Chapter 8: Exploitation of multiple model layers within LEXIS Weather and Climate Pilot: an HPC-based approach

Paola Mazzoglio[1] (0000-0002-3662-9439)
Emanuele Danovaro[2] (0000-0003-2025-6794)
Laurent Ganne[3]
Andrea Parodi[4]
Stephan Hachinger[5] (0000-0001-8341-1478)
Antonella Galizia[4,6] (0000-0003-1672-3281)
Antonio Parodi[4] (0000-0002-8505-0634)
Jan Martinovic[7] (0000-0001-7944-8956)

[1] Department of Environment, Land and Infrastructure Engineering (DIATI), Politecnico di Torino, Italy
[2] European Center for Medium-range Weather Forecasts (ECMWF), United Kingdom
[3] ATOS, France
[4] CIMA Research Foundation, Italy
[5] Leibniz Supercomputing Centre (LRZ), Germany
[6] Istituto di matematica applicata e tecnologie (IMATI), Consiglio Nazionale delle Ricerche (CNR), Italy
[7] IT4Innovations, VSB – Technical University of Ostrava, Czech Republic

Chapter Abstract

The LEXIS (Large-scale EXecution for Industry & Society) Weather and Climate Pilot is developing a system for the prediction of water-related phenomena and their associated socio-economic impacts. The system is based on multiple models chained together, as global weather models, high-resolution regional weather models, domain-specific application models (hydrological and forest fire risk forecasts) and impact models providing information (such as air quality and rainfall intensity) for key decisions and policy makers. This chapter describes the key aspect of this pilot in terms of serving model output data and products with Cloud and High Performance Data Analytics (HPDA) environments, on top of a Weather Climate Data API (WCDA), as well as the porting of models on the LEXIS Infrastructure via different virtualization strategies (as virtual machine and dockers).

8.1 Introduction: Background and Driving Forces

The H2020 Large-scale Execution for Industry and Society (LEXIS) project aims to design and develop an advanced engineering platform at the confluence of High-Performance Computing (HPC), Big Data and Cloud solutions which leverages large-scale geographically-distributed resources from existing HPC infrastructure, employing Big Data analytics solutions and augmenting them with Cloud services (Parodi et al. 2021). The emphasis of LEXIS is on the interaction between HPC and cloud systems, built on top of data sharing and methods to compose workflows of tasks running on both cloud and HPC systems, understood as LEXIS Distributed Computed Infrastructure (DCI). A platform has been developed to enable these workflows and demonstrate its abilities through three large-scale socio-economic pilots, targeting aeronautics, weather, and earthquakes and tsunamis (Parodi et al. 2021). This chapter reports the key challenges and results concerning the Weather and Climate pilot which fully benefits of the project results in terms of orchestration system, Distributed Data Infrastructure (DDI), DCI, and workflows management.

8.2 The Weather and Climate pilot

The Weather and Climate pilot focuses on a complex system to provide a set of forecasts concerning weather, flood, forest fire, air pollution and agriculture by means of several complex workflows each consisting of various meteorological, hydrological and air quality components (Parodi et al. 2021). These workflows include ingestion of conventional and unconventional observations, global weather models, regional weather models, application models and socio-economic impact models (Figure 8.1). The workflows are run across disjoint computing resources (LEXIS DCI): global weather models are executed on ECMWF's HPC in the UK whilst regional weather models run on HPCs in Italy (CIMA), Germany (LRZ) or Czechia (IT4I). Application models and socio-economic impact models are instead executed on cloud-based resources in Germany or Czechia.

LEXIS is managing several weather and climate modelling tasks, namely WRF Model, RISICO, Continuum, ADMS and ERDS, which are described below.

The Weather Research and Forecasting (WRF) Model is a proven mesoscale numerical weather prediction (NWP) system, designed to serve both operational forecasting and atmospheric research needs (Powers et al., 2017). It features multiple dynamical cores, a 3-dimensional variational (3DVAR) data assimilation system, and a software architecture allowing for computational parallelism and system extensibility (Lagasio et al. 2019a, 2019b). WRF is suitable for a broad spectrum of applications across different scales, ranging from thousands of kilometers to meters.

RISICO (RISchio Incendi e COordinamento/Fire Risk and Coordination) is a mathematical model developed by the CIMA Research Foundation to support operators in forest fire prevention activities (Fiorucci et al. 2008). RISICO processes a continuous data flow consisting of meteorological information as weather forecast and satellite records. Parameters such as the moisture content of the vegetation, the wind and the orography of the territory allow to quantitatively assess the danger resulting from the eventual triggering of a forest fire both in terms of propagation speed and linear intensity of the flame front.

Continuum is a hydrological model developed by the CIMA Research Foundation to reproduce the flow of water within a basin (Silvestro et al. 2015). The model has a reduced number of parameters and is able to work both in the pre-event analysis and forecast phase and in the monitoring stage for the active control of hydrological events taking advantage of all the information available via in situ weather stations and satellite data.

ADMS (Atmospheric Dispersion Modelling System) models are managed by NUMTECH to perform atmospheric dispersion calculation depending on meteorological conditions and emission release (Brocheton et al. 2008; Carruthers et al. 2011). Two applications are performed. The first one (industrial case) aims to forecast SO_2 impact at ground from industrial release to prevent pollution peak, while the second one (urban case) has been developed to forecast at high-scale NO_2 and PM concentration over a full city (Paris).

The Extreme Rainfall Detection System (ERDS - http://erds.ithacaweb.org), developed and implemented by ITHACA, is an early warning system for the monitoring and forecasting of rainfall events, with a nearly global spatial coverage (Mazzoglio et al. 2019a, 2019b, 2019c). More importantly, the system is able to provide alerts about heavy rainfall events using both near real-time measurements and rainfall forecasts. The information is accessible through a WebGIS application, developed in an Open Source environment.

8.3 Observational data

The Weather and Climate pilot considers different observations such as authoritative and personal weather stations as well as meteorological radar data to be assimilated in the WRF model and used for validation purposes.

Concerning authoritative weather stations, namely Italian Civil Protection Department (ICPD) dataset, ICPD is designing and managing in real-time risk reduction actions over the national territory, determined by high-impact adverse weather, through its Centro Funzionale Centrale (Central Functional Center), and coordinating a federated national early warning system, in collaboration with regional authorities. In this framework, CIMA archives and curates, on behalf of ICPD, a large number of in-situ authoritative weather stations: 6059 rain gauges, 2299 hydrometers, 4373 thermometers, 1270 barometers, about 2500 anemometers, and 2683 hygrometers (Parodi et al. 2021).

CIMA has also developed a partnership with IBM to obtain real-time measurements (hourly temperature, wind, rainfall, relative humidity and pressure) acquired by about 150000 personal weather stations (PWS) over the globe (32000 from Europe, Figure 8.2). In addition, historical data acquired by about 13000 personal weather stations over Europe are available for the time period 1 June - 30 November 2018.

Two different radar meteorological datasets are available. Reflectivity radar CAPPI (500 m vertical resolution over 500-12000 m, 2.5 x 2.5 km grid spacing) are available over France for the period 1 June - 30 November 2018. Furthermore, for the years 2018, 2019, and 2020 reflectivity radar CAPPI (2000-3000-5000 m, 1 km x 1 km grid spacing) is available over Italy.

8.4 LEXIS Distributed Data Infrastructure and Weather and Climate Data API

The LEXIS DDI is based on an iRODS (https://irods.org) data management system with fault-tolerant setup. One of the key strengths of iRODS is represented by its seamless integration of the services B2SAFE, B2STAGE and B2HANDLE of the European Collaborative Data Infrastructure EUDAT (Parodi et al. 2021). The LEXIS iRODS back-end is structured in federated zones (one per supercomputing centre, namely LRZ and IT4I) in order to ensure for the users the data accessibility independently of its actual location as well as automatic replication and migration of data on an as-needed basis.

LEXIS has designed and implemented a rich set of modern JSON-based REST APIs as a frontend to the iRODS backend to facilitate upload, download and deletion of datasets, managing access rights and staging of data sets into LEXIS DCI. Datasets (e.g., weather forecast outputs) include relevant metadata, so that data can be findable according to FAIR principles (Wilkinson et al 2016). The data volumes involved in LEXIS is managed by means of asynchronous transfers over the LEXIS DDI: an appropriate queuing system allows the user to follow the progress of their requests. The DDI security is ensured by means of the global LEXIS Authorization and Authentication (AAI) infrastructure, which is based on Keycloak4 instances, and it interfaces to via OpenID3. The setup uses token hashing, as a security extra layer, to allow iRODS and Keycloak to interact (Garcia-Hernandez and Golasowski 2020).

However, the LEXIS DDI does not exist in a vacuum and independently from other data sources requested by the different workflows: some data sets required for the Weather and Climate Large-scale Pilot have been stored (and continue to be stored) in previously-developed domain-specific storage libraries, as for example datasets at ECMWF and at CIMA Research Foundation. These systems can often provide a more feature-rich view of the data sets, due to their domain knowledge.

The LEXIS project is implementing the “Weather and Climate Data API” (WCDA) in order to deliver a state-of-the-art domain-specific storage library for curated weather and climate data (Parodi et al 2021). WCDA stores and organize weather observations from a variety of sources (including in-situ unstructured observations at the European level), as well as numerical weather prediction outputs and intermediate weather data. Data are indexed according to domain-specific metadata, to efficiently support metadata-based queries. Additionally, WCDA has been designed to provide efficient distributed access to ECMWF’s MARS, that can be considered, according to our knowledge, the largest European meteorological archive. MARS stores more than 300 PB of meteorological and climatological data, from observations to global model outputs. WCDA, in addition to MARS, utilizes a FDB (Fields DataBase) for storing, indexing and retrieving GRIB data. The FDB is an internally provided service used as part of ECMWF’s weather forecasting software stack: it operates as a domain-specific object store, designed to store, index and serve meteorological fields produced by ECMWF’s forecast model and able to support different storage systems (parallel FS, Ceph cluster). The FDB serves as a ‘hot-object’ cache inside ECMWF’s high-performance computing facility (HPCF) for accessing MARS Archive. Each instance of WCDA can provide seamless access to local and remote data by contacting other WCDA instances available on the LEXIS platform (Parodi et al 2021). The WCDA interface is RESTful and the backend is based on a fully scalable architecture with containerized components. Moreover, it can be deployed with Docker Compose or in a Kubernetes cluster. The Figure 8.4 represents the WCDA instance deployed at ECMWF.

Around 80% of MARS requests are served directly from the FDB, typically for very recently produced data (Gogolenko et al 2020). A subset of this data is later re-aggregated and archived into the permanent archive for long-term availability. Usage of the FDB allows the WCDA to meet the requirements of data sizes.

The current release of WCDA, based on FDB approach, is designed to efficiently handle global NWP model outputs, encoded in GRIB file format. GRIB (General Regularly-distributed Information in Binary form) is a concise data format widely used in meteorology to store weather data standardized by the World Meteorological Organization's Commission for Basic Systems. GRIB files are a collection of self-contained records of 2D data (GRIB messages). The individual records stand alone as meaningful data, with no references to other records or to an overall schema. Each GRIB record has two components: the first one is the part that describes the record (i.e., the header) while the second part is the actual binary data itself. The data in GRIB-1 are typically converted to integers using scale and offset, and then bit-packed while with GRIB-2 the possibility of compression is available.

Recent versions of FDB, and thus of the WCDA, have been extended to handle observational data encoded in Observation DataBase (ODB) format, a World Meteorological Organization's standard for meteorological observations. ODB is a file-based database-like system developed at ECMWF to store and retrieve efficiently large volumes of meteorological observational and feedback data.

To accommodate intermediate output files encoded in NetCDF format, the WCDA FDB instance has been further extended to store a selected set of fields generated by the WRF mesoscale NWP model and required by the downstream Risico and Continuum applications.

For intellectual properties rights motivations, the ICPD observational data will be served, solely for LEXIS project research purposes, via a set of dedicated and secured APIs directly accessing CIMA Foundation databases, these data will be transformed into input data necessary for the WRF data assimilation module. The Weather and Climate Large-scale pilot has thus two data infrastructures it can rely on – the specialised WCDA for efficient handling of meteorological data, and the DDI which facilitates data exchange as well as general-purpose sharing and publications of results. The usage of both systems will be combined for maximum efficiency.

8.5 LEXIS Orchestration System

The LEXIS platform tightly couples and federates multiple heterogeneous resources to facilitate workflows mixing HPC, IaaS-Cloud and Big Data (BD) requirements. The workflow orchestration is built on a flexible orchestration solution, Ystia (https://ystia.github.io/), developed by Atos, which combines a front-end system, Alien4Cloud (https://alien4cloud.github.io), and an orchestration engine, Yorc (Ystia orchestrator, https://github.com/ystia/yorc).

Applications and workflows are modelled using the TOSCA (Topology and Orchestration Specification for Cloud Applications, https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=tosca), an OASIS consortium standard language to describe an application made of components, with their relationships, requirements, capabilities, and operations (Atrey et al 2015). The TOSCA description of an application includes its life cycle support, as deployment workflow, execution workflow, etc.. The front-end Alien4Cloud, provides a studio allowing to create applications from an extensible catalog of TOSCA components, deploy these applications, run and monitor workflows through its back-end Yorc. Yorc provides the ability to allocate computing resources on Iaas-Cloud Infrastructures, to manage applications lifecycle on these compute instances, and to submit/run jobs on HPC infrastructures. A Yorc plugin was developed in LEXIS so that the orchestrator can access LEXIS HPC infrastructures through IT4Innovation's HPC-as-a-Service framework HEAppE (High-End Application Execution Middleware; https://heappe.eu; Svaton et al. 2019). HEAppE middleware allows to run complex calculations on HPC infrastructure via the user interface of a client application, without the necessity to connect directly to the HPC cluster. HEAppE is able to provide information about the status of submitted jobs and to ensure data transfer between HPC infrastructure and a client-side app. HEAppE also provides a mapping of LEXIS user identities to the user accounts on HPC systems. The HPC-centre account to which the LEXIS user is mapped is a particular account (functional account usually not personalized) associated with an HPC project Id as per approval of the project's principal investigator (PI). For this mapping and the HEAppE mechanism to work, the project PI does not have any affiliation with LEXIS whatsoever.

The LEXIS Orchestration System can dynamically select available and appropriate computing resources from all LEXIS sites for each modelling task of the Weather and Climate Large-scale Pilot workflow (Parodi et al 2021). This approach allows each of the modelling tasks to execute as fast as possible, while the integration with different European computing centers increases the redundancy, allowing for a reliable execution. The Weather and Climate Large-scale Pilot workflow implementation includes a set of tasks for Cloud as well as HPC systems and is thus ideally suited for execution on the hybrid platform (Parodi et al 2021). The proposed orchestration system solution features an open and user-friendly design that allows to maximise the compatibility, minimizing the effort for porting workflows to the LEXIS platform and facilitating a future integration of more computing and data centers with LEXIS. Indeed, the LEXIS orchestration solution, based on TOSCA, allows to define portable workflows (templates) that can be easily customized; any additional computing or data resources integrated in the LEXIS platform will automatically be taken into account for deploying such workflows.

8.6 Weather and Climate Pilot Workflows

As mentioned in the Introduction, the LEXIS Weather and Climate Large-scale Pilot includes modelling tasks from weather prediction to hydrological prediction, forest fire risk forecast, air quality and industrial pollution forecasting, as well as extreme rainfall identification.

The hydrological prediction workflow involves the WRF model, including a WRFDA data assimilation system, and the fully distributed hydrological model Continuum (Figure 8.5), which have already demonstrated their combined potential in research activities (see e.g., Lagasio et al 2019b). The WRF model is executed on HPC facilities at IT4I and LRZ after the preparation of initial and boundary conditions provided by the WPS (WRF Preprocessing System). The WPS task is executed, as a container, on cloud computing facilities (at LRZ and IT4I) and it allows processing both ECMWF (European Centre for Medium-Range Weather Forecasts) IFS (Integrated Forecasting System) and NCEP (National Centers for Environmental Prediction) GFS (Global Forecast System) global circulation model data to generate input fields for the WRF model (Parodi et al 2021).

The WRFDA task is a flexible, state-of-the-art atmospheric data assimilation system that is portable and efficient on available parallel computing platforms. WRFDA is a task executed on cloud computing facilities while the Continuum hydrological model is executed as sequential tasks (container) on cloud computing facilities. Both the models are executed using IT4I and LRZ HPC capabilities. Hydrological prediction workflow results are published on the MyDewetra platform (https://www.mydewetra.org), a web-based real-time system for hydro-meteorological forecasting and monitoring developed by CIMA on behalf of the Italian Civil Protection Department (ICP Department et al. 2014).

The complete forest fire risk prediction workflow involves the aforementioned meteorological task WRF, including the WRFDA data assimilation system, and the fire risk model RISICO in place of the Continuum model. Also, the complete air quality prediction workflow involves the aforementioned meteorological task WRF, including the WRFDA data assimilation component, and the ADMS, in place of the Continuum model in the hydrological workflow.

For heavy rainfall detection purposes, available capabilities of ITHACA ERDS include the analysis of both the near real-time and the forecast rainfall amounts for different lead times, with the aim to deliver extreme rainfall alerts. NASA/JAXA GPM (Global Precipitation Measurement) IMERG (Integrated Multi-satellitE Retrievals for GPM) Early run data is downloaded every hour to provide near real-time rainfall measurements over the past 12, 24, 48, 72 and 96 hours. GPM IMERG Early run data is characterized by a temporal resolution of 30 minutes, a 0.1 x 0.1 degrees spatial resolution, a spatial coverage between 90 degrees North and 90 degrees South and a ~4 hours latency (Huffman et al., 2020). GFS data are instead used as a source of global-scale rainfall forecasts, to provide longer lead-time information (up to 4 days) with a 0.25 x 0.25 degrees spatial resolution, updated every 12 hours (Mazzoglio et al., 2019c). Despite the good results achieved during the validation analysis (Mazzoglio et al., 2019a), the need of using more accurate rainfall forecasts emerges. WRF data produced by

CIMA in the framework of the LEXIS project are therefore included to provide 48-hours forecasts at 7.5 km over Europe and 2.5 km over Italy (including radar data assimilation over Italy), and some examples are described hereafter.

8.6.1 WRF-ERDS Workflow examples

The extreme rainfall detection methodology used by ERDS is based on the concept of threshold: a threshold represents the amount of precipitation needed to trigger a flood event induced by extreme rainfall. Specifically, if for a selected interval the accumulated precipitation exceeds the threshold value, an alert is provided. This set of thresholds has been calculated for every aggregation interval by using values equal to a percentage of the mean annual precipitation that affects each place of Earth's surface (Mazzoglio et al. 2019b). In other words, threshold values increase as the aggregation interval increases. Moreover, this matrix of threshold values has been calibrated on the basis of the input data to take into account possible under/overestimations of the dataset. Threshold values applied to the rainfall depths retrieved by GPM data are therefore different from those applied to GFS data or to WRF data. The system issues alerts in the form of a georeferenced raster map, allowing to obtain precise information of the locations affected by significant rainfall. Alerted areas can be therefore exploited for the definition of specific Areas of Interest, to be used for retrieving information about the affected population or for mapping purposes.

The entire WRF-ERDS workflow has been tested over different heavy rainfall events that affected Italy during 2020, both in the case of convective and stratiform rainfall events. In this chapter, the results related to two case studies are reported.

The first case study is related to a heavy rainfall event that affected Tuscany during the 4th June 2020 (Figure 8.6). More than 100 mm of rainfall was recorded in the northern part of

Tuscany (Figure 8.6a), corresponding to an estimated return period of about 200 years (Centro Funzionale della Regione Toscana 2020) in a very small area and > 50 years in a larger area (Figure 8.6b). Despite the slight underestimation of the rainfall depth (Figure 8.6c), WRF model was able to properly identify the most affected areas (Figure 8.6d). Thanks to the WRF-based analysis, information about the locations that would be affected by heavy rainfall was available in the early morning of the 4th of June, several hours before the event.

An intense convective event affected the city of Palermo (Sicily, South of Italy) during the afternoon of the 15th July 2020. According to SIAS (Servizio Informativo Agrometeorologico Siciliano) more than 130 mm of rainfall was recorded in about 2.5 hours, causing urban flooding phenomena and damages. ERDS was not able to detect the event using GFS data due to severe underestimation of the forecast. A WRF modelling experiment based on three nested domains (with 22.5, 7.5 and 2.5 km grid spacing), innermost over Italy, was executed by assimilating the national radar reflectivity mosaic and in situ weather stations from the Italian Civil Protection Department. Good results were achieved using WRF data at a 2.5 km resolution: a peak rainfall depth of about 35 mm in 1 hour and 55 mm in 3 hours were predicted roughly 30 km far from Palermo (Figure 8.7).

8.7 Conclusion

The chapter describes the first results achieved with the framework of LEXIS Weather and Climate Large-Scale Pilot. LEXIS Distributed Data Infrastructure and Weather and Climate Data API are described in Section 8.4 while LEXIS Orchestration System is described in Section 8.5. All the Weather and Climate Pilot Workflows are described in Section 8.6, together with a specific focus on WRF-ERDS workflow. Preliminary results obtained with the WRF-ERDS workflow highlight that improved rainfall forecasts obtained by using HPC resources significantly increase the performance of an early warning system as ERDS. Global-scale low-resolution rainfall dataset as the GFS one are often characterized by poor performances, especially in the case of very intense and localized convective rainfall that occurs in the summer season. Further experiments will be performed to assimilate atmospheric data from personal weather stations over Europe to increase WRF accuracy.

Acknowledgement

The authors acknowledge the use of imagery from the NASA Worldview application (https://worldview.earthdata.nasa.gov/), part of the NASA Earth Observing System Data and Information System (EOSDIS). This chapter was supported by the LEXIS project, funded by the EU's Horizon 2020 research and innovation programme (2014-2020) under grant agreement No 825532.

References

1. Atrey, A., Moens, H., Seghbroeck, G., Volckaert, B., Turck, F. 2015. An overview of the OASIS TOSCA standard: topology and orchestration specification for cloud applications. Technical Report, IBCN-iMinds, Department of Information Technology.
2. Brocheton, F., Armand, P., Soulhac, L., Buisson, E. 2008. A methodology to characterise the sources of uncertainties in atmospheric transport modelling. Hrvatski meteorološki časopis 43(43/1), 78–82.
3. Carruthers, D., Seaton, M., McHugh, C., Sheng, X., Solazzo, E., Vanvyve, E. 2011. Comparison of the complex terrain algorithms incorporated into two commonly used local-scale air pollution dispersion models (ADMS and AERMOD) using a hybrid model. Journal of the Air & Waste Management Association 61(11), 1227–1235. https://doi.org/10.1080/10473289.2011.609750
4. Centro Funzionale della Regione Toscana 2020. Disaster report 4-5 June 2020. Available online: https://www.cfr.toscana.it/supports/download/eventi/report_evento_04-

05_giugno_2020.pdf (accessed March 14, 2021).
5. Fiorucci, P., Gaetani, F., Minciardi, R. 2008. Development and application of a system for dynamic wildfire risk assessment in Italy. Environmental Modelling & Software 23(6), 690-702. https://doi.org/10.1016/j.envsoft.2007.05.008
6. Garcia-Hernandez, R.J., Golasowski, M. 2020. Supporting Keycloak in iRODS systems with OpenID authentication. Presented at the CS3 2020 - Workshop on Cloud Storage Synchronization and Sharing Services.
7. Gogolenko S. et al. 2020. Towards Accurate Simulation of Global Challenges on Data Centers Infrastructures via Coupling of Models and Data Sources. In: Krzhizhanovskaya V. et al. (eds) Computational Science – ICCS 2020. ICCS 2020. Lecture Notes in Computer Science, vol 12142. Springer, Cham. https://doi.org/10.1007/978-3-030-50433-5_32
8. Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Tan, J., Xie, P. 2020. NASA Global Precipitation Measurement (GPM) Integrat-ed Multi-satellitE Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) Version 06. https://gpm.nasa.gov/sites/default/files/2020-05/IMERG_ATBD_V06.3.pdf (accessed April 14, 2021).
9. ICP Department, CIMA Research Foundation 2014. The Dewetra platform: a multi-perspective architecture for risk management during emergencies. In: Proceedings of the First International Conference Information Systems for Crisis Response and Management in Mediterranean Countries (ISCRAM-med 2014), Toulouse, France, 15-17 October 2014, vol. 1, pp. 165–177. Springer.
10. Lagasio, M., Parodi, A., Pulvirenti, L., Meroni, A.N., Boni, G., Pierdicca, N., Marzano, F.S., Luini, L., Venuti, G., Realini, E., et al. 2019a. A synergistic use of a high-resolution numerical weather prediction model and high-resolution earth observation products to improve precipitation forecast. Remote Sensing 11(20), 2387.
11. Lagasio, M., Silvestro, F., Campo, L., Parodi, A. 2019b. Predictive capability of a high-resolution hydrometeorological forecasting framework coupling WRF cycling 3dvar and continuum. Journal of Hydrometeorology 20(7), 1307–1337.
12. Mazzoglio, P, Laio F, Sandu C, Boccardo P. 2019a. Assessment of an Extreme Rainfall Detection System for Flood Prediction over Queensland (Australia). Proceedings 18(1):1. https://doi.org/10.3390/ECRS-3-06187
13. Mazzoglio, P., Laio, F., Balbo, S., Boccardo, P., Disabato, F. 2019b. Improving an Extreme Rain-fall Detection System with GPM IMERG data. Remote Sensing 11(6), 677–677. https://doi.org/10.3390/rs11060677
14. Mazzoglio P., Laio F., Balbo S., Boccardo P. 2019c. ERDS: an Extreme Rainfall Detection System based on both near real-time and forecast rainfall measurements. Annual of the University of Architecture, Civil Engineering and Geodesy (Sofia), 52, Issue S1, 1423-1433.
https://uacg.bg/UserFiles/File/UACEG_Annual/2019/%D0%91%D1%80%D0%BE%D0%B9%20S1/19-3.pdf
15. Parodi, A., Danovaro, E., Hawkes, J., Quintino, T., Lagasio, M., Delogu, F. et al. 2021. LEXIS Weather and Climate Large-Scale Pilot. In: Barolli L., Poniszewska-Maranda A., Enokido T. (eds) Complex, Intelligent and Software Intensive Systems. CISIS 2020. Advances in Intelligent Systems and Computing, vol 1194. Springer, Cham. https://doi.org/10.1007/978-3-030-50454-0_25
16. Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., et al. 2017. The weather research and forecasting model: Overview, system efforts, and future

directions. Bulletin of the American Meteorological Society 98(8), 1717-1737. https://doi.org/10.1175/BAMS-D-15-00308.1
17. Silvestro, F., Gabellani, S., Rudari, R., Delogu, F., Laiolo, P., Boni, G. 2015. Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data. Hydrology and Earth System Sciences 19(4), 1727-1751. https://doi.org/10.5194/hess-19-1727-2015
18. Svaton, V., Martinovic, J., Krenek, J., Esch, T., Tomancak, P. 2019. HPC-as-a-Service via HEAppE Platform. In: Conference on Complex, Intelligent, and Software Intensive Systems, pp. 290–293. Springer.
19. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3, 160018. https://doi.org/10.1038/sdata.2016.18

Figure legend

Figure 8.1: LEXIS Weather and Climate complex workflows.

Figure 8.2: Weather Underground PWS on MyDewetra platform (courtesy of ICPD).

Figure 8.3: Radar data mosaic over Italy published on MyDewetra platform (courtesy of ICPD).

Figure 8.4: WCDA at ECMWF.

Figure 8.5: WRF-Continuum workflow.

Figure 8.6: 24-hours rainfall depth recorded by rain gauges during 4th June 2020 (figure a). 24-hours rainfall return period (figure b). 24-hours rainfall forecast provided by WRF model during 4th June 2020 (figure c). Heavy rainfall alerts provided by ERDS using WRF data as input (figure d). Contains rain gauge and return period information retrieved by Centro Funzionale della Regione Toscana (2020).

Figure 8.7: WRF 24-hours forecast at 2.5 km resolution (a) and heavy rainfall alerts (b) provided by ERDS using WRF data as input.