Time Series Analysis On Snow Depth In the Alps

Pablo FERREIRA – ING1

Professor: Dr Abhishek Grover

Recap – Objectives, limitations, progress

Dataset:
Gapfilled SAFRAN-CROCUS
Snow depths report
(1975-2020)

Time Series Analysis techniques used:
Naïve, ARIMA, Prophet

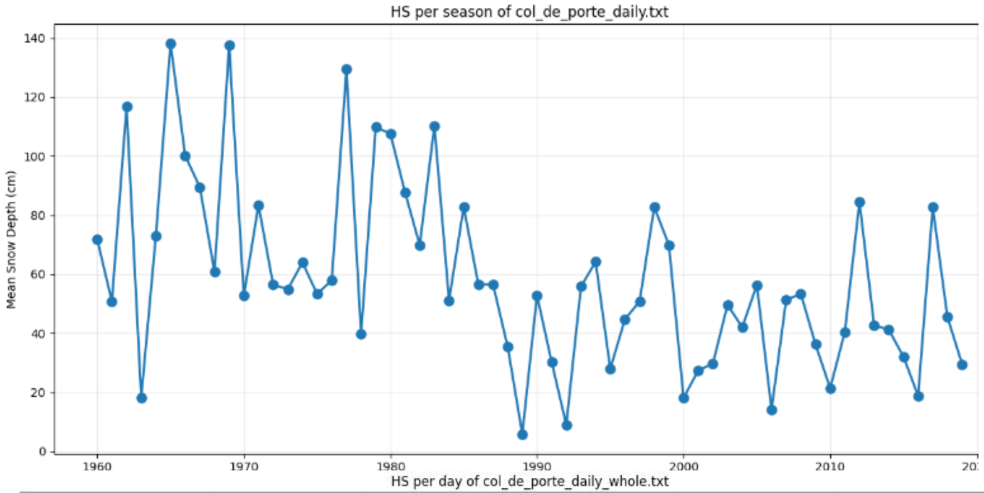
Best NMAE: ~0.4 – 0.8

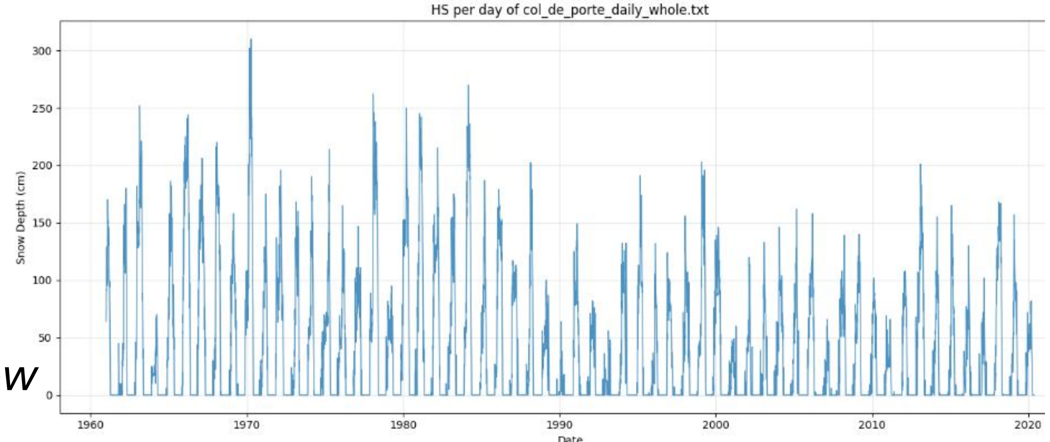
Limits found: erratic values because of no successful seasonality modelling

Dataset

- **1975 => 2019**, 1 log/day, 212 day/year (Nov – May)
- Source: MeteoFrance's **SAFRAN-CROCUS** dataset
- Stations to study:
  - Les 2 Alpes: Highest, glacier ~**3000m**
  - Tignes: north-facing, stable snowpack ~**2600m**
  - Serre Chevalier: south-facing, unstable snowpack ~**2100m**
  - Col de Porte: Lowest, humid, perfect data source ~**1300m**

- Measurements are grid-generalized per station
- Processed Day Metrics : HS = > qc(HS) => **gapfill(qc(HS))**





Col de Porte metrics, per season above and per day below

Work Done – Naïve Model

- Average of past Same Days:
  - NS(y, m, d) = $\frac{1}{y-1}\sum_{k=1}^{y-1} ds[k, m, d]$
  - NMAE = $\frac{MAE}{mean(HS\_after\_gapfill)}$
  - 2 Clear groups performance-wise: Stable/Unstable snowpack
- Note: 10 first years without cross-validation

Worst Accuracy

Best Accuracy

STATION NAME	MAE	NMAE
Tignes	30.822	0.395
Les 2 Alpes	48.345	0.404
Serre-Chevalier	25.076	0.82
Col de Porte	32.797	1.125

Baseline; basically a mean

Work Done – (AR)(I)(MA) Model

- $AR_{m,d}$(p) = $C + \sum_{k=1}^{p} \Phi_k * ds[t-1]$
- $I_{m,d}$(d) = $\nabla^d * ds[t]$
- $MA_{m,d}$(q) = $C + \sum_{k=1}^{q} \theta_k * \varepsilon_{t-k}$
- Note: We add white noise shock

For every ARIMA member as ε_t

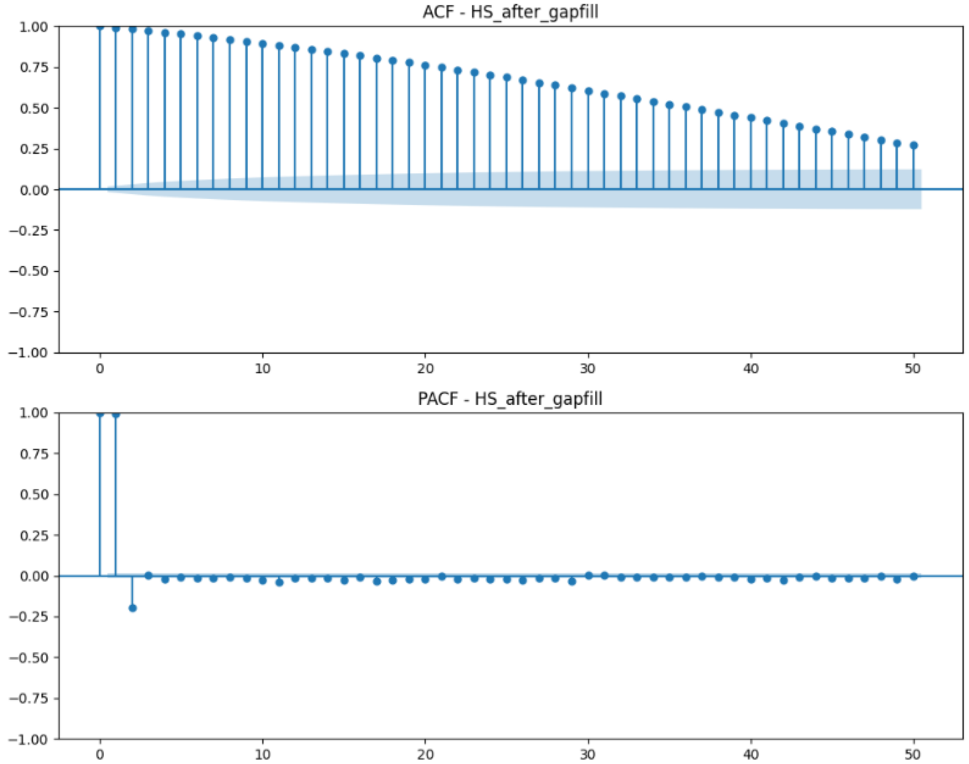
(pink)	Worse Accuracy
(green)	Better Accuracy

STATION NAME	MAE	NMAE	p, d, q
Tignes	44.769	0.52	1, 0, 1
Les 2 Alpes	64.701	0.521	1, 0, 1
Serre-Chevalier	33.173	0.905	1, 0, 0
Col de Porte	42.711	0.946	1, 0, 0

Work Done – (AR)(I)(MA) Model

Worse Accuracy

Better Accuracy




```
ADF Statistic of col_de_porte_daily.txt : -12.037385978227787
p-value: 2.78284946105129e-22
Lags used: 28
Number of observations: 12615
Critical Value (1%): -3.431
Critical Value (5%): -2.862
Critical Value (10%): -2.567
```

STATION NAME	MAE	NMAE	p, d, q
Tignes	44.769	0.52	1, 0, 1
Les 2 Alpes	64.701	0.521	1, 0, 1
Serre-Chevalier	33.173	0.905	1, 0, 0
Col de Porte	42.711	0.946	1, 0, 0

PACF stops at 2 => p = {1,2}

ACF(x) – ACF(x-1) <0.5 => q={0,1?}

ADF << Critical values & p~0 => d = 0

PACF & ADF : correlation coefficients

ADF : probability coefficient

Work Done – Prophet Model

- Linear smoothing model by Meta
- Arguments tested:
  - Seasonality (custom – 212 VS **year**)
  - Fourier order [**5**;15]
  - Changepoint prior scale [**0.05**;0.5]
  - Seasonality mode (**add** VS mult)
- Default values are **bold**

Worse Accuracy

Better Accuracy

STATION NAME	MAE	NMAE	Args
Tignes	35.516	0.389	Default
Les 2 Alpes	52.951	0.396	Default
Serre-Chevalier	26.249	0.792	Seasonality
Col de Porte	34.219	0.847	Seasonality

Work Done – Recap

(pink)	Worse Accuracy
(green)	Better Accuracy

- Naïve (d – mean(sum(d))
- ARIMA (p, d, q)
- Prophet (seasonality, fourier orders,cp prior scale, s modal)

STATION NAME	Naive	Arima	Prophet
Tignes	0.395	0.52	0.389
Les 2 Alpes	0.404	0.521	0.396
Serre-Chevalier	0.82	0.905	0.792
Col de Porte	1.125	0.946	0.847

Work Done II – SARIMAX

- Sarima(X) -> No X:

model must be based only on past snow depth data

- SARIMA (1, 1, 1)(P, D, Q, S)

- No seasonality patterns useful in SARIMAX => FAIL

- Ran on Kaggle with CPU

Worse Accuracy

Better Accuracy

STATION NAME	(1, 0, 1, 30)	(1, 0, 1, 7)
Tignes	0.868	0.892
Les 2 Alpes	0.647	0.65
Serre-Chevalier	0.943	0.96
Col de Porte	0.988	0.99

ARIMA > SARIMAX here

Work Done II, NN – LSTM / GRU

Worst Accuracy

Best Accuracy

- Type of RNN, solves vanishing gradient issue
- GRU simpler than LSTM:

  only Update + Reset Gate
- 2 sets -> light / heavy :

  seq_len, hidden_dim, epochs, batch_size, LR, patience, dropout
- Ran on Kaggle with GPU

STATION NAME	GRU Light	GRU Heavy	LSTM Light
Tignes	0.595	0.668	0.989
Les 2 Alpes	0.765	0.577	0.67
Serre-Chevalier	0.753	0.875	0.901
Col de Porte	0,857	1.005	0.81

No clear « best » RNN or hyperparemeter set

Work Done II, NN – XGBoost

- boosted decisional trees

- V1: failure (generic params)

  predicted ~0cm

- V2: good (squared-error, bigger depth)
  rewarded wilder guesses

- V3 : good+ (custom weights)

  ignored summer days where ~0cm

- *Ran on Kaggle with GPU for hist*

Worse Accuracy

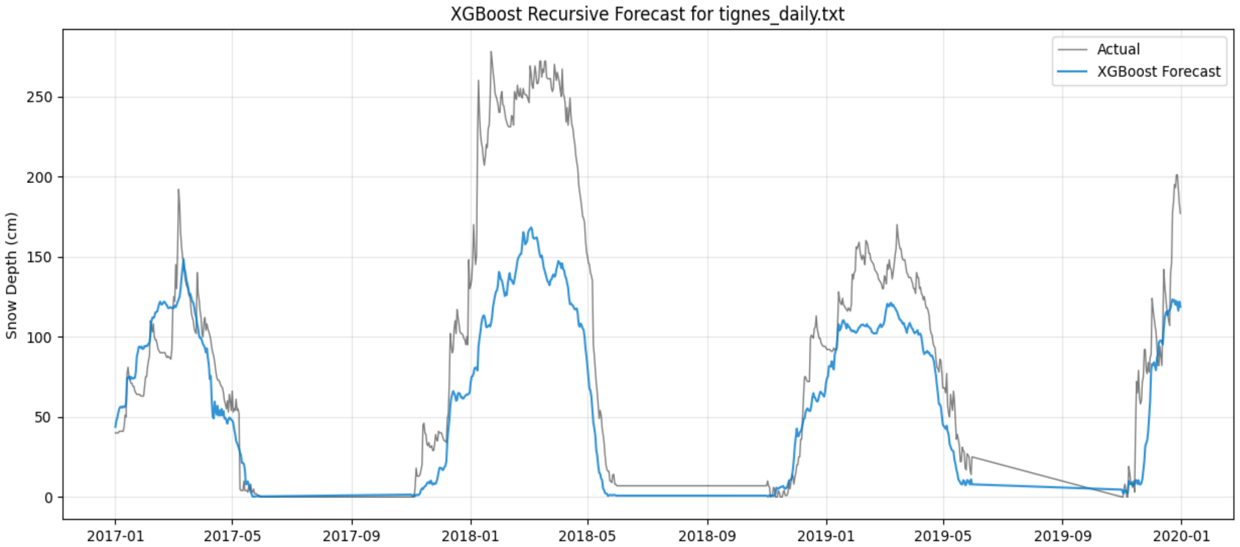
Better Accuracy

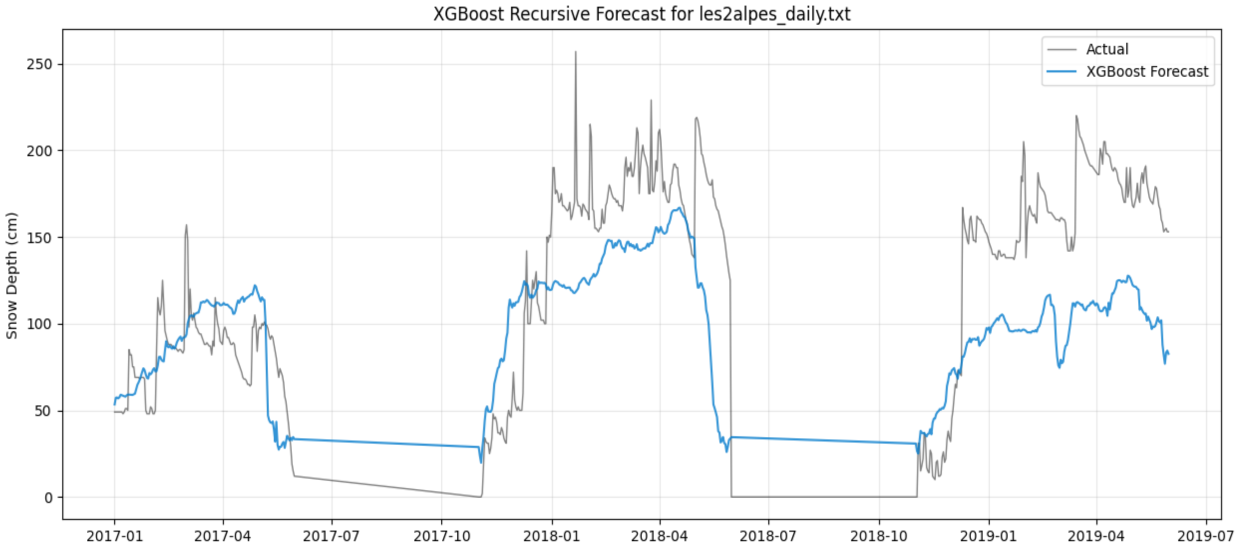
STATION NAME	V2	V3
Tignes	0.365	0.356
Les 2 Alpes	0.37	0.337
Serre-Chevalier	0.628	0.537
Col de Porte	0.525	0.535

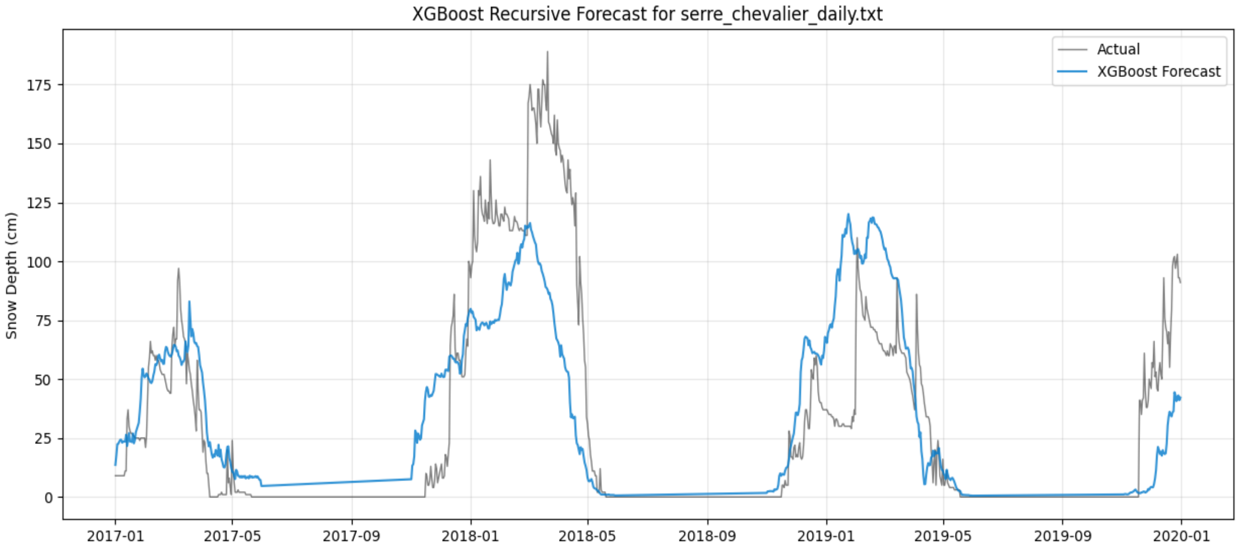
Limitations: truly eternal/erratic snow depths patterns

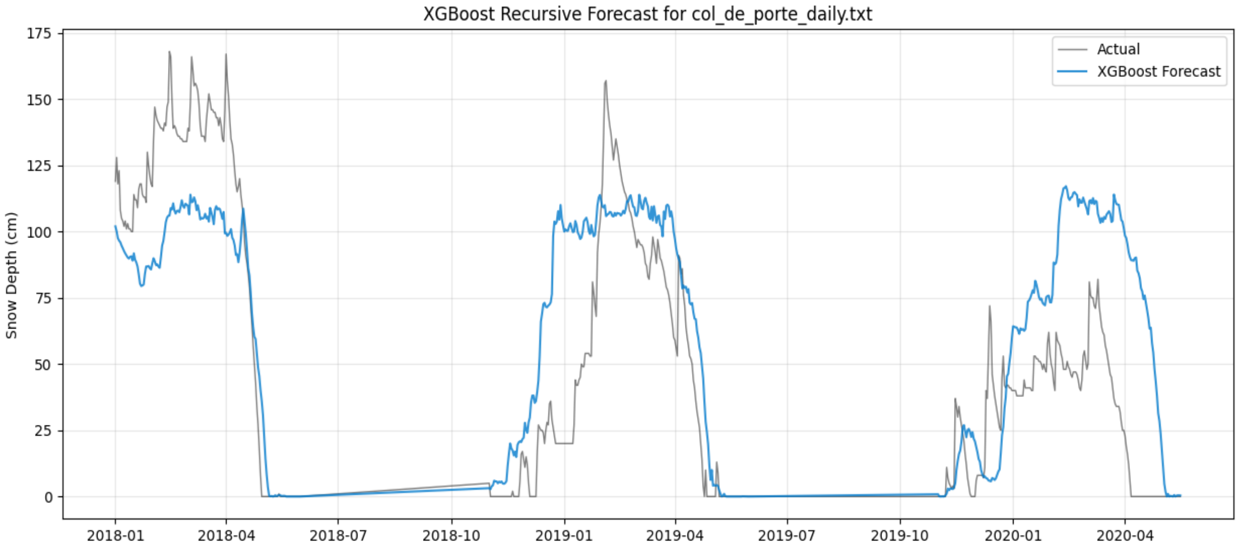
Final Results + GUI

Last 3 years shown here









CONCLUSION – What’s next

https://github.com/PapayaSupreme/snow-depth-time-series-analysis-model
https://www.kaggle.com/work/collections/17079326

Add features ?:

Temp°, radiation, humidity…

Improve GUI

Add other stations to the dataset

Thank you for your time !

Contact : pablo.ferreiraa10@gmail.com

Literature – Primary Data Source

Matiu, M., Crespi, A., Bertoldi, G., Carmagnola, C. M., Marty, C., Morin, S., Schöner, W., Cat Berro, D., Chiogna, G., De Gregorio, L., Kotlarski, S., Majone, B., Resch, G., Terzago, S., Valt, M., Beozzo, W., Cianfarra, P., Gouttevin, I., Marcolini, G., Notarnicola, C., Petitta, M., Scherrer, S. C., Strasser, U., Winkler, M., Zebisch, M., Cicogna, A., Cremonini, R., Debernardi, A., Faletto, M., Gaddo, M., Giovannini, L., Mercalli, L., Soubeyroux, J.-M., Sušnik, A., Trenti, A., Urbani, S., and Weilguni, V.: **Observed snow depth trends in the European Alps: 1971 to 2019**, The Cryosphere, 15, 1343–1382, https://doi.org/10.5194/tc-15-1343-2021, 2021.

Literature – Gap filling / Meaning

Use of the models Safran-Crocus-Mepra in operational avalanche forecasting, Coléou C* , Giraud G, Danielou Y, Dumas J-L, Gendre C, Pougatch E CEN, Météo France, Grenoble, France

Version 3.0 of the Crocus snowpack model, Matthieu Lafaysse1 , Marie Dumont1 , Basile De Fleurian1 , Mathieu Fructus1 , Rafife Nheili1 , Léo Viallon-Galinier1 , Matthieu Baron1 , Aaron Boone2 , Axel Bouchet1 , Julien Brondex1,4 , Carlo Carmagnola1 , Bertrand Cluzet1 , Kévin Fourteau1 , Ange Haddjeri1,4, Pascal Hagenmuller1 , Giulia Mazzotti1,4, Marie Minvielle2 , Samuel Morin1 , Louis Quéno1,5, Léon Roussel1 , Pierre Spandre1 , François Tuzet1 , and Vincent Vionnet1,3

Literature - Models

Aschauer, J., Bavay, M., Begert, M., and Marty, C.: **Comparing methods for gap filling in historical snow depth time series**, EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020–17211, https://doi.org/10.5194/egusphere-egu2020-17211, 2020.

Introduction to Time Series and Forecasting – Third Edition,

By Peter J. Brockwell Richard A. Davis, 2016

Durbin, James, and Siem Jan Koopman. 2012. **Time Series Analysis by State Space Methods: Second Edition.** Oxford University Press.