DEPLOYMENT AUTOMATION OF HYDROLOGICAL FORECASTING SYSTEMS ON A GLOBAL SCALE.

Arianna Onate-Paladines[1], Bart Van Osnabrugge[2], Jan Verkade[2], Albrecht Weerts[2,3]

[1] Universidad Técnica Particular de Loja, Loja, Ecuador - ayonate@utpl.edu.ec
[2] Deltares, Delft, Netherlands
[3] Wageningen University, Wageningen, Netherlands

INTRODUCTION

Hydrological forecasting systems are fundamental to support various water-related sectors. Their deployment process from scratch constitutes an important challenge: It can be costly and time-consuming and it can pose as well constrains over its maintenance and expansion once deployed. This is because these systems are traditionally tailor-made, and configuration automatization and replicability are not considered when developed. Although, some open frameworks do allow the reuse of previously built components.

OBJECTIVES

To overcome the exposed limitations, the automatic deployment of hydrological forecasting systems aims to:

- Transfer expertise from one system to the next, speeding-up and simplifying the configuration process.
- Simplify the system's maintenance process.
- Enable the straightforward addition of new data products and hydrological models, when developing large hydrological systems (Such as Deltares' GLOFFIS) as well as local systems.

METHODS

We took as example the Delft-FEWS forecasting platform, which uses XML files for system configuration.

- We started defining the forecasting chain structure of the system,
- Later, we selected the configuration components to be automatically generated (mainly related to data products and hydrological models characteristics).
- Finally, we moved towards the creation of a production system (Figure 1), composed of two main elements:

1. A relational database that compiles all the information required to build the system, (whose ER diagram is shown in figure 2), and
2. A set of python scripts, that produce the required configuration files.

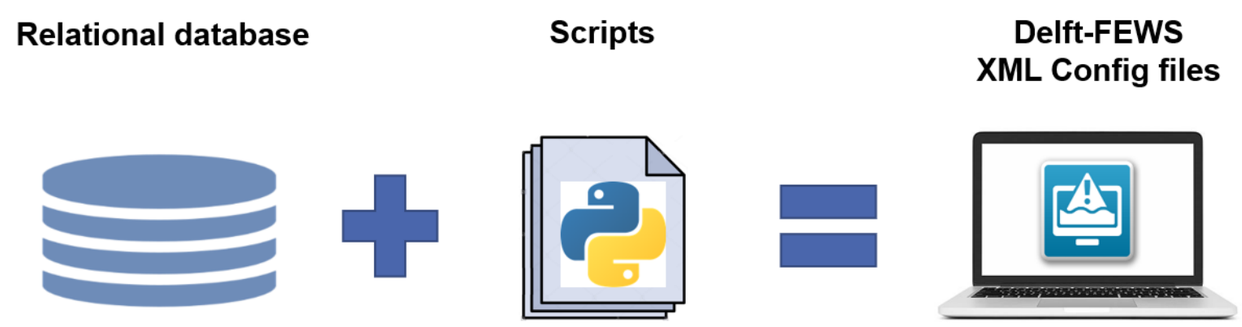


Figure 1: Elements of the Production System

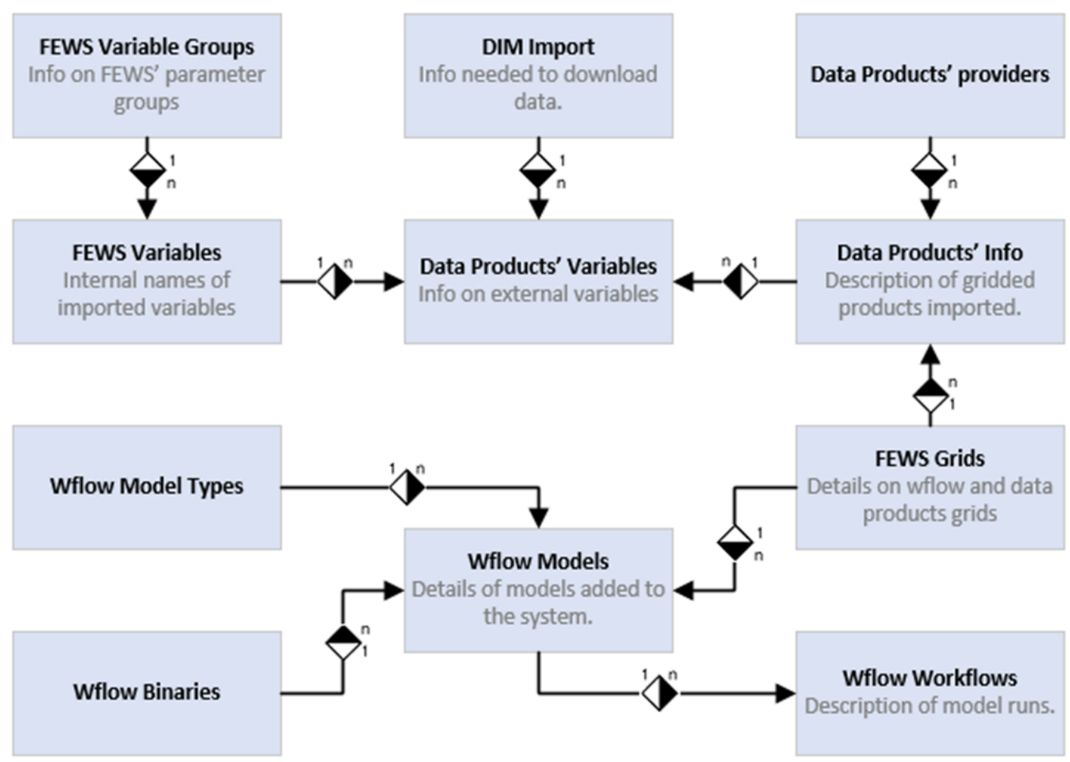


Figure 2: Simplification of the Entity Relationship Diagram of the Database

The hydrological forecasting system was developed for a sample catchment in Europe for different spatial resolutions, and the system was replicated to various other catchments. The data incorporated were DWD's ICON NWP and ECMWF's ERA5 reanalysis.

RESULTS

This automation methodology accomplished not only the individual recycling of some components, but the complete reproduction of an already designed configuration for a different location (enabling an alternative solution to deliver hydrological forecasts around the globe). It also demonstrated the capability of a rapid expansion and maintenance in terms of meteorological data products and hydrological models.

Deltares UTPL WAGENINGEN UNIVERSITY & RESEARCH

Deltares

Configuration of GLOFFIS-Prod:
Generating config through an access database and scripting

Arianna Onate-Paladines, Bart van Osnabrugge, Jan Verkade

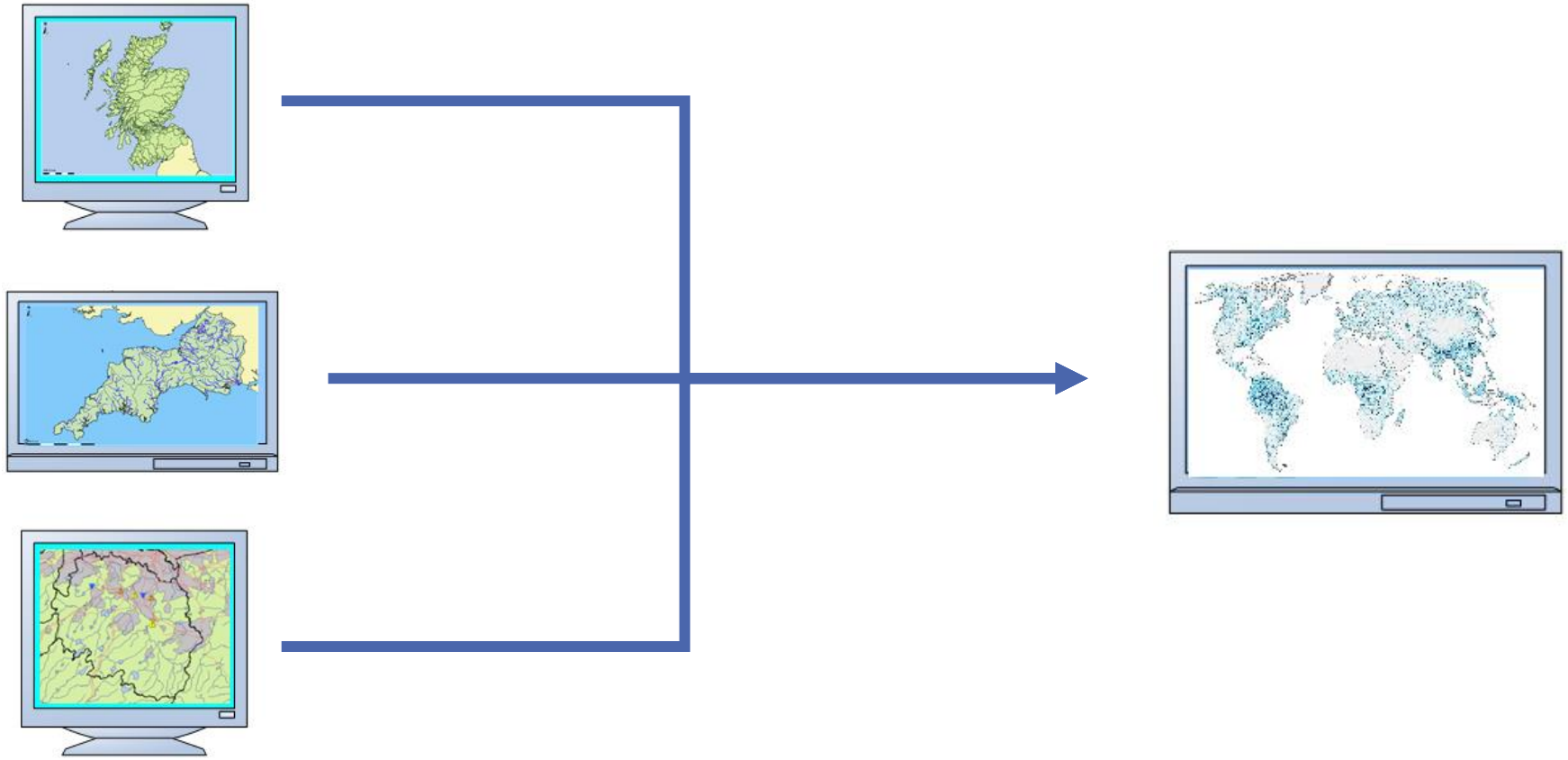
Contact: Bart.vanOsnabrugge@deltares.nl

Structure of Presentation

1. Introduction: GLOFFIS
2. Configuration of GLOFFIS-Prod
3. The Production System: Relational Database
4. The Production System: Scripts
5. wflow models incorporated to the system
6. Conclusions / Final thoughts

Introduction

Global Hydrological Forecasting as Data Services

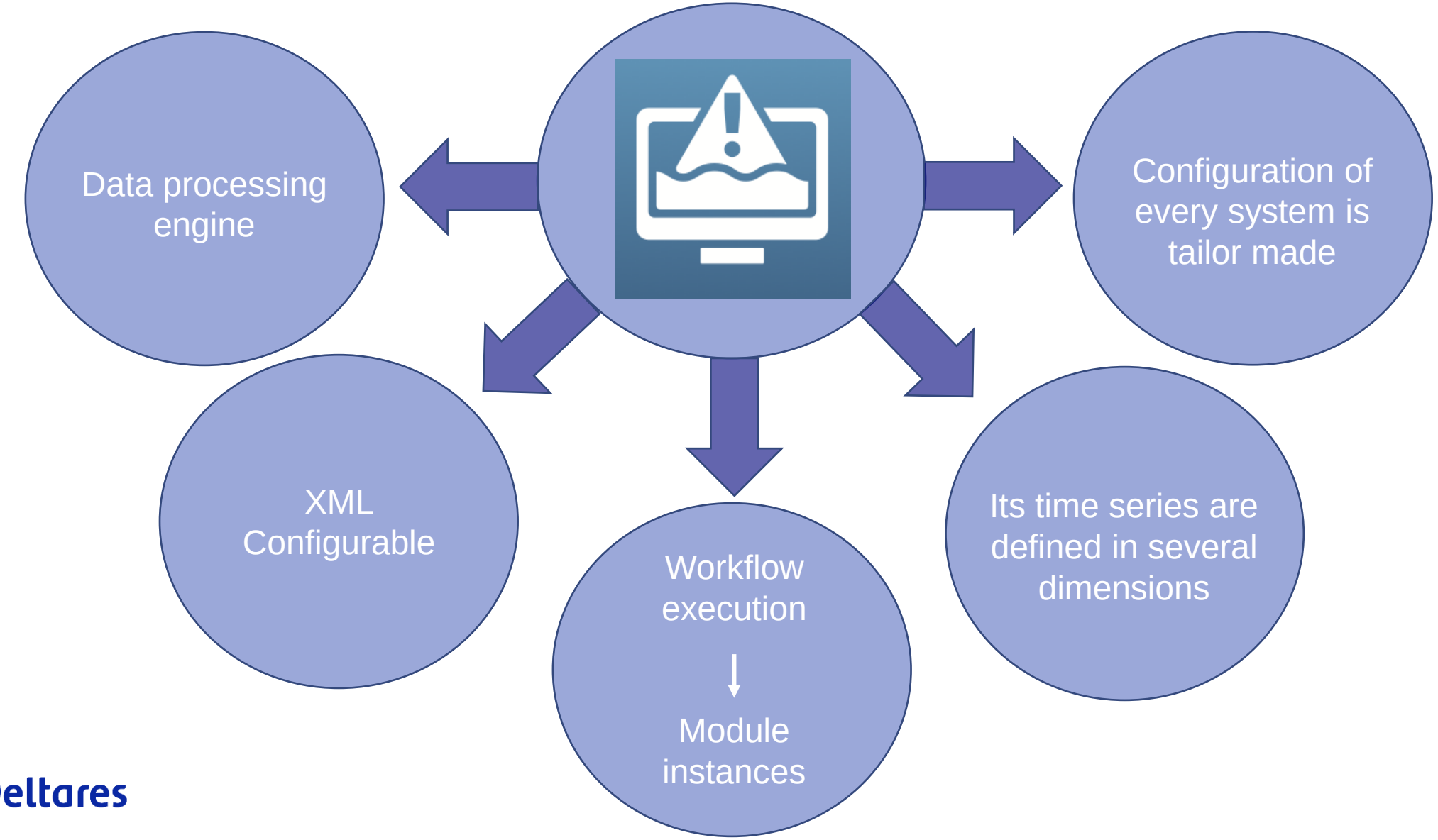


Local Hydrological Forecasting Systems

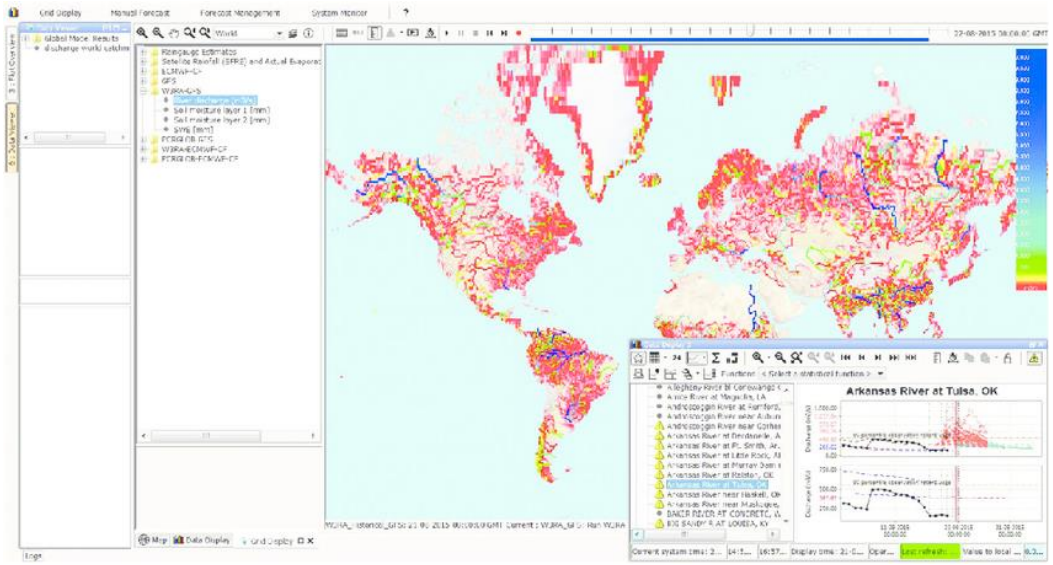
(Single) Global Hydrological Forecasting System

Deltares

Delft-FEWS as forecasting platform



Versions of GLOFFIS



- Based on Delft-FEWS.
- Runs the global hydrological models W3RA and PCRGLOB-WB (50 km resolution).
- Uses GFS, GEFS and ECMWF ENS data.

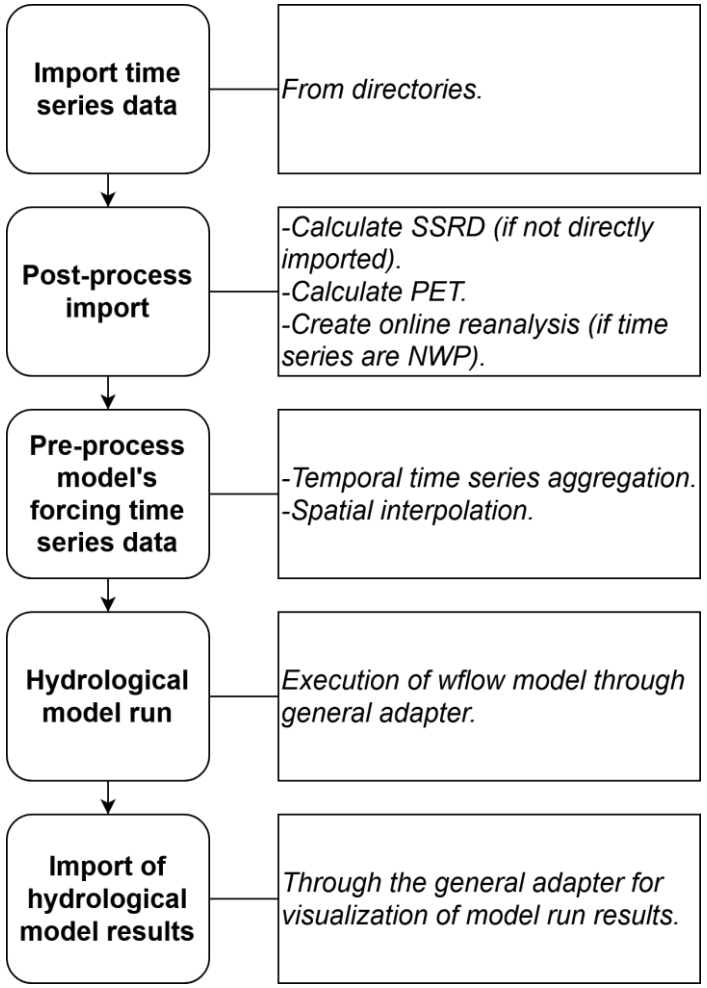
System is in redesigning phase to: (among other reasons)

- Include other NWP and historical products with regional and global coverage.
- Incorporate higher resolution wflow models (1 km) for catchments of interest.

→ Set up of a production system

Configuration of GLOFFIS - Prod

Configuration of GLOFFIS - Prod



- Gridded data products imported:
  - DWD ICON NWP family (ICON, ICON-EU, ICON-EU-EPS)
  - ECMWF reanalysis ERA5.

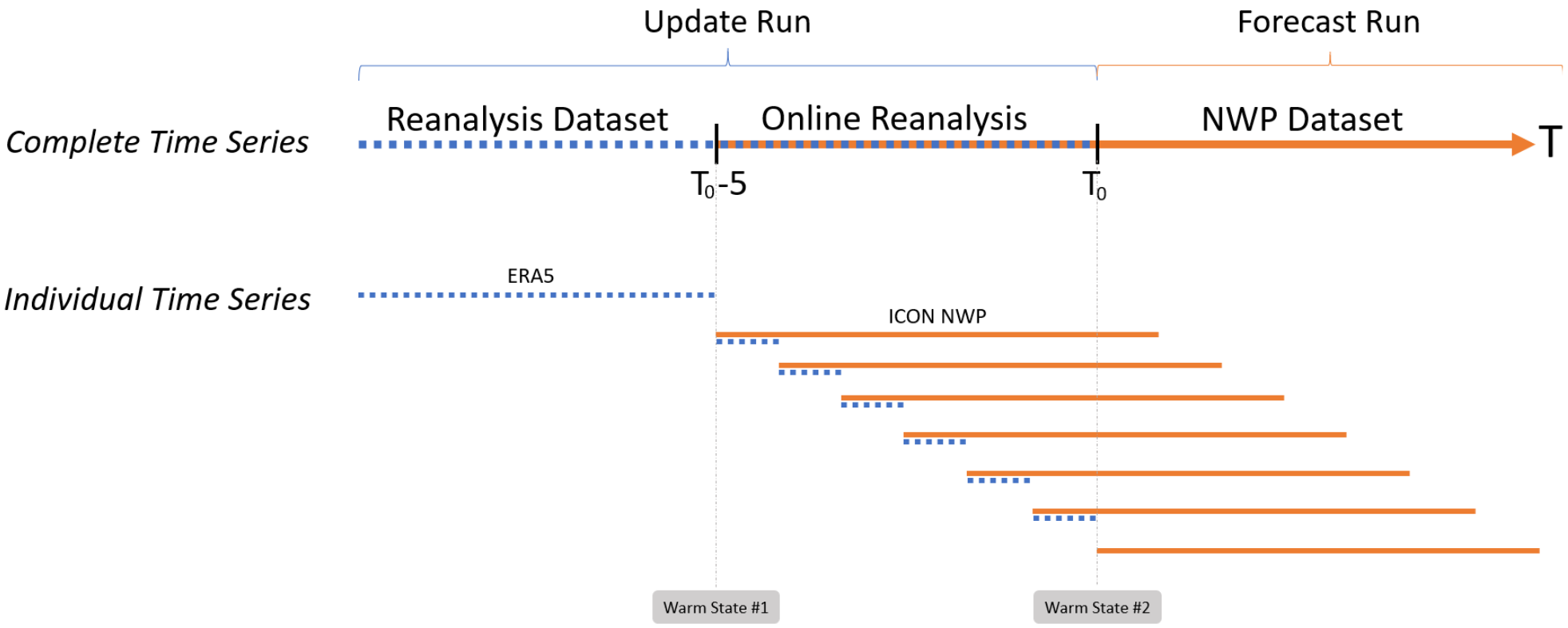
- Variables imported:
  - Total precipitation, temperature, surface solar radiation.

- PET is calculated using Makkink (1957) methodology.

> Templates for PET calculation, temporal and spatial pre-process, etc. are placed in default config folder and depend on the time series type.

Deltares

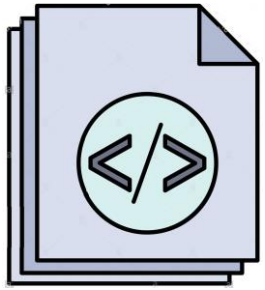
Forecast Chain considered for each model

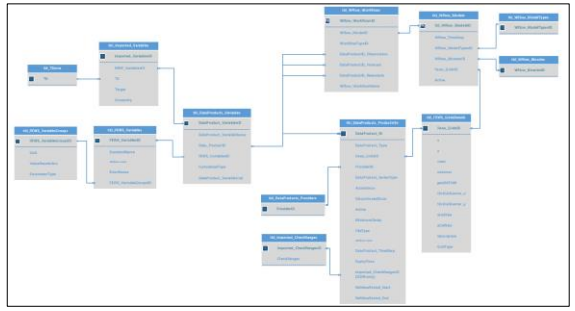


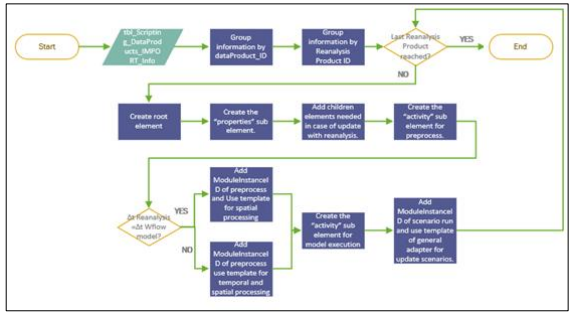
The Production System: Relational Database

Composition of the Production System

Relational database **Scripts** **XML Config files**



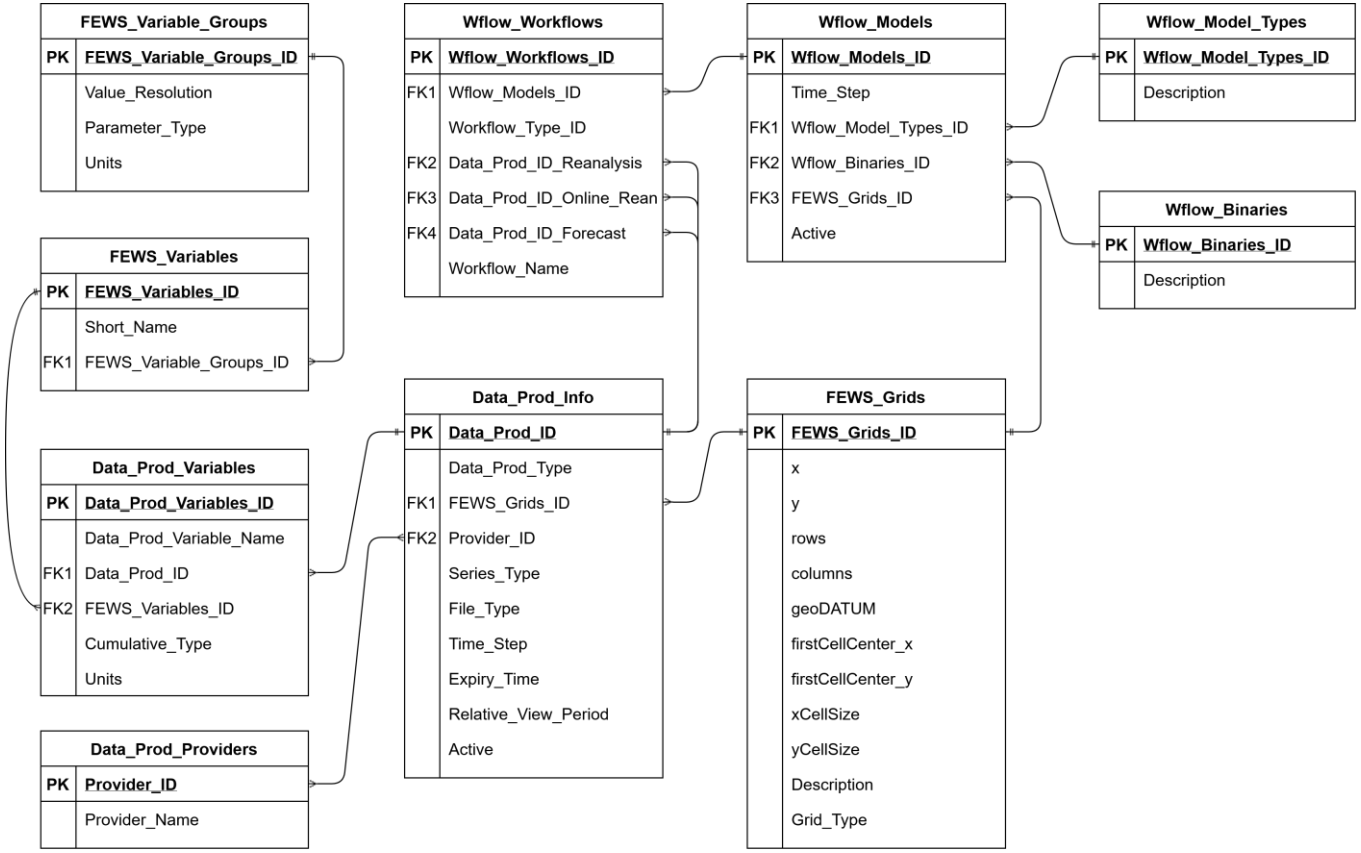




Deltares

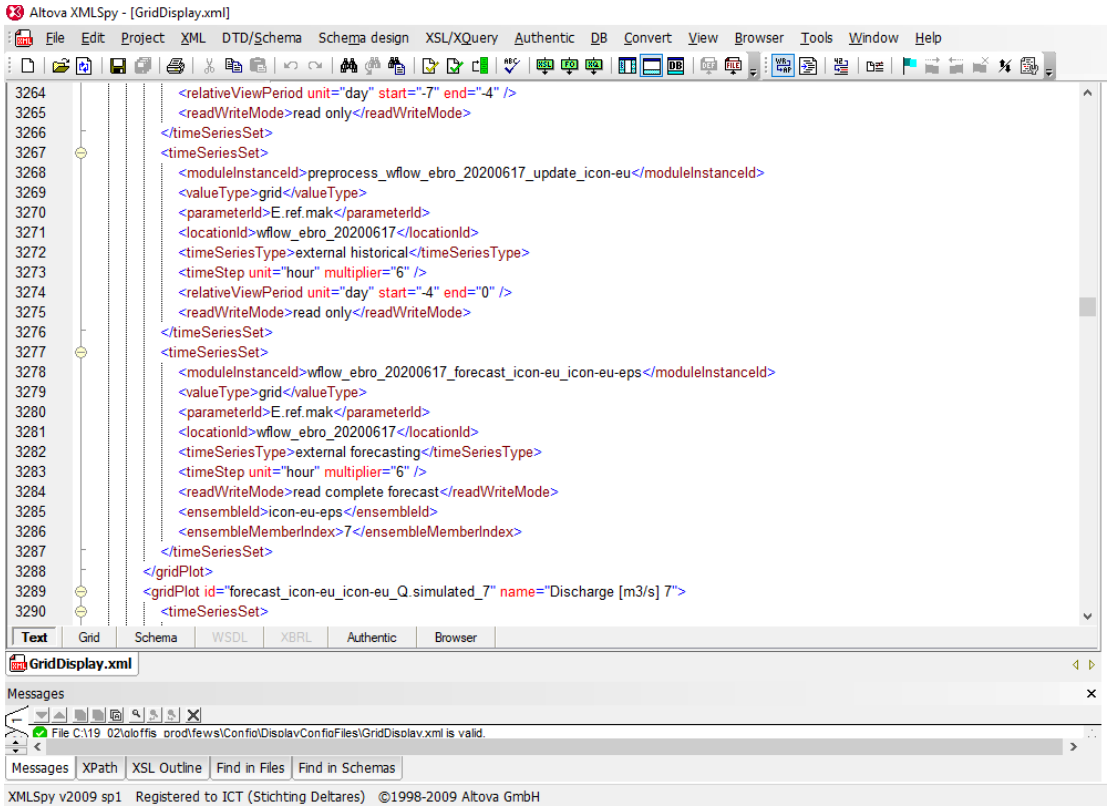
Relational Database

- Useful because:
  - Compiles and organizes info with defined relationships: Edit and manage data entries without affecting logical construction.
  - Maintains data consistency and avoids redundancy.
- Built using Microsoft Access.

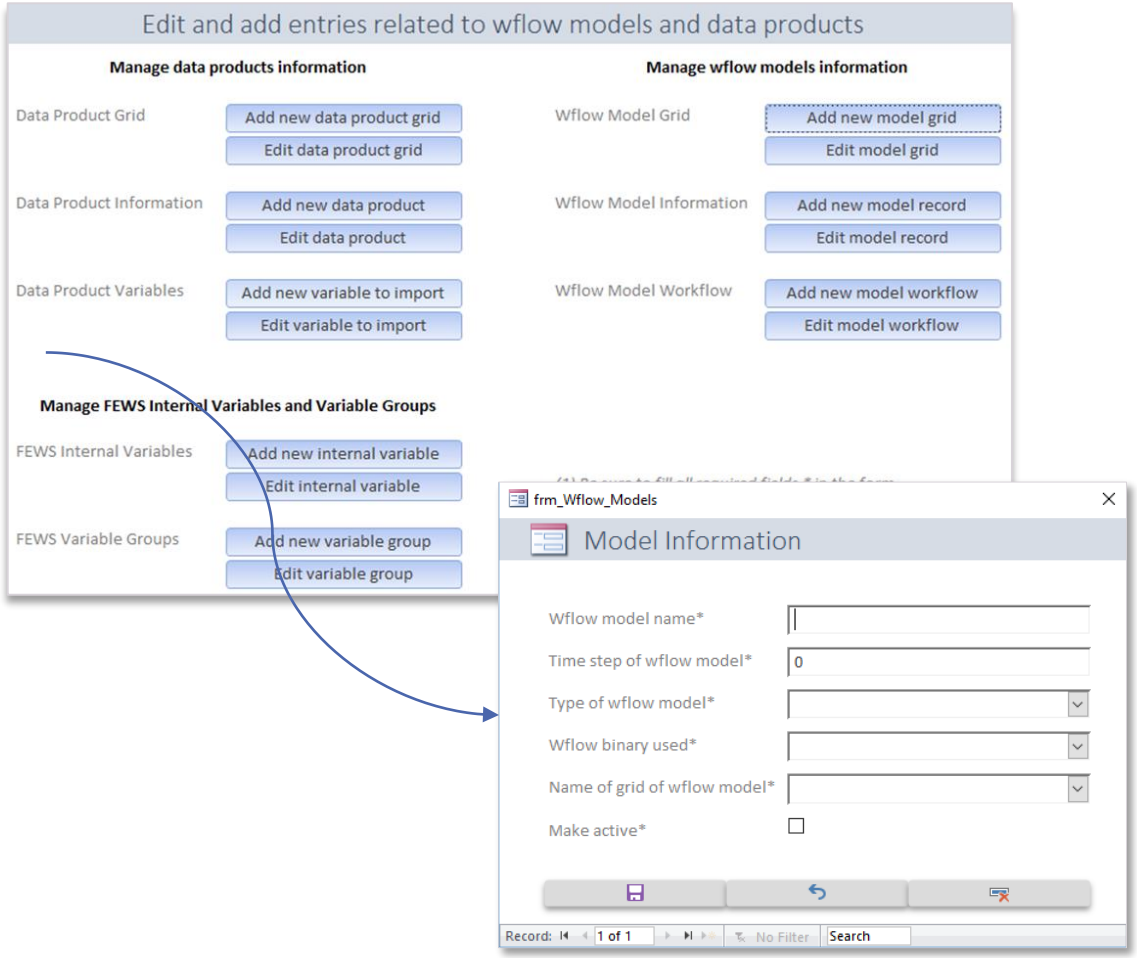


Overview of tables of relational database.

Data entry to database



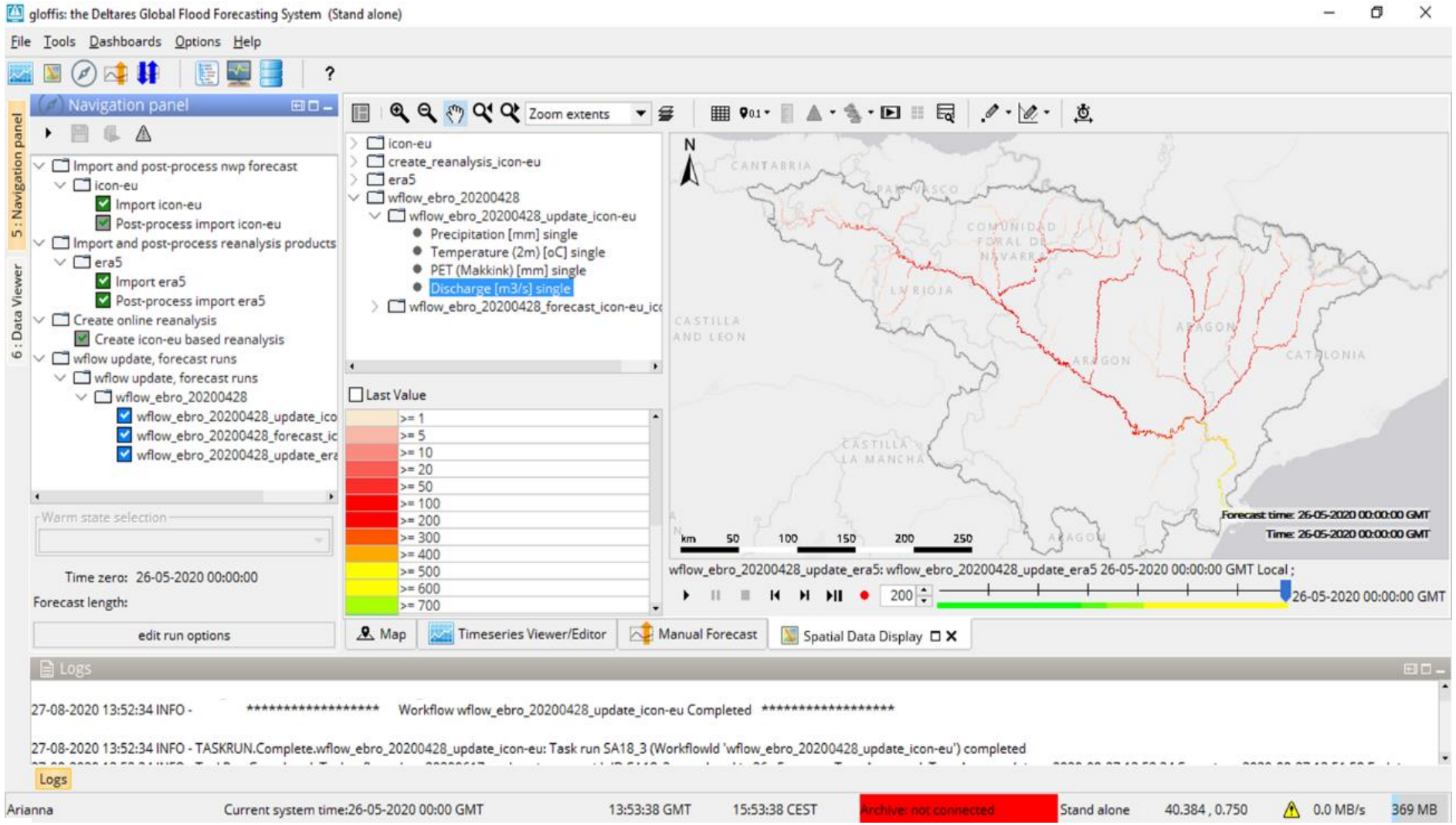
Manual writing/edition of XML files



Data entry form example - Database

Deltares

Generated configuration executed by DELFT-FEWS



Voettekst van de presentatie

The Production System: Scripts

GLOFFIS – prod directory

Final configuration (defaults + scripts).

Contains the .accdb database read by the scripts, the master script and other files needed by the scripts.

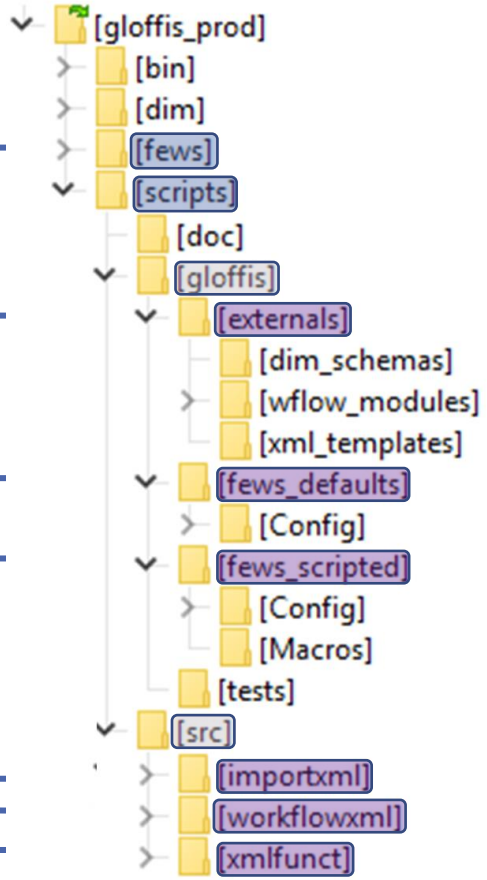
Basic configuration of the system.

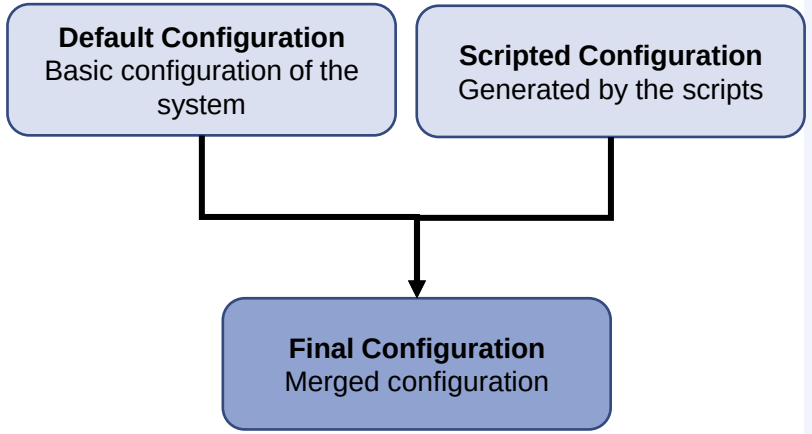
Configuration generated by the scripts.

Scripts used to the generate import-related files.

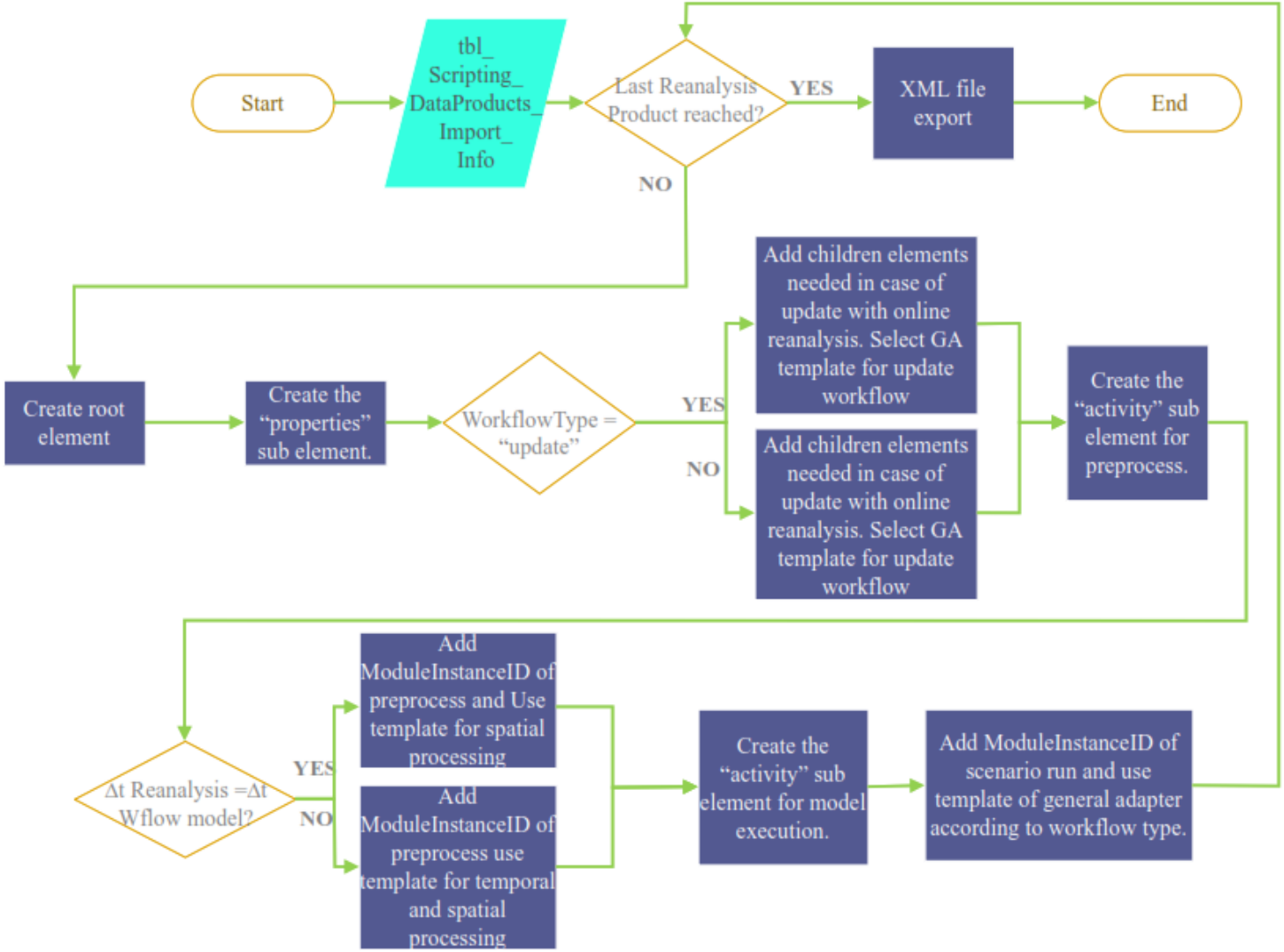
Scripts used to generate wflow models – related files.

Scripts that support the creation of xml files.





Flowchart of script (sample)

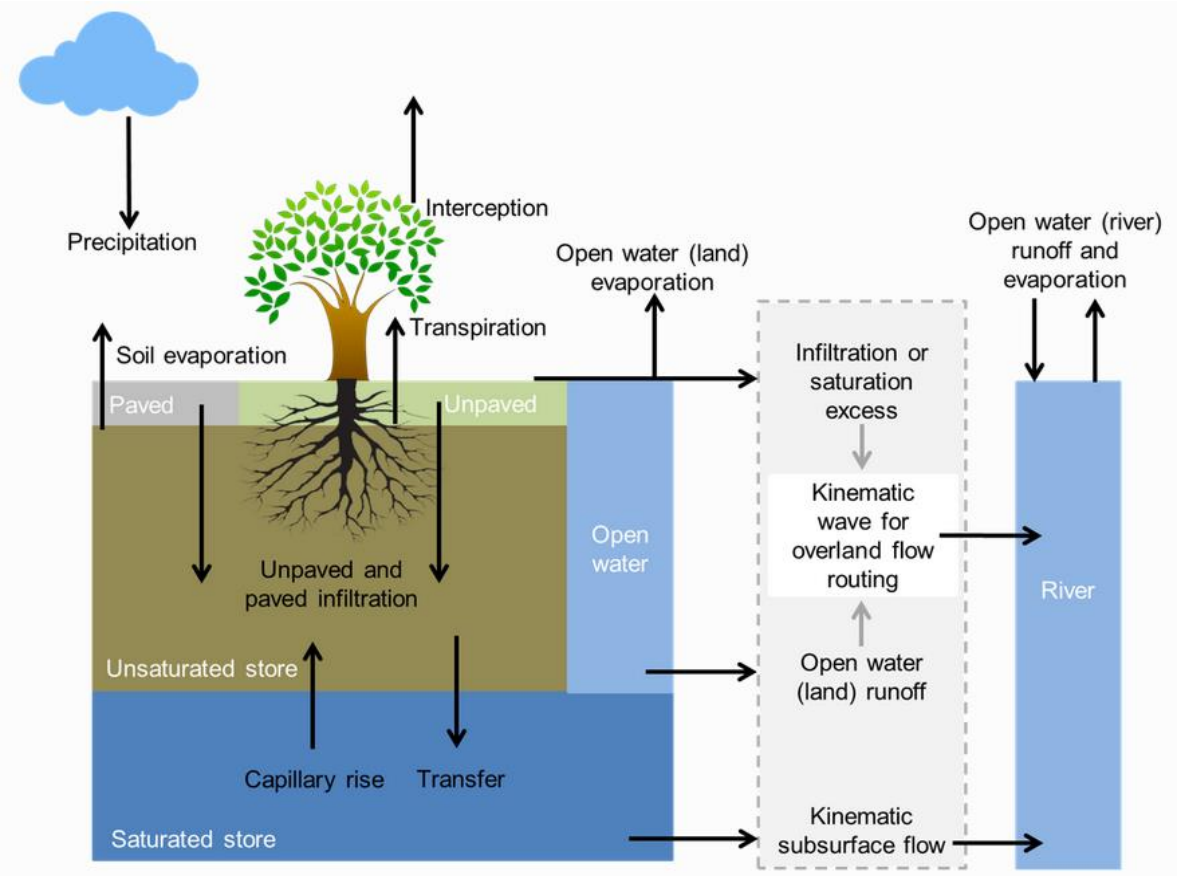


Flowchart of script to create workflow file to execute model in update mode (using online reanalysis) and forecast mode (using NWP).

Voettekst van de presentatie

wflow models incorporated to the system

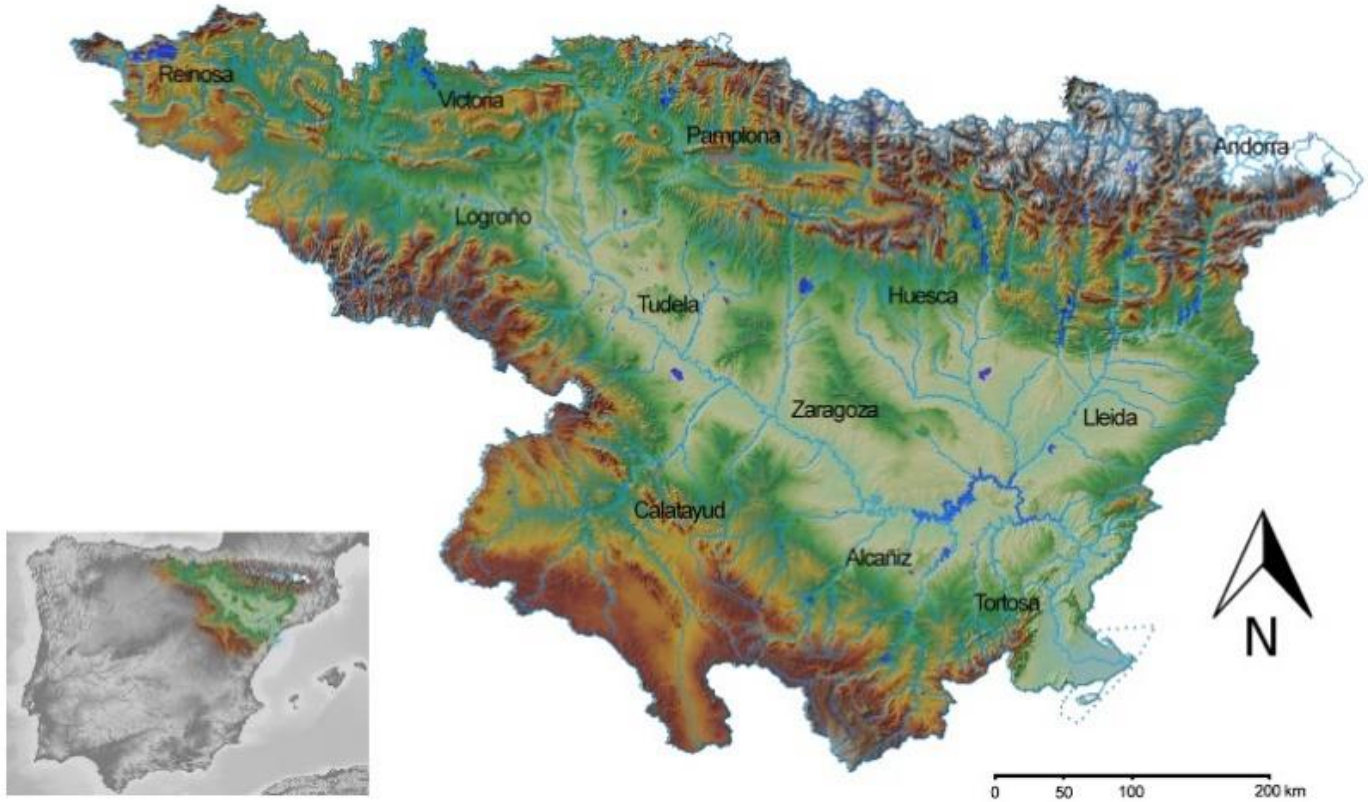
The wflow sbm model



Source: *Schellekens (2019)*

- Spatially distributed.
- Requires physical parameters:
  - DEM
  - Land cover
  - Soil type
- Meteorological input data:
  - Precipitation
  - Temperature
  - Potential evapotranspiration
- Flow is routed using the kinematic wave.

Wflow model set-up: Ebro catchment at various spatial resolutions



Source: *Chebro (2004)*

- 3 spatial resolutions were considered:
  - 1 km, 6 km and 13 km.
- Daily timestep.
- Simulation period: 01-01-2008 to 31-12-2019.
- No calibration involved.

Area: 85 534 km^2

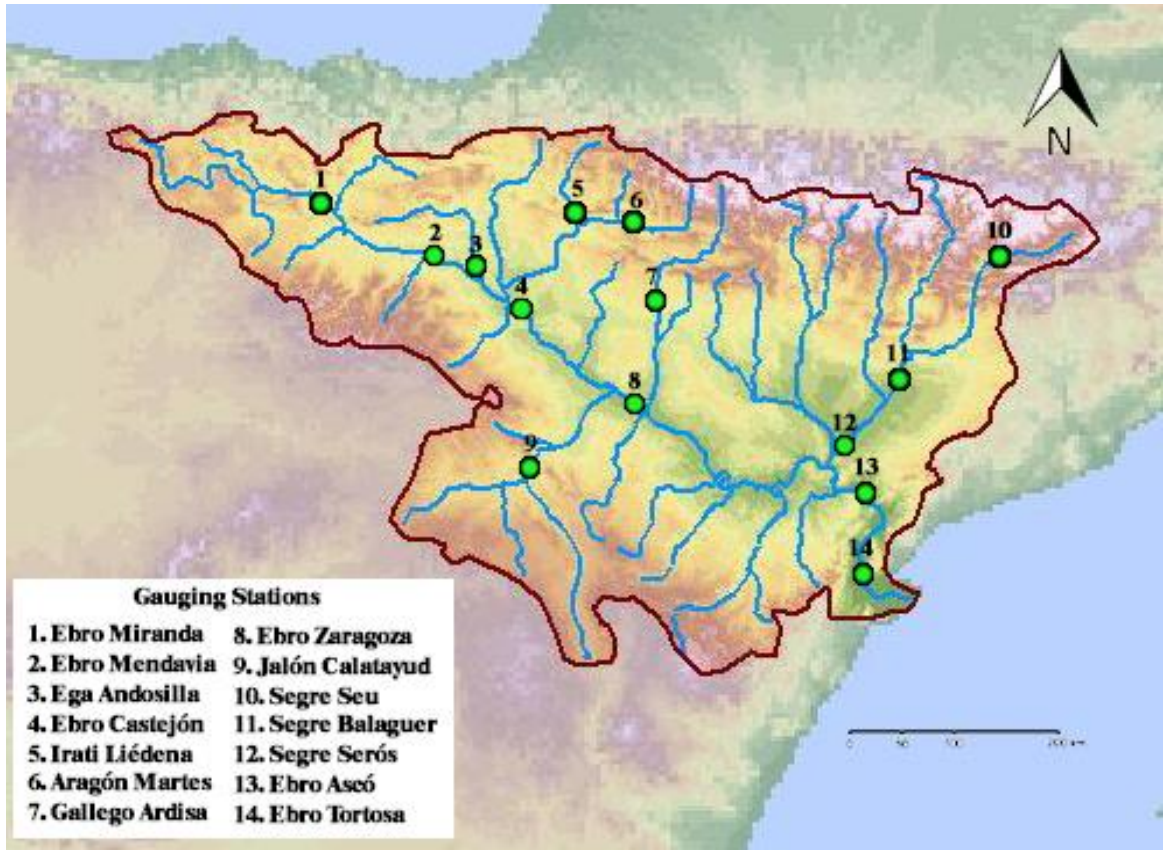
Mean precipitation: 622 mm/yr (irregular distribution)

Average annual flow: 464 m^3/s

Deltares

Wflow model performance assessment

14 gauging stations were considered for validation:



- Overall performance assessed by:
  - Nash-Sutcliffe Efficiency (NSE) coefficient

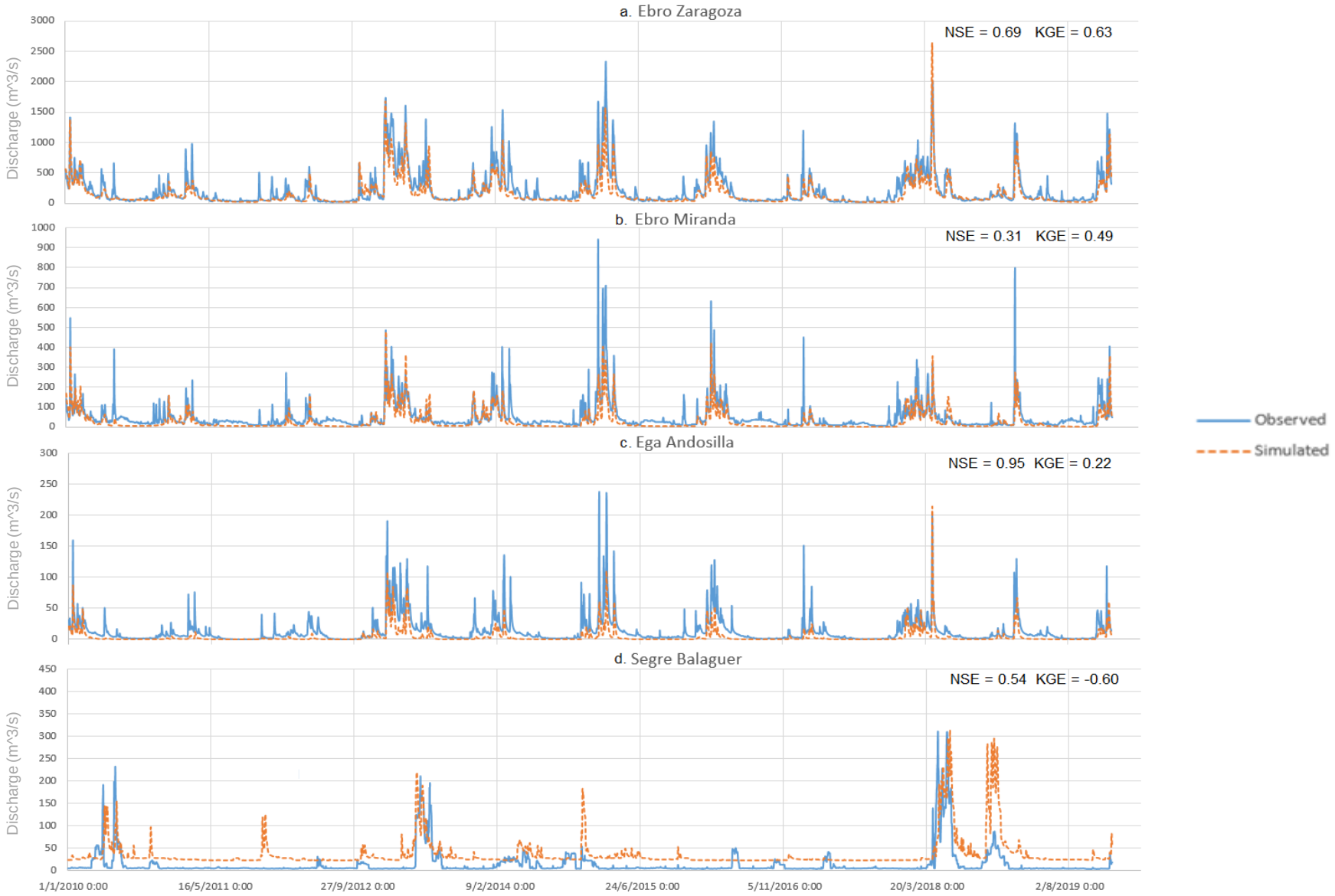
$$\mathrm{NSE} = 1 - \left[\frac{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_i^{sim}\right)^2}{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_{obs}^{mean}\right)^2}\right]$$

  - Kling-Gupta Efficiency (KGE) coefficient

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2}$$

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2}$$

Simulated and observed hydrographs (1 km)



Conclusions / Final Thoughts

Conclusions / Final Thoughts

- Setting up forecast production systems through a database and scripting is a relatively new approach with potential to greatly increase reproducibility.
- By introducing the relational database and scripting, the data management and forecast system set-up are decoupled.
- this reduces the need for manual edits of the system and thus:
  - Eases the work of system (re)production.
  - Reduces the time spent on building the system.
- It makes the configuration tasks highly dynamical, boosting the flexibility of the system, particularly in a global context. Useful to:
  - Scale-up the number of data products
  - Add higher-resolution wflow model rapidly and on the go.

Conclusion: wflow models

- For this catchment the NSE and KGE coefficients (and thus the simulation efficiency) decreased with coarser resolutions.
  - Acceptable performance throughout all resolutions in most stations over the Ebro main channel (NSE = 0.56 to 0.72 & KGE = 0.49 to 0.71)
  - Stations in tributaries for resolutions of 1 and 6 km presented statistically acceptable NSE values in most cases, however just half of these stations presented KGE values above 0.3. For the coarsest resolution (13 km), the efficiency was found poor for both coefficients.
- Overall, performance was acceptable for the higher spatial resolution models considered.

Thank you for your attention!
Any questions?

Contact: Bart.vanOsnabrugge@deltares.nl