WEB-BASED PLATFORM FOR DATA MANAGEMENT, ANALYSIS AND VISUALIZATION OF WEATHER AND ENVIRONMENTAL MONITORING STATIONS

Legarretaetxebarria A.[a], Alonso K.[b], Serrano M.[c], Aginako N.[a], Olaizola I.[a]

[a]Vicomtech-IK4, Mikeletegi Pasealekua 57, 20009 Donostia, Spain
{alegarreta; naginako; iolaizola}@vicomtech.org

[b]German Aerospace Center (DLR), Oberpfaffenhofen, D-82234 Weßling, Germany
kevin.alonsogonzalez@dlr.de

[c]Dominion, Barrio Boroa, 100, 48340 Boroa, Spain
mikel.serrano@dominion.es

Abstract

The high growth of ICT approaches brings opportunities for novel applications in the fields of meteorology and environment. The capacity of measuring more parameters and the broader access to them increases significantly the amount of data, which requires new systems and platforms for the analysis, monitoring graphical representation of geospatial information.

This paper presents a platform designed to optimize the management of all the information generated by spatially distributed weather and environmental monitoring stations. The platform is connected to the station network so it receives stores and processes all the observations. The incorporation of data aggregation systems is one of the key factors in the development of the platform due to the different nature of the sensors.

The validation of the platform has required the integration with real monitoring stations belonging to different meteorology and environmental agencies, permitting the involvement of expert knowledge.

Keywords: GIS Technologies, GIS Data Processing, Visual Analytics

INTRODUCTION

With the growth of ICT approaches, the access in real-time to all kind of data captured in environmental and weather stations has become easier. Even more, the development of ICT technologies regarding data management, analysis, visual representation and geographic localization has boosted the meteorological and environmental sectors. The capacity to store and manage all these raw and processed data is essential for government and agencies since they are used in their decision making processes including warning reports when natural hazards are happening [1].

These data gathering capacity will highly benefit weather forecasting systems which consist in the prediction of the state of the atmosphere for a future time in a given location. Surface weather observations are the fundamental input data of the atmosphere models used for forecasting weather, issue warnings or alarms of disasters worldwide [2] [3]. Weather forecasts are made by collecting quantitative data about the current state of the atmosphere and using scientific models of the atmosphere for forecasting atmospheric processes to project how the atmosphere will evolve [4].

Despite the development of ICTs, the management of data from environment and atmosphere observation stations is still an outstanding issue. Searching, accessing, and processing heterogeneous data from different sources are very time-consuming and complex tasks. But new database technologies coupled with both advanced data analysis processes and emerging web-based technologies, hold the key to lowering the cost of data management, visualization and adaptation to scientific processes **Error! Reference source not found.**.

The main objective of this work is the implementation of an integrated solution for the collection, management, processing, analysis and modeling of weather and environmental data. Besides, the project has a secondary purpose which is the graphical representation of the forecasted data by exploiting the potential of the latest technologies in 3D rendering and Visual Analytics in web environment. The main drawbacks that arise at design and implementation levels of the platforms are due to the diversity in the spatio-temporal resolution of the collected and forecasted data. This fact hinders the data interpretation process.

In this solution a particular emphasis was put on implementing the capability of visualising the collected parameters in interactive maps. The main purpose was to help users to get a better understanding of all the data, prediction and observation alike. Therefore, the architecture also considers web mapping technologies like Google Earth or Google maps as well as OpenLayers and Leaflet open technologies.

In the following sections we present the platform. The next section describes the features of the different kind of data that the platform needs to handle. After that, the paper describes the platform architecture, followed by the presentation of two real case system implementation scenarios, with validation purposes. Finally conclusion remarks are presented.

DATA

The aimed platform needs to handle weather and environmental observations and weather forecasts as source data. All these data are heterogeneous representing different values of measurement or forecast and come from variety of sources with different formats, spatial distribution and sampling time.

The observation data are air/water physical and chemical values measured by observation stations. The common environmental measurements are CO, NO, NO_2, NOX, O_2, PM_{10} and SO_2 on air and PH-s, cloudiness, temperature, suspended solids and conductivity on water. The weather measurements are parameters such as precipitation accumulation, surface temperature, the direction and speed of the wind, and so on. These values are subject to the spatiotemporal variability, the geographical distribution of the sensors and the time in which the values are measured. The observation stations have a sensor network which comprises an array of sensor nodes and a wireless communications system which allows their data to reach a centralized server.

For observation, is possible to collect the data from different meteorological and environmental agencies or directly from stations. The time ranges of measurements depend on the type of station and the spatial resolution is defined by the number of stations and the area which these stations cover. The geographic location of all the stations and measurements is also provided. This information is needed in order to provide maps with georeferenced observations.

The weather forecasts are produced by widely employed Weather Research and Forecasting (WRF) Model **Error! Reference source not found.**. It is a mesoscale numerical weather prediction (NWP) system for the simulation and prediction of the atmosphere. Therefore, the output of the model consists on a set of weather-related parameters for each of the points of a grid domain where the forecasting is processed. These output parameters, by default, are written in NetCDF data format which is a standard data form defined for the creation, access and sharing of N-dimensional array oriented scientific data. Thus, the framework will have to handle weather parameters stored in a set of NetCDF files [8].

The forecasting model generates a lot of parameters such as the surface temperature, surface pressure, convective precipitations and so on. However, there are other parameters required by the users such as thermal comfort or wind chill that have to be processed afterwards.

There are some differences between observation and forecast data. The observation data is available in real-time but the forecast data need around six hours to be available (computed in a supercomputer). Other differences are the time sampling and the spatial data distribution. The sampling of the forecasted data is of one hour and the spatial distribution is uniform. Alternatively, the observation data is received in real-time, with a frequency of at least 20 minutes but has no uniform spatial distribution.

SYSTEM DESCRIPTION

The goal of the platform is to provide a web-based platform for managing and visualizing vast volumes of both observed and forecasted data for environment and weather analyzing purposes. All these operations are accomplished in real time as a major requirement for web–based services. Therefore, the architecture of the framework is based in a client-server distributed application structure.

The measured observation data and prediction files are stored and managed at server side where the requests of the clients are processed and all the analysis and visualization operations are located. At the client side, a web interface helps users to explore the data through the visualization of the data in interactive web maps.

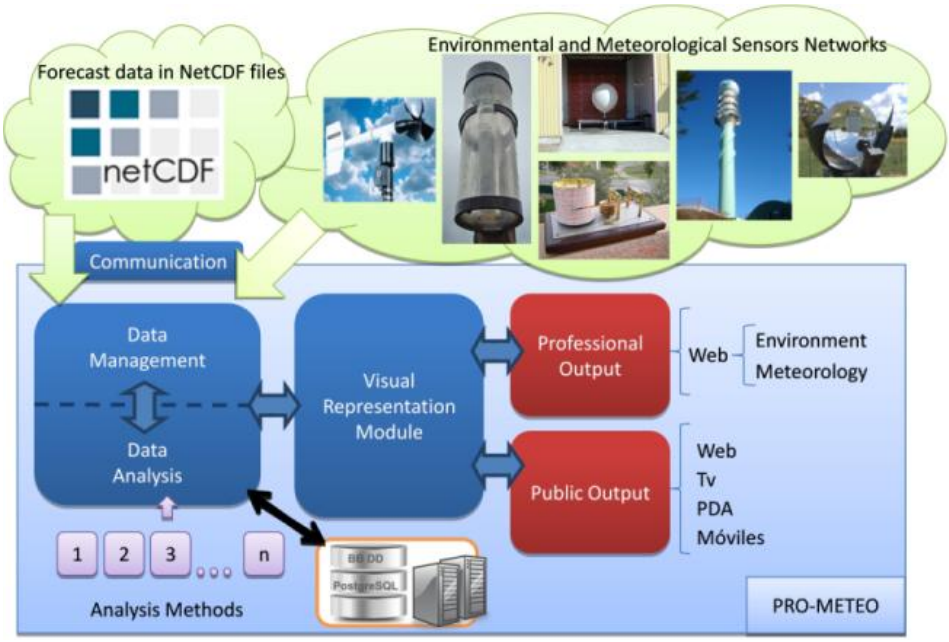


Figure 1: Architecture of the platform.

The platform is composed of four main modules. Communication module consists on the central server that is connected with the remote stations and the supercomputer where the prediction process is done. Data Management module is responsible for managing the sources of different characteristics and storing the normalized data. Data Analysis module generates new products and verifies the integrity of the data correcting or discarding wrong values. It also generates secondary weather parameters and a set of new geovisualization products as 2D images and 3D models. Visual Representation module supports the interaction of the users and handles visual representations generated in the previous module.

Communication

The communication module is responsible for collecting and storing both observation and forecasting data. Therefore it connects the platform with both the observation stations and the supercomputer that carries out the weather forecasting. Mostly, the transmission of the data from the observatories and weather forecasting units is accomplished through the standard network protocol FTP (File Transfer Protocol). Therefore, the communication module consists on a service that receives data through FTP protocol.

Data Management

Once data from different sources is received, it is necessary to pre-process and store it. In the case of observation data, due to the different nature of the sensors and the complexity of the integration process of real data, ad hoc optimizations have been performed in reading and writing functionalities for proper storage. Because of its wide usage the open source PostgreSQL object based relational database was selected to manage metadata, due to its free availability and ability to store and retrieve large datasets using its own functions. The location information of the observed data is stored through PostGIS module of PostgreSQL database.

The integration of observational data is grouped by typology providing two kind of products:

- Discrete products: are those in which the value is a measurement (speed, temperature, humidity ...)
- Raster products: those in which the value corresponds to a set of geospatially distributed discrete values (data from a weather radar, the raw file satellite image, processed image of a radar, etc.)

For the forecast data storage, an interface capable of reading a file in NetCDF format was developed. This interface generates raster data for each forecasted parameter to be subsequently integrated into PostgreSQL database. A raster data is a sequence of discrete georeferenced values, and one raster layer is generated for each parameter at a time and covering the entire spatial grid.

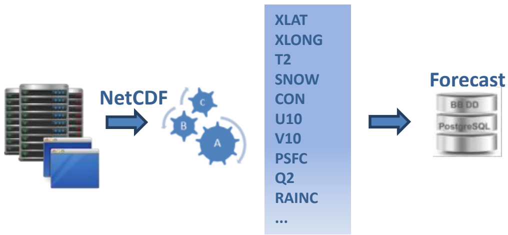


Figure 2: NetCDF file parameters stored in the database.

Data Analysis

Prior to the data storage is necessary to ensure the correctness of the data. To ensure this the system allows the validation and possible correction of the data from observation by the definition of several rules. All data marked as incorrect are not taken into account for the generation of future products that are considered the main outputs of the system. These validation rules and correction processes allow great flexibility performing the validation tasks where data from previous periods of time and other observations are also employed.

After validation and correction processes, the system permits to calculate the new products based on predefined rules or by generating complex new rules, based these on preexisting ones from a previously developed analysis. These calculations can generate products of either discrete values, such as addition, maximum, minimum, average, etc., or raster products such as images, KML / KMZ files, graphics, etc. For instance, one of the created products gives the air quality level at predefined areas and in the case that this value is out of normal range of contamination the system will activate the predefined alerts.

Regarding raster products, the georeferenced analysis module generates colour maps of the observation data. This module makes an interpolation, based on ordinary Kriging method [7], taking as input the parameter of georeferenced measured observation values and generating a georeferenced image as output.

Continuing with the description of raster products, the KMZ file generation analysis module uses raw data and geographic coordinates of the weather radar to create 3D models. Firstly, textures are created like RGBA images. After that, the KML and the layer description file are created. Finally all these images and files are compressed in a KMZ file to be saved in a database.

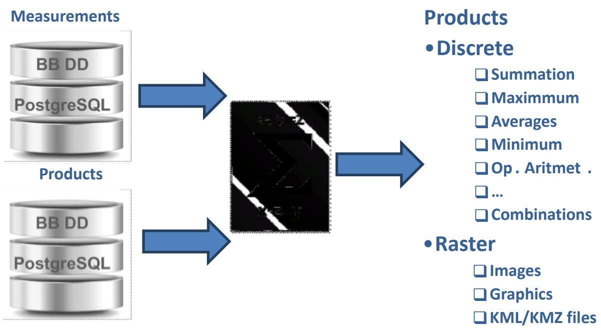


Figure 3: Discrete and raster products generation.

Besides, in the case of forecast data, it is necessary to configure the server with MapServer development environment and ncWMS web map service [8]. MapServer compiles MapCache module which is able to generate tiles from NetCDF files and store them in a cache avoiding regenerating the visualizations for the subsequent calls.

Visual Representations

This module, executed at the client side, contains the interface for handling the interaction with the users. Although the main goal of the platform is to provide the data and product visualization through a web application, the content is also adaptable to be displayed in tablets and mobile phones.

As mentioned above, the data management module stores different types of data such as maps, texts or numeric values. New ways of visual representation were developed in order to facilitate the access and interaction with the data as well as to combine data of different nature.

Graphical display

The system is able to return the required data to generate different kind of 2D graphics depending on user selection. The tool allows the user to compare values of different parameters after defining the time range by a start and end date. In addition, the system allows making product comparisons among created values like the average, maximum and minimum, and also permits comparison of the forecast values with the measured ones. All graphics enable user interactivity allowing adjustment of the visualization to suit the user and it is possible to export graphics to a Microsoft Excel or PDF files which facilitates the inclusion of the graphics in formal documents.

Visual interface requests the information to be displayed to the database through functions that are predefined in it. For example, in Figure 4 we can appreciate the comparison among measured NOX values by sensors which are geographically in different places.

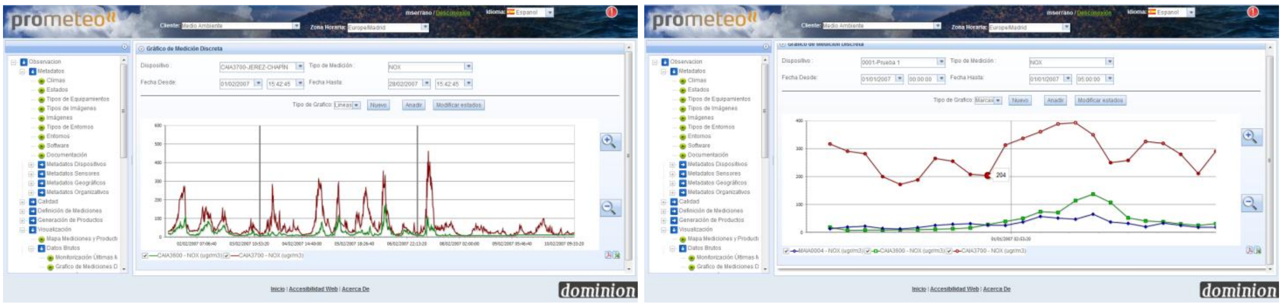
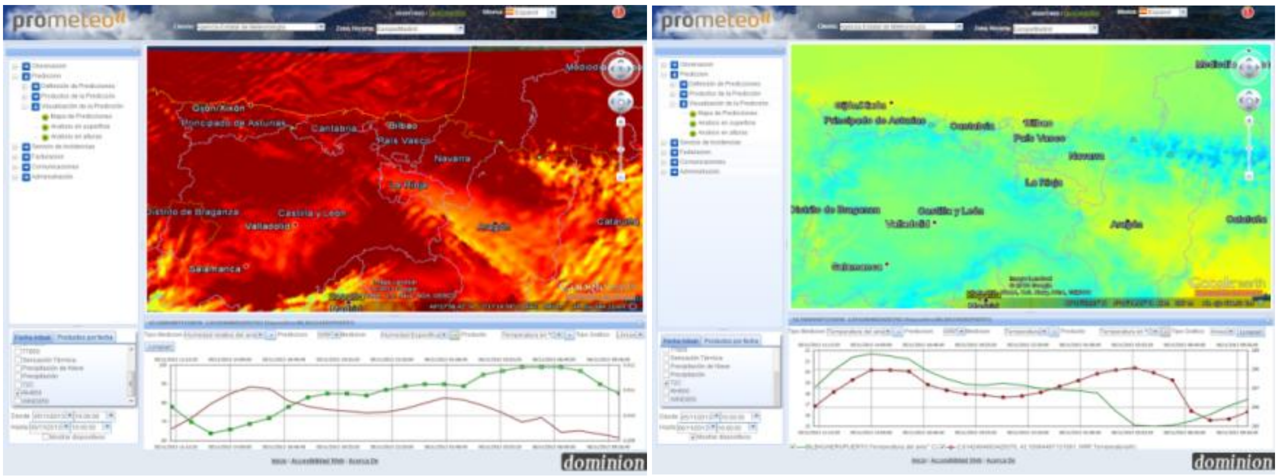


Figure 4: Comparison among measured NOX values by sensors which are geographically in different places.

The

Figure 5 shows how to perform a comparative between observation measurements and predictions. In these case, the same variable have been compared, but it is also affordable the comparison between different variables, for example temperature with the relative humidity.



a) Areas graphic, forecast RH (green line) and observed RH (red line).

b) Lines graphic, forecast Temperature (red line) and observed Temperature (green line).

Figure 5: Comparisons between forecast and observation.

Map

Maps are visual representations of a geographic area and are used to display images and 3D models generated in the data analysis module. It is necessary to add web mapping services in web interfaces, for 3D maps Google Earth has been used and regarding flat maps, Google Maps, OpenLayers and Leaflet. The main characteristic of these services is the possibility of displaying georeferenced images which brings the advantage of providing more intuitive views about the global location and other information which may be useful for the weather and environmental study. Not only professionals but also the general public takes benefits of the utilization of georeferenced solutions.

For example, one of the non-discrete products developed in the platform are color maps created by ordinary Kriging method (**Error! Reference source not found.**). In the example, image is generated from temperature observation measurements and displayed in Google Earth for better comprehension.

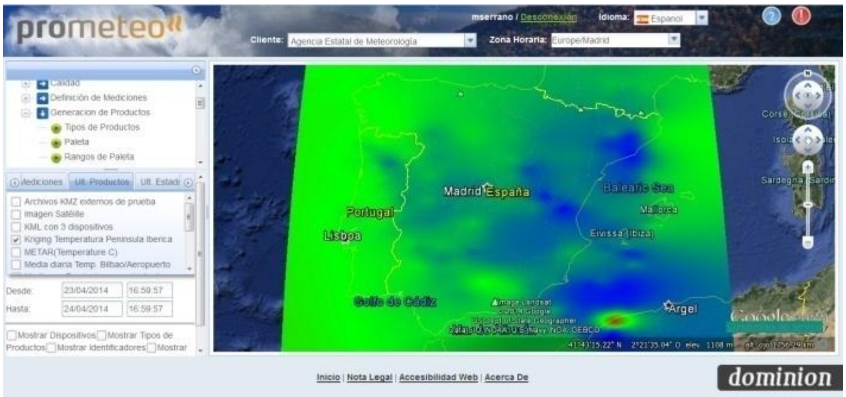


Figure 6: Temperature Colour Map.

Based on the generated KMZ files, 3D model is constructed and displayed on Google Earth web map service. This model is composed of a set of concentric cones with different elevations and the textures are mapped to ensure 3D volume appearance. This image represents precipitation location and direction, permitting the experts to differentiate between types of precipitation such as rain, snow or hail. The a) Aerial.
b) Terrestrial.

Figure 7 shows precipitation from an aerial and terrestrial point of view.

a) Aerial. *b) Terrestrial.*

Figure 7: KMZ file visualization.

All web map services have the ability to show and manage multiple layers allowing to vary the order and the transparency, facilitating simultaneous layers representation with the added information that this supposes for user. Based in the previous characteristics have been developed an interface allow comparing 2 forecast map windows overlapping images on OpenLayers, Leaflet or Google Maps flat web map services. These both windows shows the same forecast hour which is selected by user but parameters can be different in order to perform a comparative. Applying transparency to layers is possible to show more than one layer per window creating complex representations as shown in

Figure 8 and Figure 9.

a) Between wind and precipitation combination with thermic sensation filter.

(b) Between temperature filter with wind, pressure and temperature combination.

Figure 8: Forecast surface comparison at 00:00.

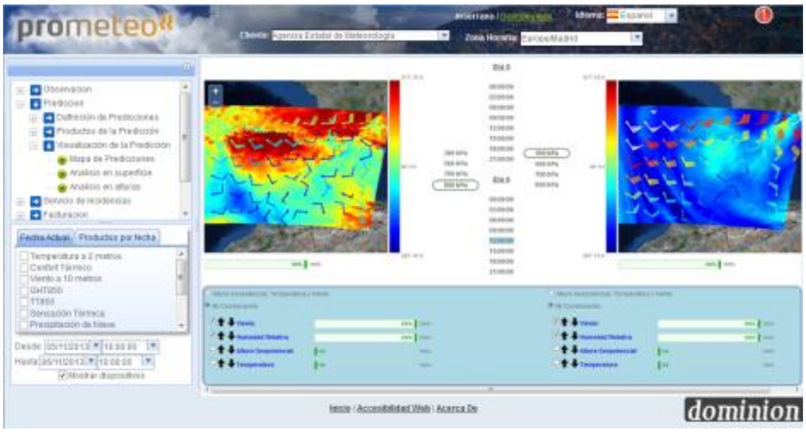


Figure 9: Forecast height comparison at 12:00. Between wind and RH combination at 850hPa with wind and RH with combination at 300hPa.

PLATFORM VALIDATION:

The developed platform has been functionally validated in both Spain and Honduras weather and environment observation national networks. The needs of the countries weather and environmental agencies vary substantially which reinforces the validation process.

Observation solution has been integrated as web-service in our platform and installed in Honduras Permanent Commission for Contingencies, known by its acronym in Spanish COPECO. This government institution was established for effective coordination of the National Risk Management, contributing to equitable and sustainable development of the country.

In this country the platform gets information from 16 meteorological stations, 26 pluviometric stations and 12 hidro-pluviometric stations in an area of 112.492 km^2 as shown in

a) AEMET.

b) COPECO.

Figure 10: Georeferenced weather observatories.

b and the time range of the measurements varies between 15 minutes and 1 hour between the case of higher periodicity and the lowest. Even though the measurements are extracted in different periods, all of them are sent to the central server every hour because the communication is via satellite. In addition, it is also part of the sensor network a weather radar with a grid spacing of 500 km, a resolution of 1000 km/pixel and 10 minutes time scale.

In Spain, both observation and prediction solutions has been integrated and tested in the platform. Regarding observation, there are 260 meteorological stations (AEMET Agencia Estatal de Metereologia) in an area of 504.645 km^2 and the time range of the measurements is 20 minutes. In addition to this, there is also a regional weather radar (EUSKALMET Euskal Metereologi Agentzia) which has a range of 300 km, a resolution of 1000km/pixel and 10 minutes time scale.

a) AEMET. *b) COPECO.*

Figure 10: Georeferenced weather observatories.

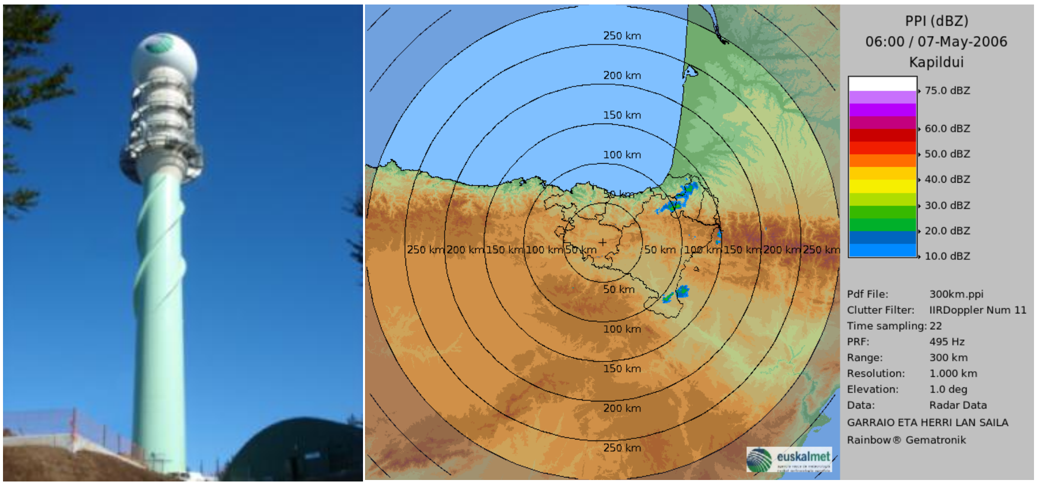


Figure 11: Radar sensor of the regional meteorological agency Euskalmet.

In the forecast case University of León created the forecast weather parameters stored in a set of NetCDF files. All NetCDF files are sent to the servers to processed and represented in developed web-based interface.

CONCLUSIONS

The article describes a web platform that has demonstrated the ability to unify the entire data flow for both observation and prediction solutions. It has been possible to develop an integrated system for the collection, management, processing, analysis and modelling of weather and environmental data. Even more, the platform has been installed and validated in real scenarios as mentioned in the previous section.

The structured and well-organized geographic information integration, which has allowed an agile and simple management of the data, has endorsed the development of the platform. Thus, analyzing and map-making processes of the subsequent modules exploit the capabilities of both software and hardware resources in order to offer extended functionalities to the user.

This paper confirms the potential of the analysis of georeferenced spatial data as well as the value of visual presentation of the results. The platform has demonstrated to be an appropriate and intuitive tool for users who want to interact with weather and environmental georeferenced data permitting the experts concentrating on the analysis of the data.

AKNOWLEDGMENTS

The work described in this paper has been funded by the Basque Government's ETORGAI (IEI09-034) and the Spanish Government's INNPACTO programs (IPT-310000-2010-022). Moreover, the authors would like to thank Euskalmet, Dominion and University of Leon for providing the meteorology and environmental background and resources needed to carry out this work.

REFERENCES

[1] Sun, X.; Shen, S.; Leptoukh, G.G.; Wang, P.; Di, L.; Lu, M. "Development of a Web based visualization platform for climate research using GoogleEarth". *Comput. Geosci.* 2011, *47*, 160

[2] R. Jeberson Retna Raj, T. Sasipraba, "Disaster Management System based on GIS Web Services", *IEEE Trans. Recent Advances in Space Technology Services and Climate Change (RSTSCC)*,pp. 252 – 261, 13-15 Nov.

[3] N. Ahmad, M. Hussain, N. Riaz, F. Subhani, S. Haider5, K. S. Alamgir, F. Shinwari, "Flood Prediction and Disaster Risk Analysis using GIS based Wireless Sensor Networks, A Review", *Journal of Basic and Applied Scientific Research (JBASR)*, pp. 632-643, 2013

[4] P. Sardar Maran and R. Ponnusamy, "Web Enabled Real Time Weather Data Analysis", *Indian Journal of Science and Technology*, Vol 6(11), 5507–5513, November 2013

[5] Sun M., Li J., Yang C., Schmidt G.A., Bambacus M., Cahalan, etal. "A Web- based geovisual analytical system for climate studies". Future Internet4(4), 1069–1085., 2012

[6] J. Michalakes, J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, and W. Wang, "The Weather Research and Forecast Model: Software Architecture and Performance", *Eleventh ECMWF Workshop on the Use of High Performance Computing in Meteorology* . 25–29 October 2004, Reading, U.K., Ed. George Mozdzynski

[7] Cressie, N. A., Statistics for Spatial Data. Wiley, New York, Chichester, Toronto and Brisbane Wiley, 1993

[8] M. Laka-Iñurrategi, I. Alberdi, K. Alonso, M. Quartulli, "Cloud based n-dimensional weather forecast visualization tool with image analysis capabilities" *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XL-7/W2, 2013

BIOGRAPHY OF AUTHORS

Aritz Legarretaetxebarria Torrealday received the B.S. degree in Telecommunications Engineering, Communication Systems specialization, at the University of Mondragon in 2007 and the M.S. degree in Computational Engineering and Intelligent Systems from Basque Country University in 2011. He has worked for Vicomtech-IK4 as a research assistant in the area of Digital TV and Multimedia Services since February 2008. His primary research interest is intelligent virtual environments.

Kevin Alonso Gonzalez studied Technical Telecommunication Engineering in the speciality of Communication Systems (2002-2005) at Engineering Technical High School of Bilbao. He studied Telecommunication Engineering (2005-2007) at the Engineering Technical High School of Bilbao. Since July of 2008 to 2013 he worked in Vicomtech-IK4 as staff researcher in the field of Digital Television and Multimedia Services, participating in national and European research projects. During 2009 and 2010 he combined his research work with the study of a Master Degree on Space Science and Technology. Since 2013 he has worked as PhD. candidate at DLR-German Aerospace Center. His research interests include Data Fusion, Machine Learning and Semantic Web technologies for the development of new algorithms and applications for multimedia asset analysis and management focused in Earth Observation Big Data scenario.

Mikel Serrano Almuña received the degree in Informatics Engineering from Basque Country University in 1997. The last 15 years he was been working in information and communications technologies and consultancy. Since 2001 he works in Dominion as a researcher and programmer analyst and is the head of the I+D area. His research interests are related with the management of Earth Observation Big Data using Web technologies.

Naiara Aginako Bengoa got a degree in Telecommunications engineering from the University of the Basque Country in 2005. From 2003 till 2005 she collaborated the Signal Processing and Communication Group of the Electronics and Telecommunications department (http://det.bi.ehu.es/). She develops and manages research projects in Vicomtech-IK4 since 2005, in the Digital Media department. Since three years and a half she is also teaching at the Politechnique School of Donostia. She is currently pursuing her PhD in image analysis and content based retrieval of images and video. Her work as a researcher counts on a number of publications and on two patents.

Dr. Igor G. Olaizola received a six years degree in Electronic and Control Engineering from the University of Navarra, Spain (2001) and a PhD in Informatics from the Faculty of Informatics of San Sebastián, in the University of Basque Country (2014). He developed his master thesis at Fraunhofer Institut für Integrierte Schaltungen (IIS), Erlangen Germany 2001 where worked for a year as research assistant on several projects related to MPEG standard audio decoding. He is member of Vicomtech technological centre since 2002. Nowadays he is the head of the Digital Interactive TV and Multimedia Services department. In 2006 he also participated as a technology consultant in Vilau company for one year. His PhD focused on automatic multimodal indexing and management of multimedia content.