Summary

The number and intensity of severe weather events is increasing, leading to loss of goods, property, and human lives. Improved weather forecasting, especially for severe weather events expected less than 6 hours in the future, also known as *nowcasting*, is expected to help mitigate the outcome of such events. The WeaMyL project aims to improve nowcasting accuracy through deep learning methods and Big Data approaches able to manage the large volume of meteorological data that is constantly being produced. National Meteorological Institutes are the project's direct beneficiaries, while the public at large is the most important, indirect beneficiary. Both Romanian and Norwegian populations will benefit by having more time to prepare and, with a reduced risk of false alarms, more confidence in meteorological alerts.

The project team includes researchers from the Babeș-Bolyai University (BBU), who contribute their expertise in machine learning. Meteorology experts from the Romanian National Meteorological Administration (NMA) provide the data and interpretation for Romania, while their colleagues at the Norwegian Meteorological Institute (MET) do the same for Norway. The Norwegian team includes software development specialists (MET-IT subteam) who are responsible for the development of the WeaMyL platform front-end, the integration layer between weather data and the software platform, as well as integration with national warning systems. The MET team also includes a team of meteorology operatives (MET-MT) who will extensively test and evaluate WeaMyL.

The main objectives set for 2020 and 2021 were completely fulfilled. The main goals for the project's startup phase were to study existing approaches and solutions in nowcasting and to define WeaMyL's requirements and general architecture. The 2021 project phase continued the technical work with major focus on development, analysis, and validation of specially tailored machine learning models.

Scientific and Technical Achievements

The first project phase took place in 2020 and was entitled ***Documentation, system requirements and architecture***. Its goals included conducting an analysis of the existing research literature, identifying the limitations of existing approaches and solutions in nowcasting, defining the functional and nonfunctional requirements for WeaMyL and establishing the platform's architecture. This included establishing user requirements as well as the main conditions for the system's proper functioning. The second project phase took place during 2021, and was entitled ***Machine learning models for weather nowcasting***. The second phase continued the work carried out in 2020, and added new objectives, including the definition of a theoretical model for nowcasting, developing specially tailored scalable ML models for accurate nowcasting, together with scientifically validating the developed models using experimental result analysis and interpretation. All planned activities were carried out and completed successfully within existing time and budgetary constraints.

In the following we describe the work that was carried out by project partners for fulfilling the scientific and technical objectives during 2020 and 2021.

The **activities carried out by the BBU team** are summarized as follows. During 2020, team members reviewed the existing machine learning (ML) methods for nowcasting, data models and architectures for meteorological systems. During the review, drawbacks of existing approaches were identified as a starting point for offering solutions that can overcome observed limitations. With this input, the BBU

team defined and structured the functional requirements considering the literature review together with the end-user requirements provided by the meteorological partners. System usage was characterized through a set of user scenarios that will be used for system development. Next, WeaMyL's main modules and interfaces were specified; this included defining a general, theoretical model for nowcasting based on relevant data sources as well as the project's end-user requirements. The last activity of 2020 saw the BBU team elaborate specifications for the data to be integrated into the system and the functional requirements related to data access.

The BBU team's activities for 2021 started with incrementally improving the system design and general architecture of WeaMyL. This included several activities that were carried out in parallel. The work started by investigating the developed computational models from an unsupervised learning perspective, as well as from classification and regression based supervised learning. Next, the team investigated raw meteorological data preprocessing techniques such as data preparation for training, detection of possible outliers as well as identification of meteorological observations and products relevant from a computational viewpoint. Additionally, this involved investigating techniques for reducing data dimensionality in order to cope with the curse of dimensionality problem. The first developed methods were the unsupervised learning methods that were aimed towards detecting patterns in the preprocessed meteorological and geographical data. These were then applied for visualization purposes and for providing some indications about the prediction task. Next, the BBU team investigated the application of supervised learning methods as well as the development of machine learning models for nowcasting based on the Xception Deep Learning architecture. The validation of the Xception model was started in September 2021. As part of the upcoming project phase, the BBU team envisions carrying out an extensive experimental evaluation of the Xception model that will include comparisons and statistical analyses of the obtained results.

In parallel with the development of ML methods, the data that must be integrated into the system and the functional requirements related to data access were specified. Towards the end of 2021, the BBU team offered support to MET-IT regarding the integration of the Xception deep learning module into the initial prototype of the forecasting platform. In addition, team members contributed to integrating historical and real-time data obtained from the meteorological partners into the Xception deep learning module of the forecasting platform.

The **NMA team's activities for 2020 can be summarized as follows**. First, the team reviewed established and ML based techniques and technologies used in nowcasting. This included current and emerging solutions that support meteorologists in determining and issuing severe weather warnings. The review process was followed up by identifying which of the available data sources were relevant for nowcasting. In addition, a structured documentation of the data types and sources currently used in weather nowcasting was provided. Next, team members provided end-user and non-functional requirements related to the WeaMyL platform and defined user stories relevant for nowcasting operational activities. Support was provided to BBU and MET-MT during the development of the nowcasting models by providing relevant data sets and meteorological expertise specific to the geographical area.

NMA activities in 2021 started by providing continuous feedback regarding the computational modeling of the nowcasting problem as a classification/regression task, as well as contributing to incrementally improving the computational models. This continued by providing meteorological insight concerning the data used and the results of the experiments conducted using the Xception model, both accuracy and performance-wise. Furthermore, NMA contributed to the development of the Annotated Atlas by providing continuous assistance regarding the integration of relevant radar, satellite and ground

observation data into the initial prototype of the data bank. The NMA team members provided feedback regarding the testing of the initial prototype of the Atlas.

Additionally, NMA contributed by providing access to its database of historic data. Radar data was extracted, annotated, validated and integrated in the initial prototype of the Annotated Atlas. Starting September 2021, the current prototype was deployed at NMA, where it will remain in use for an undetermined period of time. WeaMyL will be integrated into the Romanian severe weather warning system while the Annotated Atlas will be used in the nowcasting process for forecasts more than one hour in advance.

The work carried out by the MET teams (MT and IT subteams) is summarized in the following. At project onset, MET-IT conducted a review of existing software solutions for weather forecasting and particularly nowcasting. This was followed by defining the data sources used by different ML algorithms for nowcasting purposes (IT subteam). At the same time, the non-functional requirements for WeaMyL and security issues regarding the system's proper functioning were structured. Quantitative and qualitative measures were defined to validate user requirements and correctly identify when these were achieved based on the pilot deployments. Then, the general architecture of the system was established using the previously identified functional and nonfunctional requirements**.** Members of the MT subteam contributed to developing conceptual models of meteorological processes, in particular severe weather. During this time, the IT subteam implemented a data acquisition component to prepare the platform's meteorological data sources. This component will provide the necessary data sets further needed for training the ML algorithms included in the Forecasting Platform.

Activities of 2021 were started with the IT subteam defining the basic concepts and relationships that govern the software platform's infrastructure. In parallel, the IT team started to incrementally include the developed computational models into the software, while the MT subteam processed the raw historical meteorological data in order to bring it to a suitable form for further analyses. The MT subteam also assisted BBU in conducting the experimental evaluation of Xception, statistical analysis of the obtained results and comparison to existing solutions. Meteorological feedback was provided regarding the performance of the Xception model on meteorological data held by MET. The software team from MET continued with the development and testing of the orchestration layer between WeaMyL and meteorological data sources. The data pipeline which will be further tested and maintained during the deployment of the data acquisition component was defined and set up. In April 2021, the IT team started the development and testing of the first Annotated Atlas of Meteorological Observations, mainly intended for early feedback purposes. The incremental development, planned to result in three prototype versions – initial, advanced and final of the Forecasting Platform started in August 2021. The team worked on integrating the nowcasting module by implementing the Xception deep learning model into the Forecasting Platform. The Data acquisition component was improved according to feedback received from the meteorological teams; it also integrated the meteorological data sets required to build the initial prototype of the Annotated Atlas semantic data bank. Starting from September 2021, the system prototype was deployed at MET, where it is planned to be used for an undetermined period of time. The WeaMyL platform is expected to be integrated into the Norwegian severe weather warning system while the Annotated Atlas will be used in the nowcasting process for forecasts more than one hour in advance.

Dissemination

As the project promoter, BBU was in charge of administrative, scientific and technical project coordination. This involved coordinating the communication and collaboration between partners, identifying and monitoring risks and taking corrective measures when required. The project promoter coordinated work on deliverables and relevant artefacts to ensure the project progressed in accordance with time and budgetary constraints towards meeting its defined objectives. All consortium partners worked on disseminating the results to relevant authorities and scientific organizations. In this regard, they contributed in publishing the results of the technical and scientific activities in high-quality open access journals and conference proceedings. These efforts materialized in the publication of 7 scientific papers, 5 of which in Web of Science indexed journals and 2 in Web of Science conferences.

One of the most important activities during the first year of project implementation was the **WeADL workshop** organized on May 28, 2021. Its main goal was to raise awareness in the scientific community regarding the challenges of deep learning, machine learning or broadly artificial intelligence-based scientific research in highly interdisciplinary domains. Emphasis was placed on the purposes of the current project -- employing deep learning techniques for improving meteorological nowcasting based on historical satellite, radar and other meteorological products. The event took place over Zoom, with the programme and talks recorded and made available on the website. More than 70 people registered for the event, among which more than 30 actively participated.