Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

Veröffentlichung MeteoSchweiz Nr. 89

COSMO-2 Model Performance in Forecasting Foehn: a Systematic Process-oriented Verification

Micah Wilhelm

Herausgeber
Bundesamt für Meteorologie und Klimatologie, MeteoSchweiz, © 2012

MeteoSchweiz
Krähbühlstrasse 58
CH-8044 Zürich
T +41 44 256 91 11
www.meteoschweiz.ch

Weitere Standorte
CH-8058 Zürich-Flughafen
CH-6605 Locarno Monti
CH-1211 Genève 2
CH-1530 Payerne

Veröffentlichung MeteoSchweiz Nr. 89

ISSN: 1422-1381

COSMO-2 Model Performance in Forecasting Foehn: a Systematic Process-oriented Verification

Micah Wilhelm

Master Thesis

Supervisor: Dr. Matteo Buzzi, MeteoSwiss

Co-supervisor: Dr. Michael Sprenger, IAC ETH Zürich

Bitte zitieren Sie diese Veröffentlichung folgendermassen
Wilhelm, M: 2012, COSMO-2 Model Performance in Forecasting Foehn: a Systematic Process-oriented Verification, *Veröffentlichungen der MeteoSchweiz*, **89**, 64pp.

Herausgeber
Bundesamt für Meteorologie und Klimatologie, MeteoSchweiz, © 2012

MeteoSchweiz
Krähbühlstrasse 58
CH-8044 Zürich
T +41 44 256 91 11
www.meteoschweiz.ch

Weitere Standorte
CH-8058 Zürich-Flughafen
CH-6605 Locarno Monti
CH-1211 Genève 2
CH-1530 Payerne

COSMO-2 Model Performance in Forecasting Foehn: a Systematic Process-oriented Verification

Master Thesis

Author:
Micah Wilhelm

Supervisor:
Dr. Matteo Buzzi

Co-supervisor:
Dr. Michael Sprenger

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

IAC ETH
Institute for Atmospheric and Climate Science
Swiss Federal Institute of Technology Zurich

March 26, 2012

Abstract

Strong foehn episodes can develop into violent wind storms capable of causing severe material damages. Accurate forecasts of foehn strength and duration are essential for preventative and protective measures against losses to be effectively taken. High-resolution numerical prediction models, such as COSMO-2, show several deficiencies when forecasting both south foehn episodes in the Swiss Alps. Therefore, a series of systematic verifications was done to objectively assess the overall south foehn detection skill, biases and accuracies of individual variables, as well as the temporal lag of south foehn in COSMO-2 reanalysis data. The analysis, done over the surface and 3 lowest model levels, showed that the overall foehn frequency is estimated best at Alpine stations and is overestimated in fore-Alpine and Swiss plateau stations. The foehn detection skill was found to decrease the slowest with foehn intensity when using the second model level. This skill was optimized using a scheme combining the most accurate variables and model levels for high and low foehn intensities. However, even when all foehn hours are considered, this method shows less skill than the boosting method by Oechslin (2008). Overall model errors in wind speed and air temperature during south foehn are approximately 2.5 times greater in magnitude than the climatological errors over Switzerland. The presence of a linear relationship between the relative humidity bias and altitude error rather than with distance from the Main Alpine Ridge suggests that topographical smoothing influences relative humidity to a greater extent than inaccuracies in the PBL parameterization scheme. Temporal lag errors of foehn onset and decay were found to have a wide range of positive and negative values alike. The onset lag tends to become more negative with model level height. Except for the surface, which shows a negative mean lag, higher model levels show positive mean decay lags increasing with model height.

Contents

1 Introduction **1**
1.1 Warming mechanisms of foehn . . . 1
1.2 Foehn flow . . . 3
1.3 Foehn forecasting . . . 4
1.4 COSMO-2 model description . . . 5
1.5 Foehn forecasting deficiencies in COSMO-2 . . . 6
1.6 Motivation of research . . . 7

2 Data and Methods **8**
2.1 Verification dataset . . . 8
2.2 COSMO-2 model dataset . . . 10
2.3 Foehn detection . . . 12
2.4 Foehn detection applied to COSMO-2 . . . 14
2.5 Systematic verifications of foehn in COSMO-2 . . . 16
2.5.1 Categorical verification . . . 16
2.5.2 Classical verification . . . 17
2.5.3 Lag time of foehn onset and decay . . . 18

3 Results **20**
3.1 Overview of results . . . 20
3.2 COSMO-2 surface level . . . 24
3.2.1 Categorical verification . . . 24
3.2.2 Classical verification . . . 26
3.2.3 Lag of foehn onset and decay . . . 28
3.3 Lowest 3 COSMO-2 levels . . . 30
3.3.1 Categorical verification . . . 30
3.3.2 Classical verification . . . 32
3.3.3 Lag of foehn onset and decay . . . 34
3.4 Skill optimization . . . 36

4 Discussion **39**
4.1 General COSMO-2 foehn detection skill . . . 39
4.1.1 Foehn frequency . . . 39
4.1.2 Foehn detection skill and foehn intensity . . . 39
4.2 Performance of COSMO-2 during observed foehn . . . 40
4.2.1 Bias trends . . . 40
4.2.2 MAE and correlation . . . 42
4.3 Comparison to seasonal COSMO-2 performance . . . 43
4.4 Foehn onset and decay lag . . . 44
4.5 Optimized foehn scheme . . . 45

5 Summary and Outlook **46**

1 Introduction

The World Meteorological Institution (1992) defines foehn as "wind [which is] warmed and dried by descent, in general on the leeward side of a mountain." Air parcels are forced over a mountain range by a strong, synoptic-scale pressure gradient oriented quasi-perpendicularly to that range. In central Europe, cross-Alpine pressure gradients often occur when a low-pressure system propagates from the north-west towards the Alps while there is a high-pressure system to the south. This situation is called south foehn since the air parcels are advected from the south (figure 1). The opposite synoptic situation and flow direction occurs in the case of north foehn. For the remainder of this thesis, the term "foehn" shall refer to the phenomenon of south foehn.

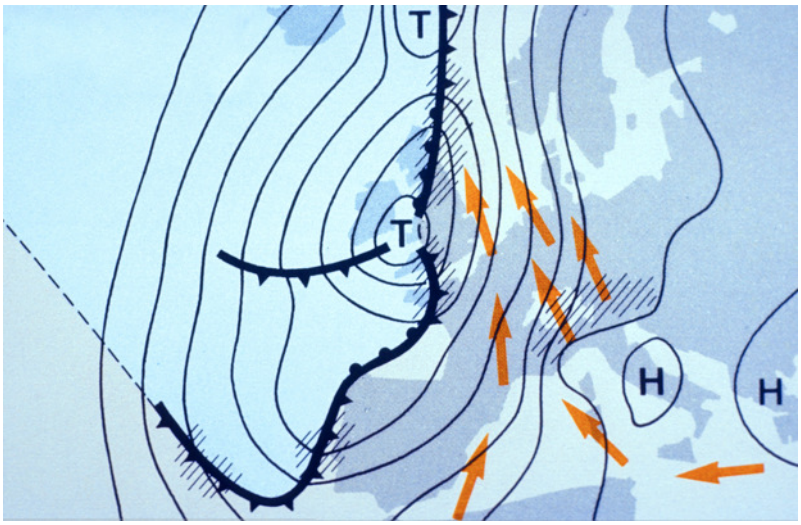

Figure 1: Surface weather analysis chart of central Europe depicting typical synoptic-scale conditions during south foehn episodes over the Alps. Solid lines are lines of constant surface pressure (isobars). Red arrows indicate approximate surface wind direction. T indicates the centre of a low-pressure system. H indicates the centre of a high-pressure system. Warm and cold fronts are shown with thick dark blue lines. Adapted from Richner and Haechler (2008).

1.1 Warming mechanisms of foehn

Two warming mechanisms of foehn can be distinguished. Both can occur simultaneously with varying amount of contribution. The degree of the contributions are a function of the static stability and the water content of the advected air parcels. In addition, local topography of the mountain range can favour one mechanism over the other.

- In the first case, air parcels are forced to ascend as they are advected up the windward slope of a mountain. As the air parcels are lifted, they expand and cool at the dry adiabatic lapse rate (-9.8°C/km altitude). Water vapour condenses once the dew-point temperature is reached at the lifting condensation level (LCL). As a result, latent heat is released (heating by condensation), and

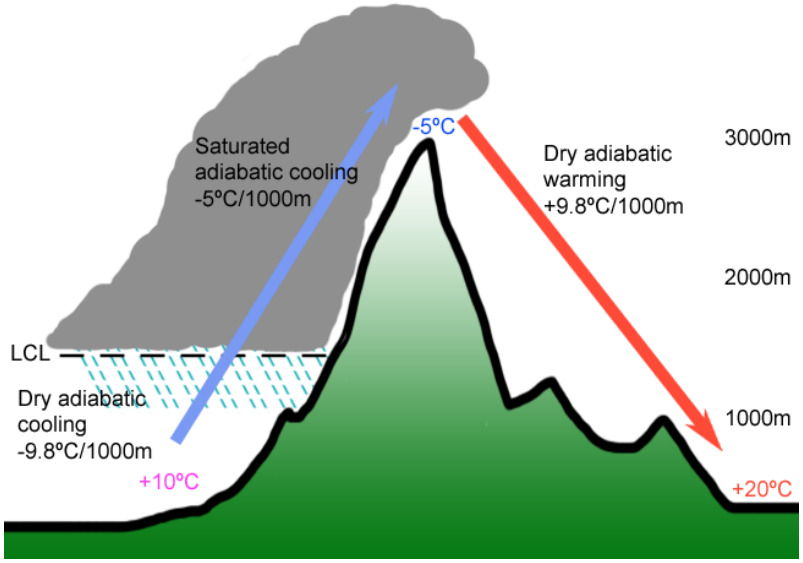

Figure 2: Cross-section of a latent heating mechanism of foehn as a result of synoptic-scale flow over a mountain ridge. The blue arrow shows ascent of the air parcels on the windward side. The red arrow shows descent of the air parcels on the leeward side. The black dashed line shows the lifting condensation level (LCL) where the temperatures of the advected air parcels reach the dew point temperature. Thus, the air becomes saturated with respect to water vapour and forms stratocumulus clouds. Adapted from Steinacker (2006).

allows air parcels to cool slower at the moist adiabatic lapse rate (-5°C/km altitude). Consequently, precipitation occurs on the windward side and lowers the water content of the advected air parcels (Ficker and De Rudder, 1943). As the air parcels then descend down the leeward slope and warms by adiabatic compression. Relative humidity is further decreased by this warming. This results in warmer air temperature and lower relative humidity on the leeward side of the mountain range (figure 2). This so-called "waterfall" foehn warming mechanism is experienced mainly in the Swiss Alps, where the mountain passes are quite high (Steinacker, 2006).

- In the second case the ascent of air parcels on the windward side of a mountain ridge is inhibited by strong static stability. Consequently, the low-level flow is blocked from advection by the mountain range (Seibert, 1990). Air parcels near the level of the inversion layer descend down the leeward slope and warm due to adiabatic compression (figure 3). A significant increase in temperature and decrease in relative humidity is observed along the leeward slope of the mountain range (Hann, 1867). This is because air near the level of the alpine crest generally has a higher potential temperature and a lower relative humidity than the air below. This second warming mechanism can theoretically contribute a larger increase in air temperature than the previously described mechanism. Moreover, this mechanism occurs more often in

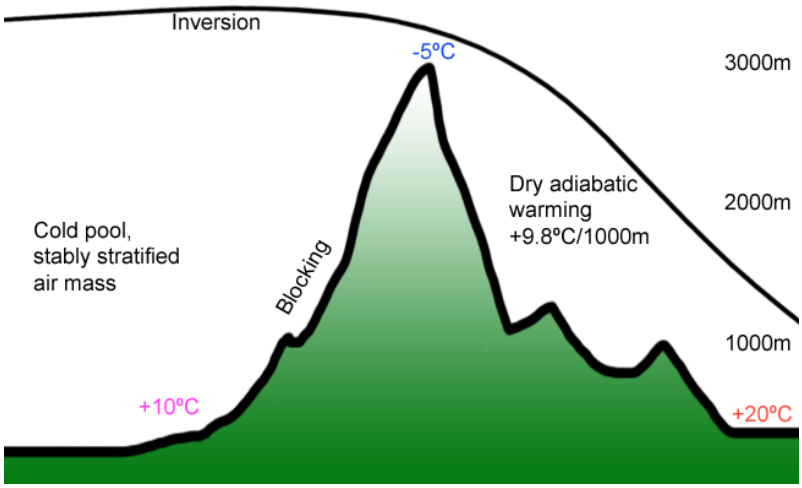

Figure 3: Cross-section of upstream blocking mechanism of foehn as a result of synoptic-scale flow over a mountain ridge in the presence of an inversion. Solid line shows the general air path trajectory along a line of constant potential temperature (isentrope). The areas of orographic blocking of the windward cold air pool and subsequent adiabatic descent of air parcels in the lee of the mountain range are labelled. Adapted from Steinacker (2006).

the Austrian Alps, where the mountain passes are lower and precipitation or clouds are rarely observed upstream of the mountain ridge (Hann, 1866).

1.2 Foehn flow

Foehn episodes are marked by strong wind speeds and rather gusty conditions in the Alpine valleys. The acceleration of the air can be explained by hydraulic theory and mountain induced gravity waves as described by Durran (1990). In addition, foehn air is channelled through gaps in the ridge and valleys within the mountain range, such as the Brenner Pass (Wipp Valley), the Gotthard Pass (Reuss Valley), the San Bernardino Pass (Rhine Valley) and the Simplon Pass (Rhone Valley) (Seibert, 1990). This flow restriction by the valley transect induces a further increase in wind speed and a decrease in pressure following the Bernoulli principle. Thus, a supercritical flow state is achieved by the foehn air as it descends in the lee of the mountain range. The foehn flow breaks away from the slope surface once it achieves the level of neutral buoyancy at the top of the "cold pool" (a cold air mass north of the mountain range). As the wind speed decreases, the foehn transitions to a sub-critical flow regime. This is marked by very turbulent conditions and a decrease in wind speed (hydraulic jump). It is here that turbulent mixing between cold pool and foehn air occurs. Depending on the strength of the foehn wind and on the stability of the cold pool, this can lead to the gradual erosion of the cold pool in the valley bottoms. This process, called "foehn breakthrough", allows the foehn air to

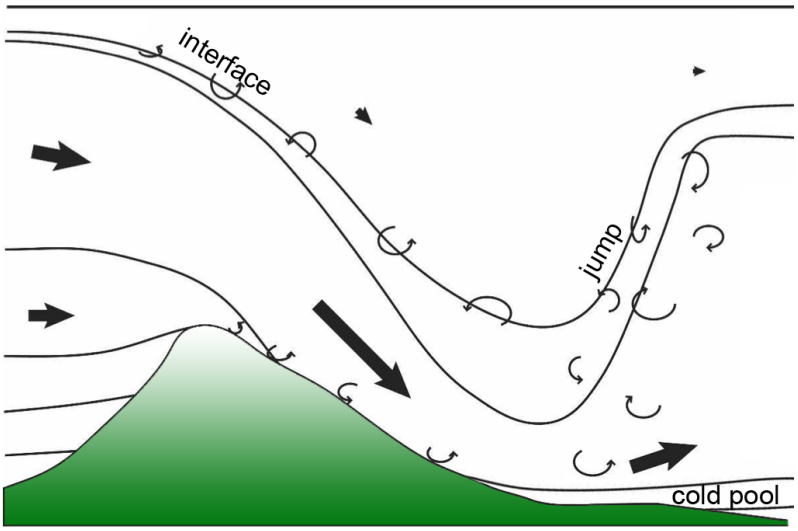

Figure 4: Cross-section of a foehn episode depicting typical flow features. Solid lines are lines of constant potential temperature (isentropes). Curved arrows indicate turbulent mixing. Solid arrows show the qualitative wind speeds. The areas of stable air-interface, upstream blocking, hydraulic jump, and cold air pool are labelled. Adapted from Drechsel and Mayr (2008).

penetrate to the valley bottom and allows the foehn front to extend further north. In areas where foehn breakthrough is not achieved, the foehn air mass acts as a capping inversion that traps the air below (Drechsel and Mayr, 2008).

1.3 Foehn forecasting

Deterministic foehn forecasting was first developed by Widmer (1966) to predict South foehn in Altdorf. Based on statistical analysis of observational data, the Widmer index was designed to forecast foehn winds as far as two days in advance. Courvoisier and Gutermann (1971) later simplified the index and applied it to numerical weather prediction (NWP) model forecast data. The newer index relies on the cross-alpine forecasted pressure gradients at 500 and 850 hPa geopotential height and has been in regular use in MeteoSwiss since the 1970's. In a similar fashion, other foehn forecast tools have been realized by applying foehn detection algorithms, such as those outlined in the previous section, to the of NWP model forecast data. Drechsel and Mayr (2008) have shown that objective, categorical foehn forecast skill is high for lead times as long as three days. Using this method and the global forecast model T511 from the European Centre for Medium-Range Weather Forecasts (ECMWF), the predictability for a γ-mesoscale ($2 - 20$ km horizontal scale) phenomenon such as foehn is high despite a 40 km spatial resolution. Higher resolution NWP models are, however, required to simulate complex topography of alpine valleys. This would enable more accurate predictions of specific foehn flow dynamics and properties.

1.4 COSMO-2 model description

Since February 2008, MeteoSwiss has operated a high-resolution model from the Consortium for Small-Scale Modelling (COSMO). COSMO is based on the primitive thermo-hydrodynamic equations that describe non-hydrostatic (vertical movement), compressible flow in a moist atmosphere. The model utilizes a rotated geographical Arakawa C-grid and generalized terrain-following height coordinates with a Lorenz vertical staggering scheme. Prognostic variables include: vertical and horizontal wind components, pressure perturbation, potential temperature, specific humidity, cloud water and ice content, rain, snow, and turbulent kinetic energy. Diagnostic variables include: total air density and fluxes of rain and snow. COSMO applies the following to all runs: a third-order spatial discretization and 2 time-level time integration based on Runge-Kutta split-explicit scheme and Rayleigh damping in the upper layers (Doms and Schaettler, 2002). COSMO is also coupled to a multi-layer soil model utilizing 8 and 6 layers for energy and moisture fluxes, respectively. Moisture values in the lowest levels are updated every 24 hours from Integrated Forecast System (IFS).

COSMO is run at two spatial resolutions at MeteoSwiss: 2.2 km (COSMO-2) and 6.6 km (COSMO-7). Several differences between the parameterizations of COSMO-2 and COSMO-7 exist. In COSMO-7, convection is parameterized while deep convection is explicitly computed in COSMO-2. Moreover, topographic effects on incident radiation are considered in COSMO-2 but not in COSMO-7. Finally, COSMO-2 has a micro-physical parameterization that uses a prognostic graupel hydrometeor class. COSMO contains a Newtonian relaxation algorithm that facilitates observational data assimilation. In essence, the simulated atmospheric fields are forced towards the observations at the time of the observation by a nudging function. The following in-situ data are assimilated (Schraff and Hess, 2011):

- Surface synoptic stations, ship and buoy measurements
  - Surface pressure
  - 2 m humidity for the lowest model level
  - 10 m wind for stations below 100 m above mean sea level
- Temperature/pilot data
  - Wind
  - Temperature
  - Humidity profiles
- Aircraft measurements
  - Wind
  - Temperature
- Wind profiler data
- Snow analysis, derived from Meteosat Second Generation (MSG) satellites combined with dense observations

Additionally, COSMO-2 assimilates Radar-derived rainfall data via latent heat nudging. This scheme utilizes high spatial and temporal resolution of rainfall observations. Rainfall data is obtained from three C-Band Radars in the Swiss Radar Network. The assimilation cycle is run in 3-hour intervals with a 45-minute observation

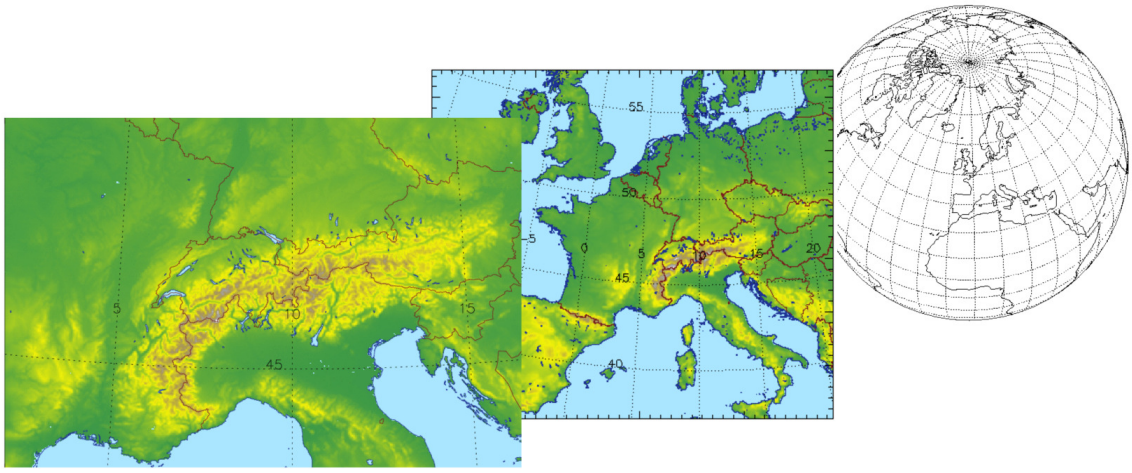

Figure 5: The three nested NWP model domains of the COSMO system as operated by MeteoSwiss. Left: COSMO-2 model describing the local scale. Middle: COSMO-7 model describing the regional scale in Europe. Right: ECMWF operates a global NWP model describing the synoptic scales. Adapted from MeteoSwiss (2012).

cut-off time. COSMO-2 is run with three nested domains with increasing horizontal resolution. The global Integrated Forecast System (IFS) model from ECMWF drives lateral boundary conditions for COSMO-7, which extends over Western Europe. COSMO-7 then provides lateral boundary conditions for COSMO-2, which covers the Alpine arch (figure 5). This configuration allows COSMO to provide short-range forecasts at mesoscale resolutions in relatively short computational times. Forecasts, which begin every 3 hours, are integrated up to 24 hours in a rapid update cycle (Schraff and Hess, 2011).

1.5 Foehn forecasting deficiencies in COSMO-2

Although upper-level wind fields are represented quite well in COSMO-2, there are several deficiencies in the simulation of foehn events within the Planetary Boundary Layer (PBL). This can often lead to forecasting problems for the duration and/or intensity both south and north foehn episodes. A foehn case study by Burri et al. (2007) indicated for example that foehn mean/gust wind speeds in COSMO-2 are underestimated despite generally accurate pressure gradients. Additionally, a negative temperature bias has been detected during foehn episodes, which is likely a result of too strong coupling of the surface air and soil temperatures in the model parameterization (Duerr et al., 2010). Lastly, foehn breakthrough, termination and geographical extension are inaccurately reproduced due to poor representation of cold pool erosion. An objective, quasi-climatological systematic assessment of these forecast errors is the basic step in order to improve the forecast skill of foehn in the COSMO-2 model.

1.6 Motivation of research

Foehn is a dynamic down-slope wind phenomenon and has captivated the interest of scientists for over a century (Hann, 1866; Wild, 1868; Schweitzer, 1952). Foehn winds occur not only as semi-2D flows over mountain ridges but also are channelled into narrow gaps through mountain passes. Consequently, the flow characteristics of foehn are highly dependent on local topography and are thus unique to each Alpine valley. This complex 3-D gap flow behaviour makes foehn a challenge to predict (Drobinski et al., 2007). Local air quality, surface air temperature and humidity are strongly influenced by the presence of foehn. Strong foehn events can produce increases in surface temperature as great as approximately 20°C. Warming of this magnitude can melt snow quickly, destabilizing snow pack stability and temporarily increasing the danger of avalanches. Frequent foehn events can lead to shallow snow cover and can have negative economic repercussions to local ski resorts. With increased temperatures and low relative humidity, the risk of forest fires also increases. In addition, strong foehn episodes can develop into violent wind storms capable of causing severe material damages (Brinkman, 1974). Therefore, accurate forecasts of foehn strength, duration and spatial extent are essential for preventative and protective measures to be effectively taken if possible.

2 Data and Methods

2.1 Verification dataset

Observational data from February 24$^{\text{th}}$, 2008 to December 31$^{\text{st}}$, 2011 were downloaded from the CLIMAP 8.2 database of MeteoSwiss. The data were collected by a national network of automated surface weather stations (SwissMetNet) that are distributed across Switzerland (figure 6). Of the 20 surface weather stations analysed (figure 7), 18 were those for which foehn thresholds were established by Duerr (2008) including the MAR reference station Gütsch (GUE). Several of these stations record hourly measurements of atmospheric variables. However, the majority of stations record measurements every 10 minutes. Since COSMO-2 is unable to reproduce atmospheric dynamics on a 10 minute time scale, hourly averages of the following foehn relevant variables were extracted:

- air temperature at 2 m above the surface (T)
- relative humidity at 2 m above the surface (RH)
- scalar wind speed at 10 m above the surface (V)
- wind direction at 10 m above the surface (V_{dir})
- surface pressure at the station altitude (p)
- surface pressure reduced to sea level (p_{red})
- gust speed (hourly V maximum) at 10 m above the surface (V_{max})

The hourly averaged potential temperature at 2 m above the surface was calculated with the following equation:

$$\theta = T(\tfrac{p_0}{p})^{\frac{R_v}{c_p}},$$

where p_0 is a reference pressure (1000 hPa), R_v is the specific gas constant of water vapour (461.51 J/kgK) and c_p is the specific heat capacity of dry air (1003.5 J/kgK). The remaining two stations, Zürich-Kloten and Magadino-Cadenazzo, were chosen to calculate the cross-Alpine pressure difference:

$$\Delta p_{alps} = p_{red}(KLO) - p_{red}(MAG)$$

Therefore, only the hourly averages of reduced surface pressures for these two stations were extracted.

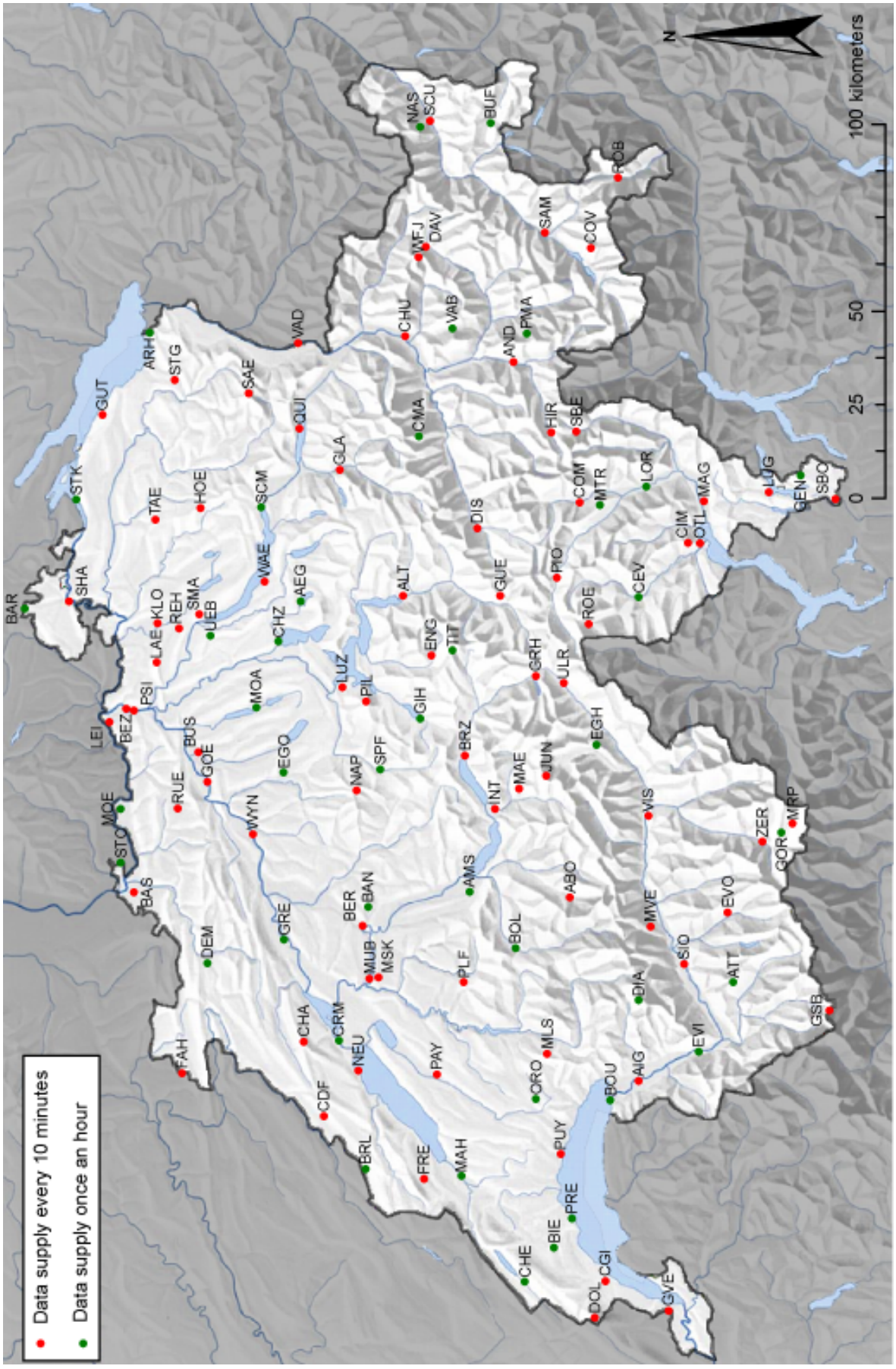

Figure 6: Locations of the SwissMetNet automatic surface weather stations operated by MeteoSwiss as of August 18th, 2009. Green dots indicate hourly data supply. Red dots indicate data supply every 10 minutes. (MeteoSwiss, 2010)

2.2 COSMO-2 model dataset

COSMO-2 reanalysis data from 01 hr on 24^{th} of February, 2008 to 23 hr on 31^{st} of December, 2011 were present in the archive of MeteoSwiss. Unfortunately, COSMO-2 underwent a pre-operational phase at MeteoSwiss prior to February 24^{th} during which no analysis data were archived. Days on which any station observed 10 minutes of foehn or longer were extracted from the archive. However, the hours of data for $19 - 23$ hr on of November 7^{th}, 2009 and 03 hr on October 31^{th}, 2010 were missing from the selected foehn days. The days were, therefore, omitted from the analysis. The extracted COSMO-2 data files were then converted from grib format to NetCDF format. The field variables that were kept for analysis are listed in table 1.

Point data were interpolated from these 2-D fields at the locations of the 20 stations. This was done by rotation of the geographical coordinates of each station in accordance to the rotated coordinate system of the COSMO-2 domain. The algorithm by Kaufmann (2008) was then applied to the surrounding grid points within a radius of 1.415 time the grid resolution (2.83 km for COSMO-2). This algorithm selects the grid point that minimizes the following equation:

$$d_{opt} = d_{hor} + |d_{vert}| \cdot f_{ve}$$

where $d_{hor} = \sqrt{\Delta x^2 + \Delta y^2}$ is the horizontal distance from the grid point to the station, $d_{vert} = \Delta z$ is the vertical distance from the grid point to the station with the vertical weight factor $f_{ve} = 500$. The use of this vertical weight factor can result in horizontal location discrepancies as large as 2.83 km for certain stations (figure 7). Time series containing the above model variables from all 4 levels were constructed for all 20 stations.

The coordinates of various passes and mountain peaks along the MAR were chosen (table 4). Linear interpolation between these points was done to form 10 line segments. The minimum distances from the MAR was then calculated for each interpolated foehn station in the model domain (figure 7).

Table 1: List of analysed COSMO-2 field variables relevant for foehn from the surface and the 3 bottom most model levels.

Variable	**Symbol**	**Model level**
Air potential temperature	θ	surface (2 m)
	${}^1\theta, {}^2\theta, {}^3\theta$	1, 2, 3
Relative humidity	RH	surface (2 m)
	${}^1RH, {}^2RH, {}^3RH$	1, 2, 3
Wind speed	V	surface (10 m)
	${}^1V, {}^2V, {}^3V$	1, 2, 3
Wind direction	V_{dir}	surface (10 m)
	${}^1V_{dir}, {}^2V_{dir}, {}^3V_{dir}$	1, 2, 3
Gust speed	V_{max}	surface (10 m)
Air pressure reduced to sea level	p_{red}	surface (0 m)

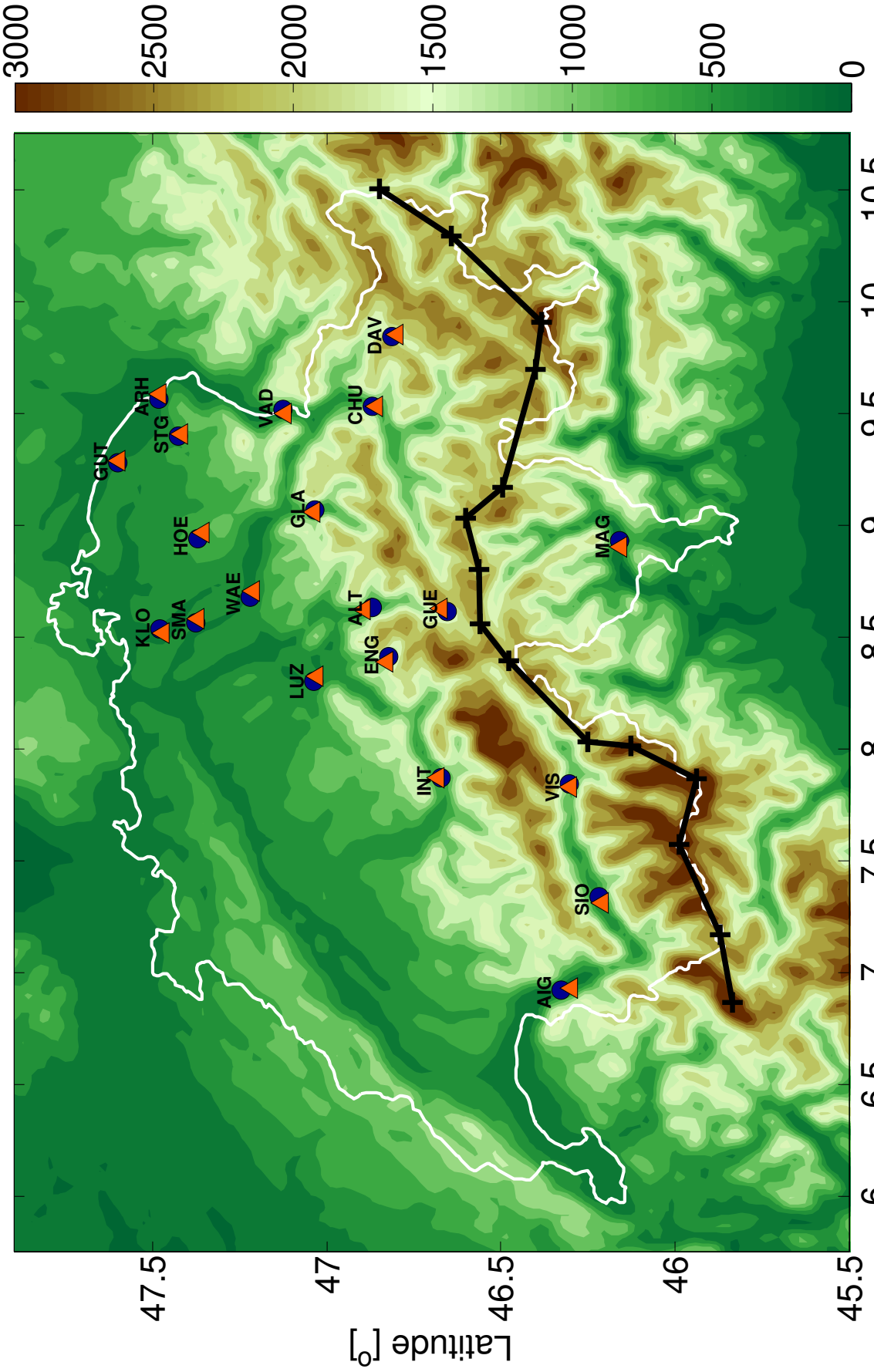

Figure 7: COSMO-2 topographical relief map of Switzerland. Approximation of MAR location from major mountain peaks and passes (back crosses) and the linear interpolations between them (black lines) are shown. Actual and interpolated station locations are indicated by blue dots and orange triangles respectively. Station names are abbreviated. For full station names and geographic coordinates see table 3 in appendix A. For geographic coordinates of the major mountain peaks and passes see table 4 in appendix G.

2.3 Foehn detection

Foehn episodes frequently manifest dramatic changes of potential temperature, humidity, wind/gust speed. Indeed, changes in these parameters serve as indicators and predictors for foehn detection and forecasting, respectively. However, distinguishing between foehn and other atmospheric phenomena such as thunderstorm outflows or strong nocturnal down-valley winds is difficult (Drechsel and Mayr, 2008). In the past, foehn detection methods were principally based on the orientation of synoptic pressure fields (Widmer, 1966; Courvoisier and Gutermann, 1971). However, the occurrence and strength of foehn in individual valleys was uncertain due to their unique orientations and topographies. In the last decade however, more objective methods of foehn detection have been developed.

Drechsel and Mayr (2008) developed a foehn detection method combining the indicators of isentropic descent and cross-barrier pressure gradients. The use of both indicators provides a mesoscale ($2 - 2\,000$ km horizontal scale) fingerprint of foehn events. Drechsel and Mayr (2008) verified these parameters against three years of observational data from surface weather stations in the Wipp Valley in the Austrian Alps. This method can also be applied to locations where topography is insufficiently resolved by general circulation models. Moreover, the joint probabilities of these parameters show a higher skill of foehn detection than the probabilities of the separate parameters.

Duerr (2008) developed another method for detecting south foehn similar to that of Vergeiner (2004). The former method differs from the latter in that it exploits the decrease in relative humidity during foehn in addition to the lee slope descent of isentropes, increases in wind/gust speed and constant southerly in wind direction. Thresholds for these variables were statistically determined from 10 years of observational data from 18 surface weather stations of the Swiss automatic monitoring network (ANETZ). These thresholds are unique to each of the 17 "foehn stations" and take the upstream and the local topographical effects on the foehn flow into account. The 18^{th} station is a reference station located close to the Main Alpine Ridge (MAR), from which the foehn begins its adiabatic descent along an isentropic trajectory. This station serves as a representative sample for all the foehn air flowing over the MAR within Switzerland. Differences in potential temperature ($\Delta\theta$) between the reference station and the foehn stations is a measure of the extent of isentropic descent. In a stably stratified atmosphere, potential temperature increases with altitude, which yields negative $\Delta\theta$ values. These values approach zero during foehn wind storms, however improper inter-calibration of temperature sensors, turbulent mixing in hydraulic jumps at the surface, and entrainment of the air aloft and from tributary valleys can cause small deviations. The average deviation from zero (offset) must, therefore, be determined statistically from a series of foehn events.

Duerr (2008) defined foehn station thresholds as the 90th percentiles of $\Delta\theta$, wind/gust speed as well as the 99th percentile of relative humidity. He also determined the main foehn wind direction thresholds for each station as a consequence of the local channelling of the foehn flow by the valley topography and its general orientation. This algorithm combines the joint probabilities of each foehn relevant variable, aside from the cross-alpine pressure gradient. It has been implemented at the Swiss Federal Office of Meteorology and Climatology (MeteoSwiss) since July

of 2008. It also has the benefit of it being fully automated and can be applied at an hourly or 10 minute time resolution. However, the validity of this method is restricted to valley floors or near valley outlets where, to some extent, foehn flow is channelled (Duerr, 2008).

2.4 Foehn detection applied to COSMO-2

The foehn detection method of Duerr (2008) was applied to the model time series to yield a binary indication of foehn occurrence at each station on an hourly basis. In the presence of a southerly wind direction at the reference station (Gütsch ob Andermatt), foehn onset occurs only when all thresholds are met at a given foehn station. Foehn can persist at a foehn station provided none of the thresholds (excluding wind/gust speed) are violated. The foehn parameter thresholds applied to all in the detection algorithm to all model levels are listed in table 2.

A main challenge in applying this algorithm to a numerical model is the constraints on wind direction. Since COSMO-2 topography is only modelled at a 2 km resolution, the wind fields in valleys can deviate from those observed. Even though the model may be producing an accurate foehn event concerning all other parameters, some wind direction thresholds might not be met. This would lead to a false diagnosis of foehn and would lead to a false lack of forecast skill. Moreover, differences in altitude of the station locations in the model can cause biases in surface pressure and air temperature. Adapting these foehn thresholds to the model topography would require the replication of the climatological analysis done by Duerr (2008) with COSMO-2 data. However, to do this would require several more years of COSMO-2 analysis data than are currently available. An additional challenge is that gust speeds are only available at the surface model level (10 m). It was therefore assumed that gust speeds in the 3 higher model levels were equal to those of the surface model level.

Table 2: List of foehn station parameter thresholds. Non-published thresholds are in italics. (Duerr, 2008, 2011)

Station	V_{dir}^{min} [°]	V_{dir}^{max} [°]	RH [%]	$\Delta\theta$ [**K**]	V[**m/s**]	V_{max} [**m/s**]
Aigle	90	210	56	-3.4	4.2	7.1
Altdorf	60	240	54	-4.0	3.7	6.2
Altenrhein	*120*	*210*	*52*	*-3.5*	*5.3*	*7.7*
Chur	140	260	57	-3.3	2.8	5.4
Davos	90	270	72	-2.0	3.4	5.6
Engelberg	10	160	61	-1.6	2.5	5.7
Glarus	60	190	56	-3.5	4.9	8.1
Güttingen	90	190	58	-3.1	2.8	4.9
Hörnli	*90*	*235*	*64*	*-1.7*	*4.9*	*8.1*
Interlaken	0	250	50	-2.1	4.9	9.2
Luzern	*90*	*190*	*56*	*-2.8*	*3.7*	*5.8*
Sion	20	150	56	-3.2	3.0	4.9
St. Gallen	90	200	61	-2.7	3.3	5.5
Vaduz	60	260	48	-2.8	4.1	7.6
Visp	60	180	56	-3.4	2.9	4.8
Wädenswil	*110*	*220*	*58*	*-3.7*	*3.5*	*5.7*
Zürich-Fluntern	*110*	*200*	*54*	*-2.5*	*3.7*	*6.7*

2.5 Systematic verifications of foehn in COSMO-2

A systematic verification of foehn in COSMO-2 was performed following three approaches. Firstly, general measures of COSMO-2 foehn forecast skill were attained through categorical verification methods. These measure the correspondence between foehn hours as binary events that occur in the observations and the model. The skill scores calculated from this class of verification summarized the collective skill of all modelled foehn variables. Secondly, a the model verification of the individual foehn variables was done using classical statistical methods. The systematic model biases were determined with this verification class. Finally, a novel approach for measuring the temporal lag of foehn onset and decay in COSMO-2 was done.

2.5.1 Categorical verification

Categorical verifications were done by comparing the foehn hours detected in COSMO-2 and the observations using the detection algorithm by Duerr (2008) as a binary indicator of foehn occurrence. Several categorical skill scores were calculated based on the foehn hour contingency in table 8. These scores were iteratively calculated by progressively discarding foehn events in both COSMO-2 and the observations with wind speeds lower than a given threshold. The frequency bias ratio (B) or Frequency Bias Index (FBI) of foehn mean is defined as:

$$B = \frac{a+b}{a+c}.$$

This is also known as the frequency bias score, which estimates the ratio of forecasted and observed foehn hours (B=1 is unbiased, $B > 1$ overestimated and $B < 1$ underestimated). However, the bias score is not a measure of correspondence. The Probability of Detection (POD) is a better measure of model skill:

$$POD = \frac{a}{a+c}.$$

This is also known as the hit rate, and represents the fraction of all observed events that were correctly forecasted. Conversely, the False Alarm Ratio (FAR) is defined as:

$$FAR = \frac{b}{a+b}.$$

This represents the fraction of forecasted events that were false alarms. However, a more balanced score is the Critical Success Index (CSI) or Threat Score (TS) and is defined as:

$$TS = \frac{a}{a+b+c}.$$

This score takes both false alarms and missed events in to consideration, unlike the POD and FAR. However, the TS is somewhat sensitive to the event climatology: it tends to yield poorer scores when events are rare.

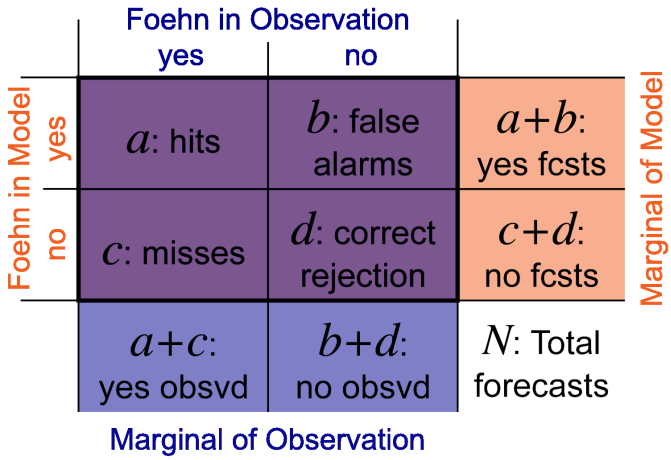

Figure 8: Contingency table for observations and COSMO-2. Adapted from Warner (2011).

2.5.2 Classical verification

At a given foehn station, COSMO-2 and observational data were discarded during hours when no foehn was detected in the observations with the algorithm by Duerr (2008). A classical model verification was then done with the remaining data. The biases, linear correlation coefficients and Mean Absolute Errors (MAE)s between the model and observations of all foehn relevant variables (table 1) were calculated during observed foehn hours. Classical model verification of the reference station Gütsch ob Andermatt and Δp_{alps} was done for hours on which any foehn occurred at any foehn station. The bias, or Mean Error (ME) is defined as:

$$bias = \frac{1}{N} \sum m_i - o_i = \bar{m} - \bar{o},$$

where m_i and o_i are the model and observation data pair for the sample $i = 1, ..., N$ and $\bar{m}$ and $\bar{o}$ are the model and observation means, respectively. A mean bias of 0 would indicate that on average the model variable in question has no deviation from the observations. However, a mean bias of 0 does not imply a perfect model, since it does not take the bias distribution into account. To measures the strength of the linear relationship between observation and model data sets the linear correlation coefficient is used:

$$r = \frac{\frac{1}{N} \sum (m_i - \bar{m}) \cdot (o_i - \bar{o})}{s_m \cdot s_o},$$

where s_m and s_o are the sample standard deviations of the model and observations, respectively. Correlation coefficients can range from -1 to +1 and are insensitive to biases or errors in variance. Importantly, the square of the correlation coefficient is a measure of potential model skill (Wilks, 2011). For circular data, such as wind direction, the following circular correlation coefficient was used (Fisher and Lee, 1983):

$$r_{circ} = \frac{\sum M_i \cdot O_i}{\sqrt{\sum M_i \sum O_i}},$$

where $M_i = sin(m_i - m_{i+1})$ and $O_i = sin(o_i - o_{i+1})$.

Amplitude and phase errors are two aspects of model-observation agreement that are relevant to this model validation. While the bias may sometimes give insight into the general direction of model errors, other times competing negative and positive errors can cancel out and yield no useful model skill information. The mean amplitude error of the model is measured by the MAE, which takes the distribution of bias errors into account. The extent of phase errors (linear correspondence) is quantified by the correlation coefficient. Comparison of MAEs from different variables is difficult because it is not a dimensionless quantity like the correlation coefficient. Therefore, the MAEs were normalized to the observed standard deviations of their respective variables:

$$MAE = \frac{1}{N} \sum \mid m_i - o_i \mid.$$

The MAE is a common measure of model skill and resolves the issue of compensating negative and positive errors. It also has the advantage of being a linear score, which means that each error has an equal weight in the overall statistic.

2.5.3 Lag time of foehn onset and decay

The temporal error in foehn occurrence in COSMO-2 was evaluated on a daily basis. The first and last hours of foehn were determined for each day (foehn onset and decay) in the observations and model time series. Foehn days in which the modelled foehn has no corresponding foehn observation (or vice versa), were discarded. Instances when foehn occurs semi-continuously between several days were also possible. Therefore, foehn onset and decay pairs that were falsely generated by the algorithm figure between consecutive foehn days were also discarded (i.e. if there was a decay at 23 hr and an onset at 01 hr the next day). An illustration of this concept can be seen in figure 9. The bias distribution of the onset and decays of the model were then calculated.

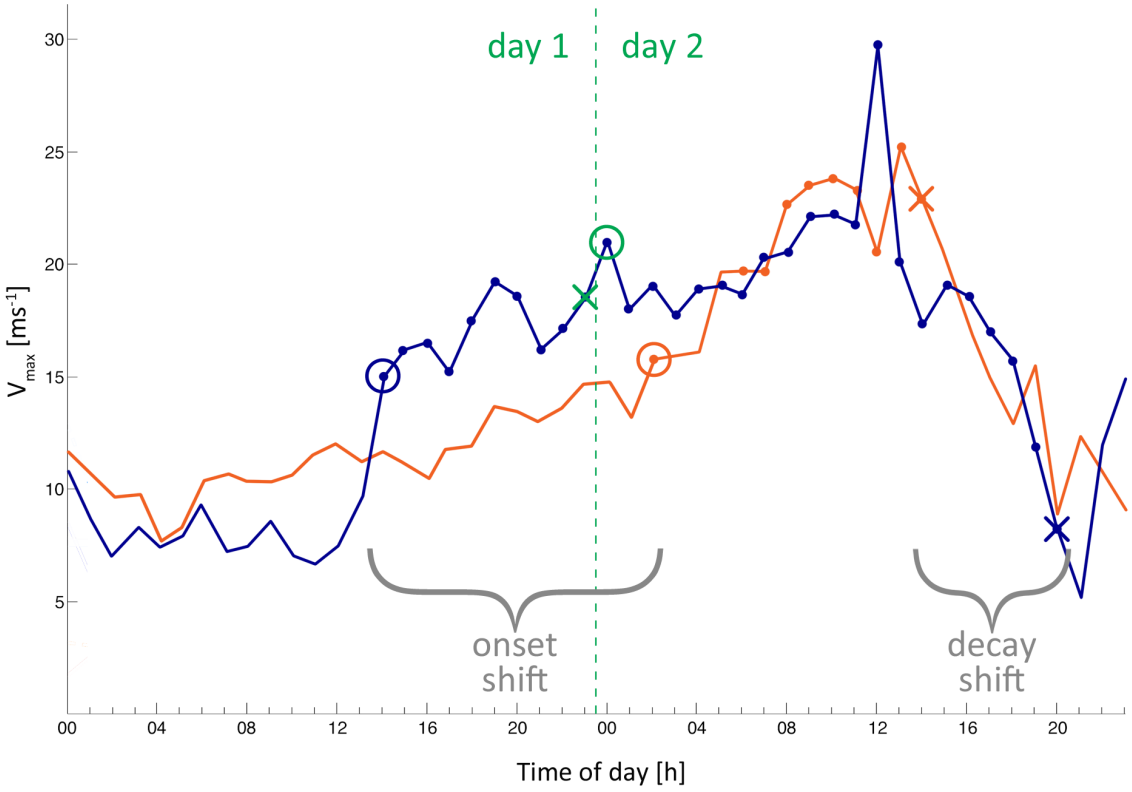

Figure 9: Lag in foehn onset and decay algorithm applied to multi-day foehn event example. The observed (blue) and modelled (orange) $\Delta\theta$ at 2 m above the surface at Vaduz station during a multi-day foehn event are shown. Foehn hours are marked with dots. The first foehn hour of the day (onset) and is marked with an 'O'. The final foehn hour of the day (decay) and is marked with an 'X'. Shown in green are the false onset/decay pairs at 23 hr and 00 hr of the following day that are removed from analysis.

3 Results

3.1 Overview of results

For a process-oriented verification, as is done in this thesis, an issue of validation scope/aim arises when attempting to quantitatively compare observed and modelled foehn. If the aim is to assess the ability of the model to reproduce foehn variables at the precise time it is observed then one can verify the model during observed foehn hours easily. However, if the aim is to asses the ability of the model to reproduce realistic foehn characteristic regardless of when they occur in relation to reality, direct classical verification of the two datasets is not possible. Classical model verifications (comparison of continuous model field variables) require identical sample sizes between the two datasets. This means that one must compare data from the same time index. If erroneous foehn occurs in the model at times when it is not observed then the model classical statistics, calculated only when foehn is observed, will not necessarily be representative of the foehn characteristics in COSMO-2. Visual inspection of figure 10 shows a greater agreement between the two data clouds when all foehn instances in the model are included regardless of the observations. Indeed, these two approaches can yield different amounts of agreement between the two datasets. In light of this, categorical verifications were included to asses the extent that this erroneously occurring foehn and its effect on the foehn detection skill of COSMO-2.

In this section the results of various model verification steps are presented. In addition to the 2D-parameterized variables at the surface (2 m and 10 m), the first 3 model levels were taken into account and are presented separately. Categorical and classical verifications as well as temporal lag analysis are discussed for both variable groups. Seasonal and overall classical verifications of amplitude (MAE) and phase errors (correlation) are shown. The individual foehn station biases are analysed in relation to distance from the MAR as well as to model altitude error. Finally, the temporal lag of foehn and an optimization of model foehn skill are discussed.

It is important to reiterate that the reference station Gütsch is assumed to be a representative sample of the conditions at the Alpine crest, to which all foehn stations are compared. Therefore, the classical verifications for Gütsch were done for hours when foehn occurred at any foehn station. Moreover, the verifications of Gütsch are presented in parallel to the foehn stations in the sections to follow, since it is not a foehn station itself.

Cumulative 4-year-long model verifications were done during foehn hours corresponding to each foehn station. However, the occurrence of foehn has a strong seasonal dependency, leading to a large range of sample sizes. In figure 11, the seasonal variation of mean foehn hours diagnosed with surface variables is shown for both observations and COSMO-2. The overall agreement with observations seems to be low - the poorest in the latter half of the 4 year period. However, the maxima tend to occur during the spring in both the observations and COSMO-2. The largest overestimation occurs during the summer months: the largest of which is 882 hours in 2011. The observed maximum in the spring of 2010 is the only exception, where the sample size is underestimated by approximately 400 foehn hours. Conversely, the sample size is much lower than observations during the winter seasons; the largest

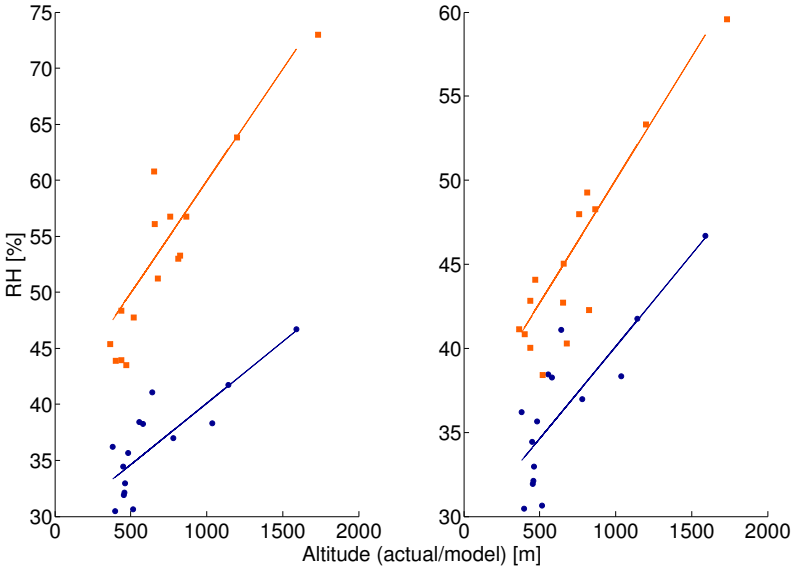

Figure 10: Scatter plots of mean RH for individual foehn stations and Gütsch relative to actual/model altitude. **Left:** Both COSMO-2 and observed θ means calculated during **observed** foehn hours. **Right:** COSMO-2 θ means calculated during **COSMO-2** foehn hours and observed θ means calculated from observed foehn hours. Observations are shown in blue and COSMO-2 data are shown in orange.

underestimation is 843 foehn hours during the winter of 2010.

A general description of the foehn variable means and variances is required in order to understand the size of the relative magnitudes of normalized forecast errors ($\frac{MAE}{s_o}$) that shall be presented. By combining data from all foehn stations, the general magnitude and spread of COSMO-2 foehn variables can be seen in relation to the observations (figure 12). When analysed separately, some distributions of individual stations suffer from very low sample sizes (less than 100 hours over 4 years). These distributions tend to be coarse and harder to interpret. For these reasons, the foehn stations Güttingen (GUT) and Zürich-Flutern (SMA) were only included in the combined foehn station analyses. Relative to the observations, the COSMO-2 variables V, θ and $\Delta\theta$ are further positively skewed. Contrarily, V_{max}, V_{dir}, and RH in COSMO-2 are skewed further to the higher values than the observations. The general shapes of the distributions are well retained by the model, the exception being RH. Additionally, the steep lower tail of observed $\Delta\theta$ is not well reproduced.

The combined density distributions of surface-level foehn variables for Gütsch as well as Δp_{alps} are shown in figure 13. The variables V_{dir} and θ are of particular interest at the reference station Gütsch, since only these two variables determine the possibility of foehn occurrence for the foehn stations (see section 2.3). V_{dir} shows larger variance in COSMO-2 but no large shift in the mean. However, large departures from the observed distribution shape of θ occur in COSMO-2 as it shifts to the left. A large peak in frequency at 292 K as well as a more distinct bimodal distribution can be seen. The synoptic cross-Alpine pressure difference is well represented in COSMO-2. Only a small reduction in variance is apparent.

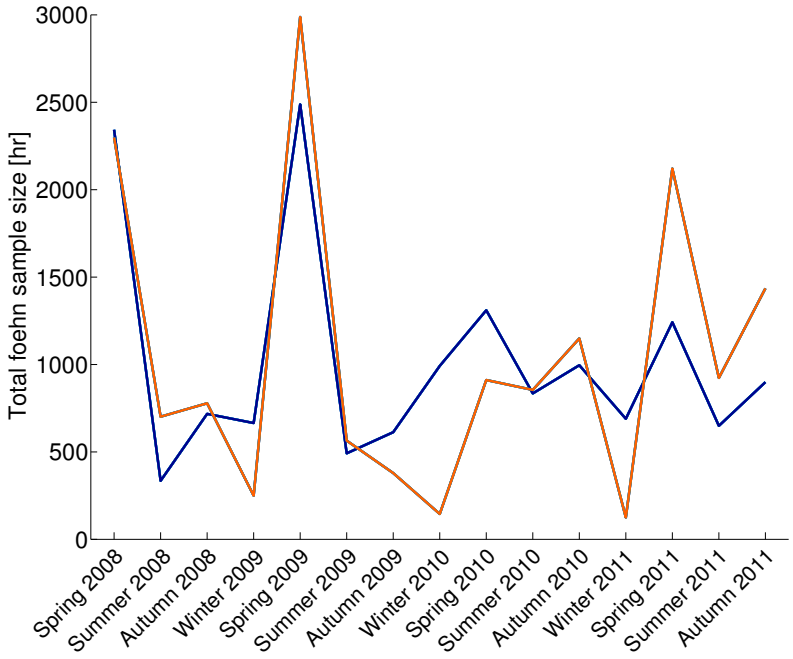

Figure 11: Total seasonal foehn hours from all foehn stations. COSMO-2 foehn hours are diagnosed from surface-level foehn variables. Observations are shown in blue and COSMO-2 data are shown in orange.

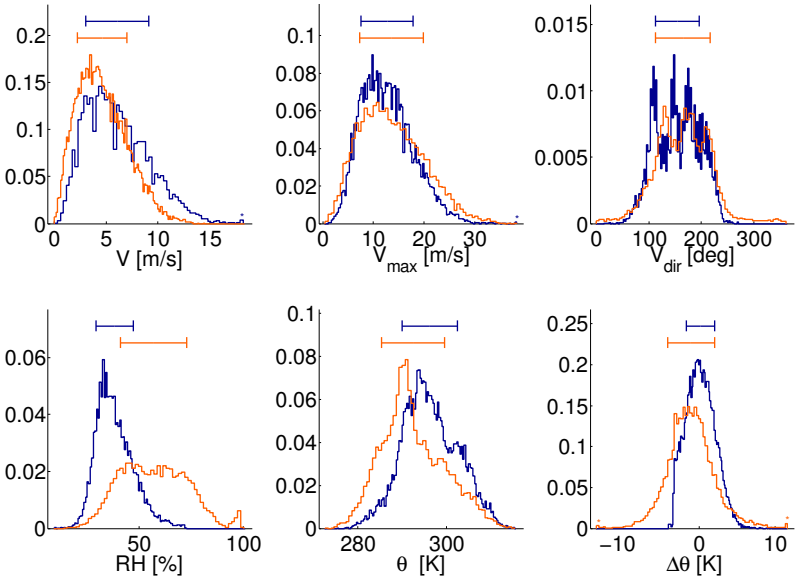

Figure 12: Combined density distributions of surface variables V, V_{max}, V_{dir}, RH, θ and $\Delta\theta$ of all foehn stations for corresponding foehn hours. Observations are shown in blue and COSMO-2 data are shown in orange. Horizontal bars show the standard errors about the means.

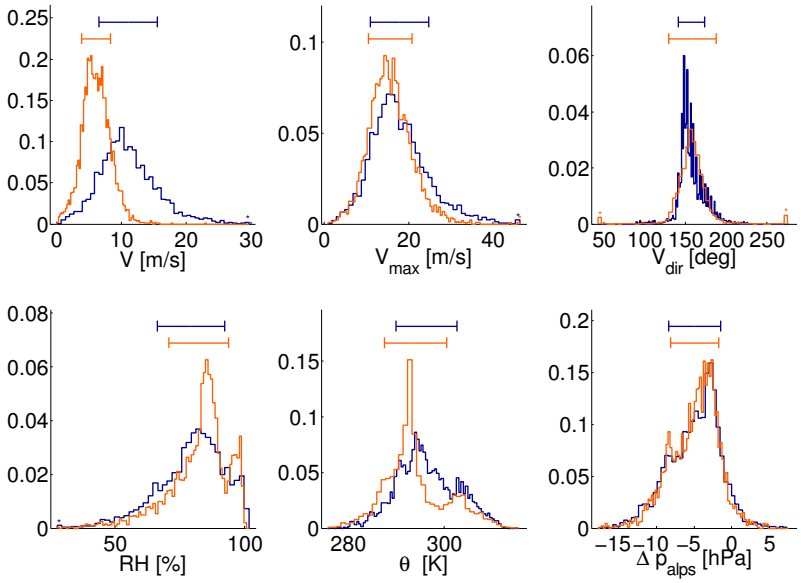

Figure 13: Density distributions of surface foehn variables V, V_{max}, V_{dir}, RH and θ for Gütsch as well as Δp_{alps} during hours when foehn occurred at any foehn stations. Observations are shown in blue and COSMO-2 data are shown in orange. Horizontal bars show the standard errors about the means.

3.2 COSMO-2 surface level

3.2.1 Categorical verification

The frequency bias scores of individual foehn stations relative to their modelled altitude are shown in figure 14. Six stations, Lucerne (LUZ), Aigle (AIG), Altenrhein (ARH), Wädenswil (WAE), Interlaken (INT) and St. Gallen (STG), all have frequency bias scores greater than one. The stations Vaduz (VAD) and Glarus (GLA) are both equal to unity. The six remaining stations, Sion (SIO), Chur (CHU), Altdorf (ALT), Visp (VIS), Engelberg (ENG) and Davos (DAV), all have frequency bias scores between 0.5 and 0.9. Lucerne has the highest value at 3.9, while Chur has the lowest at 0.5. The highest stations in altitude, ENG and DAV, have frequency bias scores of similar magnitudes. In general, stations with lower altitudes show larger variance than higher altitude stations. In particular, the station cluster LUC, AIG, ARH and WAE, with altitudes less than 500 m, shows the largest variations from the mean value of 1.6. However, stations of comparable altitude to this cluster, such as VAD (428 m) and SIO (518 m) do not deviate from the observations as strongly.

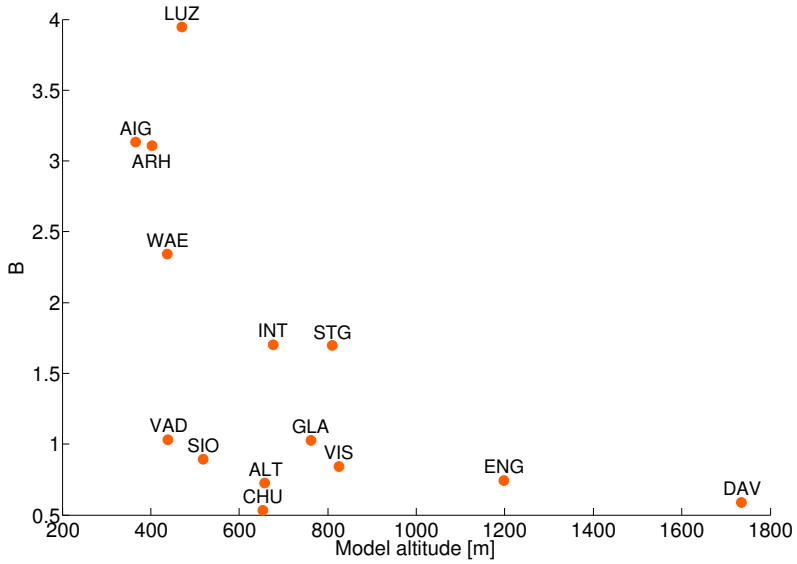

Figure 14: Surface-level frequency bias score vs. model altitude for all foehn hours at individual foehn stations. Station names are abbreviated.

The skill scores POD and FAR of all foehn stations combined are shown in figure 15 as a function of foehn intensity. A maximum POD value of 0.45 occurs when no wind threshold is applied. As the wind speed threshold is increased to approximately 17 kts, the POD decreases rapidly to 0.09. Concurrently, the FAR also decreases until approximately 8 kts, after which it increases with the wind speed thresholds. The maximum FAR is 0.63, and occurs approximately at the 18 kts threshold. After this point, both POD and FAR decrease to zero very rapidly.

Figure 16 shows the TS and the total foehn hours in COSMO-2 combing all foehn stations as a function of foehn intensity. This summarizes the information from the POD and the FAR in a single skill score: TS. The maximum TS value is 0.29 and occurs when no wind threshold is applied. TS decreases slowly with the

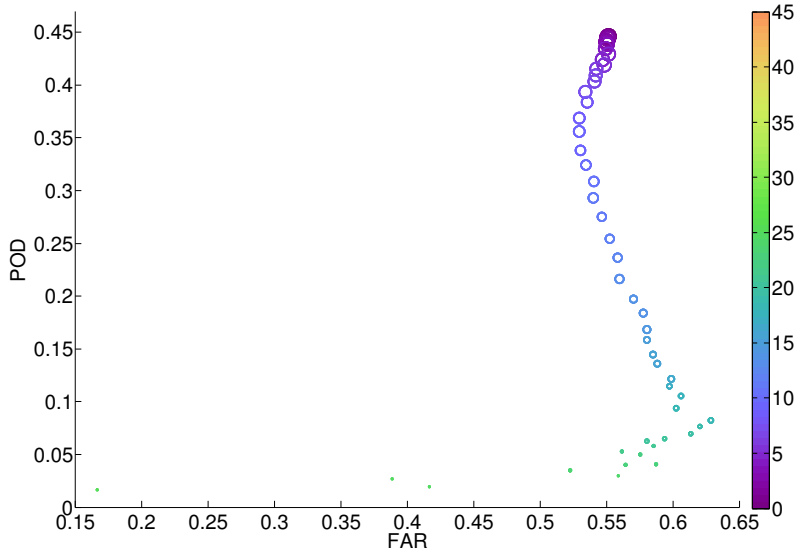

Figure 15: Surface-level POD vs. FAR for all foehn stations by foehn intensity (V [kts]). The colour bar indicates the intensity thresholds applied to the model and observational data to calculate the corresponding POD and FAR scores. The point size is linearly scaled to the total number of COSMO-2 foehn hours.

wind threshold until approximately 9 kts. After this, it decreases in an exponential fashion until abruptly reaching 0 at 24 kts.

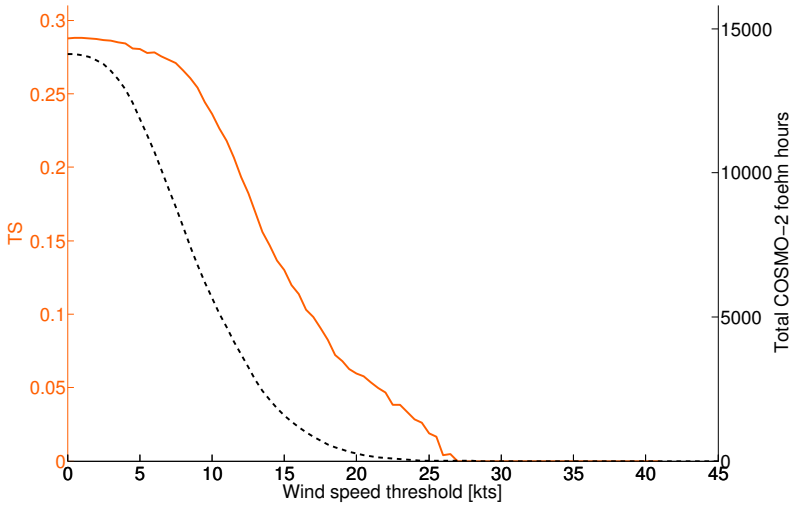

Figure 16: Surface-level TS for all foehn stations vs. foehn intensity (V [kts]). The intensity thresholds were applied to the model and observational data to calculate the corresponding TS scores.

3.2.2 Classical verification

The biases of all surface-level foehn variables are shown in figure 17. The biases of all variables are unique for every station location. Wind and gust speeds show a positive linear trend with the distance from the MAR. Stations further away tend to have positive biases, while closer stations have negative biases. However, the wind direction bias, as expected, shows no dependence on MAR distance. The mean bias is 9.5° and all stations, except for Sion, fall between -19° to 33°. Sion has the greatest wind direction bias of 61°. Note that a positive bias in wind direction is defined as a clockwise rotation. Biases of RH are all positive and increase linearly with the error in model station altitude. Gütsch has the lowest value of 3.0 % and Davos has the maximum value of 26.3 %. The spread about the main trend is greatest for Vaduz and Visp. The same holds true for θ and $\Delta\theta$, even though both variables have negative linear trends with respect to model station altitude error. Similar station clustering as described in section 3.2.2 occurs for all three thermodynamic variables (RH, θ and $\Delta\theta$). Again, the reference station Gütsch is always isolated and is either the maximum or minimum bias station. On the opposite extreme, the stations Altdorf, Chur, Davos, Engelberg and Chur are clustered. Stations with middle-range biases form the other cluster with the stations Aigle, Altenrhein, Interlaken, Lucerne, Sion, St. Gallen and Wädenswil.

The normalized MAE and correlation coefficients of each variable for all foehn stations and Gütsch are shown in figure 18. The correlation coefficients for Gütsch range from 0.08 to 0.95 corresponding to θ and wind direction, respectively. The correlation of wind direction for the combined foehn stations shows a much higher correlation (0.64) while the correlation for θ is slightly lower (0.92). The correlations of variables at Gütsch, except for wind direction and speed, are all greater than those of the combined foehn stations. In addition, the normalized MAEs of all variables at Gütsch are lower than those of the foehn stations, except for wind speed. The

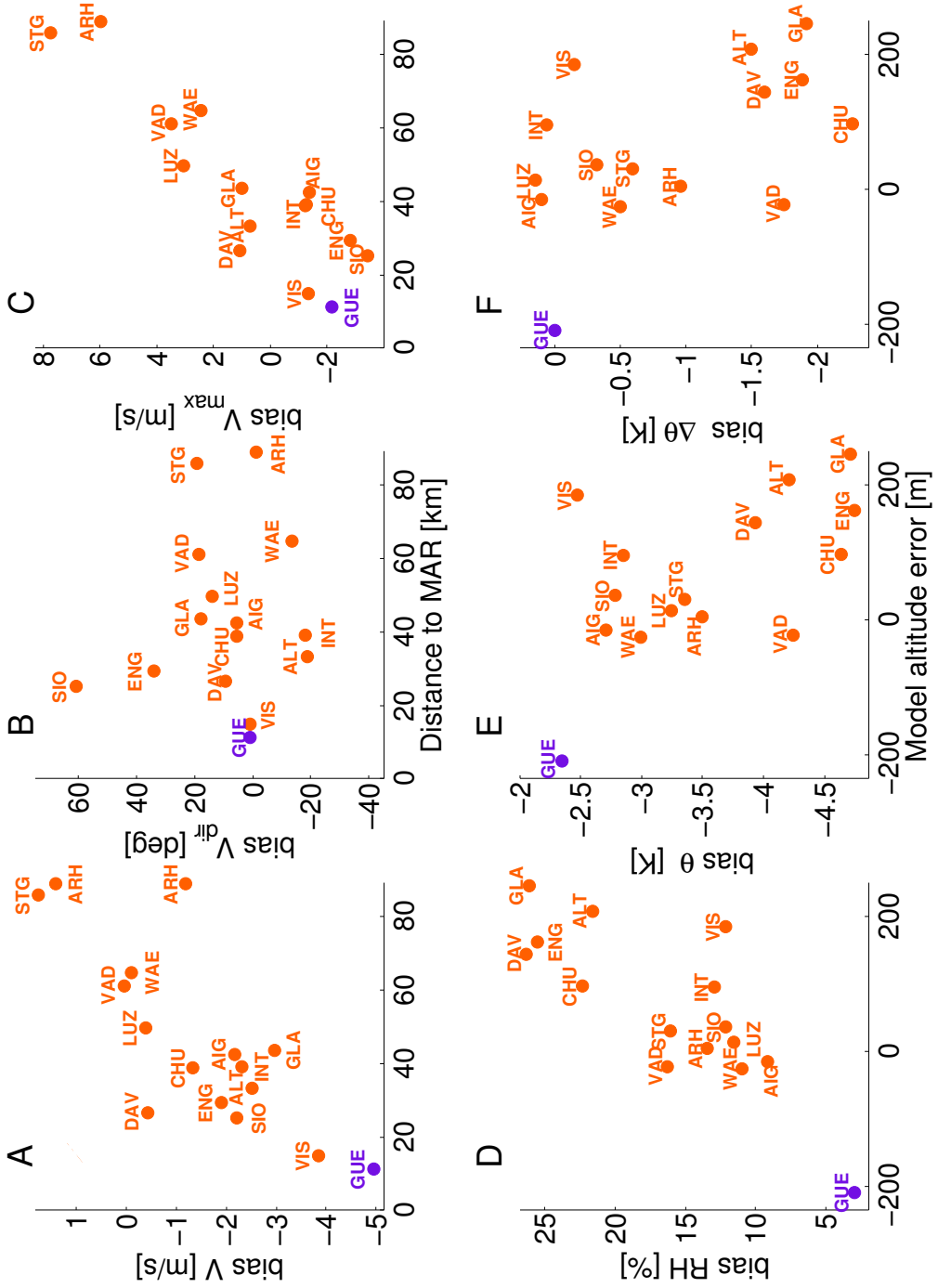

Figure 17: Biases of surface-level foehn variables for individual foehn stations (orange) and Gütsch (purple). **A, B, C:** biases of V, V_{max}, V_{dir} ,respectively, vs. station distance from MAR in COSMO-2. **C, D, E:** biases of RH, θ and $\Delta\theta$, respectively, vs. station altitude error in COSMO-2. Station names are abbreviated.

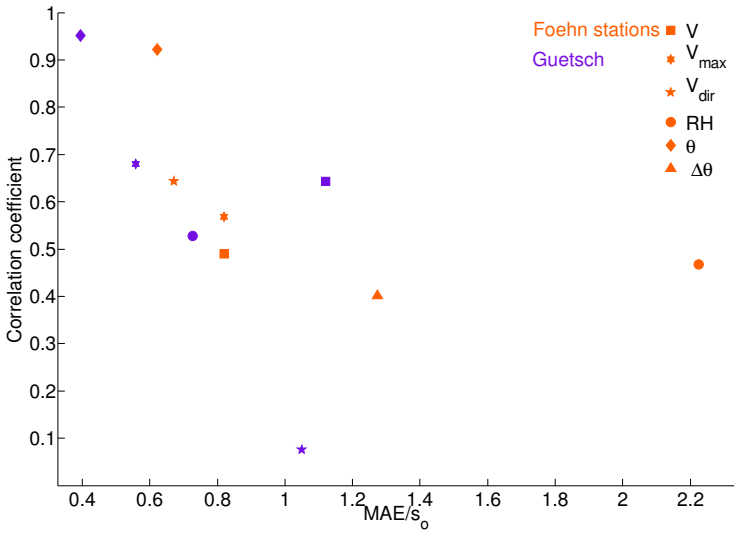

Figure 18: The correlations and normalized MAEs of all surface-level foehn variables from Gütsch (purple) and all combined foehn stations (orange). The MAEs of $\Delta\theta$ is not presented for Gütsch, since it is zero by definition.

largest discrepancy of correlation between the foehn stations and Gütsch is in wind direction (60% difference). The largest discrepancy in normalized MAE is in RH (150%). Conversely, θ shows the smallest change in both correlation (3% difference) and MAE (23% difference).

3.2.3 Lag of foehn onset and decay

The onset lag ranges between -22 and 15 hours. The decay lag ranges between -17 and 20 hours. The mean onset lag is -1.3 hours, while the mean decay lag is -0.3. The hourly time discretization is apparent in the stepped pattern of the probability plot. Both onset and decay deviate from Gaussian distributions strongly in their tails. Moreover, both tails in the onset lag distribution are equally extended, while the uppermost tail of the decay lag distribution shows the stronger extension (figure 19). Despite these differences between the onset and decay lag distributions, the standard deviation is 4.6 hours for both.

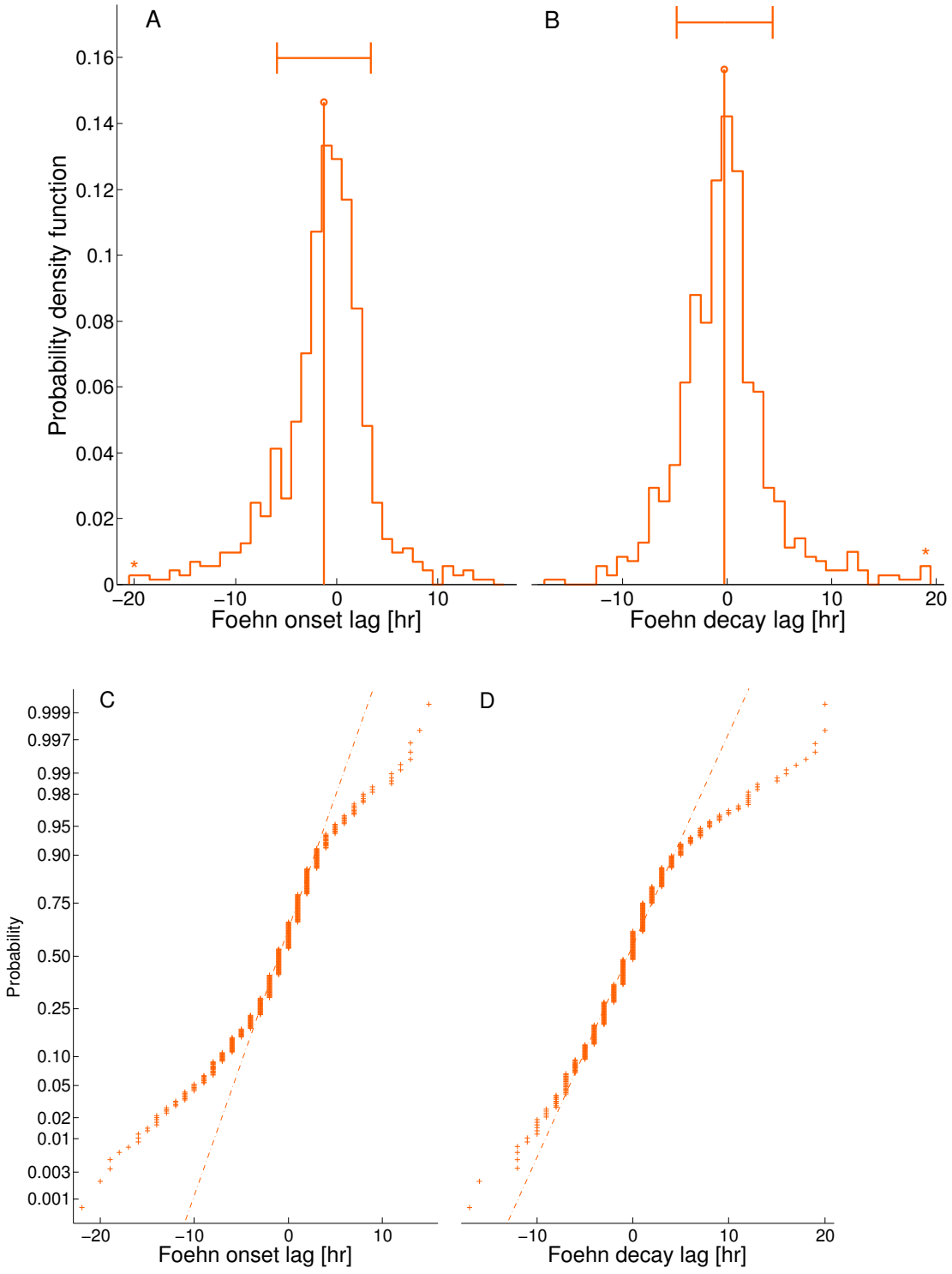

Figure 19: **Top:** combined density distributions of lag times of COSMO-2 foehn onset **(A)** and decay **(B)** for all foehn stations. Horizontal bars show the standard errors about the means, which are marked with a vertical stem. **Bottom:** probability plots of foehn onset **(C)** and decay **(D)** lag times. Normal distributions follow the dashed lines. COSMO-2 foehn hours are diagnosed from surface-level foehn variables.

3.3 Lowest 3 COSMO-2 levels

3.3.1 Categorical verification

The frequency bias scores of individual foehn stations relative to their modelled altitude are shown in figure 20. The frequency bias scores are larger at higher model levels for all foehn stations. However, the magnitudes of the increases are not constant between model levels nor foehn stations. Increases in frequency bias scores between the second and third levels tend to be smaller than frequency bias score increases between the surface and first levels. Stations with frequency bias scores near unity tend to have less variance between model levels. All stations have levels with at least one frequency bias scores greater than one, with the exception of Chur. The stations Altdorf, Davos, Sion and Visp all straddle unity and surpass it with higher model levels. Lucerne has the highest value of 5.6 at the third model level, while Chur has the lowest at 0.57. The highest stations in altitude, Engelberg and Davos, have frequency bias scores of similar magnitudes. Conversely, stations with lower altitudes show larger variance between stations than at higher altitudes. In particular, the station cluster Lucerne, Aigle, Altenrhein and Wädenswil, with altitudes less than 500 m, shows the largest variations from the mean values of 1.8, 2.3 and 2.6 for levels 1, 2 and 3 respectively. However, stations of comparable altitude to this cluster, such as Vaduz (428 m) and Sion (518 m) do not deviate from the observations as strongly.

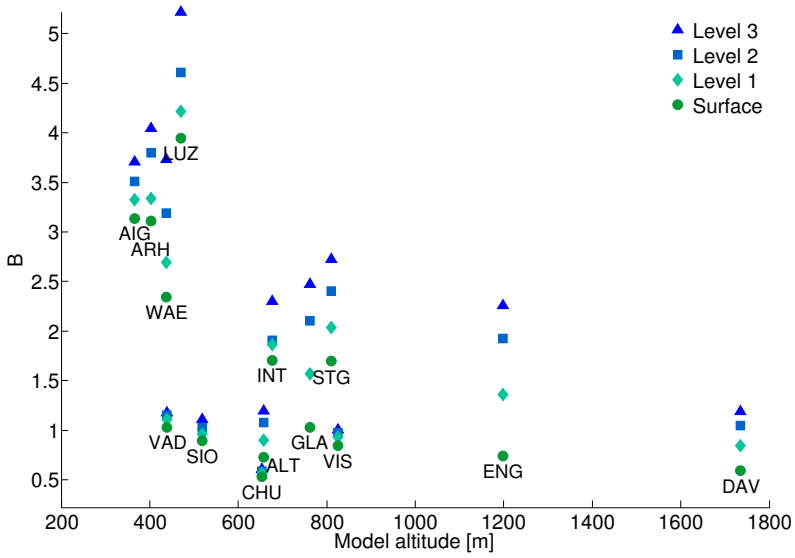

Figure 20: Frequency bias scores of all model levels vs. model altitude for all foehn hours at individual foehn stations. Station names are abbreviated and model levels are indicated by the colours and shapes listed in the legend.

The skill scores POD and FAR of combining all foehn stations are shown in figure 21 as a function of foehn intensity. For the first model level, the general shape resembles the surface-level POD and FAR pattern as described in section 3.2.1. The POD decreases rapidly to 0.07 as the wind speed threshold is increased to approximately 17 kts. Concurrently, the FAR also decreases until approximately 8 kts, after

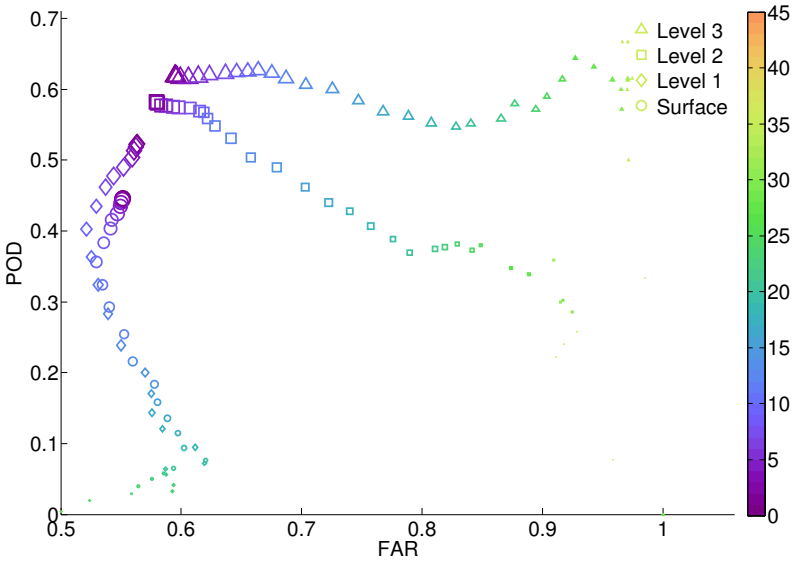

Figure 21: POD vs. FAR for all foehn stations by foehn intensity (V [kts]). The colour bar indicates the intensity thresholds applied to the model and observational data to calculate the corresponding POD and FAR scores. Model levels are indicated by the shapes listed in the legend. The point size is linearly scaled to the total number of COSMO-2 foehn hours.

which it increases with the wind speed thresholds. The maximum FAR is 0.62, and occurs approximately at the 18 kts threshold. After this point, both POD and FAR decrease to zero very rapidly. The second model level shows a roughly constant decrease in POD and increase in FAR until the 20 kts wind threshold is reached. At higher thresholds, the POD increases slightly, then decreases sporadically to zero as FAR increases to unity. The third model level maintains roughly constant POD values until approximately 30 kts, after which it increases sporadically. A maximum POD value of 0.67 occurs when an approximate wind threshold of 37 kts is applied. However, this is also accompanied by an FAR of 0.97.

Figure 22 shows the TS of all foehn stations combined as a function of foehn intensity. This summarizes the information from the POD and the FAR in a single skill score. The maximum TS value is 0.32 and occurs when wind threshold of 1.5 kts is applied to the second model level. For this level, TS decreases slowly with the wind threshold until approximately 11 kts. After this, it decreases in a roughly linear fashion until 29 kts, after which it remains constant (TS is approximately 0.075). TS drops abruptly from 0.075 to 0 at 36 kts. For thresholds higher than approximately 36 kts, the third model level shows the highest TS values of all model levels.

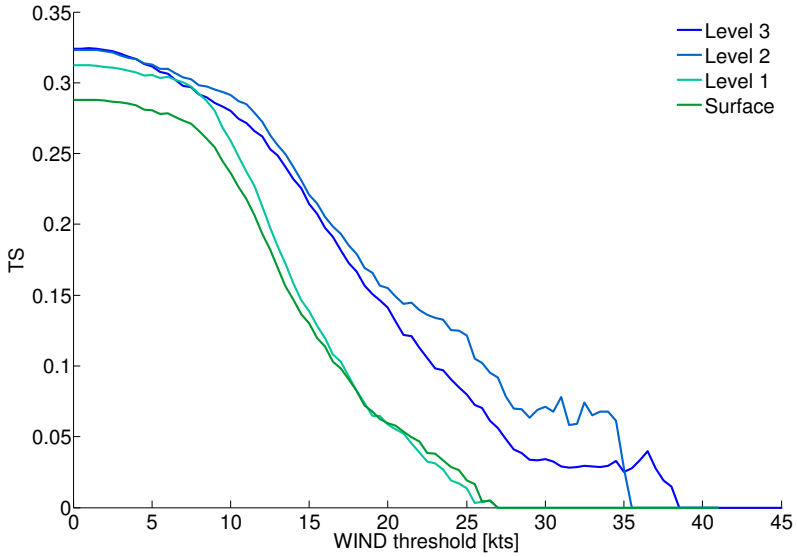

Figure 22: TS for all foehn stations vs. foehn intensity (V [kts]). The intensity thresholds were applied to both model and observational data to calculate the corresponding TS scores. Model levels are indicated by the colours listed in the legend.

3.3.2 Classical verification

The biases of all lower-level foehn variables are shown in figure 23. The biases of all variables are for every station location. Wind speed shows a positive linear trend with distance from the MAR. Stations further away from the MAR tend to have positive biases, while closer stations have negative biases. The slope of this trend is larger for higher model levels.

The wind direction bias shows no dependence on MAR distance. The mean biases are 9.6°, 10.1°, 10.3° for model levels 1, 2 and 3, respectively. All stations, except for SIO, fall within -19° to 36°. SIO has wind direction biases between 49° and 61°.

Biases of RH are all positive and increase proportionally with the error in model station altitude. RH bias decreases with model level. The spread about the main trend is greatest for Vaduz and Visp. The same holds true for θ and $\Delta\theta$, even though both variables have negative linear trends with respect to model station altitude error. Similar station clustering occurs for all three thermodynamic variables (RH, θ and $\Delta\theta$). The reference station GUE is always isolated and is either the maximum or minimum bias station. On the opposite extreme, the stations ATL, CHU, DAV, ENG, and GLA tend to cluster. Stations with middle-range biases form a second cluster (stations AIG, ALT, Int, LUZ, SIO, STG and WAE).

The correlation coefficients and normalized MAEs of each variable for all foehn stations are shown in figure 24. θ has the highest correlation coefficients of all foehn variables. The $\Delta\theta$ shows the lowest correlations. The correlation coefficients of all variables, except for wind speed, increases with model level height. Similarly, the MAEs of all variables decrease with model level height, except for wind speed. Gust speed shows the largest change between model levels in correlation (7%) and in normalized MAE (145%). θ has the smallest change in correlation (0.4%) and wind direction has the smallest change in normalized MAE (3%). For every variable, the

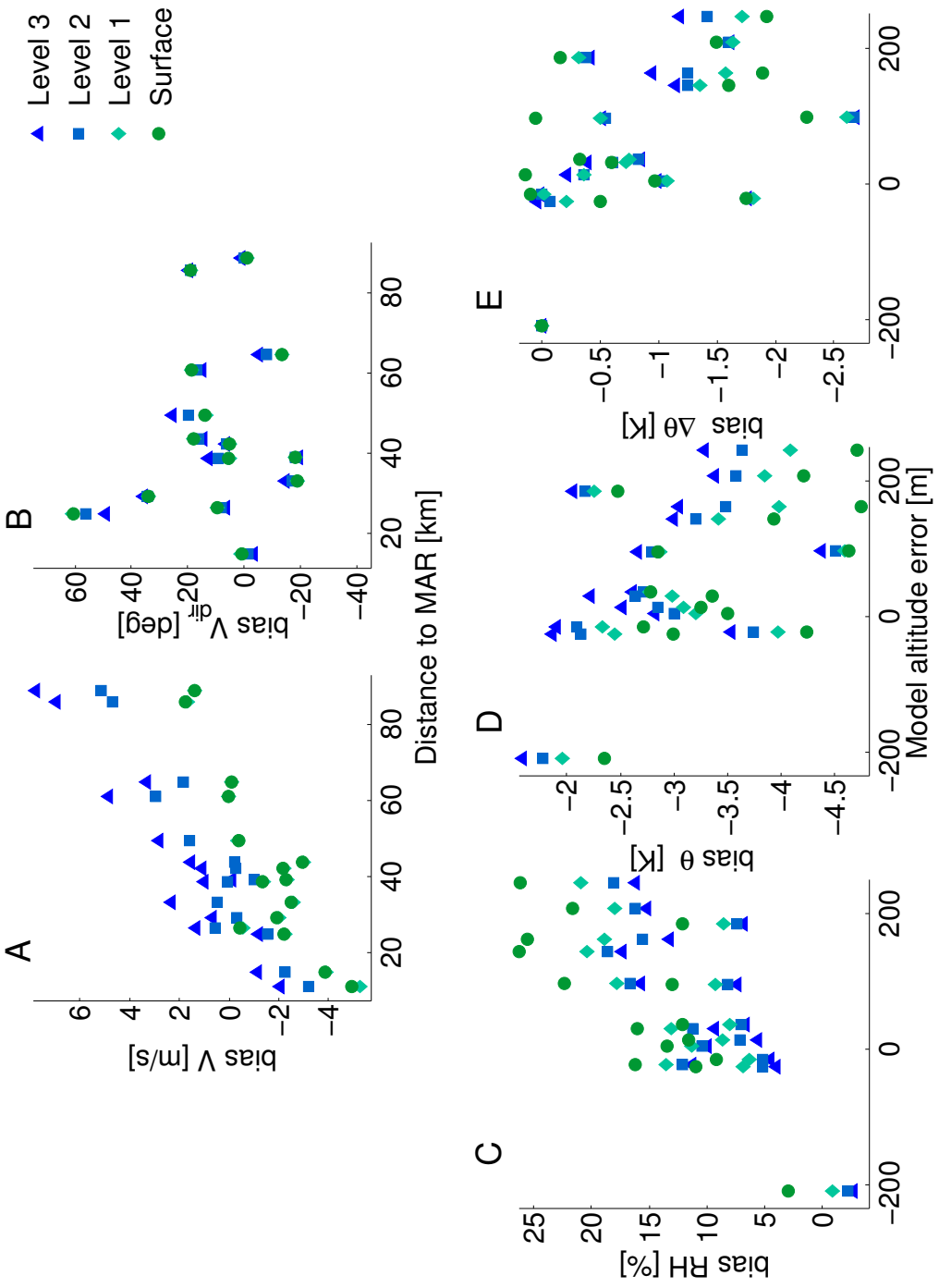

Figure 23: Biases of foehn variables from all model levels for individual foehn stations and Gütsch. **A, B:** biases of ${}^{123}V$ and ${}^{123}V_{dir}$ vs. station distance from MAR in COSMO-2. **C, D, E:** biases of ${}^{123}RH$, ${}^{123}\theta$ and ${}^{123}\Delta\theta$ vs. station altitude error in COSMO-2. Model levels are indicated by the colours and shapes listed in the legend.

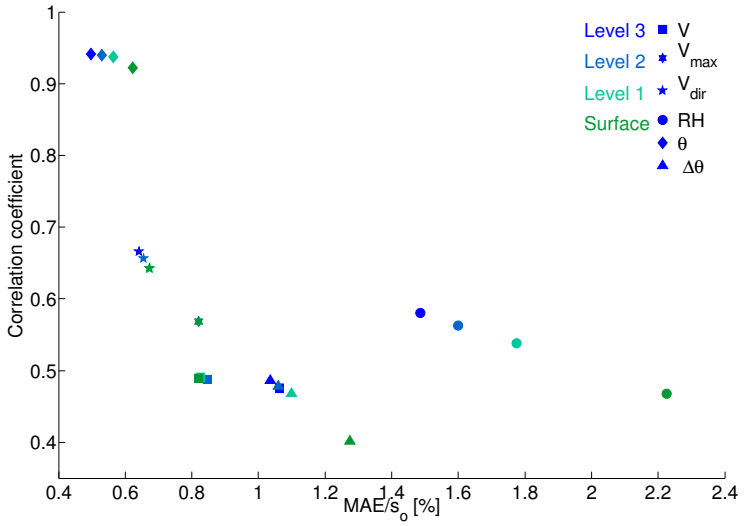

Figure 24: The correlations and normalized MAEs of foehn variables from all combined foehn stations. Model levels are indicated by the colours and shapes listed in the legend.

model levels with the highest correlation coefficients also show the lowest MAE relative to the other model levels.

3.3.3 Lag of foehn onset and decay

The onset lag ranges between -22 and 16 hours for all levels. The decay lag ranges between -17 and 21 hours. The mean onset lag decreases with model level height to a minimum value of -2.4 hours while the mean decay lag increases to 0.55 hours. The hourly time discretization is apparent in the stepped pattern of the probability plots (figure 25). Both onset and decay deviate from Gaussian distributions strongly in their tails. Moreover, both tails in the onset lag distribution are equally extended, while the uppermost tail of the decay lag distribution shows the stronger extension. The extensions of both distribution tails decrease with model level height for foehn decay lag. However, this is only true for the upper tail concerning foehn onset lag.

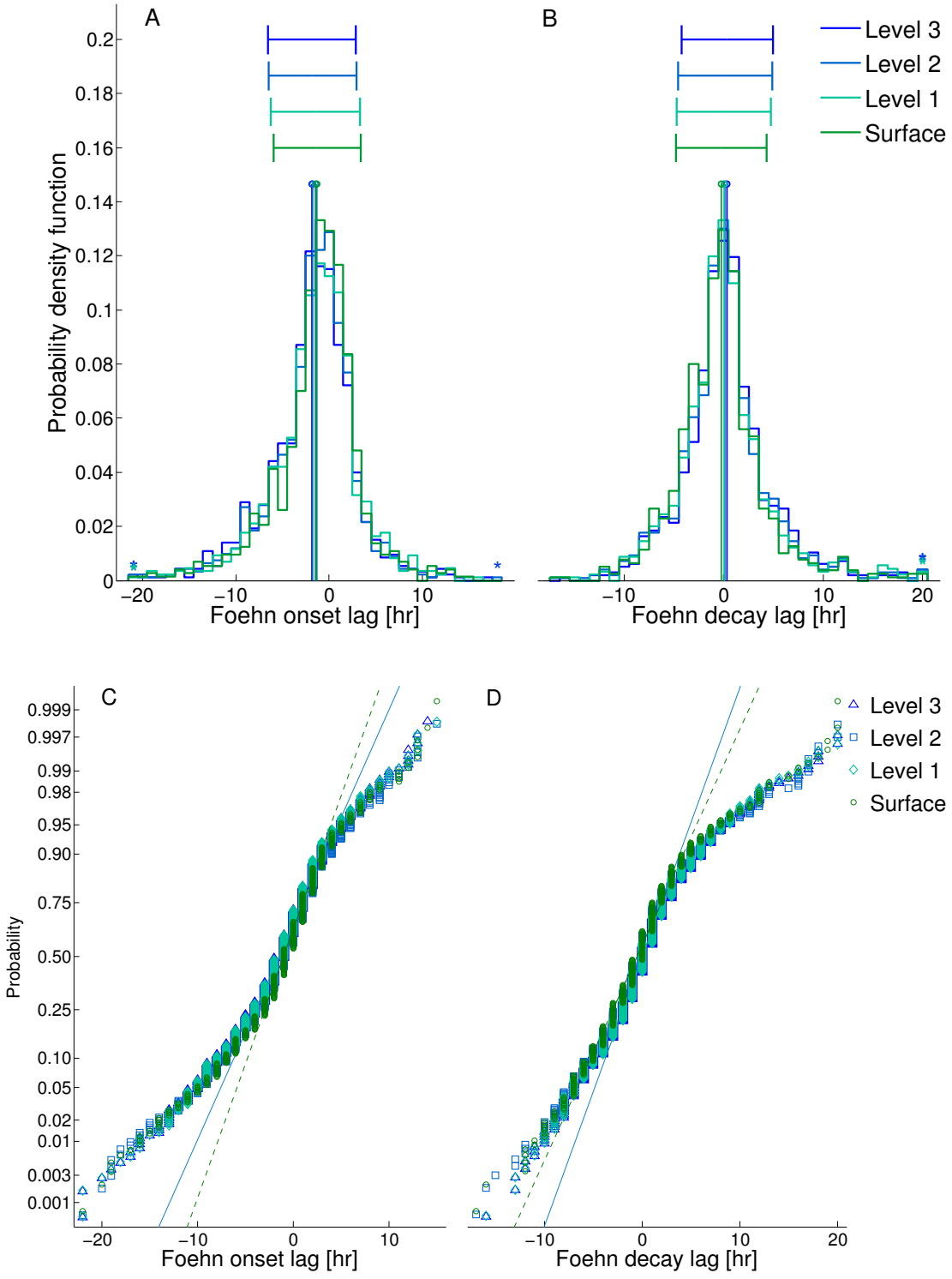

Figure 25: **Top:** combined density distributions of lag times of COSMO-2 foehn onset **(A)** and decay **(B)** for all foehn stations. Horizontal bars show the standard errors about the means, which are marked with a vertical stem. **Bottom:** probability plots of foehn onset **(C)** and decay **(D)** lag times. Normal distributions follow the dashed lines. Model levels are indicated by the colours and shapes listed in the legend.

3.4 Skill optimization

The foehn variable that show the highest skill of all model levels analysed in sections 3.2.2 and 3.3.2 are selected for an "optimized" foehn detection scheme. This scheme consists of the variables with which the foehn was diagnosed: V, V_{max}, ${}^3V_{dir}$, ${}^3RH$ and ${}^3\Delta\theta$. The frequency bias scores of individual foehn stations relative to their modelled altitude are shown in figure 26. The frequency bias scores or the optimized scheme are lower than those of the third model level for every station. However, the discrepancies between the two schemes are different for each station. The frequency bias scores for most stations are closer to the third level than the surface.

Figure 27 shows the TS of all foehn stations combined as a function of foehn intensity. The maximum TS value is 0.33 and occurs when wind threshold of 1.5 kts is applied to the optimized foehn scheme. For this scheme, TS remains roughly constant with increasing wind threshold until approximately 7.5 kts. It then decreases in an roughly an exponential fashion similar to the surface and first model levels until 27 kts. The optimized scheme shows the highest TS of all schemes. However after the 10 kts threshold is reached it is surpassed by the second model level. Subsequently, at 11 kts it is surpassed by the third model level as well.

The onset lag ranges between -22 and 17 hours for all levels. The decay lag ranges between -16 and 21 hours. The mean onset lag (-1.8 hours) lies between the surface and third model level values. The same holds true for the mean decay lag (0.16 hours). The hourly time discretization is apparent in the stepped pattern of the probability plot (figure 28). Both the onset and decay of the optimized scheme deviate from Gaussian distributions strongly in its tails. In general, both tails in the onset lag distribution are equally extended, while the uppermost tail of the decay lag distribution shows the stronger extension. The extensions of both distribution tails for foehn decay lag are greater that those of the surface and third model level. However, both tails of the optimized scheme onset lag fall between those of the surface and third model level.

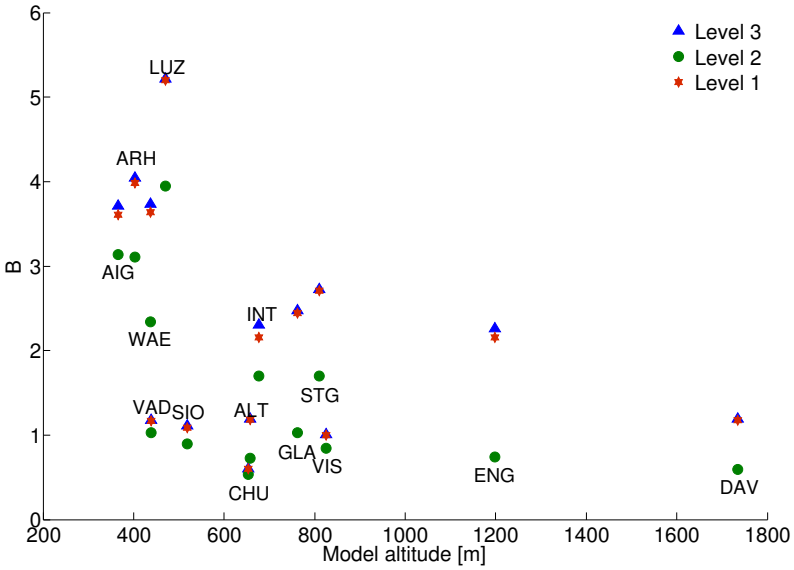

Figure 26: Optimized frequency bias score vs. model altitude for all foehn hours at individual foehn stations. Station names are abbreviated.

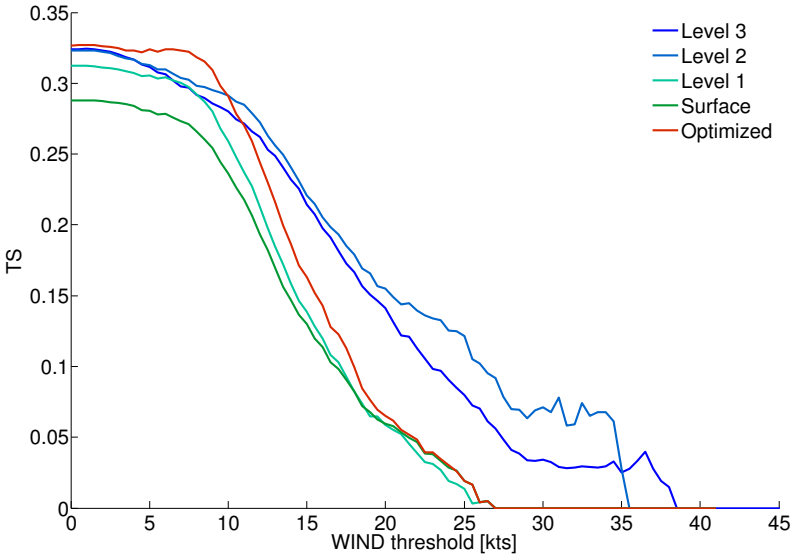

Figure 27: TS of all model levels and the optimized scheme using surface-level wind speed vs. foehn intensity (V [kts]). The intensity thresholds were applied to the model and observational data to calculate the corresponding TS scores. Model levels are indicated by the colours listed in the legend.

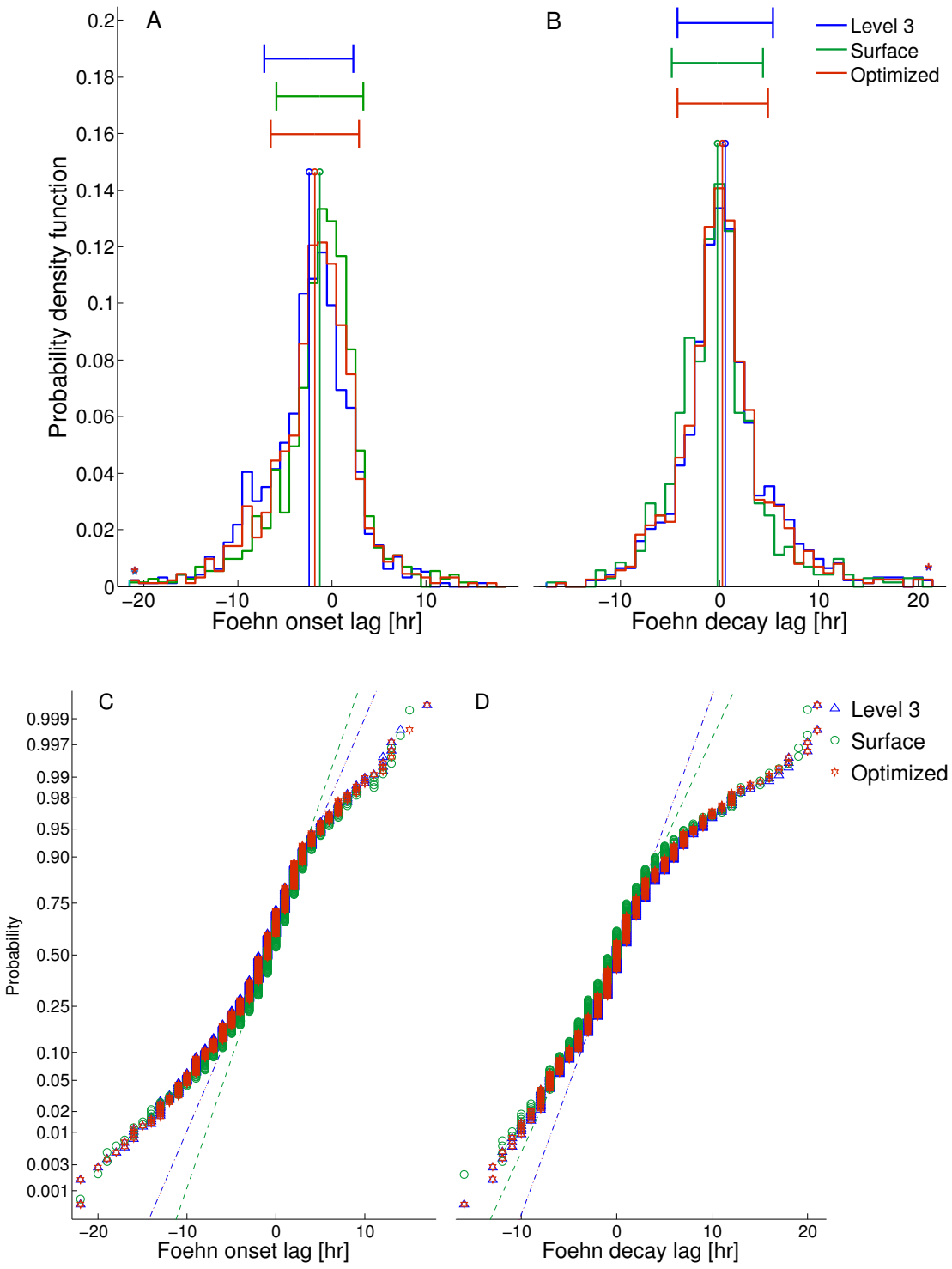

Figure 28: **Top:** combined density distributions of lag times of COSMO-2 foehn onset **(A)** and decay **(B)** for all foehn stations. Horizontal bars show the standard errors about the means, which are marked with a vertical stem. **Bottom:** probability plots of foehn onset **(C)** and decay **(D)** lag times. Normal distributions follow the dashed lines. Model levels are indicated by the colours and shapes listed in the legend.

4 Discussion

4.1 General COSMO-2 foehn detection skill

Only COSMO-2 data for days in which at least 10 minutes of foehn were observed were extracted from the MeteoSwiss archive. Very short instances of foehn, such as this, are often not detected on an hourly time-scale in both modelled and observational time series. Due to this effect, 62 days out of the 478 days in observational time series showed no foehn signal. However, COSMO-2 foehn did occur in 39 of these 62 "foehn free days". This overestimation of foehn frequency is discussed for individual stations in the next section. The sporadic overestimation of foehn days found in COSMO-2 highlights the need for caution when interpreting the categorical verification skill scores that are presented. Since the observational and COSMO-2 data sets used for the model validation did not include all dates, the model skill scores based on these data may indicate erroneously higher skill than if data from all days were analysed.

4.1.1 Foehn frequency

On an hourly time-scale, the bias of modelled foehn frequency varies between foehn stations. The magnitude of the frequency bias roughly corresponds to the topographical surroundings of each station. In this study, stations located in the Alpine regions of Switzerland (ALT, CHU, DAV, SIO, VAD and VIS) tend to have low foehn frequency biases. These stations generally show unbiased (B = 1) foehn frequencies, with the exception of CHU, where it is underestimated by approximately 50%. Fore-Alpine and plateau stations (AIG, ARH, INT, LUZ, STG and WAE) have altitudes comparable to those of the Alpine cluster, yet they exhibit much higher overestimations (B > 1). This clustering occurs at higher model levels as well. Higher model levels produce higher overestimation of foehn frequency than the surface level. The extent of this bias increase is lowest for the Alpine stations. The stations ENG and GLA, located between the Alpine and Fore-Alpine regions show unbiased scores at the surface but transition to higher values with model level height. However, it is important to reiterate that these results do not necessarily imply that foehn detection skill is highest for Alpine stations at the surface model level. This is because the bias score uses the marginals of the observed foehn and modelled foehn, with which no correspondence skill can be deduced (see section 2.5.1 and figure 8).

4.1.2 Foehn detection skill and foehn intensity

The POD was shown to increase with model level height. Unfortunately, the FAR increases as well. As discussed in section 4.1.1, the number of modelled foehn hours increases with model level height. More modelled foehn hours can mean less *misses* and thus greater detection skill. However, if the number of *false alarms* increases as well, this can negate the additional skill. Therefore, both relative increases of POD and FAR need to be assessed together in a single skill score: TS. The TS analysis shows that overall foehn detection skill is greater for the higher model levels. The third model level has the highest skill (TS of 0.324) when all foehn intensities in the

model and the observations (no wind threshold) are considered. This corresponds to a POD of 0.619 and a FAR of 0.595. This skill is quite poor in comparison to the boosting method by Oechslin (2008) which achieved a maximum TS of 0.654 using only 2 years of COSMO-7 reanalysis data. This corresponds to a POD of 0.854 and a FAR of 0.263.

The stratification of categorical skill scores assesses the ability of the model to correctly predict foehn hours with a minimum foehn intensity. Foehn intensity in observations and COSMO-2 was defined by their respective wind speeds (at various model levels in the case of COSMO-2). Wind speed thresholds were reported in knots to allow for easy comparison to conventional wind storm warning thresholds. Wind speed was chosen over gust speed as a measure of foehn intensity because it is explicitly modelled in COSMO-2 and is therefore available at each model level. Interestingly, the skill of the third level decreases with foehn intensity faster than at the second level, and is eventually surpassed by it. Low wind speeds tend to be overestimated at the second model level and to a slightly larger extent at the third model level. Unlike at the second, this overestimation continues for midrange wind speeds at the third model level (figure 32 in appendix D). This likely contributes to the faster degradation of the forecast of the detection skill with increasing foehn intensity at the third model level. If wind speeds are overestimated, a more rapid increase in FAR would occur.

Skill scores could only be calculated for foehn intensities less than 38 kts (19.5 m/s), due to low sample sizes. Unfortunately, there is little demand for accurate foehn forecasts at such low foehn intensities since there would be negligible wind damages incurred. However, observed foehn gust speed tend to be roughly 2 times faster than wind speed over the period analysed (figure 30 in appendix A). Using this ratio, a rough *ad hoc* conversion of the wind thresholds to gust speed can be done. Doing so yields TS of the values for the second model level of 0.31, 0.26, 0.20 and 0.14 for 10 kts, 25 kts 33 kts and 42 kts, respectively. The optimization of the detection skill for both the overall maximum and the strong foehn events is discussed section 4.2.

4.2 Performance of COSMO-2 during observed foehn

4.2.1 Bias trends

A strong linear relationship can be detected between the bias of wind speed and the station distance from the MAR. This relationship holds true for all model levels, with the second model level showing the least amount of scatter about the trend line. Alpine stations that are located closer to the MAR experience lower wind speeds during foehn events in COSMO-2. This may be caused by the smoothed topography in the model domain. Mountain passes and valley transects tend to be less narrow in COSMO-2 and therefore the extent of the vertical constriction of the foehn flow would be smaller. However, the tendency for faster wind speeds at stations located far from the MAR on the Swiss plateau cannot be explained by this, but rather due to the foehn front extending north earlier and then retreating south later than observed. This would subject the plateau and fore-Alpine stations to prolonged periods of gusty conditions. Another possibility is that the roughness lengths of the surface are too large in the Alpine region and too small in the Swiss

plateau region.

Overall, the Alpine stations ALT, CHU, DAV, ENG, and GLA show weaker descent of isentropes in the model than observed ($\Delta\theta$ around $-1.5\,\mathrm{K}$) and around 25% higher humidity. Exceptions to this clustering are SIO and VIS, which tend to behave more like the fore-Alpine/plateau stations. Stations in the fore-Alpine and Swiss plateau regions as well as VAD, SIO and VIS tend to be less humid than the Alpine stations, yet approximately 14% higher than reality. The fore-Alpine stations as well as VIS tend to have little to no bias in isentropic descent, while the plateau stations and SIO show slightly weaker descent of isentropes.

$\Delta\theta$ was calculated from two *random variables*: θ at a foehn station and at Gütsch. However, these two temperatures are neither *independent* nor *uncorrelated* to each other. Air temperature and other variables at a foehn station are influenced by those at Gütsch during foehn, since the foehn flow originates from the MAR nearby Gütsch. Thus, if air flowing past Gütsch is anomalously cold or warm this will also be reflected at the foehn stations. Moreover, according to the principles of random variable algebra, the difference between these two variables will not be normally distributed, since they are not independent and uncorrelated. Therefore, one cannot simply subtract the θ bias of Gütsch from those of the foehn stations to yield the bias in $\Delta\theta$.

When the $\Delta\theta$ bias is calculated properly, i.e. directly from $\Delta\theta$ values, several stations shift in relation to one another compared to the biases in θ. This means that some stations with negative θ biases have positive $\Delta\theta$ biases (AIG, INT, LUZ) and vice versa. Stations that experience higher foehn frequencies, such as DAV and ALT, also change positions relative to one another. Therefore, there is little influence of sample size on these shifts. Since these stations have negligible altitude errors, these biases are more likely caused by other factors, such as: excessive upstream entrainment of air high in θ from aloft, super-adiabatic lapse rates from excessive radiative heating, or higher passes in the MAR that force the air parcels to descend from higher altitudes.

Apart from the surface levels at AIG, INT and LUZ and the third level at WAE, all biases in $\Delta\theta$ were found to be negative. These discrepancies could be due to a combination of several factors, concerning influences of local topography: the contribution of cold air mixing from side valley tributaries could be larger in the model; the air at the station may originate from lower heights than in reality; the coupling of the soil-air temperature is too strong at the foehn stations. Many side valleys in the model topography are simply missing because of the coarse resolution. In addition, Alpine passes through the MAR are smoothed and can force air parcels from higher altitudes to descend.

The strong air-soil temperature coupling may lead to a negative systematic bias of θ at the station Gütsch and may also occur to a greater extent at a given foehn station a negative $\Delta\theta$ bias. The result would show a more stably stratified atmosphere. This is because sensible and latent heat fluxes determined by the multilayer soil module in COSMO-2 are sensitive to the soil type (ice, rock and peat) and texture (sand, sandy loam, loam, loamy clay and clay) which are inhomogeneously distributed over the model domain. The grid points chosen by the interpolation process may be ideal regarding topography, wind fields and other parameters. However,

if the soil type is not representative for that particular station it can cause a temperature bias and a loss of foehn detection skill. This bias would lessen with model level height as the air-soil coupling becomes weaker. Stations that show little bias improvement with increasing model level height (CHU, INT and SIO) may have more appropriate model soil types than other stations. Unfortunately, the analysis of the land surface-atmosphere temperature coupling was beyond the scope of this thesis and was therefore not persued in more detail.

Incorrect paramaterization of PBL processes in Alpine regions can also suppress the descent of isentropes (Zaengl et al., 2008). If precipitation is not well parameterized the liquid water content can be too high as the foehn flows over the MAR. Consequently, additional energy would be required to first contribute to evaporation before it can lead to warming of the descending air. This would cause a strong negative bias in $\Delta\theta$ for stations close to the MAR. Positive station altitude errors would have the same effect. Temperatures decrease with altitude, which leads to a lower average RH, though the water content remains the same. The analysis conducted in this thesis shows that there is a strong positive correlation of RH bias with foehn station altitude error (figure 23), not with distance from the MAR. Therefore, the topographical resolution (altitude error) seems to hinder COSMO-2 RH skill to a greater extent than PBL parameterization. It may be possible to correct for this, given the vertical temperature profile. However, only the three lowest model levels of data were available for COSMO-2, which is insufficient to correct for such large deviations in station altitude.

For most stations, the general orientation of the surrounding valley walls are well represented in COSMO-2. However, the station at SIO shows an unexpectedly high bias in wind direction. It is situated at the wide, well resolved section of the zonally oriented Rhône valley with no large deviation in local topography. Wind direction distributions at SIO show a roughly up-valley/down-valley bimodal distribution. The larger of the modes is down-valley and likely corresponds to foehn events. However, the presence of the opposing mode of wind direction shows that foehn did not always occur at thelocation where it was observed. Upon inspection of the grid point choice, the station location is shifted down the valley by approximately 2 km (figure **??** in appendix C). It is possible that foehn does not easily breakthrough this far down the valley, however the exact reason for this bias is unclear.

4.2.2 MAE and correlation

Amplitude and phase errors are two aspects of model-observation agreement that are relevant to model skill. While the bias may sometimes give insight into the general direction of model errors (RH), competing negative and positive errors can cancel each other out and yield no useful model skill information (^2V). Instead of the bias, the mean amplitude error of the model is measured by the MAE, which takes the distribution of bias errors into account. The extent of phase errors (linear correspondence) is quantified by the correlation coefficient. Inter-comparison of MAE from different variables is difficult since it is not a dimensionless quantity. To circumvent this problem, the MAEs were normalized to the observed standard deviations of their respective variables. Minimization of the MAEs and maximization of the correlation coefficients are both desirable to optimize the model skill (see section

4.5).

In terms of error amplitude, the poorest modelled variable is RH at all levels, despite improvements in performance with mode level height (approximately 70%). $\Delta\theta$ is also poorly modelled, particularly at the surface. Indeed, for all variables except wind speed and direction, which both remain approximately constant, the largest increases in model error (amplitude and phase) are seen when in transition from the surface to the first model level. Since wind gusts were assumed to be constant with model height, the jump in maximum TS between the surface and first model level must be due to the increase in RH and $\Delta\theta$ skill. A smaller increase in TS from the second to the third model levels occurs despite the largest decrease in wind speed skill. This indicates that foehn detection skill using this threshold method is most sensitive to changes in $\Delta\theta$ and RH performance. In light of this, the poor detection skill, further work to reproduce the method by Duerr (2008) with COSMO-2 data would be beneficial. This may yield internally consistent foehn thresholds that are tuned to COSMO-2 topography deviations, and PBL parameterizations effects.

4.3 Comparison to seasonal COSMO-2 performance

The performance of COSMO-2 wind speed is significantly reduced during the observed foehn events. The overall MAE of wind speed during observed foehn is 4.38 m/s at the surface. This represents the best performance of mean foehn wind speed, yet it is still 2.5 times greater than the climatological MAE of COSMO-2 over Switzerland (1.74 m/s). In addition, the seasonal variation of foehn MAE does not correspond well with the regular COSMO-2 performance at all the model levels (figure 29). Foehn wind speed performs poorly during the winter and autumn months and well as during the summer months at the surface and first model levels. The seasonal pattern of wind speed MAE is less clear at higher model levels. Contrarily, the model performance of foehn temperatures follows a similar seasonal pattern at all model levels. The overall MAE of temperature is 4.2 K - approximately 2.4 times greater than the climatological MAE of COSMO-2 over Switzerland.

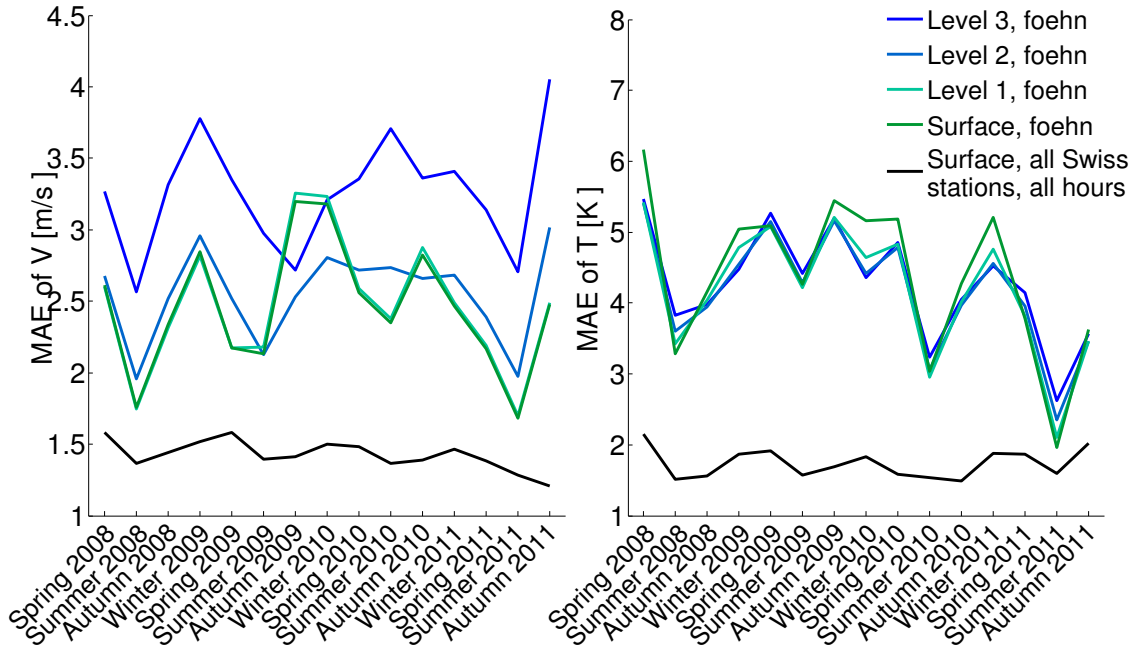

Figure 29: Seasonal MAE.

4.4 Foehn onset and decay lag

The spurious nature of foehn detection in COSMO-2 made the temporal comparison between the model and observations less than straight-forward. Observed foehn events (a foehn event being a series of consecutive foehn hours) were detected in COSMO-2 as a string of non-continuous "mini-foehn events". As illustrated earlier by the high bias scores, COSMO-2 tends to overestimate the frequency of foehn. This overestimation is not necessarily due to excessively long foehn events, but rather from the occurrence of foehn on days when none were observed. In addition, some foehn days were not reproduced in the model - particularly during the winter months. These foehn events were discarded from the lag analysis. Thus, only 728 samples of foehn onset and decay were obtained for analysis at the surface level, while the higher model levels showed less overestimation of foehn with increasing mode levels (maximum sample size of 929 at the third level).

In general, foehn onset tends to occur too early, while foehn decay tends to be too late. Foehn onset is earlier with higher model levels (-1.28 to -1.88 hours). Contrarily, foehn decay is delayed longer with higher model levels. Foehn diagnosis at the surface model level actually results in a large negative decay lag, which shifts to a slightly positive lag of -0.25 to 0.35 hours for the first model level. It is important to note that the distributions of both onset and decay lag are quite wide and contain compensating negative and positive values. Based on the skill assessment of each level thus far, this shows lower overall error amplitude for most foehn variables at higher model levels. However, the correlation of wind speed decreases with model level height. This could contribute to phase errors in the foehn detection, leading to apparent shifts in onset and decay that would be difficult to predict. The lag time distributions for individual foehn stations are shown in figure 34 in appendix F.

4.5 Optimized foehn scheme

The first optimized foehn detection scheme combines all of the foehn variables with the highest model performance: V, ${}^3V_{dir}$, ${}^3RH$ and ${}^3\Delta\theta$. This strategy was chosen in order to maximize the absolute TS maximum. For most variables, the choice of the most suitable model level was clear because both the maximum correlation coefficient and the minimum MAE occur at the same mode level. However, for wind speed, the surface model level shows the lowest error amplitude and the first model level has the highest phase error. Trial and error showed slightly higher maximum TS value using surface-level wind speeds. TS was maximized to a value of 0.33 when no wind threshold applied. However, the skill quickly decreases with stronger foehn intensities similar to the surface level skill, which reduces its usefulness for a forecast warning system. The optimized scheme also showed a decrease in foehn onset and decay lag - falling between the surface and the first model layer. In order to shift the optimization to higher foehn intensities the skill at low foehn intensities must be sacrificed slightly. ${}^2V$ is the ideal choice for second optimization scheme because it has a lower conditional bias than V, since high wind speeds will be underestimated to a lesser extent. The resulting distribution can be seen in 33 in appendix E. The mean onset and decay lags of the second optimization scheme improved by a smaller margine - falling between those of the third and second model levels.

5 Summary and Outlook

Various model verifications were done to assess the overall foehn detection skill of COSMO-2 reanalysis data. The analysis showed that the foehn frequency is estimated best at Alpine stations and is overestimated in fore-Alpine and Swiss plateau stations. The frequency overestimation is highest for higher model levels, which results in a higher FAR. POD increases to a larger extent with higher model levels than FAR, which yields an overall increase in foehn detection.

Stratification of the detection skill by wind speed was done to assess foehn detection at different minimum foehn intensity levels. The foehn detection skill was found to decrease the slowest with foehn intensity when using the second model level. The detection skill at lower foehn intensities was improved to an overall maximum by using V instead of ${}^3V$ for the third level scheme. The detection skill at higher foehn intensities was improved slightly by using ${}^2V$ instead of ${}^3V$ for the third level scheme. However, even when all foehn hours are considered, this method shows less skill than the boosting method by Oechslin (2008). This is particularly underwhelming considering that the boosting method was performed with COSMO-7 which has a coarser lower horizontal resolution and less sophisticated PBL parameterizations.

Temporal lag errors of foehn onset and decay were found to have a wide range of positive and negative values alike. The onset lag tends to become more negative with model level height. Except for the surface, which shows a negative mean lag, higher model levels show positive mean decay lags increasing with model height. The lags of onset and decay improved slightly for both optimization schemes - with both falling between the third and second model levels.

Classical verifications during observed foehn hours were performed over the 4 year period. Overall errors in V and T during foehn are approximately 2.5 times greater in magnitude than the climatological errors over Switzerland. The errors decrease with model level height, though on a seasonal basis, this is not always the case. For instance, during the seasons of autumn 2009 and winter 2010 the surface and first model levels surpassed the other levels with respect to MAE of V.

A strong linear relationship between the wind bias and the distance from the MAR was found. Wind speeds are underestimated close to the MAR and overestimated further away. The presence of a linear relationship between RH bias and altitude error rather than with distance from the MAR suggests that topographical smoothing influences RH to a greater extent than inaccuracies in the PBL parameterization scheme. Since RH shows the worst performance of all variables, further work to correct these biases may benefit the foehn detection skill. Account of topographic changes in valley orientation and upstream effects could be done in the form of COSMO-2 tuned foehn thresholds, which may improve the detection skill.

When interpreting the results of this model verification, it is important to note that the analysis was performed on COSMO-2 reanalysis data. In addition, observational and COSMO-2 data sets used did not include all days throughout the analysis period. Therefore, the verification scores and statistics based on these data may indicate erroneously higher skill than if forecast data from all days were analysed.

Appendix A

Table 3: List of foehn relevant SwissMetNet stations names, abbreviations, geographical locations and altitudes. (MeteoSwiss, 2010)

Station	**Abbrv.**	**Latitude [°]**	**Longitude [°]**	**Altitude [m]**
Gütsch ob Andermatt	GUE	46.653	8.616	2287
Aigle	AIG	46.326	6.921	381
Altdorf	ALT	46.870	8.632	449
Altenrhein	ARH	47.484	9.566	398
Chur	CHU	46.870	9.530	556
Davos	DAV	46.813	9.844	1590
Engelberg	ENG	46.822	8.411	1035
Glarus	GLA	47.035	9.068	515
Güttingen	GUT	47.601	9.279	440
Hörnli	HOE	47.371	8.941	1144
Interlaken	INT	46.672	7.871	580
Luzern	LUZ	47.036	8.301	456
Sion	SIO	46.219	7.338	482
St. Gallen	STG	47.426	9.399	779
Vaduz	VAD	47.128	9.518	460
Visp	VIS	46.303	7.843	640
Wädenswil	WAE	47.221	8.677	463
Zürich-Fluntern	SMA	47.378	8.566	556
Magadino-Cadenazzo	MAG	46.160	8.933	203
Zürich-Kloten	KLO	47.479	8.536	436

Appendix B

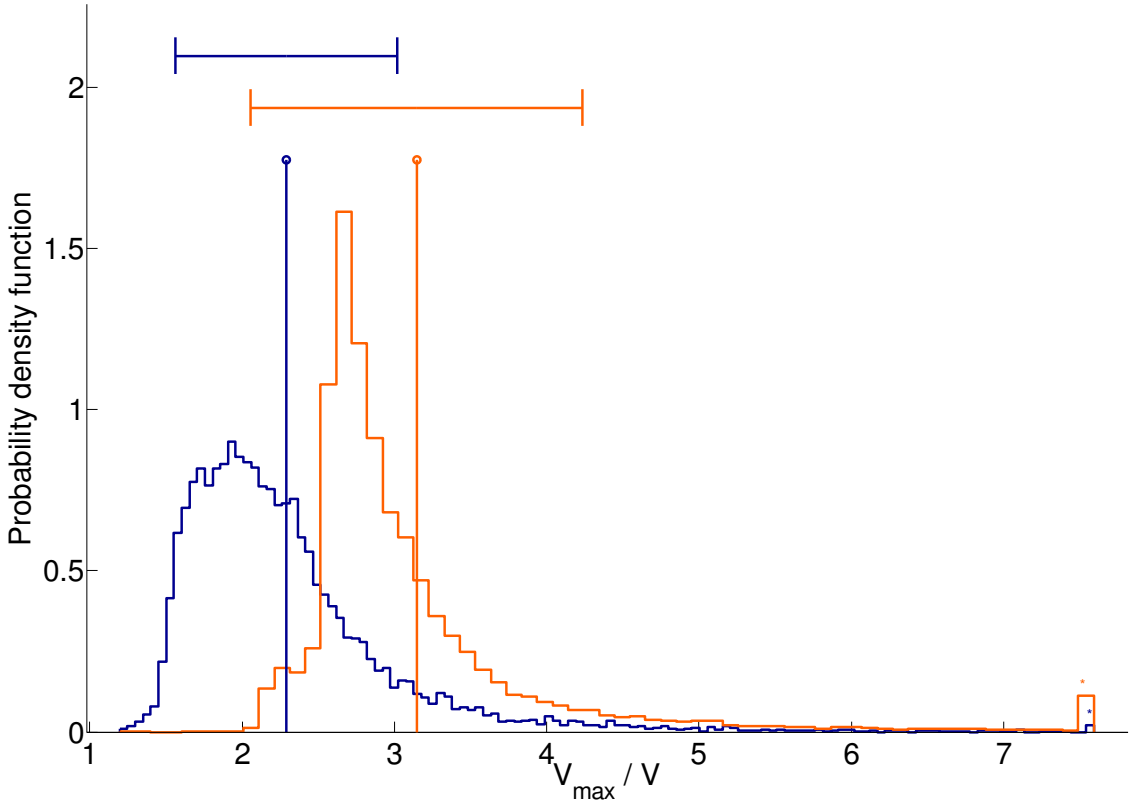

Figure 30: Ratio of gust to wind speed. Observations are shown in blue and COSMO-2 data are shown in orange. Horizontal bars show the standard errors about the means. The mean values are marked by stems.

Appendix C

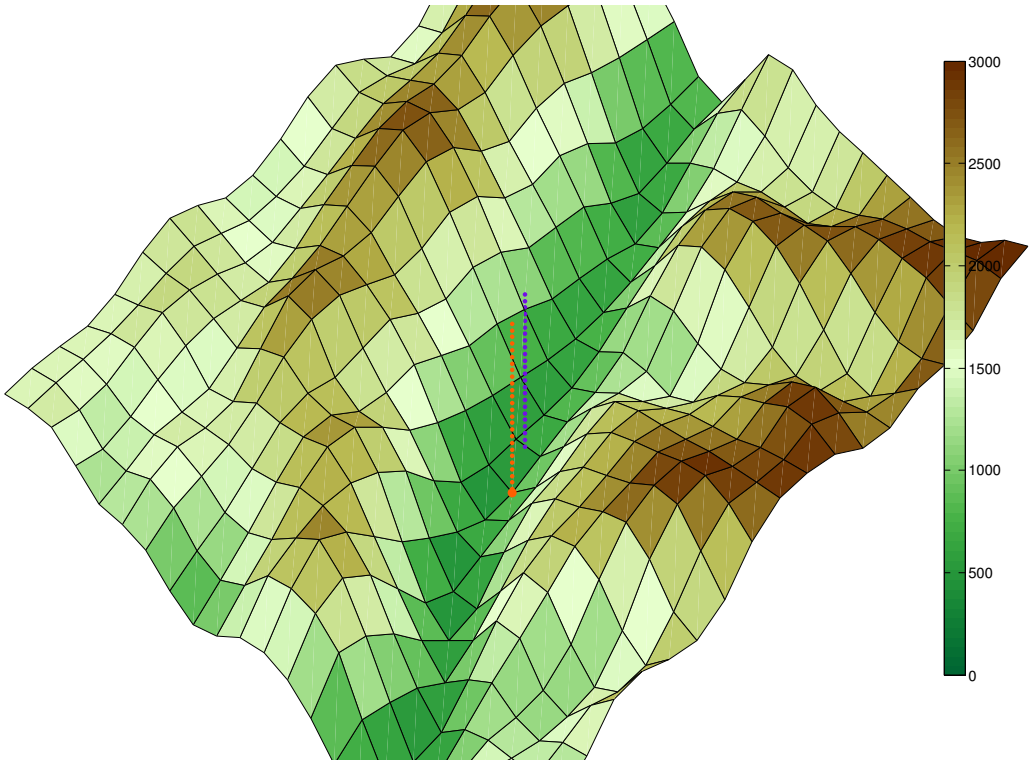

Figure 31: COSMO-2 topography of SwissMetNet station near Sion (SIO) looking to the north-east. Actual location is indicated in purple and the location of the interpolated COSMO-2 grip point is indicated in orange. The colour bar shows the topography altitude in meters above sea level.

Appendix D

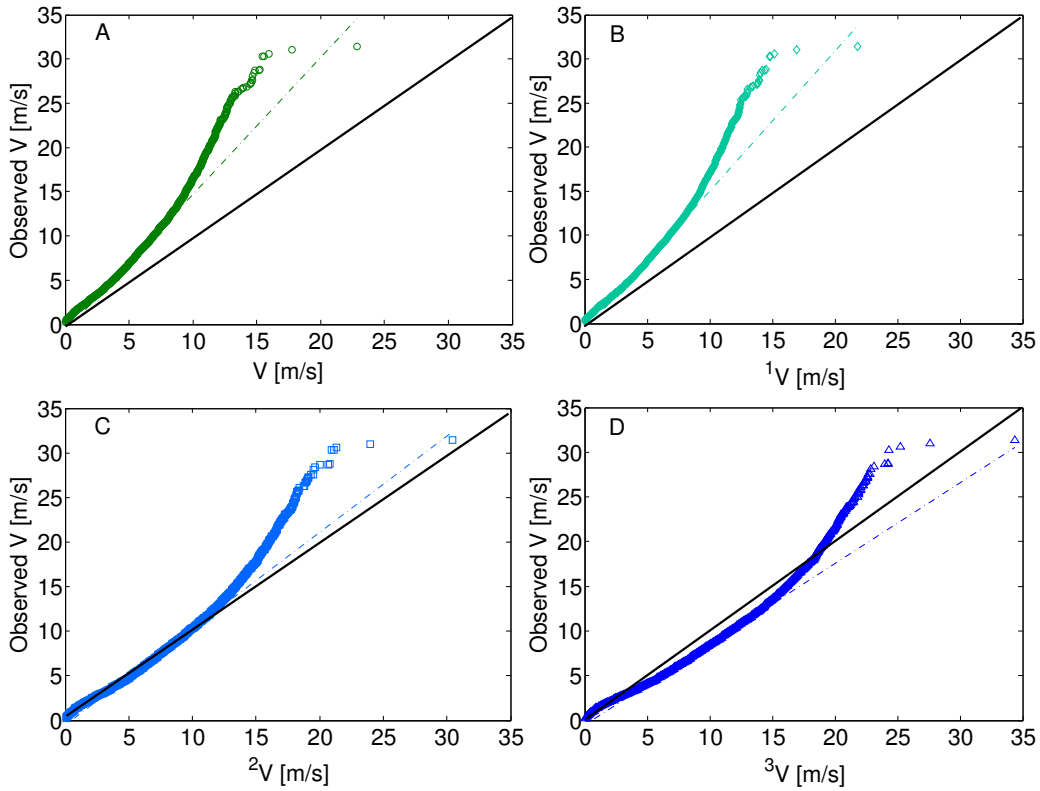

Figure 32: Q-Q plots of COSMO-2 wind speeds at the surface **(A)**, level 1 **(B)**, level 2 **(C)** and level 3 **(D)**.

Appendix E

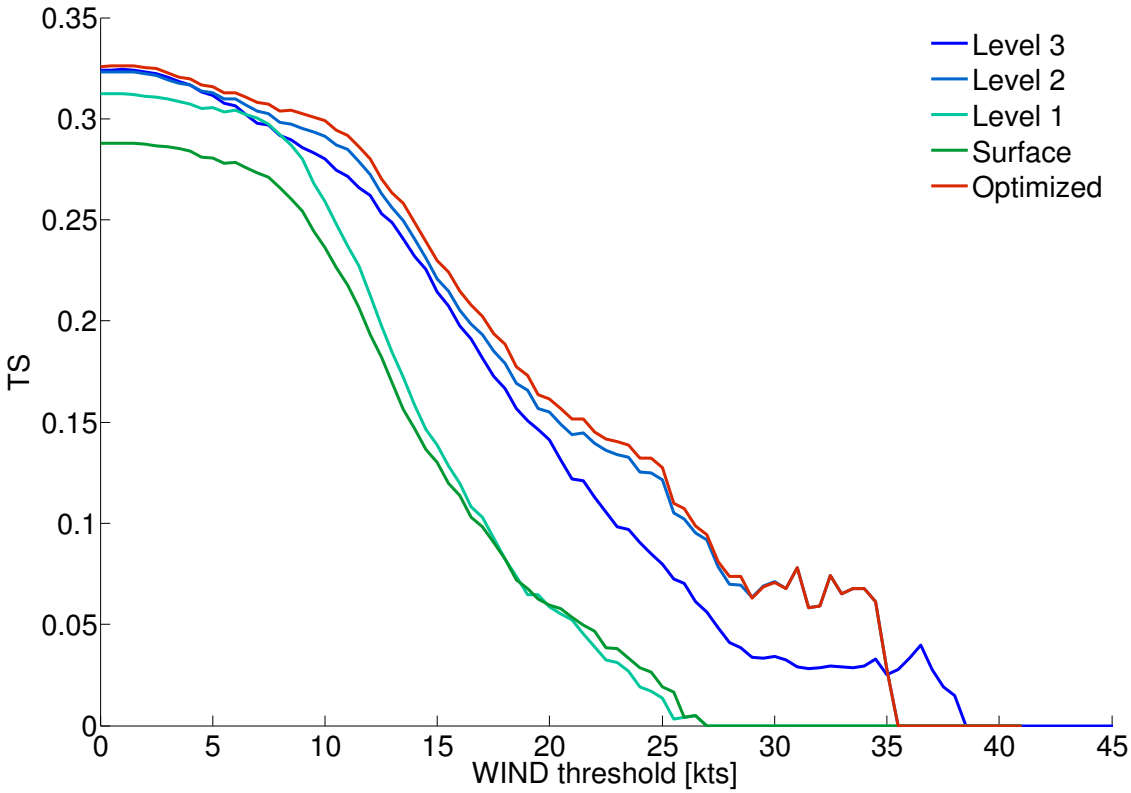

Figure 33: Optimization scheme using ${}^2V$ with the other third level variables.

Appendix F

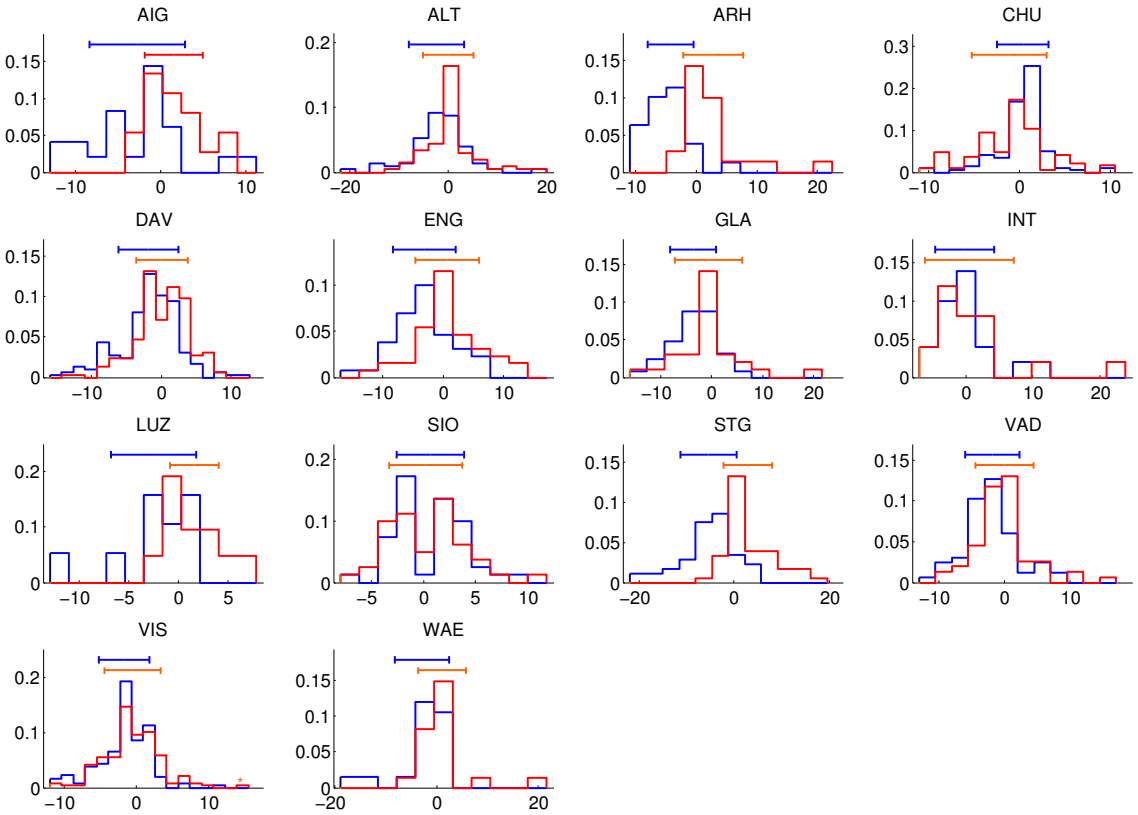

Figure 34: Density distributions of lag of COSMO-2 foehn onset and decay, diagnosed with surface-level foehn variables. COSMO-2 foehn hours are diagnosed from surface-level foehn variables. Foehn onset lags are shown in light blue and foehn decay lags are shown in red. Horizontal bars show the standard errors about the means, which are marked with a vertical stem.

Appendix G

Table 4: List of major mountain peaks and passes and their geogaphic coordinates used to approximate the location of the MAR

Peaks/Pass name	**Latitude** [$^\circ$]	**Longitude** [$^\circ$]
Mont Blanc	45.83361111	6.86500000
Col du Grand-Saint-Bernard	45.86888889	7.17055556
Tête Blanche	45.98750000	7.57500000
Pointe Dufour	45.93694444	7.86694444
Weissmies	46.12777778	8.01194444
Col du Simplon	46.25166667	8.03333333
Passo della Novena	46.47805556	8.39305556
Passo del San Gottardo	46.55916667	8.56138889
Passo del Lucomagno	46.56277778	8.80083333
Piz Terri	46.60000000	9.03388889
Passo del San Bernardino	46.49611111	9.17083333
Passo del Maloja	46.40100000	9.69500000
Piz Bernina	46.38222222	9.90805556
Pass dal Fuora	46.64138889	10.29333333
Passo di Resia	46.84805556	10.50500000

Appendix H - from presentation

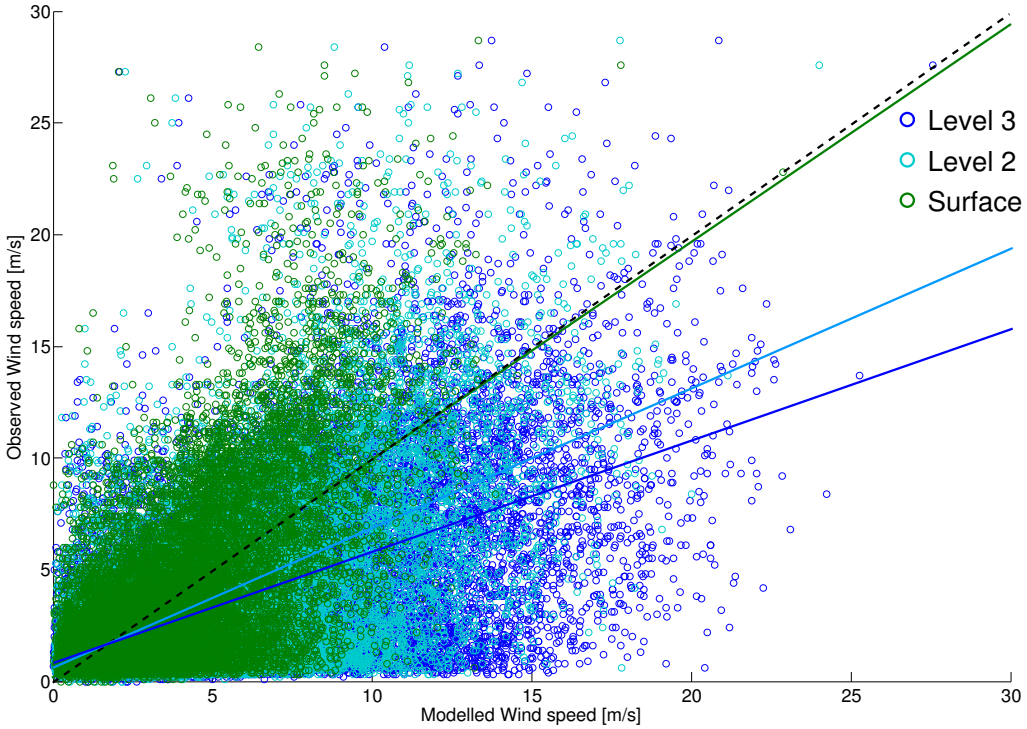

Figure 35: Conditional biases of COSMO-2 wind speed.

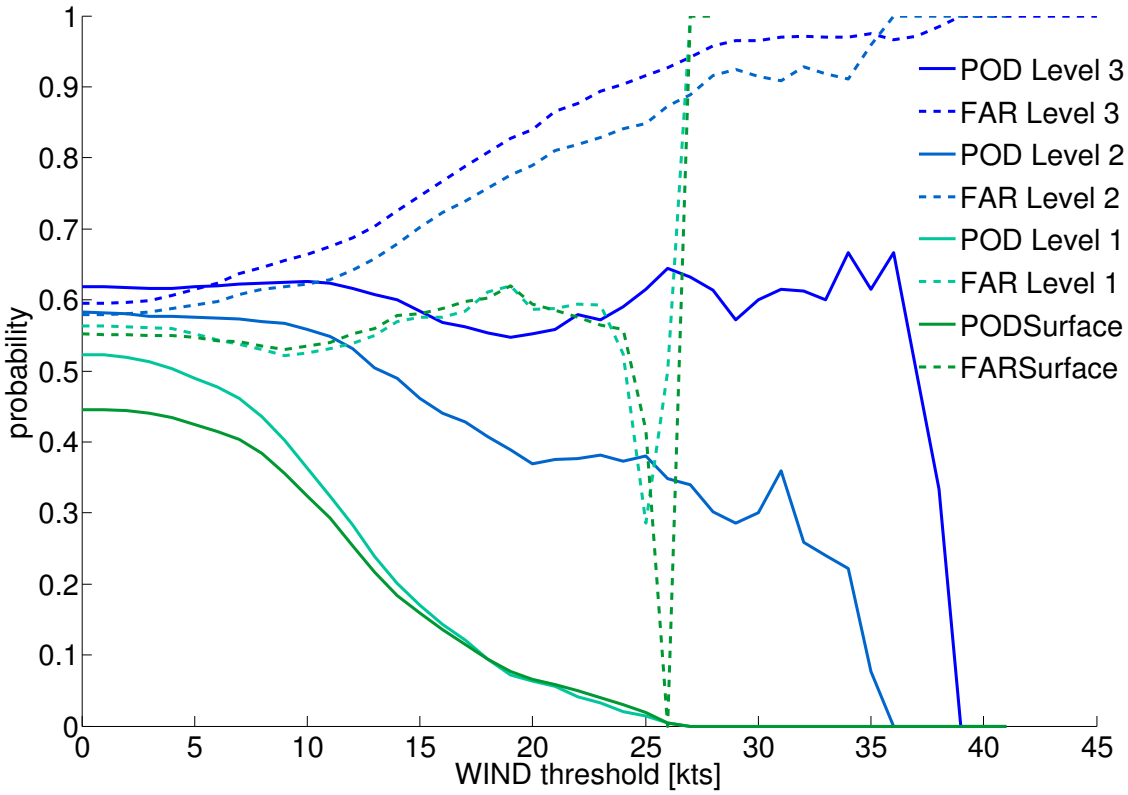

Figure 36: POD and FAR of all foehn station as a function of foehn intensity.

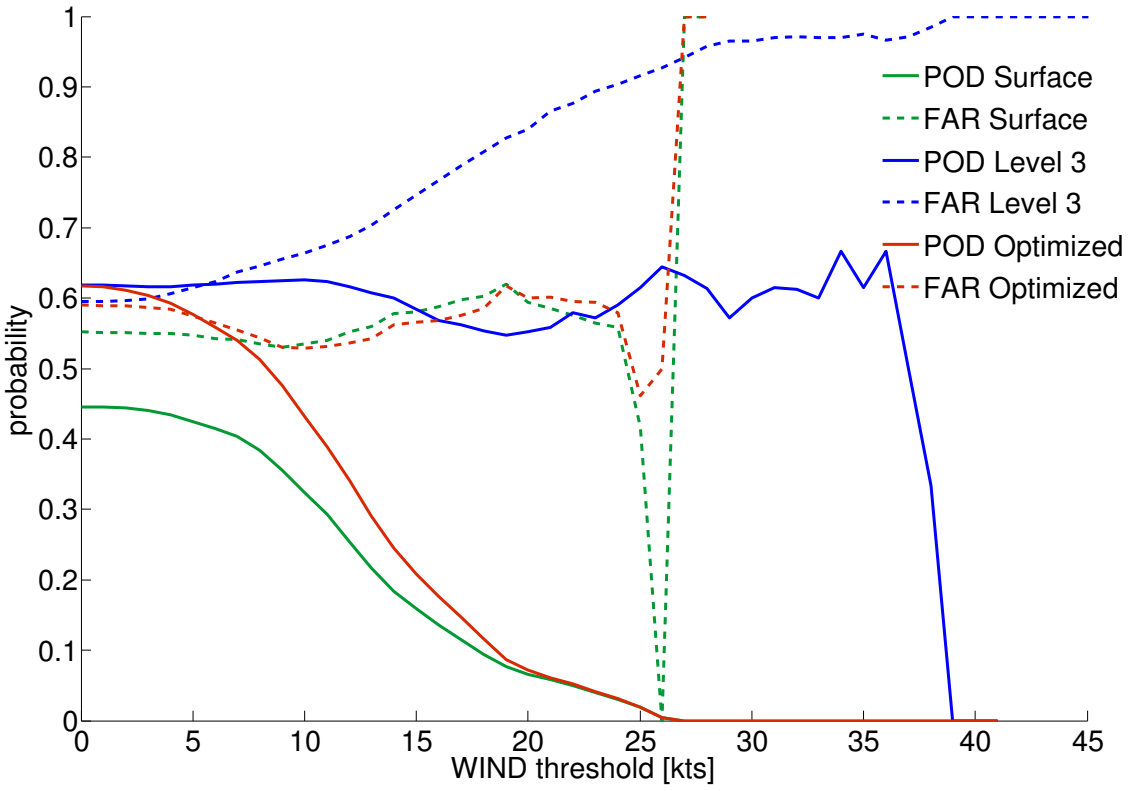

Figure 37: POD and FAR of all foehn station as a function of foehn intensity.

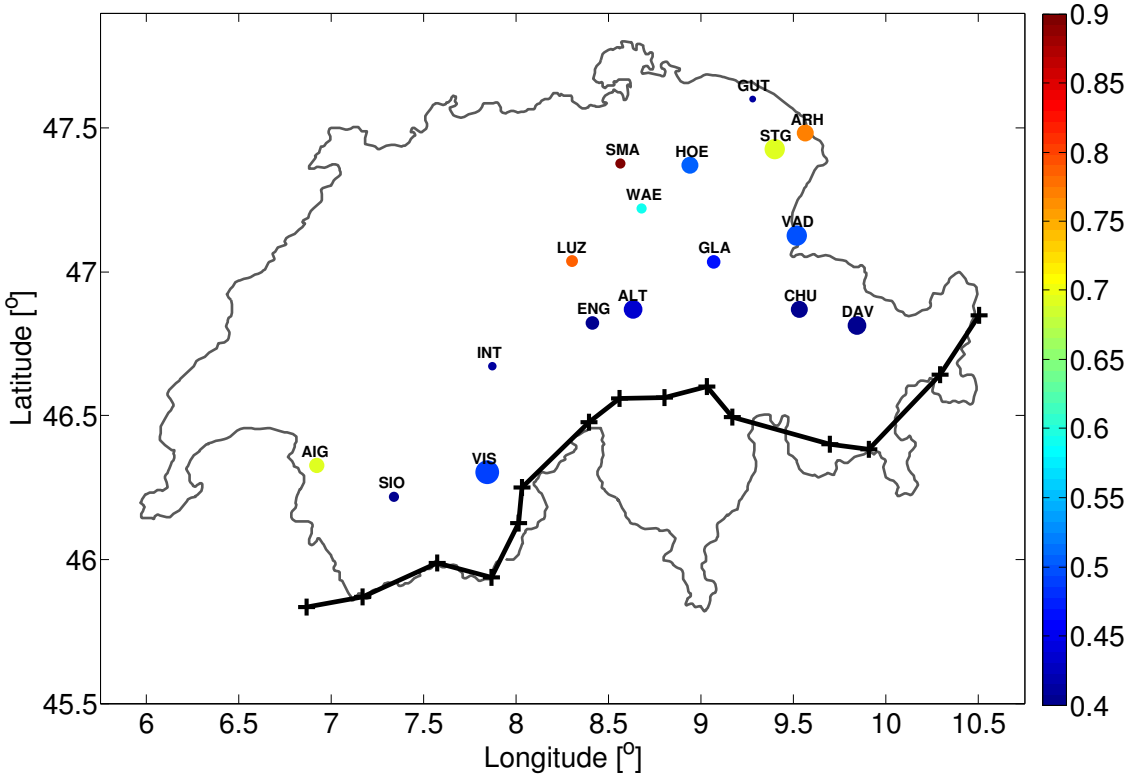

Figure 38: POD of each foehn station using surface variables.

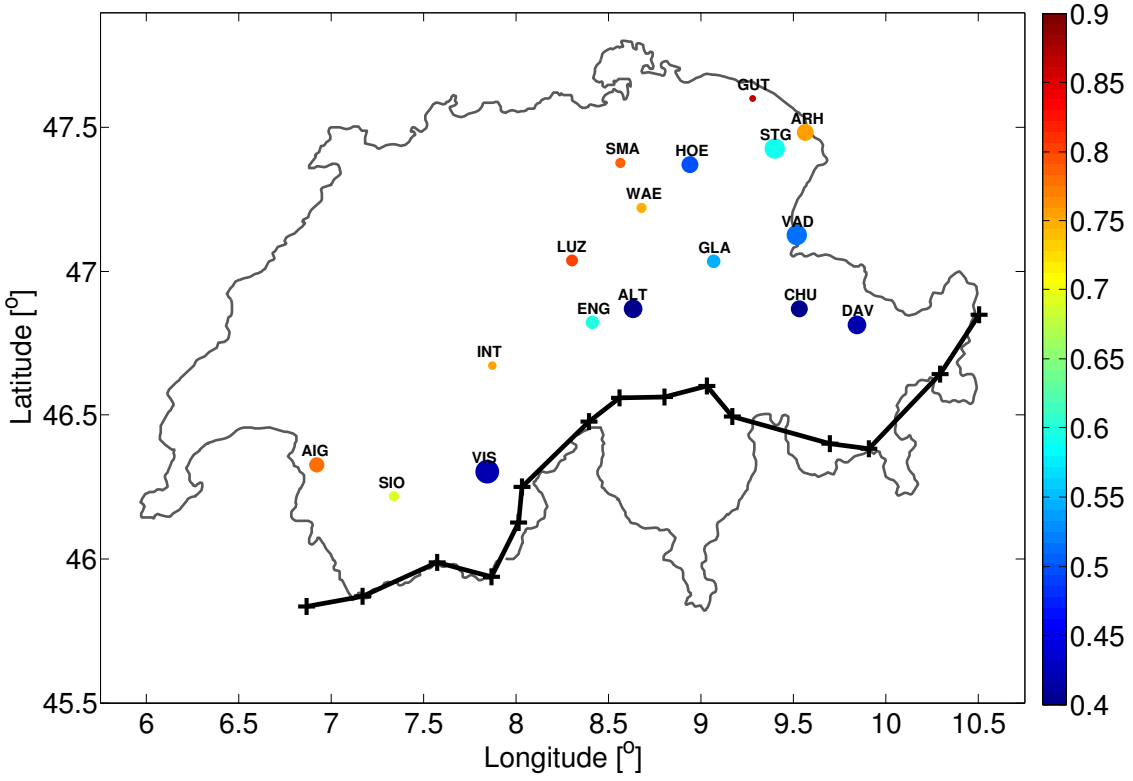

Figure 39: FAR of each foehn station using surface variables.

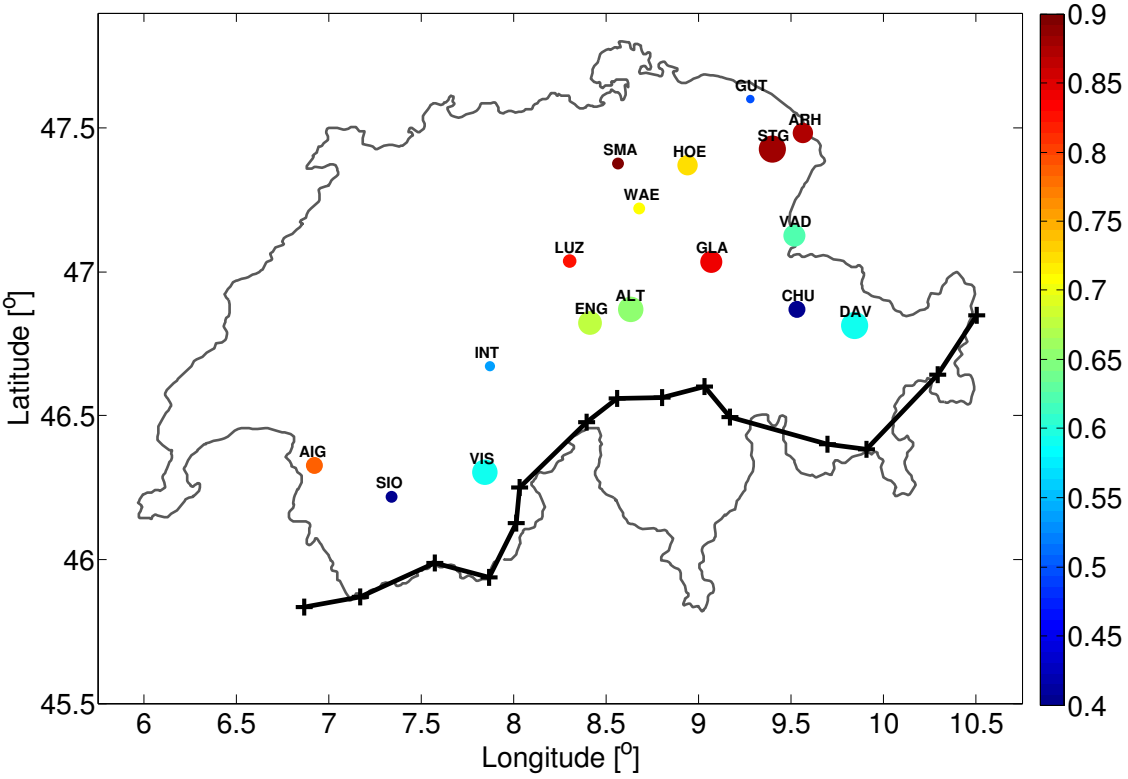

Figure 40: POD of each foehn station using optimization scheme 1 variables.

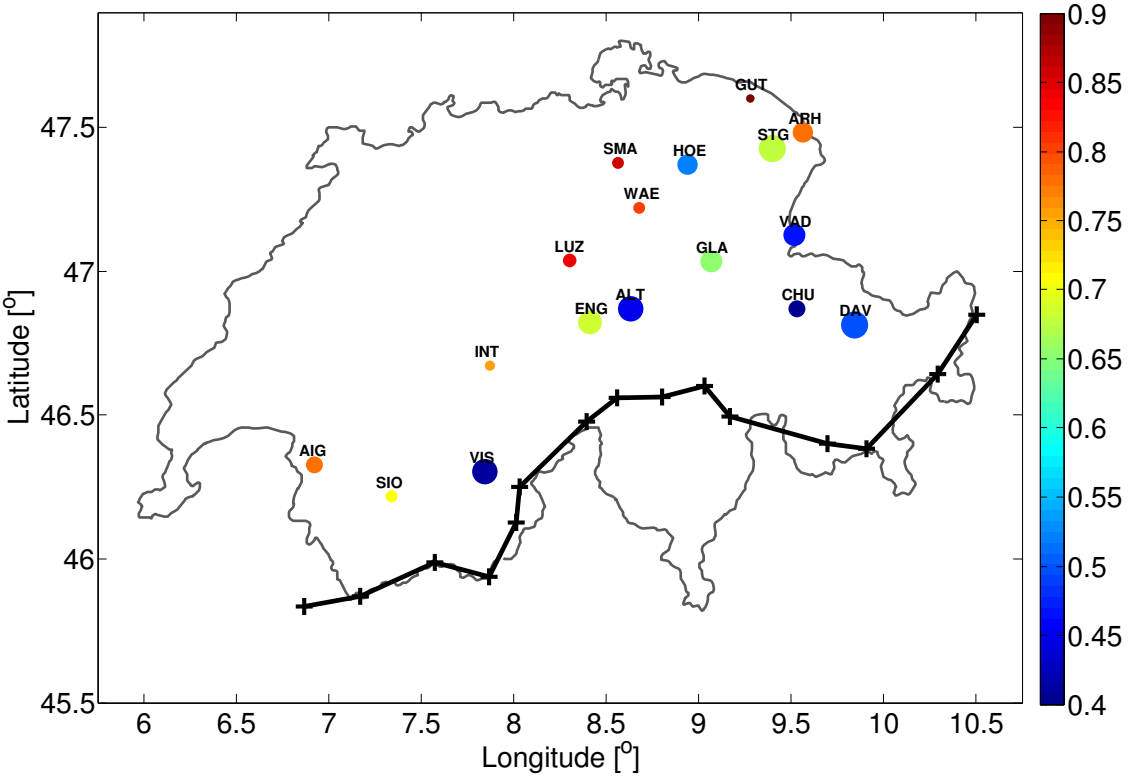

Figure 41: FAR of each foehn station using optimization scheme 1 variables.

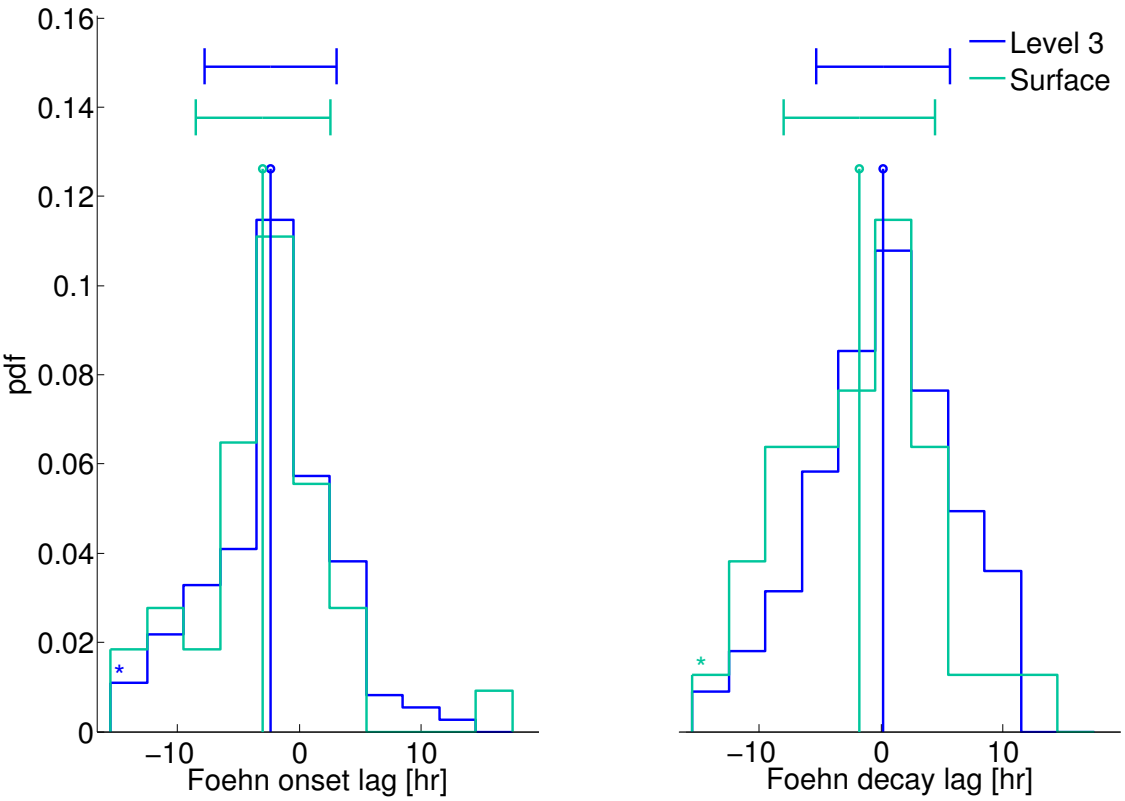

Figure 42: Foehn lag during winter months (DJF).

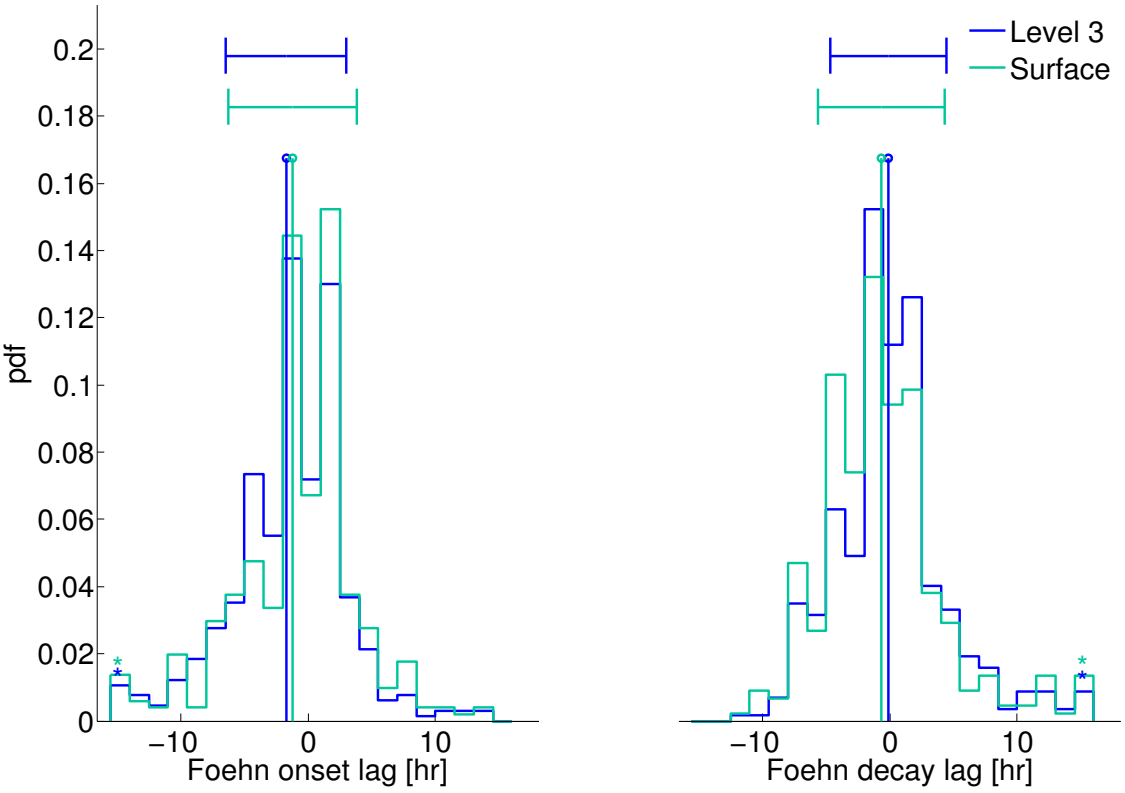

Figure 43: Foehn lag during spring months (MAM).

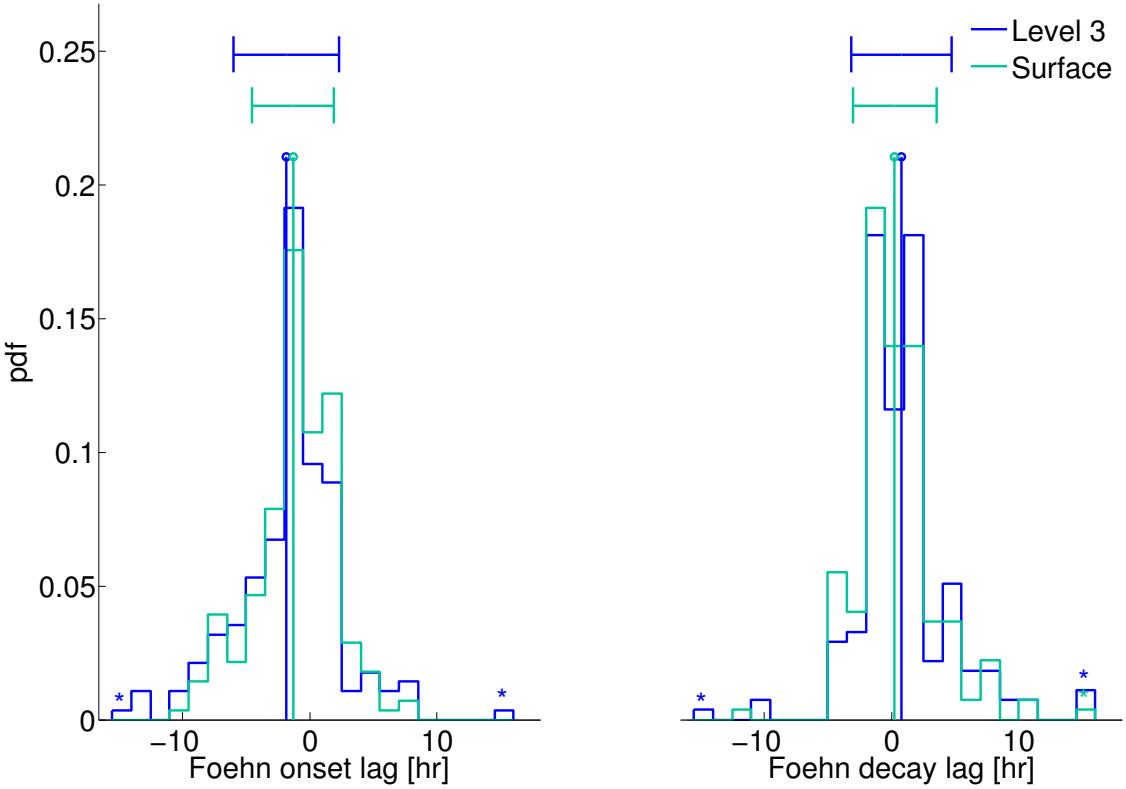

Figure 44: Foehn lag during spring months (JJA).

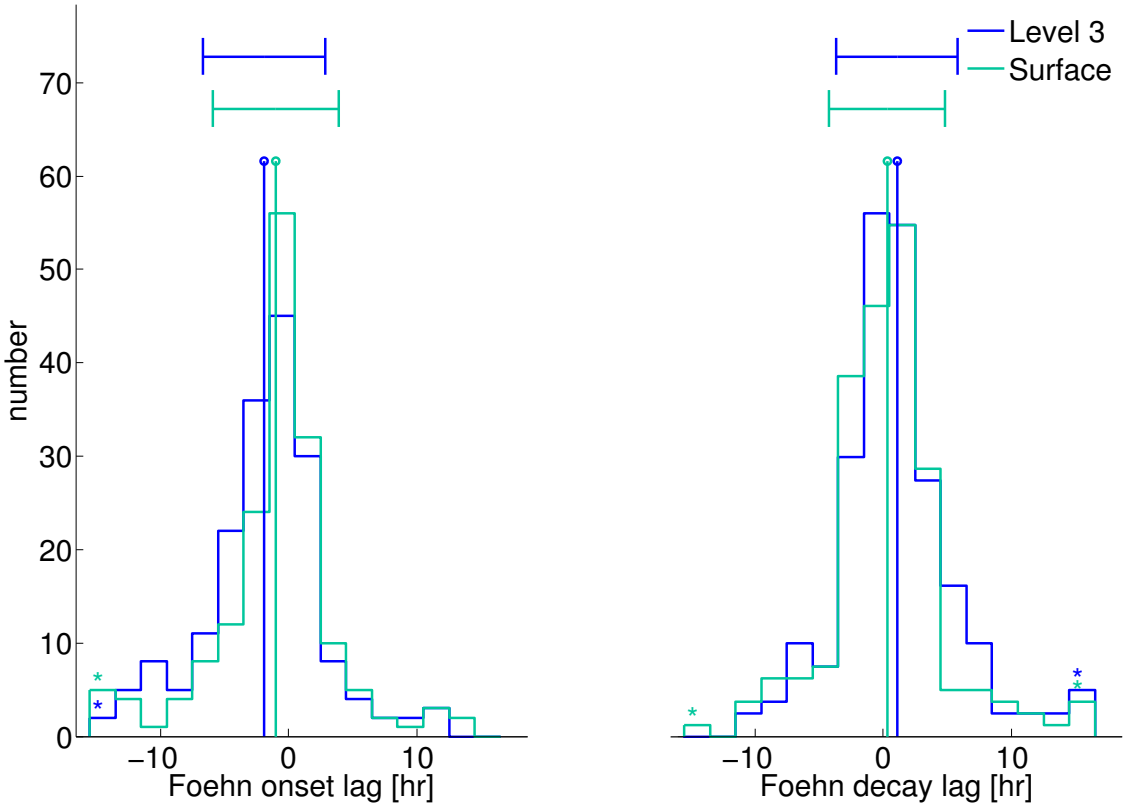

Figure 45: Foehn lag during spring months (SON).

References

Brinkman, W. (1974). Strong downslope winds at boulder. *Mon Wea Rev 102*, 592–602.

Burri, K., B. Duerr, T. Gutermann, A. Neururer, R. Werner, and E. Zala (2007, June). Foehn verification with the cosmo model. In *29th International Conference on Alpine Meteorology (ICAM)*, Chambéry, France.

Courvoisier, H. and T. Gutermann (1971). Zur praktischen anwendung des foehn-tests von widmer. Technical Report 21, MeteoSwiss. out of print.

Doms, G. and U. Schaettler (2002). A description of the nonhydrostatic regional model lm, part i: dynamics and numerics. Technical report, Consortium for Small-Scale Modelling.

Drechsel, S. and G. Mayr (2008). Objective forecasting of foehn winds for a subgrid-scale alpine valley. *Weather and Forecasting 23*(2), 205–218.

Drobinski, P., R. Steinacker, H. Richner, K. Baumann-Stanzer, G. Beffrey, B. Benech, H. Berger, B. Chimani, A. Dabas, M. Dorninger, et al. (2007). Foehn in the rhine valley during map: A review of its multiscale dynamics in complex valley geometry. *Quarterly Journal of the Royal Meteorological Society 133*(625), 897–916.

Duerr, B. (2008). Automatisiertes verfahren zur bestimmung von foehn in alpentaelern. Technical Report 223, MeteoSwiss.

Duerr, B. (2011). personal communication.

Duerr, B., M. Sprenger, O. Fuhrer, K. Burri, T. Gutermann, P. Haechler, A. Neururer, H. Richner, and R. Werner (2010, September). Foehn diagnosis and model comparison. In *10th EMS Annual Meeting, 10th European Conference on Applications of Meteorology (ECAM)*, Zuerich, Switzerland.

Durran, D. (1990). Mountain waves and downslope winds. *Atmospheric processes over complex terrain 23*(45), 59–18.

Ficker, H. and B. De Rudder (1943). *Foehn und Foehnwirkungen.* Akademische Verlagsgesellschaft Becker & Erler.

Fisher, N. and A. Lee (1983). A correlation coefficient for circular data. *Biometrika 70*(2), 327.

Hann, J. (1866). Zur frage ueber den ursprung des foehn. *Zeitschrift der oesterreichischen Gesellschaft fuer Meteorologie 1*(1), 257–263.

Hann, J. (1867). Der foehn in den oesterreichischen alpen. *Zeitschrift der oesterreichischen Gesellschaft fuer Meteorologie 2*(19), 433–445.

Kaufmann, P. (2008). Association of surface stations to nwp model grid points. Technical Report 9, MeteoSwiss.

MeteoSwiss (2010, August). List of the automated network stations of meteoswiss. http://www.meteoschweiz.admin.ch/web/en/climate/observation_systems/surface/swissmetnet/smn-stations.html.

MeteoSwiss (2012, January). The numerical weather prediction model cosmo. http://www.meteoswiss.admin.ch/web/en/weather/models/cosmo.html.

Oechslin, R. (2008). *Prediction of Foehn in the Alps with Boosting*. Ph. D. thesis, Bachelor thesis, ETH Zuerich.

Richner, H. and P. Haechler (2008, August). Understanding and forecasting alpine foehn - what do we know about it today? In *13th Mountain Meteorological Conference*, Whistler, British Columbia, Canada.

Schraff, C. and R. Hess (2011, January). A description of the nonhydrostatic regional model lm, part iii: Data assimilation. http://www.cosmo-model.org/content/tasks/operational/meteoSwiss/default.htm.

Schweitzer, H. (1952). Versuch einer erklaerung des foehns als luftstroemung mit ueberkritischer geschwindigkeit. *Meteorology and Atmospheric Physics 5*(3), 350–371.

Seibert, P. (1990). South foehn studies since the alpex experiment. *Meteorology and Atmospheric Physics 43*(1), 91–103.

Steinacker, R. (2006, March). Alpiner foehn-eine neue strophe zu einem alten lied. *Meteorologische Fortbildung (promet) 32*, 3–10.

The World Meteorological Institution (1992). *International Meteorological Vocabulary* (2nd ed.). The World Meteorological Institution.

Vergeiner, J. (2004). *South foehn studies and a new foehn classification scheme in the Wipp and Inn valley*. Ph. D. thesis, PhD thesis, University of Innsbruck.

Warner, T. (2011). *Numerical weather and climate prediction*. Cambridge Univ Pr.

Widmer, R. (1966). Statistische untersuchungen ueber den foehn im reusstal und versuch einer objektiven foehnprognose fuer die station altdorf. *Vierteljahresschrift der Naturforschende Gesellschaft in Zuerich 111*, 331–375.

Wild, H. (1868). *Ueber Foehn und Eiszeit*. Jent & Reinert.

Wilks, D. (2011). *Statistical methods in the atmospheric sciences*, Volume 100. Academic press.

Zaengl, G., A. Gohm, and F. Obleitner (2008). The impact of the pbl scheme and the vertical distribution of model layers on simulations of alpine foehn. *Meteorology and Atmospheric Physics 99*(1), 105–128.