

EDERA Product Generation Software

Calum Baugh (ECMWF), Dimitar Tasev (ECMWF), Marc Berenguer (UPC), Shinju Park (UPC), Seppo Pulkkinen (FMI), Heikki Myllykoski (FMI)

1 Introduction and scope

The second work package of the EDERA project is focussed on developing new products related to convective and pluvial flooding and compound flooding resulting from heavy rainfall and riverine processes and also enhancing the flash flood impact products developed during the TAMIR project. Achieving each of these requires the creation of new software, or the modification of existing software, to generate these products. This milestone report documents the software that has been created in order to generate these products. The software will then be used in work package 3 to set up and run a real time demonstration of the EDERA forecast products for a 15 month period.

2 Convective flood hazard forecasts over Europe

As in TAMIR, the EDERA products include nowcasts of convective hazards and impacts. Here the TAMIR software has been extended to the pan-European scale. The extended nowcast products use the OPERA radar mosaics as the main input data source. The Finland-only product implemented in TAMIR is also being generated, because it can utilize additional data sources that are not available for the whole Europe.

The product generation software combines a cell-based convective storm nowcast ([1]) with a machine learning (ML) model. The ML model is applied for predicting storm severity level (i.e. how many damage reports it is expected to cause). It is trained against severity levels derived from the European Severe Weather Database (ESWD) weather hazard reports. The severity is estimated based on the spatial and temporal distance between the reports and the corresponding storm cells using the methodology developed in [2]. A separate ML model is trained for each impact type, which currently include large hail, heavy rain, severe wind gusts and damaging lightning strikes. In the Finnish product, the hazard levels are estimated from weather-related emergency calls, as in TAMIR, so only one type of model is trained.

A schematic overview of the product generation is shown in Figure 1. The steps shown in the diagram are done each time a new set of products is generated from the most recently available data. The first step is identification of convective cells from the radar reflectivity mosaics. Then each storm is assigned a set of raster attributes. These are calculated as the minimum, mean and maximum values inside each storm cell for auxiliary raster data (such as lightning flash density). Then the current storm cells are matched to those from the previous product generation time to produce a set of continuous storm tracks. This is followed by calculation of temporal attributes from the storm tracks (e.g. cell age) and applying a Kalman filter to produce estimates of storm velocities ([1]). Severity levels are then assigned for each storm cell by using the ML model. This is done separately for each storm hazard type using the storm attributes listed in Table 1 and Table 2. The severity levels are then converted into discrete classes based on pre-selected thresholds. Finally, ellipses are fitted to each storm cell to produce a simplified representation for the product visualization. The above information is stored into a PostGIS database.

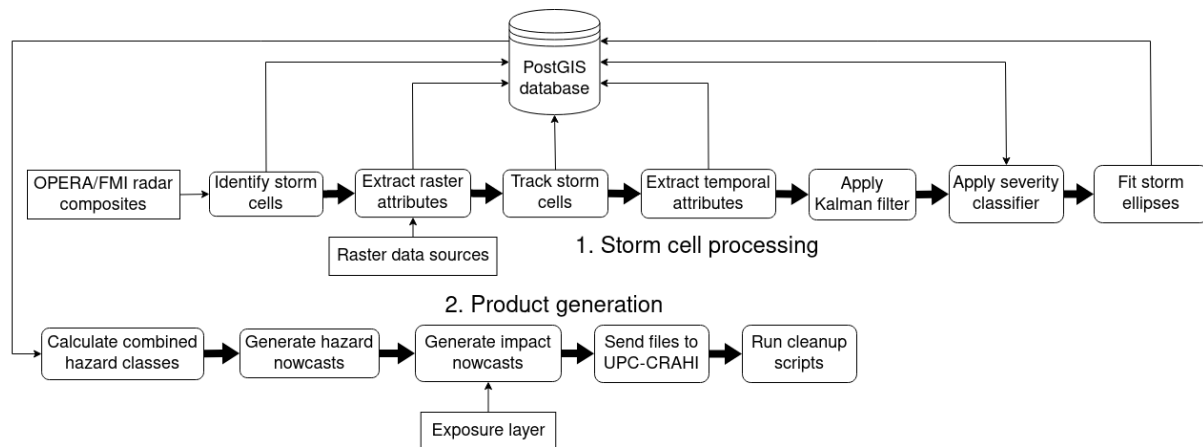


Figure 1. Overview of the EDERA convective hazard/impact nowcast generation.

In the second stage shown in Figure 1, storm cell and track information are read from the database, and hazard and impact nowcasts are generated. The products consist of GeoTIFF files containing gridded nowcasts. The hazard nowcast gives a probabilistic prediction of future storm positions together with the hazard class for each different subtype (e.g. rainfall or lightning). The current version of the EDERA platform shows the radar reflectivity (Figure 2) and only the combined hazard class (Figure 3). The classes are: no hazard (0), moderate (1), severe (2) and extreme (3) based on the selected severity thresholds. The impact nowcasts are then generated by using a separate exposure layer. In the current version, this is taken from gridded population density divided into four classes. The impact class is obtained by multiplying the hazard class (0-3) with the exposure class (0-3) in each grid pixel. This gives in total 10 distinct impact classes. Finally, the GeoTIFF files are sent to UPC-CRAHI via FTP for visualization on the EDERA platform, and a cleanup script is executed to remove old files.

Table 1. Cell/track attributes used with the ML-based model for predicting storm hazard levels.

Storm cell/track attributes	
Attribute	Description
Area	horizontal area of storm cell (km ²)
Age	storm cell age (minutes)
Speed	storm propagation speed (km/h)
Direction	storm propagation direction (degrees)

Table 2 Raster-based attributes used with the ML-based model for predicting storm hazard levels.

Raster attributes		
Quantity	Description	Data provider
Reflectivity	radar-measured reflectivity (dBZ)	The OPERA data centre (Odyssey) for Europe FMI for Finland
Accumulated rainfall	hourly rainfall accumulation (mm) calculated from the above after conversion to rain rate	
CAPE	convective available potential energy (J/Kg)	ECMWF (data obtained from the FMI SmartMet API)
CIN	convective inhibition (J/Kg)	
Flash density	gridded lightning flash density (flashes/grid cell during the last x minutes) from point observations	Vaisala Global Lightning Detection Network (GLD360) for Europe (will be used once the license has been obtained) The NORDLIS lightning network for Finland

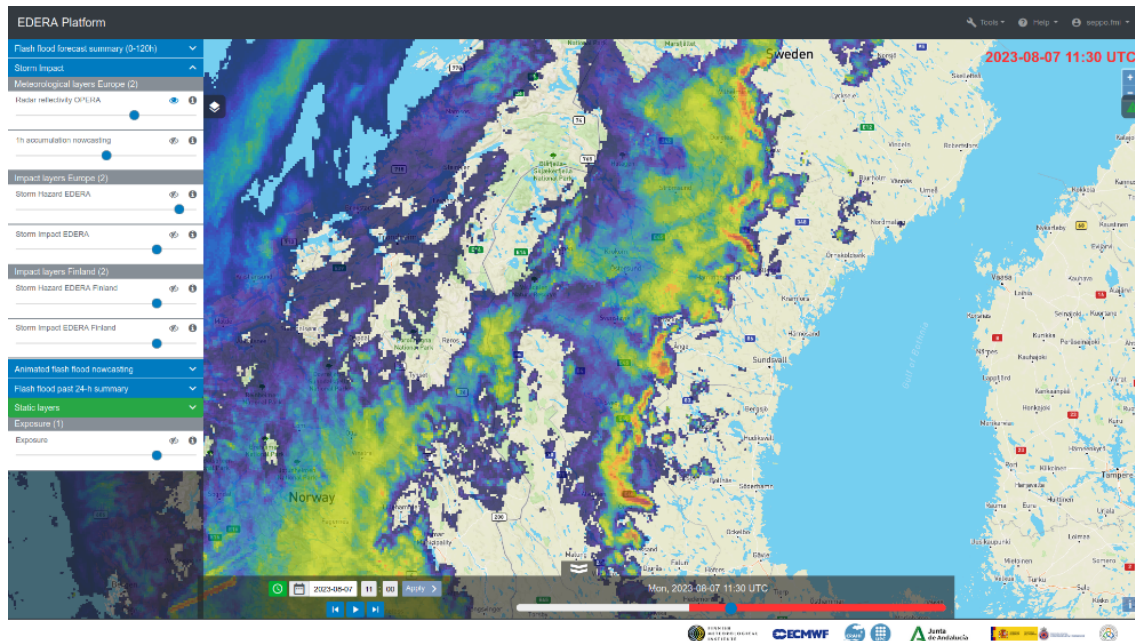


Figure 2. OPERA radar reflectivity composite shown in the EDERA platform.

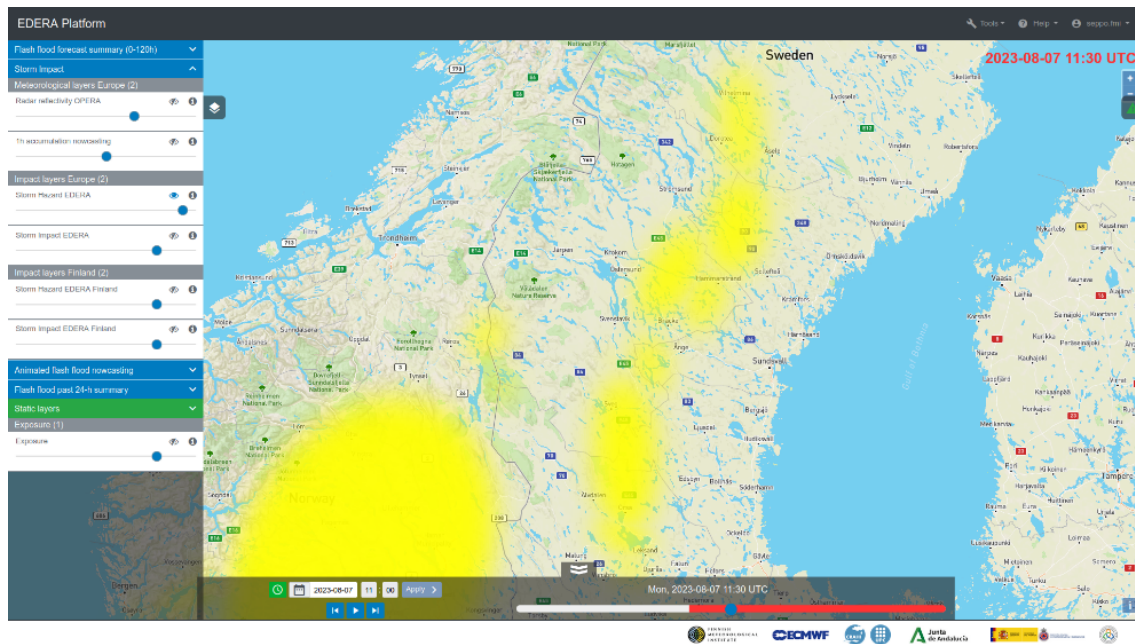


Figure 3. EDERA convective hazard nowcast generated by using the radar composite shown in Figure 2. The lead time is 30 minutes, and the yellow colour represents the overall hazard level (moderate). The coloured ellipses represent the areas potentially affected by the storms.

The product generation software is implemented in Python using the scikit-learn library (<https://scikit-learn.org>) for the ML components. Together with the real-time storm database, the software is run on a virtual server provided by the FMI. Based on the available update intervals of the radar composites, the generation is done every 15 and 5 minutes for the pan-European and the Finnish products, respectively. The maximum lead times for these products has been set to three and one hour, respectively, and their spatial resolutions are 2 km and 1 km.

3 Development of a strategy for compound heavy rainfall and flood impact forecasting

Urban flood hazard nowcasting

A simple urban flood hazard nowcasting product has been set up by assessing the return period of observed and forecasted rainfall in urban areas across Europe. The inputs used are the radar-based precipitation estimates and nowcasts (1-h and 2-km resolutions).

The assessment of the hazard of pluvial floods in urban areas has been limited to the areas with the most impermeable land cover types. Land cover information has been obtained from the CORINE Land Cover inventory, version 2018 ([3]). It has a resolution of 100 m, and the areas identified as *urban* are those with the land cover types listed in Table 1. On the 2-km grid, the pluvial floods hazard product is issued on the cells with more than 40% of urban areas.

The flood hazard (expressed in terms of a return period) is estimated by comparing the observed and forecasted 1-h accumulated precipitation with thresholds derived from 10 years of gauge-adjusted radar-based 1-h rainfall accumulations.

The pluvial floods hazard product has been integrated in real time in the EDERA production system in July 2023, and it is being displayed in the EDERA web-based platform (Figure 4).

Table 3. CLC types used for classified as *urban*.

Code	Type
1.1.1	Continuous urban fabric
1.1.2	Discontinuous urban fabric
1.2.1	Industrial or commercial units
1.2.2	Road and rail networks and associated land
1.2.3	Port areas
1.2.4	Airports

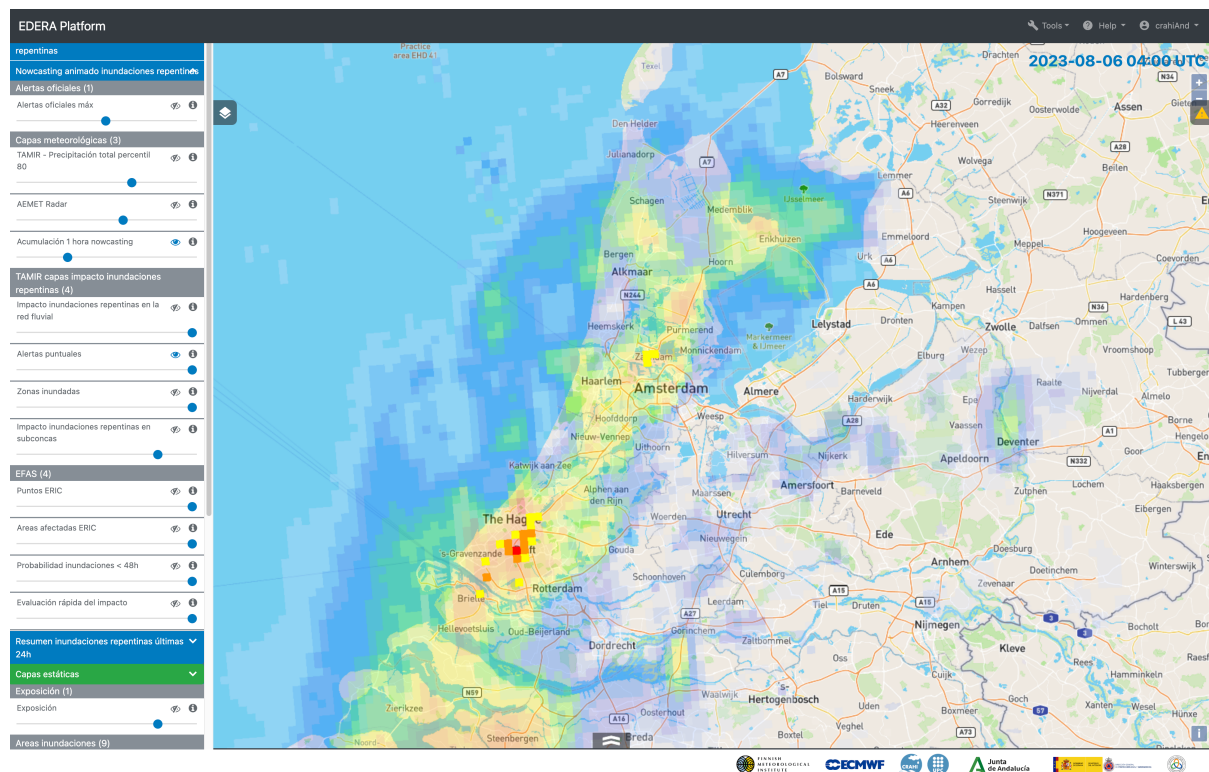


Figure 4. Example of the pluvial flood product for the intense rainfall event that affected the Netherlands on 06 August 2023.

Flash flood and river flood forecasting

The flash flood impact forecasting products used in EDERA are those developed in TAMIR, which combine the probability of exceeding the 2-years return period for flash floods with the TAMIR exposure layer through a confusion matrix ([4]). The flash flood impact forecasts have been set up for running in real time, and the data-transfer system (between the EFAS Computational Centre and UPC) has been established, which enabled their integration in the web-based EDERA platform (Figure 5).

Also, some of the river floods products from EFAS have been integrated in the EDERA platform and are displayed through the EDERA platform (Figures 6 and 7).

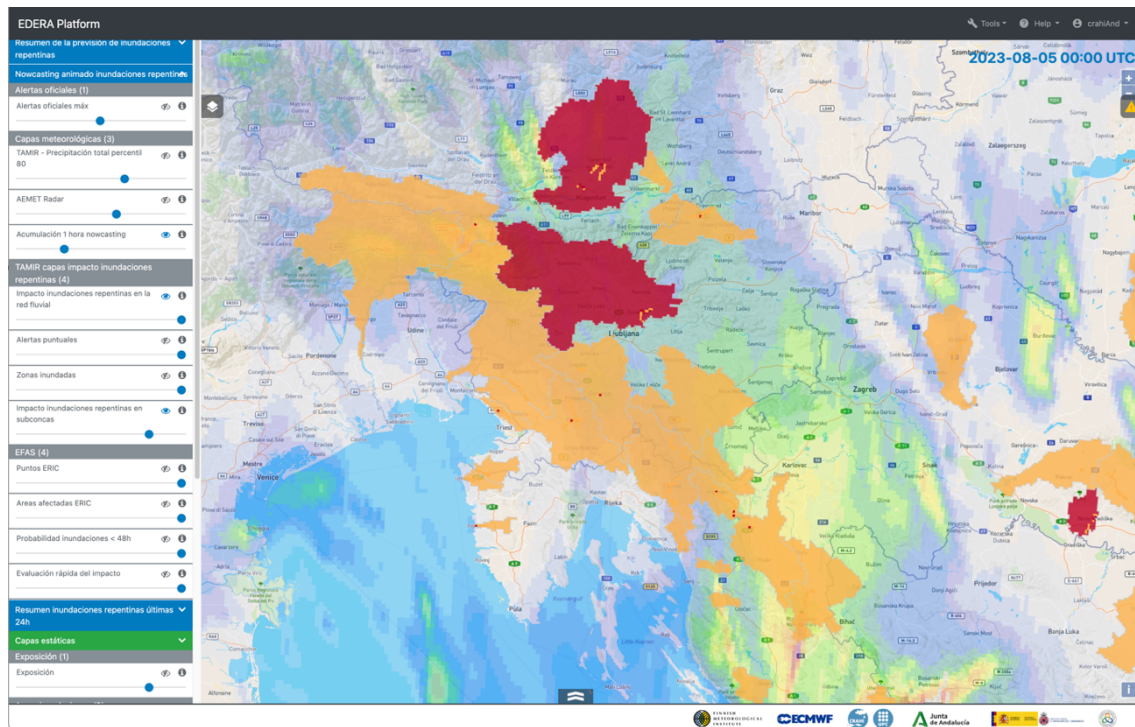


Figure 5. Flash flood impact estimated on 05 August 2023, during the floods that affected Slovenia.

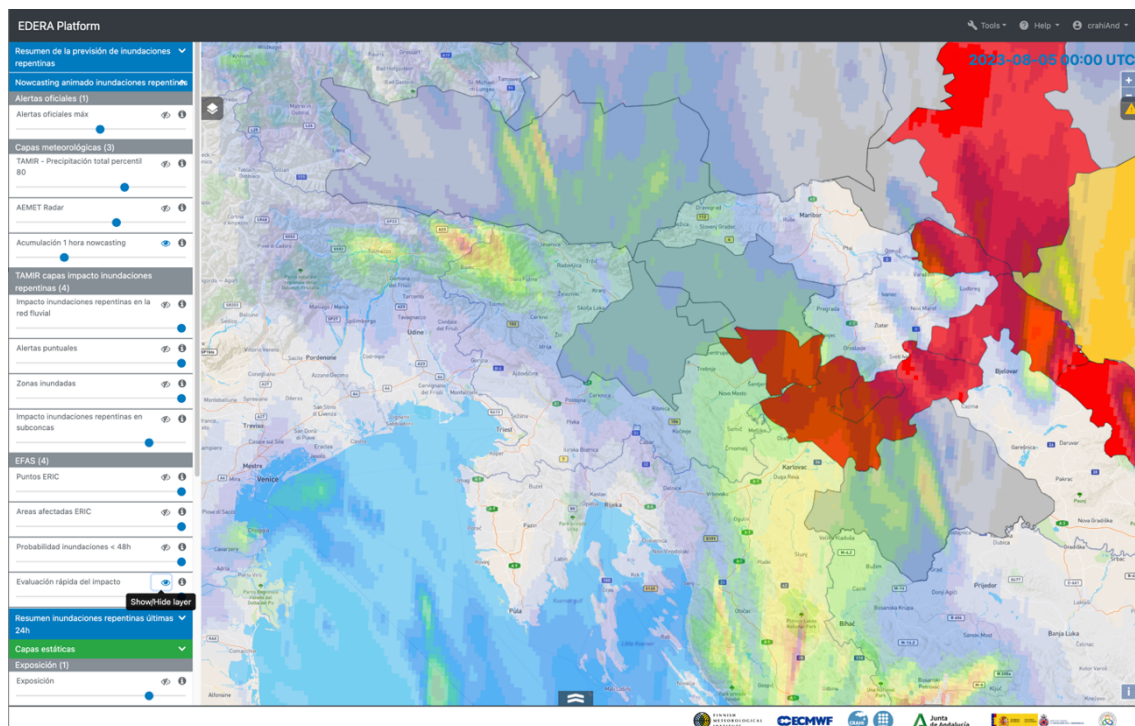


Figure 6. EFAS rapid flood impact assessment product for 05 August 2023, during the floods that affected Slovenia.

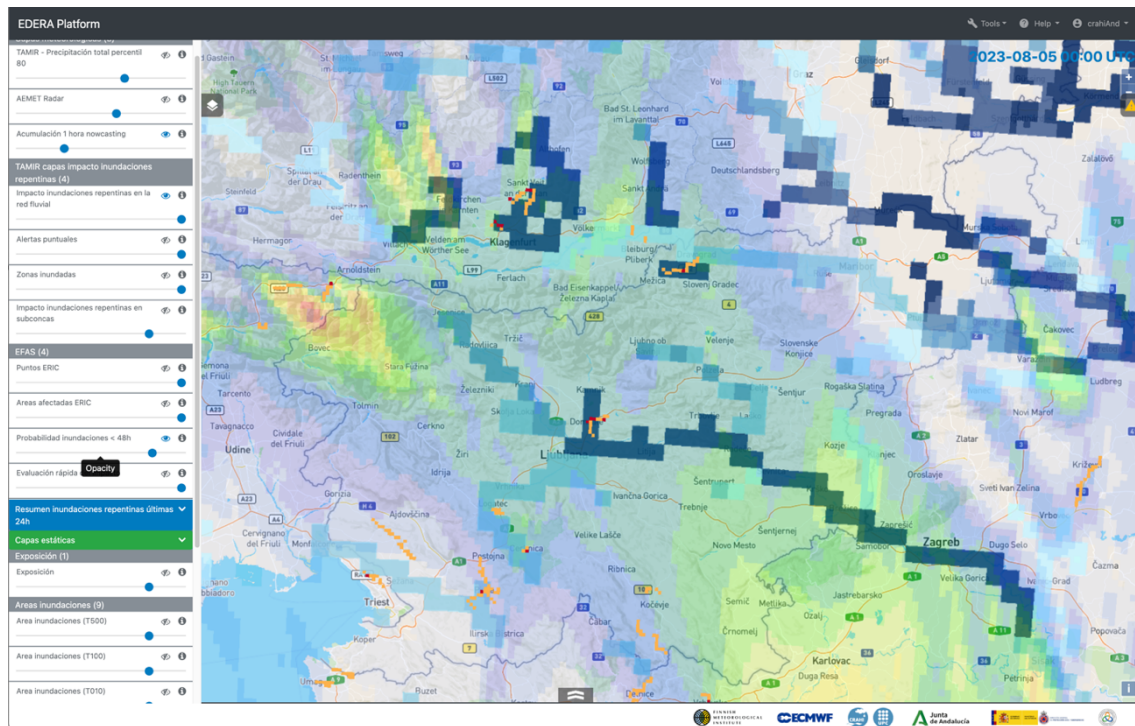


Figure 7. EFAS flood probability product (exceeding the 5-year return period) overlaid with the flash flood impact product for 05 August 2023, during the floods that affected Slovenia.

Compound floods forecasting strategy

The strategy for compound flood forecasting consists on the combination of the products for pluvial floods nowcasting, and flash floods and floods forecasting. It relies on the functionality added to the EDERA web-based platform to issue warnings adapted to the profile of the EDERA user.

The warnings for compound floods forecasting will present together the impact forecasts for the 3 types of floods summarised over NUTS-2 and NUTS-3 regions corresponding to the area of interest of each end user. In those countries for which the flood hazard and risk maps have been integrated in the EDERA processing system, the warnings will include the extent of the expected flooded area.

The different software modules involved in the compound floods forecasting strategy have been committed in UPC-CRAHI's [repository](#).

4 Enhancement of flash flood impact forecasts

Additional geographical contextual information

The flash flood impact catchment summary product, that was developed during the TAMIR project, has been enhanced to include additional information. Previously this product summarised the potential flash flood impact within each sub-catchment over four different lead time summary windows: 1) 0-6 hours, 2) 7-24 hours, 3) 25-48 hours, 4) 49-120 hours. When the product was displayed within the EFAS (European Flood Awareness System) web-viewer, a user could click on a shaded sub-catchment and a pop-out window would appear with information including; the lead

time of the event, the total population potentially affected and the number of health, education and energy facilities potentially affected (Figure 8).

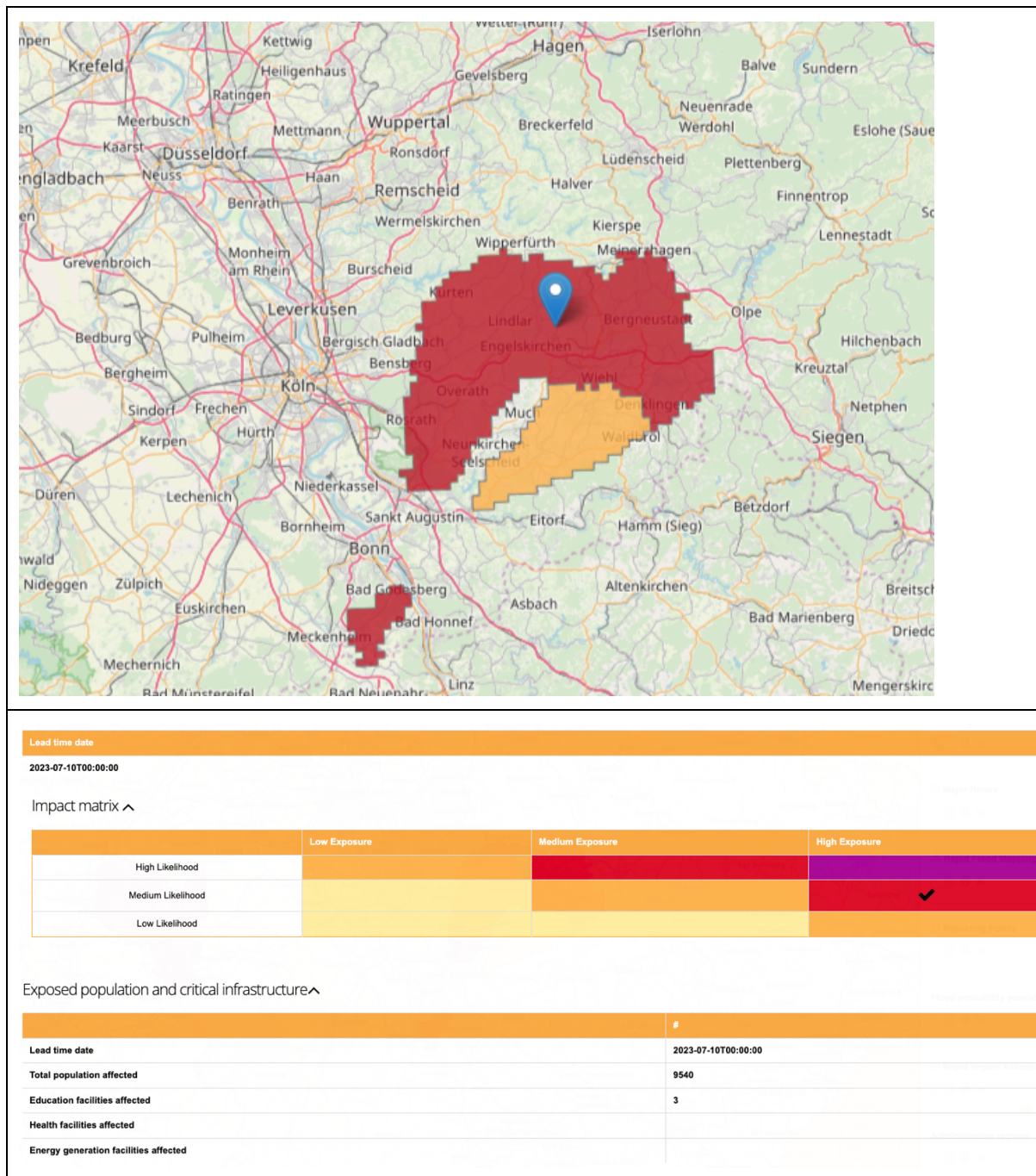


Figure 8. The Catchment summary flash flood impact product developed during the TAMIR project. Top panel shows the sub-catchments shaded according to their impact matrix classification, shown in the bottom panel. Additional information appears in a pop-out window, shown in the bottom panel.

During the EDERA project, additional information has been added to this catchment summary product:

1. The name of the river within the sub-catchment.
2. The name of the river basin within which the sub-catchment lies.
3. Names of all NUTS level 2 or 3 administration regions which intersect each sub-catchment
4. Names of all countries which intersect each sub-catchment.

River and basin names were obtained from OpenStreetMap and the European Environment Agency by the European Commission's Joint Research Centre (JRC), they were provided to ECMWF as part of their role as the computational centre of the Copernicus Emergency Management Service (CEMS) for floods. Administration region and country names were obtained from the Nomenclature of territorial Units for Statistics (NUTS) dataset ([EUROSTAT, 2022](#)). For areas outside of the European Union's member states, the region and country names were obtained from the Global Administration units database ([GADM](#)).

The above information was joined to the sub-catchments using the spatial join functionality of the GeoPandas module ([5]) within the Python programming language. Where multiple rivers intersected a single sub-catchment, the river which was closest to the outlet was identified, the name of this river was assigned to the sub-catchment. If no river name was available in a sub-catchment, the name of the basin was assigned instead. If multiple regions intersected a sub-catchment, the names of all these regions along with the two-letter ISO2 code for the corresponding country was concatenated into a single text string. The same was done if multiple countries intersected a single sub-catchment, albeit the ISO 3166-1 alpha-2 code was not included. The resulting information would be included in the pop-out window shown on the EFAS web-viewer, a mock-up of how this information might be presented is given below (Figure 9).

Geographical contextual information

	*
River	Agger
Basin	Rhine
Administration region(s)	Koln (DE)
Country	Germany

Figure 9. Mock up of how the additional geographical contextual information of river, basin, region and country names could appear in the pop-out window in the catchment summary product on the EFAS web-viewer.

The software which generates the realtime catchment summary forecast layer has been modified to include this additional information within the saved product. All changes have been committed to the [software repository](#) which is hosted on the ECMWF Bitbucket service.

Software for flood notification service

An automated flash flood notification service is one of the new developments within the EDERA project. The service will allow partners to specify their criteria, creating an object named a 'rule', that will be used to determine whether a notification should be issued about forecasted flash flood events. These criteria could include, but can be extended to include additional checks and information:

- Area of interest – as defined by a geospatial vector file (e.g. an ESRI Shapefile or GeoJSON) provided by the partner

- Lead time range
- Flash flood hazard return period exceedance probability
- Upstream area size

The software which serves the backend for the EFAS web services has been modified to include this notification facility. This has been achieved by extending the existing database software in the EFAS web services, to allow storing the rules and notifications to be issued for each partner. An additional system has been implemented to allow for the automatic execution of those rules against the latest forecast and determine what notifications should be issued.

The system can run in either “test” mode – notifications are only recorded, but not actually submitted through to the real users, or “production” mode – notifications are recorded and submitted. To achieve the submission an email server, known as an SMTP server, will be used which will perform the email submission to the users.

5 Conclusions

This milestone report has outlined the software which has been developed, or modified from previous projects, to generate the products which will be shown during the EDERA project. For convective hazards, the pre-existing software has been modified to apply the methodology at a pan-European scale. New software has been developed to generate a pluvial flood product and to obtain EFAS flood forecast data which will be used in a new compound flood product. The flash flood impact forecast product has been modified to include additional geospatial contextual information, and a new flood notification service has been developed.

References

- [1] Rossi, P.J., V. Chandrasekar, V. Hasu, D. Moiseev, 2015: Kalman Filtering-Based Probabilistic Nowcasting of Object-Oriented Tracked Convective Storms. **Journal of Atmospheric and Oceanic Technology**. 32, pp 461-477, <https://doi.org/10.1175/JTECH-D-14-00184.1>
- [2] Rossi, P.J., V. Hasu, K. Halmevaara, A. Makela, J. Koistinen, H. Pohjola, 2013: Real-time hazard approximation of long-lasting convective storms using emergency data. **Journal of Atmospheric and Oceanic Technology**. 30, pp 538-555. <https://doi.org/10.1175/JTECH-D-11-00106.1>
- [3] CLC 2018, Copernicus Land Monitoring Service. <https://land.copernicus.eu/news/corine-land-cover-now-updated-for-the-2018-reference-year>
- [4] Hansford, E., C. Baugh, M. Berenguer, S. Park, C. Prudhomme, 2021: Improved Flood Rapid Risk Assessment (RRA) tools. **TAMIR Deliverable 5.1**. 68 pp.
- [5] Jordahl, K., J. van den Bossche, M. Fleischmann, J. Wasserman, J. McBride, J. Gerrard, F. Leblanc, 2020: geopandas v.0.8.1. Zenodo. <http://doi.org/10.5281/zenodo.3946761>