

Plavcan, David; Mayr, Georg J.; Zeileis, Achim

Working Paper

Automatic and probabilistic foehn diagnosis with a statistical mixture model

Working Papers in Economics and Statistics, No. 2013-22

Provided in Cooperation with:

Institute of Public Finance, University of Innsbruck

Suggested Citation: Plavcan, David; Mayr, Georg J.; Zeileis, Achim (2013) : Automatic and probabilistic foehn diagnosis with a statistical mixture model, Working Papers in Economics and Statistics, No. 2013-22, University of Innsbruck, Research Platform Empirical and Experimental Economics (eecon), Innsbruck

This Version is available at:

<https://hdl.handle.net/10419/101080>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Automatic and probabilistic foehn diagnosis with a statistical mixture model

David Plavcan, Georg J. Mayr,
Achim Zeileis

Working Papers in Economics and Statistics

2013-22

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact Address:
University of Innsbruck
Department of Public Finance
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.

Automatic and Probabilistic Foehn Diagnosis with a Statistical Mixture Model

David Plavcan
Universität Innsbruck

Georg J. Mayr
Universität Innsbruck

Achim Zeileis
Universität Innsbruck

Abstract

Diagnosing foehn winds from weather station data downwind of topographic obstacles requires distinguishing them from other downslope winds, particularly nocturnal ones driven by radiative cooling. We present an automatic classification scheme to obtain reproducible results that include information about the (un)certainty of the diagnosis. A statistical mixture model separates foehn and no-foehn winds in a measured time series of wind. In addition to wind speed and direction, it accommodates other physically meaningful classifiers such as relative humidity or the (potential) temperature difference to an upwind station (e.g., near the crest). The algorithm was tested for the central Alpine Wipp Valley against human expert classification and a previous objective method (Drechsel and Mayr 2008), which the new method outperforms. Climatologically, using only wind information gives nearly identical foehn frequencies as when using additional covariables, making the method suitable for comparable foehn climatologies all over the world where station data are available for at least one year.

Keywords: foehn wind, foehn diagnosis, finite mixture model, model-based clustering.

1. Introduction

Foehn is ‘a wind (which is) warmed and dried by descent, in general on the lee side of a mountain’ (WMO 1992). As such it is as ubiquitous as the mountain ranges of the earth. Diagnosing when foehn blows requires distinguishing it from other downslope flows, which can be non-trivial. An accurate diagnosis is a prerequisite for studies dealing with mechanisms, climatologies and effects of foehn (e.g., on air quality). This paper describes a method to objectively diagnose both foehn occurrence and the (un)certainty of the diagnosis.

The conceptual model of foehn that fits best the results of the latest large field campaigns (Mesoscale Alpine Programme MAP, Mayr and Armi 2008; Terrain-Induced Rotor Experiment TREX, Armi and Mayr 2011) is depicted in Figure 1a. Other explanations of foehn were recently summarized in Jackson, Mayr, and Vosper (2013). The descent of upstream air is possible when the virtual potential temperature of the descending upstream air mass is equal or lower than the air in the downstream valley. Diurnal heating in the lee or upstream cold air advection may cause such a situation (Mayr and Armi 2010). While flowing over the mountain the foehn layer changes from a subcritical to supercritical state (Durran 1990; Armi and Farmer 2002). Aloft, a nearly mixed and stagnating layer separates foehn from flow further aloft.

Other wind systems, which flow down topography and might be mistaken for foehn are

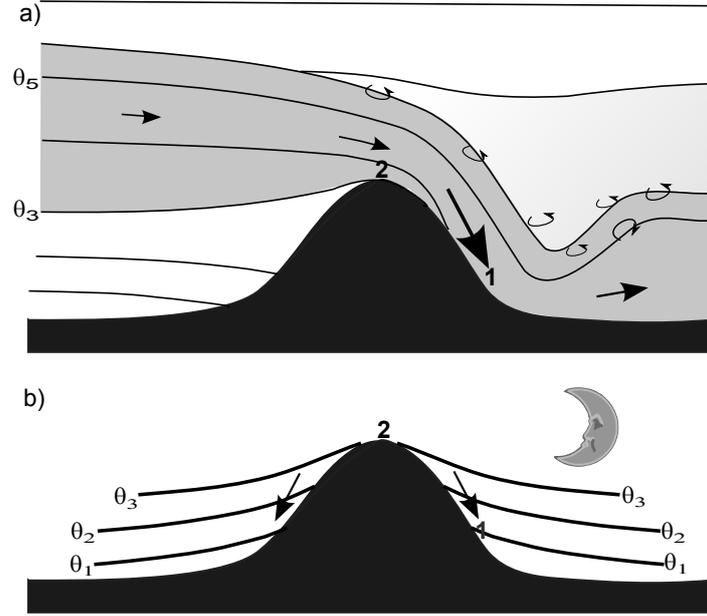


Figure 1: Schematic representation of (a) foehn and (b) radiatively driven downslope flow, respectively. In (a) dark gray shading indicates the foehn layer and the arrows indicating qualitatively the flow direction and speed. Turbulent mixing at the ground and at the upper edge of the foehn flow is indicated by curved arrows. The light gray shaded area above the foehn flow indicates the mixed residual layer. The solid lines in both parts are isentropes.

downslope/down-valley flows from nocturnal radiative cooling (Defant 1949; Whiteman 2000) depicted in Figure 1b, and shorter-lived events like frontal passages (Prandtl 1944) or convective outflows.

The traditional method for detecting the onset of foehn from observation data at a single (valley) station (location 1 in Figure 1) is to analyze temporal changes of temperature, relative humidity and wind. Wind speed must increase and direction must be down the local terrain, temperature must increase and relative humidity decrease (Conrad 1936). This method normally requires a human expert, making results non-reproducible and very time-consuming. And even a human expert has difficulty distinguishing weaker foehn flows from radiatively driven downslope flows.

Diagnostic accuracy increases with the availability of a second station further upslope at (or near) the crest (location 2 in Figure 1) with which one can exploit the physical differences between foehn and downslope winds. During foehn (Figure 1a) the upstream air mass descends so that potential temperatures at stations 1 (valley) and 2 (crest) will almost¹ be identical ($\Delta\theta = \theta_{crest} - \theta_{valley} \approx 0$). In contrast to this, radiative cooling during a nocturnal downslope wind situation (Figure 1b) leads to stable stratification ($\Delta\theta > 0$; Whiteman 2000).

The availability of a second station both improves the manual classification accuracy and enables the formulation of an objective diagnostic method. Such an objective foehn classifi-

¹Small differences may occur due to, e.g., imperfect sensor calibration, mixing in of nocturnally-cooled air or evaporation of precipitation into the downstream foehn layer. They are discussed at length in Vergeiner (2004).

cation (OFC) method was developed for the first time during MAP. It is described in detail in Vergeiner (2004) and summarized in Drechsel and Mayr (2008) and was further developed for Swiss Stations by Dürr (2008) to also include relative humidity and gust thresholds. Despite being an objective method, two parameters have to be determined manually prior to the application of the method. The determination of the topographic downslope direction range is trivial. More difficult and subjective is setting a hard threshold for the potential temperature difference $\Delta\theta$ between the crest and valley stations below which foehn will be classified. Having to determine the hard threshold individually for each location is the main drawback of the method.

Even a human expert using a second station will encounter events that are not clearly distinguishable and that are classified by the OFC without giving an indication of the high uncertainty. An improved objective classification method should consequently avoid both drawbacks by making it unnecessary to determine anything except possibly the topographic downslope direction by hand and by moving from a deterministic to a probabilistic diagnosis.

2. Foehn diagnosis with a statistical mixture model

Separation foehn from radiatively-driven downslope winds is a typical classification problem, for which statistical science provides several methods. Since we want the method to be applicable to any location we chose unsupervised classification. A mixture model allows to both estimate the *unknown* density distribution of foehn and no-foehn cases from the observed and thus *known* density of all cases and the probability that observation i belongs to one of the two classes (McLachlan and Peel 2000; Hastie, Tibshirani, and Friedman 2009). An example is shown in Figure 2.

In our case, the mixture model for the wind speed distribution $f(s)$ consists of two normally distributed components, downslope wind and foehn:

$$f(s_i) = \underbrace{(\pi - 1)\phi(s_i|\mu_1, \sigma_1^2)}_{\text{I downslope wind}} + \underbrace{\pi\phi(s_i|\mu_2, \sigma_2^2)}_{\text{II foehn}} \quad (1)$$

where

$$\phi(s_i|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(s_i - \mu)^2}{2\sigma^2}\right\} \quad (2)$$

denotes the Gaussian density function with mean μ and variance σ^2 . π is the prior probability for foehn class and the complementary probability $\pi - 1$ is the prior for downslope wind class. We can properly label these two classes since foehn is stronger than nocturnal downslope wind, i.e., $\mu_2 > \mu_1$. Note, that in general any component density can be used in place of the Gaussian or that wind speed can be transformed (e.g., taking the square root) before being used in the mixture model, which, however, did not improve our model.

The probability that one measurement of wind speed s_i belongs to the foehn cluster is given by the proportion of the probability density function for foehn to the total wind speed distribution and can be calculated by the ratio of II/(I + II) from the components of Equation 1.

Additional physically meaningful classifiers such as the potential temperature difference between valley and crest or relative humidity called “concomitant” variables (Dayton and Macready 1986) may be used to improve the estimated distributions. Then the prior probability π changes from being constant to a (in our two component case) binomial logit model.

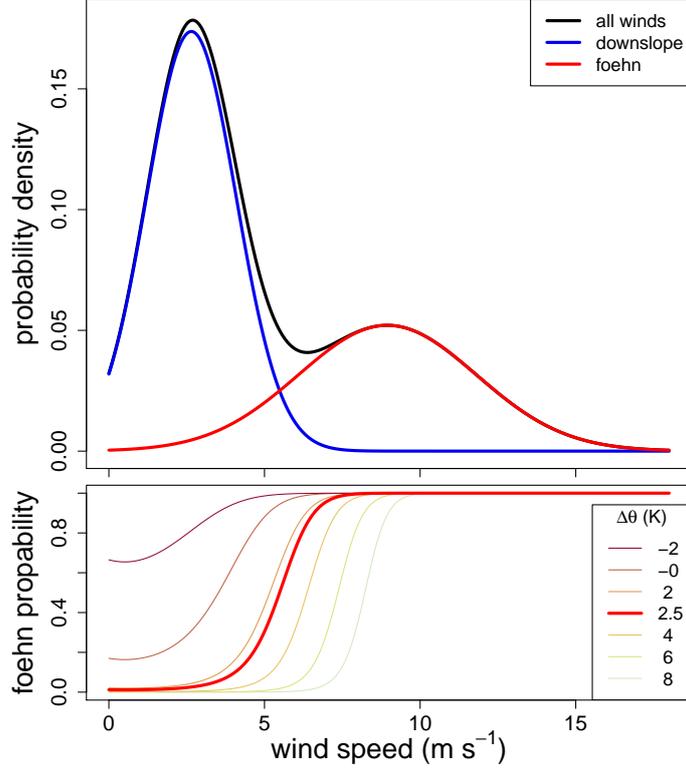


Figure 2: (top) Distribution of wind speed (black) for winds from the foehn sector at the Wipp Valley station with distributions of foehn (red) and no-foehn (blue) fitted by the mixture model M2 for wind speed with the difference in potential temperature between crest and valley station as concomitant. (bottom) Foehn probability for model M2 as function of wind speed and potential temperature difference $\Delta\theta$ (isolines). The red line represents the mean state for the whole distribution (top).

The prior for the second component (foehn) is then

$$\pi(\mathbf{x}) = \text{logit}^{-1}(\mathbf{x}^\top \boldsymbol{\beta}) \quad (3)$$

where the vector

$$\mathbf{x} = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \end{pmatrix} \quad (4)$$

contains the used concomitant variables x_i after its first component which is 1. The corresponding coefficients are written in the vector $\boldsymbol{\beta}$.

We used the **flexmix** package (Leisch 2004; Grün and Leisch 2008) in the programming language R (R Core Team 2013) to fit the mixture model. The only pre-processing was the application of a wind direction filter: Only wind from a 180° sector centered along the topographic downslope direction could be classified as ‘foehn’. The downslope direction can either be determined manually or automatically (Pelletier 2013).

3. Application

This automatic classification method was then tested for the central alpine Wipp Valley. It is a typical foehn valley in the Alps, orientated perpendicular to a gap in the main crest. The indentation in the crest line increases foehn occurrence (e.g., Jackson *et al.* 2013). It was a study area of MAP and has since then been instrumented to allow testing of different foehn classification schemes. We use one crest/gap station at 2.1 km amsl and one valley station at 1.1 km amsl, which is 21 km downstream. Mayr *et al.* (2004) describe topography and stations in detail. 14 years of measurements averaged over 10-minute intervals of wind speed and direction, temperature, relative humidity and pressure are available. Joint data availability is 91%.

Three different mixture models exploiting an increasing amount of measurement information were applied to probabilistically diagnose foehn occurrence at the valley station. The first model (M1) only uses wind speed at the valley station itself. M2 uses the measurements at the crest station to include the potential temperature difference $\Delta\theta$ between crest and valley as the first concomitant variable. Finally, M3 adds relative humidity at the valley station as a second concomitant (cf. Table 1).

The behavior of all three models will be first explored in a case study and then in a foehn climatology over the whole 14-year data set.

3.1. Case study: Shallow foehn on 27/28 October 2005

For a subjective (human-expert) verification of these foehn models more than 50 case studies were examined. Most of them, especially the stronger foehn events are well captured by all models with only minor differences. For illustration, we present one where the differences between models M1–M3 are especially *pronounced* and where the ending of the foehn period is difficult to analyze even by a human expert.

On 27 and 28 October 2005 the Alps separated a cold air mass to the south from warmer air to the north, which caused (shallow) south foehn through the lower alpine passes. At the Wipp Valley station, the foehn event started at 1400 UTC 27 October 2005 in our subjective classification (orange shading), shortly after the direction shifted from northwest (= upvalley) to southeast (= downvalley) with a strong increase in temperature and a decline in relative humidity (Figure 3). It ended the following night at 0150 UTC with reverse but

Model	Variables	No foehn (%)		Foehn (%)	
		≥ 99	50–99	50–99	≥ 99
M1	s	49.46	30.92	5.27	14.36
M2	$s, \Delta\theta$	71.82	9.60	4.29	14.30
M3	$s, \Delta\theta, rh$	71.71	8.63	4.10	15.57
OFC	$s, \Delta\theta, d$	77.67		14.51	

Table 1: Summary of mixture models M1-M3 with variables used for fitting (wind speed s and direction d , relative humidity rh and difference in potential temperature ($\Delta\theta = \theta_{crest} - \theta_{valley}$) and relative frequency of no-foehn and foehn classification with two levels of certainty each for the 14-year data set at the valley station. Additionally, results for the previous non-probabilistic objective algorithm OFC are shown.

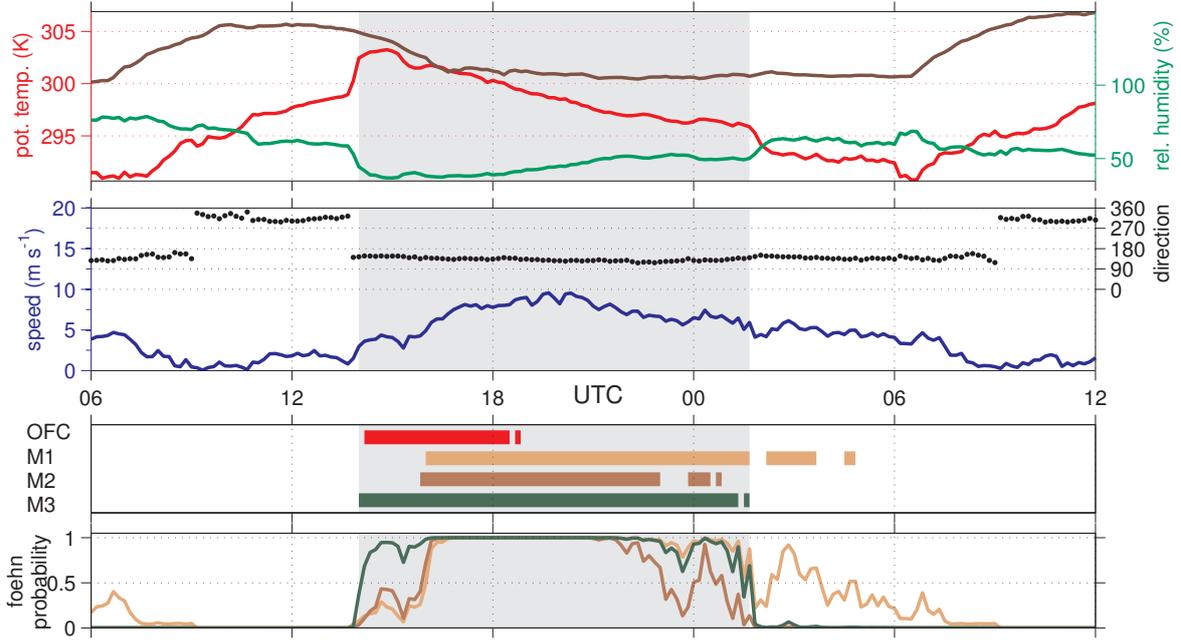


Figure 3: Time series from 06 UTC 27 October 2005 to 12 UTC 28 October 2005 for the valley station with (first row) potential temperature (red line) and relative humidity (green line) and the potential temperature of the crest station (brown line). Second row: Wind direction (black dotted line) and wind speed (blue line) for valley station. Third row: Periods of foehn diagnosed by each model M1–M3 (cf. Table 1) having a foehn probability $\geq 50\%$, and the objective foehn classification (OFC). Fourth row: Foehn probabilities from mixture models M1–M3 with same colors as in third row. Manually classified foehn period is shaded light gray in the background of all rows.

less pronounced signals in temperature and humidity. Notable is the continuous decrease of $\Delta\theta$ due to mixing-in of radiatively cooled air since 17 UTC to -4 K shortly before this foehn event ended. Afterwards, the flow continued to be downvalley and of similar strength but no longer caused by foehn but rather by radiative cooling.

Overall all models captured the core part of this foehn event (compare M1–M3 in rows 3 and 4 of Figure 3). By using only wind speed at the valley station, the mixture model M1 misses the onset by a few hours due to low foehn wind speeds and erroneously postpones the ending caused by high radiatively driven speeds. However, by including the information about the uncertainty of the classification, foehn beginning and ending are indicated correctly in row 4 of Figure 3. Adding the difference in potential temperature between mountain and valley station (M2) avoids the misclassification at the end of the foehn period but shortens the foehn period and still misses the first foehn hours – although again the foehn probability correctly increases. Adding the relative humidity at the valley station as further concomitant variable in M3 finally nearly coincides with our subjective classification.

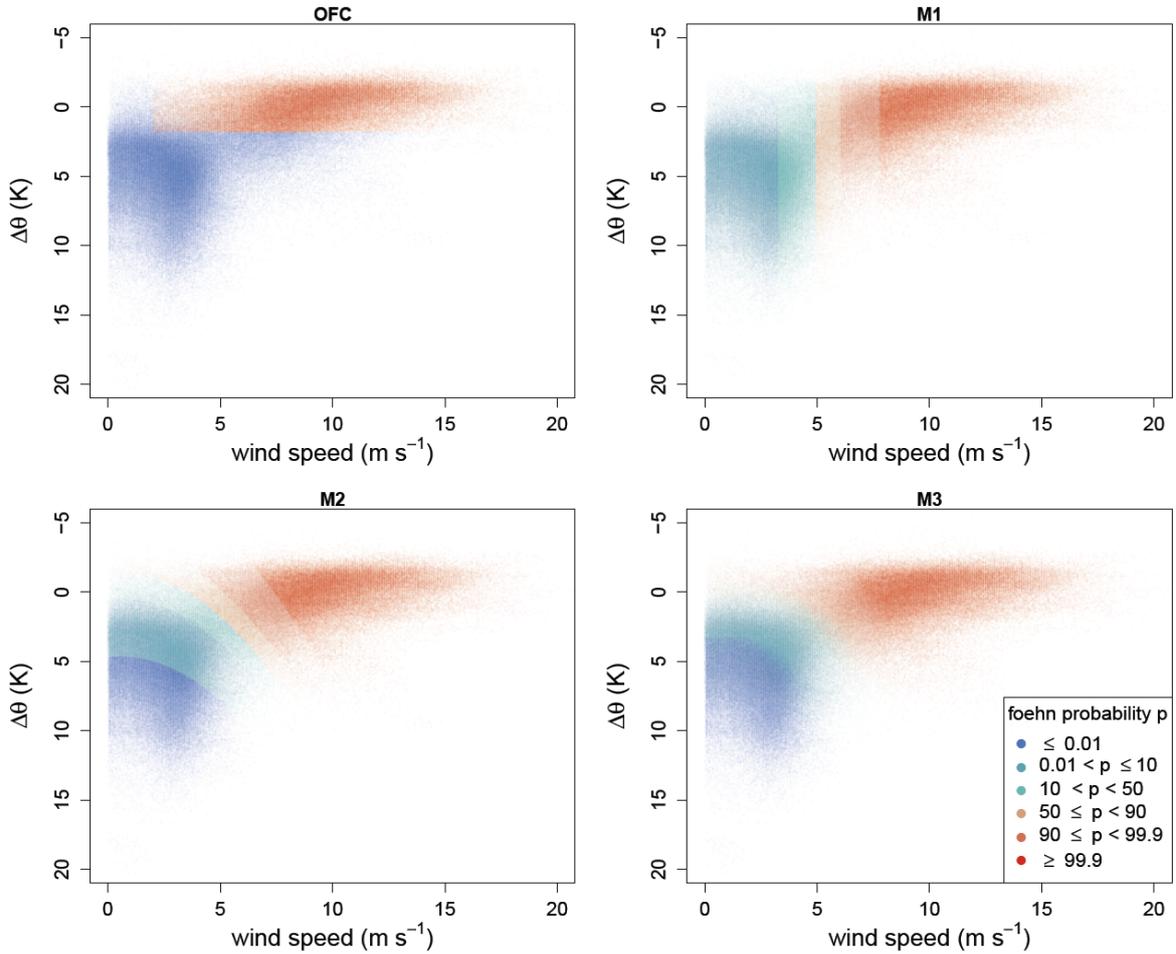


Figure 4: Likelihood of foehn (colors) for wind speed and potential temperature difference ($\Delta\theta = \theta_{crest} - \theta_{valley}$) combinations for wind from the 180° foehn wind sector. Results are shown for the previous objective method OFC (which is deterministic and only has yes/no), and the mixture models M1, M2 and M3 (see also Table 1).

3.2. Climatological and statistical aspects

Climatologically, foehn at the valley station is frequent and occurs for about one fifth of the time. All three mixture models analyze the overall foehn frequency² for the investigated 14-year period to within one percentage point (Table 1): using only wind speed gives 19.6%. Adding potential temperature difference reduces the frequency by one point to 18.6% (as in the case study), while adding relative humidity in M3 brings the frequency back to the speed-only value of 19.7%. The previous objective method, OFC, on the other hand, found only 14.5% foehn.

Figure 4 demonstrates the workings of the mixture models (b)–(d) in comparison to the previous objective method OFC (a). The point clouds have two maxima. One at low wind speeds and large potential temperature differences $\Delta\theta$ (high static stability), which indicates the down valley winds and another with $\Delta\theta$ near zero and high wind speeds, which indicates

²using a threshold of 50% foehn probability

foehn. The OFC (a) uses hard thresholds for $\Delta\theta$ and minimum wind speed and misses events of moderate-to-low speeds and moderate stability. Mixture model M1 with wind speed only (b), on the other hand includes too many of these cases but misses some low speed cases with low stability. If one were to use 50% probability for a yes/no classification, a line of 4.9 m s^{-1} would separate the classes. The second model M2 (Figure 4 (c)) divides the classes more appropriately by a curved cut, because now foehn probability also depends on $\Delta\theta$. This can also be seen in the lower part of Figure 2. At low values of $\Delta\theta$, low wind speeds are sufficient for high foehn probabilities, while for a more stably stratified atmosphere (high $\Delta\theta$) much higher speeds are necessary for reaching the deciding probability of 50%. By adding relative humidity in M3 (d) the separation conforms even more closely to the one a human would draw.

4. Discussion and conclusion

The statistical mixture model is a method to diagnose foehn automatically and probabilistically. It eliminates having to select threshold values individually for each location as required by previous automatic methods (Drechsel and Mayr 2008; Dürr 2008) and includes information how certain the diagnosis is. Another advantage is that it is possible to diagnose foehn objectively without any mountain station (M1) at the expense of some misses (mostly delayed onset times) and false positives (mostly extended duration). For our location, this modified total foehn time by only 0.04% due to about 7% misclassifications on both sides (see Table 1). The previous objective method OFC found about one fifth less foehn (Table 1), mostly by ignoring weak nocturnal foehn cases that have a substantial amount of radiatively cooled air mixed in from the slopes and side valleys thus exceeding the hard threshold for potential temperature difference between crest and valley (Figure 4a).

Prior to the fully automatic classification, one additional piece of information is still required: the selection of the foehn sector. Doing this manually only requires having to determine the valley axis direction and selecting an appropriate sector to each side. In our case this was the widest possible of $\pm 90^\circ$. For particular topographies or when the distinction between different foehn regimes (e.g., south foehn and west foehn) is needed, smaller sectors can be chosen. Alternatively, even this step can be automated using a digital elevation model and a routing algorithm (e.g., Pelletier 2013).

Incidentally, the simplest classification setup proved also to be the best when compared against the subjective human-expert classification of the 50+ events and overall foehn frequency. Other concomitant variables like, gusts, the ratio between gust and 10-minute average speed (‘gust factor’) or the temporal gradients of temperature and relative humidity, respectively, did not improve the results for data from a single (valley) station. Adding relative humidity to M1 without using $\Delta\theta$ (as for M3) even worsened the classification with too many false positives. For the combination of valley and crest station, adding the difference of mixing ratio between mountain and valley station did not improve the results.

Since wind speed distributions are typically non-Gaussian (cf. Wilks 2011) but the mixture model in its simplest form in Equation 1 assumes such a distribution, we remedied a possible violation by first transforming wind speed with logarithm and square root, respectively. However, in all cases considered by us results were worse.

Since weather station records at foehn locations have different length, an important practical

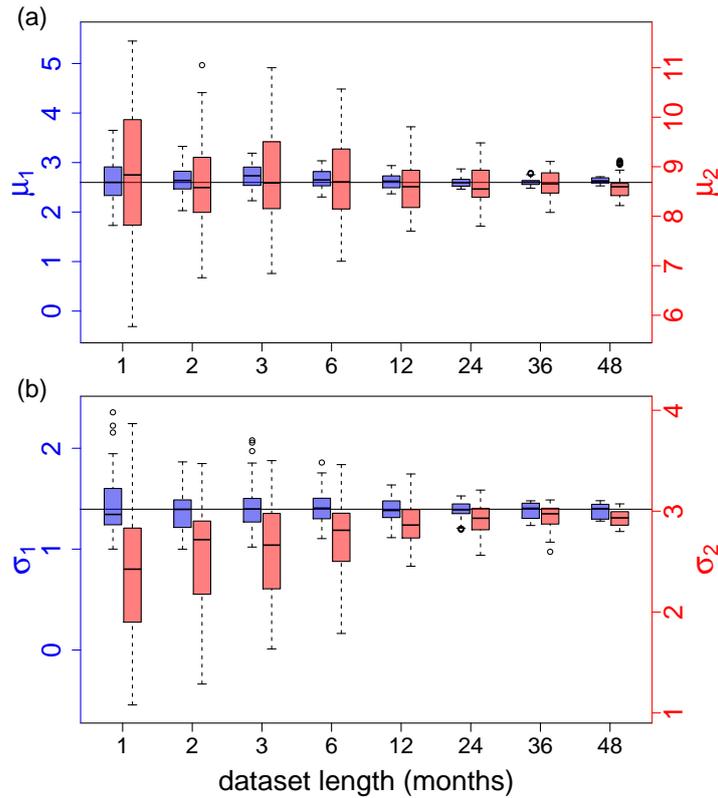


Figure 5: Effect of data length on parameter estimation of mixture model. Means μ_i (a) and standard deviations σ_i (b) were fitted separately for 100 contiguous bootstrap samples of different dataset length (months) for mixture model M3 (M1 and M2 are almost identical). Index 1 is for radiatively driven downslope winds (blue; left axis); index 2 for foehn (red; right axis). Note that scaling for each parameter is equal for both indices (ordinates). The black horizontal lines show the parameter values fitted to the whole available data set (152 months).

question is the required minimum length for the classification to become reliable. Figure 5 shows the effect of record length on four parameters of the mixture model (cf. Equation 1). For each of the record lengths (1 month to 4 years) 100 such contiguous samples were randomly drawn from the 14-year data set and the mixture models fitted only with these measurements. Parameters for no-foehn are already well captured for short periods since radiatively driven flows occur on many nights. Foehn, on the other hand, has a strong seasonal cycle due to, e.g., local cold pools (cf. Mayr *et al.* 2007) with a pronounced minimum during summer in our investigated area. Therefore short records from summer will provide too small a foehn sample for reliable parameter estimation but the same duration taken in autumn might suffice. To cover the major part of foehn wind distributions, a minimum dataset of one year is required.

Having an objective, automatic algorithm to classify foehn using just wind information from only one station means that it will be possible to compare foehn locations all over the world. This permits, e.g., unified foehn climatologies, studies of foehn mechanisms or verification of foehn forecasts. We have already started classification at other locations.

References

- Armi L, Farmer D (2002). “Stratified Flow over Topography: Bifurcation Fronts and Transition to the Uncontrolled State.” *Proceedings of the Royal Society A*, **458**, 513–538.
- Armi L, Mayr GJ (2011). “The Descending Stratified Flow and Internal Hydraulic Jump in the Lee of the Sierras.” *Journal of Applied Meteorology and Climatology*, **50**, 1995–2011.
- Conrad V (1936). *Handbuch der Klimatologie*, chapter Die klimatologischen Elemente und ihre Abhängigkeit von terrestrischen Einflüssen, pp. 1146–1166. Bornträger.
- Dayton CM, Macready GB (1986). “Concomitant-Variable Latent-Class Models.” *Journal of the American Statistical Association*, **83**, 173–178.
- Defant A (1949). “Zur Theorie der Hangwinde, nebst Bemerkungen zur Theorie der Berg- und Talwinde.” *Archiv für Meteorologie Geophysik und Bioklimatologie A*, **1**, 421–450.
- Drechsel S, Mayr GJ (2008). “Objective Forecasting of Foehn Winds for a Subgrid-Scale Alpine Valley.” *Weather and Forecasting*, **23**, 205–218.
- Dürr B (2008). “Automatisiertes Verfahren zur Bestimmung von Föhn in Alpentälern.” *Arbeitsbericht 223*, MeteoSchweiz. URL http://www.meteoschweiz.admin.ch/web/de/forschung/publikationen/alle_publikationen/ab_223.Par.0001.DownloadFile.tmp/ab223.pdf.
- Durrán DR (1990). *Mountain Waves and Downslope Winds*, volume 23 of *Meteorological Monographs*, chapter 4, pp. 59–81. American Meteorological Society.
- Grün B, Leisch F (2008). “FlexMix Version 2: Finite Mixtures with Concomitant Variables and Varying and Constant Parameters.” *Journal of Statistical Software*, **28**(4), 1–35. URL <http://www.jstatsoft.org/v28/i04/>.
- Hastie T, Tibshirani R, Friedman J (2009). *The Elements of Statistical Learning*. 2nd edition. Springer-Verlag, New York.
- Jackson PL, Mayr GJ, Vosper S (2013). *Mountain Weather Research and Forecasting: Recent Progress and Current Challenges*, chapter Dynamically-Driven Winds, pp. 121–218. Springer-Verlag.
- Leisch F (2004). “FlexMix: A General Framework for Finite Mixture Models and Latent Class Regression in R.” *Journal of Statistical Software*, **11**(8), 1–18. URL <http://www.jstatsoft.org/v11/i08/>.
- Mayr GJ, Armi L (2008). “Föhn as a Response to Changing Upstream and Downstream Air Masses.” *Quarterly Journal of the Royal Meteorological Society*, **134**, 1357–1369.
- Mayr GJ, Armi L (2010). “The Influence of Downstream Diurnal Heating on the Descent of Flow across the Sierras.” *Journal of Applied Meteorology and Climatology*, **49**, 1906–1912.

- Mayr GJ, Armi L, Banta S, Arnold RM, Darby LS, Durran DD, Flamant C, Gabersek S, Gohm A, Mayr R, Mobbs S, Nance LB, Vergeiner I, Vergeiner J, Whiteman CD (2004). “Gap Flow Measurements during the Mesoscale Alpine Programme.” *Meteorology and Atmospheric Physics*, **86**, 99–119.
- Mayr GJ, Armi L, Gohm A, Zängl G, Durran DR, Flamant C, Gaberšek S, Mobbs S, Ross A, Weissmann M (2007). “Gap Flows: Results from the Mesoscale Alpine Programme.” *Quarterly Journal of the Royal Meteorological Society*, **133**, 881–896.
- McLachlan GJ, Peel D (2000). *Finite Mixture Models*. John Wiley & Sons.
- Pelletier JD (2013). “A Robust, Two-Parameter Method for the Extraction of Drainage Networks from High-Resolution Digital Elevation Models (DEMs): Evaluation Using Synthetic and Real-World DEMs.” *Water Resources Research*, **49**, 75–89.
- Prandtl L (1944). *Führer durch die Strömungslehre*, chapter 5, pp. 326–330. 2nd edition. Friedrich Vieweg & Sohn.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Vergeiner J (2004). *South Foehn Studies and a New Foehn Classification Scheme in the Wipp and Inn Valley*. Ph.D. thesis, Universität Innsbruck. URL http://imgi.uibk.ac.at/sekretariat/dissertations/Vergeiner_Johannes_2004_Diss.pdf.
- Whiteman CD (2000). *Mountain Meteorology: Fundamentals and Applications*. Oxford University Press.
- Wilks DS (2011). *Statistical Methods in the Atmospheric Sciences*, volume 100 of *International Geophysics Series*. 3rd edition. Academic Press, Oxford.
- WMO (1992). *International Meteorological Vocabulary*, volume 182. 2nd edition. World Meteorological Organization.

Affiliation:

David Plavcan, Georg J. Mayr
Institute of Meteorology and Geophysics
Universität Innsbruck
Innrain 52
6020 Innsbruck, Austria
E-mail: David.Plavcan@student.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@R-project.org
URL: <http://eeecon.uibk.ac.at/~zeileis/>

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

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

- 2013-22 **David Plavcan, Georg J. Mayr, Achim Zeileis:** Automatic and probabilistic foehn diagnosis with a statistical mixture model
- 2013-21 **Jakob W. Messner, Georg J. Mayr, Achim Zeileis, Daniel S. Wilks:** Extending extended logistic regression to effectively utilize the ensemble spread
- 2013-20 **Michael Greinecker, Konrad Podczeck:** Liapounoff's vector measure theorem in Banach spaces
- 2013-19 **Florian Lindner:** Decision time and steps of reasoning in a competitive market entry game
- 2013-18 **Michael Greinecker, Konrad Podczeck:** Purification and independence
- 2013-17 **Loukas Balafoutas, Rudolf Kerschbamer, Martin Kocher, Matthias Sutter:** Revealed distributional preferences: Individuals vs. teams
- 2013-16 **Simone Gobien, Björn Vollan:** Playing with the social network: Social cohesion in resettled and non-resettled communities in Cambodia
- 2013-15 **Björn Vollan, Sebastian Prediger, Markus Frölich:** Co-managing common pool resources: Do formal rules have to be adapted to traditional ecological norms?
- 2013-14 **Björn Vollan, Yexin Zhou, Andreas Landmann, Biliang Hu, Carsten Herrmann-Pillath:** Cooperation under democracy and authoritarian norms
- 2013-13 **Florian Lindner, Matthias Sutter:** Level-k reasoning and time pressure in the 11-20 money request game *forthcoming in Economics Letters*
- 2013-12 **Nadja Klein, Thomas Kneib, Stefan Lang:** Bayesian generalized additive models for location, scale and shape for zero-inflated and overdispersed count data
- 2013-11 **Thomas Stöckl:** Price efficiency and trading behavior in limit order markets with competing insiders *forthcoming in Experimental Economics*
- 2013-10 **Sebastian Prediger, Björn Vollan, Benedikt Herrmann:** Resource scarcity, spite and cooperation
- 2013-09 **Andreas Exenberger, Simon Hartmann:** How does institutional change coincide with changes in the quality of life? An exemplary case study

- 2013-08 **E. Glenn Dutcher, Loukas Balafoutas, Florian Lindner, Dmitry Ryvkin, Matthias Sutter:** Strive to be first or avoid being last: An experiment on relative performance incentives.
- 2013-07 **Daniela Glätzle-Rützler, Matthias Sutter, Achim Zeileis:** No myopic loss aversion in adolescents? An experimental note
- 2013-06 **Conrad Kobel, Engelbert Theurl:** Hospital specialisation within a DRG-Framework: The Austrian case
- 2013-05 **Martin Halla, Mario Lackner, Johann Scharler:** Does the welfare state destroy the family? Evidence from OECD member countries
- 2013-04 **Thomas Stöckl, Jürgen Huber, Michael Kirchler, Florian Lindner:** Hot hand belief and gambler's fallacy in teams: Evidence from investment experiments
- 2013-03 **Wolfgang Luhan, Johann Scharler:** Monetary policy, inflation illusion and the Taylor principle: An experimental study
- 2013-02 **Esther Blanco, Maria Claudia Lopez, James M. Walker:** Tensions between the resource damage and the private benefits of appropriation in the commons
- 2013-01 **Jakob W. Messner, Achim Zeileis, Jochen Broecker, Georg J. Mayr:** Improved probabilistic wind power forecasts with an inverse power curve transformation and censored regression
- 2012-27 **Achim Zeileis, Nikolaus Umlauf, Friedrich Leisch:** Flexible generation of e-learning exams in R: Moodle quizzes, OLAT assessments, and beyond
- 2012-26 **Francisco Campos-Ortiz, Louis Putterman, T.K. Ahn, Loukas Balafoutas, Mongoljin Batsaikhan, Matthias Sutter:** Security of property as a public good: Institutions, socio-political environment and experimental behavior in five countries
- 2012-25 **Esther Blanco, Maria Claudia Lopez, James M. Walker:** Appropriation in the commons: variations in the opportunity costs of conservation
- 2012-24 **Edgar C. Merkle, Jinyan Fan, Achim Zeileis:** Testing for measurement invariance with respect to an ordinal variable *forthcoming in Psychometrika*
- 2012-23 **Lukas Schrott, Martin Gächter, Engelbert Theurl:** Regional development in advanced countries: A within-country application of the Human Development Index for Austria
- 2012-22 **Glenn Dutcher, Krista Jabs Saral:** Does team telecommuting affect productivity? An experiment

- 2012-21 **Thomas Windberger, Jesus Crespo Cuaresma, Janette Walde:** Dirty floating and monetary independence in Central and Eastern Europe - The role of structural breaks
- 2012-20 **Martin Wagner, Achim Zeileis:** Heterogeneity of regional growth in the European Union
- 2012-19 **Natalia Montinari, Antonio Nicolo, Regine Oexl:** Mediocrity and induced reciprocity
- 2012-18 **Esther Blanco, Javier Lozano:** Evolutionary success and failure of wildlife conservancy programs
- 2012-17 **Ronald Peeters, Marc Vorsatz, Markus Walzl:** Beliefs and truth-telling: A laboratory experiment
- 2012-16 **Alexander Sebald, Markus Walzl:** Optimal contracts based on subjective evaluations and reciprocity
- 2012-15 **Alexander Sebald, Markus Walzl:** Subjective performance evaluations and reciprocity in principal-agent relations
- 2012-14 **Elisabeth Christen:** Time zones matter: The impact of distance and time zones on services trade
- 2012-13 **Elisabeth Christen, Joseph Francois, Bernard Hoekman:** CGE modeling of market access in services
- 2012-12 **Loukas Balafoutas, Nikos Nikiforakis:** Norm enforcement in the city: A natural field experiment *forthcoming in European Economic Review*
- 2012-11 **Dominik Erharder:** Credence goods markets, distributional preferences and the role of institutions
- 2012-10 **Nikolaus Umlauf, Daniel Adler, Thomas Kneib, Stefan Lang, Achim Zeileis:** Structured additive regression models: An R interface to BayesX
- 2012-09 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** History repeating: Spain beats Germany in the EURO 2012 Final
- 2012-08 **Loukas Balafoutas, Glenn Dutcher, Florian Lindner, Dmitry Ryvkin:** The optimal allocation of prizes in tournaments of heterogeneous agents
- 2012-07 **Stefan Lang, Nikolaus Umlauf, Peter Wechselberger, Kenneth Harttgen, Thomas Kneib:** Multilevel structured additive regression
- 2012-06 **Elisabeth Waldmann, Thomas Kneib, Yu Ryan Yu, Stefan Lang:** Bayesian semiparametric additive quantile regression

- 2012-05 **Eric Mayer, Sebastian Rueth, Johann Scharler:** Government debt, inflation dynamics and the transmission of fiscal policy shocks *forthcoming in Economic Modelling*
- 2012-04 **Markus Leibrecht, Johann Scharler:** Government size and business cycle volatility; How important are credit constraints? *forthcoming in Economica*
- 2012-03 **Uwe Dulleck, David Johnston, Rudolf Kerschbamer, Matthias Sutter:** The good, the bad and the naive: Do fair prices signal good types or do they induce good behaviour?
- 2012-02 **Martin G. Kocher, Wolfgang J. Luhan, Matthias Sutter:** Testing a forgotten aspect of Akerlof's gift exchange hypothesis: Relational contracts with individual and uniform wages
- 2012-01 **Loukas Balafoutas, Florian Lindner, Matthias Sutter:** Sabotage in tournaments: Evidence from a natural experiment *published in Kyklos*

University of Innsbruck

Working Papers in Economics and Statistics

2013-22

David Plavcan, Georg J. Mayr, Achim Zeileis

Automatic and probabilistic foehn diagnosis with a statistical mixture model

Abstract

Diagnosing foehn winds from weather station data downwind of topographic obstacles requires distinguishing them from other downslope winds, particularly nocturnal ones driven by radiative cooling. We present an automatic classification scheme to obtain reproducible results that include information about the (un)certainly of the diagnosis. A statistical mixture model separates foehn and no-foehn winds in a measured time series of wind. In addition to wind speed and direction, it accommodates other physically meaningful classifiers such as relative humidity or the (potential) temperature difference to an upwind station (e.g., near the crest). The algorithm was tested for the central Alpine Wipp Valley against human expert classification and a previous objective method (Drechsel and Mayr 2008), which the new method outperforms. Climatologically, using only wind information gives nearly identical foehn frequencies as when using additional covariables, making the method suitable for comparable foehn climatologies all over the world where station data are available for at least one year.

ISSN 1993-4378 (Print)

ISSN 1993-6885 (Online)