WRF-Hydro hydrological model and its use in ARPA-FVG

6th Workshop on Water Resources in Developing Countries: Hydroclimate Modeling, Information Tools and Simulation Techniques

Sara Bacer,

Regional Center for Environmental Modelling, ARPA-FVG

Outline

Part I

WRF-Hydro: some general information

- Main WRF-Hydro characteristics
- Input and output data of WRF-Hydro
- Pre-processing of WRF-Hydro
- Hydrologically conditioned terrain data
- Spin-up, calibration, and validation

Part II

Application of WRF-Hydro in ARPA-FVG

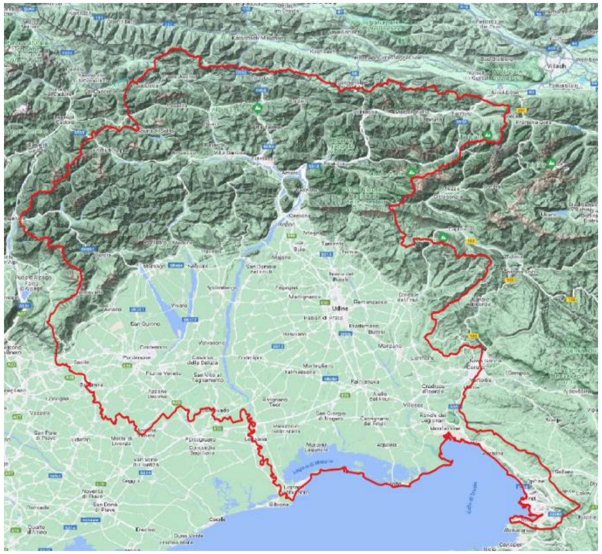
- Model setting
- Calibration
- Some specific issues

Why a hydrological model and why WRF-Hydro...

Why to use a hydrological model in ARPA-FVG

In ARPA-FVG: the **hydrodynamic model SHYFEM** is run in operational mode to produce forecasts of temperature, velocity, and sality of water in the northernmost part of the Adriatic Sea.

Boundary conditions: discharge forecasts at the river mouths.



- So far, no model considered by SHYFEM is a fully distributed one, calibrated for our region.
- In the Lagoon, also discharge of minor rivers is relevant.

We decided to use a hydrological model **in order to estimate and predict discharge at the mouths of the rivers considered by SHYFEM.**

Which hydrological model

Requirements of the model for ARPA-FVG:

- used/tested successfully according to the literature;
- good and solid support by the developers;
- high spatial resolutions;
- *fully-distributed* model to represent terrain heterogeneity;
- calibration for the local territory;
- open source;
- representation of aquifers/groundwater (deeper than 2 m);
- forcings: meteorological fields provided by the atmospheric model WRF (*Weather Reasearch and Forecasting* model).

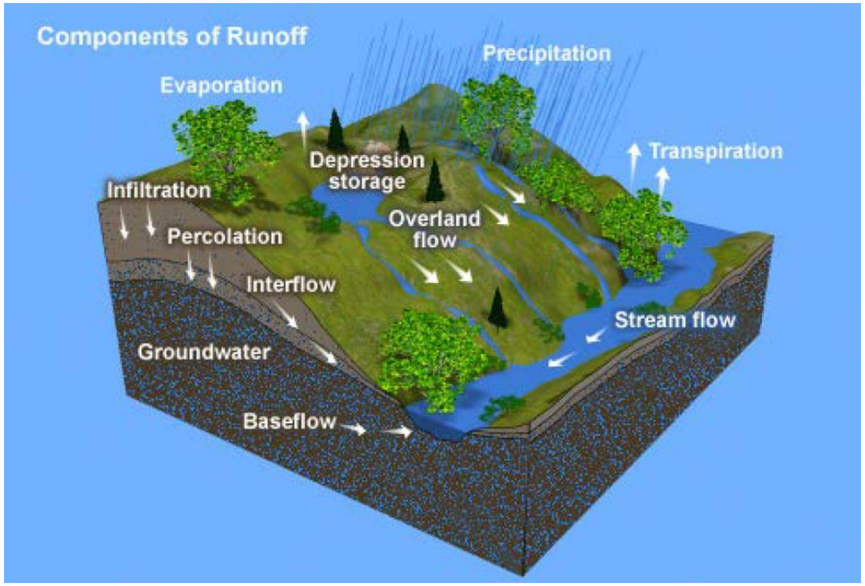
  - ARPA-FVG experience of more than 20 years;
  - WRF outputs are the meteorological forcings of SHYFEM

  => for consistency, WRF outputs should be used as meteorological forcings also for the hydrological model.

WRF-Hydro, ParFlow, HYPE,…

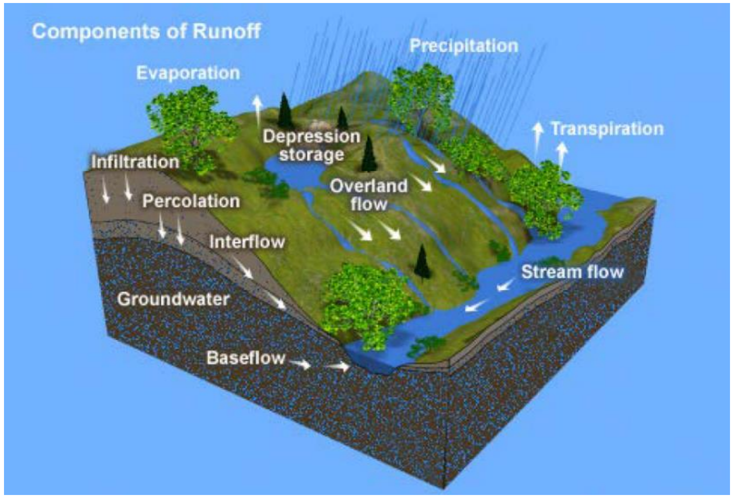
PART - I

WRF-Hydro: some general information



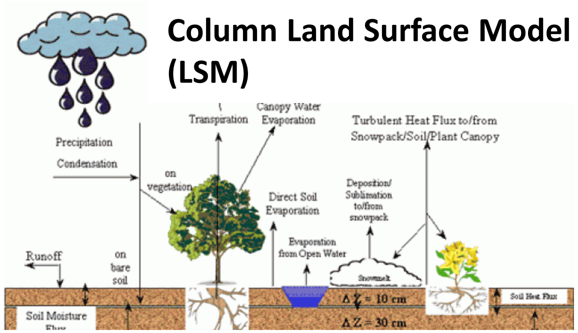
Main characteristics of WRF-Hydro

- developed as the hydrological modeling extension package of WRF (NCAR);
- used by NOAA as *National Water Model* (NWM) for US and, a configuration of NWM, by the National Weather Service;
- WRF-Hydro operates in two major modes:
  - *uncoupled* to an atmospheric model (*one-way coupling* or *offline mode*); [In ARPA-FVG: uncoupled mode with WRF]
  - *coupled* to an atmospheric model (*two-way coupling* or *online mode*);
- *processed-based* model;
- *fully-distributed* model;
- *multi-scale model capability*;
- *modularized component model*.



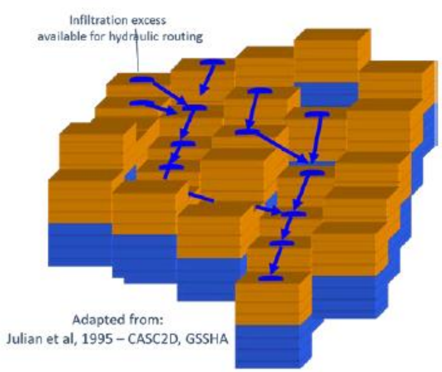
WRF-Hydro physics options

Column Land Surface Model (LSM)

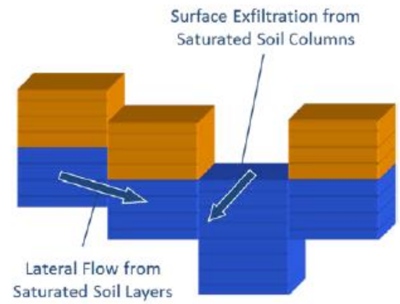


Such as the *Noah-Multiparameterization* (Noah-MP) scheme

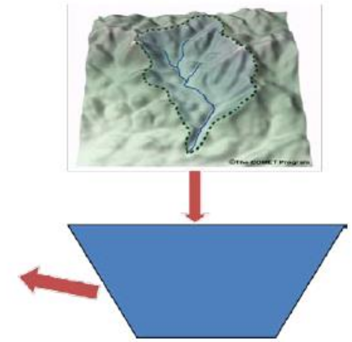
Overland Flow



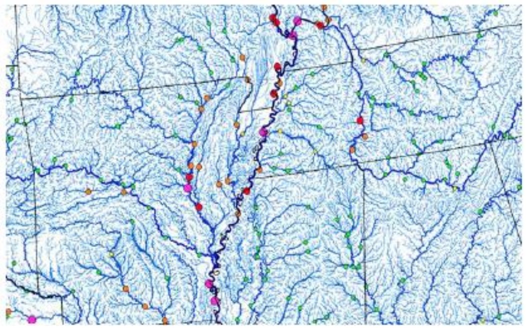
Lateral Subsurface Flow



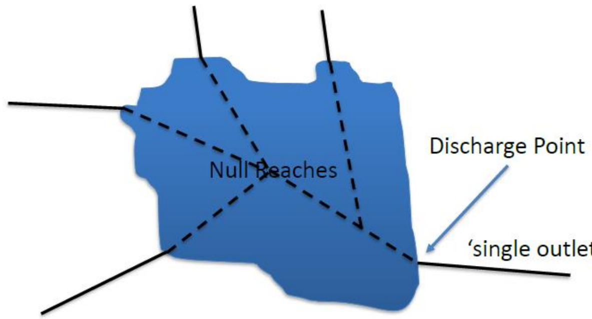
Simplified Baseflow Parameterization



Channel Hydraulics



Lake/Reservoir



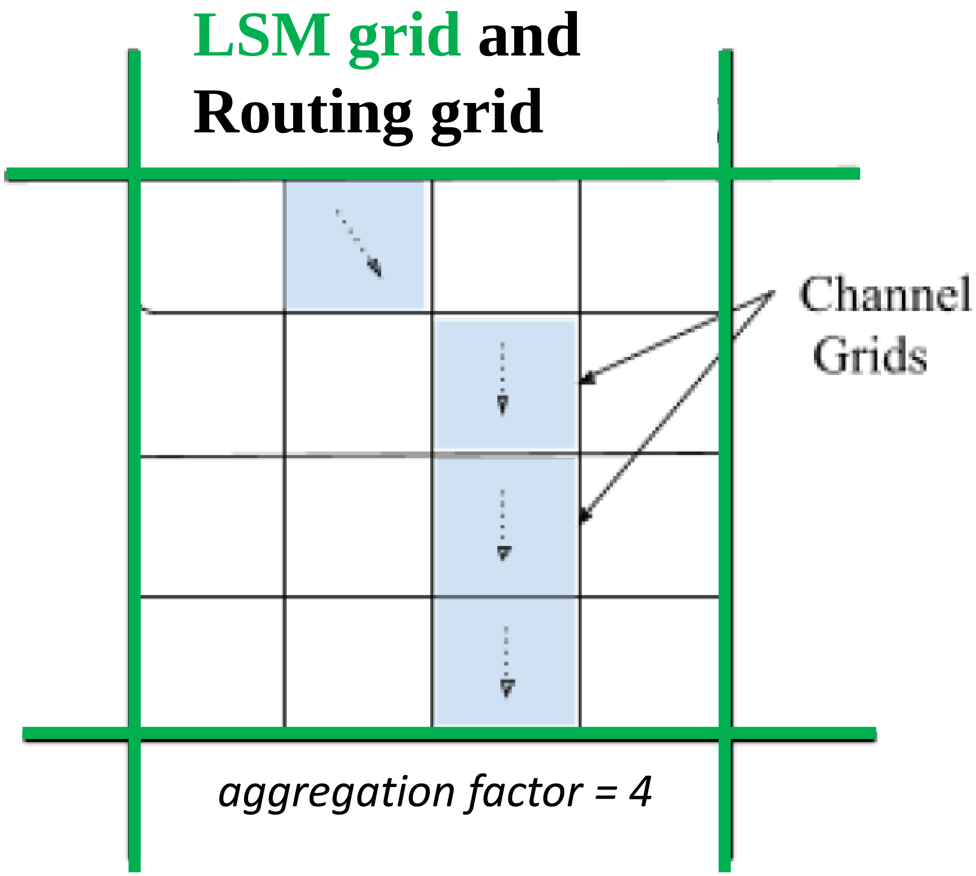
WRF-Hydro spatial description

WRF-Hydro is a multi-scale model:

- LSM grid
- Routing grid

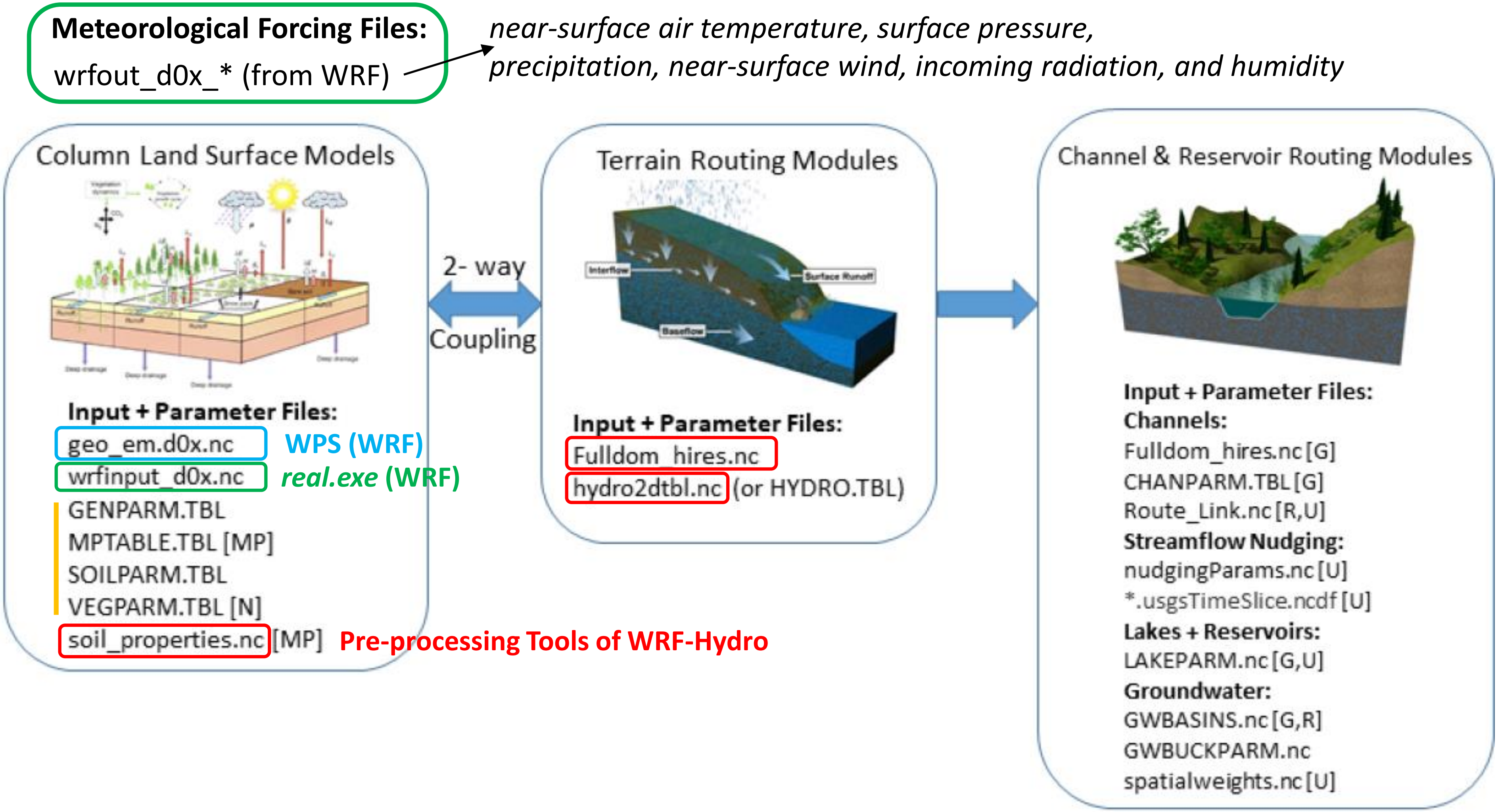
related by the *aggregation factor*

Schemes of overland and subsurface flows and channel routing operate on the routing grid thanks to the **subgrid *disaggregation* and *aggregation* routines**.



WRF-Hydro: input data

- files describing the model domain, parameters, initial conditions and forcings;
- the set of required files depends on the model configuration.



WRF-Hydro: output data

Some primary output **variables** of WRF-Hydro:
latent heat flux, surface evaporation components, soil moisture, soil temperature, surface runoff, channel discharge.

The output variables are provided as:

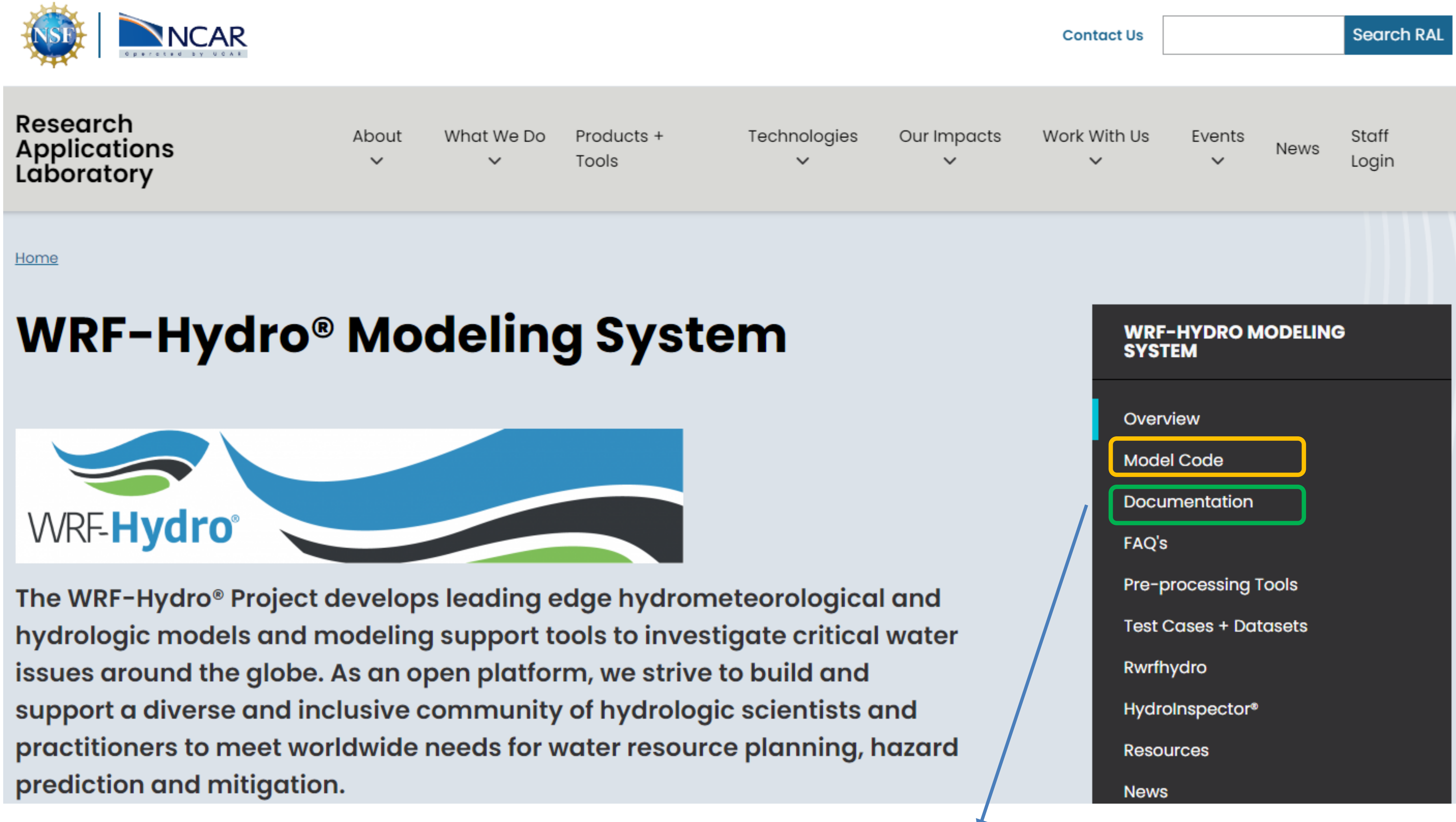
- Raster data in gridded output files on:
  - LSM grid (for LSM outputs)
  - Routing grid (e.g. discharge, soil moisture, surface head)

- Point data in output files with *features ID*:

  channel grids, lakes, gage stations.

Some practical information

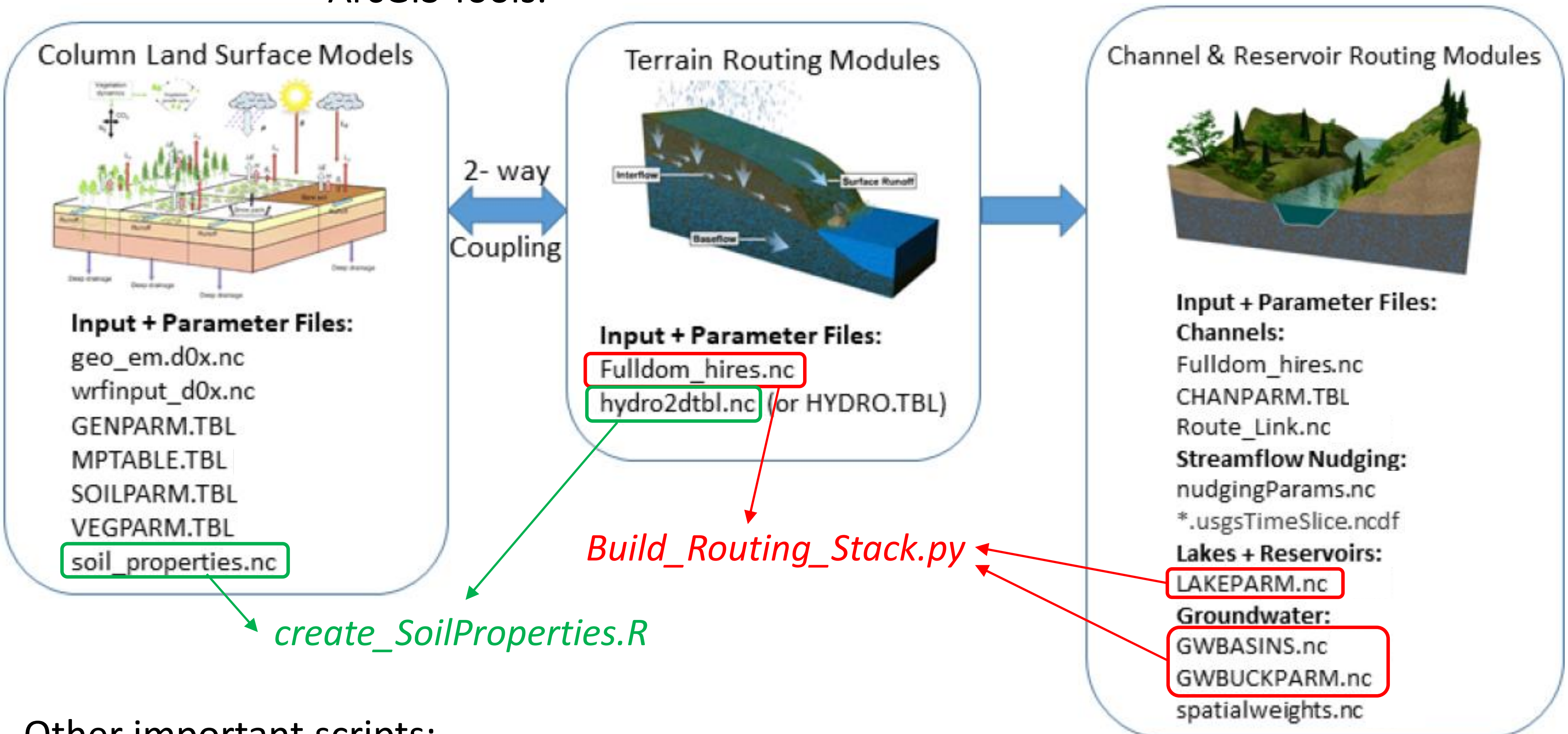
Homepage: https://ral.ucar.edu/projects/wrf_hydro



+ previous versions, GitHub repository, Users' Forum

Pre-Processing Tools

- Scripts in python, R, and NCL;
- ArcGIS Tools.



Other important scripts:

- *create_wrfinput.R* or *Create_wrfinput_from_Geogrid.py* => wrfinput_d0x.nc
- *WRF2WRFHydro_generate_weights.ncl* and *WRF2WRFHydro_regrid.ncl* => wrfout_d0x_*

Pre-Processing Tools: input files

Required input files:

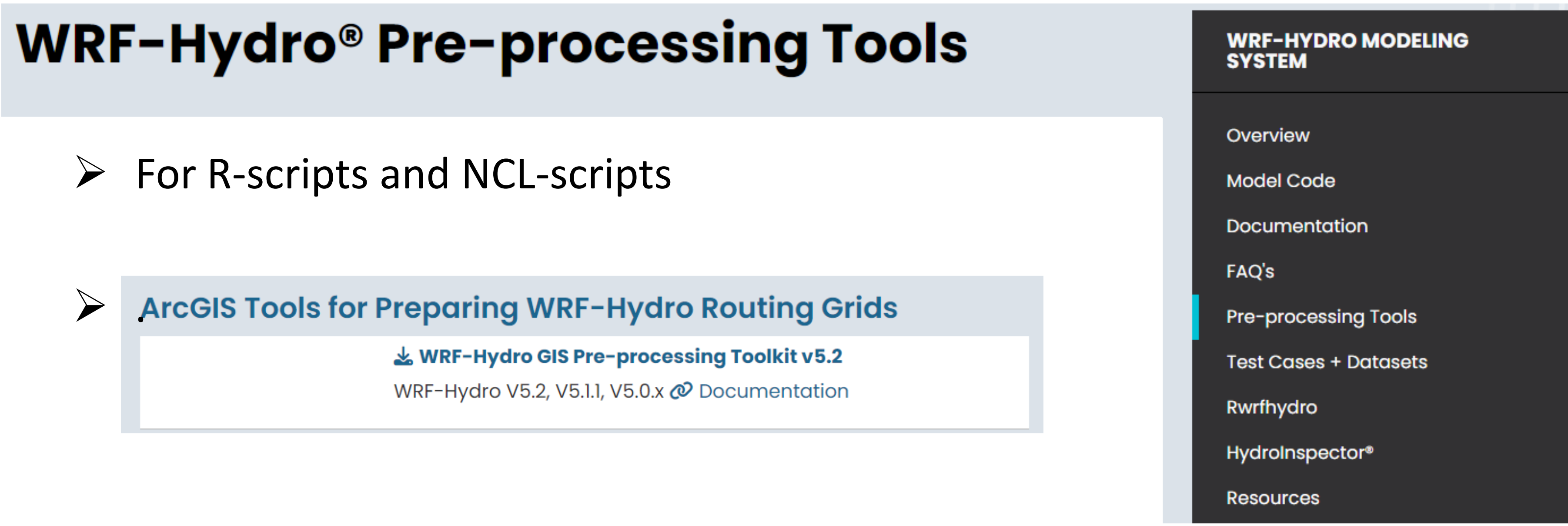
- the *geogrid file* (i.e. geo_em.d0x.nc produced by the WPS of WRF);
- a high-quality and high-resolution terrain data set*;* preferably, this data set should be a *“hydrologically conditioned” digital elevation model* (DEM).

Supplementary input files:

- tabular ascii file with the stream gage locations;
- shape file with lakes and reservoirs.

Some practical information on WRF-Hydro Pre-Processing Tools

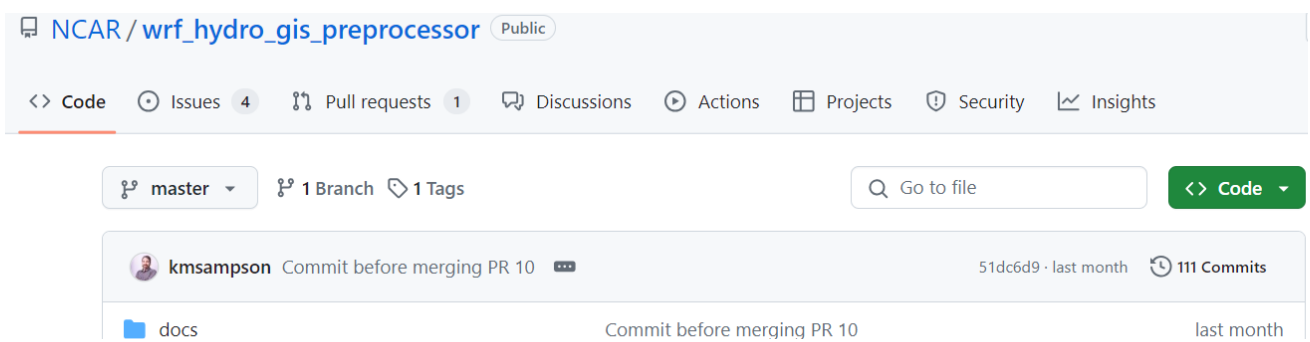
Webpage: https://ral.ucar.edu/projects/wrf_hydro/pre-processing-tools



- For R-scripts and NCL-scripts
- ArcGIS Tools for Preparing WRF-Hydro Routing Grids

GitHub repository: https://github.com/NCAR/wrf_hydro_gis_preprocessor/tree/master

- For Python-scripts



Hydrologically conditioned DEMs

DEMs (in general):

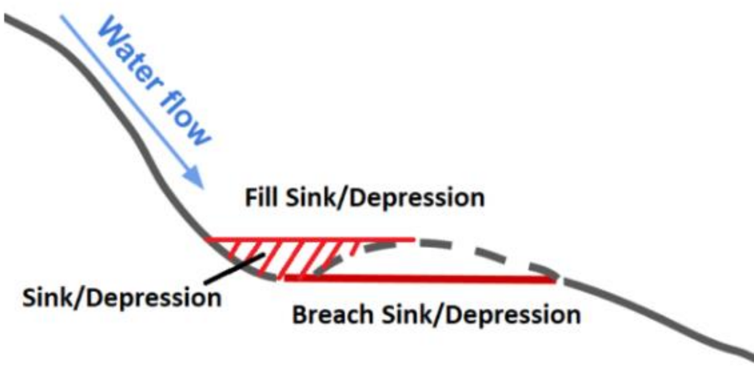
- digital representation of the terrain;
- various coverages and different resolutions;
- example: one typical product used by WRF is the terrain data derived by the *Shuttle Radar Topography Mission* (SRTM).

Hydrologically conditioned DEMs (for hydrological models):

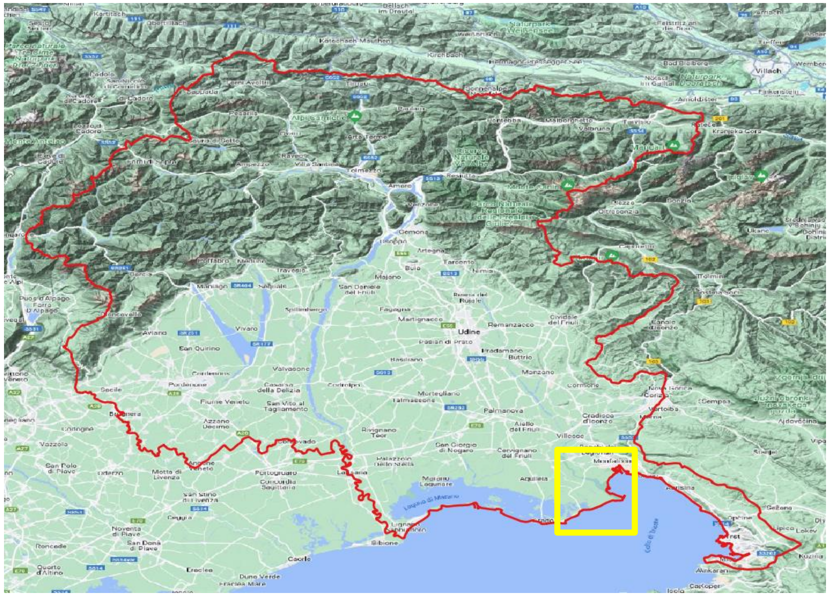
- post-elaborated DEM to reach a “hydrological connectivity”, i.e. topography is reconditioned to ensure continuity in flow directions;
- example: *Conditioned DEM* (ConDEM) of HydroSHEDS.

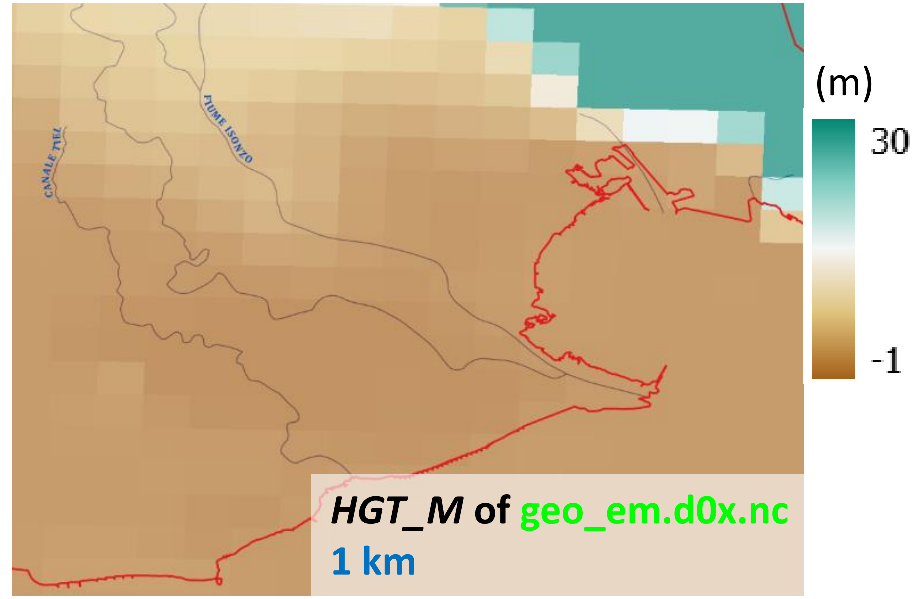
In WRF-Hydro:

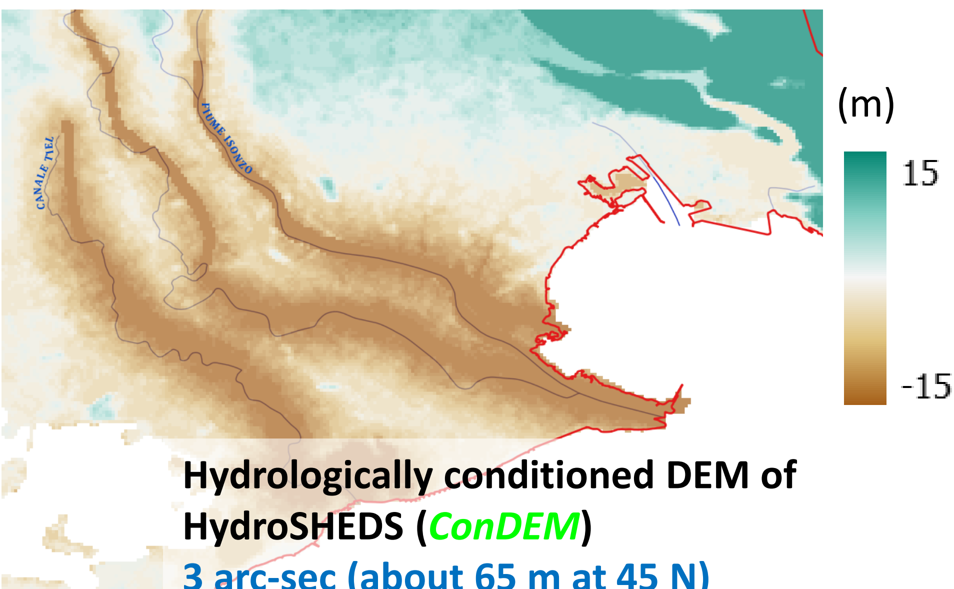
- the hydrologically conditioned DEM is regridded to the routing grid;
- algorithms in order to ensure the hydrological connectivity:
  - pits/depression filling;
  - depression breaching.

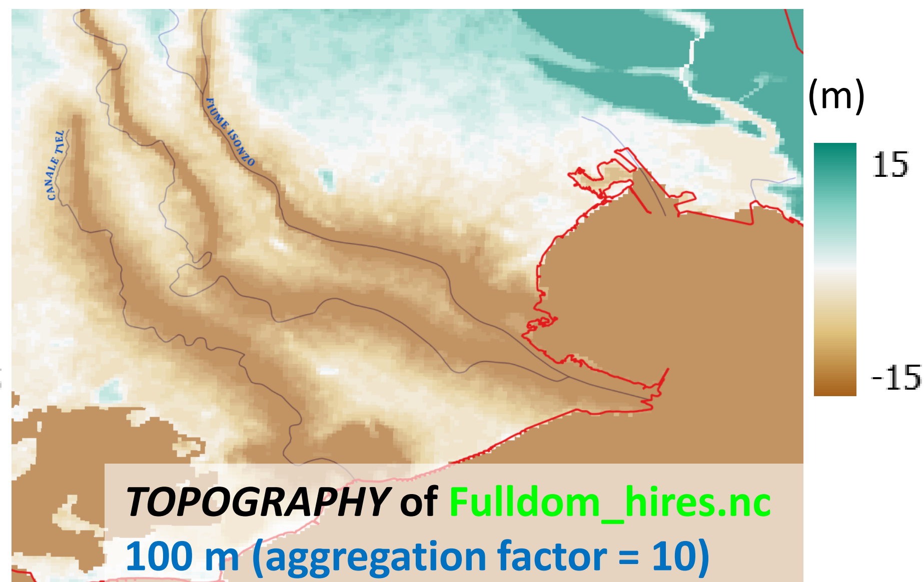


Terrain data sets









Spin-up, calibration, validation

Depending on the objective of the study, the length of the simulation (a long simulation or for specific events), the considered domain (one or more basins) and so on.

The **spin-up simulation** can vary according to:

- local climate, e.g. drier climate generally takes a longer time than a wetter climate to reach the equilibrium;
- variable of interest (a fast stabilizing state like shallow soil moisture vs. a slower evolving state like deep groundwater);
- soil types, e.g. sands have shorter memory than clays.

In general: given the long memory of some variables, such as deep soil moisture and groundwater processes, the spin-up time should be of some years (depending on the study objective).

Spin-up, calibration, validation

Calibration

- It can be performed for one or more hydrological basins.
- It is usually performed with respect to one variable (e.g. discharge).
- It can be performed for short events or long periods (it is a good practice calibrating the model to a wide range of conditions).
- It can be performed automatically with specific algorithms (e.g. DDS-Dynamically Dimensioned Search, PEST) or manually (e.g. step-wise approach).
- It is important to start the calibration after the spin-up (first, a long spin-up simulation with default values is run, then the restart files of WRF-Hydro are used for a warm start of the calibration run).
- The sensitivity parameters can be global (i.e. with a fixed value for the entire domain) or pixel-based parameters (i.e. spatially-distributed parameters); the first ones are usually defined in tabular files (*.TBL) while the second ones are defined in netcdf files.

Spin-up, calibration, validation

Validation

It is the process to estimate the model performance by comparing the calibrated hydrological model outputs with those observations that were *not* used before for the calibration process.

Performance metrics

The comparison between the model results and the observations (for calibration and validation) is based on the computation of the so-called "*performance metrics*" such as:

- root mean square error (RMSE),
- bias (BIAS),
- percent bias (PBIAS),
- Pearson correlation coefficient (PCC),
- Nash-Sutcliffe Efficiency (NSE),
- Kling-Gupta Efficiency (KGE).

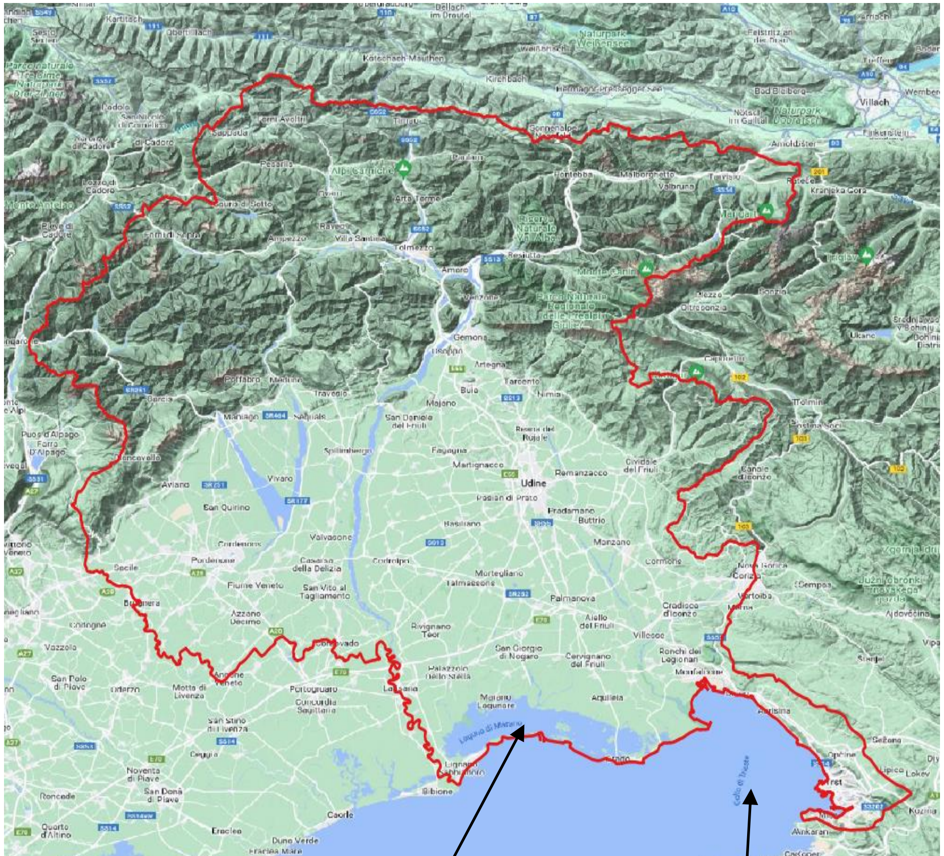
PART - II

Application of WRF-Hydro in ARPA-FVG

arpa FVG

FVG territory

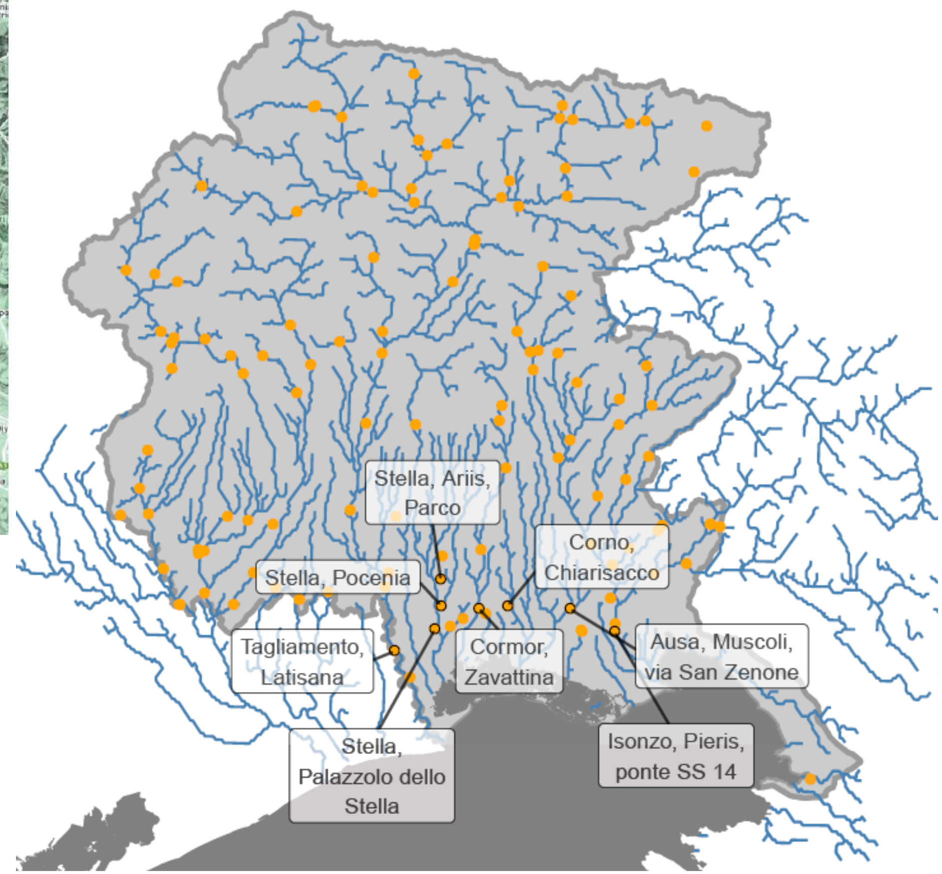
ICTP



Lagoon of Marano-Grado

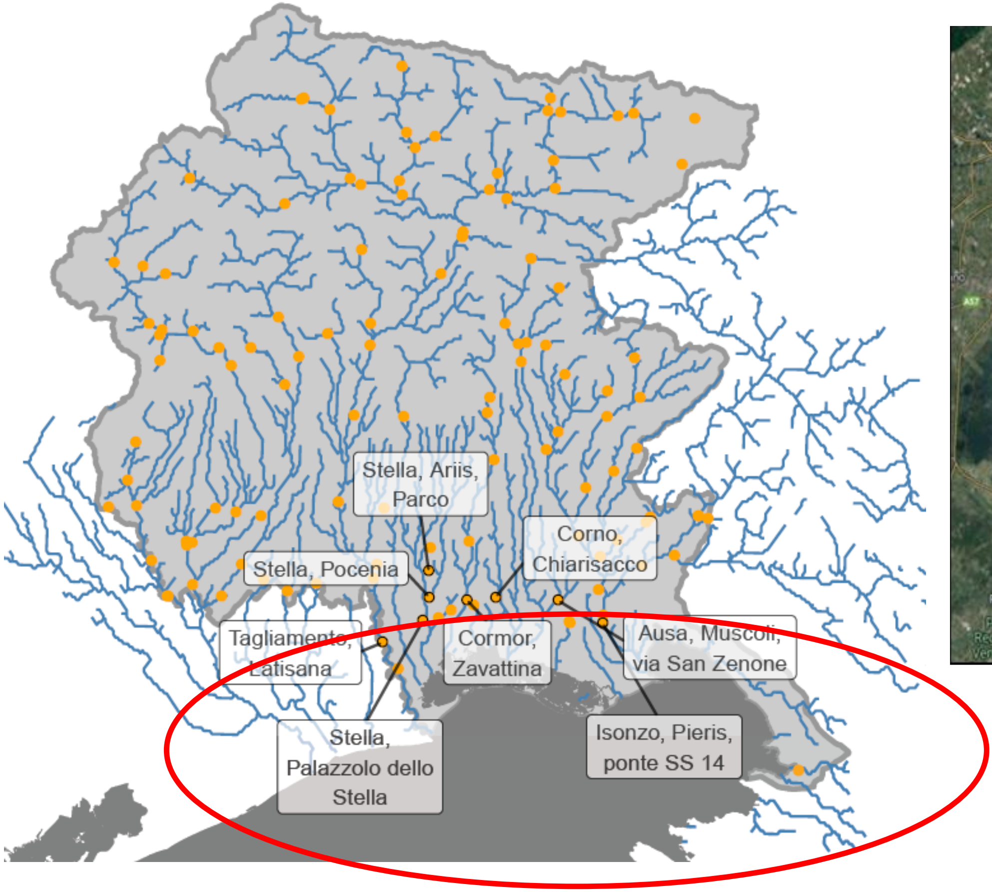
Gulf of Trieste

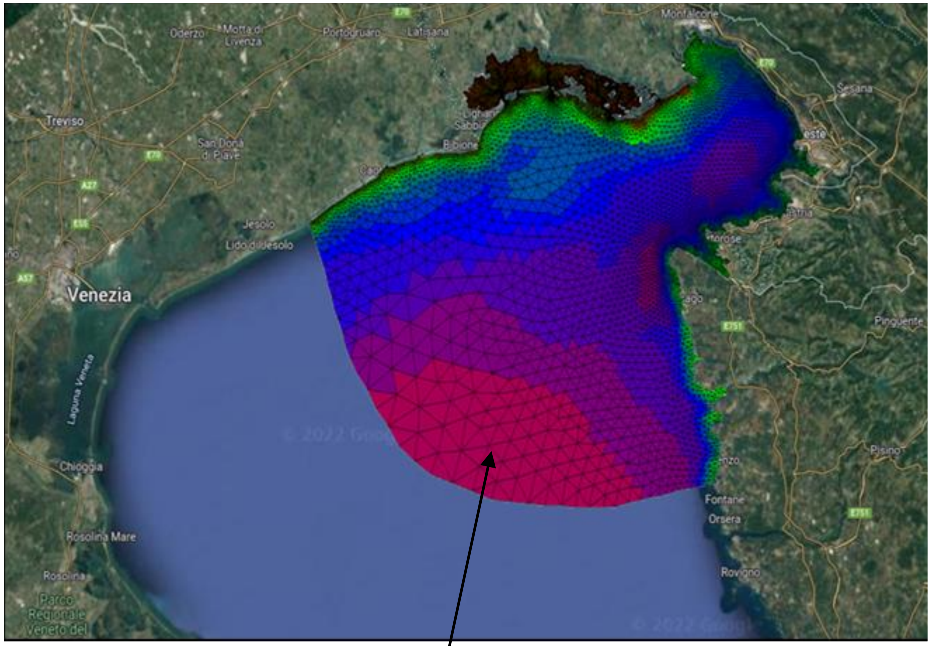
- Heterogeneity of regional territory
- Dense hydrological network



WRF-Hydro usage in ARPA-FVG

Objective: **estimating and predicting discharge at the mouths of the rivers considered by the hydrodynamic model SHYFEM.**





Domain and computational mesh of SHYFEM

WRF-Hydro setting

- Meteorological forcings: meteorological forecasts of WRF at 1 km of resolution.
- Routing grid of WRF-Hydro at 100 m of resolution (aggregation factor = 10).
- All physics options activated (i.e.: overland flow, subsurface flow, channel routing, lakes, groundwater).
- Hydrologically conditioned DEM: *ConDEM* of HydroSHEDS at 3 arc-sec.
- Period: daily simulations of 5 forcast days.
- Computational domain (number of grid boxes: 154 along x *times* 164 along y).
- Computational time (scalability tests): by using 2 nodes with 20 cores (i.e. 40 cores in total) => 5 days of simulation take about 5 minutes.

Calibration

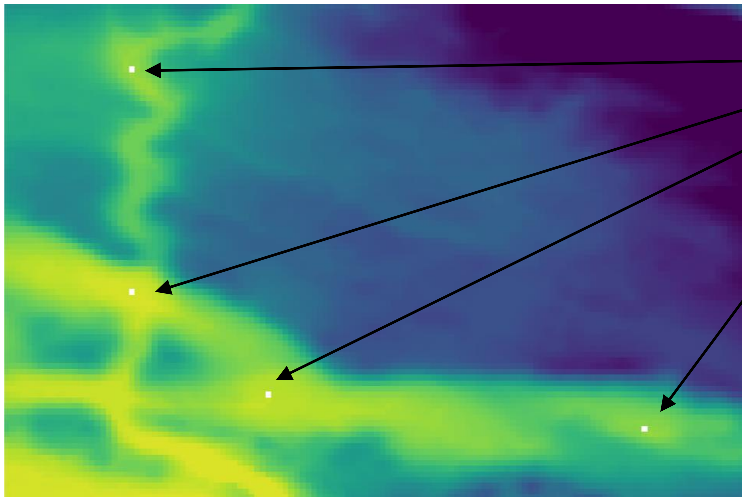
How will the calibration be done?

- By using meteorological forcings of a long WRF simulation (performed by ARPA-FVG for other purposes) with the same configuration of the WRF model used for meteorological forecasts.
- Manual calibration («step-wise approach»).
- Sensitivity parameters according to the literature, such as the infiltration of surface water into the soil (*refkdt*) and percolation parameter (*slope*).
- Inland gage observations (i.e. in mountainous/hilly areas).
- Temporal period (of a few years) that depends on the available measurements.
    - Discharge estimates derived from hydrometric levels;
    - Time series with discontinuities due to instrument changes, different reference levels, changes of the riverbeds, few human resources,…;
    - Time series of different length.

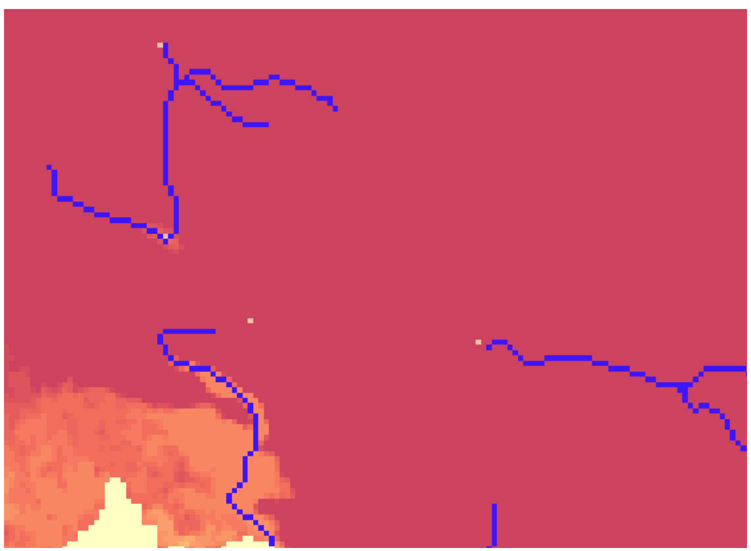
Finally, the calibrated model will be used for the hydrological forecasts.

Specific issues

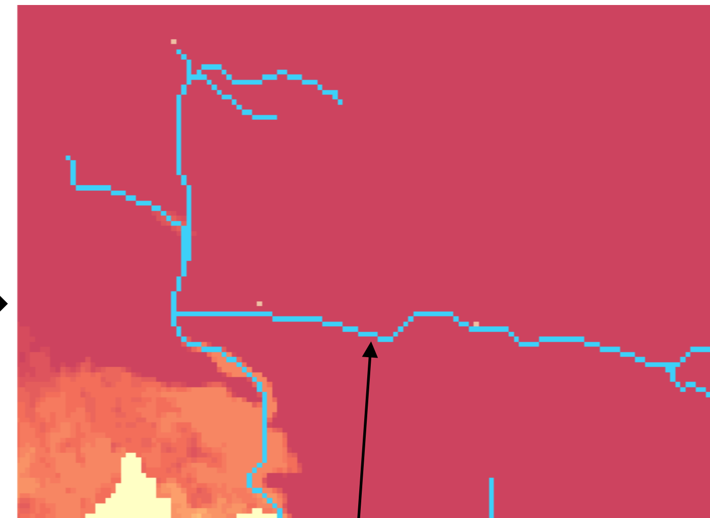
Removing pixels with «no-data» from ConDEM:



No-data in ConDEM



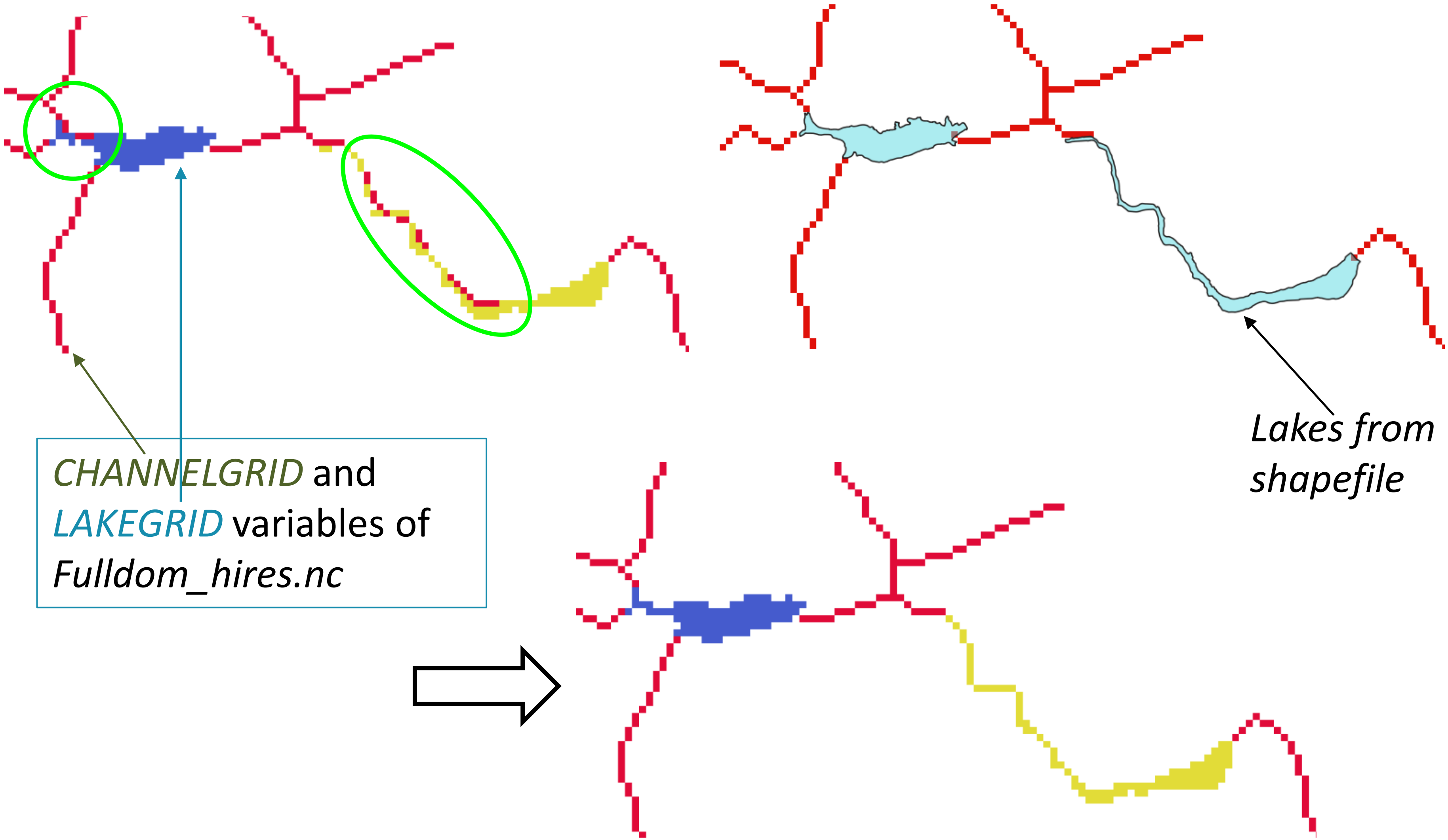
CHANNELGRID created with ConDEM including no-data pixels



CHANNELGRID created after removing no-data pixels

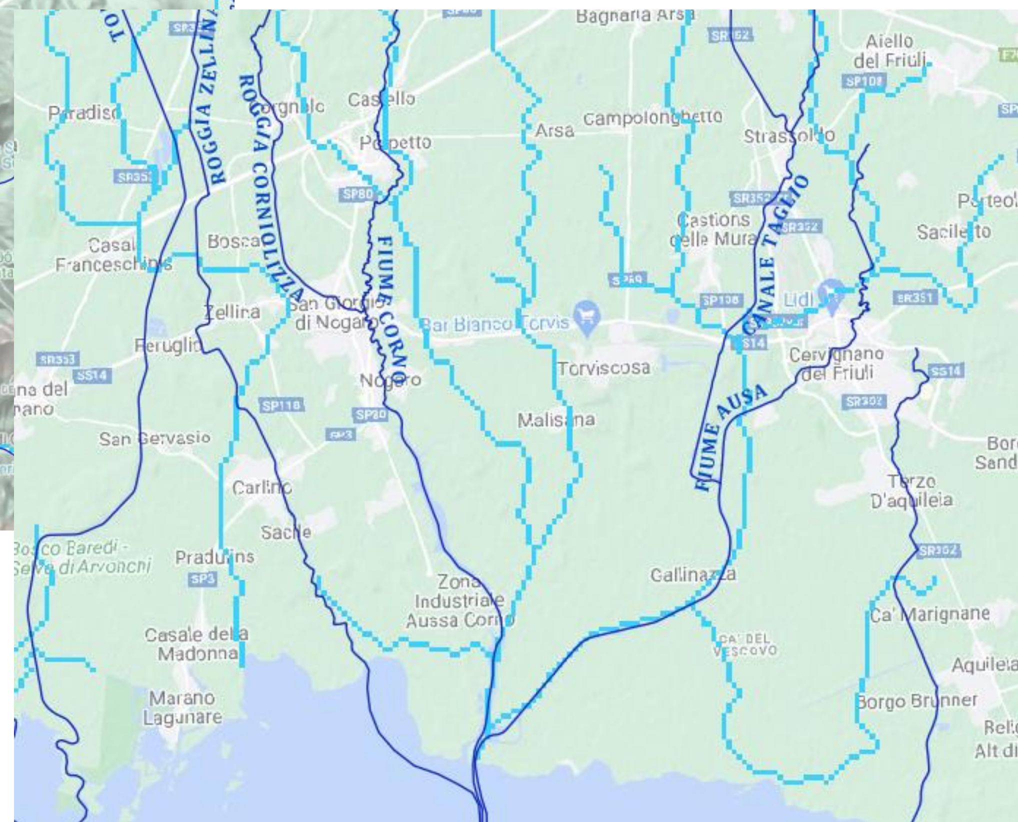
Specific issues

Removing artefacts from lakes:



Specific issues

CHANNELGRID: mountainous area vs. flat area

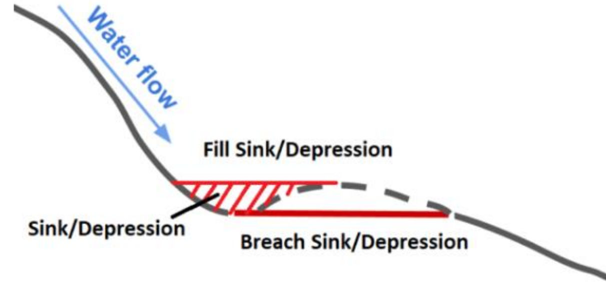


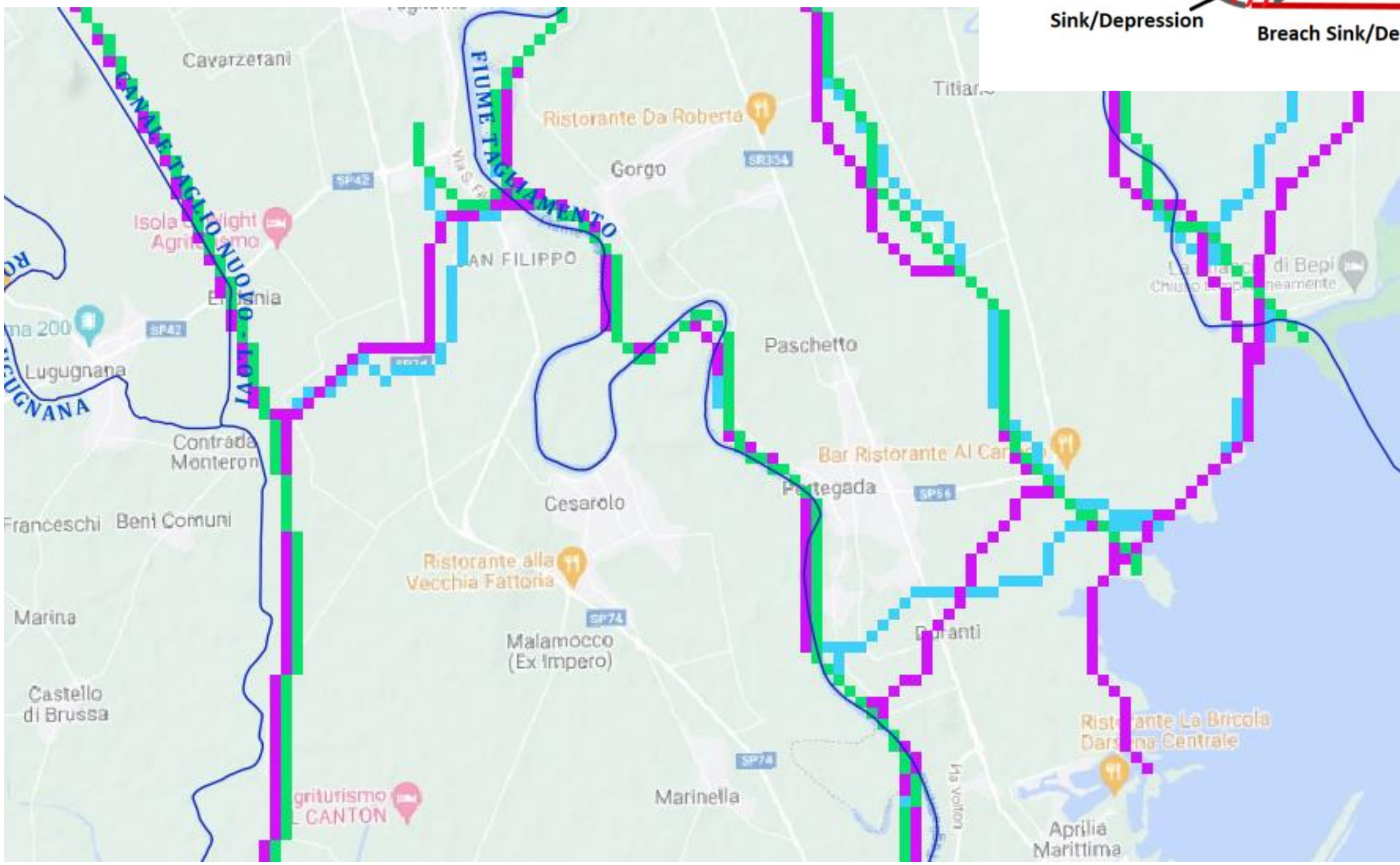
Real rivers, from shapefile

CHANNELGRID

Specific issues

Depression filling vs. depression breaching





Algorithms for depression filling

Algorithm for depression breaching

Summary

WRF-Hydro

- Main characteristics of the model, including required input data sets, main output variables and data rapresentation on different grids;
- Pre-processing phase of WRF-Hydro, with main pre-processing scripts and their input data;
- Hydrologically conditioned terrain data for hydrological models;
- Spin-up, calibration, and validation.

WRF-Hydro in ARPA-FVG

- Motivations for using a hydrological model in ARPA-FVG and the reasons for choosing WRF-Hydro;
- Model setting;
- Planned calibration;
- Some specific issues.

Thank you!!!

Useful links

WRF-Hydro homepage: https://ral.ucar.edu/projects/wrf_hydro

Technical description - User guide of WRF-Hydro:
https://ral.ucar.edu/sites/default/files/public/projects/wrf-hydro/technical-description-user-guide/wrf-hydrov5.2technicaldescription.pdf

Pre-processing Tools:

- Webpage: https://ral.ucar.edu/projects/wrf_hydro/pre-processing-tools
- Github: https://github.com/NCAR/wrf_hydro_gis_preprocessor/tree/master
- WRF Hydro GIS Pre-Processing Tools Documentation: https://github.com/NCAR/wrf_hydro_arcgis_preprocessor/releases/download/5.2.0/WRFHydro_GIS_Preprocessor_v5_2_0.pdf

Jupyternotebook lessons: https://ral.ucar.edu/events/2022/wrf-hydror-carribean-training-2022

Bias (BIAS), with optimal value = 0:

$$BIAS = \bar{Q}_{sim} - \bar{Q}_{obs}$$

Percent bias (PBIAS), with optimal value = 0:

$$PBIAS = (\bar{Q}_{sim} - \bar{Q}_{obs})/\bar{Q}_{obs} \cdot 100$$

Root Mean Square Error

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(Q_{sim,t} - Q_{obs,t})^2}$$

Pearson correlation coefficient (PCC)

$$PCC = \frac{\sum_{t=1}^{N}\left[(Q_{sim,t} - \bar{Q}_{sim}) \cdot (Q_{obs,t} - \bar{Q}_{obs})\right]}{\sqrt{\sum_{t=1}^{N}(Q_{sim,t} - \bar{Q}_{sim})^2} \cdot \sqrt{\sum_{t=1}^{N}(Q_{obs,t} - \bar{Q}_{obs})^2}}$$

Nash-Sutcliffe Efficiency (NSE)

$$NSE = 1 - \frac{\sum_{t=1}^{N}(Q_{sim,t} - Q_{obs,t})^2}{\sum_{t=1}^{N}(Q_{obs,t} - \bar{Q}_{obs})^2}$$

Kling-Gupta Efficiency (KGE)

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

where:

- r $[-1, 1]$ is the linear correlation coefficient (PCC);
- $\alpha = \sigma_{sim}/\sigma_{obs}$ $[0, \infty)$, where σ is the standard deviation (in m^3/s), is a measure of the flow variability error;
- $\beta = \bar{Q}_{sim}/\bar{Q}_{obs}$ $[0, \infty)$ is a bias error.

Thus:

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\bar{Q}_{sim}}{\bar{Q}_{obs}} - 1\right)^2}.$$

KGE can range from $-\infty$ to 1 where [Knoben2019 and Rummler2019]:

- KGE = 1 (optimal value) indicates perfect skill of the model, like NSE = 1;
- KGE > 0 indicates a good model performance, generally (*);
- KGE < 0 indicates a bad model performance, generally (*);