January 2020

Newsletter of the *Digital Earth* Project

Contributions of GFZ German Research Centre for Geosciences

This newsletter presents some specific efforts of GFZ Potsdam in activities related to the Show Cases or Work Packages of Digital Earth.

Flood Use Case Model

Bruno Merz, Stefan Lüdtke
GFZ German Research Centre for Geosciences

Flood events are cross-compartment phenomena that lead to cascading effects and complex event chains, affecting multiple compartments in a non-linear, time delay manner. The Flood Event Explorer (FEE) we develop in the Digital Earth project will allow a comprehensive assessment of flood event characteristics along the process cascade, considering atmospheric triggering events – catchment preconditions – runoff generation – flood wave routing – inundation – damage - and possibly other effects like water quality characteristics of the affected estuary. The main challenge addressed by the Flood Event Explorer is to provide tools in order to optimize monitoring designs (MOSES), predicting/projecting short-term and long-term developments and understanding process interactions across compartments.

FEE is exemplarily developed for Germany with focus areas in the Elbe catchment. Despite its exemplary character, FEE is a challenging endeavour as it comprises a variety of workflows from different scientific domains that operate on different spatial and temporal scales. The presented concept is rather general and a consolidation of topics is accomplished within the toolboxes.

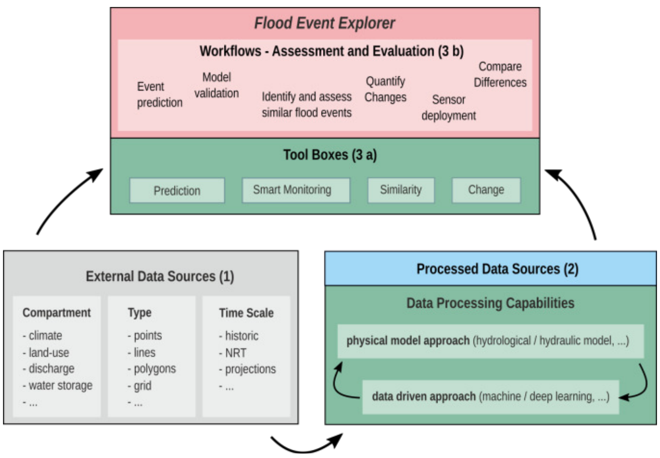

Figure 1 Concept chart of the flood event explorer and the flood use case model

The following paragraphs describe the FEE concept shown in Figure 1 in more detail.
The external data sources heavily vary for each toolbox and the corresponding scientific workflow. It is worth noting that multiple combinations across the group members may exist.

For example: The climate variable precipitation may be available as station data (type: points) or come already interpolated as a raster data set (type: gridded). Similarly, the variable might be available as a record of past observations (time scale: historic) or be a result from a climate model run (time scale: projection).

Processed data source are required as an intermediate step and primarily serve to fill data gaps in observed data from external data sources (e.g. inundation maps for past flood events). Methods to derive this additional data include process based models, machine learning algorithms or other statistical methods. In general, processed data sources can cover the same compartments, types and time scales like external data sources, but they are derived with domain specific knowledge from the different Helmholtz Centres and their working groups. Similar to the external data sources, they provide required information to accomplish a specific workflow.

Toolboxes group tools that serve a similar purpose, but at the same time, aid the realization of multiple scientific workflows. Scientific workflows are considered to be domain specific investigations for a particular research question or purpose. For example, the toolbox Smart Monitoring is able to support the workflow of identifying sensor positions shortly before a field campaign starts to ensure an optimal placement. On the other hand, it is suited to validate and assess existing data (e.g. by the validation with model results) and thus improve upcoming field campaigns in the long run.

Data Exploration Framework

Doris Dransch, Daniel Eggert
GFZ German Research Centre for Geosciences

The Flood Event Explorer can be seen as an exemplary realization / implementation of the *Data Exploration Framework* we develop in WP2 of the Digital Earth project.

Our goal with the *Data Exploration Framework* is to develop a digital research environment that supports scientists to create a more holistic view to systems, in our case the flood system, by integrating separated analytical methods in a software framework, allowing seamless / continuous digital workflows, and providing integrative views to data and data products. We develop a generic concept and implement it exemplarily for the show case flood in the "Flood Event Explorer". As a basis we use and combine several concepts from computer science. These are: a) Digital workflows which allow for implementing workflows on a scientific-conceptual, methodical and technical level. b) Software Frameworks which allow component-based, combinable and reusable software development. c) Interactive visual interfaces which provide interactive multiple linked views to data and data products. The Flood Event Explorer will consist of several digital workflows and novel methodical approaches to investigate the flood system across compartments. On a technical level it will consist of reusable and combinable software components.

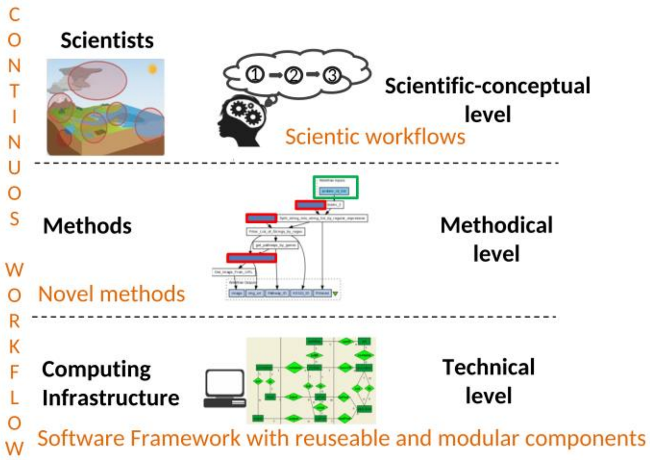

Figure 2 Digital Workflows

Current work: Scientists from the centres that are part of the Digital Earth project came together on multiple occasions like meetings and workshops, where we collected the scientific and technical requirements for the *Data Exploration Framework* / Flood Event Explorer. As one result we described five scientific workflows for the Flood Event Explorer, one is the Similarity Workflow, another one is the workflow related to Advanced data integration methods towards large-scale flood impact indicators, both are developed at GFZ and described below. A further outcome is the concrete realisation of the software framework for the technical level of the workflows. We developed the following concept and implementation: The basic idea for the general structure of our *data exploration framework* is a platform independent modularization approach as seen in Figure 3. This general structure allows the development of methods as decoupled back-end modules (fig. 1 (A), (B), (C)) in almost any desired programming language and operating system. The key to connect all initially decoupled modules is the use of a messaging system. For our implementation we use a messaging system called Pulsar, which provides a wide range of language bindings, as well as a general purpose web socket based interface.

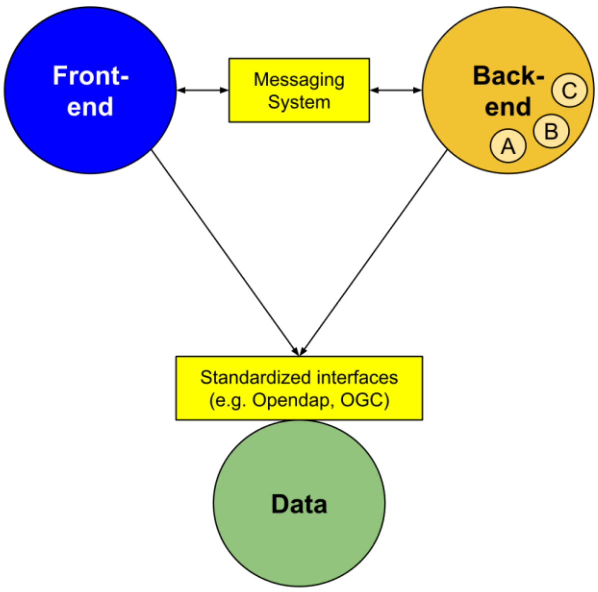

Figure 3 General technical structure of the data exploration framework

Another key factor of our proposed structure is the focus on standardized web based interfaces for data access, like OGC or Opendap services. This unifies and therefore simplifies the data access level and allows for a much easier integration of data from different sources and scientific domains and hence directly supports cross-compartment research. Since major parts of the data driven analysis workflows are going to facilitate visual data exploration approaches, the front-end, as well its components are the third key factor of our *data exploration framework*. In general the presented structure allows for arbitrary front-end modules, ranging from standalone (e.g. GIS) to web-based applications. Because of the easy accessibility and the broad availability of existing visualization libraries like D3.js and Openlayers, we decided to develop our front-end as a web based application.

Similarity Workflow

Stefan Lüdtke, Daniel Eggert
GFZ German Research Centre for Geosciences

River floods and associated adverse consequences are caused by complex interactions of hydro-meteorological and socio-economic pre-conditions and event characteristics. Floods, as exceptional events at different spatiotemporal scales, often reveal unexpected system behaviour and offer unique opportunities to gain system insights across multiple compartments and to improve existing methods and models. On the other hand, increasing flood losses indicate the immediate need of science-based facts to support flood risk management as well as disaster relief and recovery.

In our *Similarity workflow* we develop an approach to assess an arbitrary flood event against historical or future events – including the complete process chain from the triggering atmospheric conditions and hydrological pre-conditions to the flood consequences. The approach covers the last 70 years comprising a flood event catalogue for Germany. This provides the basis to identify and compare similar flood events based on, for example, antecedent catchment conditions, catchment precipitation, discharge hydrographs, and inundation maps. The *Similarity workflow* as part of the Flood Event Explorer enables the analysis of multidimensional flood characteristics including aggregated indicators (in space and time), spatial patterns and time series signatures. The similarity between the flood events is assessed using a multitude of statistical methods and machine learning algorithms incorporating the aforementioned multidimensional flood characteristics.

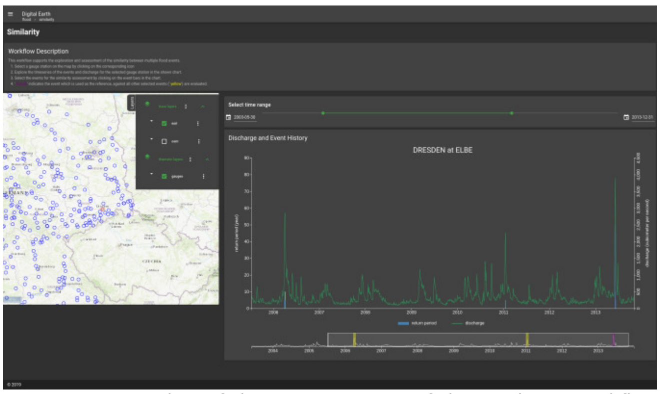

Figure 4 Screenshot of the current status of the similarity workflow implementation

Bridging PostDoc Project: "Advanced data integration methods towards large-scale flood impact indicators"

Kai Schröter[1], Diana Rechid[2]
1 GFZ German Research Centre for Geosciences
2 Climate Service Center Germany (GERICS), Helmholtz-Zentrum Geesthacht (HZG)

Floods impact individuals and communities and have large social, economic and environmental consequences. These impacts are dynamic and vary largely depending on flood characteristics, vulnerability, and value of exposed elements. The understanding of large-scale flood impacts as the outcome of complex interacting atmospheric, hydrological, inundation and socio-economic processes including human interventions is still limited and so is the knowledge about controls, feedbacks, and changes in flood impacts. An improved understanding of this interplay is needed for manifold tasks in flood risk management, for instance, emergency response and adaptation planning.

With an increasing abundance of data from a growing number and diversity of sensors, methods of data science take a key role in extracting information and knowledge from these heterogeneous data sources to deepen insights into the generation and development of floods as well as the interactions between hazard, exposure, and vulnerability in shaping flood impacts.

This bridging postdoc project addresses two major challenges in exploiting the increasing volume of data and variety of parameters from different sources available. The first problem is related to making use of novel data sources and integrating these with data from established observation systems for the application case of rapid inundation depth mapping. The second problem concerns integrating disparate and heterogeneous data sources to explore and analyse controls on floods and impacts. These challenges are addressed within a multi-variate data analytics approach of observations and simulation model outcomes of climate, hydrologic, inundation and impact variables. Data science methods are used to integrate the heterogeneous data sets and investigate the spatial and temporal dependencies of potential flood controls, their interactions as well as their causal relationships to flood impacts. The key outcomes of this project are methods and tools for the fusion of multiple data sources for rapid inundation mapping and relating climate and weather variables to flood impact indicators for enhanced flood event assessment. The scientific workflow related to *Advanced data integration methods towards large-scale flood impact indicators* will be integrated into the Flood Event Explorer framework.

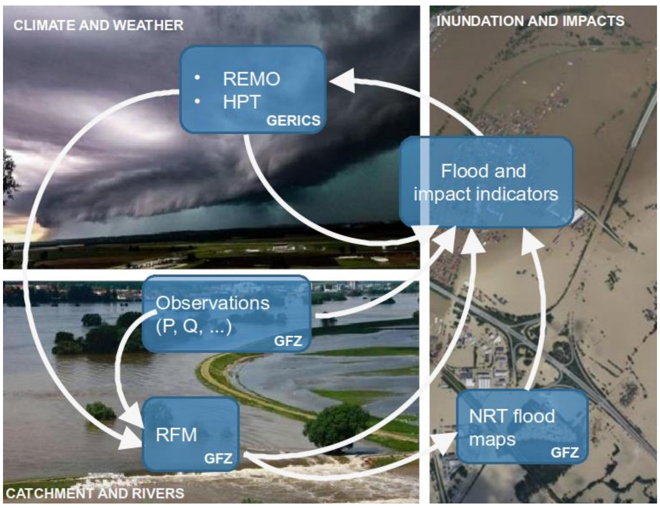

Figure 5 System sketch of the involved compartments

Bridging PostDoc Project: "Enhancing Flood Data Workflows through a Coupled Model Chain"

Dadiyorto Wendi[1], Sergiy Vorogushyn[1], Marius Oster[1], Björn Guse[1], Oldrich Rakovec[2], Stephan Thober[2], Luis Samaniego[2]
1 GFZ German Research Centre for Geosciences
2 UFZ Helmholtz Centre for Environmental Research

The reliable understanding and quantification of flood risk requires extensive datasets across multiple temporal and spatial scales. The representation of the full process chain and their associated interactions is crucial to improve the system understanding. However, gaps in observation data and the lack of interdisciplinary model integration still hinder the progress in large-scale flood risk assessment. This includes the process dynamics from the triggering factors of an atmospheric event, catchment preconditions, down to the river and floodplain hydraulics and subsequent damages.

To bridge observation data gaps, coupled simulation models consisting of a process-based mesoscale hydrological model (developed by UFZ) and a physically-based hydrodynamic inundation model (developed by GFZ) are being integrated. The hydrological model can either be driven by weather forecasts, climate scenarios or by a stochastic weather generator. The data outputs from the model chains is an extensive flood event catalogue that can be used to for operational flood impact forecasting, climate impact assessment or risk estimation

The bridging postdoc project is focused on the technical model coupling and an enhancement of cross compartment data workflows towards the integration with near real time (NRT) monitoring and weather forecasting products (e.g. COSMO-d2-eps). The outcome of this endeavour is expected to bridge existing data gaps useful for the generation of flood event catalogue containing a large set of flood characteristics along the entire process chain. The proposal is linked to the Digital Earth Showcase Floods (Flood Event Explorer) and to MOSES (Hydrological Extremes) working groups.

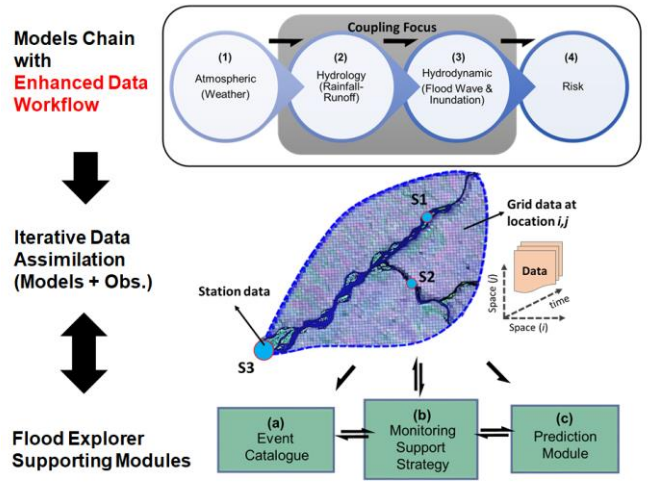

Figure 6 Conceptual overview of the enhanced cross-compartment data workflow