

Validation of a Statistical Forecast Model for Zonda Wind in West Argentina Based on the Vertical Atmospheric Structure

Federico A. Norte, Silvia Simonelli

Instituto Argentino de Nivología, Glaciología y Ciencias Ambientales (IANIGLA), CCT-CONICET, Mendoza, Argentina

Email: nortefederico@hotmail.com, fnorte@prmarg.org

Received 17 November 2015; accepted 8 January 2016; published 11 January 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Zonda is a strong, warm, very dry wind associated with adiabatic compression upon descending the eastern slopes of the Andes Cordillera in western-central Argentina. This research seeks, first, to validate the skill of a statistical forecast of zonda based on the behavior of the vertical structure of the atmosphere and, second, to describe the climatology of the vertical profile leeward of the Andes. The forecast was built for May-August 1974/1983, and was verified against a series of cases recorded in the Mendoza Aero and San Juan Aero weather stations for May-August 2005/2014. It made use of the Stepwise Discriminant Analysis (SDA) and rawinsonde data from Mendoza Aero as predictors, with the following input variables: surface pressure, temperature, dew point, and the zonal and meridional components of the wind on surface and of the fixed levels up to 200 hPa. The variables selected as predictors by the SDA were: surface pressure, dew point depression at 850 hPa, meridional wind component at 850 hPa, and zonal wind component at 400 hPa. Climatology of the vertical profile of the atmosphere leeward of the Andes was built from daily rawinsonde data from Mendoza Aero for May-August 1974/1983. Zonda markedly influences the atmospheric structure leeward of the Andes in western-central Argentina. Its maximum impact occurs at 850 to 800 hPa, with significant heating and decrease of humidity. Validation of the prediction program considered deterministic and probabilistic forecasts. Contingency tables show that probability of zonda occurrence in the plains is generally overestimated, and false alarm cases are far more frequent than surprise events. The main contribution of this paper is precisely the validation of the prediction model, which ensures forecasters one more tool to improve zonda forecasting; this, in turn, will aid decision-makers when taking steps to ameliorate zonda wind impact.

Keywords

Zonda, Argentina, Statistical Forecast, Validation, Rawinsonde Climatology

1. Introduction

Zonda is a strong, warm, very dry wind associated with adiabatic compression upon descending the eastern slopes of the Andes Cordillera in western-central Argentina (**Figure 1**), with a higher frequency in the winter and spring seasons. Zonda has specific features (such as remarkable spatial and temporal variability), and complex behavior of associated meteorological parameters (such as temperature, relative humidity, surface pressure, and wind intensity). In populated regions, it produces a range of damage according to the intensity of wind gusts (e.g. blowing off roofs, felling trees and high-voltage power lines, and interrupting power supply and communication services). In addition, it favors the ignition and propagation of fires and causes damage to crops due to the strong gusts, sudden dryness and high temperatures, and it may also be responsible for premature fruit flowering. At high altitudes in the mountains, its occurrence accelerates snow melting and evaporation, modifying the snow pack depth, contributing to avalanches, and influencing the hydrological cycle.

Initial investigations into this kind of wind focused on the Alpine region, where it is known as the “foehn”. This term became the generic name for all descending, warm, dry winds whose features and effects depend on topography, its interaction with the atmospheric flux, and the particular meteorological conditions. Such winds have different names depending on where they blow: “Chinook” in the USA and Canada east of the Rocky Mountains, “Canterbury-northwestern” in New Zealand, “berg wind” in South Africa, “afganet” in Central Asia, “ibe” in western China, “halnywiatr” in Poland and Slovakia, and “austru” in western Romania [1]-[4].

The history of zonda research in Argentina, its climatology and the forecasting problems are well explained in [1], where there is reference to a probabilistic method and the predictors obtained.

The main goal of this paper is to validate the accuracy or skill of a statistical forecast, developed by Norte [1] [2], for zonda wind based on the behavior of the vertical structure of the atmosphere.



Figure 1. Location of Mendoza and San Juan provinces in Argentina, South America.

The model made use of the Stepwise Discriminant Analysis (SDA). Thanks to the positive results obtained for predicting zonda wind [2], this statistical tool was later also applied in: [5] to compare zonda wind and snowfall prediction, [6] to predict frost using only surface data, [7] to predict frost using the vertical profile of the atmosphere, [8] to predict cyclogenesis development in the Río de la Plata estuary and surroundings, [9] to predict late frost in a particular valley (San Rafael), and [10] to build a statistical model for predicting convection.

Three sets of Probabilistic Predictors were obtained by Norte [1] [2]:

- One set was obtained analyzing the 12 UTC pressure spatial gradient where some zonal and meridional pressure differences were arbitrarily defined.
- Another analysis was conducted taking the 12 UTC pressure values (reduced to sea level) of 81 surface meteorological stations, to obtain only the ones that would best discriminate among groups. SDA selected only five stations from the total set.
- Finally, a third set of predictors was obtained from Quintero (Chile) rawinsonde data.

These Probabilistic Predictors are not analyzed in this paper. For more information, it can be seen in [1] [2].

The secondary goal of this paper is to describe the climatology of the vertical profile leeward of the Andes Cordillera for the period considered when the forecast model was created. The data analyzed in [2] were used as reference.

This paper is structured in four sections and two appendices. Following this Introduction, **Section 2** describes the data and methodology. **Section 3** includes the results obtained with the prediction program, and the rawinsonde climatology showing the vertical structure of the atmosphere leeward of the Andes Cordillera. **Section 4** presents the discussion and conclusions about the prediction model. Finally, the different statistical tools used in this paper are explained in **Appendices A** and **B**.

The relevance of the present work is twofold: not only does it validate the effectiveness of a prediction method based on rawinsonde data, but it is also a valuable tool both for forecasters and decision-makers considering that the analyzed region is protected by laws for zonda risks prevention, especially when the phenomenon can be severe or extreme. Thus, the main contribution of this paper is that, by validating the prediction model, it ensures the use of one more tool for forecasters. Additionally, improving the accuracy in zonda wind forecasts will be of aid when making decisions to ameliorate its impact.

2. Data and Methodology

The statistical forecast of zonda wind based on the behavior of the vertical structure of the atmosphere was built for the months of May, June, July and August of the 1974/1983 period. It was verified against a series of cases recorded in the Mendoza Aero and San Juan Aero weather stations for those same months of the 2005/2014 period.

Rawinsonde data from Mendoza Aero station were scarce or non-existent for the validation period, and simulated rawinsonde data produced by the ARL (Air Resource Laboratory) from the NOAA (National Oceanic and Atmospheric Administration, Department of Commerce, USA) were used.

Zonda wind forecast with computational methods employed the Stepwise Discriminant Analysis (SDA) (see **Appendix A**) and four groups were defined:

- Zonda group: days with zonda on the plains (Mendoza Aero and/or San Juan Aero);
- Previous 24 hour group: the day preceding a zonda day;
- Previous 48 hour group: the day preceding that defined as the “previous 24 hours”;
- Others group: the rest of the series, *i.e.* days that do not belong to any of the groups mentioned above.

The days or records where categories overlapped were omitted from the sample (e.g. it could happen that a day belonging to the previous 48 hour group also belonged to the zonda group).

Once these groups were selected, it was necessary to know if any of the variables indicated unequivocally when a certain day belonged to one particular group. The variables that were thus classified would be assumed as predictors, and conditions would be met to forecast the zonda event 48 hours in advance. The SDA program made the discrimination.

Rawinsonde data from Mendoza Aero were used as predictors, with the following as input variables: surface pressure, temperature, dew point, and the zonal and meridional components of the wind on surface and of the fixed levels including up to the 200 hPa layer.

As mentioned in **Appendix A**, the chosen variables that discriminate best among groups are expressed as classi-

fication functions with their coefficients and constant terms, and the amount of functions is equal to the amount of groups. Each function permits to obtain a classification score on a determined day for that group, and the day will belong to the group whose function gets the highest score. In addition, the probability of a determined day belonging to one group or another can be obtained through the classification scores. The conversion formula of classification scores into posterior probabilities is shown in **Appendix A**.

This is how the SDA program re-evaluated the 1974/83 series, from where the case-by-case classification functions had been obtained, and this was used to evaluate the method.

As the four groups were defined with a temporal criterion, and there were variables that discriminate among those groups, it is then possible to forecast a zonda wind day.

With the discriminant functions obtained, it was possible to implement the prediction program.

Forecasts were verified evaluating the series with the discriminant functions. With the classification matrix from **Table 1** and the probability used to express each forecast, they can be evaluated as deterministic or probabilistic.

The classification matrix includes information on the number of cases assigned correctly to each group. Rows correspond to observations, columns to forecasts, and the trace to the events classified correctly.

Probabilistic forecasts can be evaluated through a contingency table, and although this is not a verification method in itself, its components provide a basis for determining scores and indexes easily.

A forecast is useful when the success percentage is higher than the probability of occurrence or non-occurrence.

The “skill score” combines the information of the contingency table to obtain the relation between the number of correct forecasts, the total number of forecasts, and the expected number of correct forecasts based on a method such as chance, persistence or climatology. Skill score is 1 when all forecasts are correct, and 0 when the correct forecasts are given solely by climatological chance.

For deterministic forecasting of zonda wind: the prediction program expresses forecasts in terms of probabilities. There is always a certain probability of zonda today, zonda tomorrow, zonda the day after tomorrow, or no zonda. The summation of these four probabilities is 1. It may happen that these groups are equiprobable or not. As the probability of one group increases towards 1 (diminishing the probability of the others), it is said deterministically that the phenomenon will occur.

For example, with results as those for the dates listed in **Table 2** applied to 12 UTC data, it is deterministically forecasted YES to zonda today and the other groups are assigned NO.

This criterion was applied, day by day, to the winters of the years 1974 to 1983, and also to the 2005/2014 period. Forecasts were evaluated for each day, and the occurrence or non-occurrence of the forecasted event was verified and rated after the following classification:

- If the forecast is YES and the observed is YES then there is a HIT.
- If the forecast is YES and the observed is NO then there is a MISS or a FALSE ALARM.
- If the forecast is NO and the observed is YES then there is a MISS or a SURPRISE.
- If the forecast is NO and the observed is NO then there is a HIT.

Also, the percentage of success “A” and the “degree of success” (or “skill score” [11]) were determined. The calculations for these percentages are explained in **Appendix B**.

Table 1. Classification matrix of cases using rawinsonde data at 12 UTC from Mendoza Aero, 1974/83.

Group	% Success	No. of cases assigned to each group			
		Others	Zonda	Prev24	Prev48
Others	51.9	344	45	135	139
Zonda	77.5	1	31	6	2
Prev24	52.9	4	4	18	8
Prev48	41.9	8	2	8	13
Total	52.9	357	82	167	162

Table 2. Posterior probabilities for some zonda events between 1983 and 2015.

Date	Posterior probabilities			
	Zonda	Prev24	Prev48	Others
23 Aug 83	0.97	0.02	0.01	0.00
28 Aug 10	0.86	0.05	0.06	0.01
25 Jul 11	0.75	0.05	0.14	0.05
27 May 12	0.73	0.07	0.14	0.04
31 May 13	0.98	0.01	0.06	0.01
27 Jun 13	0.91	0.03	0.03	0.01
7 Aug 13	0.98	0.05	0.07	0.01
29 May 14	0.93	0.02	0.03	0.01
5 Jun 15	0.99	0.01	0.03	0.01

For verifying probabilistic forecasts, the methods of Cardazzo *et al.* [12] were followed. The Brier score (BR) [13] was used to assess the quality of a probabilistic forecast (Equation (1)), and the Brier Skill score (BSS) to compare the BR value obtained for the forecast against the BR value of the climatology (Equation (2)), which gives some idea of how much more information the forecast provides in relation to a reference (*i.e.* the climatology).

The BSS can be interpreted as the distance from the probabilistic forecast of the ensemble to a perfect probabilistic forecast (*i.e.* the forecast predicting a 100% probability when the phenomenon occurs, and 0% when it does not).

$$BR = \frac{1}{N} \sum_{i=1}^N (p_i - o_i)^2 \quad (1)$$

$$BSS = 1 - \frac{BR}{BR_{climatology}} \quad (2)$$

For probabilistic forecasting of zonda: the developed prediction programs calculate, for each day, the probability of zonda occurring today, tomorrow, the day after tomorrow or not occurring, according to the discriminant function obtained. The phenomenon can occur m times with a given probability x_i stated n times. The relative frequency y_i of a probability x_i is the value obtained from dividing the times the phenomenon was observed (m) by the times it was forecasted (n) with a given probability x_i (see **Appendix B**).

Besides, the climatology of the vertical profile of the atmosphere leeward of the Andes Cordillera was built with daily rawinsonde data from Mendoza Aero station for the same months of the same 1974/1983 period.

Calculations included:

1) The mean rawinsonde values belonging to the groups of days with surface zonda (*i.e.* wind blowing in the plains), days with high zonda (*i.e.* wind blowing at height) [1] [2] and total days (including surface zonda, high zonda and remaining days). In particular for Mendoza Aero, the conditions for absence of high and surface zonda were analyzed, defining this group as “no-zonda day”.

2) The anomalies of different thermodynamic parameters across groups and across locations for each group.

3) The vertical variance curves of values from Student’s t -test for validating the significance of the parameters according to levels and to groups.

4) The mean vertical profiles of the zonal and meridional wind component.

3. Results

3.1. The Prediction Program

The input data corresponded to the surface pressure values, plus the temperature and dew point values, and the

zonal and meridional components of wind on surface, and the fixed levels including up to 200 hPa.

From the 53 variables considered, SDA only selected four as relevant to discriminate and to act as predictors: surface pressure, dew point depression at 850 hPa, meridional wind component at the same level, and zonal wind component at 400 hPa.

Surface pressure was selected as a relevant variable because it decreases when zonda occurs in the plains. Dew point depression at 850 hPa is important because the maximum dryness and heating due to zonda are usually recorded at that level; moreover, in the morning of zonda days in the plains, zonda may already be blowing at that height without having reached the surface (see **Table 3**).

A north component higher than the average at lower layers was observed in the analysis of synoptic and aerological conditions connected to zonda (see [2], Chapter 4), which would explain why the meridional wind component at 850 hPa was selected. Also, the higher west component above the summits of the Andes, associated with the closeness of the jet stream on zonda days, accounts for its importance at the 400 hPa level.

From the results obtained for the 1974/83 period, contingency tables were built:

In addition, **Table 4** shows that the hits of the (ab)-type (NO forecasted/NO observed) include the cases OTHERS forecasted/OTHERS observed, OTHERS forecasted/PREV24 observed, and OTHERS forecasted/ PREV48 observed, *i.e.* the remaining days when ZONDA was neither forecasted nor observed. These cases count as “hits” regarding the forecast and observation of zonda.

This criterion may raise doubts, which is why another contingency table (**Table 5**) was built based on the discrimination, between Zonda and Others, of variables derived from Mendoza Aero rawinsonde data (but using only two groups). It can be noted that this table and the values for A, L and S do not differ significantly from those obtained with four groups.

Table 3. Selection of predictor variables using rawinsonde data from Mendoza Aero, 1974/83.

Selected variable	Fisher <i>F</i> -value
Surface pressure	51.3284
Dew point depression at 850 hPa	8.5275
Meridional wind component at 850 hPa	5.5702
Zonal wind component at 400 hPa	5.7827

Table 4. Contingency table for 1974/83 using the four groups.

Observed \ Forecasted	Forecasted		Sum
	Yes	No	
Yes	31	10	41
No	51	677	728
Sum	82	687	769

where A = 92%, L = 112, and S = 0.90 (see **Appendix B**).

Table 5. Contingency table for 1974/83 using only two groups.

Observed \ Forecasted	Forecasted		Sum
	Yes	No	
Yes	35	5	40
No	73	655	728
Sum	108	660	768

where A = 89%, L = 130.72, and S = 0.87.

Considering the little difference observed in the validation, and the fact that using all the groups allows inferring possible conditions up to two days in advance, it is clear that it is more advantageous to forecast using the four groups than with just two.

As regards the results obtained for the 2005/2014 validation period, different probability thresholds were tested, starting from threshold 0.60 to 0.75 (see [Table 6](#) to [Table 9](#)).

The values listed in these tables were taken from the classification matrices calculated by SDA. The number of times zonda was forecasted but did not occur (false alarms) is much higher than the number of times Zonda was not forecasted but did occur (surprises). Using Equations (1) and (2) from Data and Methodology, the following BSS values were obtained for the different thresholds ([Table 10](#)).

In other prediction models used by Norte [1] [2], not developed in this paper, the success probabilities show a normal distribution; however, in the method described here for the current study, the frequency of probabilities increase towards total success (see [Figure 2](#)).

Table 6. Contingency table for the 2005/14 period with threshold = 0.60.

Observed \ Forecasted	Yes	No	Sum
	Yes	24	9
No	13	127	140
Sum	37	136	173

where A = 87.3%, and S = 1.13.

Table 7. Contingency table for the 2005/14 period with threshold = 0.65.

Observed \ Forecasted	Yes	No	Sum
	Yes	23	9
No	14	127	141
Sum.	37	136	173

where A = 86.7%, and S = 1.13.

Table 8. Contingency table for the 2005/14 period with threshold = 0.70.

Observed \ Forecasted	Yes	No	Sum
	Yes	23	6
No	14	130	144
Sum	37	136	173

where A = 88.4%, and S = 1.15.

Table 9. Contingency table for the 2005/14 period with threshold = 0.75.

Observed \ Forecasted	Yes	No	Sum
	Yes	20	5
No	17	131	148
Sum	37	136	173

where A = 87.3%, and S = 1.18.

Table 10. BSS values for the different thresholds.

Threshold	BSS
75	0.99987788
70	0.99990254
65	0.99989733
60	0.99990446

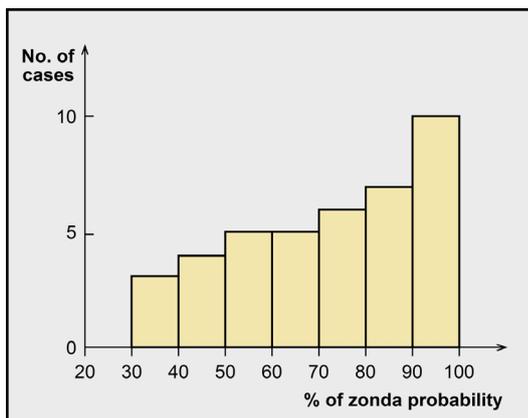


Figure 2. Frequency of zonda probability for the 1974/83 period.

3.2. Vertical Structure of the Atmosphere Leeward of the Andes Cordillera: Rawinsonde Climatology

Table 11, **Table 12** and **Table 13** were built with data of May, June, July and August of the 1974/1983 period from Mendoza Aero, and they show, respectively: mean rawinsonde values for the total of days without zonda, mean radiosonde values for the days with zonda in the plains, and the anomalies of some parameters as mentioned in [2] (Chapter 4).

The abbreviations in the following tables mean: p (atmospheric pressure), h (height), T (temperature), σT (standard deviation of temperature), Td (dewpoint temperature), σTd (standard deviation of dewpoint temperature), T-Td (dew point depression), $\sigma T-Td$ (standard deviation of dew point depression), θ (potential temperature), θ_e (equivalent potential temperature), RH (relative humidity), u (zonal wind component), v (meridional wind component).

Table 13, of anomalies, shows that:

1) Height anomalies are negative up to 200 hPa, and they decrease quite steadily towards the upper troposphere, indicating the proximity of a trough (since, after the passage of an anticyclone, the trend is negative at first on the surface and then it extends to the upper levels).

2) Temperature anomalies are positive in the entire atmosphere with a maximum at around 800 hPa due to the air descent that affected those levels. From 800 hPa to 600 hPa there is a notable increase of the vertical gradient due to the effect of zonda.

3) Humidity anomalies indicate that ΔTd values are slightly positive at very low levels, with a negative maximum at 800 hPa but turning positive again at approximately 650 hPa. Dew point depression and relative humidity anomalies are more telling, indicating that the lower troposphere is drier than normal. From these, together with the temperature and equivalent potential temperature anomalies, it can be observed that zonda usually occurs with an air mass warmer than normal, and that its effect on humidity reaches a maximum at 800 hPa.

4) Wind anomalies indicate that, except at the lower levels (which zonda has not yet reached at the time of sounding), the zonal component (u) is higher than its climatological value, with a marked increase of the vertical shear between 1500 and 3300 masl and a slight increase in the upper troposphere. Regarding the meridional

Table 11. Mean rawinsonde values from Mendoza Aero for the total of days (May to August 1974/83).

p (hPa)	h (m)	T (°C)	σT (°C)	Td (°C)	σTd (°C)	T-Td (°C)	$\sigma T-Td$ (°C)	θ (°K)	θ_e (°K)	RH (%)	u (m/s)	v (m/s)
934.5	703.0	4.8	4.50	0.5	4.4	4.2	3.4	283.2	295.1	74	0.10	0.50
900.0	1011.0	8.4	4.50	1.3	5.3	7.1	5.1	290.0	303.4	61	0.41	0.93
850.0	1481.0	7.1	5.10	-0.6	5.8	7.8	6.4	293.4	305.9	58	0.37	0.35
800.0	1977.4	6.0	6.00	-3.6	6.0	9.6	7.8	297.4	308.3	50	---	---
700.0	3061.6	2.4	5.40	-11.6	6.2	14.0	7.8	305.1	312.2	34	3.57	-1.98
600.0	4287.7	-5.3	4.40	-20.1	6.0	14.8	6.4	309.8	314.1	30	6.93	-1.73
500.0	5689.8	-15.5	4.40	-29.3	5.9	13.8	6.0	314.0	316.3	29	11.73	-0.20
400.0	7332.7	-27.6	4.10	-40.4	5.7	12.8	5.4	318.8	319.8	28	15.75	1.38
300.0	9338.2	-42.1	5.10	-53.9	6.8	11.8	4.7	325.7	326.0	24	19.78	2.08
250.0	10,548.4	-50.4	3.40	-68.2	2.8	17.7	0.6	330.8	330.9	8	21.30	2.24
200.0	11,982.5	-56.5	6.90	-73.0	7.1	16.5	1.2	342.9	343.0	8	23.37	2.04

Table 12. Mean rawinsonde values from Mendoza Aero for surface zonda days (May to August 1974/83).

p	h	T	σT	Td	σTd	T-Td	$\sigma T-Td$	θ	θ_e	RH	u	v
925.4	703.0	6.9	5.40	0.8	4.10	6.10	6.40	286.20	298.70	66.00	0.46	0.22
900.0	933.0	11.4	4.80	1.4	5.80	10.00	8.00	293.10	306.70	50.00	1.34	-1.54
850.0	1410.0	12.2	4.90	-1.4	7.00	13.70	9.00	299.10	311.10	37.00	1.80	-2.48
800.0	1916.4	11.6	4.50	-5.0	7.10	16.60	9.00	303.70	313.50	29.00	3.85	-4.76
700.0	3016.9	4.9	3.70	-12.4	7.00	17.40	7.10	307.70	314.10	25.00	8.74	-5.25
600.0	4251.2	-4.0	3.70	-19.8	6.60	15.70	6.20	311.20	315.40	27.00	11.98	-4.36
500.0	5658.6	-14.5	3.80	-28.2	6.80	13.70	6.40	314.90	317.40	29.00	18.77	-1.97
400.0	7305.4	-27.1	3.60	-39.2	6.10	12.00	5.80	319.30	320.50	29.00	25.99	-0.91
300.0	9312.2	-41.0	8.30	-52.2	10.20	11.10	5.00	325.50	325.70	8.00	32.13	0.00
250.0	10522.4	-50.1	3.70	-67.9	3.04	17.80	0.69	330.90	331.00	8.00	33.64	-1.17
200.0	11957.1	-54.5	11.50	-71.0	12.80	16.50	1.90	343.10	343.10	7.00	36.71	0.64

Table 13. Anomalies (Δ) of some parameters from Mendoza Aero (zondadays group minus total days group).

Level	Δh	ΔT	ΔTd	$\Delta T-Td$	$\Delta \theta$	$\Delta \theta_e$	ΔRH	Δu	Δv
Surface	-72.8	2.1	0.3	1.9	2.9	3.3	-9	0.36	-0.28
900.0	-70.9	3.0	0.1	2.9	3.1	3.3	-11	0.93	-2.47
850.0	-64.1	5.1	-0.8	5.9	5.4	4.9	-19	1.43	-2.83
800.0	-54.6	5.6	-1.4	7.0	6.0	5.2	-19	---	---
700.0	-38.7	2.5	-0.2	3.4	2.7	2.4	8	5.17	-3.52
600.0	-30.1	1.3	0.3	0.9	1.5	1.6	-2	5.05	-2.63
500.0	-24.0	1.0	1.0	-0.1	1.2	1.5	1	7.04	-1.77
400.0	-19.1	0.5	1.2	-0.8	0.7	0.8	2	10.24	-2.29
300.0	-12.4	1.1	1.7	-0.7	1.6	1.6	2	12.35	-2.00
250.0	-8.7	0.3	0.3	0.1	0.5	0.5	0	12.34	-3.41
200.0	-1.1	2.0	2.0	0.0	3.1	3.2	0	13.34	-2.00

component (v), although its climatological value is almost null, there is a north component in zonda events, especially between 1500 to 5500 masl, from which it decreases once again. At 5500 masl, mean climatological wind (V) corresponds to 271 or to 11 m/s, while mean wind associated to zonda at that height is 282 or 17 m/s, confirming that the wind increases when there is zonda at the height of the Andean summits (Figure 3).

Although the maximum heating produced by zonda occurs on the surface [2] (Chapter 2), soundings fail to show this. This difference could be due to the launching time of the rawinsonde weather balloon (12 UTC or 9 AM local time).

Zonda climatology [1] shows that zonda starts early in the morning in a few opportunities, but the highest onset frequency is during afternoon hours.

Also, mean rawinsonde values corresponding to the days of the series when zonda did not occur either at height or in the plains, defined as “no-zonda days”, were calculated for Mendoza Aero (Table 14).

In this case, the mean sounding values that show colder and more humid air conditions in almost the entire vertical column would be associated preferentially to polar air situations.

Figure 4 shows the vertical variations of Student’s *t*-test values (significant at 5% if $t > 1.965$) for the zonda days group against the total days group. The fact that the most important differences in temperature and dew point depression are found at 800 hPa and not on the surface would be due to the causes already mentioned regarding the time when the rawinsonde data is recorded and the onset time of zonda in the plains. The differences lose significance below the mean height of the Andean summits. As for dew point, it reaches significant values

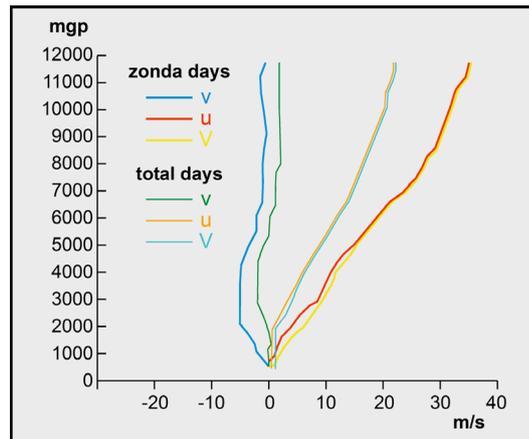


Figure 3. Mean vertical distribution of wind in Mendoza Aero at 12 UTC from May to August 1974/83.

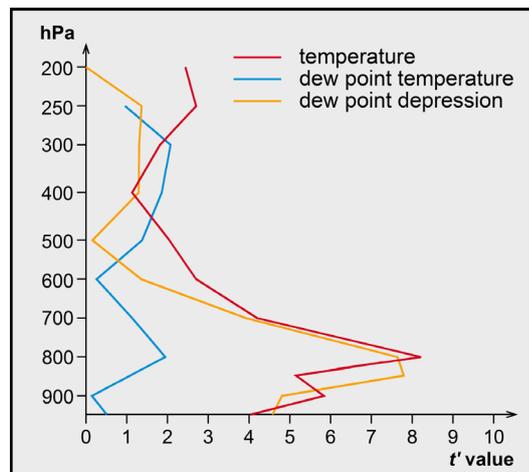


Figure 4. Vertical variation of student’s *t*-test values for zonda days group against total days group.

only at 800 hPa. The fact that the values are not statistically significant at the lower levels close to the ground (900 hPa) would indicate that the subsidence inversion is almost permanent in Mendoza Aero, at least during the winter periods.

Figure 5 shows significant differences in the entire air column, except in the surface, and allows inferring that the conditions for no-zonda days at any level are not the norm in the region during the analyzed months. The maximum values found between 800 and 700 hPa would be due to the predominance at those layers of colder and more humid air that is observed in these circumstances, in contrast with the heating and dryness produced above the inversion that can be recorded in cases of zonda occurrence.

4. Conclusions

Zonda wind occurrence markedly influences the structure of the atmosphere leeward of the Andes Cordillera in western-central Argentina.

Its maximum impact regarding temperature and dew point depression occurs at the 850 to 800 hPa levels, with significant heating and decrease of humidity.

Zonda occurs within a relatively warm air mass, with temperature anomalies being positive along the entire column. The layers affected by zonda show a notable north component in the circulation.

Table 14. Mean rawinsonde values from Mendoza Aero for no-zonda days (May to August 1974/83).

p	h	T	σT	Td	σTd	T-Td	$\sigma T-Td$	θ	θ_e	RH	u	v
937	703.0	5.3	4.5	0.8	4.6	4.5	3.3	283.6	295.7	72	0.07	1.02
900	1031.9	7.5	4.4	1.1	5.4	6.4	4.7	289.1	302.2	64	0.42	2.00
850	1499.6	5.2	4.5	-0.8	5.8	6.1	5.5	291.5	303.7	64	0.21	1.52
800	1991.8	3.0	5.2	-3.3	5.7	6.4	6.1	294.3	305.2	62	0.35	0.96
700	3065.2	0.0	5.7	-10.7	5.7	10.6	7.6	302.2	309.7	45	2.25	-1.20
600	4281.8	-6.8	4.6	-20.0	5.8	13.1	6.4	308.0	312.2	34	5.85	-1.79
500	5675.7	-17.0	4.5	-29.6	5.8	12.0	5.0	312.1	314.4	32	10.67	-0.93
400	7309.7	-28.8	4.3	-40.8	5.8	12.0	5.3	317.2	318.2	30	15.29	0.53
300	9308.1	-42.7	5.6	-53.9	7.2	11.2	4.6	325.0	325.1	8	20.88	1.09
250	10516.6	-50.6	3.6	-68.3	8.7	17.7	---	330.6	330.7	8	22.94	0.80
200	11953.7	-55.5	7.3	-72.2	7.6	16.6	---	344.5	344.6	8	26.52	0.46

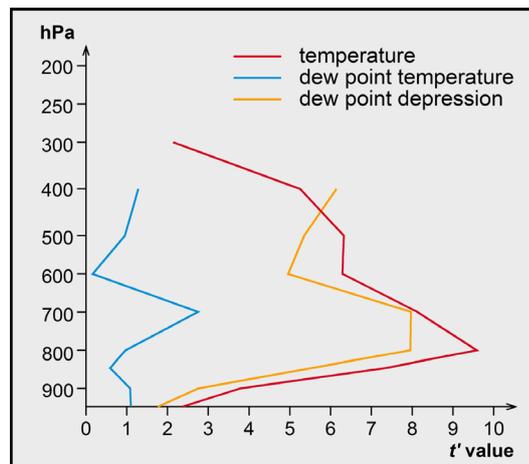


Figure 5. Vertical variation of student's t -test values for no-zonda days group against total days group.

Surface pressure was selected as a relevant variable because it decreases when zonda occurs in the plains. Dew point depression at 850 hPa is important because the maximum dryness and heating due to zonda are usually recorded at that level; moreover, in the morning of zonda days in the plains, zonda may already be blowing at that height without having reached the surface. A north component higher than the average at lower layers was observed in the analysis of synoptic and aerological conditions connected to zonda [2] (Chapter 4), which would explain why the meridional wind component at 850 hPa was selected. Also, the higher west component above the Andean summits, associated with the closeness of the jet stream on zonda days, accounts for its importance at the 400 hPa level.

The validation of the implemented prediction program considered deterministic and probabilistic forecasts, and was highly successful. Contingency tables show that the program generally overestimates the probability of zonda occurrence in the plains, and false alarm cases are far more frequent than surprise events. All in all, the prediction program is considered to be a quite effective tool.

For this region of Argentina, it is valid what Richner and Hächler [15] state for the Alps: “In practice, the skills of experienced forecasters who are familiar with the local situations are still an indispensable prerequisite for a successful forecast! They know from experience how a somewhat different wind direction might influence the onset or breakdown of foehn in a given valley, how observed wind data must be interpreted to arrive at a correct prediction. On the other hand, any tool, be it based on probability or on model output, is a welcome and appreciated support giving a first approximation which is subsequently modulated with the forecaster’s experience and skill.” The synoptic chart, particularly the surface pressure pattern, is one of the most useful tools for predicting the onset of foehn.

The main contribution of this research is precisely that, by validating the prediction model, it ensures the support of one more tool for forecasters. In addition, improving the accuracy in zonda, wind forecasts will be of great aid for decision-makers when taking steps to ameliorate its impact.

Acknowledgements

The author wishes to thank to Rafael Bottero for his help with figures and photos, and to Claudia Bottero for formatting and proofreading.

This paper has been funded by PIP 1120090100439 of CONICET (National Research Council of Argentina).

References

- [1] Norte, F.A. (2015) Understanding and Forecasting Zonda Wind (Andean Foehn) in Argentina: A Review. *Atmospheric and Climate Sciences*, **5**, 163-193. <http://dx.doi.org/10.4236/acs.2015.53012>
- [2] Norte, F. (1988) Características del Viento Zonda en la Región de Cuyo. PhD Thesis, University of Buenos Aires, 255 p. Available from Programa Regional de Meteorología (PRM), Instituto Argentino de Nivología, Glaciología y Ciencias Ambientales. (IANIGLA), Centro Científico Tecnológico (CCT), CONICET, Mendoza, Argentina. (In Spanish) <http://www.prmarg.org/tesis-doctorado>
- [3] Norte, F., Ulke, A., Simonelli, S. and Viale, M. (2008) The Severe Zonda Wind Event of 11 July 2006 East of the Andes Cordillera (Argentina): A Case Study Using The BRAMS Model. *Meteorology and Atmospheric Physics*, **102**, 1-14. <http://dx.doi.org/10.1007/s00703-008-0011-6>
- [4] Huschke, R.E., Ed. (1980) Glossary of meteorology. American Meteorological Society, Boston, 638 p.
- [5] Norte, F. and Silva, M. (1990) Comparación de la efectividad del pronóstico del viento Zonda y el de las nevadas en Mendoza (Argentina) usando Análisis Discriminante Escalonado. *Proceedings of the IV Congreso Interamericano-II Colombiano de Meteorología*, Bogotá, Colombia, 17-21 September 1990, 71-75.
- [6] Norte, F. and Silva, M. (1990) Métodos objetivos de predicción de heladas en Mendoza. Parte I: Aplicados al oasis de San Rafael con solo información de superficie. *Proceedings of the XVI Reunión Científica de la AAGG*, Bahía Blanca, 23-27 October 1990, 88.
- [7] Norte, F., Silva, M. and Cristaldo, J. (1991) Métodos objetivos de predicción de heladas en Mendoza. Parte II: Aplicados a tres oasis distintos, incluyendo información de la estructura vertical de la atmósfera. *Proceedings of CONGREMET VI*, Buenos Aires, 23-27 September 1991, 43-44.
- [8] Seluchi, M. (1993) Estudio del comportamiento de los sistemas sinópticos migratorios en la Argentina. PhD Thesis, University of Buenos Aires, 260 p.
- [9] Seluchi, M. and Norte, F. (1994) Métodos objetivos de predicción de heladas en Mendoza. Parte III: Heladas tardías en

- San Rafael, incluyendo información de la estructura vertical y espacial de la atmósfera. *Meteorologica*, **19**, 13-21.
- [10] Simonelli, S. (2000) Modelo estadístico de pronóstico de convección para la zona norte de la Provincia de Mendoza. PhD Thesis, University of Buenos Aires, 158 pp. Available from Programa Regional de Meteorología (PRM), Instituto Argentino de Nivología, Glaciología y Ciencias Ambientales. (IANIGLA), Centro Científico Tecnológico (CCT), CONICET, Mendoza, Argentina. (In Spanish) <http://www.prmarg.org/tesis-doctorado>
- [11] Brier, G.W. (1950) Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review*, **78**, 1-3. [http://dx.doi.org/10.1175/1520-0493\(1950\)078<0001:VOFEIT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2)
- [12] Cardazzo, S., Ruiz, J. and Saulo, C. (2010) Calibración de los pronósticos probabilísticos de precipitación derivados de un ensamble multi-modelo utilizando diferentes metodologías. *Meteorologica*, **35**, 41-51..
- [13] Murphy, A.H. (1973) A New Vector Partition of the Probability Score. *Journal of Applied Meteorology*, **12**, 595-600. [http://dx.doi.org/10.1175/1520-0450\(1973\)012<0595:ANVPOT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2)
- [14] Haber, A. and Runyon, R. (1973) Estadística general. Fondo Educativo Interamericano S.A., 371 p.
- [15] Richner, H. and Hächler, P. (2013) Understanding and Forecasting Alpine Foehn. In: Chow, F.K., De Wekker, S.F.J. and Snyder, B.J., Eds., *Mountain Weather Research and Forecasting* (Chapter 4), Springer Atmospheric Sciences, 219-260. http://dx.doi.org/10.1007/978-94-007-4098-3_4

Appendix A: The Stepwise Discriminant Analysis (SDA) Program

The program applied in this paper to discriminate between zonda days and no-zonda days, and to obtain a system for forecasting the phenomenon, can also generate a discriminant analysis between two or more groups of parameters.

The variables used to calculate the linear classification functions are chosen step by step, hence stepwise. At each step the variable that adds the most to the discrimination between groups is entered into (or the variable that adds the least is removed from) the discriminant function.

The output shows, in the first place, a list of the variables to be entered. Then, it indicates the specific options of the program, many of them pre-assigned, such as Fisher's F -statistics [14]—which is the minimum required for the variable to be entered in or removed from the discriminant function, the maximum amount of steps it accepts and the prior probabilities of each group.

Then, the program informs the number of cases read. A characteristic of this program is that it uses only complete cases. This means that if the value of any variable in a case is missing or out of range, the case is omitted from all computations. The program informs the number of omitted cases and of the remaining cases which are used to carry out the stepwise discrimination.

Next, the program estimates the mean values, standard deviations and variation coefficients for each variable in each group and for all groups.

Standard deviation within groups is:

$$\left\{ \sum (N_k - 1) S_k^2 / \sum (N_k - 1) \right\}^{\frac{1}{2}},$$

where Σ identifies the summation, S_k^2 is the variance and N_k the sample size of the k th group.

Step 0 corresponds to the step before any variable is entered into the discriminant function.

The program prints the F -statistic “to enter” for each variable considered. F was computed from a one-way analysis of variance on the variables of the groups used in the analysis, which allowed establishing the degree of likeness or difference between each group for each variable.

The variable with the highest F -to-enter at step 0 is entered into the discriminant functions. This is the variable that discriminates the best between groups.

Step 1 computes the following:

- 1) F -to-remove for the variable in the equation.
- 2) F -to-enter for each variable not in the equation, which is equal to the F -statistic from the one-way variance analysis of the residual variables.
- 3) Wilk's lambda or U -statistic.
- 4) Approximate F -statistic (this is a transformation of Wilk's lambda).
- 5) Mahalanobi's D -distance.
- 6) F -matrix that contains F -values computed from the Mahalanobi's D -distance.
- 7) The classification functions.

The classification functions can be used a posteriori to classify cases within groups; the case is assigned to the group whose classification function has the largest value.

This can be done in two ways:

- 1) With a prediction program of the phenomenon that uses the classification functions for new cases.
- 2) With an assessing program that classifies the same cases of the sample from which the functions were obtained (thus, the starting sample is re-classified).

The classification score for each group is computed from the discriminant function coefficients (data are multiplied by the coefficients and the constant term is added). The case is classified within the group with the highest classification score. The score can also be expressed as probabilities: q is the number of groups and S_{ij} is the classification score for the “ i th” case of the “ j th” group, so that the posterior probability that case i belongs to j group is:

$$P_{ij} = \frac{e^{S_{ij}}}{\sum_{k=1}^q e^{S_{ik}}}.$$

The prediction programs mentioned in 3.1 were thus made.

The last step, besides giving the results obtained in each step, also produces:

- 1) The classification matrix.
- 2) The jackknifed classification matrix.
- 3) A summary table: this contains a one-line summary for each step, including the F -to-enter (or F -to-remove) for the variable entered (or removed), Wilk's lambda or U -statistic, the approximate F -statistic and the degrees of freedom.

The SDA program also solves others issues that do not fall within the scope of this paper.

Appendix B: Forecast Evaluation Methods

For calculating the percentages of success "A" and the success degree or skill-score "S", the following notation was established:

N = number of forecasts

(A) = Number of times that YES was forecasted

(B) = Number of times that NO was forecasted

(a) = Number of times that YES was observed

(b) = Number of times that NO was observed

Also,

(AB) = Number of times that YES was forecasted and YES occurred (hit)

(Ab) = Number of times that YES was forecasted and NO occurred (false alarm)

(aB) = Number of times that NO was forecasted and YES occurred (surprise)

(ab) = Number of times that NO was forecasted and NO occurred (hit)

Then, success percentage is calculated:

$$A = \frac{(A)(B) + (ab)}{N} \times 100.$$

The success degree or skill-score is defined:

$$S = \frac{(AB) + (ab) - L}{N - L},$$

with

$$L = \frac{(A)(B) + (a)(b)}{N},$$

"L" being the amount of success by climatological chance, which indicates how many forecasts would a neophyte guess by chance. The success degree "S" (skill-score) filters those guesses at random.

If the amount of success of the implemented prediction methods were equal to the amount of success by climatological chance, then S would be equal to zero. Conversely, S equal to one would mean total success.

For evaluating the forecasting methods, the correlation coefficient r_p between the probabilities x_i and the observed relative frequencies y_i was calculated. It is defined:

$$r_p = \frac{\sum_{i=1}^{\gamma} x_i \cdot y_i}{\gamma \sigma_{x_i} \cdot \sigma_{y_i}}.$$

where γ is the number of the different probabilities in use, σ_{x_i} is the variance of the probabilities, and σ_{y_i} is the variance of the relative frequencies.

For the forecast to be useful, r_p must be positive. In addition, if the N number of forecast increases, the frequency of zonda wind occurrence will approach the climatological frequency; this permits obtaining a greater score of the r_p value but the forecast is less useful.

That is why the mean absolute value is calculated for the difference between the square of the forecasted probability and 1 if zonda occurs, or 0 if zonda does not occur. This index, proposed by Brier, forces the forecast to approach the extreme probabilities (probability 0 corresponds to the deterministic forecast NO, and 1 to YES).

$$\bar{D} = \frac{1}{N} \sum_{i=1}^N D_i \begin{cases} D_i = 1 - x_i^2 & (\text{zonda days}) \\ D_i = x_i^2 & (\text{no zonda days}) \end{cases}$$

The lower the mean D value is, the better the forecast will be. Absolute deterministic success is reached if mean D is equal to 0.