EPFL

crealp
Centre de recherche sur l'environnement alpin
Zentrum für alpine Umweltforschung
Research center on alpine environment

Development of a Forecasting Tool for Groundwater Levels in Valais Using Advanced Computational Techniques

Romane Collin, Corinna Frank

Design Project in cooperation with CREALP, Research Center on Alpine Environment

June 4th, 2021

Environmental Sciences and Engineering

Groundwater Prediction could be an integral part of Risk Management in Valais

Groundwater is the main water stock in Switzerland.

The canton of Valais is facing reccuring...

- flooding (e.g. of agricultural fields)
- threats by rising water table on polluted sites

Prevention & sustainability

- Need for detection of early-stage changes in groundwater table as part of the risk management
- Current monitoring network of 320 stations
- Integrate **forecasting of groundwater levels** to anticipate events

Romane Collin & Corinna Frank

Romane Collin & Corinna Frank

Two objectives are to be adressed

Goal 1
Improve the understanding of how groundwater levels are behaving in Valais

Goal 2
Develop a groundwater level forecasting model

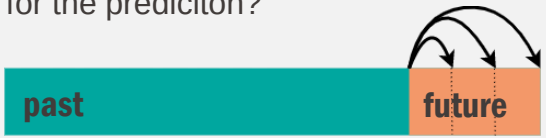
2-step Strategy towards a forecasting tool

1 Exploratory Data Analysis (EDA)

- Determine cross-correlations between groundwater levels and external variables like temperature and precipitation
- Auto-correlation
- Fourier Transform to find frequencies of the groundwater level signal
- Clustering of different types of behaviors

2 Machine Learning Model (ML)

- EDA tells us which elements are important for the forecasting
- Forecasting model based on: Random Forest Regressor
- Evaluate the prediction quality
- Which information is most valuable for the prediciton? Same as found in EDA?



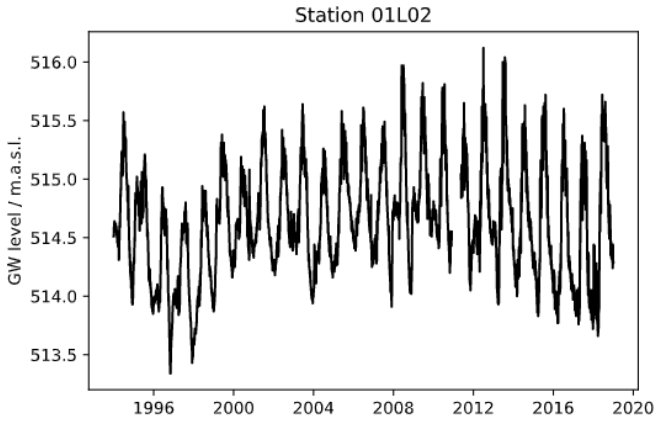
Elements impacting the Groundwater Level in Valais

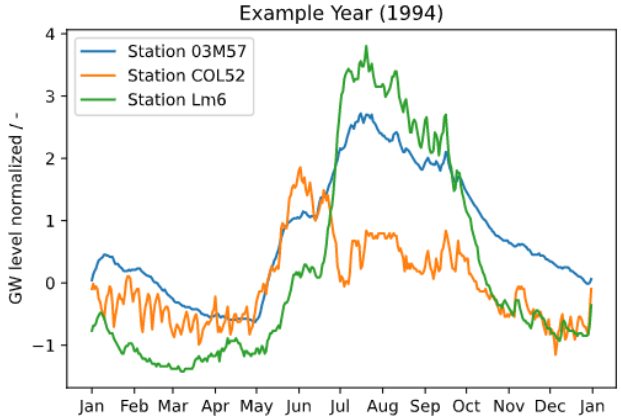
- Rhône discharge (pressure or mass exchange)
- Air temperature
- Rainwater
- Meltwater from snow & ice
- Topography
- Geology & soil
- Water withdrawal
- Land use, vegetation
- ...

Yearly patterns according to hydrological regime and seasonality.

Behavior varies between stations.

→ We will take a look at **25 years of data**.





Romane Collin & Corinna Frank

Data has been aggregated beforehand

Spatialized data over the canton is aggregated to 1 value per station.

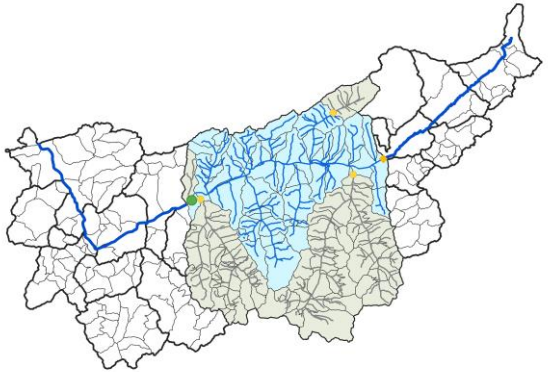
VAP

(volumetric available precipitation)

- m³/day
- Represents local available water from precipitation
- Physical elements:
  - rain
  - snowmelt delay
  - evapotranspiration

Q_{int}

(interpolated discharge)

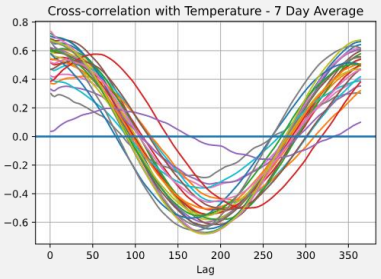


- m³/day
- Represents Rhône discharge at the height of the station
- Physical elements:
  - glaciers
  - snowmelt
  - precipitation

Romane Collin & Corinna Frank

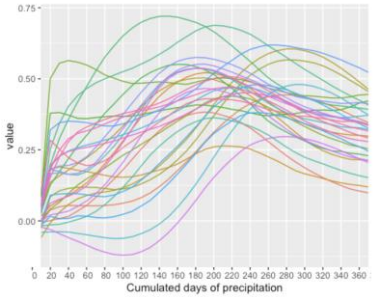
External variables are impacting Groundwater Levels on short & long scale

Air temperature



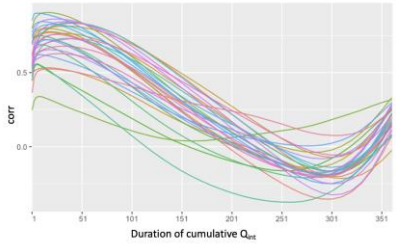
- Up to 100 days
- Gives information about underlying seasons & current weather.

VAP (volumetric available precipitation)



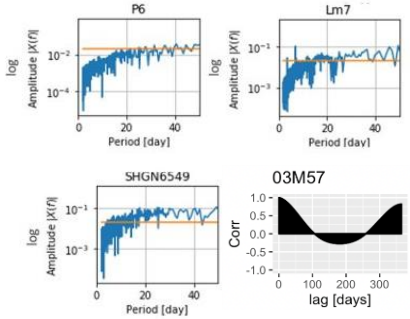
- three influences:
  - last month (20 days): low water period
  - Last 6 months: stock and release of snow
  - Annual seasonality (170 days)

Q_{int} (interpolated discharge)



- two patterns:
  - Short-term reaction (5-10 days)
  - Long-term reaction (20+ days)

Past Groundwater Level

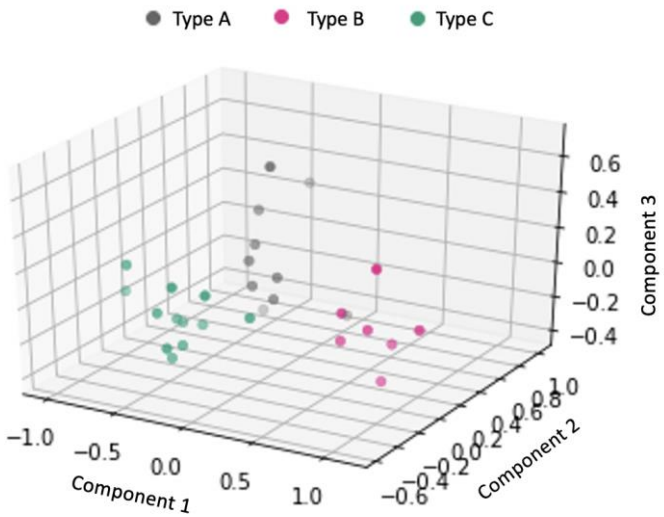


- 3 patterns from Fourier Transform: variability either noisy, smooth or in between
- Autocorrelation of groundwater levels up to 100 days

Romane Collin & Corinna Frank

Station behaviors can be clustered in 3 types

Type A	Type B	Type C
04S70	03M57	01L02
09M06	07E06	04I52
10E02	07G06	05X55
COL8	Lm5	08E6
Lm6	Lm7	COL52
Lm9b	Lm8	COL53
SHGN6543	RN14	Lm3
SHGN6545		Lm4
SHGN6546		P117
SHGN6549		P6
		P67
		S167



- Clustering using **K-means algorithm**: Unsupervised Machine Learning method which finds the number of groups autonomously
- Input: correlation values found in EDA

→ **3 clusters** were identified

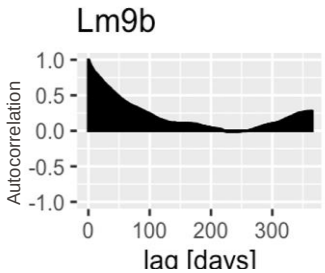
→ **What do they correspond to?**

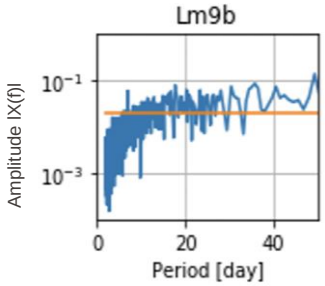
Romane Collin & Corinna Frank

Analysis of the Station Types

TYPE A

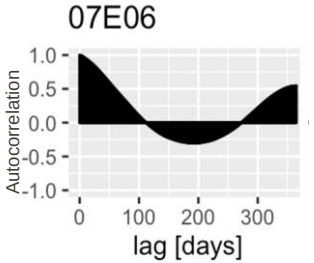
- Short-term variations are important

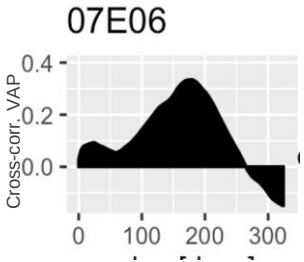




TYPE B

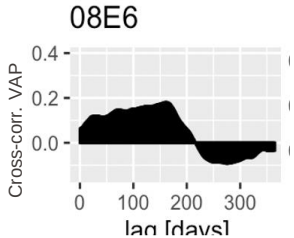
- Cyclic annual patterns are the most important

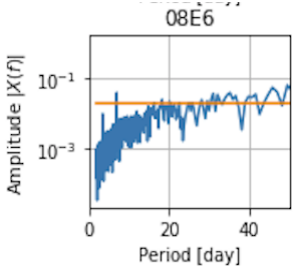




TYPE C

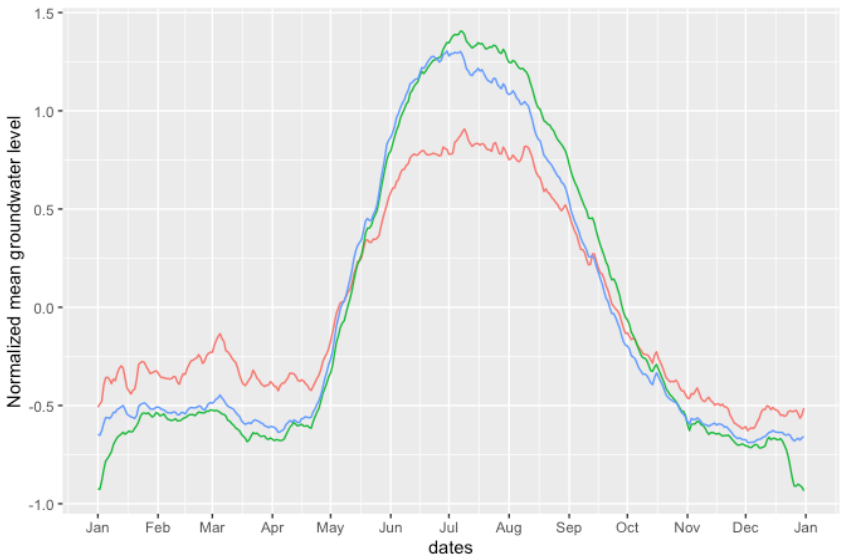
- Cyclic annual patterns are important
- Specificities of the last months as well

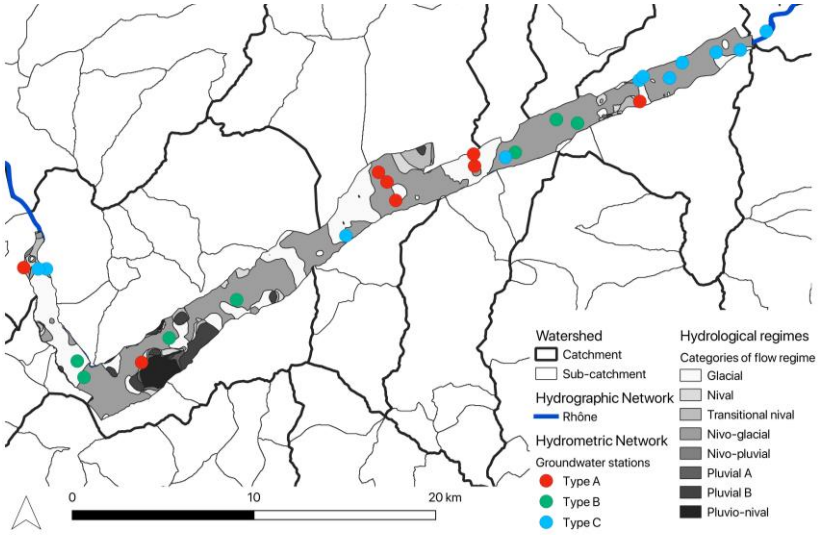




Romane Collin & Corinna Frank

Station Types differ mainly by hydrological regime





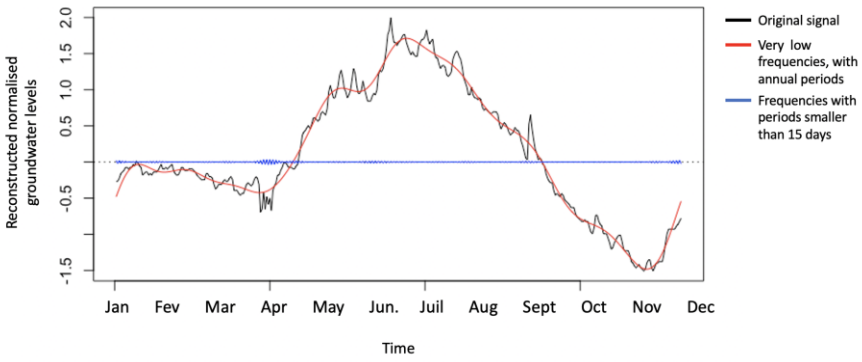
Goal 1

We identified some factors and links influencing groundwater in Valais.

Romane Collin & Corinna Frank

Groundwater Levels should be predicted every 7 days

Romane Collin & Corinna Frank

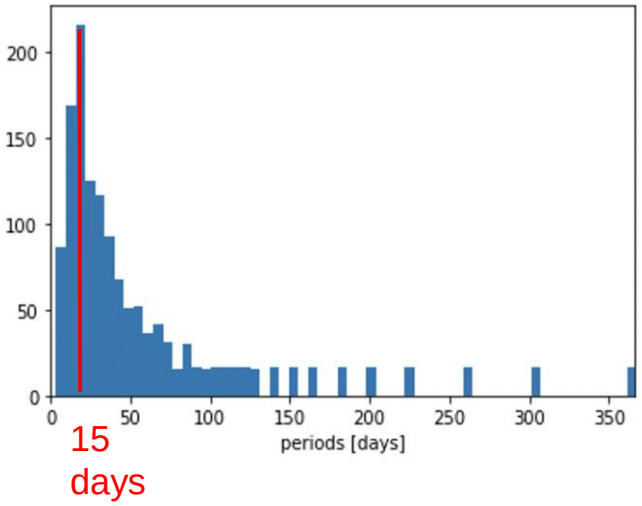


Nyquist theorem

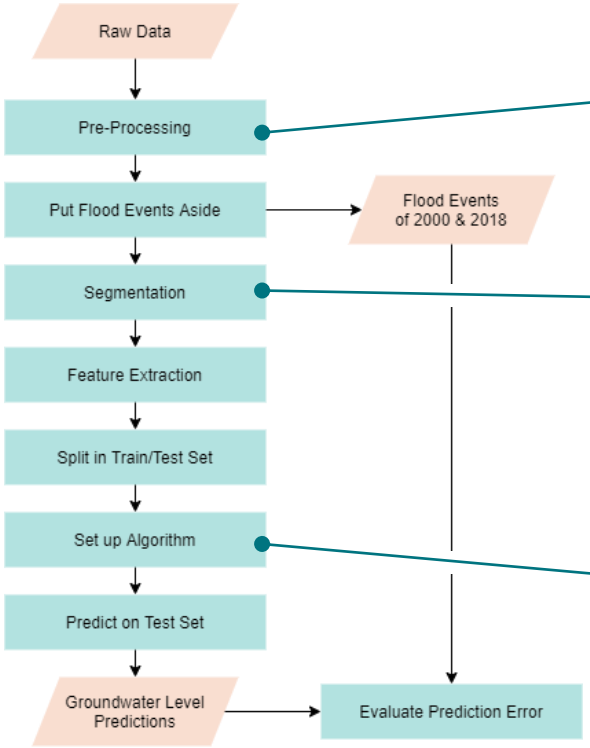
A signal may be uniquely and precisely reconstructed with a sampling rate that is equal to, or greater than, twice the highest significant frequency in the signal.

Fourier Transform analysis

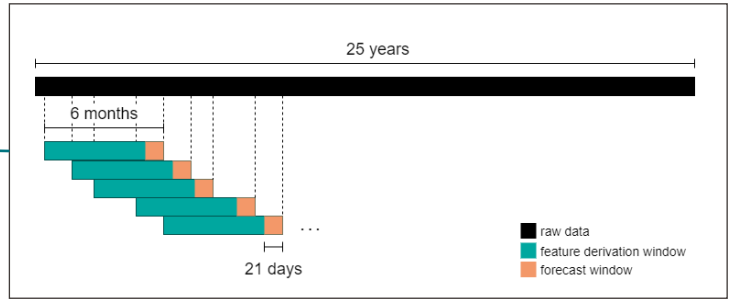
1. Forecast resolution should be of 7 days for reconstructing the groundwater level variations (Nyquist theorem).
2. Samples of past groundwater levels should be taken more often than every 7 days.



Building a Machine Learning model to predict Groundwater Levels



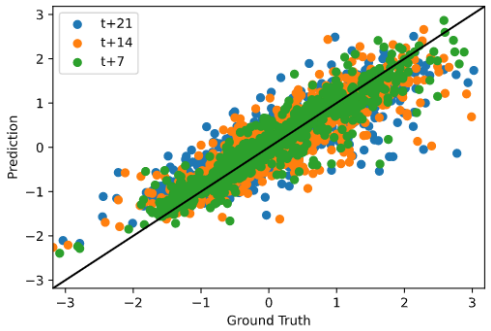
normalization since stations are on different elevations

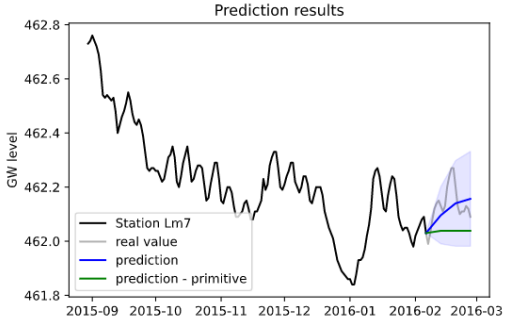


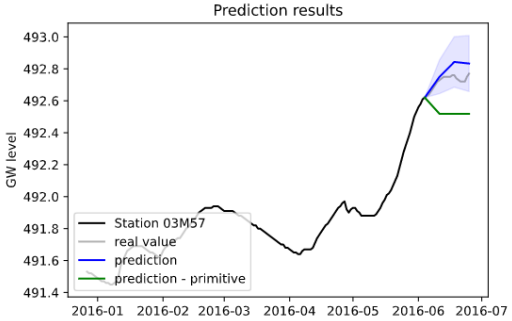
Random Forest Regressor
+ allows certain insight
+ robust (ensemble learning)
+ allows use of different feature types

Romane Collin & Corinna Frank

A simple model yields promising results for the forecasting task





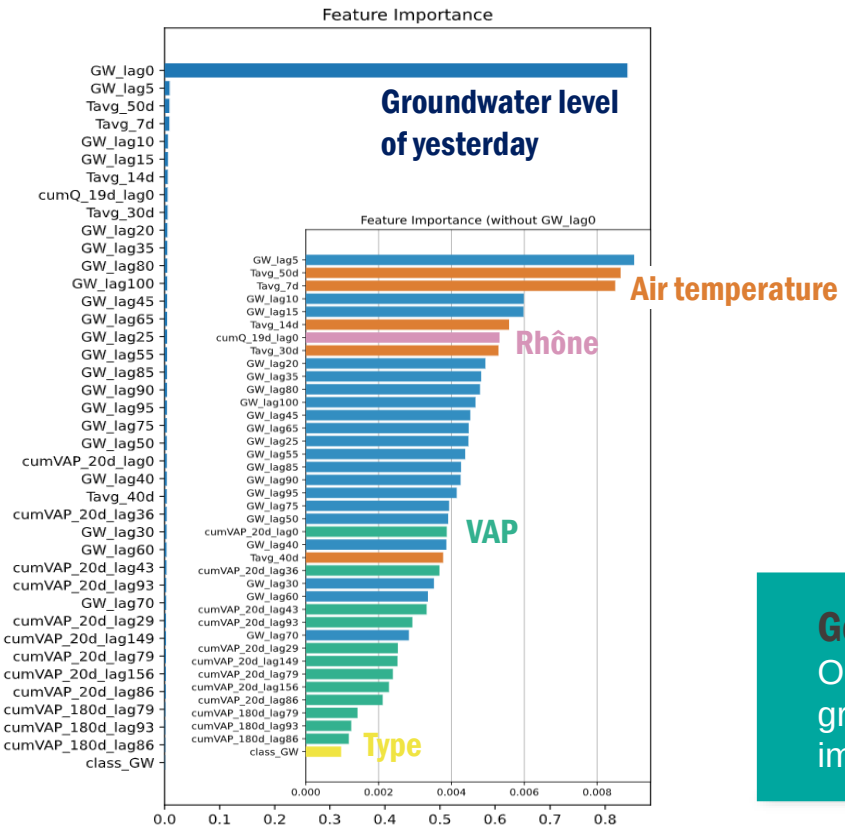


	Overall		Type A	Type B	Type C
R^2 score	**0.80**		**0.70**	0.84	0.82
	RMSE	Maximum error	RMSE	RMSE	RMSE
t+7 days	0.11 m	0.68 m	0.10 m	0.10 m	0.11 m
t+14 days	0.16 m	1.29 m	0.16 m	0.15 m	0.16 m
t+21 days	0.18 m	0.91 m	0.17 m	0.15 m	0.19 m

- Large variance in the performance between stations (R^2 from -2.86 to 0.94)
- Less precise on station type A
- Better than primitive model (R^2: 0.72)

Romane Collin & Corinna Frank

Various possibilities to improve the forecasting in the future



Our propositions

1. Apply a **high-pass filter** to remove the annual pattern.
2. Apply a **low-pass filter** to focus on trend rather than uncapturable short-term variations.
3. Use **forecasting features** (meteoSwiss, Crealp) since groundwater levels showed dependancy on recent conditions.
4. Build one **specialized model per type** of station.

Goal 2

Our model can serve as a first forecasting tool of groundwater levels in Valais, allowing further improvement in the future.

Romane Collin & Corinna Frank

Thank you…

… Pascal Ornstein (Crealp) & Javier Fluixa (Crealp)

for this project and your valuable input and assistance

… Prof. Devis Tuia

for your helpful advice!