DRIHM

DISTRIBUTED RESEARCH INFRASTRUCTURE FOR HYDRO-METEOROLOGY

D7.1 Report on requirements

Abstract: This document provides an analysis of the requirements for all subsequent WP7 activities: it is based, first of all, on the requirements provided by baseline version of the experiment suites (WP8), and on the outcomes of the DRIHMS project. An extended version will be due at month 20.

DRIHM (G.A. n° 283568) is co-Funded by the EC under 7th Framework Programme

Document Information Page

Contract Number	283568
Project Name	Distributed Research Infrastructure for Hydro-Meteorology
Project Acronym	DRIHM
Deliverable Number	7.1
Deliverable Name	Report on requirements
Work Package Number	7
Work Package Name	Use Case and Application Requirements Analysis
Deadline	01/05/2012
Version	2.0
Dissemination Level	R
Nature	PU
Lead Beneficiary	CIMA

Table of Contents

1 Executive Summary

The goal of this deliverable is the collection of the requirements for all subsequent WP7 activities. The initial version of this report, due at PM8 takes advantage of all the information provided by baseline version of the experiment suites (WP8), of the outcomes of the DRIHMS project. The final version (due at month 20) will fully account for the new requirements that will progressively arise from the in deep understanding of the innovative scenarios associated with the three DRIHM experiments (WP8). A general-purpose methodology has been adapted in order to address a DRIHM specific need that is to combine modeling techniques used in different disciplines (Meteorology, Hydrology and Hydraulics), and provide them on a distributed infrastructure. Main steps have been:

- collect information about the existing practices and methods, collect shortcomings of such methods;
- describe high level user stories from which we derived a set of new Use Cases, related to the execution of baseline version of the experiment suites 1 and 2 on the DRIHM infrastructure;
- derive requirements from shortcomings of the existing methods and from the new Use cases.

The resulting set of requirements focus on:

- data management: ability to store and serve hydro-meteo data;
- data interoperability: each hydro-meteo model should be able to operate with at least one of the most common data format (supported formats). The infrastructure will provide interoperability among different supported formats;
- workflow management: the infrastructure will provide tools for combining hydro-meteo models and running workflows on HPC resources.

The document is organized as it follows. Section 2 is devoted to a short introduction providing a context for the undertaken analysis. Section 3 describes the methodology behind the requirements collection. Section 4 describes the different baseline experiments considered in the framework of the DRIHM project and the specific models involved. Section 6 summarizes the DRIHMS legacy, and presents the expected behavior of the system (described as user stories and use cases) and related requirements. Section 7 gathers conclusions.

2 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 global certified management tools to deal with processes leading to extreme hydrometeorological events [37].

From a research point of view dealing with the forecast of flood events need to address manifold issues that involve not only hydrometeorological scientist but requires the strong connection and collaboration with the Information and Communication Technology (ICT) community of science for providing new technological solutions ([30], [31], [32]).

In this context data and models accessibility remains a key issue in HMR: while some institution are big enough to be able to face these challenges with their own resources (e.g. ECMWF, NCEP, METEOFRANCE, etc), many of the academic and research institution of small and medium dimension cannot have an easy access both to HMR resources (data and models) and to ICT solutions.

For these reasons progress must aim at ensuring persistent availability and effective sharing of data and models across scientific disciplines, institutions and national boundaries, bringing together advances in HMR and ICT. Recent European efforts in developing a platform for e-science, including EGEE (Enabling Grids for E-sciencE, www.eu egee.org), SEEGRID-SCI (South East Europe GRID e-Infrastructure for regional e-Science, www.see-grid-sci.eu, [1]), and EGI (European Grid Initiative, www.egi.eu), provide an ideal basis for the sharing of complex hydrometeorological 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.

3 Methodology and Disclaimer

This chapter provides a brief overview of the methodology followed for collecting and describing the baseline and targeted DRIHM scenarios and Use Cases. This version of the document will focus on baseline harmonization, so most of the requirements will be described as improvement on the baseline.

3.1 Methodology

This requirements collection process follows the standard scenario based methodology of deriving requirements as for example described in [18]. The starting point is the current "as-is"-situation which has been reviewed in [23] and [24]. The target situation is described first by the gap list resulting from DRIHMS project (http://www.drihms.eu) polls and by scenarios of the DRIHM baseline experiments as per [24]-especially WP8-. We derive the requirements according to the following methodology:

- The scenarios of the DRIHM baseline experiments described in [24]-especially WP8-constitute the target situation DRIHM will be aiming at. We therefore describe the models relevant for WP8;
- Since DRIHM is built upon the DRIHMS project we also take into account the aggregated outcome of the DRIHMS polls conducted during 2010 and reported in [23];

It is worth mentioning that the requirements in this deliverable reflect the results of the DRIHMS questionnaires and the model landscape as per today. Consequently, there may be requirements reported in the literature or other projects that may contradict the requirements listed here or which may not be mentioned at all in this document. However, future versions of this document may take them up.

The process utilized for requirements elicitation and selection is the following:

- we collected description of the baseline: hydro and meteo models and their "Model Use Cases", enriched with a list of their problems/limitations

- then we described scenarios of the DRIHM baseline experiments as per [24] in form of user stories and then as use cases

Each partner has elicited a list of requirements. Such requirements have been elicited according to the standard requirements elicitation techniques.

During the Kick-off meeting, held in Savona, Italy on 8-9/9/2011, the DRIHM partners have discussed the general methodology for requirement elicitation and collection of information on the baseline.

During the "DRIHM Experiment Suites" meeting, held in Garching, Germany on 27/9/2011, the DRIHM partners have discussed the experiment suites to be used as a benchmark. In particular DRIHM partners defined two experiment suites for phase 1 of the project, and an additional experiment that will be considered during phase 2.

During the "DRIHM HMR activities" meeting, held in Munich, Germany on 16/12/2011, DRIHM partners decided a template for describing the baseline (models, attributes), use cases. Limitations and problems of the baseline have been collected by each partner.

During the "DRIHM Basic Experiment Suites and basic ICT services" meeting, held in Savona on 31/01/2012-1/02/2012, DRIHM partners finalized the discussion about the baseline version of experiment suites 1 and 2 to be demonstrated in WP8 and identified three "golden cases", represented by the events that interested Catalonia France and Italy between the mid October and mid November 2011. Such golden cases will be adopted as a benchmark for project phase 2.

3.2 Conventions

3.2.1 Use Case Glossary

In the next chapters of this document, the following short glossary of terms is used (UML-based):

Actor: something with behavior, such as a person (identified by role), computer system, or organization.

Scenario: a specific sequence of actions and interactions between actors and the system under discussion (also called a use case instance). It is one particular story of using a system, or one path through the use case.

Use Case: a collection of related success and failure scenarios that describe actors using a system to support a goal. In this document we have two types of Use Case: the Model Use Case aimed to describe aspects related with a single model (Section 4); Use Case that provides a more general view (Section 5).

3.2.2 Goals

The main objectives of the Use Cases are to supply the basis for the subsequent work of finding out, describing and recording functional requirements. This is pursued by describing the baseline, writing scenarios of use of the baseline, and identifying problems and limitations.

A Use Case helps to define a set of functional requirements of how the system needs to behave. It describes an interaction between an external actor and the system and documents the specific functions that the system will perform.

A complete and detailed definition of possible Use Cases guarantees a correct development with less effort in fixing functional bugs and also a trustable guideline for tests and validation of solutions.

3.2.3 Style

As a general approach, Use Cases should be written in an essential style keeping out collateral topics such as user interfaces and technical requirements and focusing on actors' intent. This means, they should ignore when possible "how" an interaction between the actors and the system is performed and concentrating on "what" they do, i.e., which valuable result they produce.

A Use Case has at least a name and a step-by-step description of a basic course of action, including: triggering events, necessary event response, pre-conditions and post-conditions, sequence of exchanged messages and performed actions, data exchanged and non-functional constraints (reliability, performance, cost etc).

The Use Cases are developed iteratively through three steps: (a) inception, (b) elaboration, and (c) construction.

During inception the Use Case form should be the simplest one, e.g. one of the following:

- Use Case name + brief description (one to five sentences describing what the Use Case does)
- Use Case name + outline (bulleted list of Use Case steps without alternative flows)

In this phase the end-users identify the main scenarios of use of the system and compare those with the technological opportunities suggested by the software engineers.

Elaboration is the phase where architectural relevance and risk factors of the Use Case will be stated, in this phase the Use Cases is are detailed in order to enable developers to build and test derived scenarios.

During construction the full behavior of the system is filled in by specifying the remaining Use Cases' flows.

We adopted the following extended template for Model Use Case:

Name	
Summary	Description of the Use Case
Rationale	
Actors	List of primary actors involved in Use Case
Experiment type	
Preconditions	The state of the system and values for pertinent attributes before the Use Case
Basic Course of Events	
Alternative Paths	
Postconditions	The state of the system and values for pertinent attributes after

	the Use Case , no matter which flows were executed
Models / Tools used	
Data accessed / produced	
Local / Remote	
Interfaces between models	
Problems with baseline	

.

4 Description of the DRIHM baseline version of the experiment suites

One of the main outputs of the DRIHM meeting that took place in Savona on 31/01/2012-1/02/2012, was a synthesis of baseline experiments and their classification based on the usage description, as depicted in the following table, where CIMA and RHMSS feature a full hydro-meteorological forecasting chain covering both baseline suite 1 and 2, while UPM is contributing to baseline suite 2, DLR as well as CNRS to suite 1.

Partner	Case study	Baseline Suite 1	Baseline Suite 2	HMR Hot Topics
CIMA	Genoa 2011	COSMO (Deterministic) Continuum (Deterministic) RainFARM (Probabilistic)	DRIFT (semi-distributed model)	Probabilistic forecasting (full chain) Precipitation downscaling Model verification metrics
RHMSS	Serbia 2010	WRF (Deterministic)	HBV (distributed model)	Probabilistic forecasting (full chain) Precipitation downscaling
RHMSS	Serbia 2010	WRF-NMM (Deterministic)	HYPROM (shallow water)	Probabilistic forecasting (full chain) Precipitation downscaling
UPM	Catalunia 2011	-	RIBS probabilistic-near real-time	Probabilistic forecasting (hydrology)
CRNS	France 2011	AROME (Probabilistic) MESO-NH (Probabilistic)	-	Probabilistic forecasting (meteorology)
DLR	Genoa 2011	Cb-TRAM Rad-TRAM	-	Data merging/fusion Model verification metrics

Table 1: DRIHM baseline experiments classification.

Each DRIHM partner provided a brief description of the models currently adopted (baseline), and Model Use Case For each Model Use Case we highlight the problems with respect to the baseline, and such problems will become functional requirements specific on the component described in the Model Use Case.

Models are grouped in two parts: those relevant to Experiment Suite 1 are described in Section4.1, while those relevant to Baseline Experiment Suite 2 are described in Section 4.2.

A detailed description of the baseline version of the experiment suite 1 will be available on Deliverable D8.1, while baseline version of the experiment suite 1 will be available on Deliverable D8.3 (both deliverables are due on month 12).

4.1 Baseline version of Experiment Suite 1

In this section we focus on the baseline version of the experiment suite 1. Each partner provided a brief description of the adopted model (or sequence of models) and the steps required to perform the simulation (data collection/conversion, preprocessing, simulation, post-processing…)

4.1.1 CIMA - COSMO model

The Consortium for Small-Scale Modeling (COSMO) model, is a fully compressible nonhydrostatic LAM. Created in 1998 by the Deutscher Wetterdienst (DWD; the German National Weather Service), it is currently used in research and operational modes by several meteorological services and universities in Europe.

The present operational application of the model within COSMO is mainly on the meso β-scale using a grid spacing of 7 km. The key requirement is an accurate numerical prediction of near-surface eather conditions, focusing on clouds, fog, frontal precipitation, and orographically and thermally forced local wind systems. Since April 2007, a meso γ-scale version has been running operationally at DWD by employing a grid spacing of 2.8 km. It is expected that this will allow for a direct simulation of severe-weather events triggered by deep moist convection, such as supercell thunderstorms, intense mesoscale convective complexes, prefrontal squall-line storms and heavy snowfall from winter-time mesocyclones.

The primitive hydro-thermodynamical equations describing compressible nonhydrostatic flows in a moist atmosphere are integrated, without any scale approximation, over an Arakawa C/Lorenz grid. The prognostic model variables are wind vector, temperature, pressure perturbation, specific humidity, and different precipitating/nonprecipitating water categories. The COSMO model, like most of the LAMs available in literature [e.g., the fifth-generation Pennsylvania State University–National Center for Atmospheric Research (NCAR) Mesoscale Model (MM5) and the Weather Research and Forecasting (WRF) model] does not exactly conserve moist entropy for numerical modeling efficiency reasons [35] The modeling of subgrid physical processes is supported by a wide range of turbulence, surface, and microphysical schemes. Different time-integration algorithms are also available, including a leapfrog horizontally explicit and vertically implicit (HEVI) time split integration scheme, a two time-level split-explicit scheme [14], a three time-level 3D semi-implicit scheme [23], and a two-time-level third-order Runge–Kutta scheme with various options for high-order spatial discretization [10]. For a more comprehensive and detailed description of the model, the reader is referred to [36]. COSMO is made of several programs mainly written in Fortran90 that perform the different tasks needed to carry out simulation. The workflow is typically managed through Unix shell scripts and user-customizable Fortran90 namelist files, i.e., ASCII files that contain assignment statements and are read when the programs are run. COSMO has been compiled and has run successfully on a wide range of architectures ranging from personal computers to high-performance computing systems. To simulate real cases, COSMO needs initial and boundary conditions that may be provided by different global circulation models and reanalysis datasets. Physiographic data (digital elevation model and land use) are also needed.

The COSMO-I7 NWP model is the Italian version, operated by ARPA-SMR, of COSMO-Model and it has been run at its operational horizontal resolution of 7 km: CIMA Foundation receives twice-a-day (00 and 12UTC runs) the modeling outputs from the ARPA-SIMC (Hydro-Meteo-Climate Regional Service of Emilia-Romagna), in the framework of the CIMA Foundation operational activities in cooperation with the Italian Civil Protection Department (ICPD)

Name	**COSMO model**
Summary	COSMO model QPF input
Rationale	The CIMA hydrometeorological chain needs input
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 1
Preconditions	A ftp server to receive and store the data
Basic Course of Events	COSMO QPF grib file is received COSMO QPF grib file is degribbed COSMO QPF data are prepared/re-interpolated to be ingested by the RainFARM model
Alternative Paths	None
Postconditions	COSMO QPF data are ready to be ingested by the RainFARM model

Models / Tools used	Grib libraries to decode Grib QPF file GRISO interpolator to re-interpolate data for the RainFARM model
Data accessed / produced	Accessed: COSMO QPF grib data Produced: data to be ingested by RainFARM (ascii format)
Local / Remote	Local: processing of the COSMO grib QPF data Remote: ftp download of the COSMO grib QPF data from an external server, where the model is run by a thirdy party (e.g. Emilia-Romagna regional meteorological service)
Interfaces between models	Interoperability with RainFARM and/or DRIFT, but not general one with other models Not standard interfacing
Problems with baseline	Using COSMO QPF grib file: no-control on modeling parameters Running COSMO (or other meteorological models): not easy access to HPC enabling fine-resolution simulations necessary for flash-flood producing storms typical of Mediterranean watersheds (e.g Genoa 2011 event) Running COSMO (or other meteorological models): not access to probabilistic versions of meteorological models necessary for flash-flood producing storms typical of Mediterranean watersheds (e.g Genoa 2011 event)

4.1.2 CIMA - RainFARM

RainFARM, Rainfall Filtered AutoRegressive Model [26], belongs to the family of algorithms called meta-gaussian models (see, e.g. [16]) and it is based on a nonlinear transformation of a linearly correlated process. This approach is closely related to the Turning Bands Method [21] and has been used both for satellite-based rainfall measurement validation and for stochastic rainfall modelling [3][20]. The model is able to generate small-scale rainfall fields that take into account not only the total amount of precipitation predicted by the meteorological model but also its linear correlation structure and the position of the main rainfall patterns. Due to the straightforward link between the model parameters and the large-scale field, this model is suitable for operational downscaling procedures. RainFARM uses the spectral information of large-scale meteorological predictions and generates fine resolution precipitation fields by propagating this information to smaller scales. The basic idea is to reconstruct the Fourier spectrum of the small-scale precipitation field by preserving the LAM information at the scales where we are confident in the meteorological prediction. The rainfall field is seen as the superposition of a finite number of harmonics with amplitudes decreasing as spatial and temporal scales become smaller. For a given realization (the predicted field at hand) the harmonics meteorological model, and should be preserved. However, when going down to smaller scales, they are less and less reliable and can be perturbed or replaced by harmonics whose properties respect the statistical properties at small scales of such rainfall fields. This gives a set of realizations, which displays a range of uncertainty that is linked to the effect of the scales poorly resolved by the model. RainFARM is written both in fortran and in matlab.

CIMA Foundation uses the RainFARM model in the framework of its operational activities in cooperation with ICPD and ARPAL (Hydro-Meteo Regional Service of Liguria region).

Name	**RainFARM model**
Summary	RainFARM stochastic dowscaling algorithm
Rationale	To downscale COSMO model QPF to the resolution needed for hydrological applications
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 1
Preconditions	COSMO model QPF data available (or the same data from another meteorological model)
Basic Course of Events	COSMO QPF grib file is received COSMO QPF grib file is degribbed COSMO QPF data are prepared/re-interpolated to be ingested by the RainFARM model RainFARM model is run to generate multiple rainfall maps at fine resolution
Alternative Paths	None
Postconditions	Multiple rainfall maps are ready to be ingested by the DRIFT hydrological model

Models / Tools used	RainFARM model (in matlab and/or fortran coding not parallel)
Data accessed / produced	Accessed: COSMO QPF reinterpolated data (ascii format) Produced: data to be ingested by DRIFT model (ascii format)
Local / Remote	Local: RainFARM is run locally
Interfaces between models	Interoperability with COSMO model and/or DriFt, but not general one with other models Not standard interfacing
Problems with baseline	Not general interoperability with other meteorological models and/or hydrologic models

4.1.3 RHMSS - WRF

The Weather Research and Forecasting (WRF) model is a numerical weather prediction (NWP) and atmospheric simulation system designed for both research and operational applications. WRF is supported as a common tool for the university/research and operational communities to promote closer ties between them and to address the needs of both. The development of WRF has been a multi-agency effort to build a next-generation mesoscale forecast model and data assimilation system to advance the understanding and prediction of mesoscale weather and accelerate the transfer of research advances into operations. The WRF effort has been a collaborative one among the National Center for Atmospheric Research's (NCAR) Mesoscale and Microscale Meteorology (MMM) Division, the National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Prediction (NCEP) and Earth System Research Laboratory (ESRL), the Department of Defense's Air Force Weather Agency (AFWA) and Naval Research Laboratory (NRL), the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma, and the Federal Aviation Administration (FAA), with the participation of university scientists. WRF reflects flexible, state-of-the-art, portable code that is efficient in computing environments ranging from massively-parallel supercomputers to laptops. Its modular, single-source code can be configured for both research and operational applications. Its spectrum of physics and dynamics options reflects the experience and input of the broad scientific community. The principal components of the WRF system are depicted in Figure 2. The WRF Software Framework (WSF) provides the infrastructure that accommodates the dynamics solvers, physics packages that interface with the solvers, programs for initialization, WRF-Var, and WRF-Chem. There are two dynamics solvers in the WSF: the Advanced Research WRF (ARW) solver developed primarily at NCAR, and the NMM (Nonhydrostatic Mesoscale Model) solver developed at NCEP. Community support for the former is provided by the MMM Division of NCAR and that for the latter is provided by the Developmental Testbed Center (DTC). As for COSMO, WRF is made of several programs mainly written in Fortran90 that perform the different tasks needed to carry out simulation. The workflow is typically managed through Unix shell scripts and user-customizable Fortran90 namelist files, i.e., ASCII files that contain assignment statements and are read when the programs are run. In addition to COSMO (and CNRS models see later on) however WRF has also a WRF Portal (http://wrfportal.org/): WRF Portal is the graphical user interface (GUI) front end for configuring and running both WRF cores (ARW and NMM), as well as configuring and running your own programs/scripts (like post). It simplifies the selection/localization of the domain, the running and monitoring of WRF, running ensembles, and the visualization of the model's output. Written in Java to be portable to all platforms, WRF Portal includes WRF Domain Wizard. Both WRF model and WRF Portal have been already ported on grid computing: Grid enabled the climate and weather simulation WRF service (http://appdb.egi.eu/#!p=L2FwcHMvZGV0YWlscz9pZD02ODU=) and WRF model (http://appdb.egi.eu/#!p=L2FwcHMvZGV0YWlscz9pZD0yODI=). WRF has been compiled and has run successfully on a wide range of architectures ranging from personal computers to high-performance computing systems. To simulate real cases, WRF needs initial and boundary conditions that may be provided by different global circulation models and reanalysis datasets. Physiographic data (digital elevation model and land use) are also needed.

WRF-NMM model is running operationally in Republic Hydrometeorological Service of Serbia from August 2007 in operational short-range weather forecast. Boundary conditions are taken from European Centre for Medium Range Weather Forecast (ECMWF) global model.

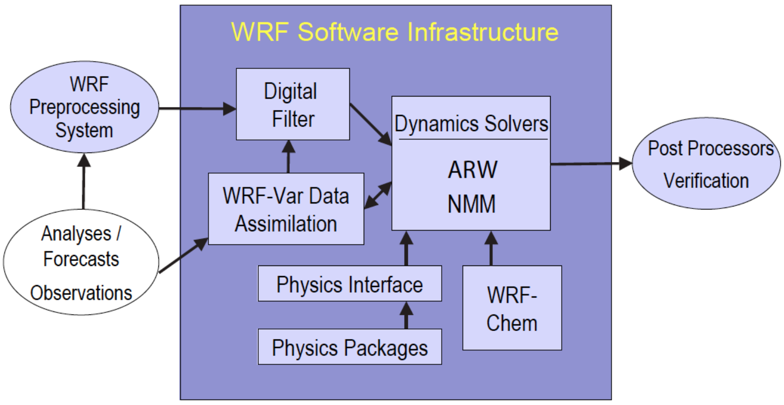

Figure 1: WRF system components.

Name	**WRF-NMM model**
Summary	WRF-NMM model quantitative precipitation and temperature forecast
Rationale	The RHMSS hydrometeorological chain needs input
Actors	Hydrometeorological operations and civil protection manager, Other end users
Experiment type	Experiment 1
Preconditions	A RMDCN connection with ECMWF to receive and store the data for WRF-NMM operational run or access to MARS archive for research.
Basic Course of Events	ECMWF grib files are received WRF-NMM model run with ECMWF's lateral and boundary conditions WRF-NMM NetCDF files are decoded WRF-NMM data are prepared to be ingested by the HBV model A ftp server to receive and store the data
Alternative Paths	DWD/GFS BLC data
Postconditions	WRF-NMM data are ready to be ingested by the HBV model

Models / Tools used	GRIB libraries to decode GRIB files WRF-NMM model NetCDF libraries to decode NetCDF files Fortran application to decode NetCDF files and calculate average QP and temperature forecast over river basin
Data accessed / produced	Accessed: ECMWF's GRIB files Produced: WRF-NMM NetCDF files data to be ingested by HVB model (ascii format)
Local / Remote	Local: running WRF-NMM model processing of the WRF-NMM NetCDF data Remote: RMDCN connection to download the ECMWF grib data
Interfaces between models	Interoperability with HBV Not standard interfacing
Problems with baseline	Not general interoperability with other hydrological models and/or stochastic downscaling models

4.1.4 CNRS - Meso-NH

Meso-NH [18] is a state-of-the-art nonhydrostatic mesoscale atmospheric model which has been jointly developed by the Laboratoire d'aérologie (UMR 5560 UPS/CNRS) and by CNRM-GAME (URA 1357 CNRS/Météo-France). The model is built around a dynamical core that is capable to simulate atmospheric motions ranging from the meso-alpha scales down to the micrometer scales; an ensemble of packages treating different physical processes in the atmosphere; a flexible file manager, and an ensemble of pre-processing tools to set up the initial conditions, either idealized or interpolated from meteorological analyses or forecasts. The basic prognostic variables are the three momentum components, the dry potential temperature, the turbulent kinetic energy, and additional scalars for multispecies and multi-phase flows. The vertical grid is defined with stretched levels [12]. Meso-NH may be run in two-way grid-nesting mode, where several smaller domains with different physics and horizontal resolutions may be nested in larger domains. Meso-NH comes with a wide variety of tools such as inline budgets, Lagrangian (backward) trajectories, and a full battery of advanced diagnostics. The ability of Meso-NH to simulate meteorological processes successfully and its efficiency in advancing the knowledge of atmospheric physics – in particular those involved in severe weather – has been demonstrated in many studies.

Meso-NH is made of several programs mainly written in Fortran90 that perform the different tasks needed to carry out a simulation or an ensemble of simulations. The workflow is typically managed through Unix shell scripts and user-customizable Fortran90 namelist files, i.e., ASCII files that contain assignment statements and are read when the programs are run. Meso-NH has been compiled and has run successfully on a wide range of architectures ranging from personal computers to high-performance computing systems such as BlueGene/P JUGENE, SGI Altix ICE, BULLX Nehalem, NEC SX8 & 9, etc.

To simulate real cases, Meso-NH needs initial and boundary that may be provided by different global circulation models and reanalysis datasets.. Physiographic data (digital elevation model and land use) are also needed. A series of dedicated Meso-NH tools is first used to prepare the simulation itself: preparation of initial conditions and coupling files, preparation of physiographic data and domain(s). Then, the main program carries out model time integrations and produces forecasts. Additional Meso-NH tools may be applied to Meso-NH forecasts to compute advanced diagnostics.

Name	**Meso-NH EPS**
Summary	Running Meso-NH EPS
Rationale	Provision of probabilistic QPF
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 1
Preconditions	Initial and boundary conditions in grib format (e.g., IFS, ARPEGE, ALADIN, or AROME files) Physiographic data (digital elevation model + land use in HDR+DIR formats) Storage and HPC resources
Basic Course of Events	Preparation of physiographic file (PGD) Preparation of the simulations (incl. retrieval, extraction and interpolation of initial and boundary conditions to be used during the simulations)

	Meso-NH runs Diagnostics
Alternative Paths	None
Postconditions	Meso-NH EPS files are available for further processing
Models / Tools used	Meso-NH and associated tools
Data accessed / produced	Accessed: Initial and boundary conditions in grib format (e.g., IFS, ARPEGE, ALADIN, or AROME files) Physiographic data (digital elevation model + land use in HDR+DIR formats) Produced: Meso-NH EPS output files (FM format, GRIB or NETCF also possible))
Local / Remote	Local: Configuration of experiment (Unix shell scripts to submit to remote resources) Remote: All the rest (computations on HPC + storage)
Interfaces between models	No generic interoperability with hydrological models No standard interface with hydrological models
Problems with baseline	No generic interoperability with hydrological models No generic tools to compare and combine with other EPSs

4.1.5 CNRS - AROME

AROME [29] is the latest operational numerical weather prediction model at Météo-France. It has been designed to improve short-range forecasts of severe events such as heavy precipitation in the Mediterranean regions (Cévenol episodes), severe storms, fog and urban heat islands during heat waves. AROME is a national and international collaborative achievement which includes the state of the art in atmospheric modelling. While its physical parameterisations are mostly inherited from the Meso-NH research model, its efficient dynamical core is derived from the ALADIN model but tailored to finer scales. Since December 2008, four daily forecasts over metropolitan France are delivered using the AROME model that thus helps improve weather forecasts for the current and the next day (since the forecast range is 30 h). As of today, AROME has a horizontal resolution of 2.5 km and 60 levels in the vertical defined on a mass-based hybrid pressure terrain-following coordinate. AROME uses twelve three-dimensional prognostic variables: the two components of the horizontal wind, temperature, specific content of water vapor, rain, snow, graupel, cloud droplets, and ice crystals, turbulent kinetic energy, and two nonhydrostatic variables that are related to pressure and vertical momentum. Hydrostatic surface pressure completes the set of prognostic variables. The lateral boundary conditions of AROME are provided by ARPEGE forecasts and coupling is typically done every hour. AROME takes its initial conditions from a three-dimensional variational (3DVar) data assimilation system derived from the ARPEGE-IFS assimilation system and tuned to the resolution of the model [5]. Besides data available for ALADIN and ARPEGE (including those from radio soundings, screen-level stations, wind profilers, GPS, buoys, ships, aircraft, and satellite data such as horizontal winds from atmospheric motion vectors (AMVs) and the QuickSCAT scatterometers, Advanced Microwave Sounding Unit (AMSU)-A and -B radiances from the NOAA-15, -16, -17, and the Aqua satellites, High-resolution Infrared Sounder (HIRS) radiances from NOAA-17, and clear-air Spinning Enhanced Visible and Infrared Imager (SEVIRI) radiances from the Meteosat-8 satellite), AROME uses fine-scale data such as reflectivity and Doppler radial velocity from the ARAMIS radar network. A prototype of ensemble prediction system based on AROME has been developed [38] and its products are available for research studies.

Name	**Baseline experiment 1 – AROME EPS**
Summary	Provision of AROME EPS QPF
Rationale	Provision of probabilistic QPF
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 1
Preconditions	A ftp server to receive and store the data
Basic Course of Events	AROME EPS native (FA) and GRIB files are received on ftp server
Alternative Paths	None
Postconditions	AROME EPS files are available for further processing
Models / Tools used	None
Data accessed / produced	Accessed: None Produced: FA and GRIB files (incl. QPF)
Local / Remote	Local: storing of AROME EPS files Remote: none
Interfaces between	No generic interoperability with hydrological models

models	No standard interface with hydrological models
Problems with baseline	No generic interoperability with hydrological models (technically speaking: needs extraction from AROME default format, interpolation, and conversion to hydrological model format) No generic tools to compare and combine with other EPSs

4.1.6 DLR - Cb-TRAM

Cb-TRAM is a fully automated tracking and nowcasting algorithm. Intense convective cells are detected, tracked and discriminated with respect to onset, rapid development, and mature phase. In addition, short-range forecasts are provided. The detection is based on Meteosat SEVIRI (Spinning Enhanced Visible and Infra-Red Imager) data from the broad band high resolution visible (HRV), infra-red 6.2 μm (water vapour), and the infra-red 10.8 and 12.0 μm channels. Areas of convection initiation, of rapid vertical development, and mature thunderstorm cells (cumulonimbus Cb) are identified. The tracking is based on geographical overlap between current detections and first guess patterns of cells detected in preceding time steps. The first guess patterns are obtained with the aid of an image matching algorithm providing complete fields of approximate differential cloud motion. Based on this so-called pyramid matcher also nowcasts of motion and development of detected areas are provided. Cb-TRAM is operated in real time and output is provided in the form of thunderstorm objects formatted in XML. The objects contain thunderstorm location (polygons), nowcast contours and some additional parameters as e.g. cell centre, cloud top temperature and trend. An example of application is presented in the figure below for thunderstorm detection and nowcasting over the Medioterranean. Cb-TRAM can be run both on observed and model generated (synthetic) satellite data. Cb-TRAM is part of the DLR Nowcasting systems operational in Germany.

Name	**Cb-TRAM**
Summary	Thunderstorm tracking, monitoring and nowcasting
Rationale	Use of observed and synthetic (model forecast generated) satellite data for thunderstorm detection and nowcasting in DRIHM case studies
Actors	Hydrometeorology research and civil protection manager

Experiment type	Experiment 1
Preconditions	A ftp server to receive and store the data
Basic Course of Events	Satellite data from DLR real time data bank received Cb-TRAM is run on LINUX-PC Detected and nowcast thunderstorm objects are provided on output
Alternative Paths	Synthetic satellite data from numerical model output (e.g. COSMO-DE) are received instead of real satellite data and Cb-TRAM is run with these synthetic data
Postconditions	Cb-TRAM output is ready for display/overlay on background geospatial data and for use in WxFUSION
Models / Tools used	Cb-TRAM
Data accessed / produced	Accessed: Meteosat8/9 satellite data; Data format HRIT; decoding with HRIT-decoder (EUMETSAT) Produced: thunderstorm object data (XML-files) to be ingested by WxFUSION or used later-on in DRIHM user display; Output available over distributed resources
Local / Remote	Local: processing of satellite data Remote: ftp download of satellite data from an external server. Upload of Cb-TRAM output XML files to DRIHM data handler
Interfaces between models	Interoperability with output from Met ensemble model forecasts as regards to synthetic satellite data (GRIB model output). Interoperability of Cb-TRAM XML-files for further use in user display and WxFUSION
Problems with baseline	Running Cb-TRAM: no access to output from meteorological ensemble models for generating forecast (synthetic) thunderstorm objects

4.1.7 DLR - Rad-TRAM

Rad-TRAM is quite similar to Cb-TRAM. It uses the same image processing tools for monitoring and nowcasting. However, in contrast to Cb-TRAM it operates with radar data (regional or composite) with the aim to detect areas of heavy precipitation. RAD-TRAM can be run both on observed and model generated (synthetic) satellite data. Rad-TRAM is part of the DLR Nowcasting systems operational in Germany.

Name	**Rad-TRAM**
Summary	Heavy precipitation tracking, monitoring and nowcasting
Rationale	Use of observed and synthetic (model forecast generated) radar data for heavy precipitation detection and nowcasting in DRIHM case studies
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 1
Preconditions	A ftp server to receive and store the data
Basic Course of Events	Radar data data from real time data bank received Rad-TRAM is run on LINUX-PC Detected and nowcast objects of heavy precipitation areas are provided on output
Alternative Paths	Synthetic radar data from model output are received and Rad-TRAM is run with these synthetic data
Postconditions	Rad-TRAM output is ready for display/overlay on background geospatial data and for use in WxFUSION
Models / Tools used	Rad-TRAM
Data accessed / produced	Accessed: Regional radar or radar composite data (DWD internal format) Produced: object data (XML-files) of heavy precipitation areas to be ingested by WxFUSION or used in DRIHM user display; output available over distributed resources
Local / Remote	Local: processing of radar data Remote: ftp download of radar data from an external server. Upload of Rad-TRAM output XML files to DRIHM data handler
Interfaces between models	Interoperability with output from Met ensemble model forecasts as regards to synthetic radar data (GRIB model output).

	Interoperability of Rad-TRAM XML-files for further use in user display and WxFUSION
Problems with baseline	Running Rad-TRAM: no access to output from meteorological ensemble models for generating forecast (synthetic) objects of heavy precipitation areas

4.1.8 DLR - WxFUSION

WxFUSION – Weather Forecast User-oriented System Including Object Nowcasting - is an integrated system which aims at combining data from real-time observations with nowcasting tools and numerical model forecasts accordingly in order to detect, nowcast (0-1 hrs) and forecast (1-24 hrs) target weather objects (TWO) which are pre-defined and specified by the user. These TWOs represent hazard areas, e.g. for air traffic like thunderstorms, or areas of heavy precipitation leading to flash floods. The former are detected from space by satellite while the latter are detected by radar. The system's central element for nowcasting TWOs is the fusion of the different data sources by using fuzzy logic, a method that deals with parameter ranges instead of fixed thresholds and allows to account for imprecise observations and forecasts. Based on conceptual models and expert knowledge mathematical functions are defined in order to characterize a particular weather hazard and to estimate its probability to occur in a defined region. For instance, the intensity of a thunderstorm is estimated by combining three measures: the cloud top temperature within its Cb top volume, the maximum reflectivity within its Cb bottom volume, and the lightning density. The use of numerical model forecasts for the time range one to several hours is accounted for by another part of WxFUSION called "forecast validation", where the quality of the forecasts is checked against observations. Thunderstorms and heavy precipitation areas analysed from synthetic satellite and radar images are compared to their real counterparts. In a user-defined window in space and time overlapping synthetic and real TWOs are searched for. If a synthetic object overlaps an observed object, specific attributes like size, moving speed, moving direction, trend and history are compared, and the quality of the forecasted TWO is assessed. This method enables the selection of the best forecast out of an ensemble, which can then be used for further predictions of the observed patterns beyond the nowcasting horizon. An example screen shot of WxFUSION graphical user interface is given in the figure below.

Name	**WxFUSION**
Summary	Comparison of observed and forecast thunderstorm or heavy precipitation areas for determining best member out of a meteorological ensemble model
Rationale	Linking met ensemble model with hydrological model in DRIHM forecasting chain
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 1
Preconditions	A ftp server to receive and store the data
Basic Course of Events	Cb-TRAM and Rad-TRAM output for real (observed) and model forecast (synthetic) data from local data bank received WxFUSION is run on LINUX-PC Info-file on selected members of met ensemble for both thunderstorm objects and heavy precipitation areas are provided on output
Alternative Paths	-
Postconditions	WxFUSION output is ready for use in hydrological forecast chain over distributed resources
Models / Tools used	WxFUSION, Cb-TRAM, Rad-TRAM
Data accessed / produced	Accessed: output from Cb-TRAM and Rad-TRAM for real and synthetic satellite and radar images Produced: info file: ensemble member information (identities, forecast times, additional parameters) which match best the observation; output from WxFUSION available over distributed resources
Local / Remote	Local: processing of Cb-TRAM and Rad-TRAM data Remote: ftp upload WxFUSION output (ASCII file) to DRIHM data handler
Interfaces between models	Interoperability with output from Cb-TRAM and Rad-TRAM XML-files
Problems with baseline	Only small (time-lagged) ensemble from only one model available

4.2 Baseline Experiment Suite 2

In this section we provide User Stories for the baseline version of the experiment suite 2. Each partner provided a brief description of the adopted model and the steps required to perform the simulation (data collection/conversion, preprocessing, simulation, post-processing…)

4.2.1 CIMA - DRiFt

DriFt, Discharge River Forecast [15][16][11], is a semi- distributed rainfall runoff model, based on a geomorphologic approach. It uses information and input (e.g. rainfall field, elevation and soil properties) distributed over the territory, while it is almost lumped in parameters and results.

A morphologic filter, based on contributing area and local slope, is used to identify hillslope and channel paths [28]. On the basis of this distinction, different typical velocities are assigned to each portion of the surface paths, either classified as hillslope or channel. These two velocity values are the kinematic parameters of the model. By summing the time that water spends respectively on the hillslopes and in the channel, the routing time of the i-th cell, τ_i, is obtained in the form:

$$\tau_i = \frac{L_{hi}}{v_h} + \frac{L_{ci}}{v_c}$$

where L_{hi} and L_{ci} are respectively hillslope and channel lengths of the i-th path, v_h is the overland flow velocity and v_c is the channeled flow velocity. These constant velocities maintain a physical implication as average values, on hillslopes and in channels. They are usually related, through the calibration phase, to a particular class of extreme events and therefore to a particular range of discharge [27]. In this way the hydrologic processes that take place on the different components of the system are coupled with basin morphology. This framework is directly linked to pluviometric and hydrometric records as the two kinematic parameters are calibrated.

In the last version of the model the infiltration process is modeled by using a modified Horton method [11] with two parameters: the final infiltration rate and the soil field capacity that needs to be calibrated, at a basin scale, using observed discharge.

The model is tied to run in an operational forecasting chain [33] and it uses as input quantitative precipitation forecasts in the form of spatial and temporal matrices produced by meteorological models. This general and flexible input data structure can be also produced from satellite or ground measurements. In this way, the spatial and temporal variability of rainfall patterns and the basin heterogeneity in morphologic, geologic and anthropic characteristics are considered. On the other hand, the model is lumped in parameters: these must be considered as mean quantities describing the catchment system and its dynamic at the basin scale. All the parameters have strong physical implications, allowing an easy and controllable calibration. From these characteristics the model is defined a semi-distributed model.

Name	**DriFt model**
Summary	DriFt semi-distributed hydrological model
Rationale	Designed for event-scale simulations and is based on a geomorphologic approach
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 2
Preconditions	COSMO model QPF data available (or the same data from another meteorological model) RainFARM model run It needs, as fundamental starting data, a digital elevation model by which the drainage network is individuated and every cell is classified as a channel or hill slope based on a filter that depends on drainage area and mean slope. Two different flow velocities are associated with channel cells and hill slope cells; as a consequence, the concentration time is defined for each point of the basin. The implemented infiltration scheme allows the modeling of “multipeak” events making it possible to simulate

	quite long periods (5–8 days) during which different events can occur.
Basic Course of Events	COSMO QPF grib file is received COSMO QPF grib file is degribbed COSMO QPF data are prepared/re-interpolated to be ingested by the RainFARM model RainFARM model is run to generate multiple rainfall maps ("story") at fine resolution DriFt is run for each RainFARM "story", each DriFt is independent
Alternative Paths	None
Postconditions	The discharge prediction in probabilistic format is available
Models / Tools used	DriFt model (fortran coding not parallel)
Data accessed / produced	Accessed: RainFARM multiple rainfall maps ("story") at fine resolution Produced: discharge prediction data (one discharge timeseries for each RainFARM input, ascii format)
Local / Remote	Local: DriFt is run locally
Interfaces between models	Interoperability with COSMO model and/or RainFARM, but not general one with other models Not standard interfacing
Problems with baseline	Not general interoperability with other meteorological models and/or stochastic downscaling models

4.2.2 CIMA - Continum

Continuum [34] is a continuous distributed hydrological model that strongly relies on a morphological approach, based on a novel way for the drainage network components identification [16]. The model has been conceived to be a compromise between models with a strong empirical connotation, which are easy to implement but far from reality, and complex physically based models which try to reproduce the hydrological processes with high detail but which introduce a hard parameterization and consequent uncertainty and lack of robust parameters identification. It is designed to be implemented in different contexts with a special focus on data scarce environments. All the main hydrological phenomena are modeled in a distributed way.

The basin is represented using a regular square mesh based on Digital Elevation Model (DEM), the flow directions are identified on the basis of the directions of maximum slope derived by the DEM. The drainage network is represented distinguishing between hillslope and channeled flow. Distinction between hillslopes and channels is made with a morphological filter [15]. This is able to describe hydrodynamic and morphological conditions in the channeled network [15]. Infiltration and subsurface flow are described using a semi-empirical, but quite detailed, methodology based on a modification of Horton algorithm [2][8] and focuses especially onto exploiting land use information and climatology to set the infiltration parameters (see for details, [11]). The energy balance is based on the so-called “force restore equation” [9] which balances forcing and restoring terms, with explicit soil surface temperature prognostic computation. The overland runoff is distributed with differentiation between hillslope and channel flow. Vegetation interception and water table flow have been also schematized.

The model shown a good ability in reproducing the observed discharge in different Italian basin and it also have very good performances in reproducing satellite observed Land Surface Temperature [34].

Name	**Continum model**
Summary	Continum continuous distributed hydrological model
Rationale	Designed for long-term simulations and is based on a geomorphologic approach
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 2
Preconditions	COSMO model QPF data available (or the same data from another meteorological model) Meteorological stations observations
Basic Course of Events	COSMO QPF grib file is received COSMO QPF grib file is degribbed COSMO QPF data are prepared/re-interpolated to be ingested by the Continum model

	Meteorological weather station data (temperature, wind, pressure) Continum is run for a certain period of time
Alternative Paths	None
Postconditions	The discharge prediction in is available
Models / Tools used	Continum model (fortran coding not parallel)
Data accessed / produced	Accessed: COSMO model prediction and/or observational data Produced: discharge prediction data (one discharge timeseries for each RainFARM input, ascii format)
Local / Remote	Local: DriFt is run locally
Interfaces between models	Interoperability not general one with other models Not standard interfacing
Problems with baseline	Not general interoperability with other meteorological models and/or observational datasets

4.2.3 RHMSS - HBV

The HBV rainfall – runoff model was originally developed at the Swedish Meteorological and Hydrological Institute (SMHI) in the early 1970s [4]. The model structure is comprised of three main components: a module for snow accumulation and melt, soils moisture accounting, and the response module. The model can be characterized as conceptual, in that the processes within and fluxes between each of the modules are governed by simplified expressions for the underlying physical processes. Processes are represented as linear or simple non-linear relationships. The model is adapted to different catchments by adjusting a number of model parameters. The physical description of the catchments is based on the catchment's area, the hypsographic curve and the land cover types. The physical processes are controlled by parameters selected during calibration. The model operates on readily available input data, i.e. time series (usually daily or hourly) of air temperature and precipitation. The climate input can be obtained from station observations or from interpolated observation or forecast grids. Evapotranspiration in the current application is estimated by HBV using average monthly values as model input.

Name	**HBV model**

Summary	Conceptual hydrological model
Rationale	Designed for continuous calculation of runoff
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 2
Preconditions	Constant fields WRF-NMM data are available
Basic Course of Events	Download all constant, initial and boundary condition fields WRF NMM preprocessing WRF NMM run WRF NMM postprocessing HBV preprocessing HBV run HBV postprocessing
Alternative Paths	None
Postconditions	HBV output is available
Models / Tools used	HBV model (Fortran, not parallel)
Data accessed / produced	Accessed: constant fields and WRF NMM model output Produced: HBV files (computed outflow and other parameters- soil moisture,evaporation,snow..,)
Local / Remote	HBV is run locally
Interfaces between models	Interoperability with WRF Not standard interfacing
Problems with baseline	Not general interoperability with other meteorological models and/or stochastic downscaling models

4.2.4 HYPROM

HYPROM is a hydrology surface-runoff prognostic model designed for real-time watershed prediction [22]. The model solves grid point-based shallow water equations with numerical approaches that include an efficient explicit time-differencing scheme for the gravity wave components and a physically based and numerically stable implicit scheme for the friction slope terms. The model dynamics (advection, diffusion, and height gradient force) are explicitly represented, whereas the model physics (e.g., friction slope) are parameterized, i.e., subgrid effects is expressed in terms of the model grid point variables. The fact that the modeling governing equations for momentum and mass are all prognostic makes HYPROM distinct to most other prognostic hydrology systems. The model uses real topography, river routing, and land cover data to represent surface influences. The HYPROM calculations can be executed offline (i.e., independent of a driving atmospheric model) or online as a callable routine of a driving atmospheric model. The model is applicable across a broad range of spatial scales ranging from local to regional and global scales. The model can be set up over different geographic domains and can run efficiently on conventional computer platforms. Finally, the model can be used either for hydrologic forecasts or climate studies if embedded as a component of an atmospheric climate model (Figure 2).

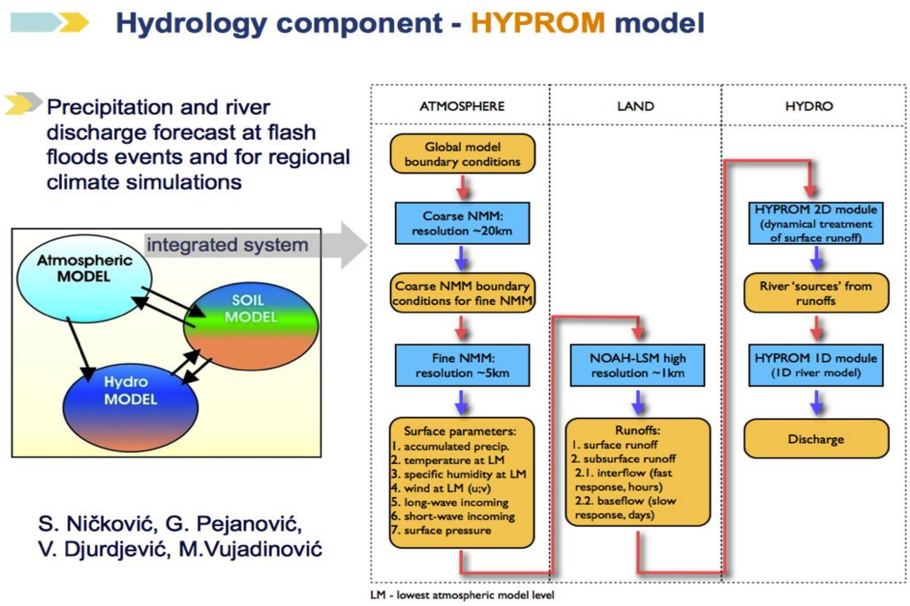

Figure 2: HYPROM system components.

Name	**HYPROM model**
Summary	HYPROM dynamical hydrological model
Rationale	Designed for simulations and forecasting overland hydrology
Actors	Hydrometeorology research and civil protection manager
Experiment type	Experiment 2
Preconditions	Constant fields and NMM output available
Basic Course of Events	Download all constant, initial and boundary condition fields

	NMM preprocessing NMM run NMM postprocessing HYPROM preprocessing HYPROM run HYPROM postprocessing
Alternative Paths	None
Postconditions	HYPROM output is available (binary files)
Models / Tools used	HYPROM model (Fortran, not parallel)
Data accessed / produced	Accessed: constant fields and NMM model outputProduced: HYPROM binary files (surface runoff, subsurface runoff, river discharge, other surface parameters)
Local / Remote	Local: HYPROM is run locally
Interfaces between models	Interperability with NMM model but not general one with other models Not standard interfacing
Problems with baseline	Not general interoperability with other meteorological models and/or stochastic downscaling models

4.2.5 UPM - RIBS

The Real-time Interactive Basin Simulator (RIBS) model is a distributed hydrological rainfall–runoff model that simulates the basin response to an event of spatially distributed rainfall. This model was designed for real-time application in medium-size basins. The model follows the structure of the grid of a digital terrain model in a matrix form. The data are stored in layers of raster-type information, which are combined to obtain the model parameters [13].

The RIBS model consists primarily of two modules. The first is a runoff-generation module that gives the spatial distribution of the infiltration capacity of the basin and represents the evolution of the saturated surface to the size of each cell. The saturation takes place mainly in two areas. First, we describe the area where rainfall intensity exceeds the infiltration capacity of the land.

The RIBS model is completed with a second runoff propagation module, with the objective of simulating the runoff routing process through the basin. We consider that a hillside path and a river-bed path form the length of runoff travel. Therefore, in order to obtain the total travel time, we must define travel velocities on the hillslope (v_h) and on the riverbed (v_s).

Name	RIBS model
Summary	RIBS distributed hydrological model
Rationale	Designed for event-based simulation and based on detailed topographical information
Actors	Hydrology research and practitioners
Experiment type	Experiment 2
Preconditions	Observed rainfall data available (either rain gauge observations of radar quantitative precipitation estimation) Observed discharge data available (one or several gauges in the basin)
Basic Course of Events	Model initialization Cyclic operation loop every ☐T Obtain rain fall and discharge observations in ☐T Read previous basin state a time t-☐T Simulate N replicas with current parameter distribution from t-☐T to t Update estimation of model parameters Simulate N replicas with updated parameter distribution from t-☐T to t+Tforecast Save basin state a time t Model end
Alternative Paths	None
Postconditions	The discharge prediction in probabilistic format is available
Models / Tools used	RIBS model (Windows executable version) RIBS model (Linux C++ code likely available)
Data accessed / produced	Accessed: Time series of rainfall and discharge measurements in gauges

	Sequence of radar raster maps Produced: Discharge data (one discharge timeseries per streamflow gauge, per ensemble member and per time step) Basin states (three raster maps of state variables per ensemble member and per time step)
Local / Remote	Local: RIBS is run locally
Interfaces between models	Not standard interfacing
Problems with baseline	Not general interoperability with other meteorological models and/or stochastic downscaling models

5 Databases for experiment suites 1 and 2

A relevant number of data are necessary to calibrate, run and validate the different models involved in experiment suites 1 and 2. A brief overview of the main hydro-meteorology databases of interest for DRIHM is provided:

- The Meteorological Archival and Retrieval System (MARS) is the main repository of meteorological data at European Centre for Medium-Range Weather Forecasts (ECMWF). It contains terabytes of operational and research data as well as data from special projects. MARS data is freely available to registered users in the Member States and Co-operating States. There is no public access to MARS. For research and commercial use, data can be obtained through MARS Data Services. For research use only, some datasets are freely available on MARS Data Server;
- The NOMADS web site provides access to the model data repository at National Climatic Data Centre (NCDC). This repository contains both archived (i.e., off-line) and on-line model data. Climate model data are provided in many way of access. It is possible to download the on-line data using traditional access methods (web-based or ftp), or using open and distributed access methods promoted under the collaborative approach called the NOMADS;
- HyMeX (HYdrological cycle in the Mediterranean EXperiment) is an international programme that aims at a better understanding and quantification of the hydrological cycle and related processes in the Mediterranean. It involves a large research community from various disciplines and methodological backgrounds. The HyMeX Database offers users full public access to its metadata catalogue through the "Browse catalogue" section. However, access to data is restricted to HyMeX registered users, as described in the HyMeX data policy section.

In a subsequent version of this document, the possible interplay with EuroGEOSS, which is a large scale integrated project in the Seventh Framework programme of the European Commission. It is part of the thematic area: "ENV.2008.4.1.1.1: European environment Earth observation system supporting INSPIRE and compatible with GEOSS". EuroGEOSS demonstrates the added value to the scientific community and society of making existing systems and applications interoperable and used within the GEOSS and INSPIRE frameworks.

5.1 MARS overview

The Catalogue web pages that users browse are dynamically built, from the content of the MARS metadata databases. This catalogue represents the exact state of the MARS archive. Every one of the billions of meteorological fields can be reached and retrieved from there whilst support for observations is not yet available.

Data Finder

Because the catalogue is organized according to the content of the MARS metadata databases, its structure may not be suitable for quickly find information. The Data Finder tries to solve this problem by fully indexing the operational part of the archive and allow users navigate this index in various ways.The Data Finder database is, on the average, updated two days after the archive. Every night, the output of all ECMWF operational models is archived as soon as it is created. At the end of the archive process, a set of web pages is automatically created that describe what meteorological fields have been added to MARS.

Parameter database

MARS content is based on GRIB tables, e.g. on one page users can see which meteorological parameters in GRIB table 128 are archived in MARS and sort them by their GRIB code or MARS abbreviation. By following the links users can also find further information on the computation of individual parameters in MARS model.

Server Activity

The MARS system serves about tens of thousands of requests daily. The duration of a retrieval depends on the number of concurrent requests the system is currently processing, the amount of data retrieved as well as the number of tape mounts involved. Users can monitor the progress of usersr requests from Server Activity page.

MARS user guide

The MARS user guide contains information about the operational archive, the description of the MARS language to retrieve data in batch, some useful hints to optimise user retrievals and information about the architecture of the MARS system.

Migration to GRIB API

The MARS client using GRIB API as its data decoder instead of GRIBEX has been made the default on all major platforms on 10 May 2011. To retrieve data in GRIB 2 format the use of this version is essential. For a list of known differences between the GRIBEX and the GRIB API version of MARS please refer to a separate page giving more details.

5.1.1 Datasets of special interest for DRIHM

Atmospheric deterministic models

Analysis: global analyses for the four main synoptic hours 00, 06, 12 and 18 UTC. They are the best gridded estimate of the state of the atmosphere (best fit to observations). For each of the synoptic hours, data is produced at the following levels:

Forecast: global 10-day forecasts based on the 00/12 UTC Analysis (the 00 UTC run started on March 2001 as an experimental suite for severe weather prediction). Forecast products are classified in the same level types as Analysis data: Surface, Model levels, Pressure levels, and Isentropic levels. Meteorological parameters are written output for every forecast time step, 3-hourly intervals from 00 to 72 hours, and 6-hourly from 72 to 240 hours.

First guess: a forecast with base time from the previous synoptic hour and a forecast time step of (usually) 6 hours. Note that since the change to 12 hour-cycling 4d-Var in the year 2000, this type of data has been discontinued.

Initialised Analysis: the best 'balanced' gridded estimate of the state of the atmosphere (initial state of the forecast). With the implementation of the three dimensional variational Analysis (3d-Var) on January 1996, the Initialised Analysis is no longer produced, and the Analysis is the best and the best 'balanced' gridded estimate of the atmosphere.

Four-dimensional variational Analysis: analyses using observations over a time window (e.g. a 6 hour 4d-Var at cycle hh will contain observations from hh-02:59 to hh+03:00).

Errors in Analysis: the assumed uncertainty of an observation is combined with the assumed uncertainty of the First Guess, resulting in an estimate of the total uncertainty in the Analysis.

Errors in First Guess: result of the uncertainty of the First Guess compared with observations

Four-dimensional variational Analysis increments: the low resolution increment which is added to the first-guess after each inner loop minimisation.

Ensemble Prediction System

ECMWF's Ensemble Prediction System has a coupled atmospheric and wave model. On 28 November 2006 the Variable Resolution Ensemble Prediction System (VarEPS) has been introduced by extending the forecast range from 10 to 15 days with a resolution of T399L62 for day 1 to day 10 (Leg 1) and T255L62 for day 11 to day 15 (Leg 2). On 11 March 2008 the Monthly Forecasting System, running once a week, has been integrated with the VarEPS. The new monthly forecast products were produced for the first time on 13 March 2008.

Control forecast: an unperturbed forecast at a lower resolution than the main 10-day deterministic forecast. Forecast runs to 15 days, with lower resolution from truncation step 240 onwards. Data is available on Surface, Model levels, Pressure levels and Isentropic levels.

Calibration/Validation forecast: VarEPS includes two constant-resolution forecasts for calibration and validation purposes which run for both resolutions from day 1 - 15. Data is available on Surface, Model levels, Pressure levels and Isentropic levels.

Perturbed forecasts: different forecasts to 10 days with perturbed initial conditions. They are numbered from 1 to N depending on the EPS setup. Data is available on the Surface and on Pressure levels and Isentropic levels.

Initial condition perturbations: the initial conditions for the EPS are designed to represent the uncertainties inherent in the operational analysis. They are created by adding perturbations to the operational analysis which produce the fastest energy growth during the first two days of the forecast period, defined using the singular vector technique.

Forecast probabilities: a statistical distribution of the weather parameters from all ensemble members is used to produce probabilistic weather forecasts. With the introduction of VarEPS this data type has been discontinued.

Event probabilities: provide the probabilities of the occurrence of weather events at each grid point. The probabilities are calculated on the basis that each ensemble member is equally likely.

Ensemble means: are means of the ensemble forecast members.

Clusters: similar ensemble members are grouped together into clusters. The mean and standard deviation of these clusters are computed (as well as the mean and standard deviation of the overall ensemble). Five sets of clusters are computed, one for the entire European area, and four for smaller areas.

Tubes: another clustering method which averages all ensemble members which are close to the ensemble mean and excludes members which are significantly different.

Extreme Forecast Index: (EFI) measures the difference between the probability distribution from the EPS and the model climate distribution.

ECMWF Re-Analysis

The ECMWF Re-Analysis (ERA) project has so far produced three datasets:

- ERA-15, a validated 15 year data set of assimilated data for the period 1979 to 1994
- ERA-40, a validated 45 year data set of assimilated data for the period 1957 to 2002
- ERA-Interim, a validated set of assimilated data starting from 1989, to be extended in near real-time

These available datasets include: analysis, forecast and forecast accumulations as output from atmospheric model, as well as analysis and forecast from a wave model re-analysis. There is also a Monthly Means data set containing data at the resolution of the data assimilation and forecast system used by each Re-Analysis.

TIGGE The Interactive Grand Global Ensemble, is a key component of THORPEX: a World Weather Research Programme to accelerate the improvements in the accuracy of 1-day to 2 week high-impact weather forecasts for the benefit of humanity. Global ensemble forecasts to around 14 days generated routinely at different centres around the world. Currently data from ECMWF, JMA (Japan), Met Office (UK), CMA (China), NCEP (USA), MSC (Canada), Météo-France, BOM (Australia), CPTEC (Brazil) and KMA (Korea) is archived.

5.1.2 ECMWF Data policy

The Licensee is authorised to use on a non-exclusive basis the Archive Products for its own purposes, including research, education and activities for commercial gain.

The Licensee shall not reproduce, distribute, license, transfer, assign, sell, disclose or otherwise forward the Archive Products, whether in their original form or as part of a service where the original numerical values of the Archive Products can be accurately rebuilt from the service, to any third party, without the prior written consent of the Licensor.

The Licensee shall not use, store or deal with the Archive Products in any way which contradicts the authorised use indicated in Condition 1 and/or the exclusions indicated in Condition 2 without the Licensor's prior written consent. The Licensee shall notify the Licensor if it becomes aware of any unauthorised use of the Archive Products.

All Intellectual Property Rights of the Archive Products owned by ECMWF or its licensors shall remain the property of ECMWF or its licensors and the Licensee acknowledges the full title and ownership by ECMWF or its licensors of all the Archive Products supplied.

The Licensee agrees to acknowledge the origin of the Archive Products in any scientific publication arising out of the use of the Archive Products. It will provide the Licensor a copy of such publication, upon request.

The Licensor makes no warranty as to the accuracy or completeness of the Archive Products. All Archive Products are provided on an "as is" basis. Any warranty implied by statute or otherwise is hereby excluded from this agreement to the fullest extent permissible by law.

Neither party shall be liable to the other under any cause of action for any loss of profit or loss of revenue (whether direct or indirect); any indirect loss; loss of goodwill, reputation or opportunity; or any loss of or corruption of data however so arising out of or in connection with this agreement whether or not that party had been informed of or was aware that there was a serious possibility of such loss.

The Licensor may terminate this Licence if the Licensee fails to remedy any breach of its obligations under this agreement within 30 days of notice from the Licensor, requesting the Licensee to remedy such breach.

The validity, construction and performance of this agreement shall be governed by the law of England. In the event of a dispute arising in connection with this agreement, the Parties shall attempt to settle their differences in an amicable manner; in the event that any dispute cannot be so settled, it shall be finally settled under the Rules of Conciliation and Arbitration of the International Chamber of Commerce by three arbitrators appointed in accordance with the said rules; unless otherwise agreed in writing, the arbitrators shall sit in London, England. The proceedings shall be in the English language. In accordance with Sections 45 and 69 of the Arbitration Act 1996, the right of appeal by either Party to the English Courts on a question of law arising in the course of any arbitral proceedings or out of an award made in any arbitral proceedings is hereby agreed to be excluded.

The Licensor makes no warranty with regard to the software provided to the Licensee to access the Archive Products. The software is provided on an "as is" basis. Any warranty implied by statute or otherwise is hereby excluded from this agreement to the fullest extent permissible by law.

The Licensee acknowledges that Direct Access to the Archive Products may be unavailable, delayed or interrupted. The Licensor takes no responsibility for, and will not be liable for, any unavailability, delay or interruption in Direct Access to the Archive Products at any time or for any period.

The Licensee undertakes to use identifiers, passwords and any other security information or device received from ECMWF for the sole purposes for which access to the ECMWF systems has been granted and not to make them available to third parties under any circumstances.

5.2 NOMADS overview

The NOMADS framework is a distributed data system that promotes the combining of data sets between distant participants using open and common server software and methodologies. Users effectively access model and observational data and products in a flexible and efficient manner from archives or in real-time through existing Internet infrastructure. The framework of the server and software used to manipulate the archive data base and distribute the real time data is the Open source Project for a Network Data Access Protocol (OPeNDAP) formally known as the Distributed Ocean Data System (DODS). OPeNDAP is a binary-level protocol designed for the transport of scientific data subsets over the Internet. Through OPeNDAP applications can open and read subsets from remote data sets. The Grid Analysis and Display System (GrADS) is an open source desktop tool that integrates the analysis and display of a wide variety geophysical data. OPeNDAP is an open source software project that provides a way to share data over the Internet. NOMADS combines these technologies to greatly expand the utility of the client and server concept. DODS/OPeNDAP-enabled clients such as GrADS, Ferret, IDL, Matlab, and end-user programs linked with various DODS/OPeNDAP client-libraries, can can access datasets and subsets of datasets over the network from DODS/OPeNDAP servers. The server can perform calculations over large amounts of data and only transmit (fairly small) results back to the client. This avoids transmitting large amounts of data to the client. With the NOMADS Web Interface users can create, display, and download plots, use ftp to download parameters and files, and use the NOMADS Distributed Data Server using OPeNDAP and the GrADS Data Server (GDS) to access the data directly from any preferred client or desktop. It is feasible also to perform remote calculations using scripts under the GDS. The available datasets are catalogued in the NCDC Model Data Inventories section in order to locate datasets of interest and dataset documentation. The following access functions are provided:

Plot – allowing the user to build a collection, then select parameters from the collection, and create plots

Grid/Scale - type, resolution and domain of grid

FTP4u - allows to building a collection, select files or parameters from the collection, and ftp selected data to NCDC anonymous ftp server

FTP – user can browse lists of files on NCDC server and download files using the ftp protocol

HTTP – user can browse lists of files on NCDC server and use any browser to download files

GDS - GrADS DODS/OPeNDAP server must be accessed to retrieve the data of this dataset. Use the GDS server to obtain the URL for files users want to access from their own desktop DODS/OPeNDAP enabled client.

OPeNDAP - DODS/OPeNDAP must be accessed to retrieve the data of this dataset. DODS server helps the user to obtain the URL for files to access from user's desktop DODS/OPeNDAP enabled client.

For data users, the currently available OPeNDAP servers include data formats for many desktop applications, and the user base for OPeNDAP enabled clients is expanding. NOMADS users have multiple options to access data from their desktops. The OPeNDAP web site maintains a list of OPeNDAP-enabled analysis applications and libraries. With a certain level of understanding of the OPeNDAP protocol, a user can also use a web browser or even a spreadsheet to access OPeNDAP data.

5.2.1 Datasets of special interest for DRIHM

NCEP High Resolution Global Forecast System (1 degree GFS), North American Mesoscale (12km NAM), and Rapid Update Cycle (20km RUC)

The table below lists the available weather model datasets both on-line or archived at NCDC. Each entry in the table below provides links to access methods for that dataset. Note: data may be delayed up to one day. For real-time access, user are referred to the NCEP R/T NOMADS.

Hi-Res NCEP NWP Model Datasets in NCDC Repository

Model	Grid/Scale	Runs/Day	Fcst Proj	Plot|FTP	FTP	HTTP	Data Server
High Resolution Analysis and Forecasts							
Historical							
Meso-ETA	218 (12km) Domain	4/day	0-60 hrs.	Plot|FTP4u	---	---	---
GFS-AVN	003 (1.0°) Domain	4/day	0-240 hrs.	Plot|FTP4u	---	---	---
Near Real-Time and Historical							
NAM	218 (12km) Domain	4/day	0-60 hrs.	Plot|FTP4u	FTP	HTTP	GDS TDS
GFS	003 (1.0°) Domain	4/day	0-180 hrs.	Plot|FTP4u 003	FTP 003	HTTP 003	GDS 003 TDS 003
	004 (0.5°)Domain			Plot|FTP4u 004	FTP 004	HTTP 004	TDS 004
RUC	252 (20km) Domain	hourly	0-3,6,9,12 hrs.	Plot|FTP4u 252	FTP 252	HTTP 252	GDS 252 TDS 252
	130 (13km)Domain			Plot|FTP4u 130	FTP 130	HTTP 130	TDS 130
Analysis-Only, Period of Record (March 02, 2004 to latest full month)							
GFS-ANL	003 (1.0°) Domain	4/day	0,(3-6 pcp.)	Plot|FTP4u 003	FTP	HTTP	GDS 003 TDS 003
	004 (0.5°) Domain			Plot|FTP4u 004			GDS 004 TDS 004
RUC-ANL	130 (13km) Domain	hourly	0,(0-1) hrs.	---	---	HTTP	TDS 130
	252 (20km) Domain						TDS 252
NAM-ANL	218 (12km) Domain	4/day	0,(3,6 pcp.)	Plot|FTP4u	FTP	HTTP	GDS TDS
NAM (WRF-NMM) - North American Mesoscale model, As of 2006-Jun-20 (Weather Research & Forecasting - Nonhydrostatic Mesoscale Model) Legacy versions and dates: ETA, Meso-ETA (Pre 2005-Feb-15), NAM (2005-Feb-15). The analysis from all these are unified under NAM-ANL.							

NCEP Climate Forecast System version 2

The Climate Forecast System (CFS) is a fully coupled model representing the interaction between the Earth's oceans, land and atmosphere. This model became operational in August 2004, and has been run in retrospective mode back 25 years to 1981. Users seeking access to

CFS archived and real-time data should visit the National Centers for Environmental Prediction's CFS homepage. Available on the site are retrospective daily and monthly data, climatologies, links to operational CFS model runs, CFS documentation, news, and other products.

Reanalysis

The NCEP Climate Forecast System Reanalysis (CFSR) was designed and executed as a global, high resolution, coupled atmosphere-ocean-land surface-sea ice system to provide the best estimate of the state of these coupled domains over the 31-year period of 1979 to 2009. It will be extended as an operational real time product into the future.Get CFSR data here. The CFSR data was developed by NOAA's National Centers for Environmental Prediction (NCEP). The data for this study are from NOAA's National Operational Model Archive and Distribution System NOMADS which is maintained at NOAA's NCDC.

5.3 HyMeX overview

HyMeX database incorporates various dataset types from different disciplines, either present or historical, either operational or research. It also contains detailed metadata about datasets relevant for the HyMeX project. Precisely, the HyMeX database will contain: -HyMeX field campaigns datasets; - Historical, LOP, EOP data; -Satellite products; -Model outputs from atmosphere, ocean or hydro-meteorological operational analysis and forecasts, and from research simulations The data policy has been endorsed by HyMeX International Governing Boards (IGB) in September 2010, and can be downloaded from the database gateway (http://mistrals.sedoo.fr/HyMeX). The data catalogue has been fully designed and developed. It is adapted to all the HyMeX data, whatever the type is. It complies with international standards (ISO 19115; Directory Interchange Format; Global Change Master Directory Thesaurus). A set of forms has been set online in order to allow HyMeX scientists to document the satellite and model products they will need or produce in the framework of the project. In situ observations can be described through a registration tool located on the HyMeX website (www.hymex.org). The database gateway has been developed using the SPIP CMS. The HyMeX database offers different tools to browse its metadata catalogue: • A search Tool, which allows browsing the catalogue using several criteria. • Sorted lists of the datasets by thematic keywords, by measured parameters (following the GCMD Science keywords), by instruments or by platform type. Data already available in the database can be accessed directly from the metadata forms. In-situ data can be ordered through a shopping-cart web interface, while satellite products, model outputs and radar data are accessible through a FTP site. Interoperability between the two data centres will be enhanced by the OpenDAP protocol associated with the Thredds software, which may also be implemented in other data centres that manage data of interest for the HyMeX project. The database web site also proposes a registration procedure including the form to be filled by the user and an administrator corner. The different databases (in situ database, satellite database, model database) are now under development in accordance with the data policy issues. Historical, Long-term Observation Period (LOP), Enhanced Observation Period (EOP) datasets including in-situ and satellite observations, as well as model outputs are already available. At the approach of the Special Observing Period (SOP) beginning, scientists are invited to communicate the datasets that should be incorporated in priority in the HyMeX database, and in particular the data that shall be available in real-time during the SOP.

5.3.1 Datasets of special interest for DRIHM

- Aircraft
  - ACAR - US aircraft reports, GTS data
  - AIREP - Manual aircraft reports, GTS data
  - AMDAR - Automatic aircraft reports, GTS data
- Atmospheric sites
  - Balearic islands site (BA)
    - Mallorca
    - Operational surface weather automatic stations over Spain -AEMET data
- Corsica site and super-sites (CO)
  - Radiosonde, Ajaccio
- Balloons
  - Operational atmospheric soundings (GTS) from land RS stations
  - Operational atmospheric soundings (GTS) from marine RS stations
  - PILOT - Upper wind from fixed land stations, GTS data
- Drifters, buoys and moorings
  - Azur Buoy
    - Anemometer, Azur Buoy
    - Operational surface weather observation stations over France - Daily data
    - Precipitation gauge, Azur Buoy
    - Pyranometer, Azur Buoy
    - Pyrgeometer, Azur Buoy
    - Thermosalinograph, Azur Buoy
  - Lion Buoy
    - Anemometer, Lion Buoy
    - Lion Buoy, mooring line, 0-200 m ocean temperature
    - Operational surface weather observation stations over France - Daily data
    - Operational surface weather observation stations over France - Hourly data
    - Precipitation gauge, Lion Buoy
    - Pyranometer, Lion Buoy
    - Pyrgeometer, Lion Buoy
    - Thermosalinograph, Lion Buoy
  - Other sites
    - Operational Buoy observations (GTS)
- Ground networks
  - Ground-based GNSS networks
    - GPS_France
  - Meteorological operational network
    - EUROPROFIL - Wind and temperature profiles from ground instruments, GTS data
    - Operational atmospheric soundings (GTS) from land RS stations
    - Operational surface weather observation stations over France - Daily data
    - Operational surface weather observation stations over France - Hourly data

        - Operational synoptic surface observations from fixed stations (GTS)
        - PILOT - Upper wind from fixed land stations, GTS data
    - OHMCV rain gauges network
        - Hpiconet
        - Hpiconet, Lavilledieu-Ecole
        - Hpiconet, Lussas-Salle-Polyvalente
        - Hpiconet, Mirabel-Ferme-Pradel
        - HPiconet, Mirabel-Les-Blaches
        - HPiconet, Mirabel-Mairie
        - HPiconet, Saint-Germain-Ecole
        - HPiconet, Villeneuve-de-Berg-College
    - Other networks
        - Radiosonde, Ajaccio
        - Radiosonde, Nimes
        - Rainfall and River Height, Magra Bassin
- Hydrometeorological sites
    - Catalan hydrometeorological observatory (CA)
        - Operational surface weather automatic stations over Spain -AEMET data
        - Operational Weather Radar AEMET, Barcelona (Corbera)
    - Central Italy (CI)
        - Precipitation and Temperature Data, Abruzzo
    - Cevennes-Vivarais hydrometeorological observatory (CV)
        - OHMCV rain gauges network
            - Hpiconet
                - Hpiconet, Lavilledieu-Ecole
                - Hpiconet, Lussas-Salle-Polyvalente
                - Hpiconet, Mirabel-Ferme-Pradel
                - HPiconet, Mirabel-Les-Blaches
                - HPiconet, Mirabel-Mairie
                - HPiconet, Saint-Germain-Ecole
                - HPiconet, Villeneuve-de-Berg-College
    - Liguria-Tuscany site (LT)
        - Hourly surface stations, Liguria
        - Rainfall and River Height, Magra Basin
    - Southern Italy
        - Salento University Campus
        - Meteorological network station, Calabria
    - Valencia site (VA)
        - Operational surface weather automatic stations over Spain -AEMET data
        - Operational Weather Radar AEMET, Valencia (Cullera)
- Ocean sites
    - Adriatic
        - Operational surface weather observation stations over France - Hourly data
- North-Western Mediterranean
    - Gulf of Lion
    - Ligurian sea
- Whole Mediterranean Sea
    - BATHY - Sea surface and soundings observations, GTS data
    - Operational atmospheric soundings (GTS) from marine RS stations

    - Operational Buoy observations (GTS)
    - Operational synoptic surface observations from ship (GTS)
    - TESAC - Sub-surface profiling floats, GTS data
- Questionnaires and post-event surveys
    - Post event survey, Gard 2002
- Radars
    - ALERIA radar 5min cumulative rainfall in mm
    - ALERIA radar 5min multi-elevation reflectivity and radial wind speed
    - BOLLENE radar 5min cumulative rainfall in mm
    - BOLLENE radar 5min multi-elevation reflectivity and radial wind speed
    - COLLOBRIERE radar 5min cumulative rainfall in mm
    - COLLOBRIERE radar 5min multi-elevation reflectivity and radial wind speed
    - European Radar reflectivity composite (Odyssey)
    - French Radar composite 5min cumulative rainfall in mm
    - MONTCLAR radar 5min cumulative rainfall in mm
    - MONTCLAR radar 5min multi-elevation reflectivity and radial wind speed
    - NIMES radar 5min cumulative rainfall in mm
    - NIMES radar 5min multi-elevation reflectivity and radial wind speed
    - Operational Weather Radar AEMET, Almeria (Nijar)
    - Operational Weather Radar AEMET, Barcelona (Corbera)
    - Operational Weather Radar AEMET, Madrid
    - Operational Weather Radar AEMET, Mallorca (Llucmajor)
    - Operational Weather Radar AEMET, Murcia (Fortuna)
    - Operational Weather Radar AEMET, Valencia (Cullera)
    - Operational Weather Radar Province Autonome Trento e Bolzano (Macaion)
    - Operational Weather Radar SMER, Capofiume
    - Operational Weather Radar SMER, Gattatico
    - OPOUL radar 5min cumulative rainfall in mm
    - OPOUL radar 5min multi-elevation reflectivity and radial wind speed
    - SEMBADEL radar 5min cumulative rainfall in mm
    - SEMBADEL radar 5min multi-elevation reflectivity and radial wind speed
- Ships
    - BATHY - Sea surface and soundings observations, GTS data
    - Operational synoptic surface observations from ship (GTS)
    - TESAC - Sub-surface profiling floats, GTS data

5.3.2 HYMEX Data policy

The HyMeX data policy covers the data collected during the HyMeX Long Observing Periods, Extended Observing Periods, and Short Observing Periods. It may be revised in order to tit the HyMeX programme and database evolutions at best, as deemed necessary by the HyMeX IGB. At the term of this ten-year period, the HyMeX IGB will be responsible for defining how data archive and data access shall be handled.

The restricted access granted to non public data through this policy allows the sole and direct use of the data for scientific and educational activities.

Since the data are originally collected for specific purposes, no warranty is given as to their suitability for the use intended by the recipient. The database centers and the data providers have no liability for any loss, damage, claim, demand, cost or expense directly or indirectly arising from any use, receipt or supply of data under this agreement. - Data owners are the agencies or institutes funding the data collection.

- Principal investigators are associated with an instrument or site from instrument deployment and data collection, to data processing and transfer to the database. A principal investigator is the scientist responsible for the instrument or site or any person (collaborator, student) that he/she may suggest.

- Data providers provide data to the database. They are either data owners or principal investigators.

- Core users are HyMeX data providers and scientists from institutions providing funding or in-kind support to HyMeX, according to the criteria set by the HyMeX ISSC. Core users can access all the data. They are granted exclusive access to the HyMeX data for a default period of 2 years from data deposit deadline on the database as defined in paragraph 3.

- Associated scientists are not directly involved in HyMeX but in scientific studies designed to meet HyMeX objectives and educational activities related to the dissemination of HyMeX science. They can access data open to research activities and HyMeX data once the period of exclusive access is over.

6 Requirements Collection

6.1 The DRIHMS Legacy

The DRIHMS project (www.drihms.eu) was initiated to provide a detailed analysis of the needs for such a step forward in the cooperation between HMR and ICT scientists, aiming at the creation of a HMR Virtual Research Community, and to provide a foundation for the timely deployment of an e-Science environment for hydrometeorology research on extreme events, as well as on climate change impacts and adaptation.

The project organized a set of networking activities, involving both HMR scientists and ICT scientists related to Grid Computing, whose main objectives were to:

- Understand how to overcome the current limitations in the sharing of data, tools and knowledge in the European hydrometeorology research (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 hydrometeorological 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. The consultation process, based on the Delphi method [7] used also other networking instruments such as consultation meetings, open meetings, and communities polling methodologies.

The purpose of the HMR questionnaire was to define important topics in Hydro-Meteorology Research (HMR) that can benefit from advances in Information and Communication Technologies (ICT), and thus set a joint HMR-ICT research and applications agenda. These HMR hot topics are: data merging/fusion; probabilistic forecasting; model verification metrics; precipitation downscaling.

The questionnaire respondents were requested to score the importance of the aforementioned HMR hot topics for the Hydro Meteorological Research in the next 3 years:

- Probabilistic forecasting is the top HMR topic considering the full audience and represents the most relevant argument within the Meteorology sub-community;

- The model verification metrics appears to be the second reference topic for the full audience and the top one for the Hydro-Meteorology and Hydrology sub-communities;
- The data merging/fusion and precipitation downscaling, in contrast, obtained in general lower scores and correspondingly ranking, both among the full audience and within the different sub-communities.

Through the survey, also the accompanying ICT priorities for each research hot topic were identified.

Concerning data merging/fusion the following issues are perceived as the most important:

- The definition of common interchange data formats is the 'hottest' topic for all communities;
- The availability and reliability of high capacity communication channels from/to data sources is among the most relevant for all the sub-communities and it is ranked second when the full audience is addressed;
- The definition of libraries for merging and analyzing data from multiple sources is also among the most relevant for all the sub-communities and it is ranked third when the full audience is addressed;

The following issues are considered as the most important for the probabilistic forecasting:

- Definition of common interchange data formats (like netcdf, grib, bufr) and protocols (e.g., atmospheric-hydrologic models coupling/interfacing);
- Availability and reliability of high performance computing resources for probabilistic forecasting;
- The definition of libraries for merging and analyzing data from multiple sources is also among the most relevant for all the sub-communities and it is ranked third when the full audience is addressed.

About model verification, the availability of model output and observations in compatible formats for easy verification is the topmost topic for the full audience and for the Hydro-Meteorology and Meteorology sub-communities, while it ranks second for the Hydrology one.

Furthermore not surprising, since we are dealing with model verification:

- The availability of well documented metrics of verification (single output or ensemble) and codes for computation of those metrics is perceived as very important;
- The availability of standards, in this case to perform model verification, is recognized as an important issue at the full audience scale and also by many sub-communities.

Finally for precipitation downscaling:

- The definition of common interchange data formats and protocols is the reference topic for the full audience and a large amount of the sub-communities here considered;

The definition and dependability of catalogues of available algorithms and codes for downscaling is perceived among the most important aspect both at the full audience scale and for reference sub-communities as hydrometeorology and meteorology.

6.2 User Stories

6.2.1 Meteo Prediction

1. Select the initial and boundary condition from the global circulation model, available on the platform (for example provided by MARS ECMWF and NOMADS databases).
2. Initial and boundary conditions are selected by specifying a temporal window and domain regions (optionally, nested domains). The domain region is identified with an interactive map (like google maps), and the related ancillary data (DEM, land use, soil type, vegetation type) are prepared.
3. Select WRF as Numerical Weather Prediction Model (NWP-Model). Available models on the platform are Cosmo, WRF or Meso-NH.
4. Configure the NWP-Model parameters: grid spacing, timestep, timeframe, output intervals, physics, dynamics and numeric options, restart files, …
5. Start the execution of the simulation.
6. As soon as the results are available, a notification is received.

6.2.2 Hydro Model Calibration

1. Select DRIFT as an hydro model. Available models on the platform are DRIFT, Continuum, RIBS, HBV, HYPROM
2. Select Bisagno river catchment from a GIS database.
3. Collect static data (DEM, land use, soil type, vegetation type)
4. Calibrate the model on a few events. For each event:
   a. Collect point time series data for model calibration (for example from HyMeX database).
   b. Select the time period for model calibration and assess calibration results.
   c. Fix the calibration parameters
5. Store the calibration parameters for the DRIFT model on Bisagno river catchment.

6.2.3 Hydro Prediction

1. Select DRIFT as an hydro model. Available models on the platform are DRIFT, Continuum, RIBS, HBV, HYPROM
2. Select Bisagno river catchment with static data and configuration parameters for DRIFT hydro model on such a catchment (As in Subsection 6.2.2)
3. Collect point time series data for model execution (for example from HyMeX database).
4. Select the time period.
5. Start the execution of the simulation.
6. As soon as the results are available, a notification is received.

6.3 Use Cases

6.3.1 Meteo Prediction

Name	Meteo Prediction
Summary	The procedure required to compute a meteorological prediction adopting a Numerical Weather Prediction Model
Actors	Hydro-Meteorologist, Numerical Weather Prediction Model, Global Circulation Model
Preconditions	Availability of a global circulation model
Basic Course of Events	1. Login on the DRIHM platform 2. Select the initial and boundary condition from the global circulation model. The domain region is identified with an interactive map 3. Prepare related ancillary data (DEM, land use, soil type, vegetation type) 4. Select a Numerical Weather Prediction Model (NWP-Model). Configure the NWP-Model parameters: grid spacing, timestep, timeframe, output intervals, physics, dynamics and numeric options, restart files, … 5. Start the execution of the simulation.
Alternative Paths	Instead of a single domain region, nested domains can be selected
Postconditions	Multiple rainfall, temperature, wind and pressure maps ready for analysis or further processing (e.g. hydrological simulation)
Models / Tools used	NWP-Model such as COSMO, WRF or Meso-NH
Data accessed / produced	Accessed: initial and boundary conditions from the global circulation model Produced: Multiple rainfall, temperature, wind and pressure maps - usually very large datasets
Local / Remote	Execution can be local or remote, on single core or HPC

Interfaces between components	Input interface should be able to read initial and boundary condition from global circulation models. Generated maps are to be produced / translated in a standard format (NetCDF, grib, grib2)

6.3.2 Hydro Model Calibration

Name	**Hydro Model Calibration**
Summary	The procedure required to calibrate an hydrological model on a river catchment
Actors	Hydro-Meteorologist, Hydro Model, GIS database
Preconditions	Availability of static data (DEM, land use, soil type, vegetation type) Observed initial condition. Point time series on a regular or irregular mesh (e.g. rainfall, temperature as input and measured discharge for output validation)
Basic Course of Events	1. Select a hydro model. 2. Select a river catchment from a GIS database 3. Collect static data (DEM, land use, soil type, vegetation type) 4. Calibrate the model on a few events. For each event: 4.1. Collect point time series data for model calibration. 4.2. Select the time period for model calibration and assess calibration results. 4.3. Fix the calibration parameters 5. Store the calibration parameters for the selected hydro model on the river catchment.
Alternative Paths	None
Postconditions	Calibrated Hydro model on a river catchment
Models / Tools used	Hydro Model such as DRIFT, Continuum, RIBS, HBV, HYPROM

Data accessed / produced	Accessed: static data (DEM, land use, soil type, vegetation type) from a GIS database Historical point time series on a regular or irregular mesh (beside rainfall, depending on the hydro model, temperature, wind, pressure may be required) Historical discharge measured data for output validation Produced: hydro model calibration data
Local / Remote	Execution can be local or remote, usually on single core HYPROM can be executed on HPC
Interfaces between components	None

6.3.3 Hydro Prediction

Name	**Hydro Prediction**
Summary	The procedure required to perform an hydro prediction
Actors	Hydro-Meteorologist, Hydro Model, Meteo Model, Field sensors
Preconditions	Availability of a calibrated hydro model on a desired river catchment. Measured or predicted point time series on a regular or irregular mesh (rainfall and, according to the hydro model, soil moisture, temperature, wind and pressure maps).
Basic Course of Events	1. Select a hydro model. 2. Select a river catchment with static data and configuration parameters for the hydro model on such a catchment (As in Subsection 6.2.2) 3. Select the time period. 4. Collect point time series data for model execution. 5. Set initial condition according to the hydro model. 6. Start the execution of the simulation.
Alternative Paths	Ensemble grid data instead of deterministic grid data
Postconditions	Discharge prediction on a river catchment

Models / Tools used	Hydro Model such as DRIFT, Continuum, RIBS, HBV, HYPROM
Data accessed / produced	Accessed: hydro model calibration data, static data, rainfall and, according to the hydro model, soil moisture, temperature, wind and pressure maps. Produced: Discharge prediction data
Local / Remote	Execution can be local or remote, usually on single core Ensemble computation can be parallelized
Interfaces between components	Deterministic and ensemble grid data can be generated by a meteo prediction, or obtained from observation (on field, satellite, radar, …).

6.4 Requirements

Given the DRIHMS Legacy, and the use cases described in Section 6.3, we derived four sets of requirements: Data Management, Model Management, Workflow Management and Post processing.

6.4.1 Data Management

The first set of requirements focus on data management, and mainly refers to services for facilitating interaction between existing data and baseline methods.

ID	Name
D1	Ingest Modelled Deterministic Grid Data
D2	Ingest Modelled Ensemble Grid Data
D3	Ingest Measured Point Series Data
D4	Ingest Measured Coverage Data
D5	Ingest Measured Track Data
D6	Allow Interoperability of Deterministic Global Circulation Model Data
D7	Allow Interoperability of Deterministic Grid Data Inputs
D8	Allow Interoperability of Ensemble Grid Data Inputs
D9	Allow Interoperability of Point Series Data Inputs
D10	Allow Interoperability of Coverage Data Inputs

D11	Allow Interoperability of Track Data Inputs
D12	Allow Interoperability of Observed Grid Data
D13	Store and Serve Grid Data
D14	Store and Serve Point Series Data
D15	Store and Serve Coverage Data
D16	Store and Serve Track Data
D17	Store and Serve Output Data
D18	Store and Serve Hydro Model Calibration
D19	Provide General Background Information

In the following we define the file formats relevant for models interoperability

As modeled Deterministic Global Circulation Model Data we consider the data provided by the European Integrated Forecasting System (IFS-ECMWF) stored in formats such as GRIB, GRIB2 and/or the data provided by the NCEP Global Forecasting System (GFS) stored in formats such as GRIB and NetCDF.

As modeled deterministic/ensemble grid data we consider predicted time series on a regular mesh (rainfall and optionally soil moisture, temperature, wind and pressure maps) such as the data provided by COSMO model, stored in GRIB and GRIB2 and/or the data provided by WRF, stored in NetCDF format.

As measured point series data we consider point time series on a regular or irregular mesh (rainfall and optionally soil moisture, temperature, wind and pressure maps).
Up to now no standards are widely adopted. Possible data formats currently available are WaterML and O&M (Observations and Measurements).

As coverage data we consider static data (DEM, land use, soil type, vegetation type) from a GIS database such as ESRI.

As Observed Grid Data we consider the data provided by meteorological radars and satellites on regular mesh (e.g. reflectivity) stored in formats such as BUFR, HDF, HDF4 and HDF5

6.4.2 Model Management

The second set of requirements focus on model management, and mainly refers to services for interoperability of Meteorological and Hydrological models.

ID	Name
M1	Allow Interoperability of "Large" Scale Meteorological Models
M2	Allow Interoperability of "Small" Scale Meteorological Models
M3	Run Ensemble Hydrologic Model on HPC
M4	Allow Interoperability of Hydrologic Models

6.4.3 Workflow Management

The third set of requirements focus on workflow management, i.e. the ability to chain two or more model, in order to perform complex simulation. Workflow management requirements heavily depend on Data and Model requirements.

ID	Name
W1	One Way Connection from "Small" Scale Deterministic Meteorological Models to Hydrologic Models
W2	One Way Connection from "Small" Scale Ensemble Meteorological Models to Hydrologic Models
W3	Two Way Connection Between Deterministic Hydrologic Models
W4	Handle Model State (for re-start etc.)
W5	Manual Construction of Workflows
W6	Adaptation Across Differing Model Grids / Meshes
W7	Adaptation Across Differing Temporal Definitions
W8	Adaptation Across Other Differing Features
W9	Access Control for Individual Users / User Groups
W10	Saving and Re-loading Workflows (including Models and Adaptors Used and Versions)
W11	Track User Usage (for Future Accounting and Billing)
W12	Component Licence Management (for Future Accounting and Billing)

DRIHM
DISTRIBUTED RESEARCH INFRASTRUCTURE FOR HYDRO-METEOROLOGY

6.4.4 Post Processing

The fourth set of requirements focus on basic services for checking model execution and to export/analyze the outcome of models execution

ID	Name
P1	Production and Display of Model Run Statistics
P2	Production and Display of Model Run Error Codes
P3	Visualisation of Numeric Results
P4	Visualisation of Charted Results
P5	Visualisation of Mapped Results
P6	Export of Results to Common Formats (e.g. csv, shapefile)

6.5 Requirements dependencies

D2,D7 and D8 depends on D1

D8 depends on D2

D9 depends on D3

D10 depends on D4

D11 depends on D5

M4 depends on M3

W1 depends on D7, M2, M4, W5 and W6

W2 depends on D8, M2, M4, W5 and W6

W3 depends on M4

6.6 Requirements mapping to baseline

The following tables represent how baseline models are affected by the requirements.

	COSMO	WRF	MESO-NH	DRIFT	CONTINUUM	RAINFARM	RIBS	HBV	HYPROM	RAD-TRAM	CB-TRAM
D1	-	-	-	X	X	X	X	X	X	X	X
D2	-	-	-	X	X	X	X	X	X	X	X
D3	-	-	-	X	X	-	X	X	X	-	-
D4											
D5				X	X	X	X	X	X		
D6	X	X	X								
D7	-	-	-	X	X	X	X	X	X	X	X
D8	-	-	-	X	X	X	X	X	X	X	X
D9	-	-	-	X	X	-	X	X	X	-	-
D10	X	X	X	X	X	X	X	X	X	X	X
D11	X	X	X	X	X	X	X	X	X	X	X
D12	X	X	X	X	X	X	X	X	X	X	X
D13											
D14	X	X	X	X	X	X	X	X	X	X	X
D15	-	-	-	X	X	-	X	X	X	-	-
D16	X	X	X	X	X	X	X	X	X	X	X
D17	-	-	-	X	X	-	X	X	X	-	-
D18											
D19											

	COSMO	WRF	MESO-NH	DRIFT	CONTINUUM	RAINFARM	RIBS	HBV	HYPROM	RAD-TRAM	CB-TRAM
M1	X	X	X	-	-	-	-	-	-	-	-
M2	X	X	X	-	-	-	-	-	-	-	-
M3	-	-	-	X	X	X	X	X	X	-	-

M4	-	-	-	X	X	X	X	X	X	-	-

	COSMO	WRF	MESO-NH	DRIFT	CONTINUUM	RAINFARM	RIBS	HBV	HYPROM	RAD-TRAM	CB-TRAM
W1	X	X	X	-	-	-	-	-	-	-	-
W2	X	X	X	-	-	-	-	-	-	-	-
W3	-	-	-	-	-	-	-	-	-	-	-
W4	X	X	X	X	X	X	X	X	X	X	X
W5	X	X	X	X	X	X	X	X	X	X	X
W6	X	X	X	X	X	X	X	X	X	X	X
W7	X	X	X	X	X	X	X	X	X	X	X
W8	X	X	X	X	X	X	X	X	X	X	X
W9	X	X	X	X	X	X	X	X	X	X	X
W10	X	X	X	X	X	X	X	X	X	X	X
W11	-	-	-	-	-	-	-	-	-	-	-
W12	-	-	-	-	-	-	-	-	-	-	-

7 Conclusion

The initial version of this report, due at PM8 takes advantage of all the information provided by baseline version of the experiment suites (WP8), of the outcomes of the DRIHMS project. The final version (due at PM 20) will fully account for the new requirements that will progressively arise from the in deep understanding of the innovative scenarios associated with the three DRIHM experiments (WP8).

This version of the document focus on the development of an infrastructure offering those services required to run experiment suites 1 and 2. The elicitation process started from an analysis of the outcomes of the DRIHMS project, a careful description of existing methods (the baseline), and their shortcomings, and thanks to User Scenarios describing the execution of a meteorological model and the calibration and execution of an hydro model.

The baseline version of the experiment suites 1 and 2 can be considered as a benchmark for the infrastructure outlined in this version of the requirement document.

8 References

[1]. Balaž, A., Prnjat, O., Vudragović, D., Slavnić, V., Liabotis, I., Atanassov, E., Jakimovski, B. and Savić, M., 2010. Development of Grid e-Infrastructure in South-Eastern Europe, J. Grid. Computing, 10.1007/s10723-011-9185-0.

[2]. Bauer, S., 1974. A modified Horton equation during intermittent rainfall. Hydrol. Sci. Bull., 19, 219–229.

[3]. Bell, T. L., and P. K. Kundu, 2003: Comparing satellite rainfall estimates with rain gauge data: Optimal strategies suggested by a spectral model. J. Geophys. Res., 108.4121, doi:10.1029/2002JD002641.

[4]. Bergström, S., 1976. Development and application of a conceptual runoff model for Scandinavian catchments. (Department of Water Resources Engineering, Lund Institute of Technology/University of Lund, Bulletin Series A, No. 52, 134 pp.).

[5]. Bradner, S. (ed.): Key Words for Use in RFCs to Indicate Requirements Levels. The Internet Engineering Task Force Best Practice, March 1997. http://www.ietf.org/rfc/rfc2119.txt

[6]. Brousseau P., L. Berre, F. Bouttier, G. Desroziers, 2011: Background error covariances for a convective scale data assimilation system: AROME 3D-Var. Quart. J. Roy. Meteor. Soc., 137, 409.

[7]. Dalkey N. An experimental study of group opinion, 1969: The Delphi method. Futures, 408–26.

[8]. Diskin, M. H., Nazimov, N., 1994. Linear reservoir with feedback regulated inlet as a model for the infiltration process, J. Hydrol., 172, 313–330.

[9]. Dickinson, R. E., Henderson-Sellers, A., Kennedy, P. J., 1993: Biosphere-Atmosphere Transfer 25 Scheme (BATS) Version 1e as Coupled to the NCAR Community Climate Model, National Centre for Atmospheric Research, Boulder, Colorado, NCAR/TN-387+STR, 72 pp.

[10]. Doms, G., and J. Förstner, 2004: Development of a kilometer-scale NWP-System: LMK. COSMO Newsletter, No. 4, Deutscher Wetterdienst, Offenbach, Germany, 159–167.

[11]. Gabellani, S., Silvestro, F., Rudari, R., Boni, G., 2008. General calibration methodology for a combined Horton-SCS infiltration scheme in flash flood modeling. Nat. Hazards Earth Science., 8, 1317 - 1327.

[12]. Gal-Chen, T., and R. C. J. Sommerville, 1975: On the use of a transformation for the solution of the Navier-Stokes equations. J. Comput. Phys., 17, 209-228.

[13]. Garrote, L., R.L. Bras, 1995, A distributed model for real-time flood forecasting using digital elevation models J. Hydrol., 167, 279–306.

[14]. Gassmann, A., 2005: An improved two-time-level split-explicit integration scheme for non-hydrostatic compressible models. Meteor. Atmos. Phys., 88, 23–38.

[15]. Giannoni, F., Roth., G., Rudari, R., 2000. A Semi – Distributed Rainfall – Runoff Model Based on a Geomorphologic Approach. Physics and Chemistry of the Earth, 25/7-8, 665-671.

[16]. Giannoni, F., Roth., G., Rudari, R., 2005. A procedure for drainage network identification from geomorphology and its application to the prediction of the hydrologic response. Advances in Water Resources, 28, 6, 567

[17]. Guillot, G., and T. Lebel, 1999: Disaggregation of Sahelian mesoscale convective system rain fields:

Further developments and validation. J. Geophys. Res., 104, 31533–31551.

[18]. I. Jacobson, M. Christerson, P. Jonsson: Object-Oriented Software Engineering. A Use Case Driven Approach; ACM Press 1992; ISBN 978-0201544350

[19]. Lafore, J.-P., J. Stein, N. Asencio, P. Bougeault, V. Ducrocq, J. Duron, C. Fischer, P. Héreil, P. Mascart, J.-P. Pinty, J.-L. Redelsperger, É. Richard, and J. Vila-Guerau de Arellano, 1998: The Meso-NH atmospheric simulation system. Part I: Adiabatic formulation and control simulations. Annales Geophysicae, 16, 90-109

[20]. Lanza, L. G., 2000: A conditional simulation model of intermittent rain fields. Hydrol. Earth Syst. Sci., 4, 173–183.

[21]. Matheron, G., The Intrinsic Random Functions and Their Applications, Advances in Applied Probability, 439-468, 1973.

[22]. Nickovic, S., G. Pejanovic, V. Djurdjevic, J. Roskar, M. Vujadinovic, 2010: HYPROM Hydrology surface-runoff prognostic model, Water Resour. Res., 46, p. W11506 http://dx.doi.org/10.1029/2010WR009195

[23]. Parodi, A., et al.: *DRIHMS: The White Paper*; ISBN 978-88-906068-0-9, 2011

[24]. Parodi, A. et al.: DRIHM Description of Work (DoW), 2011

[25]. Read, P. L., N. P. J. Thomas, S. H. Risch, 2000: An Evaluation of Eulerian and Semi-Lagrangian Advection Schemes in Simulations of Rotating, Stratified Flows in the Laboratory. Part I: Axisymmetric Flow. Mon. Wea. Rev., 128, 2835–2852. doi: http://dx.doi.org/10.1175/1520-0493(2000)128<2835:AEOEAS>2.0.CO;2

[26]. Rebora N, Ferraris L, von Hardenberg J, Provenzale A. RainFARM: rainfall downscaling by a filtered autoregressive model. J Hydrometeorology 2006;7(4):724–38.

[27]. Rosso, R., A. Peano, I. Becchi and G. A. Bemporad (editors), ADVANCES IN DISTRIBUTED HYDROLOGY, Water Resour. Pubblication, Highlands Ranch, Col. 1994.

[28]. Roth, G., P. La Barbera and M. Greco, On the description of the basin drainage structure, J. Hydrol., 187, 119-135, 1996.

[29]. Seity, Y., Brousseau, P., S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson, 2011: The AROME-France convective scale operational model. Mon. Wea. Rev., 139, 976-99, DOI: 10.1175/2010MWR3425.1

[30]. Shapiro, Melvyn, and Coauthors, 2010: An Earth-System Prediction Initiative for the Twenty-First Century. Bull. Amer. Meteor. Soc., 91, 1377–1388. doi: http://dx.doi.org/10.1175/2010BAMS2944.1

[31]. Shukla, J., R. Hagedorn, B. Hoskins, J. Kinter, J. Marotzke, M. Miller, T. N. Palmer, and J. Slingo, 2009: Revolution in climate prediction is both necessary and possible: A declaration at the World Modelling Summit for Climate Prediction. Bull. Amer. Meteor. Soc., 90, 175–178.

[32]. Shukla, J., T. N. Palmer, R. Hagedorn, B. Hoskins, J. Kinter, J. Marotzke, M. Miller, and J. Slingo, 2010: Toward a new generation of world climate research and computing facilities. Bull. Amer. Meteor. Soc., 91, 1407–1412.

[33]. Silvestro, F., Rebora, N. and Ferraris, L.: Quantitative flood forecasting on small and medium size basins: a probabilistic approach for operational purposes. Journal of Hydrometeorology, 12(6), 1432-

1446, 2011

[34]. Silvestro F., Gabellani S., Delogu F., Rudari R., Boni G., 2012. Continuum: A Parsimonious distributed hydrological model for water management and flood forecasting. Submitted.

[35]. Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. Mon. Wea. Rev., 132, 3019–3032.

[36]. Steppeler J, Doms G, Schättler U, Bitzer HW, Gassmann A, Damrath U, Gregoric G (2003) Meso-gamma scale forecasts using the nonhydrostatic model LM. Meteorol Atmos Phys 82:75–96

[37]. UNISDR, Disasters in numbers, http://www.unisdr.org/eng/terminology/terminology-2009-eng.html, 2009.

[38]. Vié, B., O. Nuissier, V. Ducrocq, 2011: Cloud-Resolving Ensemble Simulations of Mediterranean Heavy Precipitating Events: Uncertainty on Initial Conditions and Lateral Boundary Conditions. Mon. Wea. Rev., 139, 403–423. doi: http://dx.doi.org/10.1175/2010MWR3487.1