ENVIRONMENTAL RISK ASSESSMENT USING GEOSPATIAL DATA AND INTELLIGENT METHODS

Nataliia Kussul, Sergii Skakun, Oleksii Kravchenko

***Abstract**: In this paper, we describe intelligent methods and technologies for environmental risks assessment using geospatial data. The risk assessment process is based on fusion of data acquired from different sources: models, in-situ observations and remote sensing instruments. The ensemble approach is used for data processing. Several real-world applications are described to demonstrate efficiency of the proposed approach, namely numeral weather prediction (NWP), land biodiversity assessment, vegetation state assessment, fire monitoring and flood mapping. These applications are being implemented within international projects within the UN-SPIDER Regional Support Office (RSO) in Ukraine.*

***Keywords**: intelligent methods, risk assessment, remote sensing from space, satellite data processing, environmental monitoring, vegetation state assessment, fire monitoring, UN-SPIDER.*

***ACM Classification Keywords**: D.2.12 [Software Engineering] Interoperability; Information Systems; H.1.1 [Models and Principles] Systems and Information Theory; H.3.5 [Information Storage and Retrieval] Online Information Services; I.4.8 [Image Processing and Computer Vision] Scene Analysis - Sensor Fusion.*

Introduction

At present, global climate changes on the Earth made rational land use, environmental monitoring, and prediction of natural and technological disasters the tasks of great importance. The basis for the solution of these crucial problems lies in integrated use of multisource data of different nature, in particular modelling data, in-situ measurements and observations, and indirect observations such as airborne and spaceborne remote sensing data [GEOSS, 2005].

In particular, models can be used to fill in gaps in data by extrapolating and estimating necessary parameters to the site of interest, to better understand and predict different processes occurring in the atmosphere, land, ocean and sea. The models can also help to interpret measurements and to design new observing systems. In-situ measurements are often used for calibration and validation of modelling and remote sensing data, and usually assimilated into models. Satellite observations have an advantage of acquiring data for large and hard-to-reach territories, as well as providing continuous and human-independent measurements. Many important applications such as environmental monitoring, agriculture monitoring, monitoring and predictions of natural disasters heavily rely on the use of Earth observation (EO) data from space. For example, both spaceborne microwave and optical data can provide means to detect drought conditions, estimate drought extent and assess the damage caused by the drought events [Kogan et al, 2004; Wagner et al, 2007]. To assess vegetation health/stress, which is extremely important for agriculture applications, optical remote sensing data can be used to derive biophysical and biochemical variables such as pigment concentration, leaf structure, water content at leaf level and leaf area index (LAI), fraction of photosynthetically active radiation absorbed by vegetation (FPAR) at canopy level [Liang, 2004]. The satellite-derived flood extent [Kussul et al, 2011] is very important for calibration and validation of hydraulic models to reconstruct what happened during the flood and determine what caused the water to go where it did [Horritt, 2006]. Information on flood extent provided in the near real-time (NRT) can also be used for damage assessment and risk management, and can benefit to rescuers during flooding.

The EO domain is characterized by the large volumes of data that should be processed, catalogued, and archived [Shelestov et al, 2006]. The processing of satellite data is carried out not by the single application with a monolithic code, but by the distributed applications. This process can be viewed as a complex workflow that is composed of many tasks: geometric and radiometric calibration, filtration, reprojection, composites construction, classification, products development, post-processing, visualization, etc. Dealing with EO data, we have to also consider the security issues regarding satellite data policy, the need for processing in NRT for fast response within international programs and initiatives, in particular the International Charter “Space and Major Disasters” and the International Federation of Red Cross. It should be also noted that the same EO data sets and derived products can be used for a number of applications. For example, information on land use/change, soil properties, and meteorological conditions is important for droughts identification, vegetation state assessment and floods. Therefore, once we develop interfaces to discover and access the required data and products, they can be used in a uniform way for different purposes and applications. This represents one of the important tasks that are being solved within the development of the Global Earth Observation System of Systems [GEOSS, 2005] and European initiative Global Monitoring for Environment and Security [GMES, 2008]. Services and models that are common for different EO applications (e.g. flood monitoring and crop yield prediction) are shown in Figure 1.

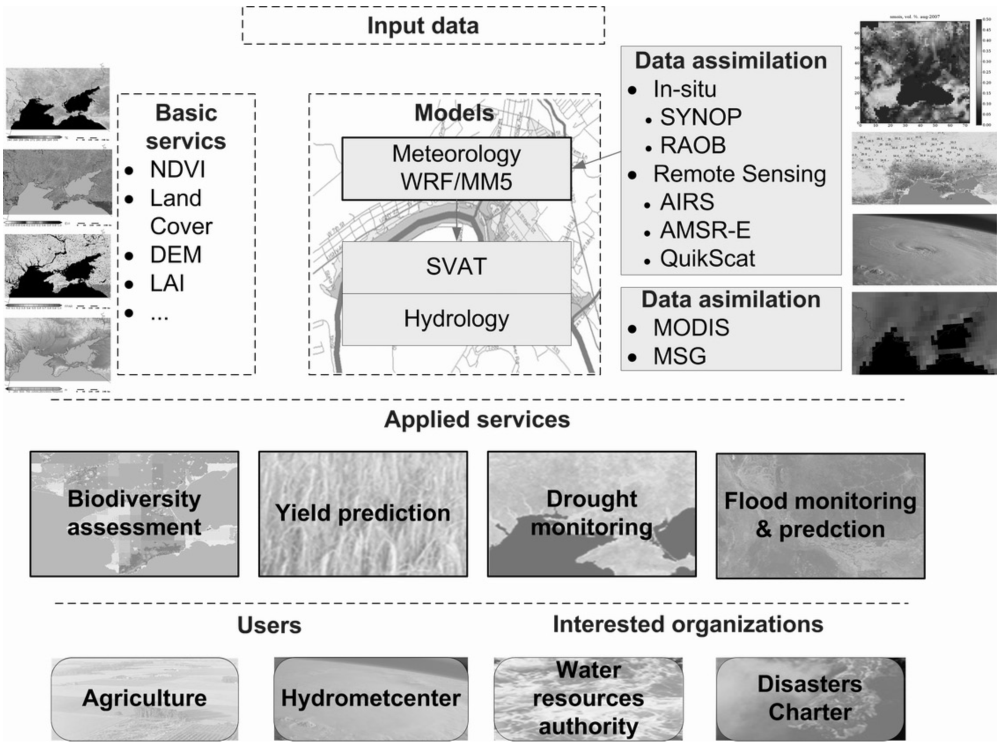

Figure 1. Common services and models for different applications

A considerable need therefore exists for intelligent methods and appropriate technologies that will enable the integrated and operational use of multi-source heterogeneous data for different application domains, and in particular environmental risk assessment.

In this paper, we describe intelligent methods and technologies for environmental risks assessment using geospatial data. The risk assessment process involves a fusion of data acquired from different sources: models, in-situ observations and remote sensing instruments. Several real-world applications are described to demonstrate efficiency of the proposed approach, namely *numeral weather prediction* (NWP), *land biodiversity assessment*, *vegetation state assessment*, *fire monitoring* and *flood mapping*. Most of these applications are being implemented within international projects within the UN-SPIDER Regional Support Office (RSO) in Ukraine (http://un-spider.ikd.kiev.ua).

Environmental Risk Assessment using Geospatial Information

Usually, risk represented as a combination of the likelihood of an occurrence of a hazardous event or exposure(s) and the severity of injury or ill health that can be caused by the event or exposure(s) [OHSAS, 2007]. Mathematically, risk R often simply defined as a function f of disaster probability and expected loss (hazards): $R = f(probability, loss)$.

Event probability could be estimated using a neural network (forecast) model [Haykin, 1999] based on data acquired from remote and in-situ observations (data fusion approach) [Kussul et al, 2009]. To identify the neural forecast model we use risk functional minimization theory developed within a theoretical framework known as computational learning theory [Bishop, 2006] or statistical learning theory [Vapnik, 1998]. Within this approach there are three types of empirical risk minimization problems: classification problem, regression retrieval problem and problem of indirect experiments interpretation. For each of the problems a specific loss function is determined.

To estimate event probability density function information from different sources is integrated (Figure 2).

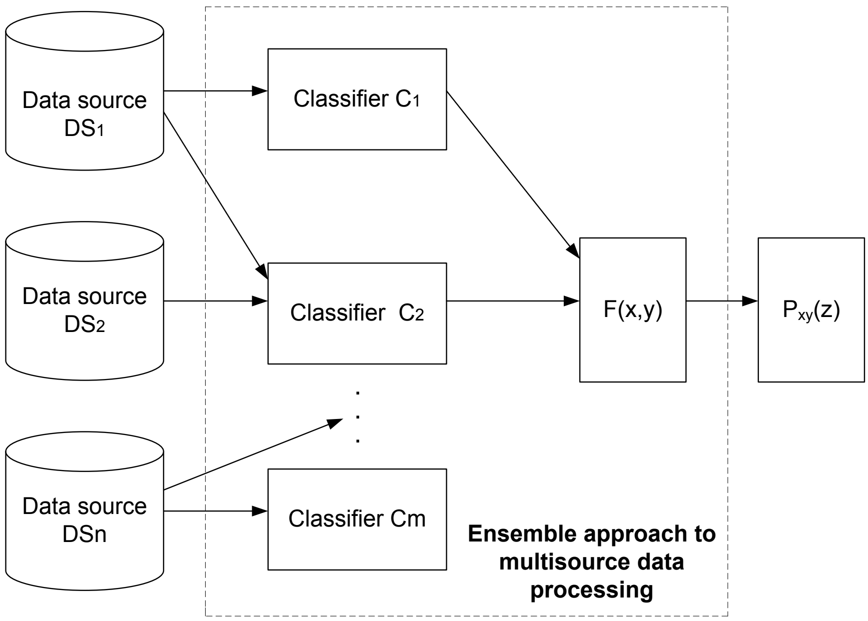

Figure 2. Event probability density estimation from multisource data using ensemble approach

Each classifier (can also be referred as an *expert*) provides an opinion on the event using corresponding data source (geospatial information, point observations). Their outputs are combined through a generalized rule *F*. Such a framework is known as a *mixtures-of-experts model* [Jacobs et al, 1991]. In the following sections we describe how multisource data are combined using this approach for applied problems solving in different domains.

Applications

Numerical Weather Modelling (NWP). Prediction of meteorological parameters represents one of the core services for a number of applications (e.g. floods, droughts, agriculture, etc). Currently, we run the Weather Research and Forecasting model (WRF) [Michalakes et al, 2004] in operational mode for the territory of Ukraine. The meteorological forecasts are generated every 6 hours with a spatial resolution of 10 km. Forecast range is for 72 hours in advance. The horizontal grid dimension is 200 by 200 cells with 31 vertical levels. We use forecasts from the Global Forecasting System (NCEP GFS) for boundary conditions. This data is available via Internet though the National Operational Model Archive & Distribution System (NOMADS).

The workflow of the model run is composed of the following steps: data acquisition; data pre-processing, computation of forecasts using WRF model and data post-processing; visualization of the predicted parameters.

To run WRF model, it is necessary to obtain boundary and initial conditions for the territory of Ukraine. This data can be extracted from the GFS model forecasts. To get the required data, the dedicated script was developed. This script downloads global forecasts every 6 hours. To decrease the data volume, our script uses a special Web-service capable of selecting subsets of the GFS data for the territory of Ukraine. The acquired data is transferred to the storage subsystem and marked as unprocessed (i.e. it has to be processed by the WRF model). After the GFS data has been downloaded, the Karajan script initialises a workflow for data pre-processing, WRF run, and data post-processing.

Data pre-processing step is intended to transform the downloaded data into the format that is used to run the WRF model. GFS data is delivered in the GRIB format in geographical projection. This data is transformed into the internal WRF format by the grib_prep.exe command, warped into the Lambert Conformal Conic projection (by executing hinterp.exe command) and vertically interpolated using the vinderp.exe command. These utilities (grib_prep.exe, hinterp.exe and vinterp.exe) are tools from the WRF Standard Initialization (SI) package. The results of these transformations are stored in the netCDF format. After that, the real.exe command is used to produce initial and boundary conditions for WRF model run. The inputs to real.exe command are GFS data in netCDF format and WRF configuration file (namelist.input).

Data processing step consists in running WRF model using wrf.exe command. The outputs of the command are forecasts of the meteorological parameters. This is the most computationally intensive task. After WRF model run, post-processing step is carried out. For specified weather parameters and for each forecast frame (3 hours), a graphic representation (in PNG format) of spatial distribution is created. Additionally, special files containing georeferencing information are created (files with *.wld extension). The results of the post-processing phase are used to visualize the WRF forecasts via the mapping service. This service provides to the users animations of the weather forecasts (Figure 3). The service provides tools to select a forecast time, forecast frames (up to 72 hours in advance), and weather parameters to display. Selected by the user information is packed into the request to the server. To process the request, all required data (in PNG and WLD formats) is retrieved from the storage subsystem and passed to the mapping server in order to create the maps. Maps are further processed by the script to generate weather animation in GIF format. Finally this animation is presented at user side.

We have also tested the performance of the WRF model in dependence of the number of computational nodes of the supercomputer SCIT-3. For test purposes, we used the WRF model version 2.2 with a model domain identical to those used in operational NWP service (200x200x31 gridpoints with horizontal spatial resolution 10 km). We observed almost linear productivity growth within increasing number of computation nodes. For instance, 8 nodes of the SCIT-3 cluster gave the performance increase in 7.09 times (of 8.0 theoretically possible) when compared to the single node. The use of 64 nodes increases the performance 43.6 times [Kussul et al, 2009].

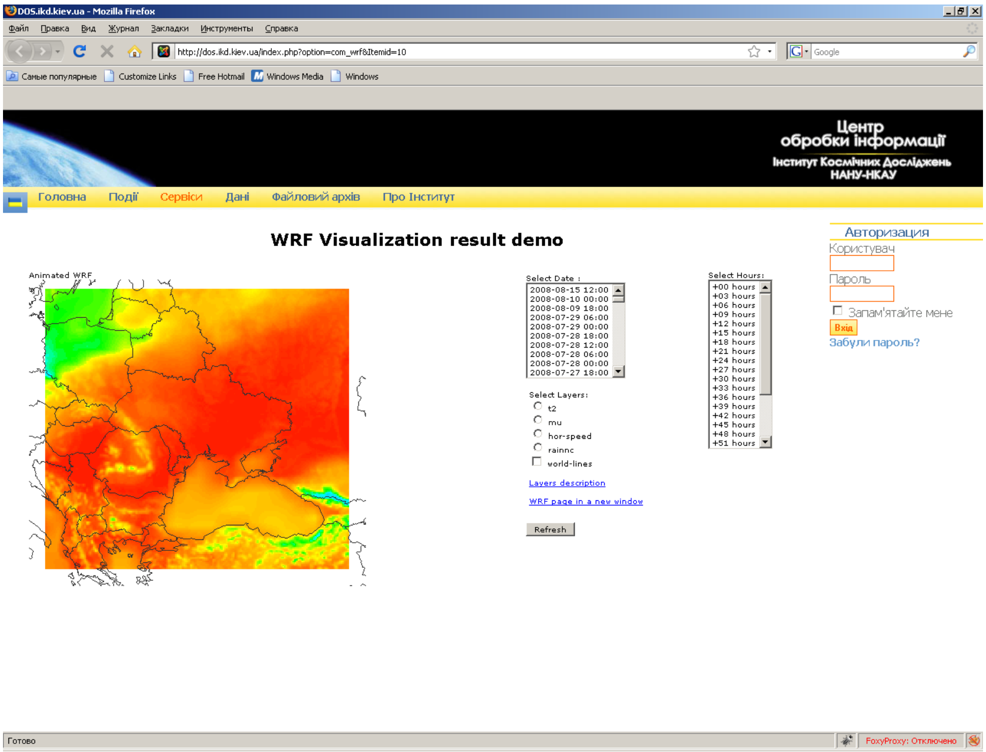

Figure 3. Example of land temperature forecasts using WRF model

Land biodiversity assessment. We have developed a Web service for biodiversity assessment for the Pre-Black Sea region of Ukraine using EOS data products [Popov et al, 2008]. Biodiversity is associated with a number of abiotic and biological factors that can be identified using remote sensing data. These factors include: landscape types, geographical latitude/altitude, climate conditions (such as mean daily temperatures, humidity, etc), structure and primary productivity of a vegetation mantle [Hansen and Rotella, 1999]. These factors can be estimated using EO data from space [Popov et al, 2008]. The workflow for biodiversity estimation consists of the following steps: data acquisition, data processing, and visualization. Figure 4 shows the overall architecture of the service with information flows and integration modules.

Special system was developed in order to acquire multisource satellite data on a regular basis. This system operationally monitors for the new products and provides automatic data acquisition from different sources: Level 1 and Atmosphere Archive and Distribution System (LAADS), Land Processes Distributed Active Archive Center (LP DAAC) and National Snow and Ice Data Center (NSIDC). The acquired data are stored in the data archive of Space Research Institute.

After the required data has been acquired, the data is re-projected to a conical Albert projection and scaled to the spatial resolution of 250 m. Since we use data from multiple sources different tools were applied for the re-projection and scaling purposes. In particular, we used MODIS Swath Reprojection Tool, MODIS Reprojection Tool, and GDAL library (Geospatial Data Abstraction Layer, http://www.gdal.org). Since biodiversity index represents a parameter that is estimated for the time range, it is required to calculate average values for the parameters influencing biodiversity. For this purpose, average composites of images were created. Using these composites and solar irradiation acquired from SRTM DEM v2, we estimated the biodiversity index using a fuzzy model [Popov et al, 2008]. The resulting product is a georeferenced file in GeoTIFF format showing biodiversity index over the given region. The workflow of the data processing step is controlled by the Karajan engine while the data are processed on the computational resources of the Grid system using the GRAM service [Shelestov et al, 2006; Kussul et al, 2009; Hluchy et al, 2010].

The proposed Web service is implemented on the basis of OGC standards, Web Map Service 1.1.1 (http://www.opengeospatial.org/standards/wms) and Web Coverage Service 1.0 (http://www.opengeospatial.org/standards/wcs). The developed Web service is accessible via Internet through the address http://inform.ikd.kiev.ua/biodiv/ (Figure 5). It represents current distribution of the potential biodiversity and allows monitoring each of the factors that influence biodiversity.

Vegetation state assessment. A cascade of models is used to vegetation state assessment (Figure 6). This includes a regional NWP model WRF that was described in the previous subsection and comprehensive land surface model (Noah). Remote sensing observations along with ground measurements are assimilated into these models to derive meteorological parameters (temperature, rainfall), land and soil parameters (moisture and temperature). Additionally, satellite-based products (for example, vegetation indices) are used monitor vegetation state.

Such an approach was used to monitor sever droughts that hit Ukraine in spring-summer 2007. Consequences were catastrophic: 1,4 million ha of crops totally destroyed, 8,5 million ha of crops damaged, 100 million of U.S. dollars losses. The use of the proposed approach allowed us to identify regions that were mostly affected by the disaster, and estimate potential losses. Figure 7 shows comparison of vegetation index of 2007 and 2006.

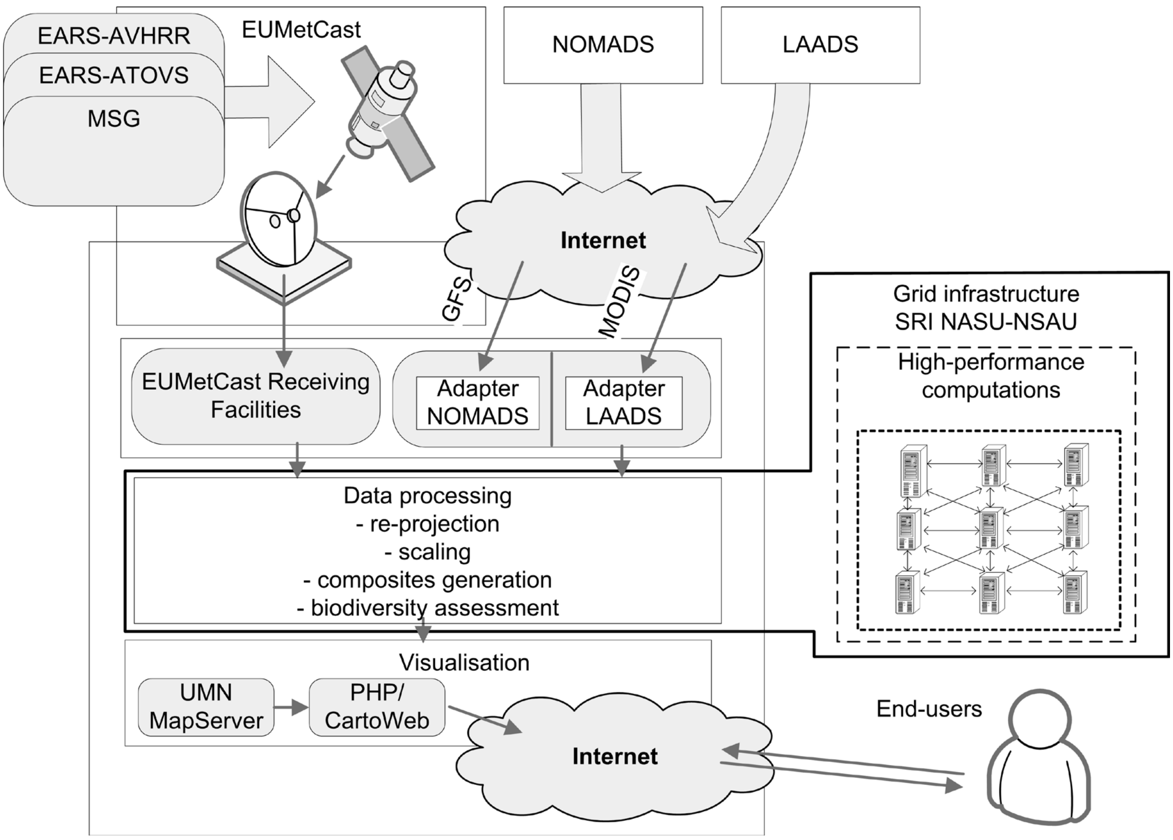

Figure 4. Overall architecture of the service with information flows

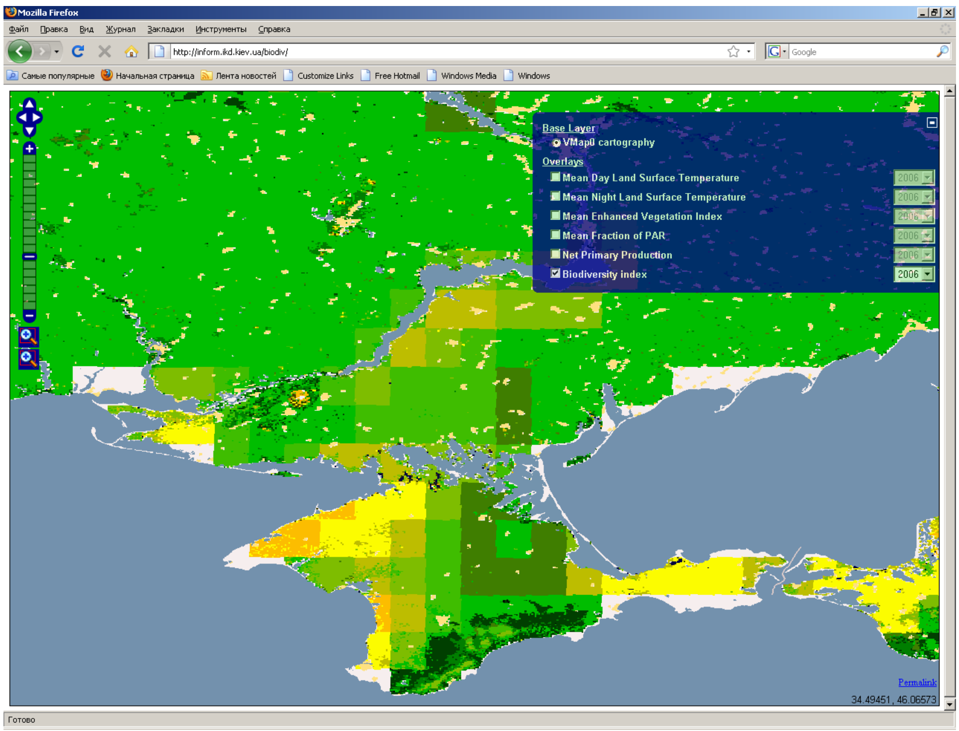

Figure 5. Demonstration of Web service for biodiversity assessment using EOS data products for the Pre-Black Sea region of Ukraine

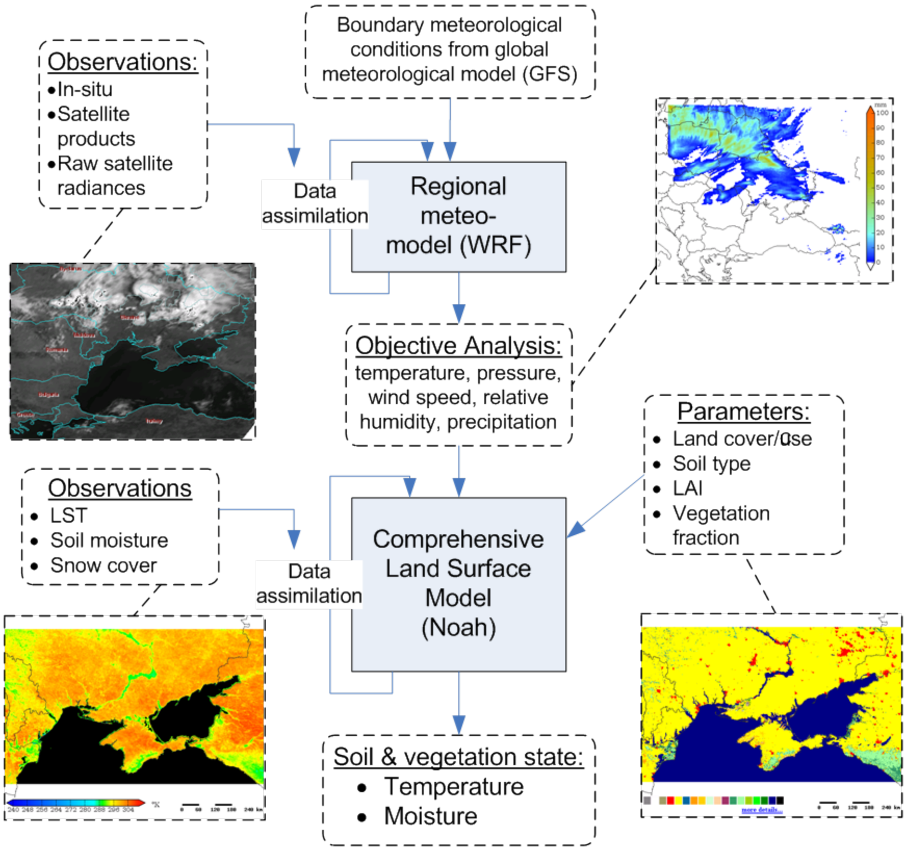

Figure 6. Modelling cascade for drought monitoring in Ukraine

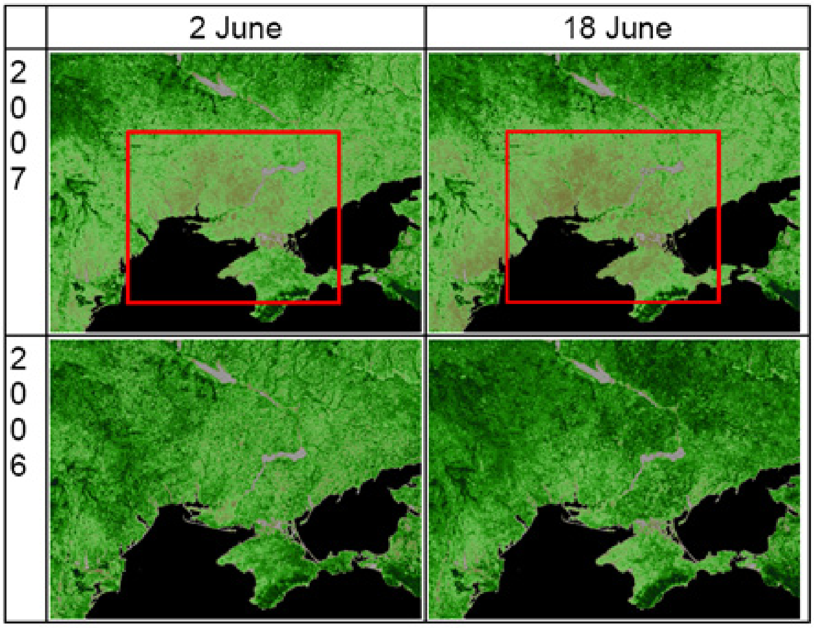

Figure 7. Evolution of Enhanced Vegetation Index (EVI) in the 2006 and 2007 vegetation seasons. Drought affected territories are highlighted by a rectangle

Fire monitoring. In July-August, 2010, Ukraine suffered from fires due to extremely high temperature: +35-39 C in Eastern regions and +40-42 C in South regions. On average 200 fires per day were detected. There was high risk of forest fires and fires approaching ammunition depots. Operational monitoring of fires was carried out using the following datasets:

- EO-1/ALI data acquired through Sensor Web prototype (date: 14.08.2010 08:15UTC)
- Landat-5/TM (date: 02.08.2010 08:15UTC)
- ZKI Fire Service that is available on daily basis and is using MODIS instrument onboard Terra & Aqua satellite.

The data products were extracted specifically for the territory of Ukraine. MODIS products were operationally delivered twice per day while other products were delivered on demand for the regions with the highest risks of fires. Cross-validation of MODIS and Landsat-5 products was done and shoed good correspondence between data (Figure 8).

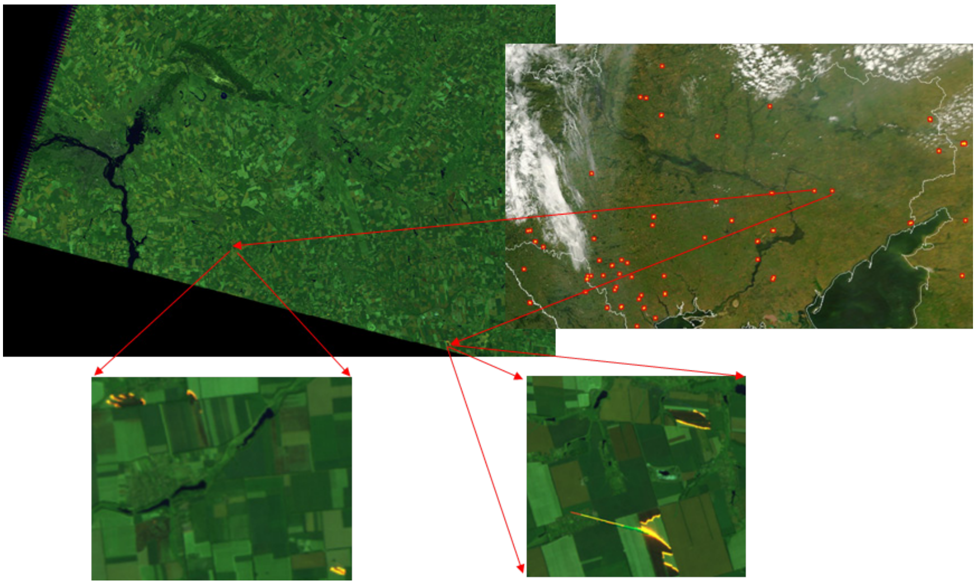

Figure 8. Cross-validation of fire products from MODIS and Landsat-5

International projects within UN-SPIDER RSO in Ukraine

UN-SPIDER is the United Nations Platform for Space-based Information for Disaster Management and Emergency Response and aims at providing universal access to all types of space-based information and services relevant to disaster management. The UN-SPIDER Regional Support Office (RSO) in Ukraine was established on basis of Space Research Institute in 2010. The RSO in Ukraine provides expertise in satellite data processing and product generation, operational delivery of services in case of emergency situations, and training activities. The RSO in Ukraine is actively involved in international projects. One of such a project is the Namibian Pilot on integrated flood management and water related vector borne disease modelling. Within this project on of the main tasks is flood risk assessment based on heterogeneous data.

These data are (Figure 9):

- Satellite imagery: synthetic-aperture radar (Envisat/ASAR, Radarsat-2), optical (EO-1, MODIS, Landsat-5), TRMM
- Modelling data: meteorological data (numerical weather prediction), hydrological data (river catchments).
- In-situ observations and river gauges: rainfall and river flow rate
- Statistical data: statistical information on floods for previous years.

The integration of different products is shown in Figure 10.

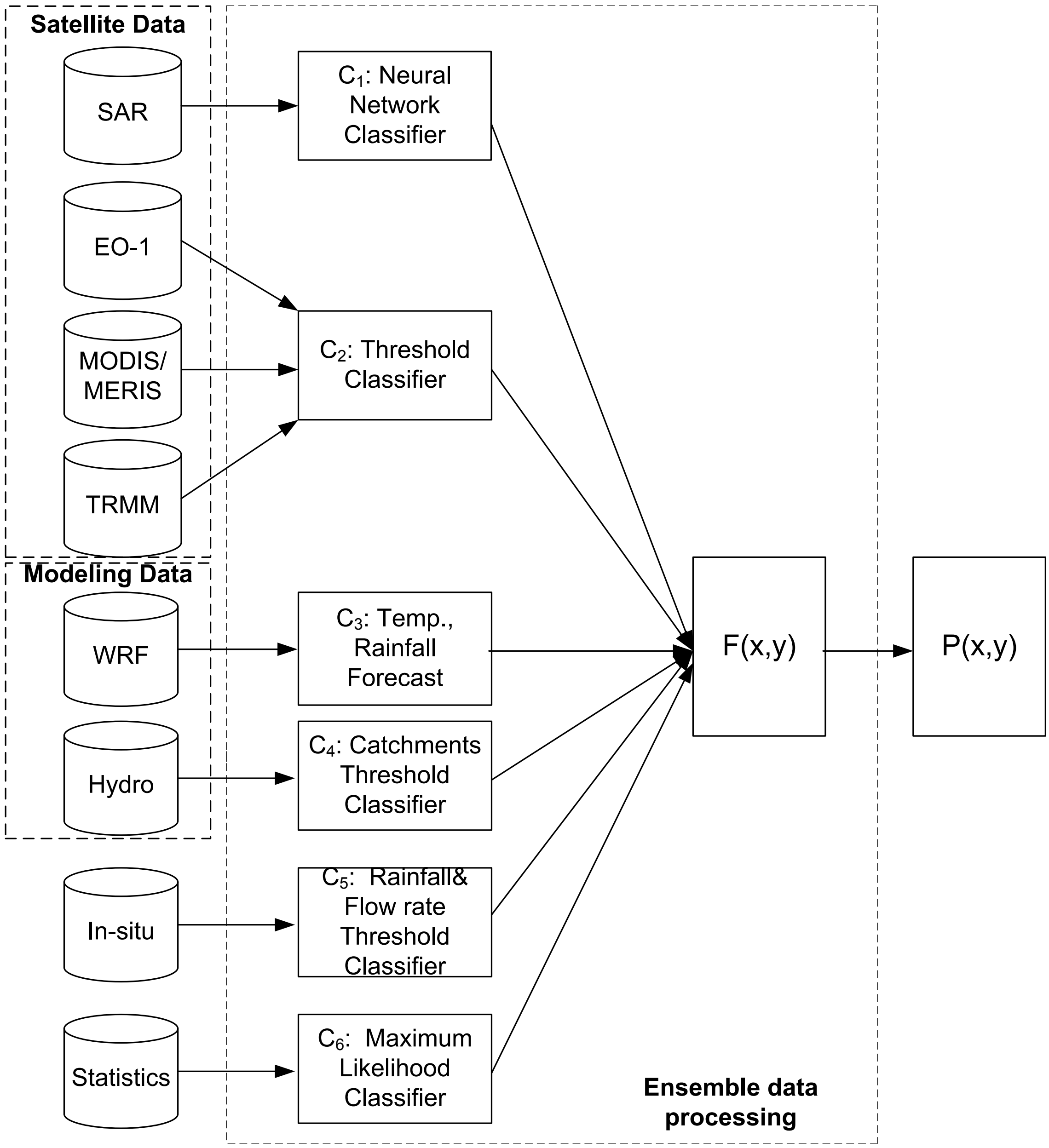

Figure 9. Integration of multisource data to flood risk assessment for the Namibian project

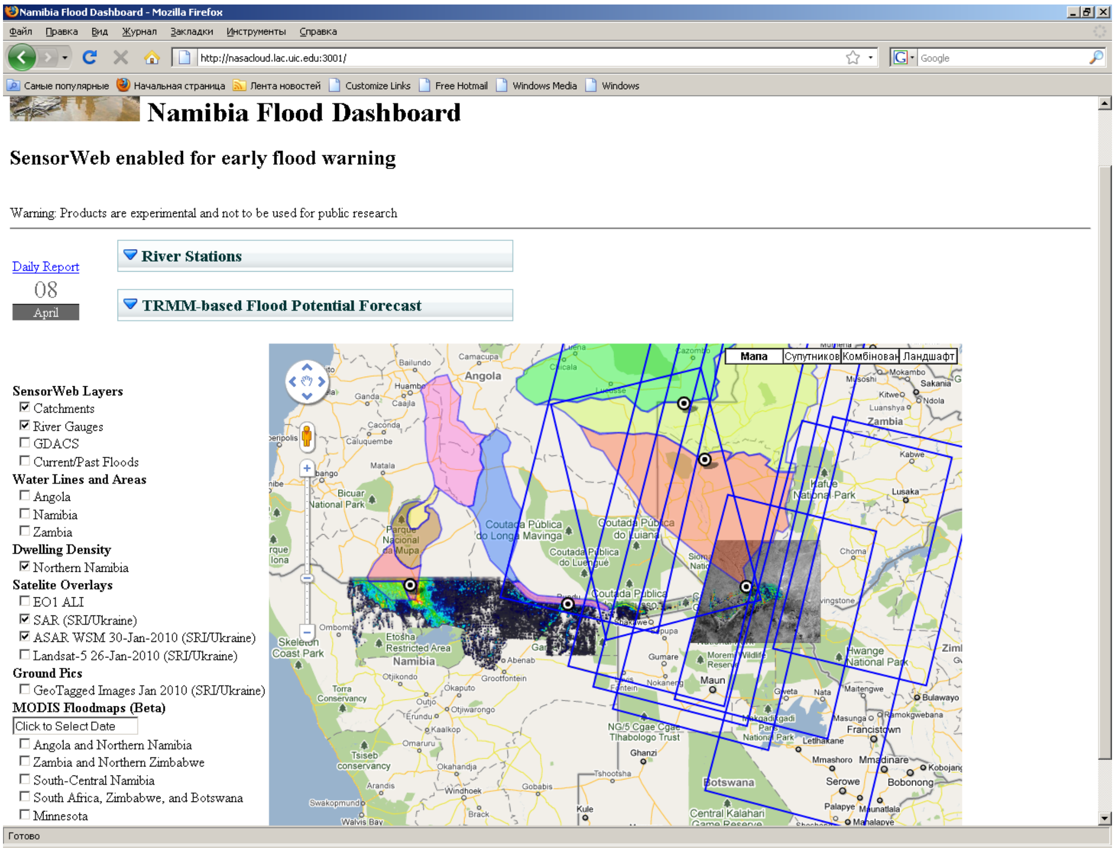

Figure 10. Namibian pilot project portal

Conclusions

In this paper we presented intelligent methods and corresponding technologies for environmental risk assessment. The risk assessment process is based on fusion of data acquired from different sources: models, in-situ observations and remote sensing instruments. The concept where the same data sets are applied for different applications is used. Therefore, once interfaces to discover and access the required data and products are developed, they can be used in a uniform way for different purposes and applications. This provides a basis for effective and operational exploitation of data.

The mixtures-of-experts concept for environmental risk assessment is introduced. Different experts provide a partial decision on the event using corresponding data, and their opinions are combined through some generalized rule. This allows for the problem to be broken into smaller sub-problems, and these sub-problems might be easier to solve than the overall problem.

Several real-world applications are described to demonstrate efficiency of the proposed approach, namely *numeral weather prediction* (NWP), *land biodiversity assessment*, *vegetation state assessment*, *fire monitoring* and *flood mapping*. Most of these applications are being implemented within international projects within the UN-SPIDER Regional Support Office (RSO) in Ukraine (http://un-spider.ikd.kiev.ua).

Bibliography

[Bishop, 2006] C.M. Bishop. Pattern Recognition and Machine Learning. New York: Springer Science+Business Media, 2006. 738 p.

[GEOSS, 2005] GEOSS, Global Earth Observation System of Systems. 10-Year Implementation Plan: Reference Document. ESA Publication Division, Netherlands, 2005.

[GMES, 2008] Global Monitoring for Environment and Security (GMES):we care for a safer planet. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Brussels, COM (2008) 748 final.

[Hansen and Rotella, 1999] A.J. Hansen, J.J. Rotella. Abiotic factors. In: Hunter ML (ed) Maintaining biodiversity in forest ecosystems, Cambridge: Cambridge University Press, pp 161-209, 1999.

[Haykin, 1999] S. Haykin. Neural Networks. A comprehensive Foundation. — New Jersey: Prentice Hall, 1999.

[Hluchy et al, 2010] L. Hluchy, N. Kussul, A. Shelestov, S. Skakun, O. Kravchenko, Y. Gripich, P. Kopp, E. Lupian. The Data Fusion Grid Infrastructure: Project Objectives and Achievements. Computing and Informatics, 29(2), pp. 319-334, 2010.

[Horritt, 2006] M.S. Horritt. A methodology for the validation of uncertain flood inundation models. J. of Hydrology, 326, pp. 153-165, 2006.

[Jacobs et al, 1991] R.A. Jacobs, M.I. Jordan, S.J. Nowlan, G.E. Hinton. Adaptive mixtures of local experts. Neural Computation, Vol. 3, pp. 79-87, 1991.

[Kogan et al, 2004] F. Kogan, R. Stark, A. Gitelson, E. Adar, L. Jargalsaikhan, C. Dugrajav, S. Tsooj. Derivation of Pasture Biomass in Mongolia from AVHRR-based Vegetation Health Indices. Int. J. Remote Sens, 25(14), 2889-2896, 2004.

[Kussul et al, 2009] N. Kussul, A. Shelestov, S. Skakun. Grid and sensor web technologies for environmental monitoring. Earth Science Informatics, 2(1-2), pp. 37-51, 2009.

[Kussul et al, 2011] N. Kussul, A. Shelestov, S. Skakun. Flood Monitoring on the Basis of SAR Data. In F. Kogan, A. Powell, O. Fedorov (Eds.) “Use of Satellite and In-Situ Data to Improve Sustainability”. - NATO Science for Peace and Security Series C: Environmental Security, pp. 19-29, 2011.

[Liang, 2004] Liang S. Quantitative Remote Sensing of Land Surfaces. Wiley, 2004.

[Michalakes et al, 2004] J. Michalakes, J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, W. Wang. The Weather Research and Forecast Model: Software Architecture and Performance. In: Proc of the 11th ECMWF Workshop on the Use of High Performance Computing In Meteorology (25-29 October 2004, Reading U.K), 2004.

[OHSAS, 2007] OHSAS 18001:2007 - Occupational Health and Safety Management Systems Requirements Standard, 2007.

[Popov et al, 2008] M. Popov, N. Kussul, S. Stankevich, A. Kozlova, A. Shelestov, O. Kravchenko, M. Korbakov, S. Skakun. Web Service for Biodiversity Estimation Using Remote Sensing Data. Int. J. of Digital Earth, 1(4), pp. 367-376, 2008.

[Shelestov et al, 2006] A.Yu. Shelestov, N.N. Kussul, S.V. Skakun. Grid Technologies in Monitoring Systems Based on Satellite Data. J. of Automation and Information Science, 38(3), pp. 69-80, 2006.

[Shelestov, 2008] A. Shelestov. Workflow Modelling in Grid System for Satellite Data Processing. International Journal on Information Theory and Applications, 15(3), pp. 225-231, 2008.

[Vapnik, 1998] V. Vapnik. Statistical Learning Theory. New York: Wiley, 1998.

[Wagner et al, 2007] W. Wagner, C. Pathe, D. Sabel, A. Bartsch, C. Kuenzer, K. Scipal. Experimental 1 km soil moisture products from ENVISAT ASAR for Southern Africa. ENVISAT & ERS Symposium, Montreux, Switzerland, 2007.

Authors' Information

Nataliia Kussul – *Deputy Director, Space Research Institute NASU-NSAU, Glushkov Prospekt 40, build. 4/1, Kyiv 03680, Ukraine; e-mail: inform@ikd.kiev.ua*

Sergii Skakun *– Senior Scientist, Space Research Institute NASU-NSAU, Glushkov Prospekt 40, build. 4/1, Kyiv 03680, Ukraine; e-mail: serhiy.skakun@ikd.kiev.ua*

Oleksii Kravchenko *– Scientific Researcher, Space Research Institute NASU-NSAU, Glushkov Prospekt 40, build. 4/1, Kyiv 03680, Ukraine; e-mail: oleksiy.kravchenko@gmail.com*