universität innsbruck

Tree-Based Boosting Models for Low-Visibility Forecasts at Different Lead Times

Sebastian J. Dietz (*sebastian.j.dietz@gmail.com*), Philipp Kneringer, Georg J. Mayr, and Achim Zeileis

poster number X5.11

In a nutshell

Probabilistic forecasts for low-visibility conditions, relevant for flight planning, are developed using a statistical tree-based boosting model. The forecasts are designed to provide the air traffic controllers with information for short-term regulation, and for the air traffic managers to improve flight plan construction.

Introduction

Safety operations with low visibility:
Instrument landing approach
Increased **spacing distances** ⟹
Raised **taxi times**

Results:
Decreased **capacity**
Flight **delays**
Economic **loss**

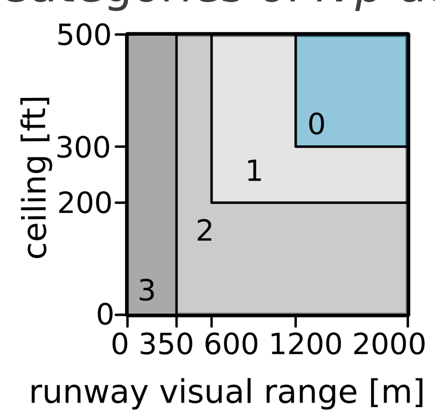
Accurate forecasts of low visibility allow:
Short-term **regulations** ⟶ rapid nowcasts (1–2 h)
Flight plan **reorganizations** ⟶ nowcasts (3–18 h)
Better long-term **flight planning** ⟶ medium-range (1–14 d)

Lead times required:
rapid nowcasts (1–2 h)
nowcasts (3–18 h)
medium-range (1–14 d)

Low-Visibility Procedure (*lvp*) States

- Define **safety procedures** during low visibility that **reduce airport capacity**
- Occur mainly with **fog**, **low ceiling**, or **heavy precipitation**

Categories of *lvp* at **Vienna Airport**:



lvp	capacity	occurrence
0	100%	89.7%
1	70%	1.7%
2	60%	7.1%
3	40%	1.5%

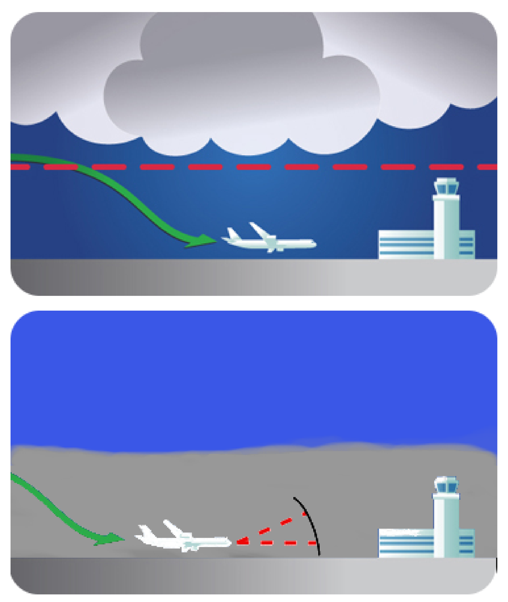


Figure 1: Illustration of **ceiling** (top) and **runway visual range** (bottom).

Explanatory Variables

Observations*

Variable	Unit	Description
lvp	[0,1,2,3]	low-visibility procedure state
rvr	[m]	runway visual range
vis	[m]	visibility
cei	[ft]	ceiling
dpd	[°C]	dew point depression
dts	[°C]	temp. difference to surface
sza	[°]	solar zenith angle

NWP model outputs (DET, ENS)**

Variable	Unit	Description
bld	[Jm^{-2}]	boundary layer dissipation
blh	[m]	boundary layer height
e	[m.w.e]	evaporation
cdir	[Jm^{-2}]	clear sky direct solar radiation
dpd_{model}	[°C]	dew point depression
dts_{model}	[°C]	temp. difference to surface
lcc	[0 – 1]	low cloud cover
shf	[Jm^{-2}]	sensible heat flux
tp	[m]	total precipitation

* the observations are chosen using the results of Kneringer et al. (2017)

** the NWP models used are the ECMWF deterministic high resolution model (**DET**) and the ensemble prediction system (**ENS**). From the ENS only mean and spread is used. The predictors are selected with the results of Herman and Schumacher (2016) and meteorological expertise. The maximum lead time of the DET is 10 days; for the ENS it is 15 days

Tree-Based Boosting Models

Figure 2: Schematic illustration of a boosting tree.

Model Development

1. Develop a single **decision tree**
2. Compute **residuals*** of the model
3. Fit a **new tree** on the residuals
4. **Add** new tree to previous ones
5. **Repeat** recursively steps 2–4

* negative gradient vector of the likelihood function

Models are developed with data from **5 cold seasons** (2011–2017).

Nowcasts

The **statistical nowcasts outperform** persistence and climatology at each lead time (Fig. 3). **Observations** have highest influence on **shortest lead times**. The impact of **DET model outputs increases** with **mid-term nowcasts** after 2 hours lead time. They **control the predictions** with lead times longer than 8 hours, when the range of **long-term nowcasts** starts. Most important **observations** are the ***lvp* state** and **dew point depression**. From the **DET model dew point depression** and **evaporation** have highest impact (Fig. 4).

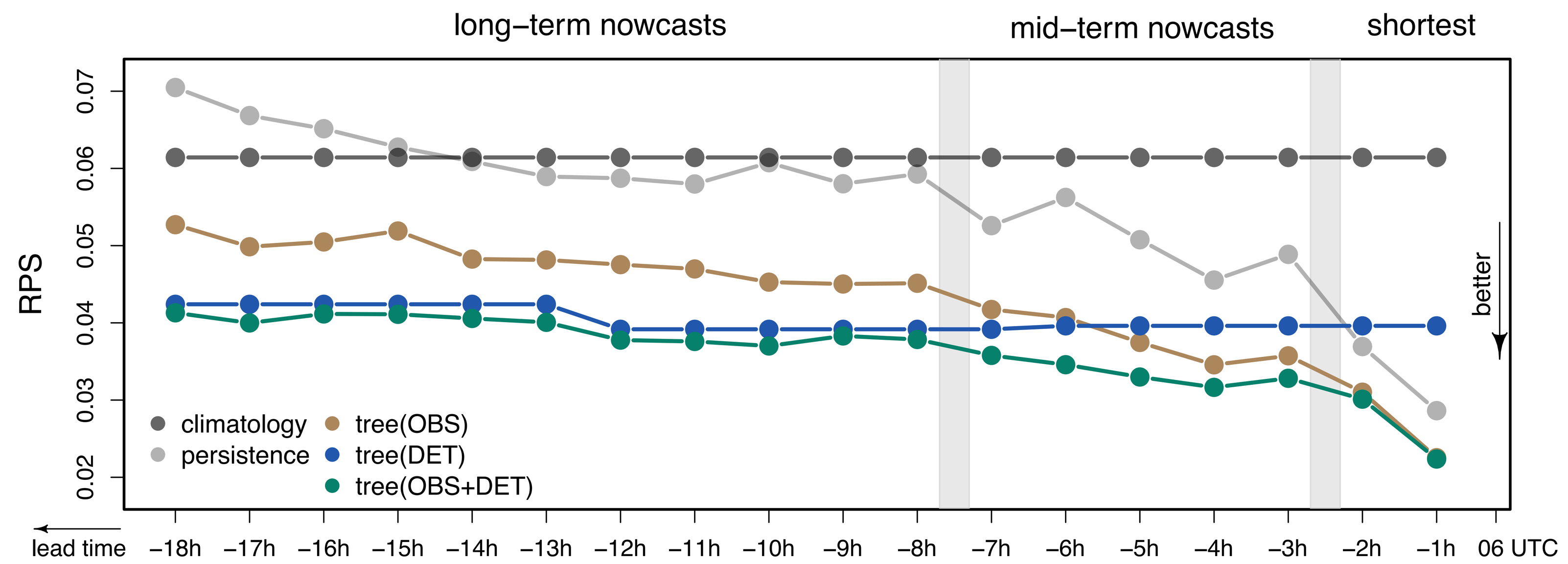


Figure 3: **Nowcast performance** of persistence, climatology, and boosting trees with different inputs.

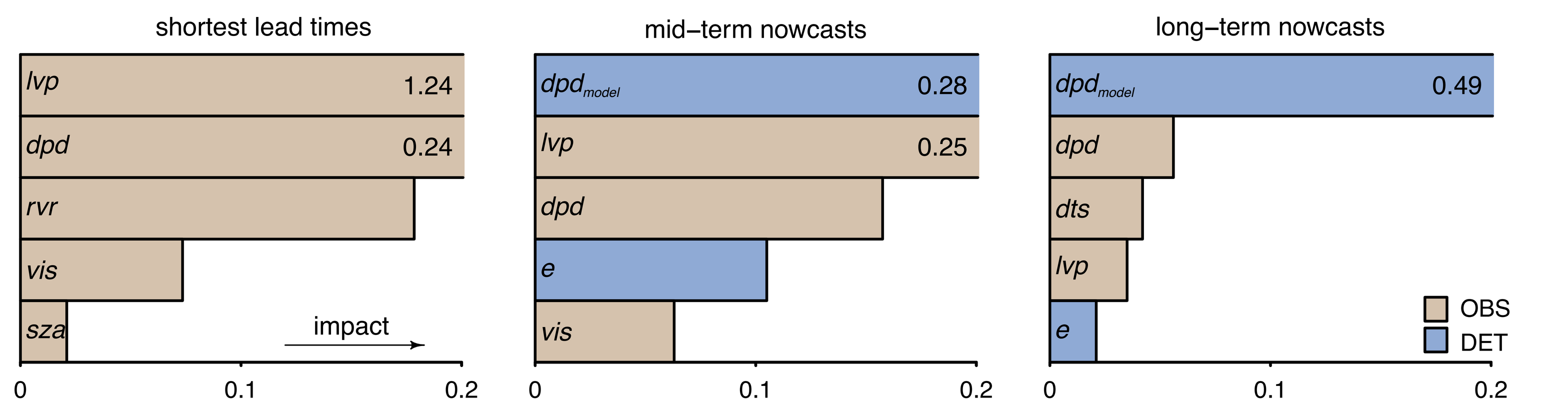


Figure 4: Explanatory variables with **highest impact** on *lvp* nowcasts. The impact is computed by variable permutation tests. Higher numbers concern to a stronger impact.

Medium-Range Forecast and Forecast Horizon

The performance of the **statistical models converges** strongly to **climatology** for lead times longer than **8 days**. After this time the **forecast horizon** is approximately reached. Postprocessed outputs of **DET** and **ENS** perform similarly until **5 days lead time**. Afterwards models with **ENS** information have **higher benefit**. **Raw ENS performs similarly** to statistical models between the lead times **1 day** and **5 days**; **raw DET** is **outperformed** at each lead time (Fig. 5).

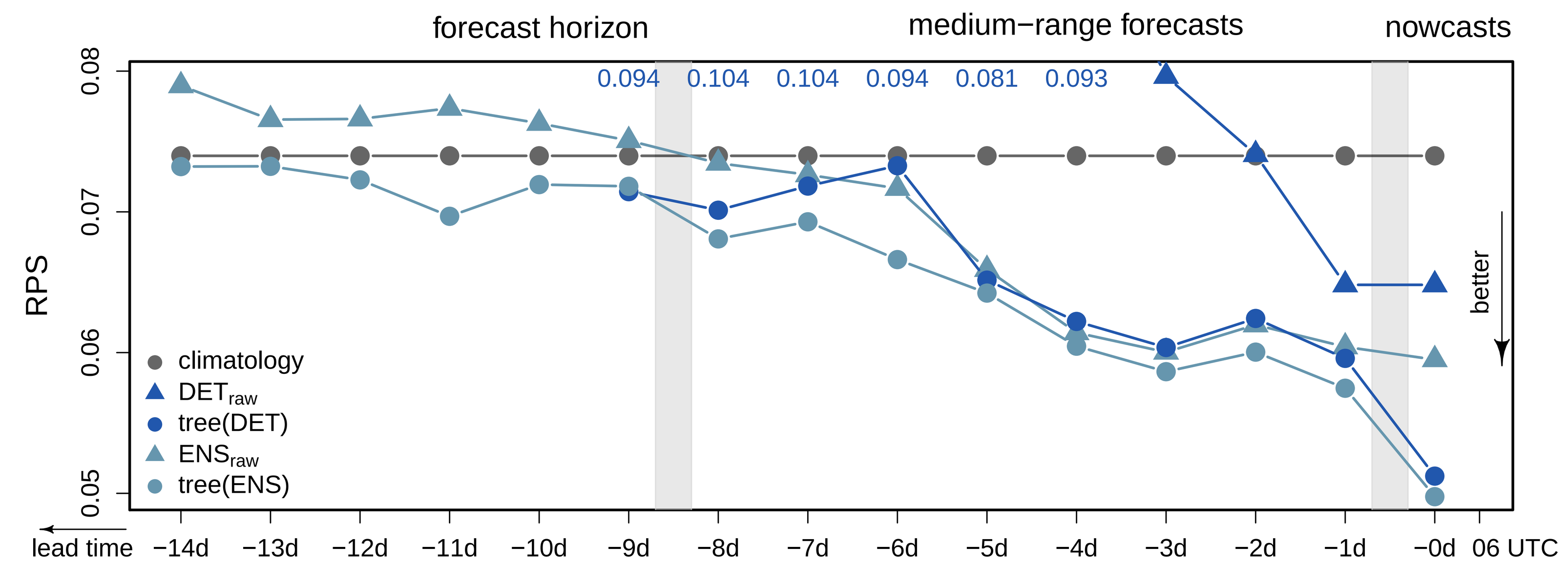


Figure 5: **Medium-range performance** of models with deterministic and ensemble information.

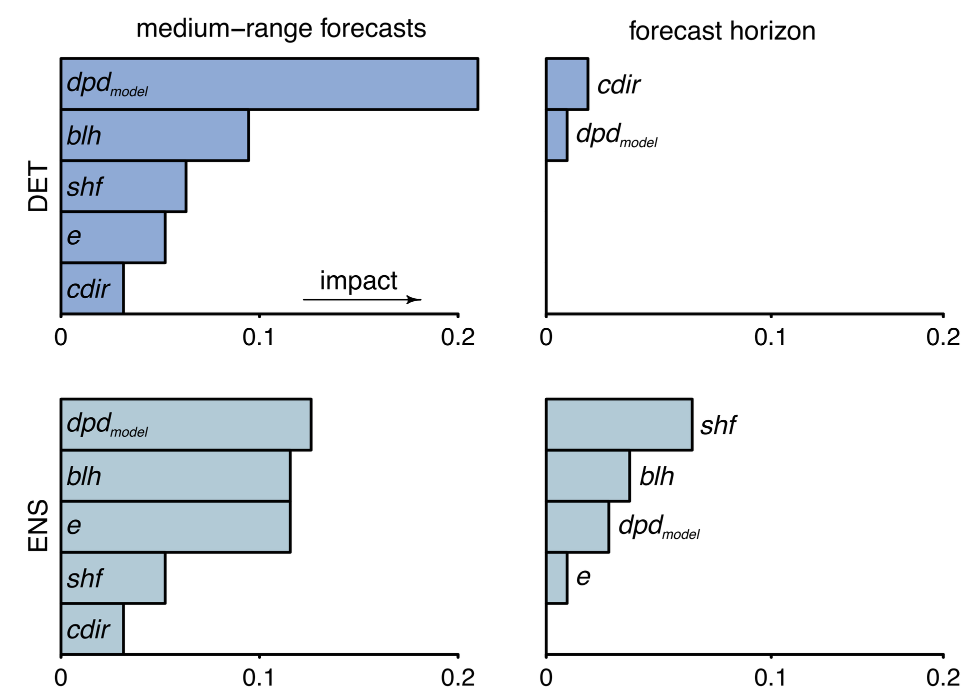


Figure 6: Variables with **highest impact** for medium-range forecasts of deterministic and ensemble based models.

Raw ENS and raw DET are computed with the NWP outputs **visibility** and **ceiling**. The **data archive** of these variables, however, is too **short** for the usage as predictors.

The **impact of individual predictors** on the forecast **decreases** strongly after reaching the **forecast horizon** (Fig. 6).

Take Home Message

- **Probabilistic** *lvp* forecasts have a **benefit over persistence, climatology, and raw DET outputs** until **8 days lead time**
- Most important inputs for *lvp* state **nowcasts** are
  - **Observations** of *lvp* and dew point depression
  - **DET outputs** of dew point depression and evaporation
- Most important predictor variables for **medium-range forecasts** are dew point depression, boundary layer height, evaporation, and sensible heat flux from **NWP models**

References:

Herman G. R. and Schumacher R. S., 2016: Using Reforecasts to Improve Forecasting of Fog and Visibility for Aviation. Weather and Forecasting, **31**, 467-482, https://doi.org/10.1175/WAF-D-15-0108.1.

Kneringer P., Dietz S.J., Mayr G.J., and Zeileis A., 2017: Probabilistic Nowcasting of Low-Visibility Procedure States at Vienna International Airport during Cold Season. Working Paper 2017-21, Working Papers in Economics and Statistics, Universität Innsbruck, http://EconPapers.RePEc.org/RePEc:inn:wpaper:2017-21.

Acknowledgments:

Ongoing project funded by the Austrian Research Promotion agency (FFG) 843457 and the PhD scholarship of the University of Innsbruck. We thank the Austro Control GmbH for access to the observation data and the ZAMG for the ECMWF data.