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
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 recurring...
- 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
Two objectives are to be addressed

**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 prediction? 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.
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
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
Station behaviors can be clustered in 3 types

- 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?
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
Station Types differ mainly by hydrological regime

Goal 1
We identified some factors and links influencing groundwater in Valais.
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

Raw Data
  ↓
Pre-Processing
  ↓
Put Flood Events Aside
  ↓
Segmentation
  ↓
Feature Extraction
  ↓
Split in Train/Test Set
  ↓
Set up Algorithm
  ↓
Predict on Test Set
  ↓
Groundwater Level Predictions
  ↓
Evaluate Prediction Error

Flood Events of 2000 & 2018

normalization since stations are on different elevations

Random Forest Regressor
+ allows certain insight
+ robust (ensemble learning)
+ allows use of different feature types
A simple model yields promising results for the forecasting task

- Large variance in the performance between stations (R² from -2.86 to 0.94)
- Less precise on station type A
- Better than primitive model (R²: 0.72)

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Type A</th>
<th>Type B</th>
<th>Type C</th>
</tr>
</thead>
<tbody>
<tr>
<td>R² score</td>
<td>0.80</td>
<td>0.70</td>
<td>0.84</td>
<td>0.82</td>
</tr>
<tr>
<td>RMSE</td>
<td>Maximum error</td>
<td>RMSE</td>
<td>RMSE</td>
<td>RMSE</td>
</tr>
<tr>
<td>t+7 days</td>
<td>0.11 m</td>
<td>0.68 m</td>
<td>0.10 m</td>
<td>0.10 m</td>
</tr>
<tr>
<td>t+14 days</td>
<td>0.16 m</td>
<td>1.29 m</td>
<td>0.16 m</td>
<td>0.15 m</td>
</tr>
<tr>
<td>t+21 days</td>
<td>0.18 m</td>
<td>0.91 m</td>
<td>0.17 m</td>
<td>0.15 m</td>
</tr>
</tbody>
</table>
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 uncatchable short-term variations.

3. Use **forecasting features** (meteoSwiss, Crealp) since groundwater levels showed dependency 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.
Thank you...

... Pascal Ornstein (Crealp) & Javier Fluixà (Crealp)
for this project and your valuable input and assistance

... Prof. Devis Tula
for your helpful advice!