

DRIHMS

DISTRIBUTED RESEARCH INFRASTRUCTURE
FOR HYDRO-METEOROLOGY STUDY

THE WHITE PAPER



DRIHMS is co-Funded by the EC
under the 7th Framework Programme

www.drihms.eu



DRIHMS
is co-Funded by the EC
under the 7th Framework Programme



CIMA Research Foundation

Antonio Parodi
Marina Morando
Nicola Rebora
Eva Trasforini
Franco Siccardi

DLR

George C. Craig
Arnold Tafferner

LMU-MNM

Dieter Kranzlmüller
Michael Schiffers

IMATI

Andrea Clematis
Daniele D'Agostino
Antonella Galizia
Alfonso Quarati

Coordination

Antonio Parodi

Project & Design

Punto a capo Coö'ee

Print

SXS, April 2011

Information in this document is compiled and provided by the DRIHMS consortium. The consortium accepts no liability whatsoever with regard to the information in this document. The consortium may make changes to contents of this document at any time without notice. The consortium assumes no responsibility for the contents of external sites the document links to. Reproduction of this document is authorized, provided the source is acknowledged.

Copyright © 2011, CIMA Foundation, Italy, on behalf of the DRIHMS consortium

EXECUTIVE SUMMARY

EXECUTIVE SUMMARY

Although international agencies like United Nations (UN) or the European Commission (EC), national government organizations and local authorities increasingly ask for globally certified management tools to deal with extreme events of precipitation and floods, the scientific community is still reduced to the ability to communicate "scenes" to urban, regional and national decision makers. Unfortunately, such "scenes" are typically both information-light and emotion-heavy. Even scientists sometimes do realize that "scenes" are a "degradation" of scientific data. A "scenario" on the other hand is a time series of data up to a given moment - which require a reliable knowledge of the underlying phenomena - physics, chemistry, biophysics, sociology to trace back the causes of the observed effects. Or track forward, i.e. predict, how the scenario may evolve. The construction of a "scenario" must therefore consider the temporal dynamics and the interactions between processes in order to mitigate risks and disastrous vulnerabilities. Risk - defined as the probability that a

foreseen event, despite the actions taken to contrast it - instigates a certain degree of effects on the population and the environment. It is typically assessed by the amount of damage people and properties are exposed to. Consequently, for Agencies and Governmental Authorities it is of utmost importance to be able to issue fact-based warnings to possibly affected areas. Important prerequisites of such ability are observed data and formally sound models to supplement them. What we require in addition, however, is a reliable access to data archives and computational technologies for providing new experimental data. Without the support of an adequate Information and Communication Technologies (ICT) infrastructure such an undertaking will be difficult if not impossible. Therefore, the main objectives of the Distributed Research Infrastructure for Hydro-Meteorology Study (DRIHMS) project (www.drihms.eu) have been first to better understand the mindsets of both the Hydro-Meteorological (HM)

and the ICT communities and second to propose a way of closing perceived and existing gaps between the communities.

Both the mindsets and the gaps were identified and analyzed by a set of questionnaires to both communities. The Hydro-Meteorology Research (HMR) part of the survey clearly shows the paramount importance of solving probabilistic forecasting issues and of providing formally sound model verification metrics. Interestingly, real-time data processing was given only a secondary priority. The ICT part of the survey emphasized the importance of data management, workflow management and High Performance solutions over portals and management of Virtual Organizations (VO). Based on the survey results, complemented by an open meeting consultation and targeted interviews to well-recognized HMR and ICT experts, a rationale for an HMR e-Science environment can be formulated which can then be mapped to an HMR e-Science architecture. The e-Science environment consists of resources, simulation models, methods and tools to be provided by various institutions over European e-Infrastructures like European Grid Infrastructure (EGI), Partnership for Advanced Computing in Europe (PRACE) and the national National

Grid Initiatives (NGIs). On the other hand, however, the HMR e-Science environment will also have specific HMR requirements not yet or only partly addressed within the actual existing e-Infrastructures. Closing the identified gaps between both communities requires a respective long-term roadmap with short-term and mid-term objectives to achieve as synchronization points. The roadmap as layered out in this document proposes to first address the issues of community building, then to remove the blocking factors for deploying HMR applications onto European e-Infrastructures, and finally to really exploit the infrastructure.



2.0

INTRODUCTION

WHY DRIHMS?

INTRODUCTION

Extreme precipitation and flooding events are among the greatest risks to human life and property. UN agencies, national governments responsible for large regions, and local governments for their local needs, ask for globally certified management tools to deal with processes leading to extreme hydro-meteorological events. The response of the scientific community is today limited to the ability to communicate the processes through maps of observed or predicted variables, that is, the ability to communicate "scenes".

The images, or "scenes", need to be circulated rapidly over the media in an easily understandable form, in order to reach those responsible for taking decisions to support affected populations. A "Scene" must therefore be "light" from the information and communication technology perspective, with minimal complexity and data volumes. It is often not realized, even by scientists, the extent to which "scenes" are a degradation, for circulation purposes, of the scientifically sound observation and prediction data, which

can be better referred to as "scenarios". In contrast to a "scene", a "scenario" is a time series of spatially distributed data of a process up to a given moment. These data provide a basis to trace back the causes of what is observed, or track forward, i.e. predict, how the scenario will evolve. Not only that: the dynamics of the process derived from the scenario can be compared with other processes to draw conclusions about additional phenomena not directly observed. For example, estimates can be made of the impact of a flood on vulnerable structures, agricultural production, or local economy that will remain even after the inundation has retreated.

The construction of a "scenario" must then consider the temporal dynamics and the interaction between different physical and other processes - chemical, biophysical, social - and convey this information. To circulate a "scenario" requires distribution of certified data and certified models, and perhaps certified model outputs. The burden of such a task is heavy for the hydro-meteorological community, but becomes impossible

without the support of an ICT structure able to deal with the massive amount of information needed.

An "extreme rainfall scenario" also has uses beyond response to individual events. The social perception of flooding and/or landslide risk is poor, and as a result population density and urbanization continue to increase in flooding and landslide prone areas around the planet, uninhibited by regulations to adapt the use, to the flood regime, the stability of slopes and the risk. All over the world and in the EU especially, new environment policies are presently making efforts towards more vigilant regulations in managing the territorial risk. Civil Protection, or Civil Protection-like organizations, are calling for limitations on urban development, together with environmentally sustainable renovation of watercourses and slopes.

However, such limitations are still perceived by land owners as a huge economic burden, and conflicts increase. Only a few citizen-scientists share the concerns and accept restrictive policies.

Agencies dealing with territorial risk, in a modern, post-industrial society, have two major tasks. The first one is sociological, in the sense that the new

organizations must closely interact with other social institutions in order to get consensus on the restrictions of land use and additional limitations. The second task is technological, in the sense that Civil Protection organizations must make use of the most advanced and efficient tools to predict, forecast and observe the ground effects affecting people and property. Sophisticated "warning scenarios" not only serve immediate needs in a crisis, but also establish the credibility of the organizations, furthering the development of consensus on needed regulations.

These needs fall within the scope of the Lisbon strategy (European Council, 2005), which identified knowledge and innovation as the engines of sustainable growth and stated that it is essential to build a fully inclusive information society.

In parallel, the World Conference on Disaster Reduction (Kobe, Hyogo, 2005), defined among its thematic priorities the improvement at a worldwide level of trans-national cooperation in basic and applied hydro-meteorology research activities for the prevention and mitigation of risk associated with the occurrence of extreme hydro-meteorological and hydrogeological events.

Progress will require bringing together advances in hydro-meteorological research HMR and information and communication technology (ICT). Hydro-meteorological science has made strong progress over the last decade at the European and worldwide level: new modelling tools, post processing methodologies and observational data are available. Meanwhile, recent European efforts in developing a platform for e-Science including EGEE (Enabling Grids for E-science), SEEGRID-SCI (South East Europe GRID e-Infrastructure for regional e-Science), and EGI (European Grid Initiative), provide an ideal basis for the sharing of complex hydro-meteorological data sets and tools. Despite these early initiatives, however, the awareness of the potential of the Grid technology as a catalyst for future HMR is still low and both the adoption and the exploitation have been slow, not only within individual European Community member states, but also on a European scale.

The DRIHMS project was initiated to provide a detailed analysis of the needs for such a step forward in the cooperation between HMR and ICT scientists, and to provide a foundation for the timely deployment of an e-

Science environment for hydro-meteorology research on extreme events, as well as on climate change impacts and adaptation.

In the following we will discuss first where we are and what gaps must be filled, centered on the views expressed by outstanding scientists and policy makers in the two fields of hydro-meteorology and information and communication technology. We then define a set of requirements for the establishment of a reliable hydro-meteorology research Grid, oriented to sound prediction, both for real time warnings and for planning. Finally we present a roadmap for the hydro-meteorology research Grid and evaluate the impact on such a research community.

3.0

GAP ANALYSIS

HYDROMETEO RESEARCH NEED
A NEW E-INFRASTRUCTURE

3.1 A SCENARIO OF RISK

We begin with a general formulation of the task, avoiding for the moment the jargon of HMR or ICT science.

Risk, in the present context, is defined as the probability that a foreseen event, despite actions taken to counteract it, results in certain effects on the population and its property, and in some cases on the whole environment. Risk is typically estimated on the basis of the vulnerability and the value of the

exposed elements. The scenario of risk is the amount of damage people and property could suffer over a specific area in a specific time after the occurrence of a foreseen physical event. The probability of the occurrence of a scenario of risk is the conditional probability of the occurrence of that amount of damage, to people and properties, given the occurrence of the event. It is therefore necessary to establish, on one hand, the probability of the scenario of risk conditioned to the

CONCEPT OF AN EVENT AND ITS PROBABILITY OF OCCURRENCE

Consider, for example, a specific river basin, say the Tiber River, and in particular the section of Ponte Milvio, the Milvio Bridge, in Rome. Flood waves are possible at the Ponte Milvio: the river flow and its stage can increase in a few days' time to reach the maximum, stay at this level for a period and then gradually decrease to moderate flow values. When a specific stage is exceeded the river inundates the centre of Rome, as it did for centuries.

The local Authorities follow an established procedure: every day of the year they observe the river stage and the corresponding flow

rate. Then they register the maximum flood peak for the year in the book of hydrological records. For each individual year they obtain a specific realization of the maximum annual flow rate of the Tiber River at Ponte Milvio. The set of all possible results defines the sample space of the event of a possible flood of the Tiber River at the Ponte Milvio. The data is then sorted in increasing order. If the number of the experiments - i.e. years of observation -- is enough, they infer the probability that in the next experiment which has not been carried out yet, i.e. next year, the flow rate will be larger than a given

threshold. Such a probability is estimated as the observed frequency of the maximum annual flow rate that exceeded the threshold in the past. Frequently they fit a model to the data, in order to be able to extrapolate to rare events they did not observe. The extrapolation introduces the first component of the hydrological uncertainty, growing with the relative rarity of the event the hydrologists would like to estimate: if they extrapolate too much the outcome is more a hydrological gentlemen agreement than a hydrological prediction.



event, and, on the other, the unconditioned probability of the event. That an Agency or a Governmental Authority be able to issue warnings to the potentially affected population on the basis of a sound probabilistic procedure is of foremost importance from a societal point of view, and furthermore the formal establishment of a probabilistic procedure is a safeguard for decision makers with respect to their liability. Finally, knowing that the warnings are issued based on sound procedures, increases confidence in the Authority. The box "Concept of an event and its probability of occurrence" clarifies the

concept of an event and its probability of occurrence using a very simple example. In reality, the concept of events is significantly more complex. The real situation of possible flooding of a generic river in Europe, in terms of real-time prediction and also for planning purposes, is even more complex, since the hydraulic evaluations are only part of the problem. In fact, an event should be defined as the ensemble of many possible realizations of atmospheric conditions, hydrological conditions of the catchments targeted by the storm, and hydraulic conditions of the streams and rivers that drain the catchments.

THE COMPLEXITY OF THE CONCEPT OF EVENTS

The physical quantities that must be observed to describe hydrological events include not only the maximum flow rate but many more: the evolution in time of the flow, the hydrograph, the correlated rate of bed load, which importantly affects the river bed; the correlated rate of load floating on the surface, which also importantly affects the capacity under bridges such as Ponte Milvio. Therefore, the flooding event cannot be described by a single scalar variable. But in many cases the desired measurements are not available for the events of the past. It is therefore impossible to infer their probability on the basis of the observed frequency. In such conditions hydrologists estimate the conditional probability of observing, for example, high surface transport,

given the value of the flow rate. The term "probability" must be understood in such case as being based on the expectations of a knowledgeable expert, instead of one which is based on the observed frequency of the phenomenon. In fact, the hydraulic engineer expects, in the case of modest peak flow rates - i.e. after rainfalls of small entity with moderate saturation of the mountain slopes - the rate of load carried by the current on the bottom to be very small and the rate of load floating on the surface to be negligible. In both cases the competent expert considers the occurrence of landslides due to slope sliding and contributing solid material and vegetation to the riverbed to have very low probability. However, when the flow rate is high the knowledgeable expert

expects surface slides to be much more likely, the phenomena of erosion and transport to be more relevant, and the on-surface transport of shrubs, tree trunks and brushwood to be extremely likely. The real situation of possible flooding on a perfluvial area of a generic river in Europe, when required for predicting in real time but also for planning purposes, is even more complex. The hydraulic evaluations are only part of the problem. In fact we should define an event as the ensemble of many possible realizations of atmospheric conditions, of hydrological conditions of the catchments targeted by the storm, and of hydraulic conditions of the streams and rivers that drain the catchments.

3.2 MODELING INSTEAD OF OBSERVING

No case exists in the world in which observations have been so complete and, most of all, the experiments have been repeated so often at one individual site, to obtain correct statistics that allow inference of the probability of each one of the components conditioned to the occurrence of the previous one. Therefore, from what has been described so far, we are compelled to replace the observation of the results of the same experiment over a specific section with a simulation of the chain of physical processes in which, starting with a predicted weather pattern, the distribution of a conditioned probability of a flood of extension and depth above a critical threshold can be quantified. For each flood-prone area of each country, the procedure described above could be repeated as often as necessary, with different expressions and values of the parameters of probability distributions, in order to build by simple enumeration the probability-conditioned distributions of the desired random variables. Alternatively, probability-conditioned distributions of the parameters could be estimated by real-time simulation of the chain of the physical processes, from meteorology to flooding.

It seems necessary that the scientific

communities of meteorology and hydrology should establish a set of tools to address the challenge of global modeling of hydro-meteorology ground effects. Centers of Excellence in each of the two disciplines must be able to cooperate by sharing most advanced models and observational datasets. This White Paper proceeds from a number of hypotheses regarding our ability to meet the challenge of hydro-meteorological prediction.

The first hypothesis is that we have sufficient knowledge, including observational and modeling data to perform such a task.

The second is that we lack fast and reliable enough access to the necessary archiving and computational technologies to allow us to combine this data into reliable prediction systems.

The third hypothesis is that scientists in the two discipline areas, HMR and ICT are not sufficiently aware of the requirements and expectations of the other group, but that there is a great willingness to improve the situation.

The building of a system in which centers of excellence, in hydro-meteorology and other Earth science related disciplines, can work together, exchanging output of models and observational data, has the potential to enable exponential growth in the research capacity of the cooperating centers. If the system continues to be

open to the participation of new research groups, provides for the needs of stakeholders not primarily interested in research, from operational agencies to the citizen-scientist, the system will satisfy a demand that until now has not found satisfaction.

These hypotheses have been examined and validated through a series of analyses that are summarized in the following sections, leading finally to a roadmap for the development of a new ICT infrastructure for HMR.

3.3 THE DRIHMS APPROACH

In order to address these hypotheses, the DRIHMS project organized a set of networking activities, involving both HMR scientists and ICT scientists related to Grid Computing.

The objectives were to:

- understand how to overcome the current limitations in the sharing of data, tools and knowledge in the European HMR community;
- create a common understanding of what is available;
- map out an e-Science based path to the generation of new knowledge from the latest generation of hydro-meteorological observing and modeling systems.

An important means of collecting the opinions and suggestions from the HMR and the ICT communities was

through a series of questionnaires, aimed at obtaining feedback regarding the topical areas and their technological requirements. This White Paper represents a synthesis of and a reflection on these feedbacks. Identifying the Grid-related gaps for innovative hydro-meteorological research was a cumbersome task as it firstly required an understanding of the real HMR needs, and secondly required a thorough understanding of the strengths and weaknesses of modern Grid (and Cloud) technologies: Figure 1 shows the consultation process, using networking instruments (consultation meetings, questionnaires, open meeting, and communities polling methodologies, competitive analysis with respect to other projects) and driven by the following work packages (WP) of the DRIHMS project:

- **HMR hot topics identification** (WP2), devoted to the identification of hydro-meteorological hot research areas that require a network-based and distributed approach, in terms of hydro-meteorological data and software sharing;
- **Grid Support for HMR** (WP3), devoted to the revision and identification of the key components of HMR applications in the hot areas identified in WP2 and representing its technological counterpart;

- **Prioritization and White Paper production (WP4)**, representing a point of reflection and ongoing consolidation of the DRIHMS project activities, based on HMR Grid usage patterns and ICT questionnaire results.

Process-wise, the DRIHMS consortium involved very early the international HMR community, the international Grid community and selected experts the DRIHMS Advisory Board composed of named HMR and Grid scientists- in order to leverage the experience which has already been gained in related work.

A number of hydro-meteorology

research projects had been financed at the European level in recent years (see table in next page).

Some of them are still active, but all of them involved a number of Centers of excellence that agreed to contribute to HMR survey (see the following table). Taken together the aforementioned projects provide a cross-section of the current and future needs and challenges of HMR at the European and global level. These needs and challenges can be categorized in terms of:

- **Observation strategy:** monitoring of all relevant atmospheric, oceanic, hydrological and bio-chemical

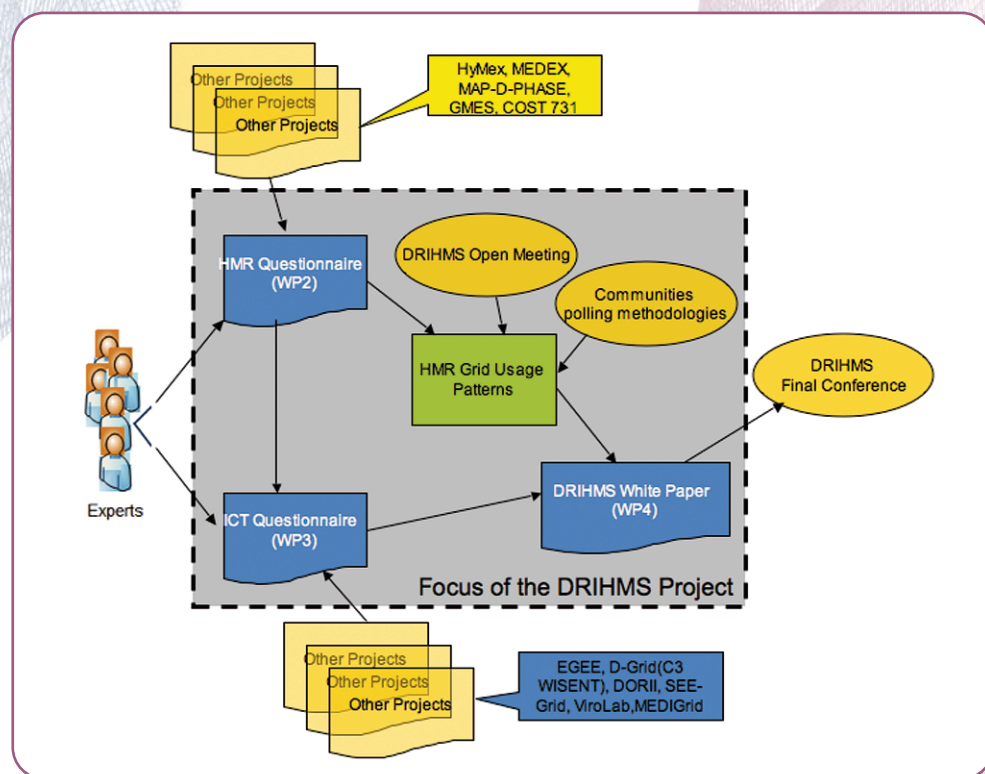


Figure 1: Consultation process rationale.

PROJECT NAME	PROJECT DESCRIPTION	WEB-SITE
HyMeX	HYdrological cycle in the Mediterranean eXperiment aiming at a better understanding and quantification of the hydrological cycle and related processes in the Mediterranean, with emphasis on high-impact weather events, inter-annual to decadal variability of the Mediterranean coupled system, and associated trends in the context of global change	http://www.hymex.org/
COST 731	Propagation of uncertainty in advanced meteo-hydrological forecast systems - aiming at addressing issues associated with the quality and uncertainty of meteorological observations from remote sensing and other potentially valuable instrumentation, as well as their impacts on hydro-meteorological outputs from advanced forecasting systems	http://cost731.bafg.de/servlet/is/Entry.9691.Display/
MAP-D-PHASE	Mesoscale Alpine Programme - Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events aiming at demonstrating some of the many achievements of the Mesoscale Alpine Programme (MAP), in particular the ability of forecasting heavy precipitation and related flooding events in the Alpine region. It addressed the entire forecasting chain ranging from limited-area ensemble forecasting, high-resolution atmospheric modelling (km-scale), hydrological modelling, and nowcasting to decision making by the end users, by setting up an end-to-end forecasting system	http://www.map.meteoswiss.ch/map-doc/dphase/dphase_info.htm
MEDEX	MEDiterranean Experiment designed to contribute to the better understanding and short-range forecasting of high impact weather events in the Mediterranean, mainly heavy rain and strong winds	http://medex.aemet.uib.es/

PROJECT NAME	PROJECT DESCRIPTION	WEB-SITE
FP6 PREVIEW	Prevention Information and Early Warning looking for new techniques to better protect European citizens against environmental risks and to reduce their consequences	http://www.preview-risk.com
CLIVAR	CLimate VARIability and Predictability: a programme of the World Climate Research Programme aiming at developing our understanding of climate variability, to apply this to provide useful prediction of climate variability and change through the use of improved climate models, and to monitor and detect changes in our climate system	http://www.clivar.org
FLOOD	Integrated Flood Risk Analysis and Management Methodologies covering the physical, environmental, ecological and socio-economic aspects of floods from rivers, estuaries and the sea. It considers flood risk as a combination of hazard sources, pathways and the consequences of flooding on the "receptors" - people, property and the environment	http://www.floodsite.net
GMES	Global Monitoring for Environment and Security GMES services are dedicated to the monitoring and forecasting of the Earth's subsystems. They contribute directly to the monitoring of climate change. GMES services also address emergency response (e.g. in case of natural disaster, technological accidents or humanitarian crises) and security-related issues (e.g. maritime surveillance, border control)	http://www.gmes.info/
IMPRINTS	IMproving Preparedness and Risk maNagemenT for flash floods and debris flow events - devoted to contribute to the reduction of loss of life and economic damage through the improvement of preparedness and operational risk management of flash flood and debris flow generating events, as well as contributing to sustainable development through reducing damages to the environment	http://www.imprints-fp7.eu/

PROJECT NAME	PROJECT DESCRIPTION	WEB-SITE
FloodProBE	Flood Protection of the Built Environment, Technologies for Improved Safety of the Built Environment in Relation to Flood Events aiming at developing technologies, methods and tools for flood risk assessment and for the practical adaptation of new and existing buildings, infrastructure and flood defenses leading to a better understanding of vulnerability, flood resilience and defense performance	http://www.floodprobe.eu/
PRISM	Partnership for Research Infrastructures in earth System Modeling devoted to help share the development, maintenance and support of standards and state-of-the-art software tools to assemble, run, and analyse the results of Earth System Models based on component models (ocean, atmosphere, land surface, etc..) developed in the different climate research centers in Europe and elsewhere	http://www.prism.enes.org/

variables during a long observation period, with additional means deployed during enhanced and special observing periods. In particular field campaigns in the near future should provide a unique high-resolution database to validate the new-generation of high-resolution models that have recently been introduced by national meteorological centers in their operational Numerical Weather Prediction (NWP) suite. Such data include: microphysical properties obtained by polarimetric radars or aircraft measurements, marine boundary layer characteristics and air-

sea flux measurements by buoys or by instrumented research vessels, novel or high-resolution moisture measurements (GPS delays on board ships, radar refractivity, water vapor from lidar, etc);

- **Modeling strategy:** The improvement of convective-scale deterministic forecast systems for better predicting high-impact weather events over complex orography areas (i.e. Mediterranean); the design of high-resolution ensemble modeling systems dedicated to the study of the predictability of severe hydro-meteorological events; the coupling of these ensemble forecast systems with

hydrological models to issue probabilistic forecasts of the impact in terms of hydrological response; the production of new hydro-meteorology knowledge in terms of new process modeling, parameterization development, novel data assimilation systems for the HMR community and related Earth science disciplines. These strategies result in important challenges from the computing, archiving and data processing viewpoints, as discussed in the next section.

In contrast to the HMR community, the relevant contact points in the ICT community were not determined by their involvement in research projects. Instead, it is more important to cover special topics like Grid infrastructure experience, Grid middleware know-how, and Grid operational expertise. Consequently, the expertise of the ICT Grid community was sought in areas including fast networking via Grid middleware (Globus, UNICORE - Uniform Interface to Computing Resources, gLite-Lightweight Middleware for Grid Computing) and via basic services (authentication, authorization), to complex services like data management, workflow management, and the management of VO.

The general methodology for interrogating the communities is described in specific reports available from the project's home page; the

structure and the contents of the HMR and ICT questionnaires, are also available at <http://www.drihms.eu/survey>.

The questionnaires are intended as a way to collect the opinion of experts involved by mean of email invitation. As these respondents build a non-random sample, the results are not projectable to any population other than the experts themselves. However, since the results showed a significant correlation between the individual answers, the sample is justifiably representational.

The list of contributing Institutions is provided in the Annex.

3.3.1 HMR QUESTIONNAIRE RESULTS

The purpose of the HMR questionnaire was to define important topics in HMR that can benefit from advances in ICT, and thus set a joint HMR-ICT research and applications agenda.

The questionnaire aimed to capture perceptions at the European and worldwide levels concerning the increasing challenges in HMR, including the ability to retrieve and access data from different sources and different countries, stored in different formats, to use in a seamless way suites of hydrological and meteorological models each with its own input/output formats, to exploit computational resources for on-line operations by individual

groups, to implement objective methodologies for downscaling model outputs to local conditions, compare model performance objectively, to store data from field campaigns in standard formats for easy access by the larger research community, to facilitate operational forecasting, etc. Just under 200 questionnaires were collected (from 40 countries): 82% of questionnaires were received from European Institutions, while the remaining 18% came from overseas, mainly from the USA. The distribution of sources within the EU was influenced slightly by the nationality of the coordinating partner, but all major branches of Hydrology and Meteorology are well represented. At the European level, the leading countries in terms of number of collected questionnaires were: Italy (20%), Germany (11%), France (9%), Spain (9%) and UK (4%). Most of the respondents were from the fields of Hydro-Meteorology (40%) or Meteorology (43%), with a smaller but still significant contribution from Hydrology (10%). About half of the respondents were from research institutions (47%), with the remainder from institutions with both research and operational responsibilities (38%) or purely operational institutions (15%). The survey respondents revealed clear choices of hot topics and

accompanying ICT priorities:

- The most important hot topics for HMR research were identified as probabilistic forecasting (particularly among meteorologists) and model verification metrics (particularly among hydro-meteorologists and hydrologists);
- For probabilistic forecasting, the most important ICT challenges were the definition of common interchange formats, definition of libraries of tools for data handling and visualization, and the availability and reliability of high-performance computing resources;
- For model verification metrics, the key ICT challenges were availability of model outputs and observations in compatible formats, and the availability of libraries of well-documented verification metrics;
- The replies regarding current practices yielded a large variety of methods of working, clearly showing the central role of the processing and communication of large data sets in all aspects of hydro-meteorological research;
- Interestingly, real-time data processing was given only a secondary priority by this research-oriented community, and accounting and billing issues seem to be regarded as insignificant.

3.3.2 ICT QUESTIONNAIRE RESULTS

The purpose of the ICT questionnaire was twofold: first, to understand the ICT implications of the HMR hot topics; second, to explore the current ICT possibilities for their applicability for HMR.

The questionnaire attracted 81 respondents from Europe (including Cyprus), Ukraine, Brasil, Korea, Taiwan and the US.

The most important results related to the HMR hot topics are:

- The respondents perceive data management as very important but they do not see significant progress in the next years.
- High Performance Computing (HPC) is perceived important and they expect significant progress within the next years.

- Workflow management is perceived important but no significant progress is expected even short term.
- Portals and user interfaces are perceived important and the existing solutions seem to fulfill most of the requirements already.
- Virtual Organization (VO) management is perceived to be less important but sufficiently mature already.

Consequently, while VO management seems to be of second concern, the ICT requirements for HMR should reflect the specific needs for cross-organizational workflows and cross-organizational data management and it must address the various interoperability issues (see also Figure 2 where higher numbers on the axes denote "more important" or "more mature", respectively).

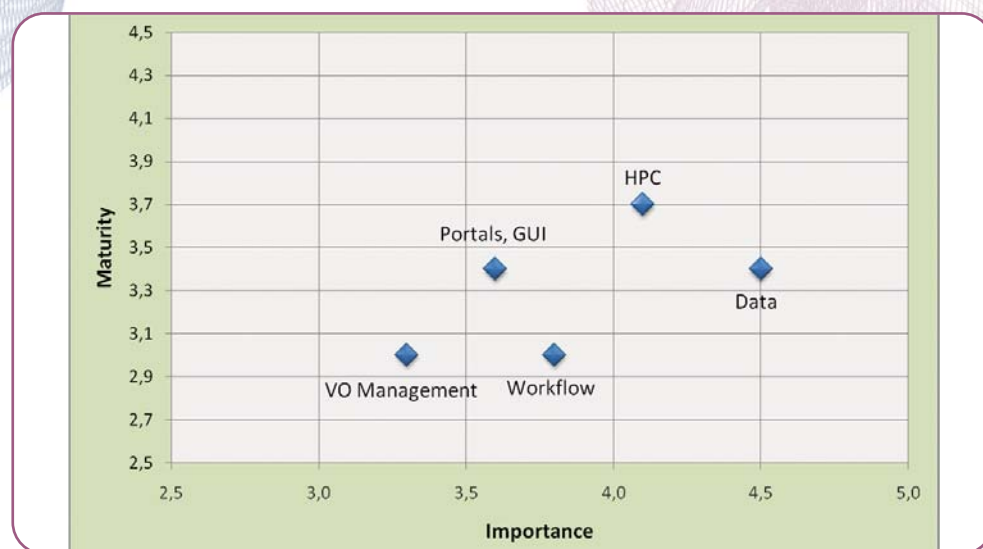


Figure 2: Correlation between importance and availability of hot topics

4.0

SELECTED HMR GRID USAGE PATTERNS

MORE ACCESSIBILITY,
MORE SHARING,
MORE COMPUTING POWER:
YOUR DREAMS ARE OUR DRIHMS!

Based upon the advice of respondents experts (see section 3.3), we identified a number of Hydro-Meteorological Research HMR hot topics, which are driving future developments in hydro-meteorological forecasting. In order of surveyed importance, these are:

1. probabilistic forecasting, particularly the use of large ensembles of forecasts to characterise uncertainty;
 2. model verification metrics, including advanced spatial error measures that consider structures of the forecast fields;
 3. data merging/fusion, the combination of large data sets with diverse properties;
 4. precipitation downscaling, statistical or dynamical methods to add realistic small-scale variability to forecasts.
- These trends are accompanied by a dramatic increase in the quantity and complexity of the tools and data sets available for hydro-meteorological forecasting. There are three reasons for this development:
- remote sensing observations from satellites and from ground-based radars that provide complete three-dimensional coverage of the atmospheric and land surface state, vastly increasing the quantity of data;
 - ensembles forecasting methods, that combine multiple numerical weather prediction and hydrological models to quantify the uncertainty in the forecast, multiplying the

computational costs;

- increasing recognition of the need to understand the entire forecasting chain, from observations through to Civil Protection response, resulting in complex work flows that must be able to combine different data sets, models and expertise in a flexible manner.

To develop a vision for how these trends can be managed and exploited in the future, we start by outlining in detail the concept of hydro-meteorological probabilistic forecasting chain.

4.1 HYDROMETEOROLOGICAL PROBABILISTIC FORECASTING CHAIN

Many attempts have been made in many countries to build up a sound hydro-meteorological probabilistic chain, notably in Southern Europe and the Western US. Major issues of uncertainty in the simulation of single physical processes, and the propagation of these uncertainties through the chain, have arisen. The scientific community and the Civil Protection Administrations have not, at least so far, constructed a sufficiently numerous sample of events for each region of interest. A major obstacle to progress is the need to combine in probabilistic form many different meteorological forecasts, either produced by the same

organization (ensembles) or by different organizations (poor man ensembles), which can then be used to produce ensembles of ground effects conditioned to the numerical weather predictions.

However, there is an increasingly pressing need, driven by societal and economical expectations, to predict possible impending floods, with a quantitative measure of uncertainty, so as to take the necessary steps of for mitigation. Some

THE UNCERTAINTY IN HYDRO-METEOROLOGY

The scheme on figure 3 represents the propagation of uncertainty of an ensemble of meteorological predictions. Each member of the meteorological ensemble (for simplicity, only four are shown here), is represented by a red line connecting the small yellow oval on the left to the large one immediately on the right, which contains the space-time conditions - white

dots- predicted by each ensemble member. From each of them, four different space-time precipitation histories - black lines - lead to different space-time distributions of saturation of the slopes of different watersheds, their space denoted by different shades of green-blue. From them, possible realizations of the riverbed flow originate, leading, further to the right, to different space-time

hydrographs where different peak flow value occur. From each of them, different realizations of floods produce, in different space-time conditions, different areas with different severity. The picture shows how the original uncertainty is propagated and increased through each successive physical process, leading to an ever-increasing number of possible outcomes.

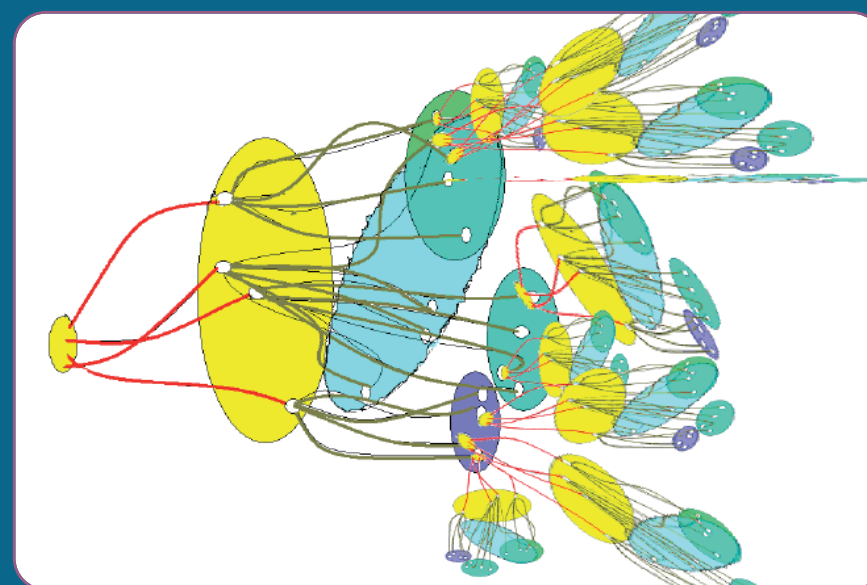


Figure 3: The scheme for the increasing uncertainty [1]

[1] Result of the interview with the expert Stefano Tibaldi - Arpa Emilia-Romagna.

research and operational institutions are attempting to build a predictive chain to be used daily for the purpose of alerting people of impending inundation and diffuse landslide processes.

Basically, the idea is that of carrying out one of the many numerical experiments of the above mentioned simulations of the chain of physical processes: the interested reader is referred for details to box: The uncertainty in Hydro-Meteorology. The end result of the chain is a prediction of a hydrological quantity such as river run-off and water level, but a large variety of models and data sources feed into the prediction. Figure 4 shows a schematic diagram of the forecast chain, working backwards from the forecast, in this case the water level in a certain section of the river or in a flooded area.

On a conceptual level a hydro-meteorological forecasting chain consists of 3 layers:

THE RAINFALL LAYER

The **rainfall layer** pertains to the combination of different (NWP) models to form a high-resolution multi-model ensemble (ranked first among HMR hot topics) together with stochastic downscaling algorithms (ranked fourth among HMR hot topics) to enable the production of quantitative rainfall predictions for severe meteorological events.

The structure of this layer is explained as follows:

- Numerical prediction of intense precipitation events, over complex orography regions, requires the use of high-resolution forecast models, and it is affected by uncertainty because of the inherently limited predictability of the rapidly developing weather phenomena that must be forecast and the sensitivity of the forecasts to initial conditions and the details of the forecast model;
- To deal with this uncertainty, there is active development of ensemble forecasting systems that can be used to make probabilistic precipitation forecasts;
- Ensemble forecasting systems: multiple numerical predictions are conducted using slightly different initial conditions that are all plausible given the past and current set of observations, or measurements. Sometimes the ensemble of forecasts may use different forecast models for different members corresponding to a multi-model ensemble approach;
- The multiple simulations are conducted to account for the two sources of uncertainty in weather forecast models:
 - 1) the errors introduced by chaos or sensitive dependence on the initial conditions;
 - 2) errors introduced because of imperfections in the model, such as the finite Grid spacing.

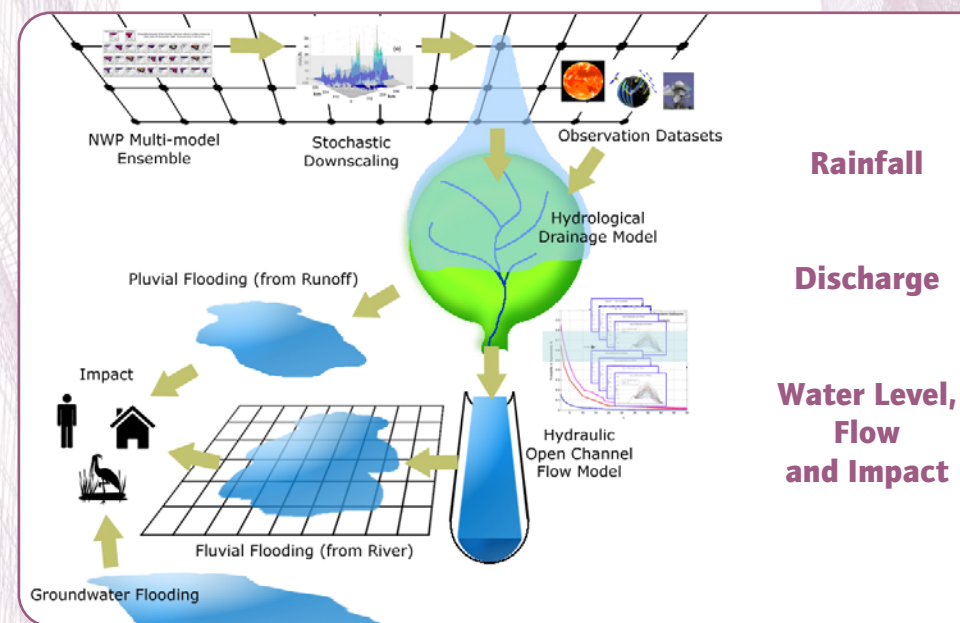


Figure 4: Graphics rationale of the hydro-meteorology probabilistic forecasting chain [2]

THE DISCHARGE LAYER

The **discharge layer**, instead, concerns the fusion/combination (ranked third among HMR hot topics) of rainfall predictions (from the **rainfall layer**) with corresponding observations, which are input into multiple hydrological models to enable the production of river discharge predictions.

The structure of this layer is explained as it follows:

- Prediction of river water levels and floodplain inundation following an extreme precipitation event requires integration of precipitation forecasts with a terrain-based hydrologic and hydraulic model. The reliability and

credibility of flood forecasting and warning systems cannot be guaranteed without properly incorporating the sources of uncertainty into the end-to-end forecasting and warning systems;

- The various sources of uncertainty associated with the model outputs can be classified as: model uncertainty, input uncertainty (due to imperfect forecasts of future precipitation, evaporation, slope runoff, etc.), and parameter uncertainty (due to imperfect assessments of model parameters);
- This problem becomes critical in small mountain catchments (or the

[2] Result of the interview with the expert Quillon Harpham - HR-Wallingford.

headwaters of bigger river systems) and urban areas, where flood prediction requires the forecast of the precipitation field down to scales of a few square kilometers and tens of minutes;

- In such situations, it becomes necessary to quantify how uncertainties introduced by the various steps in the forecasting chain propagate through the overall system, and how they affect the estimates of river runoff and peak discharge.

THE WATER LEVEL, FLOW AND IMPACT LAYER

Finally, the **water level, flow and impact layer** addresses the execution of hydraulic model compositions in different modes to assess the water levels, flow and impact created by the flood events and to compare them against observations through proper modeling verification metrics (ranked second among HMR hot topics). The structure of this layer is explained as it follows:

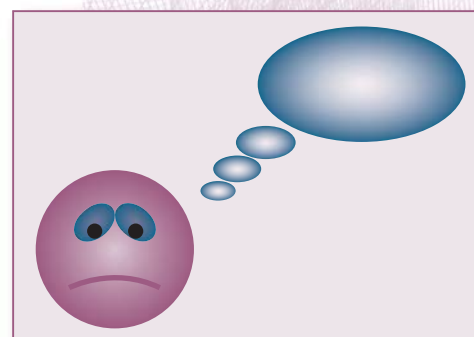
- Completing the flood forecasting chain that is initiated in the rainfall layer and continued in the discharge layer with the formulation of flooding risk scenarios that are relevant from early-warning and Civil Protection perspectives;
- Undertaking flood risk analysis for the purposes of long term strategic

planning, through investigation of the impacts of climate change on each of the case study sites.

4.2 AN IDEAL ENVIRONMENT FOR HYDRO-METEOROLOGY RESEARCH AT THE EUROPEAN LEVEL

The goal of hydro-meteorological research is to understand the physical processes underlying hydro-meteorology probabilistic forecasting chain, and to develop new and improved tools for the various layers. The daily work of HMR scientists falls generally into two categories: preparatory activities and science activities.

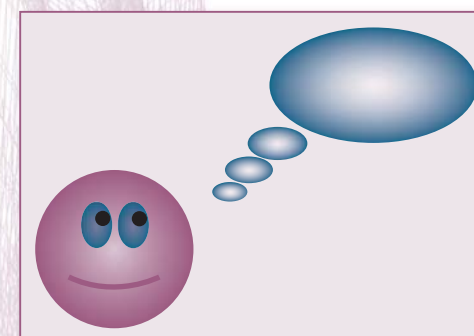
Preparatory activities include:



- find data;
- retrieve high volume data;
- learn formats and develop readers;
- extract parameters, identify quality and other flags/constraints;
- perform filtering/mask, develop analysis and visualization tools;

- select the in-house hydro-meteorological models (if available);
 - develop ad hoc post-processing tools.
- While these tasks are necessary, they are often time-consuming, tedious, and generally regarded as diversions from the scientists' primary work.

Science activities, in contrast, include:



- exploration
- initial analysis
- user the best data for the final analysis
- derive conclusions
- write the report/paper

These activities contribute in a direct way to the scientific objectives.

An ideal environment for research would take advantage of technology to minimize the first category of activities, and maximize the second.

Minimizing preparatory activities involves addressing the three challenges listed at the beginning of the section, data, models and workflows. Large data sets, whether from remote sensing instruments such

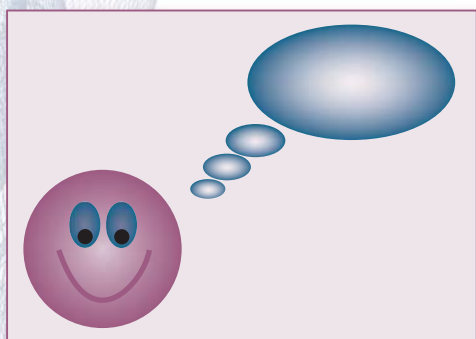
as radar networks, or from ensembles of numerical weather forecasts, need to be easily available. Availability implies easy to locate, easy to obtain the necessary permissions for use, and appropriate tools to read the different formats and meta-data associated with different data types.

Models, both meteorological and hydrological should also be available with a minimum of effort. As with data sets, substantial efforts are required to transport and use models at other institutions, and a better solution may be to provide transparent access to these tools at the facilities where they are maintained. Since some of the models are very computationally intensive, this also requires access to high performance computing to be managed alongside the tools.

Workflow becomes a problem with the increasing need to treat the hydro-meteorological forecast chain as a whole. Infrastructure is needed to allow output from alternative in tools each layer to be input into the next layer, and compare different configurations, without the need to continually develop new interface and analysis tools.

Maximizing science activities can be more than simply freeing up time. If infrastructure for handling large distributed data sets, a plethora of models and flexible workflows is

available, new possibilities for collaboration are opened up. Virtual communities involving HMR and related Earth Science disciplines can promote interdisciplinary research and open up the research world to a broader cross-section of society. Thus an ideal environment for HMR at the European level is expected to:



- Allow and improve access to HMR models, tools and data without the constraints of distance, access, usability and scientific barriers;
- Change and adopt the HMR discovery process in terms of computing, simulating, data processing and visualization;
- Enable a growing number of HMR (and related Earth Science) users to engage in open, cross-border and cross-discipline collaboration;
- Promote HMR cost-efficiency and related resource utilization;
- Attract non-specialist users such as members of the public services (Civil Protection agencies) who will be potential beneficiaries of faster and

better HMR results, and citizen-scientists who are interested in learning more about Hydro-Meteorology.

The hydro-meteorology probabilistic forecasting chain discussed in section 4.1 represents a natural framework to prove the full extent of HMR e-Science capability and potential. On the one hand, it addresses the interdisciplinary and international challenges of HMR in forecasting severe hydro-meteorological events over complex orography areas and assessing their impact. On the other hand, it also provides laboratories for the integration of new ICT infrastructure and development of new working practices in the forthcoming HMR e-Science environment.

4.3 DESIGN OF THE HMR E-SCIENCE ENVIRONMENT

The DRIHMS survey results and the corresponding reference context for hydro-meteorology research formulated in section 4.2 suggest a general rationale of the HMR e-Science environment as depicted in Figure 5, with resources (e.g., hardware, design and simulation methods and tools) provided by various institutions, services that are aiming at organizing the use of resources and applications, and application services accessible via portals.

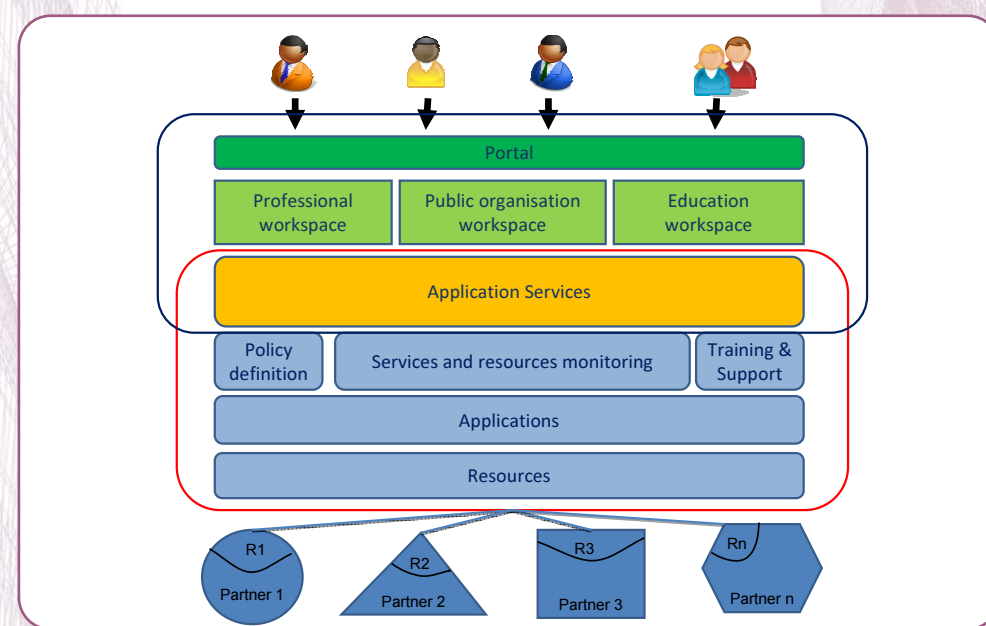


Figure 5: Conceptual view of the HMR e-Science environment [3]

The currently existing European e-Infrastructures are designed to handle the processing of a large amount of distributed data, exploiting the sharing of computing, storage and network resources in a coordinated way. The HMR applications necessary for running the hydro-meteorology probabilistic forecasting chain described in Figure 4, for example, would certainly benefit from the above design and the close interaction with the European EGI and PRACE projects. On the other hand, the HMR e-Science environment will also have specific HMR requirements not, or only partly, addressed within the

actual existing e-Infrastructures. Typical examples are the integration of sensor networks and data acquisition systems, the sharing of different modeling and post-processing tools, the data policy, and the provision of training/support in all related aspects

Analyzing the DRIHMS survey responses, the respondents require four application service families (see also Figure 6):

Simulation services: enabling users to run simulation tools, as meteorological models, hydrological and hydraulic models, stochastic downscaling and data fusion codes,

[3] Result of the interview with the expert Pierre-Henri Cros - CERFACS.

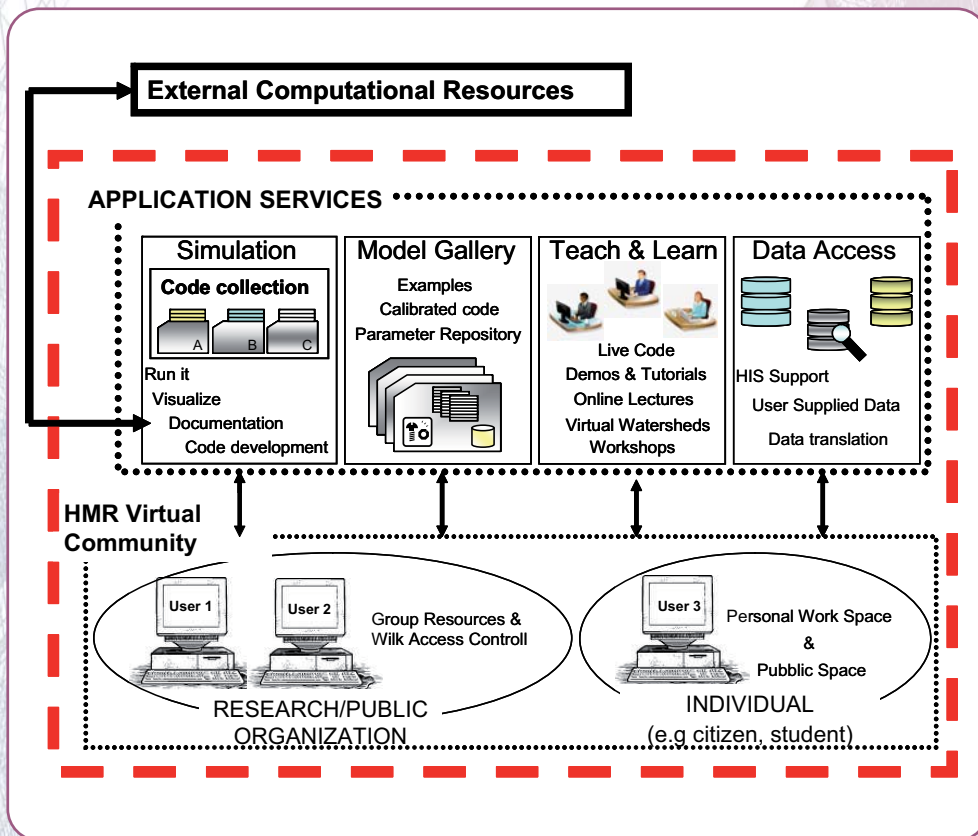


Figure 6 : HMR e-Science environment application services [4]

according to the rationale of the hydro-meteorological probabilistic forecasting chain;

Model Gallery services: allowing visualization of the results provided by HMR modelling and post-processing tools. Gallery services will enable the user to understand how to use and setup a particular model to solve a particular HMR problem: it will represent an example for others who were interested in solving similar problems;

Teaching and Learning services: giving access to a library of multimedia educational information. The simulation tools will be supported by tutorials, manuals and demonstrations, along with printed, video and audio media describing modeling techniques and principles. This material will include interactive demonstrations targeted at members of the public and other non-specialists.

[4] Result of the interview with the expert Richard Hooper [CUAHSI].

Material from workshops and conferences, and continuing education courses, will also be included;

Data Access services: including dedicated services, such as the CUAHSI Hydrologic Information System (CUAHSI HIS) for time series data and the Thematic Real-time Environmental Distributed Data Services (THREDDS) for accessing dynamic Grid data corresponding to severe hydro-meteorological events into an HMR e-Science environment data repository and needed for calibration, validation, or evaluation of the suitability of different modeling tools.

4.4 ARCHITECTURE OF THE HMR E-SCIENCE ENVIRONMENT OVER A GRID INFRASTRUCTURE

In the previous section a conceptual view of the HMR e-Science environment was outlined. As such, it does not describe the technical building blocks and how they interact over a Grid infrastructure. This is the purpose of the architectural view to be presented in this section. The HMR architecture is organized by the well-known layer pattern for implementing the inevitable separation of concerns. *The Resource Layer* (see Figure 7) contains all physical resources (computing

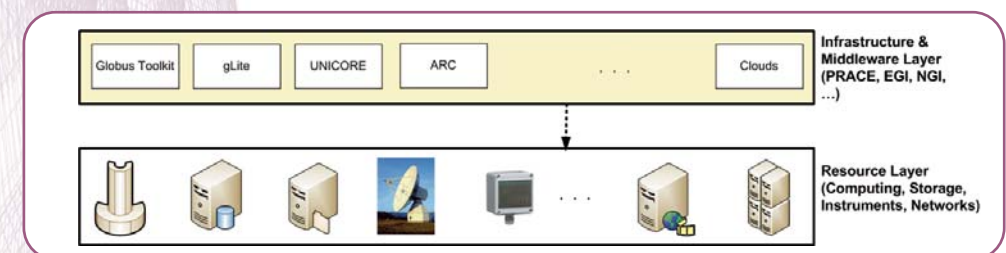


Figure 7: Resource and middleware layer

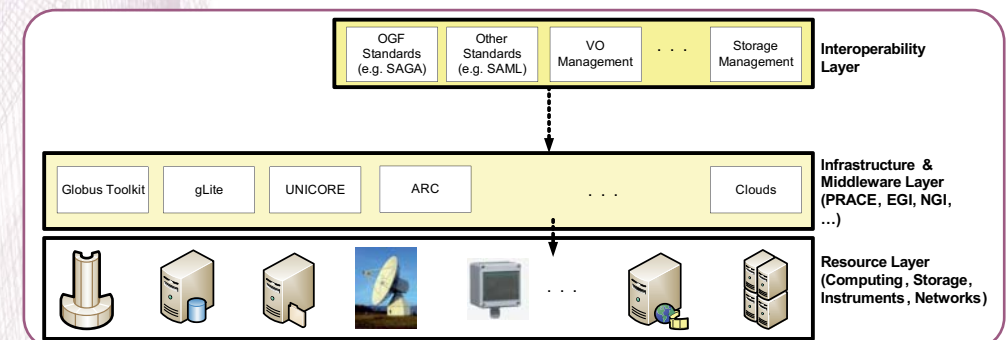


Figure 8: Interoperability layer

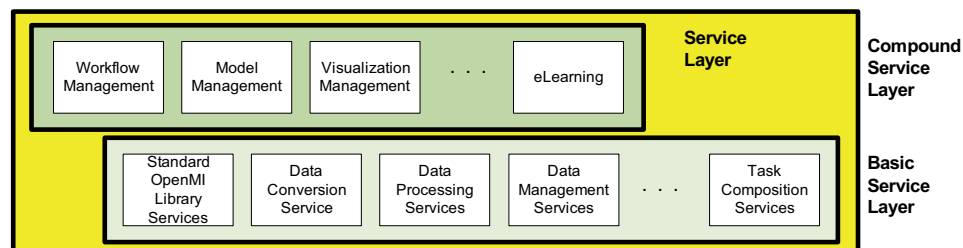


Figure 9: Service layer

elements, storage elements, instrument elements, sensor networks, communication networks) which are provided by the various partners in the HMR e-Science environment.

Typically, these resources are provided over European e-infrastructures like European e-infrastructures like Distributed European Infrastructure for Supercomputing Applications (DEISA) or the national NGIs, but operated locally by the providers. For accessing them a Grid middleware system like the (Globus, UNICORE - Uniform Interface to Computing Resources, gLite-Lightweight Middleware for Grid

Computing) is required. Some of the resources may also be available in a Cloud (public or private). The middleware and the access services are part of the **Infrastructure and Middleware Layer**.

In order to cope with the inherent heterogeneity of Grid resources and Grid middleware, an **Interoperability Layer** assists in transparently accessing the resources via the underlying layers (see Figure 8). The interoperability layer is based on standards like those defined in the Open Grid Forum (OGF, <http://www.ogf.org/>) or in the Organization for the Advancement of



Figure 10: Application layer

Structured Information Standards (OASIS, <http://www.oasis-open.org/>) like the Security Assertion Markup Language (SAML).

The HMR related services are comprised in the **Service Layer** which itself is separated in the **Basic Service Layer** and the **Compound Service Layer** (see Figure 9). The Service Layer typically interacts with the Interoperability Layer to transparently access the HMR resources. However, it may also interact with the Middleware Layer directly. The Basic Service Layer provides all services which are fundamental for the HMR community. Examples are data conversion, model access or task composition. As opposed to the Basic Service Layer, the Compound Service Layer provides more complex services assembled from basic services. Typical examples in the HMR context are the

creation and the management of workflows, the management of model sets, or the visualization of simulation results.

On top of the Service Layer the Application Layer contains the specific HMR applications for simulations, experiments and e-Learning (see Figure 10). The Application Layer uses (basic or compound) services from the Service Layer. It may, however, also interact directly with the Grid infrastructure. The Application Layer is accessed via the Access Layer which provides the required capabilities to authorize the access to the HMR applications via various portals or clients and through an authentication interface (see Figure 11).

Figure 12 depicts the complete layered architecture including the usage-relations between the layers.

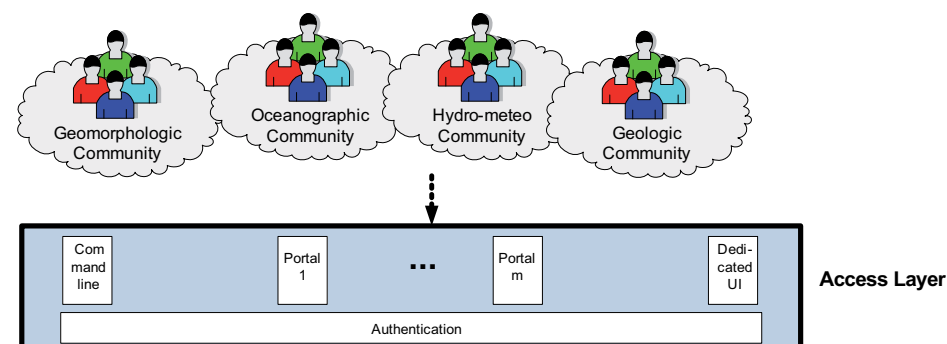


Figure 11: Access layer

5.0

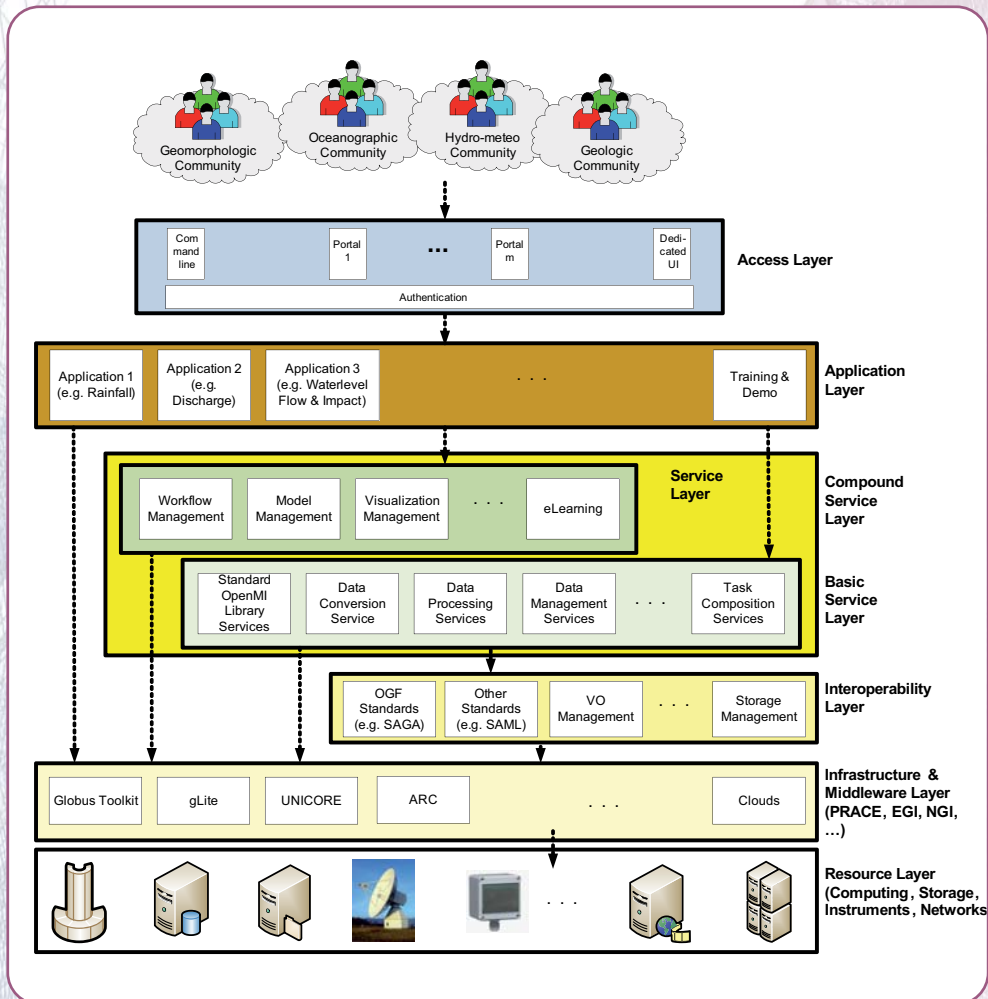


Figure 12: The complete architecture

A ROADMAP FOR HMR GRID SUPPORT

HYDRO-METEO AND ICT COMMUNITY:
TOGETHER TO BUILD A MODEL
WHICH SATISFIES THE NEEDS OF RESEARCHERS

The requirements related to the development and deployment of an HMR e-Science environment call for a well-defined set of activities to define a roadmap for HMR Grid support:

- **networking activities**, devoted to promote the cooperation between the HMR and ICT scientists involved in the development of the HMR e-Science environment, as well as between the scientific communities benefiting from the novel hydro-meteorology research infrastructures;
- **service activities**, devoted to provide specific ICT support related to the research infrastructure to the scientific community;
- **join research activities**, aiming at exploring new fundamental technologies or techniques underpinning the efficient and joint use of the HMR e-Science environment.

Among the **networking activities**, it is necessary to distinguish between the technical and administrative management of the HMR e-Science environment on the one hand and the dissemination and outreach of the solutions and results obtained by operating the HMR e-Science environment on the other hand. The latter is achieved through adequate information material, participation in conferences and workshops, media and press releases, support and training and

the organization of the HMR Virtual Research Community in cooperation with EGI.

Among the **service activities**, it is necessary to distinguish between infrastructure/data operation and management, devoted to the definition of the e-Infrastructure basic architecture, the provisioning of all necessary services and facilities, and the application services, devoted to model execution and integration, data integration, workflow and inventories of models, data sources and tools. For **joint research activities** it is necessary to distinguish between basic hydro-meteorology applications and HMR experiment suites to support representative experimental applications.

It is obvious that the provisioning of these services is not ad hoc. Rather, it is a long-term process with short-term and mid-term milestones. These milestones define what is generally called a roadmap. The roadmap described in the following aims at explaining the key steps necessary for efficiently supporting HMR from a Grid infrastructure perspective. The roadmap is adapted from similar efforts (like roadmaps for EGI, CoreGrid - European Research Network, eIRG - e-Infrastructure Reflection Group, and others). However, compared to these works, the focus here is on using

Grid infrastructures rather than on Grid deployments per se.

The starting point for the roadmap is implicitly given by the respective questionnaire results (see sections 3.3.1 and 3.3.2), whereas the direction of the roadmap follows the requirements analysis performed in section 4.

The macro objectives of the roadmap as presented in sections 4.3 and 4.4 are instrumental to the final overall objective, i.e. the provisioning of an HMR Grid infrastructure to the diverse stakeholders. These objectives are intended to be achieved in a stepwise manner, involving several sometimes transversal goals. For example, it is clear that all objectives require some kind of formal description of data. These are used for identifying data sources and formats owned and hosted by different organizations, for increasing the number of available models and for possibly changing forecasting data on-the-fly. In a similar manner, solving the data access issues for HMR applications, migrating model-aware applications to Grids and facilitating the access to non-local data for the citizen scientists are all mandatory to reach the overall objectives.

The HMR Grid support roadmap is arranged along some key categories, specifying their expected achievements and the time period (short-term, mid-term, long-term). In addition, we want

to point out some means that have been identified during the questionnaires as being of fundamental importance for user communities and for the technical aspects. For each time period, we also summarize the corresponding expected achievements of major communities.

SHORT-TERM OBJECTIVES

One of the major objectives to be addressed in the short-term is community building. To achieve this objective, several activities have been started or are being started. Examples are the work that has been undertaken in standardization bodies like the Open Grid Forum (OGF) (www.ogf.org) and the Outreach and Community Adoption Program which is part of the Open Geospatial Consortium (OGC) (<http://www.opengeospatial.org/ogc/programs/ocap>).

As community building has a strong governance component, for HMR Grid support to achieve efficiency of scale, many organizations need to participate (by granting access to data or other scientific equipment or by installing Grid nodes, Grid access middleware, tools, datasets, etc.). Without legislation and community specific directives this will probably not happen. For this reason we urge the key actors on the political and

funding levels to engage in a lasting dialogue aimed at addressing the key issues raised by this roadmap. In close relationship to community building activities, specialized Grid training programs are needed with a focus on HMR specific issues. Apart from building new or extending existing communities, HMR-related science communities need technical and operational benefits from Grid infrastructures and related HMR services. This includes (but is not limited to) single sign-on authentication to all data and services and to policy-based authorization of accessing restricted data (support for user roles, directory services, etc.). This includes the definition and implementation of a standardized approach to access required metadata. For a successful deployment of HMR services at large, the identification and discovery of data and models across communities is needed. A short-term goal is thus a standardized way not only to describe, to address and to discover such entities but also to describe, to address and to discover HMR related competence centers and respective specialized Service Providers. It should be noted that in parallel to this top-down approach, a demand-driven bottom-up approach is

necessary as well. Demonstrations of any kind (videos, live demonstrations, booths at fairs, etc) can create interest and momentum for a more thorough adoption of Grids which in turn may result in an elevated interest at other levels. Thus, this intertwined demand generation for a standardized “HMR Grid” needs to be backed up in the short- to medium-term by a number of initiatives and the development of respective components on both the Grid -and the HMR-side, by the creation of demonstrators, and by claiming compliance to standards. The table below summarizes the short-term objectives.

MID-TERM OBJECTIVES

The main concern of the mid-term period is solving the outstanding blocking issues for HMR applications deployed on a Grid, the porting of existing measurement and sensing applications from different disciplines, and facilitating the easy access for all stakeholders to HMR Grid services. The main mid-term objectives are summarized in the table below. Another issue that needs thorough mid-term investigation is the handling of data. Aspects that need to be examined are data retrieval, access to storage at near real-time and data transfer. Especially the latter is very

Short-term objectives	<ul style="list-style-type: none"> • to be achieved within the next 3 years
Expected Achievements	<ul style="list-style-type: none"> • expand the HMR Grid community in terms of number of organisations, number of projects, number of applications • reduce the technological gaps by providing a set of formal data descriptions, a set of interface descriptions and a set of basic services (see also Figure 9) • port HMR applications to use the Grid (initially mainly for disseminations purposes, but gradually moving to production-type applications) • fine-grained HMR Grid usage records can be used for accounting and billing • creation of synergies between various Grid-dependent communities and continuous promotion and dissemination activities
Means (examples)	<ul style="list-style-type: none"> • definition of a first version of the API for data access, integrated single-sign-on • definition of a first draft of metadata ontologies • definition of a first version of HMR usage records

Mid-term objectives	<ul style="list-style-type: none"> • to be achieved within the next 2 to 5 years
Expected Achievements	<ul style="list-style-type: none"> • Identification of and solutions to outstanding blocking issues for using Grids for HMR • including data access and quality of service • Porting of existing HMR applications (also from different research disciplines) to use identified standards • publication of scientific results which could not have been achieved without a standardized Grid infrastructure. • facilitate the easy access for all stakeholders to the HMR-augmented Grid infrastructure.
Means (examples)	<ul style="list-style-type: none"> • increase the number of HMR applications to use the Grid and let them benefit immediately for having access to sensors, data, archives, etc. • expose and publish the interfaces and standards • define an HMR-related profile to comply with for interoperability purposes • definition of an enhanced version of the API for data access, integrated single-sign-on enhanced definition of metadata ontologies • define dynamic authentication and authorization policies for accessing data and other services to assure Quality-of-Service levels • define methods for easy installation HMR-related Grid resources and services • finalize the specification of HMR usage records

important as on the one hand there is the necessity to transfer huge amounts of data and on the other hand there is a great variety of data formats and access protocols that need to be considered.

LONG-TERM OBJECTIVES

The long-term objectives are fundamental for a real exploitation of HMR-Grids. However, due to the time range involved, they can only be recommendations. Nonetheless, they cannot be achieved as long as critical mid-term objectives are still open. The main long-term objectives are summarized in the table below. The effort required for achieving these objectives is more about end user engagement than it is about the

deployment of any particular technology but nevertheless it should be kept in mind that the development of a standard-compliant middleware and its deployment is not an easy task and therefore will also require some work to bring the basic technologies into a useful state. Thus, special emphasis should be given to the ease of adoption, the ease of use, the hiding of complexity from the user, and the support of more complex scientific workflows for HMR.

Long-term objectives	<ul style="list-style-type: none"> • to be achieved within the next 4 to 10 years
Expected Achievements	<ul style="list-style-type: none"> • establish a standardized HMR Grid platform fully integrating the defined services as per chapter 4 and compliant with the profile to be defined mid-term • this platform facilitates an easy and seamless access to distributed and diverse data and services
Means (examples)	<ul style="list-style-type: none"> • definition of an HMR Grid platform (based on the architecture in Figure 12) • make services and data readily available across organizational boundaries via Grid services • HMR applications running on the Grid can seamlessly be used while integrating "foreign" data from other application domains • establish within 10 years a framework to provide timely data of interoperable Grid services for local, national, regional, and international policy makers • this requires a capacity building strategy for data and models that will significantly strengthen the capability of all countries

6.0

THE IMPACT ON HMR

WE HAVE A DREAM,
OVER DRIHMS

The DRIHMS survey results calls for a new way of doing HMR at European and possible global levels through the development and deployment of a dedicated e-Science environment. In particular, the following social and scientific facts promote this vision:

- On the social side, by the expectation of safety from the society and the responsibility of the local, state or national governance to provide that safety. All over the world, the social demand for safety under changing environmental conditions, including climatic perturbations and land-use changes, is growing. Recent failures of Civil Protection Agencies in most developed countries (see for example the Federal Emergency Management Agency (FEMA) re-organization after the New Orleans disaster in the USA or the 2007 fire season in the Central Mediterranean) put an increased pressure on scientists for more accurate and timely predictions, as well as on governments for effective emergency response systems to those predictions;
- On the scientific/research side, by the fact that present and future field campaigns and targeted modeling experiments will make available an unprecedented quantity of observational data and modeling tools that can enhance the study and prediction of severe hydro-

meteorological events. In the USA, the data and models are shared seamlessly across the nation, with still a lot of room for improvement. On the contrary, in Europe, it is amply clear by now that the fragmentation of data and model availability by countries is a serious limitation for further improving predictions and the response to extreme events, highlighting the need for a data and model sharing culture which will benefit all involved. This strong fragmentation poses a serious threat and a strong barrier to the fully exploitation at European and possible global levels of the different sets of instruments, data archives and computing facilities in order to efficiently perform HMR activities.

The forthcoming HMR e-Science environment is expected to support the scientific community by facilitating the development of new ideas and tools, coordinating HMR community efforts, disseminating research findings and products, and improving our ability to communicate with other scientists, students and the public, and also with public and private decision makers, all together contributing to the HMR virtual community.

The innovative solutions requested by the development and deployment of such e-Science environment will help the HMR and ICT communities to go

beyond the status quo by increasing the value of a huge body of electronic resources. These resources add extra value by supporting the production of innovative results and improving accessibility to the developed achievements.

In this respect, the progresses beyond the state of the art in HMR potentially enabled by the forthcoming e-Science environment can be summarized in:

- Horizontal comparison: Research institutions in different countries with their own systems and applications will be able to apply other models and exploit other data sets. At the same time they will be able to experiment with their own systems in different environments and conditions so that the limitations can be identified, understood and corrected;
- Ensemble approach: Grid technology offers a unique opportunity for sharing models, algorithms and data. It shortens the computing time needed for extended runs which is always a practical limitation. Dealing with uncertainty in HMR applications has been identified as a fundamental requirement for delivering usable forecasting products to the end users. HPC resources provide a perfect environment to perform extensive ensemble experiments;
- Multi-sensor extensive validation: the ready availability of different types of

data enables extensive validation of models. This will be facilitated by the Grid technology. The breakthrough would be the full and effective use of the voluminous high resolution data provided by remote sensing instruments, both satellite and ground-based (e.g., cloud microphysics retrieved by weather radars);

- Translational science: overcoming the barriers to the incorporation of research results into operational tools that make these results more useful to society.

The following major impacts of the forthcoming HMR e-Science environment are foreseen

- It will allow optimizing the performances, the use and access to an eco-system of HMR services which has the potential to enable a level of international research collaboration between HMR and related disciplines that has been impossible in the past. Distributed Research Infrastructure for Hydro-Meteorology (DRIHM) will also foster a strong and unprecedented collaboration and partnership between hydro-meteorological scientists and ICT scientists, thus serving cross-disciplinary needs;
- The activities will lead to a further exploitation of the ICT resources provided across Europe. Such exploitation is on the one hand

economically desirable, on the other hand it is also of technical importance as it provides more insight into highly loaded systems;

- It will enable collaborations with complementary worldwide e-Infrastructure, as Hydrological HUB (HydroHUB) and Earth System Modeling Framework (ESMF), in order to develop a comprehensive and reference model of hydro-meteorology virtual research community at the European and global levels. The DRIHM experience will foster the adoption of e-Infrastructures also by actors from HMR related disciplines and scientific communities, enhancing their global relevance and increasing the level of trust and confidence by the users;
- It may provide an important contribution to the development of new modeling, post-processing and analysis tools to be translated in the medium-term in operational facilities for extreme hydro-meteorological events risk prediction, prevention and mitigation;
- It will promote the interaction with other scientists in HMR and related disciplines through the HMR Virtual Research Community, enabling the collaboration with other Earth science communities (e.g. for the prediction of droughts and landslide hazards, as well as of atmosphere-ocean processes);

- One of the reasons that make it sustainable is the focusing on the better predictability of severe hydro-meteorological events. The most recent customers of the services provided by the meteorology and hydrology sector are the Civil Protection organizations. The cost of the applied research step that it is proposed is a minor fraction of the Civil Protection organizations budget, which is still increasing in spite of the present global critical period.

7.0

SUSTAINABILITY

HOW TO FINANCE A DREAM
SO THAT IT BECOMES REALITY

The development of DRIHMS action has recorded that there is the need for hydro and meteo scientists to rely upon modern ICT facilities in order to advance their specific fields of interest and to promote new multidisciplinary approaches. These advancements would represent the key factor to provide new, and more predictable forecast capacities and improve disasters prevention.

Overall these are the fundamental reasons that should make a future HMR e-infrastructure a sustainable one:

- Researchers need an environment that builds upon their previous innovations and that would permit cross contamination and the rising of new scientific knowledge;
- Civil Protection organizations have the expectation that this knowledge will improve, make more reliable and more affordable the hydro-meteorology services they use to reduce risk and prevent disaster.

On the other side DRIHMS has recorded the existence of available ICT technologies and strategies that represent a suitable support for Hydro Meteo Research. New developments and new tools will be necessary, but the basic elements are already here and it is possible to indicate them:

- Distributed Computing Infrastructures (DCI) such as DEISA/PRACE and the European

Grid Infrastructure (EGI) are available or are growing up. To support HMR will be a new challenge but also a new reason to exist for them, and HMR will find there the resources to grow up along new and promising directions.

- Open source software represents a powerful mean to provide improved computational models, and to increment their capabilities hence achieving new scientific knowledge. At the same time it will permit the development and deployment of customized and sustainable middleware for HMR infrastructure support.
- Established and emerging standards will represent the key to share data, but also to solve interoperability problems and combine together in a easy and effective way model components in new HM workflows.
- Co-operations and networking activities among different projects concerned with HMR and the use of advanced ICT capabilities will be another key factor to obtain sustainability.

These reasons and elements exist beyond DRIHMS, and DRIHMS has primarily registered them, highlighting a trend that will increase in the next few years. It is our expectation that a pan-European infrastructure for HMR research would

be set up in the near future. However we think that DRIHMS has represented a great opportunity to move the first steps in the good direction. With a small group of partners, a rewarding but small financial support, and in a short time frame it has been possible to collect a impressive set of data, to set up new networking relations between HM and ICT community, to involve top level scientists from both sides in face to face meeting and public conferences. Moreover it has been possible to clearly define what should be the next steps towards the implementation of a sustainable HMR e-infrastructure. We may mention also specific achievements such the on going effort to set up a HM Virtual Research Community within EGI. We are almost sure that DRIHMS itself is a small but essential step towards sustainability of this new growing community.

8.0

ACRONYMS & ABBREVIATIONS

ACRONYMS & ABBREVIATIONS

CERFACS - European Centre for Research and Advanced Training in Scientific Computation researches efficient algorithms for solving scientific problems

CLIVAR - CLimate VAriability and Predictability

COREGrid - European Research Network on Foundations, Software Infrastructures and Applications for large scale distributed, GRID and Peer-to-Peer Technologies

COST 731 - Concerted Research Action 731, is a European effort which deals with the quantification of forecast uncertainty in Hydro-Meteorological forecast systems

CUAHSI HIS - Consortium of Universities for the Advancement of Hydrologic Science Inc., Hydrologic Information System

DCI - Distributed Computing Infrastructures

DEISA - Distributed European Infrastructure for Supercomputing Applications

DRIHM - Distributed Research Infrastructure for Hydro-Meteorology

DRIHMS - Distributed Research Infrastructure for Hydro-Meteorology Study

EGEE - Enabling Grids for E-science

EGI - European Grid Initiative/ Infrastructure

ESMF - Earth System Modeling Framework

FEMA - Federal Emergency Management Agency USA

FloodProBE - Flood Protection of the Built Environment

FLOODsite - Integrated Flood Risk Analysis and Management Methodologies

GIS - Geographic Information System

gLite - Lightweight Middleware for Grid Computing

Globus - Research project developing a software infrastructure for distributed computing on a world-wide scale

Globus Toolkit - an open source software toolkit used for building Grids. It is being developed by the Globus Alliance and many others all over the world

GMES - Global Monitoring for Environment and Security

GPS - Global Positioning System

HM - Hydro-Meteorological

HMR - Hydro-Meteorological Research

HPC - High Performance Computing

HyMeX - HYdrological cycle in the Mediterranean eXperiment

ICT - Information and Communication Technology

IMPRINTS - IMproving Preparedness and Risk maNagemenT for flash floods and debris flow events

MAP - Mesoscale Alpine Programme

MAP-D-PHASE - Mesoscale Alpine Programme

MEDEX - MEDiterranean Experiment

NGIs - National Grid Initiatives

NWP - Numerical Weather Prediction

OASIS - Organization for the Advancement of Structured Information Standards

OGC - Open Geospatial Consortium

OGF - Open Grid Forum

PRACE - Partnership for Advanced Computing in Europe, is a unique

persistent pan-European Research Infrastructure for High Performance Computing

PREVIEW - PREvention Information and Early Warning

PRISM - Partnership for Research Infrastructures in earth System Modeling

SAGA - System for Automated Geoscientific Analyses, is a Geographic Information System Free Open Source Software. It has been designed for an easy and effective implementation of spatial algorithms

SAML - Security Assertion Markup Language

SEEGRID-SCI - South East Europe GRID e-Infrastructure for regional e-Science

THREDDS - Thematic Real-time Environmental Distributed Data Services

UNICORE - Uniform Interface to Computing Resources, offers a ready-to-run Grid system including client and server software. It makes distributed computing and data resources available in a seamless and secure way in intranets and the internet

VO - Virtual Organizations

WP - Work Packages



ANNEX

INSTITUTIONS CONTRIBUTING
TO THE QUESTIONNAIRES

INSTITUTIONS CONTRIBUTING TO THE QUESTIONNAIRES

- 1 Academia Sinica, Taiwan
- 2 Association for Computing Machinery (ACM), USA
- 3 Aston University, United Kingdom
- 4 Axceleon, Inc., USA
- 5 Boston University, USA
- 6 Brasil Space Agency, Brasil
- 7 Bundesanstalt für Wasserbau, Germany
- 8 Catalonia Meteorological Service, Spain
- 9 Cemagref, France
- 10 Center for Earth Observation and Digital Earth (Chinese Academy of Sciences), China
- 11 Centre for Marine and Atmospheric Sciences (ZMAW), Germany
- 12 Centre of Excellence in Information and Communication, Belgium
- 13 Centre of Supercomputing of Galicia, Spain
- 14 CIMA Research Foundation, Italy
- 15 CIMH (Caribbean Institute of Hydrology and Meteorology), Barbados
- 16 CRS4, Italy
- 17 Deltares, Netherlands
- 18 Deutsches Zentrum für Luft- und Raumfahrt, Germany
- 19 DHI (Danish Hydraulic Institute), Danmark
- 20 DMI (Danish Meteorological Institute), Denmark
- 21 DWD (German Meteorological Service), Germany
- 22 ECMWF (European Center Medium range Weather Forecast), United Kingdom
- 23 Emilia-Romagna Region Hydro-Meteorological Service, Italy
- 24 European Centre for Research and Advanced Training in Scientific Computation (CERFACS), France
- 25 Forschungszentrum Jülich, Germany
- 26 Fraunhofer-Gesellschaft, Germany
- 27 Friuli-Venezia-Giulia Region Hydro-Meteorological Service, Italy
- 28 HR-Wallingford, United Kingdom
- 29 IiPA, Italy

- 30 Indiana University, USA
- 31 Institut Pierre Simon Laplace (IPSL), France
- 32 Institute of Atmospheric Physics (Academy of Sciences of the Czech Republic), Czech Republic
- 33 Institute of Atmospheric Sciences and Climate (ISAC) of the Italian National Research Council (CNR), Italy
- 34 Institute of Cancer Research, United Kingdom
- 35 Institute of Meteorology and Water Management, Poland
- 36 Instituto Canario de Ciencias Marinas, Spain
- 37 Istanbul University, Turkey
- 38 JPL (Jet Propulsion Laboratory-NASA), USA
- 39 Korea Institute of Science and Technology Information (KISTI), Korea
- 40 Laboratoire Bordelais de Recherche en Informatique, France
- 41 Laboratoire d'Aerologie, France
- 42 Laboratory of Galaxy Physics, France
- 43 Leibniz Supercomputing Centre, Germany
- 44 Liguria Region Hydro-Meteorological Service, Italy
- 45 Ludwig-Maximilians-Universität München, Germany
- 46 Masaryk University, Brno, Czech Republic
- 47 Max-Planck-Gesellschaft, Germany
- 48 Meteo France, France
- 49 Meteoswiss, Switzerland
- 50 NACAD-UFRJ, Brasil
- 51 National Academy of Science, Ukraine
- 52 National Institute of Nuclear Physics and Particle Physics, France
- 53 National Laboratory of Computational Sciences, Brasil
- 54 National Meteorological Service of Bosnia-Herzegovina, Bosnia-Herzegovina
- 55 National Meteorological Service of Bulgaria, Bulgaria
- 56 National Observatory of Athens, Greece
- 57 National Research Council (CNR), Italy

- 58 Netherlands Royal Meteorological Institute (KNMI), The Netherlands
- 59 Norway National Meteorological Service, Norway
- 60 Piemonte Region Hydro-Meteorological Service, Italy
- 61 Polish Academy of Science, Poland
- 62 Politechnic University of Lyon, France
- 63 Politechnic University of Madrid, Spain
- 64 Politechnic University of Turin, Italy
- 65 Posnan Supercomputing Center, Poland
- 66 Princeton University, USA
- 67 RHMSS (Republic HydroMeteorological Service of Serbia), Serbia
- 68 Romania National Meteorological Service, Romania
- 69 Russia National Meteorological Service, Russia
- 70 Science and Technology Facilities Council (STFC), United Kingdom
- 71 Slovak Academy of Sciences, Slovakia
- 72 Spain National Institute of Meteorology, Spain
- 73 Swedish Meteorology and Hydrology Institute (SMHI), Sweden
- 74 Technical University of Dortmund, Germany
- 75 T-Systems Research, Germany
- 76 UCLA, USA
- 77 UK MetOffice, United Kingdom
- 78 Universidad de Extremadura, Spain
- 79 Universidade Federal de Campina Grande, Brasil
- 80 Universität Innsbruck, Austria
- 81 University College Dublin, Ireland
- 82 University Corporation for Atmospheric Research (UCAR), USA
- 83 University of Arizona, USA
- 84 University of Balear Islands, Spain
- 85 University of Barcelona, Spain
- 86 University of Belgrade, Serbia
- 87 University of Berlin, Germany
- 88 University of Bonn, Germany
- 89 University of Bruxelles, Belgium
- 90 University of Cardiff, United Kingdom
- 91 University of Cologne, Germany

- 92 University of Cyprus, Cyprus
- 93 University of Freiburg, Germany
- 94 University of Geneva, Switzerland
- 95 University of Göttingen, Germany
- 96 University of Graz, Austria
- 97 University of Grenoble, France
- 98 University of Hohenheim, Germany
- 99 University of Kobe, Japan
- 100 University of L'Aquila, Italy
- 101 University of Lisboa, Portugal
- 102 University of Madeira, Portugal
- 103 University of Madrid, Spain
- 104 University of Miami, USA
- 105 University of Minnesota, USA
- 106 University of Montreal, Canada
- 107 University of Murcia, Spain
- 108 University of Naples, Italy
- 109 University of Nevada, USA
- 110 University of New Delhi, India
- 111 University of New Zealand, New Zealand
- 112 University of Nice, France
- 113 University of Nicosia, Cyprus
- 114 University of Odessa, Ukraine
- 115 University of Perugia, Italy
- 116 University of Pisa, Italy
- 117 University of Reading, United Kingdom
- 118 University of Savoy, France
- 119 University of Science and Technology, Cracow, Poland
- 120 University of Split, Croatia
- 121 University of Tel Aviv, Israel
- 122 University of Timisoara, Romania
- 123 University of Tunisie, Tunisia
- 124 University of Tuscia, Italy
- 125 University of Utah, USA
- 126 University of Vienna, Austria
- 127 University of Warsaw, Poland
- 128 University of Zagreb, Croatia
- 129 Veneto Region Hydro-Meteorological Service, Italy
- 130 WMO (World Meteorological Organization), Switzerland



INDEX

INDEX

	page
1.0 EXECUTIVE SUMMARY	4.
2.0 INTRODUCTION Why DRIHMS?	7.
3.0 GAP ANALYSIS Hydrometeo research need a new e-infrastructure.	11.
4.0 SELECTED HMR GRID USAGE PATTERNS More accessibility, more sharing, more computing power: your dreams are our DRIHMS!	23.
5.0 A ROADMAP FOR HMR GRID SUPPORT Hydro-meteo and ict community: together to build a model which satisfies the needs of researchers	37.
6.0 THE IMPACT ON HMR We have a dream, over DRIHMS	42.
7.0 SUSTAINABILITY How to finance a dream so that it becomes reality	46.
8.0 ACRONYMS & ABBREVIATIONS	51.
ANNEX Institutions contributing to the questionnaires	55.

ACKNOWLEDGEMENTS

The DRIHMS Consortium is grateful to the Joint Research Centre - JRC - European Commission, International Advisory Board members (Prof. E. Foufoula-Georgiou, Prof. Evelyne Richard, Prof. Marios D. Dikaiakos and Prof. Thierry Priol) for enlighting discussions, useful comments and valuable contributions to the Project activities. Gratitude is also extended to all the international experts contributing to the DRIHMS survey activities.



DRIHMS is co-Funded by the EC
under the 7th Framework Programme



www.drihms.eu

THE DRIHMS TEAM:



CIMA Research Foundation
www.cimafoundation.org



Institute for Atmospheric Physics (DLR)
www.dlr.de



Ludwig-Maximilians-Universität (LMU) München
www.nm.ifi.lmu.de



Institute of Applied Mathematics and Information Technology (IMATI)
www.ge.imati.cnr.it