical point of view understanding of data assimilation techniques may lead to new methods for model improving. On computer engineering side, efficient resource utilization with both energy efficient and storage efficient way using cloud computing will help to access the system over the internet.

### 4.5 Problems to be Solved using the System

1. Weather Forecasting
2. Agriculture Management
3. Water Management
4. Urban Planning.
5. Pollution Monitoring
6. Ocean and Ice Glacier Monitoring.

## 5. CONCLUSION

Data assimilation techniques like statistical interpolation, Kalman Filter, 4d-Var, Ensemble Kalman filter, Optimal Interpolation, Nudging, Analysis Correction and Successive correction can be used to improve the model. But the question is how to get the real time data and improve the model, since to develop any model and to incorporate the huge amount of real time data into the model huge amount of computing

resources is necessary. Cloud Computing provide the resources as required in agile way with its characteristics like elasticity, broad network access and resource pooling. The integration of cloud computing and data assimilation will help to build new applications to solve the above problems and get the instant access to those applications on the internet.

## 6. REFERENCES

- [1] William Y.Y., Cheng Yubao, Liu Alfred, J. Bourgeois, Yonghui Wu Sue and Ellen Haupt. Short-term wind forecast of a data assimilation/weather forecasting system with wind turbine anemometer measurement assimilation. *Renewable Energy* Volume 107, July 2017, Pages 340-351.
- [2] Philipp Schneider, Nuria Castell, Matthias Vogt Franck , R. Dauge William A.Lahoz Alena Bartonova. Mapping urban air quality in near real-time using observations from low-cost sensors and model information. *Environment International* Volume 106, September 2017, Pages 234-247.
- [3] Carlos Roma *et. al.* Correcting satellite-based precipitation products through SMOS soil moisture data assimilation in two land-surface models of different complexity: API and SURFEX. *Remote Sensing of Environment* Volume 200, October 2017, Pages 295-310.
- [4] Chaowei Yang Manzhu, Yu Fei Hu, Yongyao Jiang and Yun Li. Utilizing Cloud Computing to address big geospatial data challenges. *Computers, Environment and Urban Systems*, Volume 61, Part B, January 2017, Pages 120-128
- [5] Amir Javaheri, Mohammad Nabatian, Ehsan Omranian, Meghna Babbar-Sebensand and Seong Jin Noh Merging Real-Time Channel Sensor Networks with Continental-Scale Hydrologic Models: A Data Assimilation Approach for Improving Accuracy in Flood Depth Predictions. *Hydrology* 2018, 5(1),9; doi:10.3390/hydrology5010009.
- [6] Chong Chen Dan, ChenYingnanYan, Gaofeng Zhang Qingguo and Zhou Rui Zhou. Integration of numerical model and cloud computing. *Future Generation Computer Systems* Volume 79, Part 1, February 2018, Pages 396-407.
- [7] Brian K.Blalock John D.Horel and Samuel T.Liston, Cloud archiving and data mining of High-Resolution Rapid Refresh forecast model output. *Computers & Geosciences* Volume 109, December 2017, Pages 43-50.
- [8] Jun Shen, Zuqin Ji, Yubin Zhu, and Jinjin Huang. An Analytical Method of Network Service Scalability. Special section on recent advances on radio access and security methods in 5g networks, Volume 6, 2018.
- [9] Brynmor Lloyd Evans PhD, Bethan Paterson MSc, Steve Onyett PhD, Ellie Brown MSc, Hannah Istead MSc, D Clin Psych, Richard Gray PhD, Claire Henderson PhD and Sonia Johnson DM. National implementation of a mental health service model: A survey of Crisis Resolution Teams in England. *International Journal of Mental Health*, Volume 27, Issue 1, February 2018 Pages 214-226.
- [10] Community Corrections service delivery model: An evidence-based approach to reduce reoffending, R Caruana- *Judicial Officers Bulletin*, 2018.
- [11] N Becker and M Fidler. A non-stationary service curve model for estimation of cellular sleep scheduling. *IEEE Transactions on Mobile Computing* 2018.
- [12] R Singh, HS Sandhu, BA Metri and R Kaur. Supply Chain Management Practices, Competitive Advantage and Organizational Performance: A Confirmatory Factor Model. *Operations and Service*, 2018 –igi-global.com.
- [13] M Potschin-Young, R Haines-Young and C Görg. ] Understanding the role of conceptual frameworks: Reading the ecosystem service cascade. *Ecosystem Services*, 2018 – Elsevier.