Experimental Study of Time Series Forecasting Methods for Groundwater Level Prediction

Michael F. MBouopda, **Thomas Guyet**
Nicolas Labroche, Abel Henriot

7th Workshop on Advanced Analytics and Learning on Temporal Data

23 September 2022, Grenoble, France

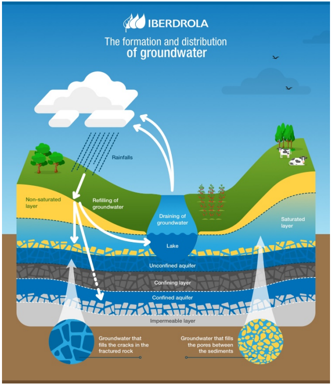
Motivation: Water and groundwater

Ground water

- ground water = water bodies entrapped in soil
- originate from rainfall that penetrates through soil, rocks, etc.
- preserved from pollutants ⇒ freshwater

A very complex dynamic system

- complex flows of water (including evapotranspiration)
- complex interactions between water stocks
- depends on a lot of physic/geographic parameters



source: https://www.iberdrola.com/

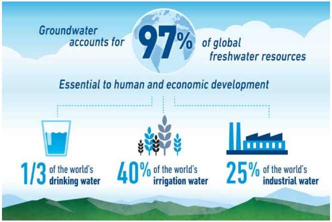
Motivation: Manage the groundwater resource

Highly demanded resources

- for human consumption (1/3 of drinking water comes from groundwaters)
- for agriculture
- for industry (cooling, washing, etc.)

A stressed common resources

- climate change and over-consumptions of water $\rightsquigarrow$ drought
- different potential problems
  - sanity issues
  - continuous access to drinking water
  - economical activities



source: https://www.worldbank.org/

$\Rightarrow$ **Need for a better management of water resources**

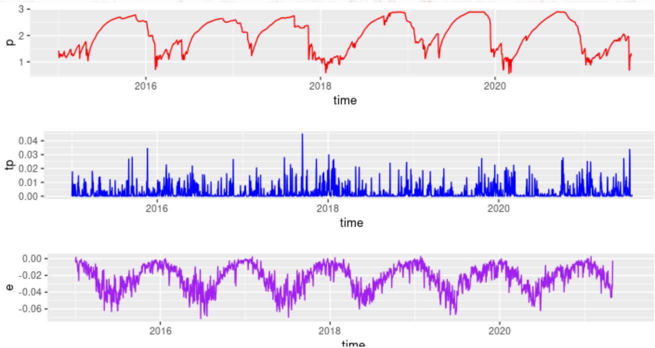
Motivation: Forecasting groundwater level

Need for a better management of water resources

- dynamically adapt consumption to available resources
  - depends on rainfalls, weather and human activities

⇒ **forecasting the groundwater level** is an essential tool to improve water management

  - long term forecasting are awaited (several months)



Motivation: Forecasting groundwater level

Need for a better management of water resources

- dynamically adapt consumption to available resources
  - depends on rainfalls, weather and human activities

⇒ **forecasting the groundwater level** is an essential tool to improve water management

  - long term forecasting are awaited (several months)

State of the art approaches

- hydrological modeling of water flows [1]
  - accurate forecasting
  - very time consuming (a new modeling for each location, require complex measurements)
- data driven modeling: [4, 2]
  - less time consuming ... only to get data about groundwater level
  - less accurate
  - only applied to small datasets of groundwater level measurements, with small horizons

Motivation: Forecasting groundwater level

Need for a better management of water resources

- dynamically adapt consumption to available resources
  - depends on rainfalls, weather and human activities

⇒ **forecasting the groundwater level** is an essential tool to improve water management

  - long term forecasting are awaited (several months)

Research questions

⇒ is it possible to make accurate long term predictions of groundwater levels (3 months)?

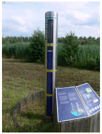
⇒ which time series forecasting model is the better?

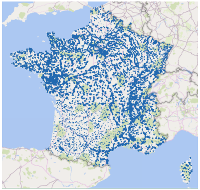
⇒ would it be possible to learn a *global* model?

Data sources

Hub'eau network of sensors

- sensor for groundwater level = *piezometer*
- hundreds of piezos make daily measurements all across the French mainland (see map)
- some sensors are installed since several decades
- data are freely available





Our prepared dataset

- 1,026 sources all over the French mainland,
- period from January 2015 to January 2021 (2,221 days) (with less than 50 missing data)
- completed with daily climate variables (from Copernicus ERA5):
  - rainfall
  - evapotranspiration

→ dataset available upon request

Visualizations

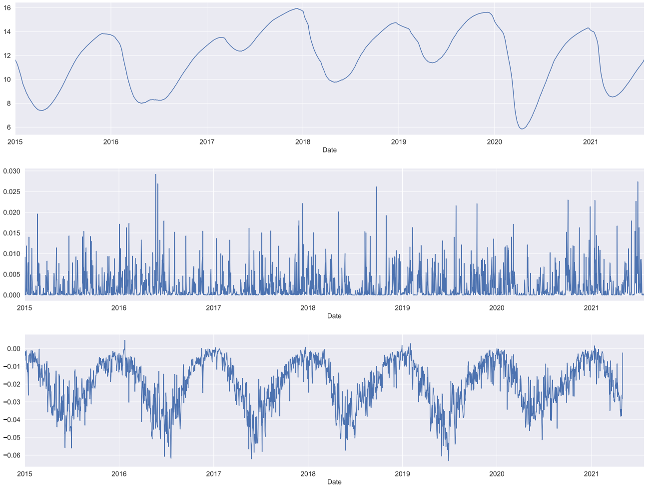


Figure: Top to bottom: daily groundwater, rainfall and evapotranspiration time series for the piezometer BSS000EBL

Formalization

Data available for the i-th piezometers

- Daily historical groundwater level: $y^i_{t=1..n}$.
- Daily exogenous data (rain, evapotranspiration): $\mathbf{z}^i_{t=1..n}$

Predict groundwater level at horizon h

$$\hat{y}^i_{t=(n+1)..(n+h)} = \varphi(y^i_{t=1..n}, \mathbf{z}^i_{t=1..n}, \tilde{\mathbf{z}}^i_{t=n+1..n+h})$$

- $\tilde{\mathbf{z}}$: exogenous data approximation for future dates, estimated by the mean of the previous years

Root mean square scaled error

$$RMSSE = \sqrt{\frac{\frac{1}{h}\sum_{t=t_0+1}^{t_0+h}\left(y_t - \hat{y}_t\right)^2}{\frac{1}{t_0-1}\sum_{t=2}^{t_0}\left(y_t - y_{t-1}\right)^2}}$$

Time series forecasting models

Tree types of forecasting models

- **Autoregressive models** (detailed later)
- **DeepAR-based models** [5]
- **Prophet-based models** [6, 7]: Prophet, NeuralProphet, NeuralProphet

Time series forecasting models: auto-regressive models

Generalization of autoregressive models

- forecasting the next value from the previous r values,

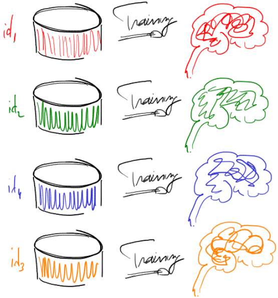
$$\hat{y}^i_{t_0+1} = \varphi(y^i_{t_0-r..t_0}, \mathbf{z}^i_{t_0-r..t_0+1})$$

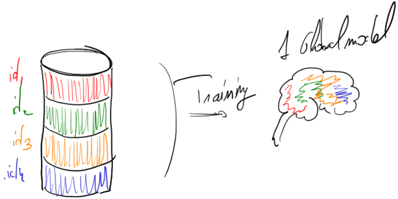
- different classes of φ functions
  - linear function (LM): classical AR model
  - XGB, LM, RF, SVR: modeling non-linear dependencies
- forecasting at horizon h apply recursively *varphi* for $k = 1..h$:
$\hat{y}^i_{t_0+k} = \varphi(y^i_{t_0-(r-k)..t_0} \odot \hat{y}^i_{t_0..t_0+k-1}, \mathbf{z}^i_{t_0-(r-k)..t_0} \odot \tilde{\mathbf{z}}^i_{t_0..t_0+k})$

We experimentally found that an history length $r = 100$ leads to better performance.

Learning strategies

- **Local forecasting**: train a model for each piezometer and use it to predict the future values of this piezometer only.
- **Global forecasting** (or cross-learning) [3]: train a single “big" model from all piezometers and use it to forecast any piezometer level.





Learning strategies

- **Local forecasting**: train a model for each piezometer and use it to predict the future values of this piezometer only.
- **Global forecasting** (or cross-learning) [3]: train a single "big" model from all piezometers and use it to forecast any piezometer level.

Pros/Cons

- *modeling capabilities*: local forecasting requires less modeling capabilities while the global model has to capture all dynamics
- *available data*: each local training has less available data than a global training
- *new piezometers*: a local model can not be used for newly installed piezometers
- global modeling requires more resource intensive training (local modeling can easily be distributed)
- global modeling can be feed with additional non-temporal descriptive features (soil nature)

Experiments and results

Settings

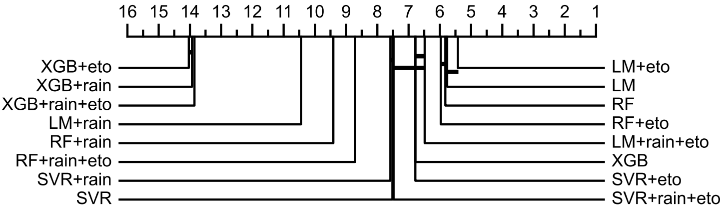
- $1,026$ multivariate time series
- $t_0 = 2,221$ days (5 years and 9 months)
- $h = 93$ days (3 months)
- Critical difference diagrams of *RMSSE* ($\alpha = 0.5$)
- $r = 100$ for auto-regressive models (determined experimentally)
- approximated exogenous data: mean of the five first years

Naming convention: `[forecaster-L,G]+exo`

- forecaster: XGB, LM, DeepAR, Prophet, etc.
- L: Local training, G: Global training
- exo $\in \emptyset$, `eto`, `rain` `eto+rain` when exogenous data are considered.

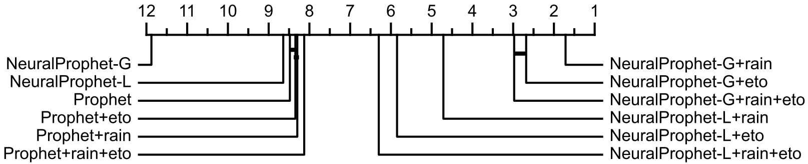
Example: `NeuralProphet-G+exo`

Comparisons of generalized autoregressive models (local training)



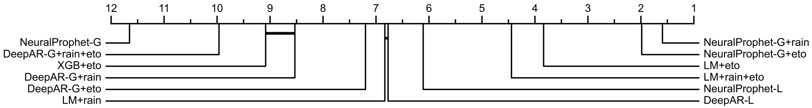
- the best performing configurations use either only evapotranspiration, or no exogenous data at all,
- Linear models are among the best models

Prophet-based models



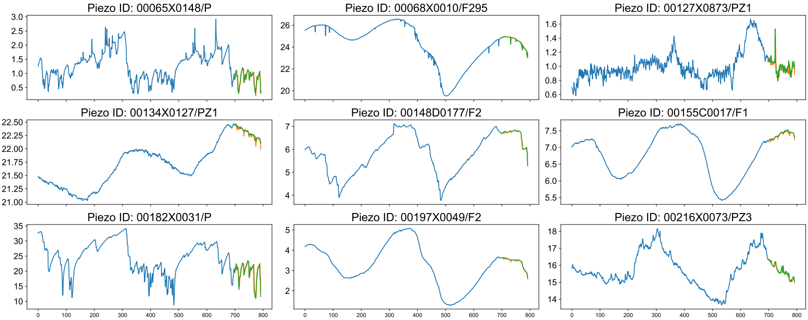
- `NeuralProphet-G` is the worst approach, but the best when combined with exogenous data,
- Unlike `DeepAR`, `NeuralPhophet` works better when executed globally than locally, in particular when exogenous data are used,
- In the absence of exogenous data, `Prophet` significantly outperforms `NeuralProphet`

Comparing the three groups of models



- `LM` and `NeuralProphet-G` respectively produce the most accurate and the least accurate forecasts when exogenous data are absent,
- `NeuralProphet-G` significantly outperforms the others when exogenous data are used,
- precipitations impact groundwater level more than ETO (see `NeuralProphet-G+rain` and `NeuralProphet-G+eto`).

Visualization of some forecasts



- blue: historical data
- green: ground truth
- orange: forecasts

Conclusion

Conclusions

- We performed the most extensive experimental study of groundwater level forecasting (1,026 sources, 9 methods),
- globally trained forecasters seems to outperform the locally trained (but are more resource intensive consuming)
- precipitation tends to have more impact on the accuracy of the forecasts than evapotranspiration.

Data preparation and source code available

https://github.com/frankl1/piezoforecast

Perspectives

- making the dataset available in a standard format
- hybrid strategy: identify groups of piezometers with "similar" dynamics to make a concise, accurate and readily applicable model
- integrate additional exogenous data (water consumption) and features (soil nature)

References

J. Brédy, J. Gallichand, P. Celicourt, and S. J. Gumiere.
Water table depth forecasting in cranberry fields using two decision-tree-modeling approaches.
Agricultural Water Management, 233:106090, 2020.

A. Ibrahem Ahmed Osman, A. Najah Ahmed, M. F. Chow, Y. Feng Huang, and A. El-Shafie.
Extreme gradient boosting (XGBoost) model to predict the groundwater levels in Selangor Malaysia.
Ain Shams Engineering Journal, 12(2):1545–1556, 2021.

T. Januschowski, J. Gasthaus, Y. Wang, D. Salinas, V. Flunkert, M. Bohlke-Schneider, and L. Callot.
Criteria for classifying forecasting methods.
International Journal of Forecasting, 36:167–177, 2020.

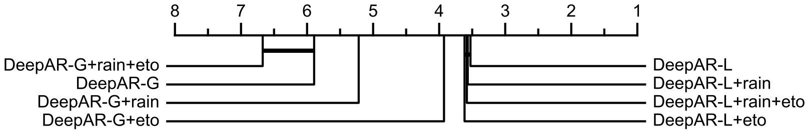
P. C. Nayak, Y. S. Rao, and K. Sudheer.
Groundwater level forecasting in a shallow aquifer using artificial neural network approach.
Water resources management, 20(1):77–90, 2006.

D. Salinas, V. Flunkert, J. Gasthaus, and T. Januschowski.
DeepAR: Probabilistic forecasting with autoregressive recurrent networks.
International Journal of Forecasting, 36:1181–1191, 4 2017.

S. J. Taylor and B. Letham.
Forecasting at scale.
PeerJ Preprints, 2017.

O. Triebe, H. Hewamalage, P. Pilyugina, N. Laptev, C. Bergmeir, and R. Rajagopal.
NeuralProphet: Explainable forecasting at scale.
arXiv:2111.15397v1, 2021.

DeepAR-based models



- Local training strategy for DeepAR is better than a global one,
- Exogenous data improve global DeepAR, but have insignificant impact on local DeepAR.