Processing AROME vertical profiles with Machine Learning methods to diagnose aeronautical ceiling for TAF messages

P. Crispel, P. Jaunet S. Moisselin and A. Drouin

Météo France DSM/AERO

Context: IniTAF project

The **IniTAF project** is intended to provide TAF initialization
→ freeing up time to focus on forecasting the evolution of the most critical parameters (i.e horizontal visibility in case of fog formation, etc.).

Several inputs are required to initialize TAFs from model data: wind, cloud cover, ceiling, etc.

Some of them are directly computed in NWP models: wind speed and direction.

Others require further developments: visibility, ceiling.

Ceiling: Aeronautical definition and thresholds

Annex 3 - Meteorological Service for International Air Navigation, ICAO

Aeronautical ceiling is the lowest cloud layer base height with:
- cloudiness greater than 50% in a radius of 8 km from the airport
- cloud base height ≤ 5000 ft

Ceiling significant thresholds for TAF evolution groups

100 ft	**200 ft**	500 ft	1000 ft	**1500 ft**	**5000 ft**
	Low Visibility Procedure (e.g. Roissy Charles de Gaulle airport)			Significant low ceiling in TAF evolution groups	CAVOK threshold

Note that ceilings upper than 5000 ft do not have to be mentioned in TAF messages.

Numerical weather prediction model: AROME

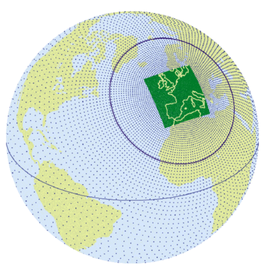
REGIONAL NWP MODEL AROME (Météo-France model)

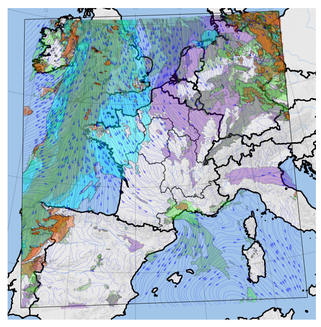
→ Input: Global model ARPEGE (Météo-France)

→ Resolution: 1.3km

→ Runs at 00 UTC, 03UTC 06UTC, 12UTC, 18UTC

→ 1h time step





State of the art: « NWP CEILING DIAG »

Benchmark ceiling diagnosis based on cloud fraction outputs from AROME (cloud fraction > 50%) - *Seity, Y., et al. "Cloud and microphysical schemes in ARPEGE and AROME models."*

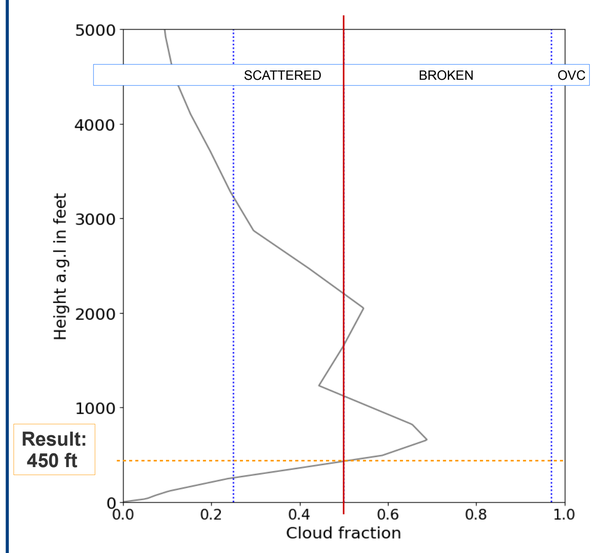


Figure: Cloud fraction vertical profile. NWP ceiling diagnosis (orange line) corresponds to the height a.g.l. where cloud fraction exceeds 50 %.

Forecast Vs. hourly METAR ceilings
Year 2017 - 66 French airports

Main ceiling thresholds	Hit rate	False Alarm Ratio
< 100 ft	46 %	70 %
< 500 ft	51 %	56 %
< 1000 ft	52 %	42 %
< 1500 ft	54 %	35 %
< 5000 ft	53 %	18 %

Table: NWP CEILING DIAG. Binary scores for several thresholds. Runs 00h, 06, 12, 18. Lead times: 6 to 11h.

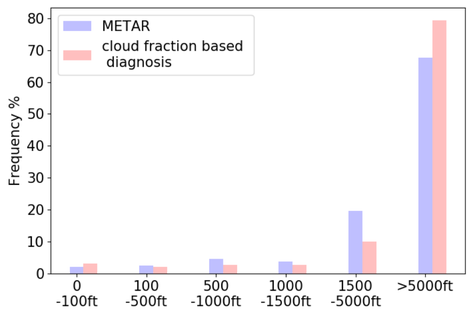


Figure: NWP CEILING DIAG. Frequency of observed (blue) and forecasted ceilings (red). Runs 00h, 06, 12, 18. Lead times: 6 to 11h.

Machine Learning dataset

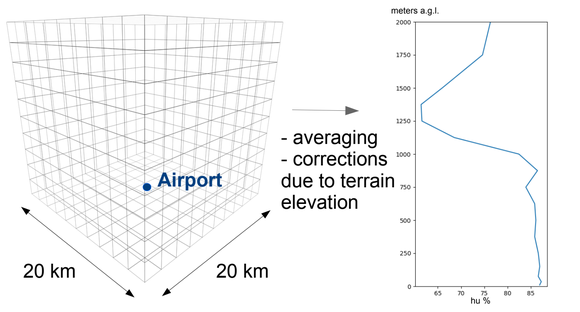
PREDICTORS

Meterological Parameters:

- humidity, wind speed, temperature
- turbulent kinetic energy
- microphysics parameters (cloud water water / ice contents, snow content)

Runs: 00, 06, 12, 18UTC **Lead times**: 6-11h

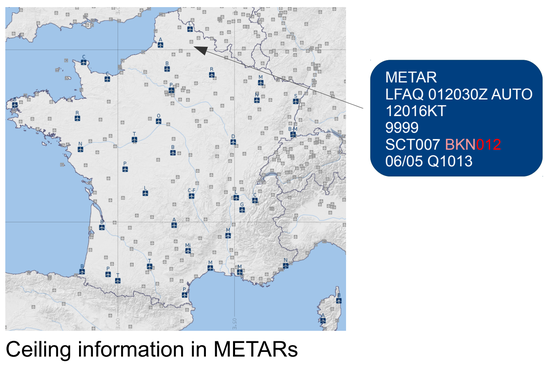
→ Hourly extracted on a **3D grid** around the airport (20 km x 20 km x 6000 ft a.g.l).



3D grid information around the airport

Median vertical profile at the airport.

OBSERVATIONS

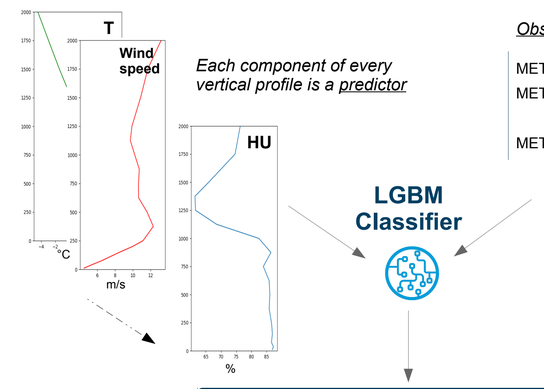


Ceiling information in METARs

- **learning**: year 2016 – 66 french airports hourly data ~ 10^5 observations.

- **verification**: year 2017 – 66 french airports hourly data ~ 10^5 observations.

Machine Learning methodology & LGBM classifier



Each component of every vertical profile is a predictor

Observed ceiling information is extracted from METARs

METAR LFAQ 011500Z AUTO 12012KT 9999 OVC007 06/06 Q1018
METAR LFAQ 011600Z AUTO 12011KT 9999 BKN007 OVC011 06/06 Q1017
……..
METAR LFAQ 011900Z AUTO 12012KT 9999 OVC012 07/06 Q1015

Light Gradient Boosting Machine (LGBM)

Ke et al. 2017
Compared to other gradient boosting methods (e.g. XGboost), lightGBM:

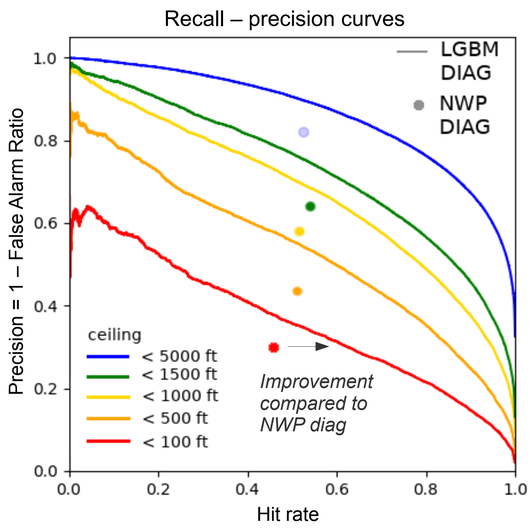
→ handles large datasets
→ handles high feature dimension
→ negative: a lot of parameters to tune.

Comparison: LGBM CEILING DIAG Vs NWP CEILING DIAG (benchmark)

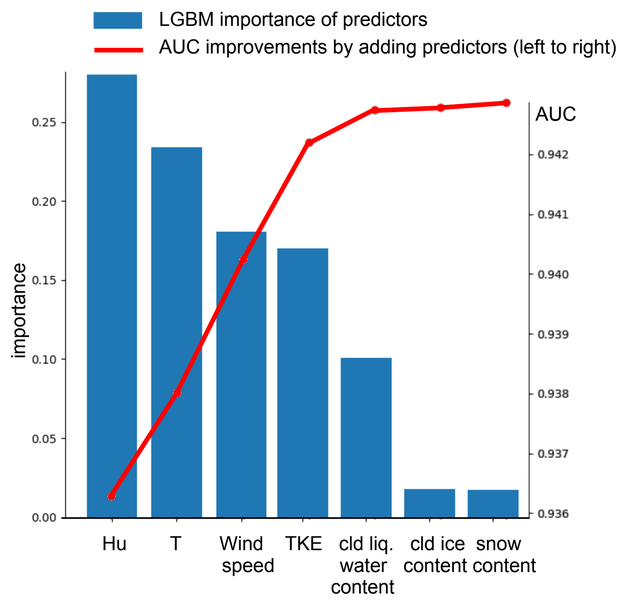
Scores

Forecast Vs. hourly METAR ceilings
Year 2017 - 66 French airports

Binary predictions for different thresholds



Predictor importance LGBM method



Main ceiling thresholds	Hit rate		False Alarm Ratio	
< 100 ft	46 %	+0	62 %	-8
< 500 ft	55 %	+4	48 %	-8
< 1000 ft	65 %	+12	39 %	-3
< 1500 ft	70 %	+16	36 %	-1
< 5000 ft	80 %	+27	24 %	+6

Table: LGBM CEILING DIAG. Binary scores for several thresholds. Runs 00h, 06, 12, 18.
Lead times: 6 to 11h.
Green/red numbers refer to improvement/loss compared to NWP ceiling diagnosis.

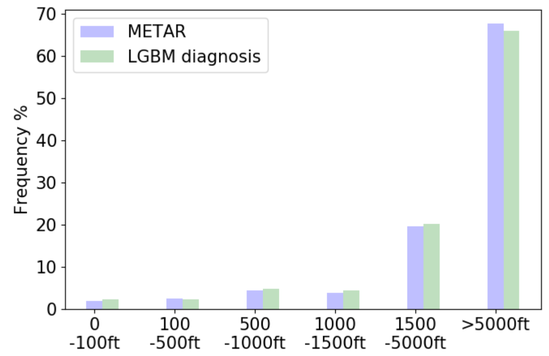
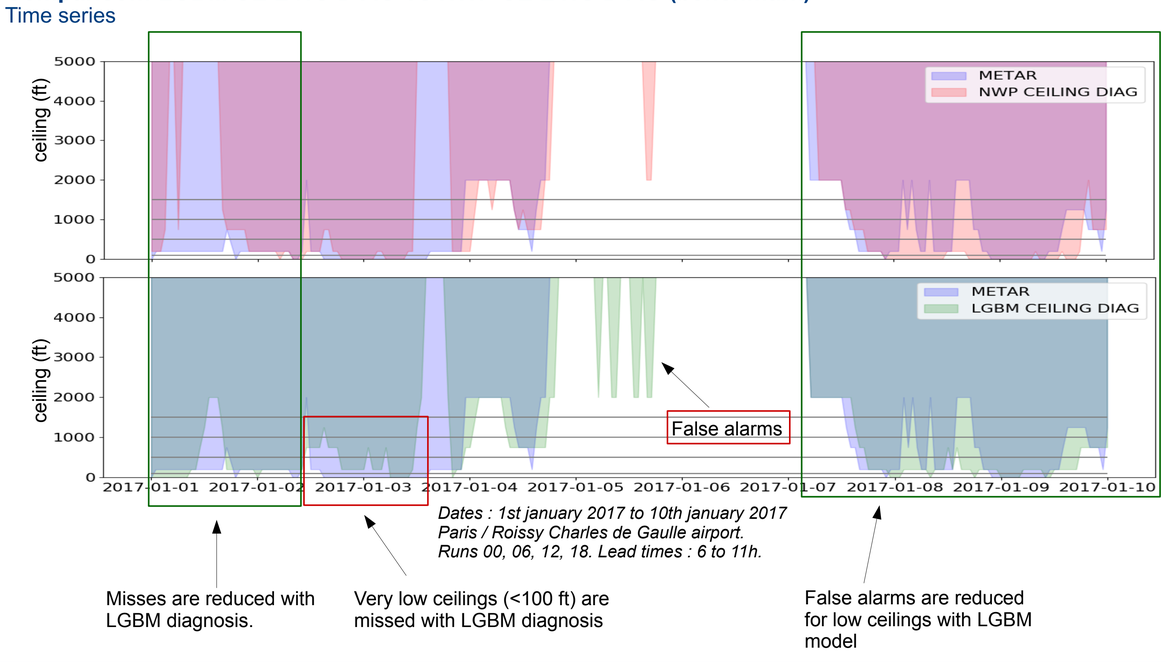


Figure: LGBM CEILING DIAG. Frequency of observed (blue) and forecasted ceilings (green). Runs 00h, 06, 12, 18.
Lead times: 6 to 11h.

Comparison: LGBM CEILING DIAG Vs NWP CEILING DIAG (benchmark)

Time series



Conclusion – future work

- Binary scores show better results for LGBM ceiling diagnosis than a direct cloud fraction based diagnosis (NWP CEILING DIAG) which was used as a benchmark.

→ Results are significantly improved when ceiling threshold rises.

- Work on case studies has to be continued to improve the classification (discrimination for very low clouds).

- Use of direct 3D grid information should be explored with convolutional neural networks (CNN).