



Probabilistic nowcasting of low-visibility procedure states at Vienna International Airport during cold season

Philipp Kneringer, Sebastian J. Dietz,
Georg J. Mayr, Achim Zeileis

Working Papers in Economics and Statistics

2017-21

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
Research platform "Empirical and Experimental Economics"
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 7171
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
<http://eeecon.uibk.ac.at/wopec/>

For a list of recent papers see the backpages of this paper.

Probabilistic Nowcasting of Low-Visibility Procedure States at Vienna International Airport During Cold Season

Philipp Kneringer
Universität Innsbruck

Sebastian J. Dietz
Universität Innsbruck

Georg J. Mayr
Universität Innsbruck

Achim Zeileis
Universität Innsbruck

Abstract

Airport operations are sensitive to visibility conditions. Low-visibility events may lead to capacity reduction, delays and economic losses. Different levels of low-visibility procedures (*lvp*) are enacted to ensure aviation safety. A nowcast of the probabilities for each of the *lvp* categories helps decision makers to optimally schedule their operations. An ordered logistic regression (OLR) model is used to forecast these probabilities directly. It is applied to cold season forecasts at Vienna International Airport for lead times of 30-min out to two hours. Model inputs are standard meteorological measurements. The skill of the forecasts is accessed by the ranked probability score. OLR outperforms persistence, which is a strong contender at the shortest lead times. The ranked probability score of the OLR is even better than the one of nowcasts from human forecasters. The OLR-based nowcasting system is computationally fast and can be updated instantaneously when new data become available.

Keywords: aviation meteorology, low visibility, probabilistic nowcasting, statistical forecasts, ordered logistic regression.

1. Introduction

Low-visibility conditions at airports impact air traffic regarding aviation safety and economic efficiency of airports and airlines. The low-visibility procedures (*lvp*) come into force when horizontal and/or vertical visibility fall below airport-specific thresholds. Additional measures, e.g., increasing spacing between approaching and taxiing aircraft, ensure safe operations but also reduce the capacity of the airport. Consequently planes might be delayed, diverted to alternative airports or prevented from taking off. Hence, reliable low-visibility forecasts are needed for tactical planning of the aircraft movements within the next few hours.

The two major weather phenomena producing low-visibility conditions are fog and low ceiling. Fog development and dissipation generally depend on temperature, the humidity and the available condensation nuclei of an air mass. Radiative cooling, change of the total water amount due to precipitation or mixing of air masses are only a few processes changing these

important parameters (Gultepe *et al.* 2007). However, despite various field experiments (e.g., Gultepe *et al.* 2009; Haeffelin *et al.* 2009; Bergot 2016) the overall effect of some physical processes is still unknown. The complexity of the processes driving low visibility makes predictions challenging.

Physical modeling with numerical weather prediction models and statistical modeling are two general low-visibility forecasting approaches (Gultepe *et al.* 2007). The statistical approaches are data-driven and computationally faster. Model parameters are estimated on an archive data set and then applied to new data to forecast. The choice of the statistical model depends on the desired form of the forecast variable (continuous or categorical) and the type of the forecast output (deterministic or probabilistic). Regression models were among the first statistical forecasting approaches applied to continuous variables with linear regression (Bocchiari and Glahn 1972) and were extended to binary and multinomial categorical variables (e.g. Hilliker and Fritsch 1999; Herman and Schumacher 2016). Some machine learning methods have also been used for low-visibility forecasts; for example tree-based methods (Dutta and Chaudhuri 2015; Bartoková *et al.* 2015; Dietz *et al.* 2017) and artificial neural networks (Bremnes and Michaelides 2007; Marzban *et al.* 2007; Fabbian *et al.* 2007).

Since risk management and decision making depends heavily on the probabilistic information (Murphy and Winkler 1984) probabilistic forecasts are essential to make safe and economic decisions especially for an application like air traffic regulation. One option for airport visibility forecasts is to predict the horizontal and vertical visibility separately and determine *lvp* afterwards. However, since horizontal and vertical visibility are not statistically independent there is no obvious way to obtain the combined probability. Instead, the variable of interest to aviation end-users, the *lvp* categories, should be forecast directly.

In this paper we present a new way to generate probabilistic *lvp* state forecasts for the next two hours with a statistical regression method. The method of choice is ordered logistic regression (OLR) to capture the categorical and ordered nature of the end-user forecast variable, which is based on fine-grained visibility and ceiling thresholds. Due to the interest in short lead times the nowcasting system is exclusively based on point measurements (see Vislocky and Fritsch 1997). The performance of the nowcasting system is compared to climatology, persistence, and human forecasts. The methodology to develop the nowcasting system is shown in Sect. 2. Sect. 3 examines the area of investigation and the data used. Sect. 4 presents the results of the nowcasting system which are discussed in the final section.

2. Methods

2.1. Ordered Logistic Regression

The method used to develop a probabilistic *lvp* nowcasting system is OLR. This method allows prediction of the probabilities for all categories of an ordered response, like *lvp*, within one consistent model. Additionally, OLR has the benefit of providing a fast update cycle with low computational costs.

The OLR model describes an ordered categorical variable by assuming an underlying continuous variable mapped with thresholds to the categories. The threshold coefficients and predictor coefficients are determined during model estimation. An arbitrary number of predictors is possible, similar to multiple linear regression. The occurrence probability of the

individual categories can be determined by evaluating the chosen error distribution function at the lower and upper thresholds of a category. In mathematical notation the OLR model is described as follows:

Each observation of the response falls into a ordered category $j = (1, 2, \dots, J)$. The ordering is from no impact to highest impact on airport operations in this case. These ordinal response y_i is modeled by assuming a continuous auxiliary variable y_i^* capturing visibility. For this variable a linear model

$$y_i^* = \mathbf{x}_i^\top \beta + \varepsilon_i \quad (1)$$

holds, where $i = (1, 2, \dots, n)$ is the index over the observations. But y_i^* is not observed directly, therefore the observation i is modeled to category j by the thresholds

$$\alpha_{j-1} > y_i^* \geq \alpha_j \quad (2)$$

Eq. 1 shows the deterministic component as $\mathbf{x}_i^\top \beta$ and the random term ε_i which is assumed to be i.i.d. with zero mean. The vector $\mathbf{x}_i = (x_{i,1}, \dots, x_{i,m})$ includes all the m predictors and $\beta = (\beta_1, \dots, \beta_m)$ includes their coefficients. The thresholds $\alpha = (\alpha_0, \dots, \alpha_J)$ are determined together with the predictor coefficients β when estimating the OLR model. The lowest and highest threshold values are fixed at the values $\alpha_0 = -\infty$ and $\alpha_J = \infty$. To access the probabilities of the categories an error distribution needs to be selected. Typical distributions are the standard normal and the logistic distribution. The logistic distribution accounts better for observations in the tails of the distribution (Winkelmann and Boes 2006). We select the logistic distribution with its cumulative distribution function

$$H_{logit}(\cdot) := \frac{\exp(\cdot)}{1 + \exp(\cdot)}. \quad (3)$$

with $H(-\infty) = 0$ and $H(\infty) = 1$. The probabilities for the categories are derived with the cumulative probability model

$$\sum_{s=1}^j \pi_{is} = H_{logit}(\alpha_j - \mathbf{x}_i^\top \beta) = \frac{\exp(\alpha_j - \mathbf{x}_i^\top \beta)}{1 + \exp(\alpha_j - \mathbf{x}_i^\top \beta)}, \quad (4)$$

where $\sum_{s=1}^j \pi_{is}$ is the cumulative probability that an observation y_i falls into category j or lower. The probability for the individual category $\pi_{ij} = P(y_i = j | x_i) = H(\alpha_j - \mathbf{x}_i^\top \beta) - H(\alpha_{j-1} - \mathbf{x}_i^\top \beta)$ is the difference between the cumulative probabilities at the associated thresholds (Fig. 1).

A notable advantage of this method is the computational speed of the model estimation. It is almost as fast as a linear regression and can be performed instantaneously using standard software. We use the function `clm()` from the R package **ordinal** (Christensen 2015), which implements the OLR model with maximum likelihood optimization.

2.2. Predictor Selection

The first task in estimating the models is to decide which of the input variables to select and which to omit. We use stepwise selection and verify the forecasts with the ranked probability score (RPS see Sect. 2.3). The initial step of the algorithm estimates the climatology as a first guess. In the next step the variable that most improves the RPS of the model is added.

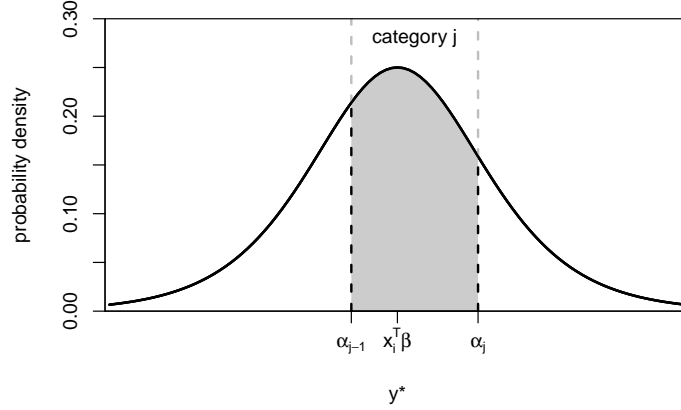


Figure 1: Probability density function of an OLR model. The shaded area highlights the probability for the category j .

Subsequently this model is used as the new best guess and all remaining variables are tested again. This procedure is repeated until either the model skill does not improve anymore, or no additional variable is left. The variable configuration of the final best-guess model is used to produce the low-visibility forecasts.

2.3. Verification

To test the model and to cover the uncertainty within the model estimation we perform ten-fold cross validation. Therefore we split the data set into ten parts, use nine parts for training, and do out-of-sample predictions on the remaining test data set. The test data set is exchanged with one part of the previous training data set and again the model estimation and out-of-sample predictions are performed. This is repeated until we have out-of-sample forecasts for all ten parts of the previously split data set, based on ten different models with slightly different training data sets.

To determine the skill of the probabilistic categorical ordered forecasts a proper scoring rule is required. The ranked probability score (RPS) is such a metric [Wilks \(2011\)](#). It compares the cumulative distribution function of the forecasts and the observations. The RPS of a forecast i is given by

$$RPS_i = \frac{1}{J-1} \sum_{s=1}^J \left[\sum_{j=1}^s (y_{ij} - o_{ij}) \right]^2 \quad (5)$$

with y_{ij} the predicted probabilities and o_{ij} the observations for each category $j = 1, 2, \dots, J$. While the predicted probabilities can have continuous values between 0 and 1, the observation is either 0 or 1. The RPS can be interpreted as the normalized shift in categories between the forecast and the observation. In addition, the RPS is normalized by the number of categories $J - 1$ to obtain scaled values within the interval $[0, 1]$. A perfect forecast has an RPS of zero, the worst forecast has an RPS of 1.

The quality of the prediction is determined by calculating the RPS of all individual observation-prediction pairs within the out-of-sample prediction data set, and averaging them. Bootstrapping is used to estimate the uncertainty due to the limited sample size. Hence, we take 500 random samples of the RPS values from the out-of-sample prediction data set. Each of this

Table 1: Predictors used within the statistical nowcasting approach.

Name	Description	Unit
<i>lvp</i>	low-visibility procedure state	ordered [0, 1, 2, 3]
<i>rvr</i>	runway visual range	meter
<i>cei</i>	ceiling height	feet
<i>vis</i>	horizontal visibility	meter
<i>rh</i>	relative humidity	%
<i>dpd</i>	dew point depression	degree Celsius
<i>rr12</i>	precipitation in the last 12 hours	factor [yes, no]
<i>dt.tow</i>	temperature difference 2m – tower	degree Celsius
<i>dt.surf</i>	temperature difference 2m – surface	degree Celsius
<i>ff</i>	wind speed	m s ⁻¹
<i>dd.se</i>	wind direction from sector southeast	factor [yes, no]
<i>dd.n</i>	wind direction from sector north	factor [yes, no]
<i>sza</i>	solar zenith angle	degree

samples is taken with replacement and has the size of the full data set. Now we calculate the mean RPS of each random sample. These 500 RPS values are the basis of the results shown in Sect. 4 including the model uncertainty.

2.4. Reference Forecasts

The OLR models are compared to three reference forecasts. The first one is the climatology, which uses the climatological occurrence probability of each category (Fig. 2b) as forecast. The second reference is the persistence forecast, which assumes that the *lvp* state at the forecast initialization remains. This state is predicted with a probability of 100% and all other states with 0%. It needs to be mentioned that persistence is already a benchmark at short lead times (Vislocky and Fritsch 1997). As a third reference we compare the OLR to the operational human forecasts at Vienna International Airport (VIE). Human forecasters use all available information from observations and numerical weather predictions and produce operational forecasts at most airports.

3. Data

VIE is selected to develop the statistical low-visibility nowcasting tool. The airport with its two runways is located in the Vienna Basin 20 km southeast of downtown Vienna. The basin is bounded by the Alps to the west and by the Carpathian Mountains to the east. The Airport is surrounded by many humidity sources. Moisture advection from the southeast (Lake Neusiedl and wetlands) and the north (Danube River) favor low-visibility conditions.

3.1. Definition of the Low-Visibility Procedure States

Low-visibility events occur when horizontal and/or vertical visibility drop below a set of thresholds. The runway visual range (*rvr*) is used as horizontal visibility measure. It is

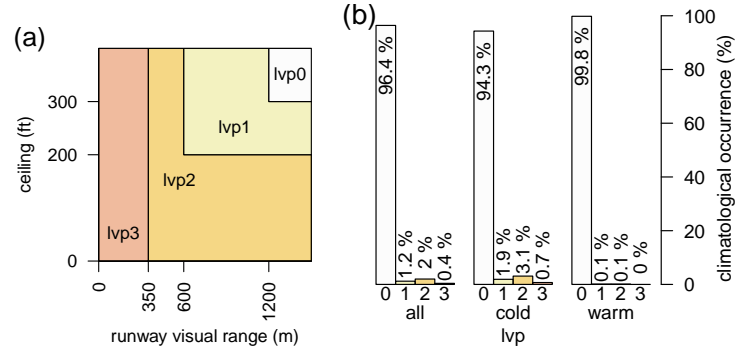


Figure 2: (a) Ceiling and runway visual range (*rvr*) thresholds for the *lvp* states. (b) Climatological occurrence of the three *lvp* and non-*lvp* (*lvp* 0) states at VIE for the whole year (all), and for the cold (September to March) and the warm season, respectively, over the period July 2008 until March 2017.

Table 2: Definition of *lvp* at VIE and associated capacities relative to a maximum of 40 aircraft per hour.

<i>lvp</i> state	<i>rvr</i>	Ceiling	Capacity
0			100 %
1	<1200 m or	<300 ft	75 %
2	<600 m or	<200 ft	60 %
3	<350 m		40 %

defined as the range over which the pilot of an aircraft on the center line of a runway can see the runway surface markings or the lights delineating the runway or identifying its center line (World Meteorological Organization 2006). The *rvr* is closely correlated to the horizontal visibility but is truncated with an upper limit of 2000 meters for higher visibilities. Vertical visibility is determined by the ceiling height. It is defined as the height of the lowest cloud layer covering more than half of the sky. The ceiling height is measured by ceilometers but finally determined by human observers while the *rvr* is measured automatically by transmissiometers. At VIE, *lvp* states from 1 to 3 depend nonlinearly on *rvr* and ceiling (see Tab. 2). The categories are ordered but not equidistant. The state *lvp* 0 indicates good visibility and ceiling and no restrictions for aviation with a maximum capacity of 40 aircraft per hour. At *lvp* 3 the airport operates at 40% of its maximum capacity (Tab. 2).

3.2. Climatology of the Low-Visibility Procedure States

A climatology of the *lvp* states was compiled to be used as one of the reference models. Fig. 2b shows the small proportion (about 4.5%) of *lvp* states 1 or higher for the whole year. Basically no low-visibility events occur during warm season. This seasonal pattern is typical for continental climates with topographical properties similar to the Vienna Basin (Egli *et al.* 2017). The challenge is to capture these relatively rare *lvp* events with a statistical approach. The significant annual cycle is also visible in Fig. 3 showing the occurrence probability of any

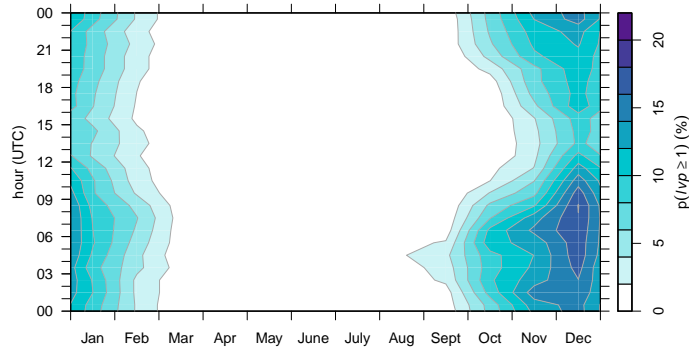


Figure 3: Contour plot of the annual and diurnal cycle of the climatological probability for poor visibility (lvp states ≥ 1) during the period July 2008 until March 2017.

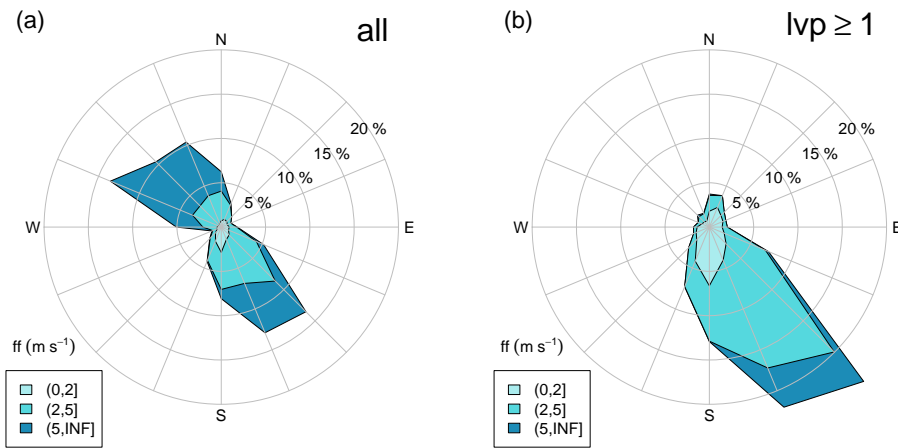


Figure 4: Wind rose plots of VIE for September to March (a) all cases and (b) lvp states ≥ 1 . The percentage of the different wind directions (color shadings) is divided into three wind speed clusters (see legend).

lvp 1-3. The annual maximum is during December. Furthermore, the diurnal maximum is during the morning hours, which coincides with one of the airports rush hours. Therefore these lvp state events have a high impact on airport operations. The diurnal minimum occurs during afternoon. The correlation of the lvp states with wind and thus advection from humidity sources is shown in Fig. 4. The two primary wind directions are from the southeast and the northwest (Fig. 4a). Low visibility states are associated predominantly with winds from the southeast, with a small peak in the northern sector, both of which are known moisture source regions (Fig. 4b). There is also a shift from higher to lower wind speeds (color shading) between the full data set (Fig. 4a) and the lvp data set (Fig. 4b), as lower wind speed generally favor the development of radiation fog. As a result of this climatology we focus only on the cold season beginning in September until the end of March.

3.3. Measurements and Model Configurations

Modeling and forecasting the lvp state with statistical methods requires information related

to low-visibility formation and dissipating processes. Since we are interested in lead times up to few hours only, we exclusively use standard meteorological measurements available at all larger airports as basis of this nowcasting system. The data come from the observation system of VIE in a 30-min resolution and are available from September 2008 to March 2017.

All potential predictors for the use in this statistical nowcasting system are shown in Tab. 1. The variable setup includes visibility measures, humidity measures, vertical temperature differences, wind information and climatological information. Most of the predictors are continuous variables except the *lvp* state, the precipitation variable, and the wind direction sector variables. These three variables are categorical variables with two or more levels (see Tab. 1). Precipitation is used to get information about humidity input to the atmosphere over the preceding 12 hours. The wind direction is split into two factor variables to capture the main fog wind directions based on the *lvp* climatology (Fig. 4b). Three visibility measures are used. The vertical visibility is captured by the ceiling and the horizontal visibility by the *rwr* and the meteorological visibility. The ceiling variable is set to 25000 feet for the cases where no ceiling was detected due to the absence of clouds. The *rwr* and the meteorological visibility provide similar information but *rwr* is truncated at 2000 meters whereas the meteorological visibility is truncated at 20000 meters. Two further variables are the relative humidity and the dew point depression which are non linearly related humidity measures. Information about the vertical stratification of the lowest air layers is provided by two vertical temperature differences. One is the difference between the 2 meter and surface temperature. The other is the difference between 2 meter and the tower temperature roughly 100 meter above ground. The only climatological variable used is the solar zenith angle, which represents the diurnal cycle of the *lvp* events. Dependent on the desired lead time we use the associated lag of the above shown predictors. For instance, if the lead time of +30-min is of interest the 30-min lags of the measurements are used as predictors, respectively for other lead times.

With measurements from 9 cold seasons in 30-min resolution we obtain a data set with roughly 85000 observations. Thus, we have a data set which is sufficient to estimate and verify the statistical models despite the rare occurrence of the low-visibility events (about 5000 observations). The forecasts are produced over the whole day with four lead times from +30 to +120-min. Four models, one for each lead time provide the forecasts. The model skill is based on out-of-sample verification by using ten-fold cross validation and estimates the model variance with bootstrapping (Sect. 2.3).

4. Results

4.1. Predictor Selection and Effects

The results of the stepwise model selection (Sect. 2.2) are shown in Fig. 5. The box plots are computed by the bootstrapped RPS mean values from the (a) +30-min and (b) the +120-min forecasts. Each box plot shows one forecast model. The model size increases from left to right by adding the annotated variable. The initial best guess is the intercept (=climatological) model (not shown in the graphic) which has the same skill for all lead times with an RPS ~ 0.032 . The stepwise algorithm indicates the best input variable combination and also shows the importance of the individual variables. The variable *lvp* is selected first within all models and is the most important variable. This confirms the findings from (Vislocky

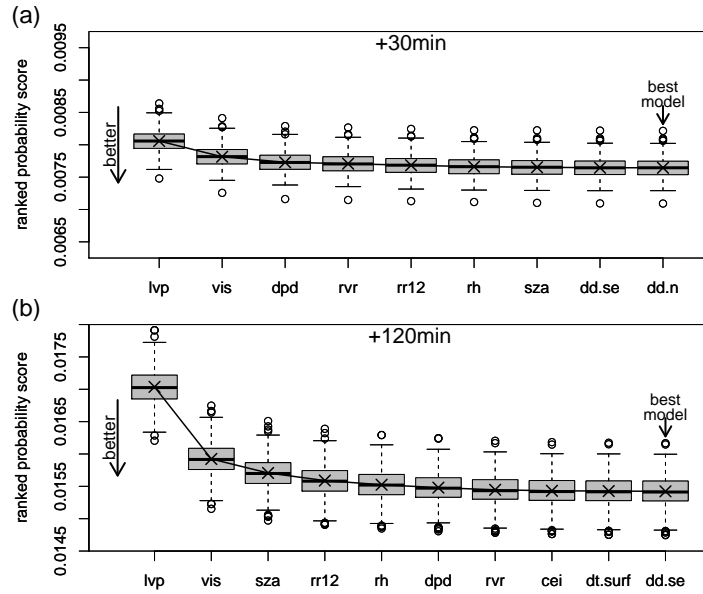


Figure 5: RPS model skill due to parameter selection with the forward stepwise selection method for the two lead times (a) +30-min and (b) +120-min. The model size increases from left to right adding the predictor (see Tab. 1) which improves the model most. The black line shows the median of the RPS. The last model indicates the best model where the stepwise algorithm stops.

and Fritsch 1997) that persistence is already a good benchmark for *lvp* forecasts. Horizontal visibility and humidity measures are selected next. The ceiling has a minor importance and is only selected in the +90 and +120-min model. The wind direction (*dd.se*) is selected in all models but improves the models only marginally. Of note is the increasing importance of solar zenith angle (*sza*) with lead time. The vertical temperature stratification (*dt.tow*) and wind speed (*ff*) were never selected as predictors. The variable selection implies that the information about upcoming low visibility is lead-time dependent and splits to a larger number of predictors for longer lead times.

The model coefficients β from the best models using the step wise algorithm are shown in Tab. 3. The sign of the coefficients indicate the sign of the effects, i.e. positive coefficients indicate higher probabilities of high *lvp* states when the predictor values increase. The first three parameters are the threshold coefficients α (see Eq. 2) between the four different *lvp* states. They can be interpreted as intercepts for the three thresholds. A further special feature of these models is the predictor *lvp*, which is an ordered variable with four levels. Within the model, *lvp* is replaced by three categorical variables which indicate if the latest *lvp* is higher *lvp* 0, higher *lvp* 1 and higher *lvp* 2. The predictors for the different lead times vary just slightly and remain the same for the lead times +90-min and longer. As already shown in Fig. 5 the model for the shortest lead time performs best with fewer predictors compared to longer lead times.

Fig. 6 shows the variation in *lvp* forecasts due to a individual predictor when all other predictors are kept constant at representative values. Representative values (red marks in Fig. 6) are values typical before low-visibility events (see Tab. 4). Fig. 6a shows that high *lvp* categories become more likely when *rvr*, visibility (*vis*) and dew point depression (*dpd*) decrease

Table 3: Coefficients β of the OLR models for the four different lead times. Contrary from Tab. 1 units from *rvr* and *vis* in km, unit from *cei* in kft.

	+30-min	+60-min	+90-min	+120-min
α_1	10.177	11.006	11.800	12.436
α_2	12.379	12.479	12.972	13.440
α_3	16.778	15.972	16.022	16.200
<i>lvp</i> _{1 0}	2.773	1.870	1.507	1.252
<i>lvp</i> _{2 1}	1.930	1.134	0.846	0.698
<i>lvp</i> _{3 2}	2.684	1.749	1.290	0.907
<i>rvr</i>	-0.668	-0.523	-0.426	-0.372
<i>cei</i>	–	–	0.010	0.012
<i>vis</i>	-0.316	-0.338	-0.313	-0.281
<i>rh</i>	0.134	0.131	0.131	0.132
<i>dpd</i>	-0.721	-0.731	-0.714	-0.649
<i>rr12</i>	-0.451	-0.511	-0.543	-0.573
<i>dt.surf</i>	–	0.063	0.071	0.059
<i>dd.se</i>	0.195	0.065	0.060	0.046
<i>dd.n</i>	0.099	–	–	–
<i>sza</i>	0.003	0.005	0.007	0.009

or when relative humidity (*rh*) or solar zenith angle (*sza*) increase. Decreasing values of *sza* indicate stronger radiative cooling and therefore favor low visibility. The predictor *lvp* has the strongest effect, which again confirms the importance of the persistence for short lead times. The effects for *lvp*, visibility and humidity parameters are nonlinear while all the other predictors produce linear effects. The factor variable *dd.se* with winds from the southeast sector has only a minor effect. However, the probability for high *lvp* states increases with winds from southeast compared to other directions. Surprisingly, the probability for high *lvp* states increases with higher ceiling height (Fig. 6b). An interpretation of this negative effect is that a low ceiling increases the long-wave downward radiation and prevents strong radiative cooling of the ground. Thus, it is less likely for a fog layer to develop. The negative effect of the precipitation within the preceding 12 hours (*rr12*) is also unexpected, but again a possible explanation is that the remaining precipitation clouds reduce radiative cooling of the ground.

Comparing the effect plots from the +30-min (Fig. 6a) and +120-min (Fig. 6b) forecasts illustrate in general lower probabilities for the higher *lvp* states at the +120-min forecasts. This indicates that the OLR model forecasts tend to converge to climatology with increasing lead times. Furthermore, the large effect of the predictor *lvp* (see Fig. 6a) indicates that the forecast information is concentrated on this predictor at the lead time +30-min. In contrast at the lead time +120-min the effect of *lvp* becomes weaker while the other effects remain similar. This behavior is consistent with the model coefficients shown in Tab. 3, where the coefficients for *lvp* decrease with increasing lead times while the coefficients of several other predictors, e.g. *sza*, *dpd*, or *dt.surf*, increase or at least remain approximately unchanged.

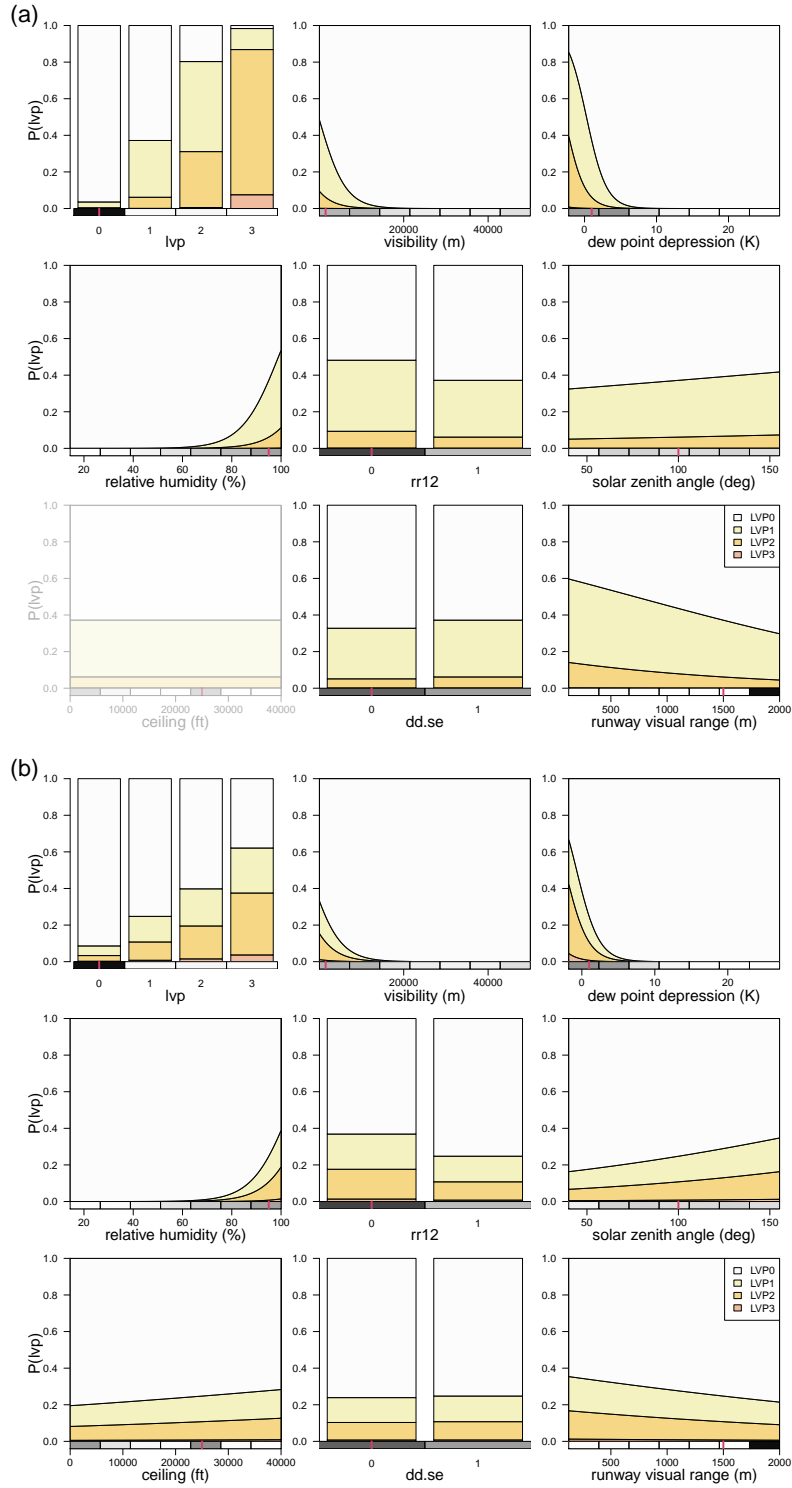


Figure 6: Effect of the predictors for the (a) +30-min and (b) +120-min forecast. The four colors indicate the probability for each *lvp* state dependent on the particular predictors. The observed distribution of the particular predictors is shown as bar with gray shading below each effect plot. Typical conditions before an *lvp* event (red markers see Tab. 4) are chosen as representative values for the marginal effects. The effect plot with the fade out indicate that these predictors have not been selected in the particular model and therefore have constant effects.

Table 4: Fixed values for the marginal effects shown in Fig. 6.

parameter	value	parameter	value
<i>lvp</i>	0	<i>rr12</i>	yes
<i>rvr</i>	1500	<i>dt.surf</i>	1
<i>cei</i>	25000	<i>dd.se</i>	yes
<i>vis</i>	1500	<i>dd.n</i>	no
<i>rh</i>	95	<i>sza</i>	100
<i>dpd</i>	1		

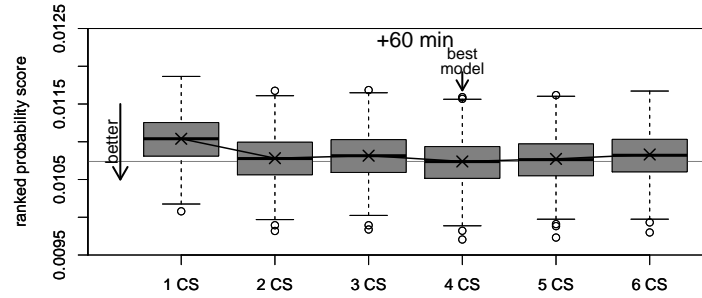


Figure 7: Model stability of the OLR due to variation of training data set length from 1 to 6 cold seasons (CS). RPS values are out-of-sample on a 3 cold season test data set for the +60-min forecast. Persistence is shown as reference model.

4.2. Model Stability

To estimate the sensitivity of the OLR models to the training data set length Fig. 7 shows RPS of +60-min forecasts based on 1 up to 6 cold seasons training data. The forecast are evaluated out-of-sample for 3 cold seasons. The box plots for the OLR models indicate a slight model improvement with an increasing training data set from 1 cold season to 2 cold seasons. The OLR models based on 2 or more cold seasons have similar skill but the model based on 4 cold seasons has the best RPS. The shortest training data set length of 1 cold season includes about 5000 observations which is already an appropriate data set to estimate an OLR model with the applied model configuration. Taken together, forecast accuracy improves with increasing training data set length up to 2 cold seasons and remains similar for longer training data sets.

4.3. Model Performance

The results of the final *lvp* forecasts are shown in Fig. 8. The models are based on data of 9 cold seasons using the best model configuration resulting from the stepwise predictor selection. Fig. 8a illustrates the verification for OLR compared to the climatology and the persistence forecasts. The OLR outperforms climatology and persistence for each lead time. The skill of all forecasts decreases with increasing lead time but the benefit of OLR over persistence increases. OLR and persistence converge to climatology but within the lead times considered here, the performance of climatology as a forecast of *lvp* is clearly worse. The spread of the RPS indicates the uncertainty due to the model estimation determined by bootstrapping (see

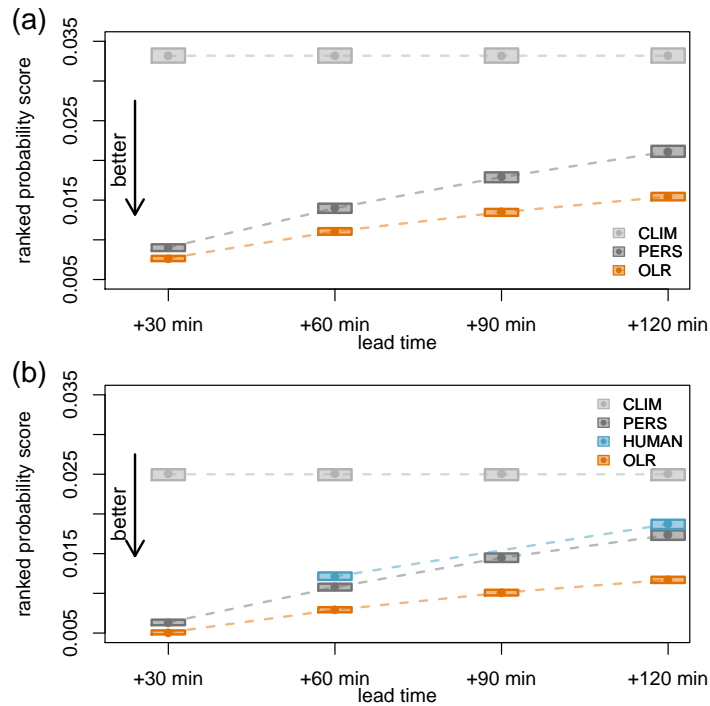


Figure 8: Model skill based on RPS as function of the lead time for (a) the full data set and (b) for the subset where the human observations are available. The dots represent the 50 percent quantile, the shadings in the lighter color show the 95% confidence interval. The model acronyms stand for climatology (CLIM), persistence (PERS), ordered logistic regression (OLR), and the human forecaster (HUMAN). The human forecasts are available from September 2012 to March 2017 for the lead times +60-min and +120-min between 5 and 20 UTC only.

Sect. 2.3). The uncertainty is smallest for the OLR model as compared to persistence and climatology and this uncertainty does not vary substantially with lead time.

Human forecasts are available as third reference (Fig. 8b), albeit only over a shorter period (September 2012 to March 2017), part of the day (5 to 20 UTC) and 60 minutes lead time intervals instead of 30 minutes. The human forecast is comparable to persistence, therefore OLR also outperforms the human forecasts. Due to the rare occurrence of *lvp* events, and the larger number of *lvp* 0 during the day, when considering only the daytime OLR, persistence and climatology perform better compared to the results on the full data set in Fig. 8a.

5. Discussion and Conclusion

Predicting *lvp* (low-visibility procedure) states is very challenging because small changes in the physical drivers of low-visibility formation and dissipation processes have a large impact on *lvp*. The OLR (ordered logistic regression) model has been used to develop a probabilistic nowcasting system producing direct airport-specific, low-visibility forecasts tested at Vienna International Airport.

The OLR model leads to promising forecasts and outperforms persistence and climatology,

although persistence offers predictions to be reckoned with especially at the very short lead times. This nowcasting approach is also competitive with human forecasts. A benefit of this method is to produce probabilistic forecasts of a variable relevant for to the end user within one consistent model. The probabilistic forecasts of such a model allows different end users, e.g., air traffic controllers or air traffic managers, to extract the information regarding their requirements.

The variable influencing the forecasts most is *lvp*. Additional important variables are the horizontal visibility and the humidity measures. Contrary to our expectations wind speed and the temperature stratification of the first 100 meters above the ground did not contribute information for future *lvp* states at Vienna International Airport.

The presented data-driven nowcasting system is transferable to other airports but it needs to consider site-specific predictors. For instance, considering the *lvp* climatology guides us to focus on the cold season only and justifies omission of the warm season, during which low-visibility events almost never occur. The nowcasting system uses the available measurements on the site and its vicinity and objectively selects the important variables to form the best model. Dependent on the regime of the site (e.g., coastal or continental) different physical effects drive the development and dissipation of low-visibility conditions. These effects need to be adequately represented in the model variables in order to achieve accurate forecasts. The usage of additional model inputs and increase of the forecast temporal resolution would have an additional gain for the operational use of the nowcasting system (cf. Dietz *et al.* 2017).

In summary, these findings suggest that the OLR model is a suitable tool for computationally fast probabilistic visibility forecasts which can provide useful guidance to aviation decision makers.

6. Acknowledgments

This study has been supported by the Austrian Research Promotion Agency FFG 843457 and benefits greatly from Markus Kerschbaum, Andreas Lanzinger and the operational personnel of Vienna International Airport. The authors thank the Austrian Aviation Agency ACG for providing access to the data.

References

- Bartoková I, Bott A, Bartok J, Gera M (2015). “Fog Prediction for Road Traffic Safety in a Coastal Desert Region: Improvement of Nowcasting Skills by the Machine-Learning Approach.” *Boundary-Layer Meteorology*, **157**(3), 501–516. ISSN 0006-8314, 1573-1472. doi:10.1007/s10546-015-0069-x.
- Bergot T (2016). “Large-Eddy Simulation Study of the Dissipation of Radiation Fog.” *Quarterly Journal of the Royal Meteorological Society*, **142**(695), 1029–1040. ISSN 1477-870X. doi:10.1002/qj.2706.
- Bocchieri JR, Glahn HR (1972). “Use of Model Output Statistics for Predicting Ceiling Height.” *Monthly Weather Review*, **100**(12), 869–879. ISSN 0027-0644. doi:10.1175/1520-0493(1972)100<0869:UOMOSF>2.3.CO;2.

- Bremnes JB, Michaelides SC (2007). “Probabilistic Visibility Forecasting Using Neural Networks.” *Pure and Applied Geophysics*, **164**(6-7), 1365–1381. ISSN 0033-4553, 1420-9136. doi:10.1007/s00024-007-0223-6.
- Christensen RHB (2015). “ordinal—Regression Models for Ordinal Data.” R package version 2015.6-28. <http://www.cran.r-project.org/package=ordinal/>.
- Dietz SJ, Kneringer P, Mayr GJ, Zeileis A (2017). “Forecasting Low-Visibility Procedure States with Tree-Based Statistical Methods.” *Working papers, Faculty of Economics and Statistics, University of Innsbruck*. URL <https://eeecon.uibk.ac.at/wopec2/repec/inn/wpaper/2017-22.pdf>.
- Dutta D, Chaudhuri S (2015). “Nowcasting Visibility During Wintertime Fog over the Airport of a Metropolis of India: Decision Tree Algorithm and Artificial Neural Network Approach.” *Natural Hazards*, **75**(2), 1349–1368. ISSN 0921-030X. doi:10.1007/s11069-014-1388-9.
- Egli S, Thies B, Drönner J, Cermak J, Bendix J (2017). “A 10 Year Fog and Low Stratus Climatology for Europe Based on Meteosat Second Generation Data.” *Quarterly Journal of the Royal Meteorological Society*, **143**(702), 530–541. ISSN 1477-870X. doi:10.1002/qj.2941.
- Fabbian D, de Dear R, Lelleyett S (2007). “Application of Artificial Neural Network Forecasts to Predict Fog at Canberra International Airport.” *Weather and Forecasting*, **22**(2), 372–381. ISSN 0882-8156. doi:10.1175/WAF980.1.
- Gultepe I, Hansen B, Cober SG, Pearson G, Milbrandt JA, Platnick S, Taylor P, Gordon M, Oakley JP (2009). “The Fog Remote Sensing and Modeling Field Project.” *Bulletin of the American Meteorological Society*, **90**(3), 341–359. ISSN 0003-0007. doi:10.1175/2008BAMS2354.1.
- Gultepe I, Tardif R, Michaelides SC, Cermak J, Bott A, Bendix J, Müller MD, Pagowski M, Hansen B, Ellrod G, Jacobs W, Toth G, Cober SG (2007). “Fog Research: A Review of Past Achievements and Future Perspectives.” *Pure and Applied Geophysics*, **164**(6-7), 1121–1159. ISSN 0033-4553, 1420-9136. doi:10.1007/s00024-007-0211-x.
- Haefelin M, Bergot T, Elias T, Tardif R, Carrer D, Chazette P, Colomb M, Drobinski P, Dupont E, Dupont JC, Gomes L, Musson-Genon L, Pietras C, Plana-Fattori A, Protat A, Rangognio J, Raut JC, Rémy S, Richard D, Sciare J, Zhang X (2009). “PARISFOG: Shedding New Light on Fog Physical Processes.” *Bulletin of the American Meteorological Society*, **91**(6), 767–783. ISSN 0003-0007. doi:10.1175/2009BAMS2671.1.
- Herman GR, Schumacher RS (2016). “Using Reforecasts to Improve Forecasting of Fog and Visibility for Aviation.” *Weather and Forecasting*, **31**(2), 467–482. ISSN 0882-8156. doi:10.1175/WAF-D-15-0108.1.
- Hilliker JL, Fritsch JM (1999). “An Observations-Based Statistical System for Warm-Season Hourly Probabilistic Forecasts of Low Ceiling at the San Francisco International Airport.” *Journal of Applied Meteorology*, **38**(12), 1692–1705. ISSN 0894-8763. doi:10.1175/1520-0450(1999)038<1692:A0BSSF>2.0.CO;2.

- Marzban C, Leyton S, Colman B (2007). “Ceiling and Visibility Forecasts via Neural Networks.” *Weather and Forecasting*, **22**(3), 466–479. ISSN 0882-8156. doi:10.1175/WAF994.1.
- Murphy AH, Winkler RL (1984). “Probability Forecasting in Meteorology.” *Journal of the American Statistical Association*, **79**(387), 489–500. ISSN 0162-1459. doi:10.1080/01621459.1984.10478075.
- Vislocky RL, Fritsch JM (1997). “An Automated, Observations-Based System for Short-Term Prediction of Ceiling and Visibility.” *Weather and Forecasting*, **12**(1), 31–43. ISSN 0882-8156. doi:10.1175/1520-0434(1997)012<0031:AAOBSF>42.0.CO;2.
- Wilks DS (2011). *Statistical Methods in the Atmospheric Sciences*. Academic Press, Oxford; Waltham, MA. ISBN 978-0-12-385022-5 0-12-385022-3.
- Winkelmann R, Boes S (2006). *Analysis of Microdata*. Springer, Berlin. ISBN 3-540-29605-0 978-3-540-29605-8.
- World Meteorological Organization (2006). “Guide to Meteorological Instruments and Methods of Observation.” *Tech. rep*, Genf, Schweiz: Secretariat of the WMO.

Affiliation:

Philipp Kneringer, Sebastian J. Dietz, Georg J. Mayr
Institute of Atmospheric and Cryospheric Sciences
Faculty of Geo- and Atmospheric Sciences
Universität Innsbruck
Innrain 52
6020 Innsbruck, Austria
E-mail: Philipp.Kneringer@uibk.ac.at, Sebastian.Dietz@uibk.ac.at, Georg.Mayr@uibk.ac.at

Achim Zeileis
Department of Statistics
Faculty of Economics and Statistics
Universität Innsbruck
Universitätsstraße 15
6020 Innsbruck, Austria
E-mail: Achim.Zeileis@uibk.ac.at

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

<http://eeecon.uibk.ac.at/wopec/>

- 2017-21 **Philipp Kneringer, Sebastian J. Dietz, Georg J. Mayr, Achim Zeileis:** Probabilistic nowcasting of low-visibility procedure states at Vienna International Airport during cold season
- 2017-20 **Loukas Balafoutas, Brent J. Davis, Matthias Sutter:** How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior
- 2017-19 **Martin Geiger, Richard Hule:** The role of correlation in two-asset games: Some experimental evidence
- 2017-18 **Rudolf Kerschbamer, Daniel Neururer, Alexander Gruber:** Do the altruists lie less?
- 2017-17 **Meike Köhler, Nikolaus Umlauf, Sonja Greven:** Nonlinear association structures in flexible Bayesian additive joint models
- 2017-16 **Rudolf Kerschbamer, Daniel Muller:** Social preferences and political attitudes: An online experiment on a large heterogeneous sample
- 2017-15 **Kenneth Harttgen, Stefan Lang, Judith Santer, Johannes Seiler:** Modeling under-5 mortality through multilevel structured additive regression with varying coefficients for Asia and Sub-Saharan Africa
- 2017-14 **Christoph Eder, Martin Halla:** Economic origins of cultural norms: The case of animal husbandry and bastardy
- 2017-13 **Thomas Kneib, Nikolaus Umlauf:** A Primer on Bayesian Distributional Regression
- 2017-12 **Susanne Berger, Nathaniel Graham, Achim Zeileis:** Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R
- 2017-11 **Natalia Danzer, Martin Halla, Nicole Schneeweis, Martina Zweimüller:** Parental leave, (in)formal childcare and long-term child outcomes
- 2017-10 **Daniel Muller, Sander Renes:** Fairness views and political preferences - Evidence from a large online experiment
- 2017-09 **Andreas Exenberger:** The Logic of Inequality Extraction: An Application to Gini and Top Incomes Data

- 2017-08 **Sibylle Puntscher, Duc Tran Huy, Janette Walde, Ulrike Tappeiner, Gottfried Tappeiner:** The acceptance of a protected area and the benefits of sustainable tourism: In search of the weak link in their relationship
- 2017-07 **Helena Fornwagner:** Incentives to lose revisited: The NHL and its tournament incentives
- 2017-06 **Loukas Balafoutas, Simon Czermak, Marc Eulerich, Helena Fornwagner:** Incentives for dishonesty: An experimental study with internal auditors
- 2017-05 **Nikolaus Umlauf, Nadja Klein, Achim Zeileis:** BAMLSS: Bayesian additive models for location, scale and shape (and beyond)
- 2017-04 **Martin Halla, Susanne Pech, Martina Zweimüller:** The effect of statutory sick-pay on workers' labor supply and subsequent health
- 2017-03 **Franz Buscha, Daniel Müller, Lionel Page:** Can a common currency foster a shared social identity across different nations? The case of the Euro.
- 2017-02 **Daniel Müller:** The anatomy of distributional preferences with group identity
- 2017-01 **Wolfgang Frimmel, Martin Halla, Jörg Paetzold:** The intergenerational causal effect of tax evasion: Evidence from the commuter tax allowance in Austria
- 2016-33 **Alexander Razen, Stefan Lang, Judith Santer:** Estimation of spatially correlated random scaling factors based on Markov random field priors
- 2016-32 **Meike Köhler, Nikolaus Umlauf, Andreas Beyerlein, Christiane Winkler, Anette-Gabriele Ziegler, Sonja Greven:** Flexible Bayesian additive joint models with an application to type 1 diabetes research
- 2016-31 **Markus Dabernig, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Simultaneous ensemble post-processing for multiple lead times with standardized anomalies
- 2016-30 **Alexander Razen, Stefan Lang:** Random scaling factors in Bayesian distributional regression models with an application to real estate data
- 2016-29 **Glenn Dutcher, Daniela Glätzle-Rützler, Dmitry Ryvkin:** Don't hate the player, hate the game: Uncovering the foundations of cheating in contests
- 2016-28 **Manuel Gebetsberger, Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Tricks for improving non-homogeneous regression for probabilistic precipitation forecasts: Perfect predictions, heavy tails, and link functions
- 2016-27 **Michael Razen, Matthias Stefan:** Greed: Taking a deadly sin to the lab

- 2016-26 **Florian Wickelmaier, Achim Zeileis:** Using recursive partitioning to account for parameter heterogeneity in multinomial processing tree models
- 2016-25 **Michel Philipp, Carolin Strobl, Jimmy de la Torre, Achim Zeileis:** On the estimation of standard errors in cognitive diagnosis models
- 2016-24 **Florian Lindner, Julia Rose:** No need for more time: Intertemporal allocation decisions under time pressure
- 2016-23 **Christoph Eder, Martin Halla:** The long-lasting shadow of the allied occupation of Austria on its spatial equilibrium
- 2016-22 **Christoph Eder:** Missing men: World War II casualties and structural change
- 2016-21 **Reto Stauffer, Jakob Messner, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Ensemble post-processing of daily precipitation sums over complex terrain using censored high-resolution standardized anomalies *published in Monthly Weather Review*
- 2016-20 **Christina Bannier, Eberhard Feess, Natalie Packham, Markus Walzl:** Incentive schemes, private information and the double-edged role of competition for agents
- 2016-19 **Martin Geiger, Richard Hule:** Correlation and coordination risk
- 2016-18 **Yola Engler, Rudolf Kerschbamer, Lionel Page:** Why did he do that? Using counterfactuals to study the effect of intentions in extensive form games
- 2016-17 **Yola Engler, Rudolf Kerschbamer, Lionel Page:** Guilt-averse or reciprocal? Looking at behavioural motivations in the trust game
- 2016-16 **Esther Blanco, Tobias Haller, James M. Walker:** Provision of public goods: Unconditional and conditional donations from outsiders
- 2016-15 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** Predictive bookmaker consensus model for the UEFA Euro 2016
- 2016-14 **Martin Halla, Harald Mayr, Gerald J. Pruckner, Pilar García-Gómez:** Cutting fertility? The effect of Cesarean deliveries on subsequent fertility and maternal labor supply
- 2016-13 **Wolfgang Frimmel, Martin Halla, Rudolf Winter-Ebmer:** How does parental divorce affect children's long-term outcomes?
- 2016-12 **Michael Kirchler, Stefan Palan:** Immaterial and monetary gifts in economic transactions. Evidence from the field

- 2016-11 **Michel Philipp, Achim Zeileis, Carolin Strobl:** A toolkit for stability assessment of tree-based learners
- 2016-10 **Loukas Balafoutas, Brent J. Davis, Matthias Sutter:** Affirmative action or just discrimination? A study on the endogenous emergence of quotas published in *Journal of Economic Behavior and Organization*
- 2016-09 **Loukas Balafoutas, Helena Fornwagner:** The limits of guilt
- 2016-08 **Markus Dabernig, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Spatial ensemble post-processing with standardized anomalies
- 2016-07 **Reto Stauffer, Jakob W. Messner, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Spatio-temporal precipitation climatology over complex terrain using a censored additive regression model
- 2016-06 **Michael Razen, Jürgen Huber, Michael Kirchler:** Cash inflow and trading horizon in asset markets
- 2016-05 **Ting Wang, Carolin Strobl, Achim Zeileis, Edgar C. Merkle:** Score-based tests of differential item functioning in the two-parameter model
- 2016-04 **Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Non-homogeneous boosting for predictor selection in ensemble post-processing
- 2016-03 **Dietmar Fehr, Matthias Sutter:** Gossip and the efficiency of interactions
- 2016-02 **Michael Kirchler, Florian Lindner, Utz Weitzel:** Rankings and risk-taking in the finance industry
- 2016-01 **Sibylle Puntischer, Janette Walde, Gottfried Tappeiner:** Do methodical traps lead to wrong development strategies for welfare? A multilevel approach considering heterogeneity across industrialized and developing countries

University of Innsbruck

Working Papers in Economics and Statistics

2017-21

Philipp Kneringer, Sebastian J. Dietz, Georg J. Mayr, Achim Zeileis

Probabilistic nowcasting of low-visibility procedure states at Vienna International Airport during cold season

Abstract

Airport operations are sensitive to visibility conditions. Low-visibility events may lead to capacity reduction, delays and economic losses. Different levels of low-visibility procedures (*lvp*) are enacted to ensure aviation safety. A nowcast of the probabilities for each of the *lvp* categories helps decision makers to optimally schedule their operations. An ordered logistic regression (OLR) model is used to forecast these probabilities directly. It is applied to cold season forecasts at Vienna International Airport for lead times of 30-min out to two hours. Model inputs are standard meteorological measurements. The skill of the forecasts is assessed by the ranked probability score. OLR outperforms persistence, which is a strong contender at the shortest lead times. The ranked probability score of the OLR is even better than the one of nowcasts from human forecasters. The OLR-based nowcasting system is computationally fast and can be updated instantaneously when new data become available.

ISSN 1993-4378 (Print)

ISSN 1993-6885 (Online)