A comparative analysis of physics-based and machine learning methods for sustainable aquifer management in the Emilia-Romagna region (Italy)

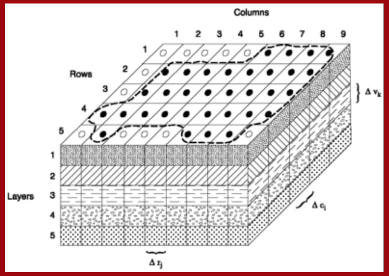
EGU25-660

I. Delfini[1], D. Zamrsky[2], A. Montanari[1]
1. Alma Mater Studiorum – University of Bologna, DICAM. Bologna, Italy
2. Department of Physical Geography, Utrecht University. Utrecht, The Netherland

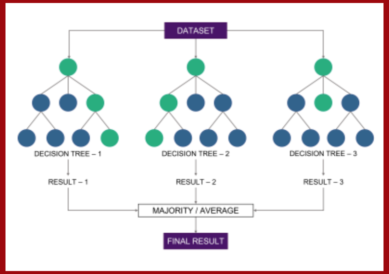
Vienna, 1 May 2025

Objectives

Predict the possible **evolution** of **groundwater** conditions in a portion of the **Emilia-Romagna** region (Italy) under the effects of **climate change** and **human activity**.

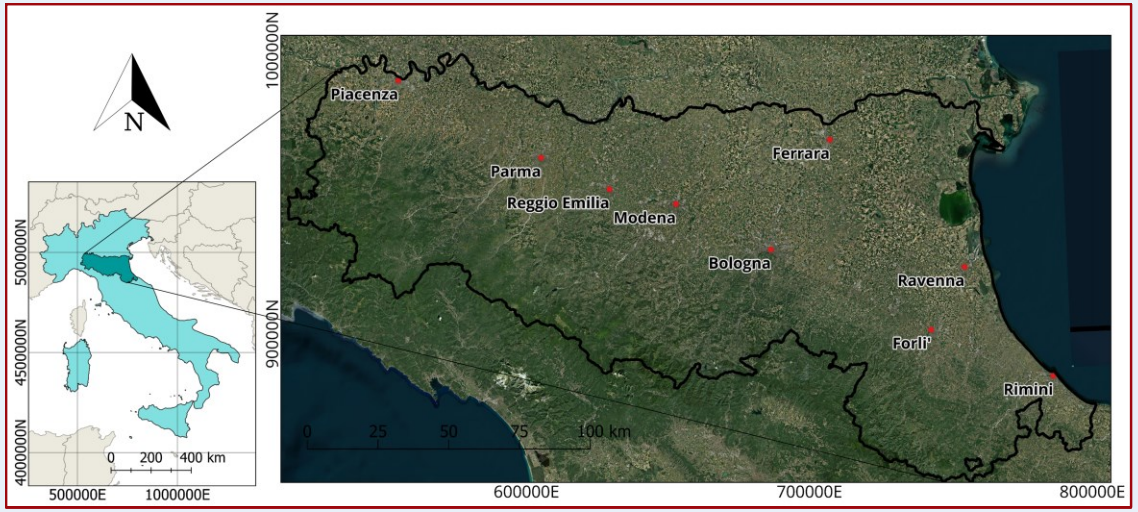


MODFLOW 6



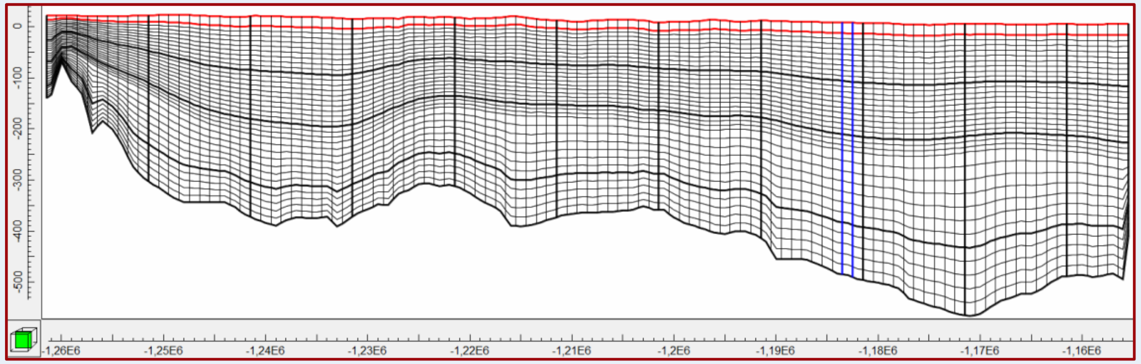
RANDOM FOREST

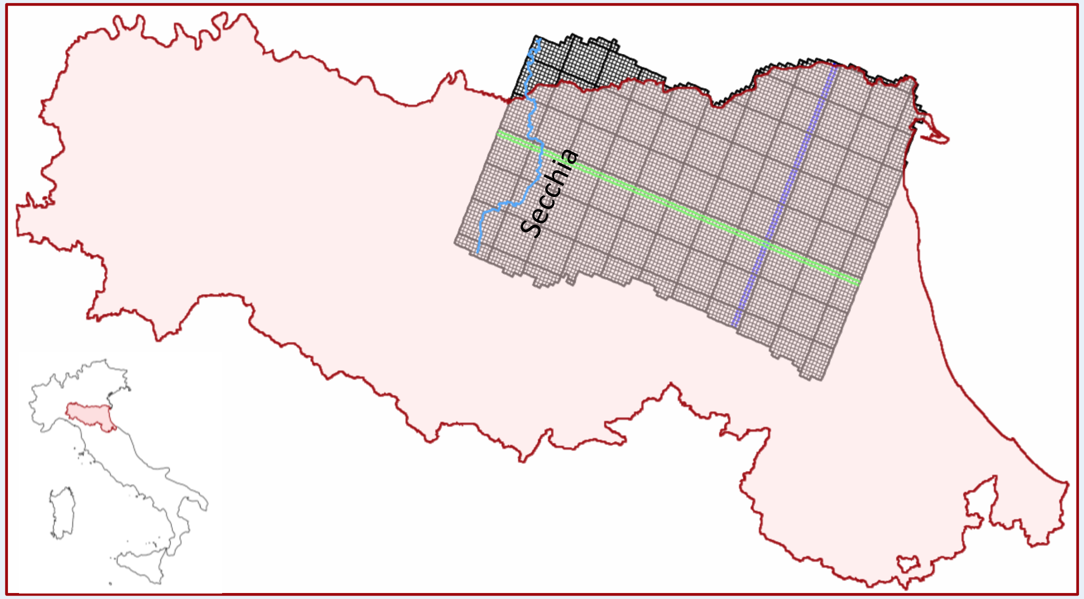
- **Compare** their **performance** in simulating **historical** and **future** groundwater head values over the **same study area**.
- Assess the **resilience** of the regional multi-layered aquifer system to **long term drought** events



Study area and data

- Portion of the **Emilia-Romagna** region (Italy).
- Simulation period: **2002-2018**. This multi-year simulation interval allows the representation of hydrometeorological seasonal variations.
- Extension: **7000 km^2**.
- Spatial resolution: **1 km x 1 km** cells, **35 layers** of variable thickness.
- **Aquifer Groups A** (superficial) and **B** (deep).



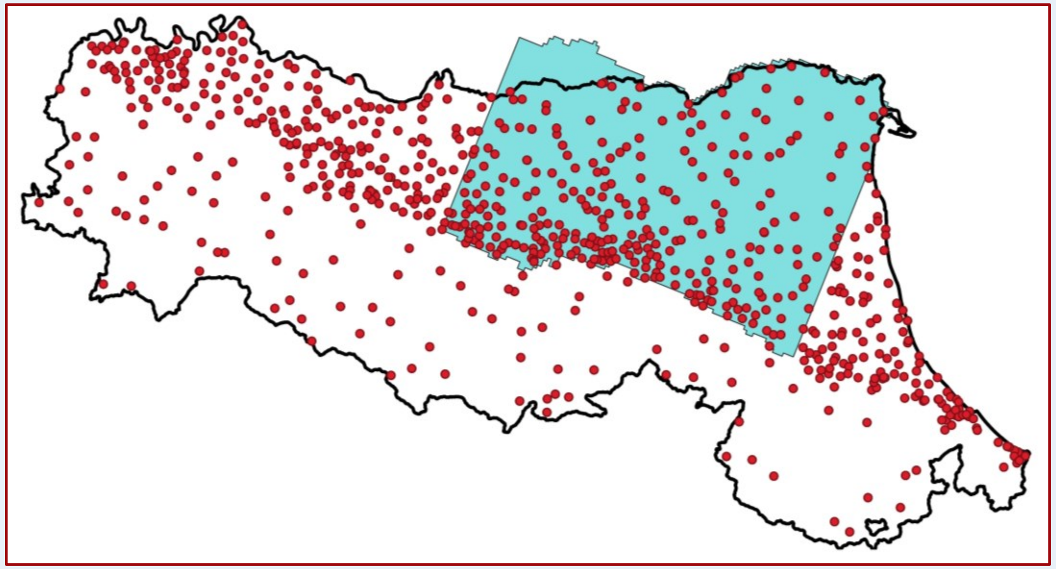


Data are mainly available from:

A. a **Modflow application** to the whole groundwater flow system of **Emilia-Romagna** by **ARPAE** (the Regional Agency for Prevention, Environment and Energy):
   - **geometry** and **hydrogeologic properties** of the aquifers;
   - **extraction rates**;

B. freely accessible datasets on the **Emilia-Romagna Region** and **ARPAE repositories**:
   - **rainfall** at several raingauges
   - **water stage** in the main rivers.

Numerical model calibration



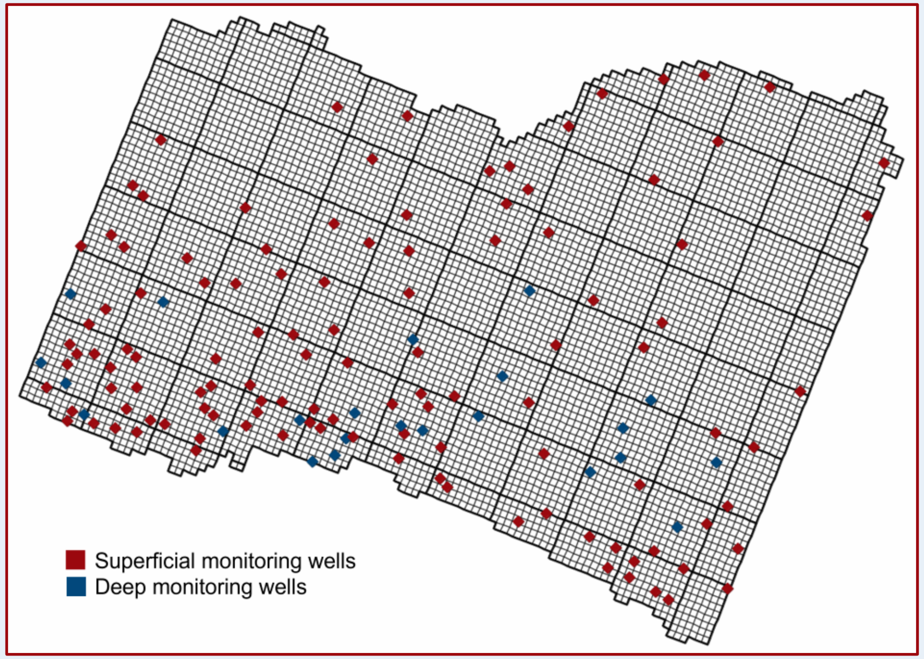
- Comparison of **simulated** and **observed groundwater head values** (2010-2018).
- **127** monitoring wells from the regional monitoring network.
- **2 observations** per **year**

~2260 observations
available for calibration

- **Parameters adjustments** to improve the match between observed and simulated groundwater head values.
- Performance evaluation: $\mathbf{R^2}$.
- **Groundwater budget consistency** with (i) **seasonal variations** and (ii) groundwater budgets from the **regional model** (Arpae).

Random Forest model setup

- Based on **groundwater head observations** from the regional monitoring network.
- Simulation of **groundwater head anomalies**.
- Distinction between **shallow** and **deep** wells, considered **separately** in the calibration process.
- Setup for training and testing datasets:
  - ***Training***: data from 2010 to 2015 *(~70%)*;
  - ***Testing****: data from 2016 to 2018 (~30%)*.



TYPE/GROUP	VARIABLES
Hydrogeology	Vertical and horizontal hydraulic conductivity Specific storage Specific yield
Topography	DEM
Water	Vertical and horizontal distance to the nearest water body
Climatic	Precipitation Average temperature Potential evapotranspiration
Land cover	Land use coefficient (0-1)
Coordinates	X utm Y utm
	Groundwater abstraction
	Time

- **Hyperparameters**:
  - *ntree*: number of trees to grow in the forest;
  - *mtry*: number of features randomly sampled at each split;
  - *nodesize*: minimum node size.
- **Different values** of *ntree*, *mtry,* and *nodesize* are tested to assess the combination that provides the **best model performance**.

Goals:

1) Estimate the **effects** of a **precipitation reduction**;
2) Get an insight into the **combined effects** of **variations** in **natural** and **artificial stresses** on the regional aquifer system;
3) **Compare** the groundwater head **predictions** by the **two methods** under the **same** scenario **conditions**.

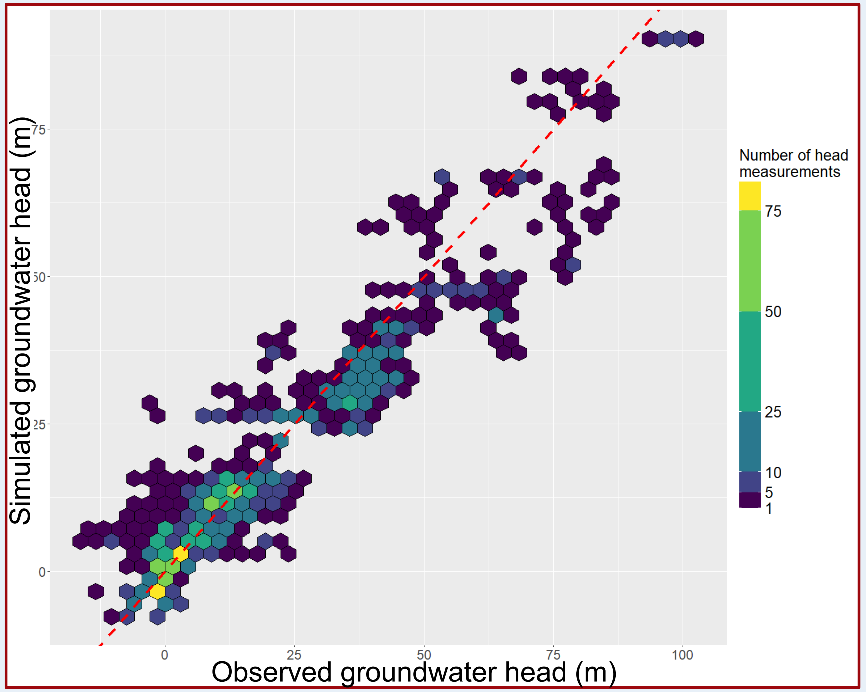
Month	Precipitation Rate Reduction (%)
January	16.3
February	12.4
March	17.1
April	20.5
May	11.8
June	11.4
July	12.8
August	23.5
September	15.1
October	12.7
November	14.0
December	21.0

From **statistical analysis of meteorological droughts** in Emilia-Romagna over the last two centuries.

Scenarios (2019-2030)	Characteristics
Reference (R)	Time dependent input parameters are **constant** at the seasonal scale, and equal to their average over the years 2014-2018
1	Monthly average **precipitation reduction**
2	Monthly average **precipitation reduction** + Groundwater **extraction rates increment** (+20%)
3	Monthly average **precipitation reduction** + Groundwater **extraction rates reduction** (-20%)

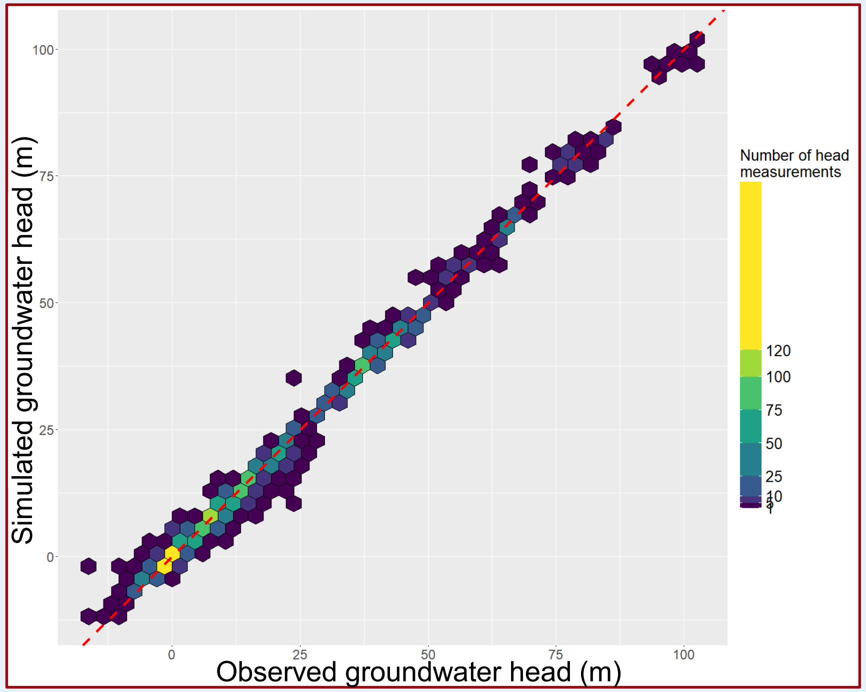
Calibration plots

MODFLOW



R^2 = 0.887
RMSE = 7.492 m
MAE = 5.216 m

RANDOM FOREST - ALL WELLS

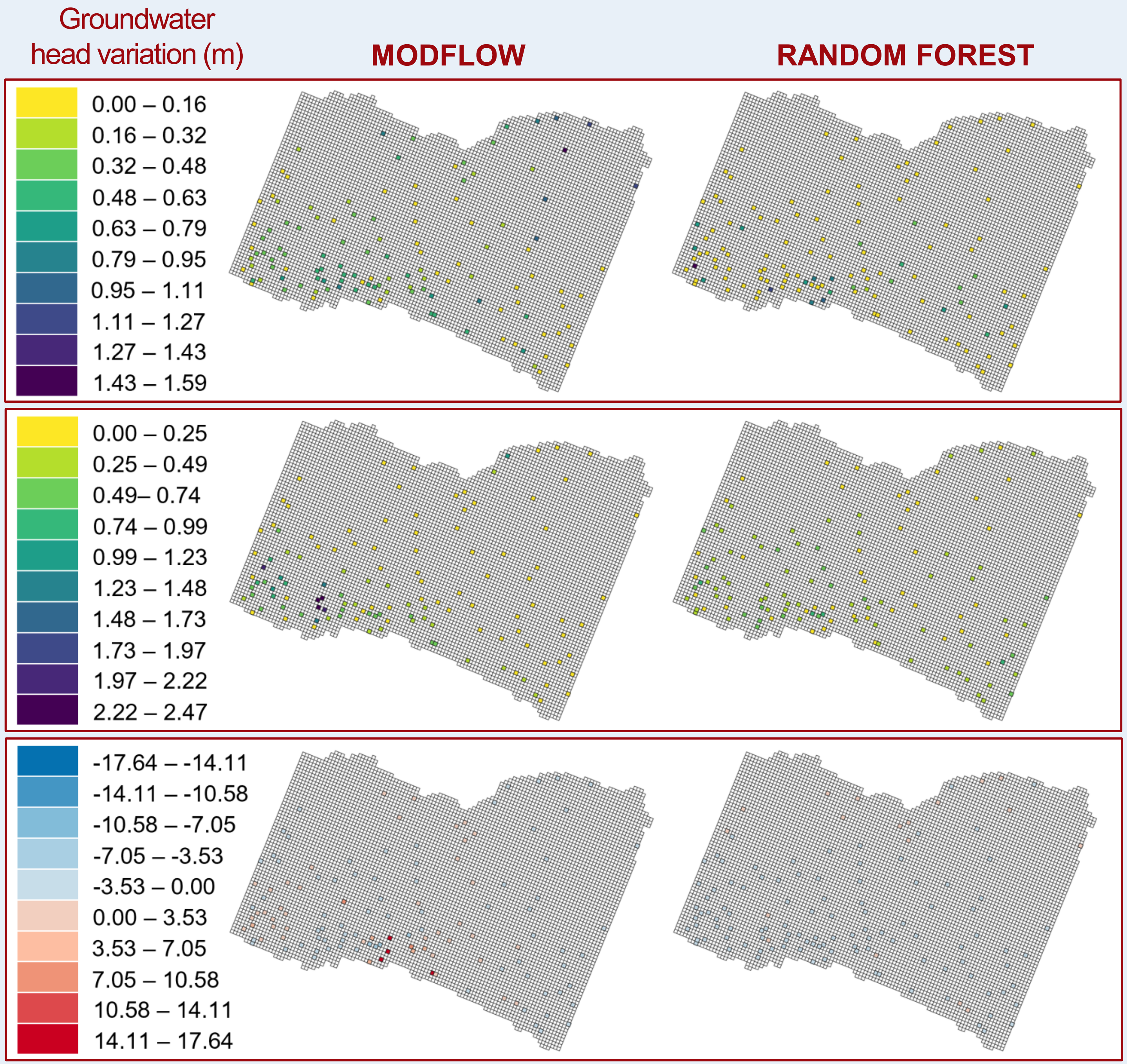


R^2 = 0.995
RMSE = 1.555 m
MAE = 0.913 m

The **random forest** algorithm produces **less scatter** in replicating **historical** groundwater head values.

Random forests, **trained** on **observed data** to **minimize simulation error**, are more **flexible** than physics-based models and can potentially **better replicate historical** groundwater levels.

Scenario analysis



Difference between scenario R and scenario 1

↓

Precipitation reduction

Difference between scenario 1 and scenario 2

↓

Groundwater abstraction increment

Difference between scenario 1 and scenario 3

↓

Groundwater abstraction reduction

- **Groundwater head reduction** almost all over the model domain.
- Groundwater head decreases **up to 60 cm** are **barely concerning** in terms of the sustainability of aquifer exploitation.
- **Most critical impacts**: southern portion of the study area (also north-eastern for the numerical model).
- The combined effects of **increased** groundwater **abstraction** and **precipitation reduction exacerbate** the **stress** over the aquifers.
- **Reductions** are more **significative** for the **numerical model**.
- Groundwater head **increments** (blue) in 60% of the monitoring locations.
- Some areas still see a decrease in groundwater head (red).
- **Range of groundwater head variations**: narrower for the random forest algorithm.

Conclusions

▶ The **random forest** algorithm more **accurately** reproduces **historical** groundwater head values than the numerical model.

▶ The **numerical model** and the **random forest** provide **coherent** results, although the random forest is generally more optimistic:

- the regional aquifer system is projected to **suffer** the **consequences** of a potential **future precipitation reduction**;
- an **increment** in **groundwater abstraction** may **exacerbate** the effects of a precipitation reduction, particularly in the **southern** and south-western regions of the study area;
- a **reduction** in **extraction** rates may partially **compensate** the groundwater head reduction due to **precipitation deficit**.

▶ Identifying **areas most affected** by **precipitation** and **pumping** changes highlights the importance of considering both **regional** and **local scales** to design climate change **mitigation** and **adaptation** strategies.

▶ The concerning results in the **south-western** part of the study area highlight the need for a **more accurate modeling process** to better assess the consequences of the considered stresses on the aquifer system.

▶ Even though a model refinement could lead to slightly different results, it is likely that **mitigation** and **adaptation strategies** should focus the most in the **south-western** area.

Sharing is encouraged

Thank you for your time and attention

Ilaria Delfini

Department of Civil, Chemical, Environmental, and Materials Engineering (DICAM)
Alma Mater Studiorum – Università di Bologna

ilaria.delfini2@unibo.it

Sharing is encouraged

Thank you for your time and attention

Ilaria Delfini

Department of Civil, Chemical, Environmental, and Materials Engineering (DICAM)
Alma Mater Studiorum – Università di Bologna

ilaria.delfini2@unibo.it