

# Cloud Service for Analysis and Interactive Visualization of Weather Data in Armenia

Hayk A. Grigoryan and Rita M. Abrahamyan

Institute for Informatics and Automation Problems of NAS RA  
Armenian State Hydrometeorological and Monitoring Service  
e-mail: hayk.grigoryan.a@gmail.com

## Abstract

The Lesser Caucasus Mountains are crossing through the territory of Armenia, creating vast differences in altitude, terrain, temperature and precipitation in provinces and towns. Even Armenia's lowlands are 500 to 1500m above sea level. Armenia's highlands extend up to Aragats mountain at 4090m where, 75% of the territory is above 1000m, 50% is above 2000m, and 3.4% is above 3000m. This paper presents a cloud service with interactive visualization and analytical capabilities for weather data in Armenia by integrating the two existing infrastructures for observational data and numerical weather prediction.

The weather data used in the platform consist of near-surface atmospheric elements including air temperature, relative humidity, pressure, wind and precipitation. The visualization and analytics have been implemented for 2m air temperature. Cloud service provides the Armenian State Hydrometeorological and Monitoring Service with analytical capabilities to make a comparative analysis between the observation data and the results of a numerical weather prediction model for per station and region for a given period.

**Keywords:** Cloud service, Weather data, Observational data, Data analysis, Numerical weather prediction, WRF, Spatial OLAP.

## 1. Introduction

Armenia occupies the north-eastern part of Armenian plateau and central part of Lesser Caucasus range (latitude 38.51' to 41.18' North, longitude 43.29' to 46.37' East), with the area of about 30 000 sq.km. The geographical location of Armenia and complex mountainous relief have led to the diversity of natural conditions across the country. Armenia is on the northern edge of the subtropical zone, in latitudes characterized by an arid and continental climate. Due to the mountainous relief, different climatic zones exist, and the weather may have high spatial gradients. High fluctuations in annual and daily temperatures are typical for the Armenian climate. The presence of six climatic zones from dry subtropical to rigorous high mountainous and from everlasting snowcaps to warm humid subtropical forests and humid semi-desert steppes make additional challenges on weather forecasting and climate prediction for the Armenian State Hydrometeorological and Monitoring Service (AHMS).

The meteorological data, received from 47 meteorological stations, among which only four stations provide historical data and monthly updates to the Global Climate Observing System Surface Network and three meteorological stations provide synoptic data are being used in the gridded analysis data set, which serves as an input for the global atmospheric models to produce weather forecasts at the global scale.

The observation data received from the meteorological stations and data received from a global model serve as inputs and outputs to the high-resolution numerical weather prediction models to produce outputs of temperature, precipitation, and other meteorological elements from the ground to the top of the atmosphere [1, 2]. The high-performance computational (HPC) resources of the Armenian e-infrastructure are used to better resolve mesoscale weather events, and hence to give reasonably accurate forecasts in a short range [3]. The limitations in such forecasts lie in the availability of initial conditions at model resolutions.

The article aims to present the weather data interactive web-based visualization and analytical platform<sup>1</sup>. The platform has been developed for the weather data in Armenia by integrating the three existing platforms for observational data and numerical weather prediction. The platform provides a way to compare the output of forecasting model with the observation data gathered from different stations for a chosen frame of time. The suggested platform is essential for a wide range of applications, such as urban area management, sustainable development and nature protection, regional and local planning, agriculture, forestry and fisheries, health, civil protection, infrastructure, transport and mobility or tourism.

The remainder of this paper is divided into the following sections: Section 2 introduces the infrastructure, Section 3 represents the discussions and analyses and finally Section 4 is the conclusion.

## 2. Cloud Service Framework

The suggested cloud service consists of 5 primary layers (see Fig. 1). The base layer provides HPC and data resources, which is particularly important for digital models [4]. The resources of the Armenian e-infrastructure are utilized, which is a complex national IT infrastructure consisting of both communication and distributed computing infrastructures.

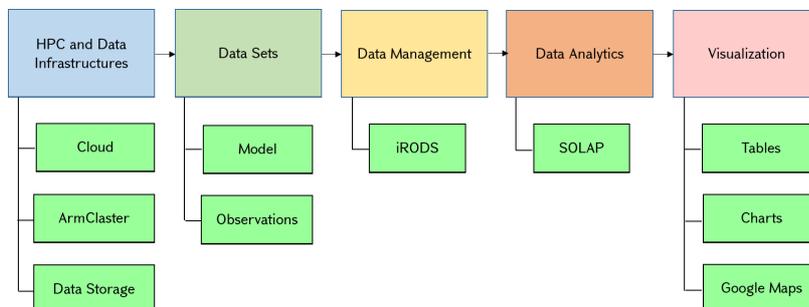


Fig 1. The framework of the platform.

<sup>1</sup>Cloud service for analysis of numerical weather prediction model accuracy using visualization technique: <http://meteo.grid.am>

The datasets layer combines two types of data platforms for further analysis:

- Model output: outputs of weather prediction models;
- Observation data: meteorological stations observations, as a base to investigate the deviation values with other model outputs.

The Data management layer provides intelligent tools to transfer raw data to data analytics layer. The Integrated Rule-Oriented Data System (iRODS) provides a middleware between the physical data storage systems and the user interface [5].

As soon as data reaches data analytics layer, it is processed, and only several indexes are left from the enormous amount of starting crude information. Finally, the top layer and final destination of the already processed data is the visualization layer, where the result lists are changed to more easy to understand graphs or tables. Besides, the advantages of Google Maps are used to map these indexes with their real location on the map.

## 2.1 Data Gathering

Observational datasets provide the observed meteorological data from different locations acquired from SYNOP (surface synoptic observations) messages issued by official ground weather stations. SYNOP reports are regularly sent every three hours, which consist of groups of numbers describing general weather information of the weather station, such as temperature, barometric pressure and visibility at a weather station, and so on.

The numerical weather prediction models are initialized using NCEP (National Centers for Environmental Prediction) Global Forecast System analysis and forecasts at 0.5 deg horizontal resolution [6]. Information created during pre-processing and simulations of the model is in the Lambert conformal projection, which is appropriate for mid-latitude domains.

The model setup (see Fig. 2) comprises a parent D01 domain (with a common center located at longitude 44.7, latitude 40.0), partly covering Europe, all the Caucasus, parts of the Central Asia and the Middle East and the nest domain d02, covering the entire region of Armenia.

## 2.2 Numerical Weather Prediction Model

The mesoscale Weather Research and Forecasting (WRF) model [9, 10], which is adapted to the territory of Armenia, is used in operational weather forecasting. Designed to serve both research and operational needs, it has grown to offer a spectrum of options and capabilities for a wide range of applications. The D01 domain with a common center located at longitude 44.7, latitude 40. has horizontal resolution 18-km with 202x202 grid points , the d2 nest with 9770 grid points has 6-km horizontal grid increment . The weather forecasts are performed on a daily basis, using the following 1-way nesting strategy. The model uses vertical 31 eta\_levels and the geographic data resolution is 30 seconds. The model was initialized with the initial and boundary conditions of Global Forecast System (GFS) at 0000 UTC (local time on 04:00) for 31-day period, namely 1-31 January 2016. The applied version of WRF-ARW model includes the WRF Single Moment 6-class scheme for cloud microphysics with ice, snow and graupel processes. Sub-grid parameterization of deep and shallow convection is based on the KainFritsch scheme applying mass flux approach with downdrafts and CAPE removal time scale (Kain and Fritsch 1993). In this study the MellorYamadaJanjic (MYJ)

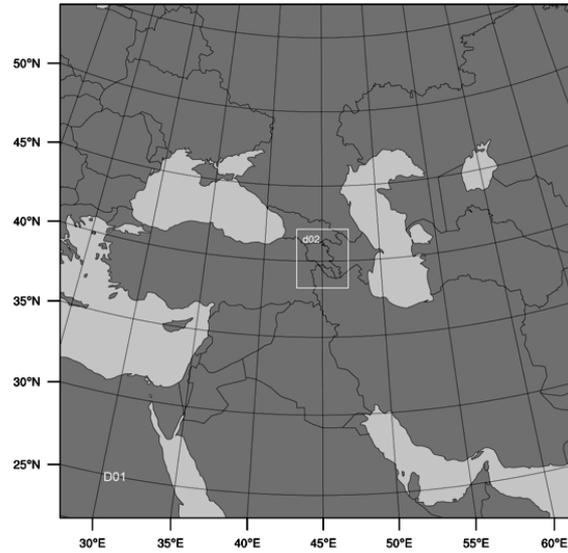


Fig 2. Parent (D01) and nested (d02) domains used in the model.

PBL scheme with Eta similarity surface layer was used (Mellor and Yamada 1982). The rapid radiative transfer model (RRTM) scheme was used for longwave radiation, while shortwave radiation processes were represented by the Dudhia scheme. Due to the limited computing resources, we have considered the inner domain with 6 km resolution, and the range of every run was 24 hours starting always at 0000 UTC of each day.

Observable 2-metres long temperatures from 42 operational stations in Armenia are used to study the accuracy of forecasting air temperature by the model WRF for January 2016. It must be noted that the inversion was observed in Armenia frequently in January 2016 causing low temperatures in inversion affected regions. The surface inversion, trapped by cold air near the ground, causes fog below the inversion layer.

### 2.3 Data Management

Everyday rapid growth of data and need to analyze these data pushed the development of analytical processing tools. OLAP (Online analytical processing) is a model for accessing multidimensional data in data warehouses [11]. Data cubes and OLAP session are main concepts in OLAP. A data cube is a collection of facts and dimensions organizing the data of a data warehouse according to different analysis axes and aggregation measures. OLAP provides a set of operations (such as drill-down and slice-and-dice) that transform one multidimensional query into another, which provide high querying. OLAP queries are formulated as sequences called OLAP sessions. For analyzing data with spatial and georeferenced components the Spatial OLAP (SOLAP) technology is used, which allows rapid and easy navigation within spatial databases and that offers many levels of information granularity, many themes, many epochs and many display modes synchronized or not: maps, tables and diagrams [12]. It allows a tight integration of GIS and OLAP systems. A SOLAP system supports three types of spatial dimensions: the non-geometric spatial dimensions, the geometric spatial dimensions and the mixed spatial dimensions. During an OLAP session, the

user analyzes the results of a query and, contingent upon the particular information, applies an activity to determine a new query that will give a superior comprehension of information.

## 2.4 User Interface and Visualization

The implemented service consists of 3 main intelligent blocks, and the data flows between these pieces. The first block is the Database. PostgreSQL with its PostGIS extension is used to store data sets with geometrical information. The following block is implementing data analytic logic. The special type of OLAP approach is used (Spatial OLAP) to process the required data. As a query language for OLAP, the MultiDimensional eXpressions (MDX) [13] is used to interact and perform tasks with multidimensional databases (OLAP Cubes). Afterwards, the handled information is exchanged to User Interface where user can make different graphs, tables and see the output on the map (see Fig. 3).



Fig 3. Web-based visualization and analytical platform consist of the following sections:  
1- query form, 2 - temperature chart, 3 - coefficients table

The main screen includes many parts, which are marked with red numbers.

- ⇒ Query form - user can pick out the station, start and end dates and the period (by default it takes all hours from 0 to 21) for requesting the wished records for plot. After the plot action, the charts and the table will be updated with the corresponding values;
- ⇒ Temperature - shows the observation and model temperature information lines correspondingly for each period;
- ⇒ Coefficients table - shows the RMSE, BIAS and R correlation coefficients.

Platform provides the ability to transfer new information to the database by utilizing API endpoints. At present, only the administrator of the platform has access to these activities. For observation data, SYNOP documents (the records are in ascii format) must be used, which will be parsed utilizing a script written in JavaScript language. For WRF model outputs, which are produced in netCDF file format, the Python script was made for finding relating values based on the stations information stored in the database (see Fig. 4).

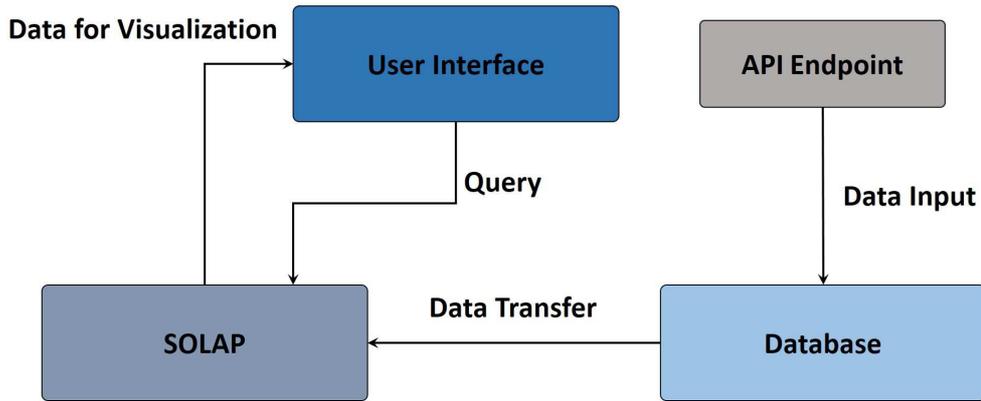


Fig. 4. Data workflow consists of API endpoint, database, SOLAP and user interface steps.

### 3. Evaluation of Results

Various statistical and object-oriented methods are used in the suggested web-based analytical platform to investigate the characteristics of model-forecast, which is important for giving helpful guidance to end-users.

As a case study, the 2 metre temperature has been investigated using the observational and regional high-resolution WRF model data for January, 2016. For the studied period, the RMSE, BIAS and R correlation coefficients are calculated for the observational and model-forecast data for 42 observation points distributed in the territory of Armenia (see Table 1).

Table 1: Mean estimates of verification of temperature forecasts for several meteo stations per region.

Region	Station	Height above sea-level, m	RMSE	BIAS	R
Tavush	Ijevan	732	7.61	-4.47	-0.40
	Bagratashen	453	7.56	-6.02	-0.31
Shirak	Ashotsk	2012	2.52	-1.10	0.91
	Gyumri	1513	3.91	-3.04	0.82
Gegharkunik	Masrik	1934	5.13	-4.14	0.8
	Gavar	1960	2.74	-1.61	0.89
Kotayq	Sevan Lake	1917	2.59	1.74	0.90
	Hrazdan	1765	2.15	-0.36	0.91
Aragatcotn	Fontan	1780	1.82	0.2	0.92
	Eghvard	1336	3.5	-2.96	0.88
Yerevan	Ashtarak	1090	3.57	-2.93	0.84
	Zvartnots	853	4.44	-3.91	0.82
Ararat	Arabkir	1113	3.24	-1.92	0.77
	Ararat	818	5.65	-4.83	0.62
Syunik	Artashat	829	5.69	-5.10	0.73
	Meghri	627	9.89	-7.37	-0.70
	Kapan	705	9.69	-7.81	-0.43

The table shows that the average difference between the forecast data and observations is about 4-5°C. For all the stations studied, an acceptable correlation coefficient has been obtained, which allows one to judge the adequacy of the model.

The analyses, which have been carried out for 42 stations (some high-altitude stations are not considered), show that the model data for almost all stations are overstated. RMSE values are about 1.6-2.5°C for the stations located at altitudes above 1500m-2000m, 4.6-9.8°C for the stations below 1000m. It means that the model predicts well the temperature values at a height of 2m for stations located above 1500m and gives unsatisfactory results for stations below 1500m. The worst results are obtained for stations below 1000m, such as stations located in Meghri, Kapan, Ijevan, Bagratashen, and the valley of Syunik and Tavush region.

The January 2016 temperature for Ararat Valley and Yerevan is also projected unsatisfactorily. In Ararat Valley RMSE 5.6-6.4°C, the worst result was obtained for Ararat station. In Yerevan, a good result was obtained for Yerevan-Arabkir station, which is located at an altitude of 1113m, and the worst result for Yerevan-Zvartnots. The correlation coefficient varies from 0.80-0.97 for mountain and foothill areas, as for valleys it is from 0.7 to 0.25 and from -0.7 to -0.31.

From all that has been described above, it can be concluded that the WRF model with the described configuration, for a cold period of time, gives a positive forecast for a variable air temperature of 2 m for the mountain and foothill regions of the republic (Shirak, Kotayk, Gegharkunik, Lori, mountain and foothill areas of Aragatsotn) and at the same time an unsatisfactory forecast for the Valley of Syunik, Tavush, Ararat Valley and Yerevan.

Analyzing all the factors that make up the temperature, we see that there were fogs in the valley during the selected period, which was the reason for low temperatures. Therefore we can conclude that the model WRF is not predicted by low temperatures of occurrence of the lowland, effect of the fog at surface inversion. Such a result does not satisfy, therefore, it is necessary to improve the accuracy of the model data, through a proper tuning, it would be useful to test the impact of resolution given, because the Armenian terrain has a rather complex orography and is characterized by several land-category types.

#### 4. Conclusion

The suggested platform enables to integrate the already available observational and model-forecast informations and use these sources for studies and analyzes in a web-based visualization environment. The interactive comparison charts for 2m air temperature allows to visually analyze and gather the information about model accuracy. It enables to adjust the forecasting results with additional methods by implementing statistical analyses and provides a fairly high result in cases where the model's sensitivity is low.

It is planned to enhance the functionality of the platform by adding a new source of satellite image processing, the land surface temperature, which will give an opportunity to analyze locations where the meteorological stations do not exist. It will include also new visualization tools of various formats, such as to analyze and compare other near-surface atmospheric elements. Different nowcasting methodologies based on artificial intelligence will be implemented for the development of a hazardous hydro-meteorological phenomena alarm system.

## References

- [1] A. Gevorgyan, “Summertime wind climate in Yerevan: valley wind systems”, *Climate Dynamics*, vol. 48, no. 5–6, pp. 1827–1840, 2017.
- [2] A. Gevorgyan, H. Melkonyan, R. Abrahamyan, Z. Petrosyan, A. Shachnazaryan, H. Astsatryan, V. Sahakyan and Yu. Shoukourian, A Persistent Surface Inversion Event in Armenia as Simulated by WRF Model, *in IEEE Proceedings of the International Conference on Computer Science and Information Technologies, CSIT’2015*, pp. 105–110, 2015.
- [3] H. Astsatryan, V. Sahakyan, Y. Shoukourian, P.-H. Cros, M. Dayde, J. Dongarra and P. Oster, “Strengthening Compute and Data intensive Capacities of Armenia”, *in IEEE Proceedings of 14th RoEduNet International Conference - Networking in Education and Research, NER’2015*, pp. 28–33, 2015.
- [4] H. Astsatryan, Yu. Shoukourian and V. Sahakyan, “The ArmCluster Project: Brief Introduction”, *in Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications*, pp. 1291–1295. 2004.
- [5] M. Hedges, T. Blanke and A. Hasan, Rule-based curation and preservation of data: A data Grid approach using iRODS , *Future Gener. Comput. Syst.*, vol. 25, no. 4, pp. 446–452, 2009.
- [6] W. S. Jeffrey , T. M. Hamill, X. S. Yucheng and Z. Toth, Ensemble data assimilation with the ncep global forecast system, *Monthly Weather Review*, vol. 136, pp. 463–482, 2008.
- [7] J. A. Sobrino, J. C. Jimenez-Munoz and L. Paolini, Land surface temperature retrieval from Landsat TM 5, *Remote Sensing of Environment*, vol. 90, pp. 434–440, 2004.
- [8] M. Neteler, M. Bowman M. Landa and M. Metz, Grass gis: A multi-purpose open source gis, *Environmental Modelling & Software*, vol. 31, pp. 124–130, 2011.
- [9] W. C. Skamarock and J. B. Klem, A time-split non-hydrostatic atmospheric model for weather research and forecasting applications, *Computational Physics*, vol. 227, no. 7, pp. 3465–3485, 2008.
- [10] J.G. Powers, J.B. Klemp, et. al, The weather research and forecasting model: Overview, system efforts, and future directions, *Bulletin of the American Meteorological Society*, vol. 98, no. 8, pp. 1717–1737, 2017.
- [11] S. Chaudhuri and U. Dayal, An Overview of Data Warehousing and OLAP Technology, *SIGMOD Record*, vol. 26, no. 1, 1996.
- [12] S. Aissi, M. S. Gouider, T. Sboui and L.B. Said, Enhancing spatial data warehouse exploitation: a solap recommendation approach, *In: Computer and Information Science*, Springer, pp. 131–147, 2016.
- [13] M. Whitehorn, R. Zare and M. Pasumansky, *Fast Track to MDX*, Springer-Verlag London, 2006, DOI: 10.1007/1-84628-182-2

## Հայաստանում եղանակային տվյալների համար վերլուծության և ինտերակտիվ վիզուալիզացիայի ամպային ծառայություն

Հ. Գրիգորյան և Ռ. Աբրահամյան

### Անփոփում

Փոքր Կովկասյան լեռները անցնում են Հայաստանի տարածքով, ստեղծելով տեղանքի բարձրության, ջերմաստիճանի և տեղումների մեծ տարբերություններ: Հայաստանի հովիտները գտնվում են ծովի մակարդակից 500-ից 1500մ բարձրության վրա: Լեռնաշղթաները տարածվում են մինչ Արագած լեռը՝ 4090մ, որտեղ տարածքի 75% -ը գերազանցում է 1000մ-ը, 50% -ը՝ 2000մ-ից բարձր է, իսկ 3.4% -ը՝ 3000 մ բարձրությունից: Աշխատանքի նպատակն է ներկայացնել Հայաստանում եղանակային տվյալների համար վերլուծության և ինտերակտիվ վիզուալիզացիայի ամպային ծառայություն, որը բաղկացած է ցամաքային կայանների դիտարկումների և եղանակի թվային կանխատեսման մոդելի ենթակառուցվածքներից:

Մշակված համակարգում օգտագործվող եղանակային տվյալները բաղկացած են մակերևույթին մոտ մթնոլորտային բաղադրիչներից, ներառյալ օդի ջերմաստիճանը, հարաբերական խոնավությունը, ճնշումը, քամին և տեղումները: Վիզուալիզացիան և վերլուծությունը իրականացվել են 2մ բարձրության օդի ջերմաստիճանի համար: Ամպային ծառայությունը Հայաստանի պետական հիդրոօդերևութաբանական և մոնիտորինգի ծառայությունը հնարավորություն է տալիս որոշակի ժամանակահատվածում յուրաքանչյուր կայանի և տարածաշրջանի համար կատարել դիտարկումներ, ինչպես նաև եղանակի կանխատեսման թվային մոդելի արդյունքների համեմատական վերլուծություններ:

## Облачный сервис для анализа и интерактивной визуализации данных о погоде в Армении

А. Григорян и Р. Абрамян

### Аннотация

Малые Кавказские горы, пересекая территорию Армении, создают большие перепады высот местностей, температур и осадков. Равнины Армении находятся на высоте от 500 до 1500м над уровнем моря. Армянское нагорье простирается до горы Арагац(высота 4090 м), где 75% территории превышает 1000м, 50% - свыше 2000м, а 3,4% - свыше 3000м. Цель данной работы представить облачный сервис для данных о погоде в Армении, разработанный путем интеграции двух существующих инфраструктур - наблюдения с наземных станций и численная модель прогнозирования погоды.

Погодные данные, используемые в разработанной системе, состоят из атмосферных элементов в приземном слое, включая температуру воздуха, относительную влажность, давление, ветер и осадки. Анализ и визуализация проводились для температуры воздуха на высоте 2м. Облачный сервис предоставляет возможность государственной службе по гидрометеорологии и мониторингу Армении в определенный период времени для каждой станции и местности провести сравнительный анализ наблюдаемых данных с результатами данных от численной модели.