Intercomparison between METAR- and SYNOP-based fog climatologies

Melina Sol Yabra (myabra@smn.gob.ar)
Servicio Meteorológico Nacional
Ramón de Elía
Servicio Meteorológico Nacional
Luciano Vidal
Servicio Meteorológico Nacional
Matilde Nicolini
Centro de Investigaciones del Mar y la Atmósfera (CIMA), CONICET – UBA

Research Article

Keywords: SYNOP, METAR, fog, aviation, meteorology, climatology

Posted Date: April 3rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2762918/v1

Abstract

Weather observations and forecasts play a crucial role in the safety and economics of aviation, particularly in the presence of hazards such as fog. Climatological studies of this phenomenon are hence standard information for the aviation industry. Some aeronautical meteorological stations located at airports are also combined with synoptic stations and have to provide local and current weather conditions in two different types of reports: SYNOP for synoptic purposes and METAR for aeronautical purposes. When climatological studies are to be carried out, the question arises of which type of report is more appropriate for the task. In this work, a comparison is made between these two types of hourly meteorological observation reports, with the aim of identifying the conditions under which differences can have a significant impact on the final results. For this purpose, SYNOP and METAR reports for the period 2000-2019 from 13 Argentine airports were selected from diverse databases. A perfect agreement between the two reports is not found, but a varying degree of similarity depends on the variable. A key difference is found in "present weather" mainly due to different reporting rules. In this context, six different selection criteria using visibility, dew point depression, and present weather are used to evaluate the impact of the differences found in the fog climatology. In conclusion, SYNOP data could be used instead of METAR in statistical studies if the need arises, but with the caveat that fog frequency of occurrence should not be interpreted too strictly.

1. Introduction

Despite technological advances to enable flight operations, the state of the atmosphere continues to significantly affect the safety and the economy of the aviation industry (Gultepe et al., 2019). It remains then, crucial to have access to reliable meteorological observations and forecasts to support operations at airports (WMO, 2014). Currently, surface weather stations around the world report more than 800,000 in-situ observations per day (Haiden et al., 2018), and some of these are located at airports. These last are called "aeronautical meteorological stations" and have different rules of procedures than the synoptic stations. In many cases aeronautical and synoptic stations are combined, and observers must report current weather observations in two different ways, each one for a different purpose. On the one hand, synoptic observation messages (SYNOPs) can be reported every 6, 3, or 1 hours depending on the country and the weather station. The meteorological data included must be representative of the state of the atmosphere in a wide area whose dimensions can vary between 2,000 and 10,000 km^2 depending on the terrain type (WMO, 2010). On the other hand, aeronautical meteorological reports (METARs) must be reported at least every hour, 24 hours a day, unless the agreement between the meteorological authority, the competent authority of the Air Traffic Service (ATS), and concerned airlines state otherwise (WMO, 2018a). Unlike SYNOPs, the content in these messages must be representative of the airport's area and its surroundings up to 16 km from the aerodrome reference point, which implies an area around 10 times smaller (WMO, 2010). As a general rule, the same regulations established for synoptic observations are applied to aeronautical observations, although different procedures have to be applied for some of the variables (WMO, 2010). The observations necessary for the preparation of the SYNOP message are specified in the Manual of the Global Observing System (WMO, 2015), and those for the METAR in the Technical Regulation No. 49 of the World Meteorological Organization (WMO; WMO, 2018a). Once completed, both reports are exchanged via the WMO Global Telecommunication System (GTS).

Presently, meteorological observations can be carried out by fully automated weather stations, or by semi-automatic systems that still retain the need for a human observer for the completion of some activities (WMO, 2014). In this last kind of station, although the measurement of certain variables is carried out automatically, it is the meteorological observer who carries out the necessary coding, prepares and sends the reports. Measurements and coding of other variables, such as visibility and present weather, are carried out purely by the human observer. Additionally, the expansion and improvement of airports also brought new regulations in the meteorological observation process. For example, at airports equipped with Instrument Landing Systems (ILS) categories I, II, and III in at least one of its runways, automatic weather observation systems must be installed adjacent to these runways (ICAO, 1969). These devices must be integrated with automatic systems for obtaining, processing, disseminating, and presenting in real time the meteorological variables that influence landing and takeoff operations (WMO, 2018a). This context has led to heterogeneity in observing systems between different airports in the world, and even within the same country and the same airport over time.

One of the main meteorological variables that influence daily airport operations is visibility, with fog being one of the most disruptive meteorological phenomena of air traffic (Gultepe et al., 2009). This weather phenomenon can cause large economic losses due to cancellations, flight delays, diversions to alternative destinations and, in the most adverse cases, can cause accidents that generate both material and human losses. In the international literature on the occurrence of fog at airports, numerous works used specifically SYNOP data (Houssos et al., 2012; Chan, 2016; Afonso et al., 2019, Smith et al., 2022) while others METAR data (Stolaki et al., 2009; van Schalkwyk and Dyson, 2013; Weston et al., 2021, Koyuncu et al., 2022). In many cases, the sources of the used observations are not explicit. Among the literature for Argentinian airports, some climatological studies were carried out using SYNOP data (Schonholz, 2015; Ruiz et al., 2018) and others using METAR data (Vasques and Ribero, 2015; Lapido, 2019, Yabra et al. 2021a, 2022).

Recently, Yabra et al. (2021a, 2023) pointed out that difficulty in the analysis and comparison of climatological results between different Argentinian airports is mainly caused by the lack of a comprehensive METAR database. This prompted an exploration into the use of the more complete SYNOP database as an alternative to fill in those hours with missing METAR data or to use the SYNOP database directly. When discussed with local observers and forecasters specialists in aviation meteorology, this last solution aroused certain skepticism mostly based on the argument that SYNOP observations do not take the operational framework into account (Yabra et al., 2021b).

In this context, this work aims to analyze in detail the similarities and differences between METAR and SYNOP data from the perspective of a fog case study. The results presented in this article could help future research on aviation meteorology while taking advantage of the existence of abundant SYNOP data. With that purpose, this work looks forward to determining some of the conditions under which METAR and SYNOP bases can be taken as essentially indistinct and those in which the differences found can have a considerable impact on the results.

This paper is organized as follows: the METAR and SYNOP datasets available at Argentinian airports are presented in section 2. The merging of different databases, the quality control applied and the format unification methodologies are described in section 3. The comparison between METAR and SYNOP data and the impact of using each report type on fog climatology are analyzed in section 4. Lastly, the discussion and the main conclusions of this work are provided in section 5.

2. Data

The National Meteorological Service (SMN) of Argentina has a digitized and curated database of hourly synoptic observations since 1991 in 123 surface stations, built by decoding SYNOP messages. Around 50 of these stations are located within airports and also report METAR messages (Fig. 1). The database acquiring METAR messages has not received the same attention, with data being available only from 2016 and without quality control. Fortunately, other international institutions partially made up for this lack. However, as of today, there is no single database that gathers all data from Argentinian airports in a homogeneous way, which makes it necessary to extract partial information from various databases that are far from complete (Yabra et al., 2023).

The database built using SYNOP messages (henceforth SMN-SYNOPs) followed a quality control process. Contrarily, the database built using METAR messages (henceforth SMN-METAR) contains the complete raw METAR message just as it leaves the weather station and has not received any quality control treatment.

In order to increase the amount of information present in SMN-METAR, METAR reports from different databases were used: the Integrated Surface Data Base (ISD, Smith et al., 2011) from the National Center for Environmental Information (NCEI) of the US National Oceanic and Atmospheric Administration (NOAA), and the Spanish base OGIMET (since 2005. Ballester Valor, G., 2019). The airports selected for this study are those considered most affected by fog according to Yabra et al. (2021a): Ezeiza (SAEZ), Aeroparque (SABE), San Fernando (SADF), Mar del Plata (SAZM), Rosario (SAAR), Resistencia (SARE), Iguazú (SARI), Córdoba (SACO), Neuquén (SAZN), Bariloche (SAZS), Río Gallegos (SAWG), Río Grande (SAWE) and Ushuaia (SAWH). As Yabra et al. (2023) has shown, the availability of METAR data varies over time and individual databases contribute very differently, with most of the data coming from NOAA (85.5%). The study period runs between 2000 and 2019.

SYNOP messages from the same aeronautical meteorological stations (SMN) stored in the SMN-SYNOP database were also used. Figure 2 shows the average time evolution of the number of SYNOP messages available for all stations. It shows relative completeness of the data at all hours and years, except for a small decrease in this amount between 1 and 5 Argentine Official Time (HOA, UTC-3) in the period 2010–2012, and between 10 p.m. and 6 a.m. in the year 2000. In total, 0.4% of data were missing.

As mentioned in the introduction, the same regulations established for synoptic observations apply to aeronautical observations although, for some of the variables, different procedures apply for METAR and SYNOP reports (WMO, 2015; 2018a). These variables and their specifications are detailed in Table 1. Note that only wind direction and speed are represented identically (although coding conventions differ), while the rest of the variables differ from each other even though they refer to the same meteorological condition. This study concentrates on the variables directly linked to fog dynamics: air temperature, dew point temperature, pressure at mean sea level or QNH, visibility, wind speed, wind direction, and present weather.

3. Methodology

3.1. Merging METAR data from different sources

As mentioned above, the availability of METAR data has been maximized by building a single METAR database merging different available sources. First, a single table was created whose registers correspond to all the hours (UTC) existing between 1/1/2000 at 00:00 UTC and 12/31/2019 at 23:00 UTC. Each row was completed with the respective METAR reports provided by the NOAA base. Then, the hours that remained empty were selected and completed with the corresponding METAR provided by OGIMET. This last process was repeated again, completing the lacking hours with the data stored in the SMN. The order of integration of the databases was chosen according to the total amount of information provided by each source in total (see Fig. 2 from Yabra et al., 2023). In this way, the METAR report base was prepared as shown in Fig. 3, for the case of Rosario airport. In this figure, it is also possible to appreciate the number of hours with missing reports that were not found in any source. Figure 4 shows the proportion of data provided by each source at each airport, with NOAA and OGIMET being the ones that make the greatest contribution at each airport, while SMN-METAR barely helps to reduce the amount of missing data. This last database was unable to provide any additional data for most of the airports except for Río Grande's airport (approx. 0.1%), Neuquén ´s and Río Gallegos' (both approx. 0.01%). The proportion of missing data is dependent on the airport and makes it difficult to compare weather characteristics between them, a conclusion already addressed by Yabra et al. (2023). Nevertheless, this amount of missing data does not preclude research based purely on METAR data at airports such as Ezeiza, Aeroparque, Córdoba, Resistencia, Iguazú, and Río Gallegos. Finally, the available METAR reports were decoded using the python Metar module (Pollard, 2005) and the values of the variables mentioned in Table 1 were extracted.

3.2. Quality control of the merged METAR database

Although not always clearly stated in the databases used to build this merged METAR database, it is assumed that raw messages had no quality control at all. The authors of this study had to choose between two options: either to transform this collection of messages into a full-fledged new database with the application of a thorough quality control procedure or to perform a minimal control to satisfy the requirements of this study. We have chosen this second option, which influences in part some of our conclusions. The creation of a curated database of METARs for Argentina will remain an interesting objective for the SMN.

Within the reconstructed database there were miscoded METAR reports, generally due to human mistakes at the time of writing, that were eliminated. Observation values were also controlled regarding the reasonableness of their range, this defined between 0.25% and 99.75% of the distribution of all data from all weather stations. The amount of errors relative to the amount of METAR data is represented in Fig. 5a, and the amount of missing information in each variable relative to the same amount is illustrated in Fig. 5b, both at each airport. The relative frequency of errors does not exceed 0.5% in any variable and airport –what is expected by the conditions of the range imposed–, but its highest values are found in the pressure coding (Table 2). The second variable with the highest number of extreme errors is wind speed and the third-highest percentage of extreme values is found in the RVR.

Regarding the amount of missing data per variable, the only ones that exceed 10% are the RVR and the cloud base height, variables whose report is subject to specific weather conditions. In the first case, the regulation indicates that this variable must be reported if its value or that of visibility is less than 1500 m. In the second case, it is a variable that is only reported in case of cloudiness (it is not necessarily reported at all hours).

3.3. METAR and SYNOP format unification methodology

As mentioned in the previous section, some of the weather data contained in METAR and SYNOP messages are encoded in different ways. Therefore, a way of bringing them to a common scale must be chosen in order to make a fair comparison between them. Temperature, dew point temperature, wind speed, wind direction, and pressure at mean sea level (MSL) values are used just as indicated in the METAR and SYNOP messages because no transformation had been applied. Present weather is also included in this group, but it has to be considered that these codes correspond to different coding families, so the comparison cannot be made solely in a quantitative way. Instead, visibility, cloud base height, and cloud cover values need to pass a transformation process as described below:

- Visibility: the value indicated in the METAR message is used. The SYNOP value is decoded according to the mentioned criteria (Table 1) to obtain the numerical value in kilometers and then is multiplied by 1000 so that its unit becomes meters. For example, if the value 93 is found in a SYNOP message, the corresponding visibility is 0.5 km or 500 m.
- Cloud ceiling: SYNOP value is decoded according to the mentioned criteria (Table 1) to obtain the numeric interval in feet (ft). METAR value is multiplied by 100 and grouped into the same intervals.
- Cloud cover: SYNOP numerical values are grouped according to the METAR coding regulation.

4. Results

4.1. Comparison between METAR and SYNOP data

Figure 6 shows two-dimensional histograms between the METAR and SYNOP values for the measured variables considering all airports: temperature, dew point temperature, wind direction, wind speed, visibility, pressure reduced to MSL, ceiling or height of cloud base and cloud cover. Dew point temperature, wind direction, wind speed, visibility, and cloud ceiling values depict very low dispersion. In the cases of temperature and pressure values a certain dispersion is observed but with a clear dominance of the identity between the variables. On the contrary, cloud cover values show larger differences.

To go deeper into these differences, we show the distributions of the errors of each variable calculated as the difference between the METAR and SYNOP values (Fig. 7). It can be seen that all distributions are centered, and bounded to values close to zero, which represents the equality between both messages. It is interesting to analyze the behavior of those extreme differences that are not found between 0.25% and 99.75%, commonly called “outliers”. The greater number of negative outliers in the distribution of temperature differences indicates higher values in the SYNOP with respect to the METAR. In the case of visibility and wind direction, these extreme differences are equally found in positive and negative values. An exploration of the data indicated the presence of typing errors in the preparation of the METAR, approximation errors in both messages, and extreme values in variables derived from observations (such as pressure reduced to mean sea level). These pairs of values from each variable were deleted from the database.

Another variable of interest in the study of reduced visibility by fog and mist is the present weather since it provides information on the phenomenon responsible for the reduction of visibility. Figure 8 shows the distribution of present weather reported in SYNOPs and METARs reports when a single meteorological phenomenon is reported in the latter (up to 3 phenomena can be reported in a METAR. WMO, 2014). These amounts were normalized by the total number of hours with each code reported in the METAR, thus, all rows add up to 100%. SYNOP codes that did not have cases of fog that matched the fog reported in the METAR were eliminated (note that in the 'FG' row there are no null frequencies). In the same way, METAR codes that did not match the remaining SYNOP codes were also removed (note that all rows have at least one non-zero frequency). In this way, it is possible to study the relationship between both messages focused on the hours with the occurrence of fog (at least reported in aeronautical observations). The first thing that can be appreciated is that the relationships are not direct, that is, in general, a phenomenon in the METAR is not represented by a single code in the simultaneous SYNOP. For example, of the total METAR with fogs (FG), between 5 and 10% correspond to fogs (10), between 20 and 30% to banks or surface fogs (11), between 1 and 5% to fog in the preceding hour but not at the time of observation (28), between 1 and 20% to different types of fog (40-49) and at least 0.1% to drizzle (50-59), rain (60 .69), snow (70-79) and storms (91-99) in the respective SYNOP. Particularly in the cases of mist (BR), confusion with fog can be observed since they coincide in several values. On the other hand, numerous SYNOPs were found whose present weather indicates rain or drizzle (50-69), and the respective METAR indicates mist (BR) and, conversely, METARs that indicate volcanic ash (VA) and the respective SYNOPs indicate mist (10).

Figure 9 highlights the relationships found in Figure 8 but for those codes whose coincidences are more frequent. For this reason, it only includes those SYNOP codes (columns) that contribute at least 1% to the fog code in the METAR (FG) of Figure 8 and those rows that will sum more than 10%. Different values were tried as thresholds, with 1% and 10% producing a reasonable compromise between a substantial simplification and keeping the key features of present weather codes in both data types. In this way, it was found that 85% and 91% of the hours with reduced visibility due to mist and fog reported in the METAR, respectively, correspond to codes 10, 11, 28, 40, 41, 42, 43, 44, 45 and 46 reported in the SYNOP with greater precision in the case of mist (72% reported with code 10). These SYNOP codes were also used in a relatively large amount but in very few cases to report funnel clouds (FC). It is conjectured that this fact may be a product of confusion between the letters 'G' and 'C' when reporting fog (FG) since only 5 reports totally reported this phenomenon.

Some form of moisture indicator is a key element for defining fog and mist detection criteria (see for example Dutta and Chauduri, 2014; Isaac et al., 2020; Zhang et al., 2020). Figure 10 shows the distribution of dew point depression for different visibility intervals in hours with fog and mist reported in present weather in both databases (as 'BR' and 'FG' exclusively in METARs and as 10, 11, 28, 40, 41, 42, 43, 44, 45, 46 and 47 in SYNOPs). An analogous behavior between both types of data and the clear impact of the approximation of the air temperature and the dew point to its integer part in the METAR report can be observed. At least 75% of the reports for any visibility value are at or below 2°C dew point depression in both report types.

4.2. Impact on fog climatologies generated by using different types of meteorological reports

As shown in Figures 8 and 9, the variable present weather appears substantially different depending on which dataset –SYNOP or METAR– is considered. Therefore, a fog hour selection criterion using only this variable would not be completely efficient and other variables may need to be considered to guarantee this equivalence. Given the results discussed in the previous section, it is proposed to analyze whether the climatological characteristics of fogs and mists found from METARs coincide with those detected in SYNOPs from different selection criteria:

- Criterion 1: use only present weather coded as 'BR' and 'FG' in METARs and as 10, 11,28, 40, 41, 42, 43, 44, 45 and 46 in SYNOPs.
- Criterion 2: use an upper threshold for the dew point depression less than or equal to 2°C.
- Criterion 3: use an upper threshold for visibility less than 5,000 m.

To avoid errors associated with the difference in the number of cases between bases, the comparison was carried out considering only the hours in which there are both METAR and SYNOP at each airport. This condition guarantees the same amount of data in both report types (2,010,170 hours in total). Table 3 shows the number of absolute and relative frequencies of fog and mist hours at all airports using each criterion. A greater number of hours with coincidence between METAR and SYNOP can be clearly seen using Criteria 2 and 3. Using Criterion 1 instead, almost twice as many SYNOP reports were selected as METAR, agreeing only in 36% with respect to the first mentioned. The greater number of hours with fog reported in SYNOP compared to METAR could be related to the more extended area that SYNOPs must represent because the present weather reported in the METAR should be representative of conditions at the airport and, just for certain specified present weather phenomena, in its vicinity (as discussed above in section 1). Therefore, the present weather alone does not guarantee the equivalence between hours with fog detected in each type of report. In contrast, Criteria 2 and 3 presented more than 60% of simultaneous hours selected and both absolute frequencies of occurrence are similar. It is necessary to mention that these criteria could include the occurrence of other phenomena. For example, in Criterion 3 visibility could be reduced by any type of precipitation or lithometeor suspension. Similarly, the moisture required in Criterion 2 does not ensure visibility reduction and therefore its number of selected hours is 5 times greater than Criterion 3. The disadvantage of not using the present weather implies paying a hefty price due to selecting hours without fog.

To overcome these limitations, the combination of two or more criteria is also investigated.The combination of Criteria 1 and 3 (Criterion 1+3) was specifically studied because it was the criterion used by Yabra et al. (2023) to build fog climatologies over 13 Argentine airports (same airports as this study). Although the number of hours selected simultaneously is substantially lower than each criterion separately, it selects a similar number of hours in both report types. This new combined criterion is useful to select the hours when fog is present over the airport and affects visibility and therefore flight operations. The mentioned authors found some hours with fog and reduced visibility, but with high dew point depression values, and concluded that sometimes fog banks were in or near the airport, but not over the weather station. Therefore, these visibility and humidity values were not related to the same local weather conditions, and this could lead to erroneous conclusions regarding fog characteristics. Criterion 1+2 was not studied because the hours selected by using it would not necessarily affect aviation activity since there is no restriction on the visibility value. Using the combination of Criteria 2 and 3 (Criterion 2+3), practically the same hours as using Criterion 3 were selected. This means there are only a few hours with dew point depression greater than 2ºC and visibility less than 5000 m, a result already observed in Figure 10. However, this combination could be selecting cases of intense precipitation where visibility is obstructed and air humidity increases due to the evaporation of raindrops, among other examples. To ensure that the phenomenon responsible for the reduced visibility is fog, Criterion 1 is added to the previous combination (Criterion 1+2+3). The inclusion of the present weather reduces approximately to half of the hours that were selected without considering it (Criterion 2+3) and decreases the percentage of simultaneity. Although this value does not exceed 70% in any type of report, the quantity of selected hours is practically similar (note that both represent the same proportion of the total hours). These results suggest that the selection criteria using these 3 variables together guarantee around 60% of simultaneous hours with fog and mist fully reported in METAR and SYNOP. In this instance, it is necessary to study if those selected hours that satisfy each fog criteria using SYNOP data and METAR data have similar climatological behavior.

Figure 11 shows yearly mean frequencies of fog and mist occurrences at all airports using different selection criteria based on METAR and SYNOP. Despite the difference in selected hours shown in Table 3, it can be seen that all criteria produce a similar climatological pattern whose highest frequencies are close to sunrise and are maximized in the winter season (Fig. 11). Criterion 1 does not seem to be reproducing a different regime since the annual and daily occurrence patterns are similar but with a different behavior between summer nights (more SYNOP reports. Fig. 11a) and winter dawns (more METAR reports. Fig 11b), despite having very different amounts of selected hours. In the most restrictive combination of criteria (Criterion 1+2+3), even where similarity in nighttime hours stands out, greater differences in daytime hours during winter months and close to dawn in summer months can be observed.

In addition to mean values at all airports, it is instructive to explore the impact on individual sites. Figure 12 depicts the same as in Figure 11 for the airport of Rio Grande, and Figure 13 the associated wind directions. This airport was chosen because it showed the greatest differences between the climatological wind speed and direction (Fig. 13m,n) and the average wind speed and direction, considering only the hours of fog and mist (Yabra et al., 2023). Particularly at this airport, it is important to mention the low number of observations between 00 and 06 HOA, mainly due to a period in which the airport was not operating during the night hours. METAR and SYNOP selected using Criteria 1 and 3 seem to show similar fog behavior with higher frequencies during cold months (between May and September) without having a preferred time range, except for dawn hours in the case of Criterion 3 (Fig. 12). Although their main wind directions are not the same, they agree in the directional spectrum between northwest and east, which corresponds to onshore wind given the orientation of the Atlantic Ocean coast at Río Grande (see Fig. 1). This agrees with Yabra et al. (2021b), where experts stated that the most common mechanism responsible for fog formation at this airport is advection from the sea. On the contrary, Criterion 2 shows a pattern that seems to be more related to the radiation cycle (Fig. 12) and wind direction and speed are more related to the climatological ones (Fig. 13). The larger number of hours compared to Criteria 1 and 3 suggests that low values of dew point depression are not always related to fog. However, the combined selection criteria allowed to find similar behaviors in the climatological wind roses during the hours with fog or mist from both report types (Fig. 13). Despite slightly lower values in the wind speed by SYNOPs (more noticeable in Criterion 1+2+3), different meteorological report types and the application of similar selection criteria do not imply a change in the wind climatology even using partially different hours (Table 3).

5. Discussion And Conclusions

This work carried out a comparison between two types of hourly meteorological observation reports generated by the same weather stations and by the same personnel: SYNOPs, containing surface synoptic observations that are representative of a wide area around the meteorological station, and METARs, containing surface observations that are representative of the airport area and its near surroundings. SYNOP messages for the 2000–2019 period from 13 Argentinian airports were selected from the database of the SMN, which follows a quality control process. In the case of METAR reports, the SMN has only stored raw data since 2016. This fact led to the use of alternative databases such as those of NOAA and OGIMET. Due to the amount of information provided by each one and the way they were merged, the database that provided the largest amount of data is NOAA, followed by OGIMET. However, in all airports, there were at least 5% of missed METAR reports that were not found in any database. All METAR reports were downloaded as they had been originally written and sent by the meteorological observer and, therefore, there may exist coding errors that were not amended. Preliminary quality control was carried out in order to identify those questionable extreme values in some variables (outliers). Inspection of the data revealed that work remains to be done on the quality control of METARs using the procedures and rules detailed by the WMO (WMO, 1950; 1993; 2010; 2018b).

Although both messages have different purposes and formats, both the observer and meteorological conditions to which they refer are the same. Perfect correspondence between both messages was not found, but a varying degree of similarity depending on the variable: for visibility, values are near identical; for temperature, dew point, pressure and wind, differences in approximations plus some form of error (typing, calculation, conversion) make the differences small but non-negligible. For the case of present weather -- a key variable in this study -- and cloud cover, comparison cannot be performed directly and need some form of manipulation, In this context, the authors attempted to answer which weather report type is more appropriate for fog studies related to aviation activity in Argentina. Our results suggest that this question could be answered in at least two ways considering different points of view. Considering quality control, the information contained in the SYNOP message would be more reliable since the SMN carried out a series of procedures to eliminate possible wrong data. Despite their lack of in-depth quality control, METAR messages respond to specific observations present at the airport and are checked by forecasters and other operators at the time of publication, and can instantly contribute to the correction of errors made in the report. Furthermore, experienced aeronautical meteorological observers report that in operational settings, the observer's subjectivity in key variables is frequently under the pressure exerted by airline members and pilots (Yabra et al., 2021b). This is the case of visibility or cloud ceiling base' reports since these values have a direct impact on airport operations.

Present weather was shown to be a particularly complex variable in the comparison between METAR and SYNOP: SYNOP contains many more categories to describe the weather, but METAR allows for more than one category to be used simultaneously. In addition, METAR is only concerned with the weather conditions at the aerodrome. The differences found between both report types could be due to the real combination of two or more meteorological phenomena occurring at the same time and the limitation of only being able to report one in the SYNOP. To avoid this inconvenience, this analysis was carried out considering only hours in which one present weather code was reported in the METAR, therefore, the same phenomenon would be expected to be reported in the SYNOP. Another possible difference lies in the meteorological phenomena that affect the normal operation of the airport but are not present in the environs. Perhaps the complexity of the observed phenomenon does not allow a simple translation between present weather from METAR to SYNOP or vice versa. This suggests that the "quantification" made by the observer with the available information does not strictly correspond to reality. Nevertheless, limiting the analysis to fog and mist cases, it was possible to find a few present weather codes reported in METARs and SYNOPs that ensure a correspondence of 85% and 91% for mist and fog hours, respectively.

With the objective of evaluating the impact of the differences found between both report types in the fog climatology (such as the one presented in Yabra et al., 2021a, 2023) six detection criteria were built: 3 including separately present weather, dew point depression (as humidity indicator) and visibility, and 3 combining the previous ones. The lowest amount of simultaneous fog hours was detected by the criterion that uses only present weather since it selected twice as many hours in SYNOP data compared to METAR. Therefore, the use of this variable as a fog indicator does not allow to find the same hours between both data types and it is necessary to include other variables to restrict the selection. The combined criterion using humidity, visibility, and present weather restricts the number of cases but ensures that purely fog hours are being selected in both report types. Although this criterion does not select entirely the same hours from METAR and SYNOP, it does guarantee the same climatological characteristics in terms of annual and daily frequency of occurrence, and wind speed and direction.

This work partially confirms the specialists' intuition regarding the clear difference between carrying out studies of aeronautical nature using METAR or SYNOP. The uncertainty remains whether these results are a consequence of data quality and local observers' habits or whether the same thing occurs in other countries. Regarding the meteorological conditions from a specific hour, it seems that the SYNOP cannot replace the METAR to ensure what was happening at the airport. Further, the difference found between METARs and SYNOPs could call into question the rigidity with which we accept the METAR values as ground truths in terms of aeronautical meteorology. However, regarding these datasets, SYNOP data could be used instead of METAR in statistical studies, such as the one discussed here (to extract the general characteristics of fog), but with the caveat of not interpreting the frequencies of occurrence in a too strict fashion.

Declarations

Funding

The author Melina Yabra has received a fellowship from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) from the period 2020-2025 to pursued her PhD studies. The resting authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Melina Yabra. The first draft of the manuscript was written by Melina Yabra and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgments

This work was carried out within the framework of a Doctoral Thesis in Atmospheric and Ocean Sciences of the Faculty of Exact and Natural Sciences of the *Universidad de Buenos Aires*, Argentina, based at the *Servicio Meteorológico Nacional,* Argentina (SMN), and funded by *the Consejo Nacional de Investigaciones Cientificas y Técnicas* (CONICET) of the Ministry of Science, Technology and Innovation, Argentina.

References

1. Afonso, J. M. S., Levit, V., Fedorova, N. (2019). Study of low visibility at Porto Alegre airport: synoptic and thermodynamic processes (*in Portuguese*). *Revista Ibero Americana de Ciências Ambientais*, v.10, n.6, p.131-145, 2019.
2. Baars, J. A., M. Witiw, A. Al-Habash. (2003). Determining fog type in the Los Angeles basin using historic surface observation data. *Proc. 16th Conf. on Probability and Statistics in the Atmospheric Sciences, Long Beach, CA, Amer. Meteor. Soc., CD-ROM, J3.8.*
3. Ballester Valor, G. OGIMET. (2019). https://www.ogimet.com/. Accessed: 31 July 2019.
4. Chan, P.W. (2016). A test of visibility sensors at Hong Kong International Airport. *Weather, 71:* 241-246. https://doi.org/10.1002/wea.2772
5. Dutta, D., Chaudhuri, S. (2015). Nowcasting visibility during wintertime fog over the airport of a metropolis of India: decision tree algorithm and artificial neural network approach. *Natural Hazards 75,* 1349–1368 (2015). https://doi.org/10.1007/s11069-014-1388-9
6. Gultepe, I., G. Pearson, J.A. Milbrandt, B. Hansen, S. Platnick, P. Taylor. (2009). The fog remote sensing and modeling (FRAM) field project. *Bulletin of American Meteorological Society, 90,* 341–359. DOI: 10.1175/2008BAMS2354.1
7. Gultepe, I., Sharman, R., Williams, P.D. (2019). A Review of High Impact Weather for Aviation Meteorology. *Pure Appl. Geophys. 176*, 1869–1921. https://doi.org/10.1007/s00024-019-02168-6
8. Haiden, T., Dahoui, M., Ingleby, B., de Rosnay, P., Prates, C., Kuscu, E., Hewson, T., Isaksen, L., Richardson, D., Zuo, H., Jones, L. (2018). Use of in situ observations at ECMWF. *ECMWF Tech. Memo 834,* 28 pp., https://doi.org/10.21957/dj9lpy4wa.
9. Houssos, E.E., Lolis, C.J., Gkikas, A., Hatzianastassiou, N., Bartzokas, A. (2012). On the atmospheric circulation characteristics associated with fog in Ioannina, north-western Greece. *Int. J. Climatol., 32:* 1847-1862. https://doi.org/10.1002/joc.2399
10. ICAO, 1969. Annex 6: Operation of Aircraft - Part I. International Commercial Air Transport, Aeroplanes. Updated in 2022. https://elibrary.icao.int/reader/290990/&returnUrl%3DaHR0cHM6Ly9lbGlicmFyeS5pY2FvLmludC9ob21lL3Byb2R1Y3QtZGV0YWlscy8yOTA5OTA%3D?productType=ebook&themeName=Blue-Theme
11. Isaac, George A., T. Bullock, J. Beale, and S. Beale. (2020). Characterizing and Predicting Marine Fog Offshore Newfoundland and Labrador. 347-364, J. *Weather Forecasting, Vol. 35.*
12. Koyuncu, R., Deniz, A., & Özdemir, E. T. (2022). Ankara Esenboga International Airport (Turkey) fog analysis and synoptical investigation of the fog event dated 17–19 December 2019. *International Journal of Climatology, 1– 22*. https://doi.org/10.1002/joc.7728
13. Kutty, S.G., Agnihotri, G., Dimri, A.P. (2019). Fog Occurrence and Associated Meteorological Factors Over Kempegowda International Airport, India. *Pure Appl. Geophys. 176,* 2179–2190 (2019). https://doi.org/10.1007/s00024-018-1882-1
14. Lapido, B., 2019. Preliminary study of fog in the airport of the city of Rosario *(in Spanish).* Licenciatura thesis, Departamento de Ciencias de la Atmósfera y los Océanos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.

15. Pollard, T. (2005). Python-Metar Package. https://github.com/python-metar/python-metar
16. Ruiz, J., T. Schonholz, C. Saulo. (2018). Generation of probabilistic visibility forecasts from numerical retrospective forecasts and observations *(in Spanish)*. *Meteorologica Vol 43 N°1 (2018)*, 73-96. http://www.meteorologica.org.ar/wp-content/uploads/2018/07/Ruiz_y-otros_Vol43N1.pdf
17. Schonholz, T., 2014. Development of an objective technique for the generation of probabilistic forecasts of visibility thresholds using retrospective forecasts at the Ezeiza station *(in Spanish)*. Licenciatura thesis, Departamento de Ciencias de la Atmósfera y los Océanos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.
18. Smith, A., N. Lott, R. Vose. (2011). The Integrated Surface Database: Recent Developments and Partnerships. *Bulletin of the American Meteorological Society, 92,* 704–708, doi:10.1175/2011BAMS3015.1
19. Smith, D. K. E., Dorling, S. R., Renfrew, I. A., Ross, A. N., Poku, C. (2022). Fog trends in India: Relationships to fog type and western disturbances. *International Journal of Climatology, 1– 19.* https://doi.org/10.1002/joc.7832
20. Stolaki, S. N., Kazadzis, S. A., Foris, D. V., Karacostas, Th. S. (2009). Fog characteristics at the airport of Thessaloniki, Greece. *Nat. Hazards Earth Syst. Sci., 9,* 1541–1549, https://doi.org/10.5194/nhess-9-1541-2009, 2009.
21. Tardif, R., and Rasmussen, R. M. (2007). Event-Based Climatology and Typology of Fog in the New York City Region, *Journal of Applied Meteorology and Climatology, 46(8),* 1141-1168. https://doi.org/10.1175/JAM2516.1
22. van Schalkwyk, L., and Dyson, L. L. (2013). Climatological Characteristics of Fog at Cape Town International Airport, *Weather and Forecasting, 28(3),* 631-646. https://doi.org/10.1175/WAF-D-12-00028.1
23. Vasques Ferro, R., C. Ribero. (2015). Fog formation at Aeroparque Jorge Newbery *(in Spanish). XII CONGREMET, Mar del Plata, Argentina.* http://hdl.handle.net/20.500.12160/857
24. Veljović, K., Vujović, D., Lazić, L. (2015). An analysis of fog events at Belgrade International Airport. *Theor Appl Climatol 119,* 13–24 (2015). https://doi.org/10.1007/s00704-014-1090-6
25. Weston, M., Temimi, M., Burger, R., Piketh, S. (2021). A Fog Climatology at Abu Dhabi International Airport, *Journal of Applied Meteorology and Climatology, 60(2),* 223-236. https://doi.org/10.1175/JAMC-D-20-0168.1
26. WMO (1950). N°8. Guide to Instruments and Methods of Observation. Updated in 2021. https://library.wmo.int/doc_num.php?explnum_id=11386
27. WMO (1993). N° 305. Guide on the Global Data-processing System (GDPS). Updated in 2001. https://library.wmo.int/doc_num.php?explnum_id=10702
28. WMO (2010). N°488. Guide to the Global Observing System. Updated in 2017. https://library.wmo.int/doc_num.php?explnum_id=4236
29. WMO (2014). N° 731. Guide to Meteorological Observing and Information Distribution Systems for Aviation Weather Services. 2014 Edition. https://library.wmo.int/doc_num.php?explnum_id=8627
30. WMO (2018a). N°49. Technical Regulations, Basic Documents No. 2, Volume II – Meteorological Service for International Air Navigation. Updated in 2021. https://library.wmo.int/doc_num.php?explnum_id=10733
31. WMO (2018b). N°100. Guide to Climatological Practices. 2018 Edition. https://library.wmo.int/doc_num.php?explnum_id=5541
32. Yabra, M. S., R. de Elia, L. Vidal, M. Nicolini. (2021a). Climatological study of visibility reduced by fog and mist in Argentinian airports *(in Spanish)*. *Technical report SMN 2021-106.* http://hdl.handle.net/20.500.12160/1698
33. Yabra, M. S., R. de Elia, L. Vidal, M. Nicolini, R. Vasques Ferro, C. Ribero, L. Chiaparri, E. Fernández, C. Campetella, O. Bonfili, M. Ceballos, G. Barrera, N. Troche, V. López, M. Schizzano, N. Bentancor, L. Berengua, M. Steven. (2021b). Fogs at Argentinian airports: literature review and forecasters' perspective (*in Spanish*). *Technical report SMN 2021-89.* http://hdl.handle.net/20.500.12160/1540
34. Yabra, M. S., de Elía, R., Vidal, L., Nicolini, M.. (2023). Climatological study of fog in Argentinian airports (*in Spanish*). *Meteorológica, Vol 48, N°1, 2023.* DOI: 10.13140/RG.2.2.32543.02728
35. Zhang, J., Pengguo Zhao, Xiuting Wang, Jie Zhang, Jia Liu, Bolan Li, Yunjun Zhou, Hao Wang (2020). Main Factors Influencing Winter Visibility at the Xinjin Flight College of the Civil Aviation Flight University of China. *Advances in Meteorology, vol. 2020, Article ID 8899750*, 13 pages, 2020. https://doi.org/10.1155/2020/8899750

Tables

Table 1: Specifications in the observation and coding of surface meteorological variables in SYNOP and METAR reports in aeronautical weather stations.

Variable	SYNOP	METAR
Atmospheric pressure	Pressure reduced to mean sea level (MSL). Rounded to the first decimal (0.1 hPa)	QNH[1] and/or QFE[2]. Rounded to the unit (1 hPa)
Air temperature	Rounded to the first decimal (0.1°C)	Rounded to the unit (1°C)
Dewpoint temperature	Rounded to the first decimal (0.1°C)	Rounded to the unit (1°C)
Wind (horizontal)	Direction rounded to the tenth (10°) and speed rounded to the unit (1 kt)	Direction rounded to the tenth (10°) and speed rounded to the unit (1 kt)
Present weather	It describes the weather phenomena that are present over the station and its surrounding areas at the time of the observation and during the time between observations. It is coded into 99 categories that can include one or more weather phenomena simultaneously.	It describes the weather phenomena that are present in the airport. It is coded into 22 categories that refer to a single phenomenon, but up to 3 can be mentioned simultaneously in each report.
Past weather	It is coded into 9 categories that can include one or more weather phenomena simultaneously.	Not reported
Cloud cover and cloud type (low)	The cloud type is coded into 9 categories and the cloud cover is rounded to the unit (1/8ths)	Only Cumulonimbus and Towering Cumulus cloud types are reported. The cloud cover is coded into 4 categories based on the number of 1/8ths of the sky that is covered.
Cloud base (low)	The cloud base is visually estimated or measured with a ceilometer (depending on the station) and is coded into 9 categories according to the interval of its height above the station (m) from 0 m to 2500 m.	The cloud base is visually estimated or measured with a ceilometer (depending on the station) and is rounded to the nearest hundred (100 ft). Although 3 groups corresponding to low, medium, and high clouds can be reported, in this work only the lowest layer of clouds (up to 3000 m) is used.
Visibility	The minimum visibility is visually estimated using obstacles and is coded into 99 categories according to its value (km) up to 70 km.	The minimum visibility is visually estimated using obstacles and its value is approximated to the nearest 50 m, 100 m, or 1000 m depending on its magnitude up to 10 km but always reported in units (1 m)
Runway Visual Range (RVR)	Not reported	It is only reported when the visibility or this variable is lower than 1500 m. Its value is approximated to the 50 m or 100 m depending on its magnitude.
Precipitation amount	It is only reported at 12 UTC in Argentina in most of the study period (1 mm).	Not reported in the study period
Precipitation intensity	Not reported	It is included as a component of the present weather in subjective terms.

[1] QNH: It is the pressure at runway level reduced to MSL according to the standard atmosphere.

[2] QFE: It is the atmospheric pressure at the runway level.

Table 2: Thresholds defined for the METAR values quality control.

Variable	Lowest value	Highest value
Air Temperature (°C)	-20	50
Dewpoint Temperature (°C)	-25	40
Pressure (hPa)	950	1040
Wind Direction (°)	0	360
Wind Speed (kt)	0	50
Visibility (m)	0	9999
RVR (m)	0	1500
Cloud base (ft)	0	25000

Table 3: For each criterion: the upper line is the total amount of fog hours selected using SYNOP and METAR data, with the proportion that each amount represents within the total amount of data (between brackets). The lower line is the number of fog hours simultaneously selected in SYNOP and METAR reports, with the proportion that this quantity represents within the total number of cases selected from each report type (between brackets).

		METAR	SYNOP
Criterion 1	Total	85363 (4.2%)	153817 (7.7%)
	Simultaneous	(63%) 53734 (35%)	
Criterion 2	Total	618845 (31%)	515616 (26%)
	Simultaneous	(68%) 420541 (81%)	
Criterion 3	Total	90445 (4.5%)	92049 (4.6%)
	Simultaneous	(67%) 60736 (66%)	
Criterion 1+3	Total	44257 (2.2%)	40217 (2.0%)
	Simultaneous	(55%) 24372 (60%)	
Criterion 2+3	Total	82661 (4.1%)	82973 (4.1%)
	Simultaneous	(66%) 54453 (66%)	
Criterion 1+2+3	Total	42721 (2%)	37874 (2%)
	Simultaneous	(58%) 22984 (61%)	

Figures

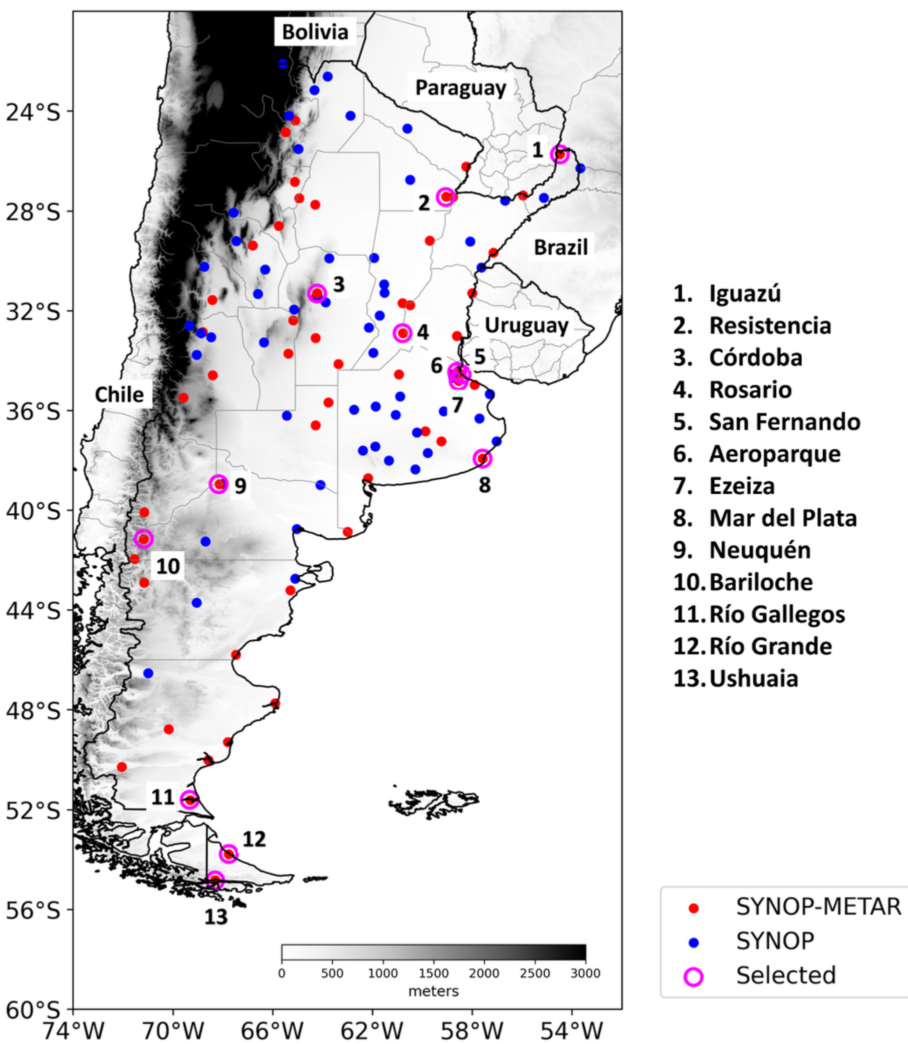

Figure 1

Current surface weather stations from the SMN network that transmit SYNOP reports (red, blue, and magenta markers), those that transmit METAR reports that are later stored in the SMN (red and magenta markers) and from all of them, those that were selected for this study (magenta markers). The topographic information comes from the Digital Elevation Model (DEM) and HydroSHEDS from the SRTM mission (NASA Shuttle Radar Topography Mission) (Rodríguez et al., 2005) with a horizontal resolution of 90 m.

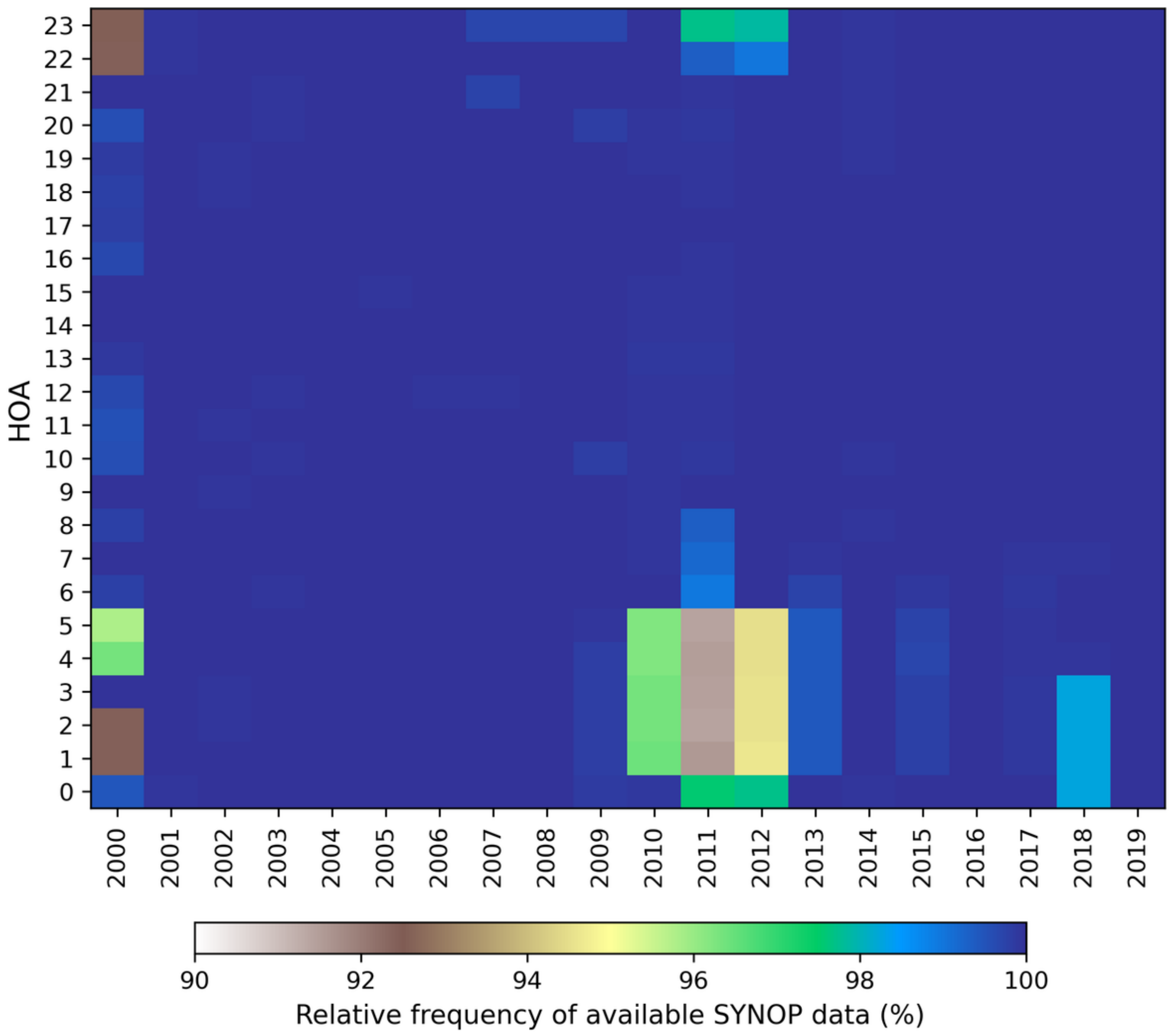

Figure 2

Mean relative frequency (%) of SYNOP reports per hour at Argentine Official Time (HOA, average of the airports studied) generated by the SMN meteorological observers and stored in the SMN database and its annual evolution during the study period.

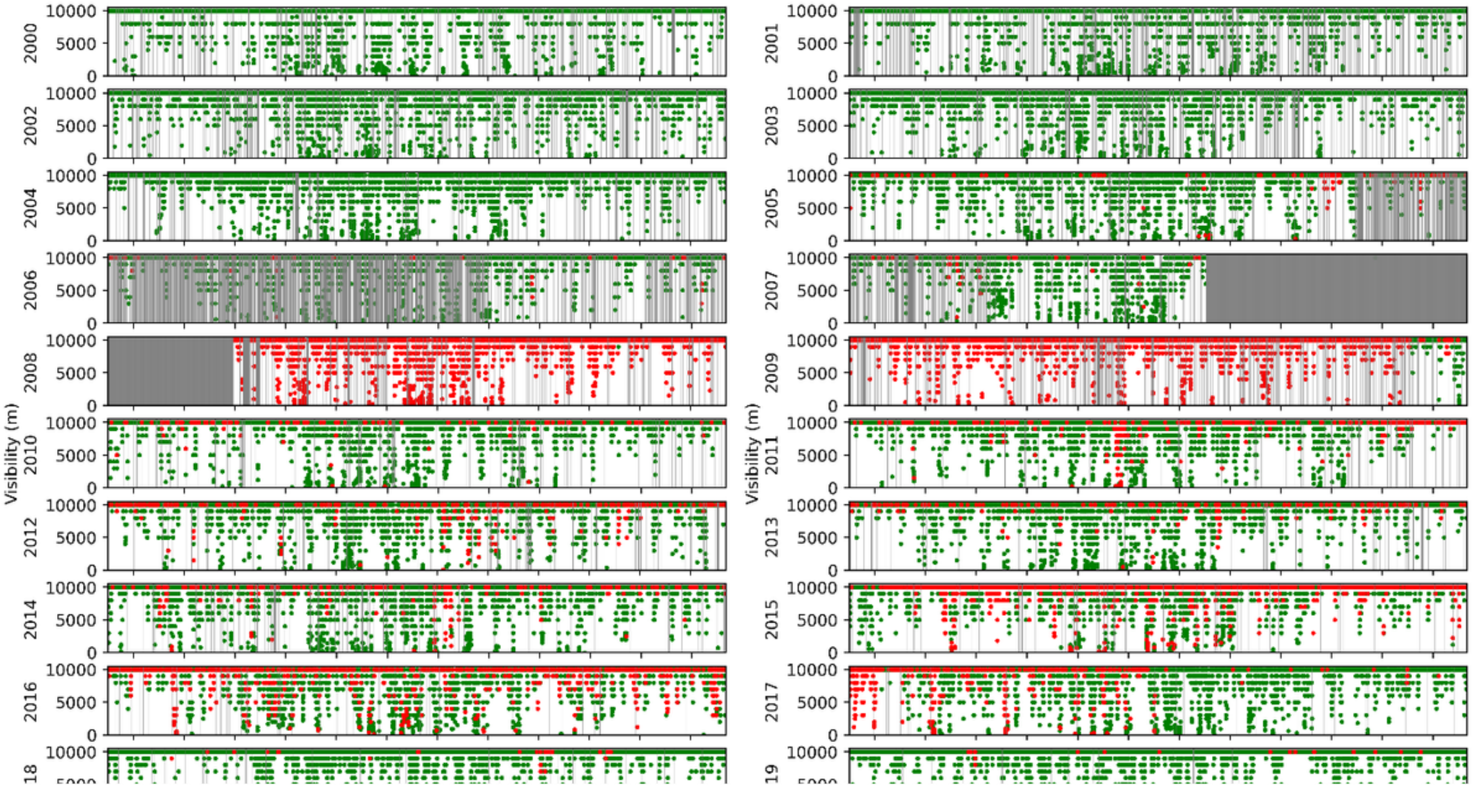

Figure 3

Rosario Airport (SAAR) annual visibility time series (1000 m) extracted from METAR reports provided by the NOAA (green dots, 71%), OGIMET (red dots, 12%), and SMN (black dots, 0%). Gray vertical lines represent hours without information (17%). A summary of the data origin source at other airports can be seen in Fig. 4.

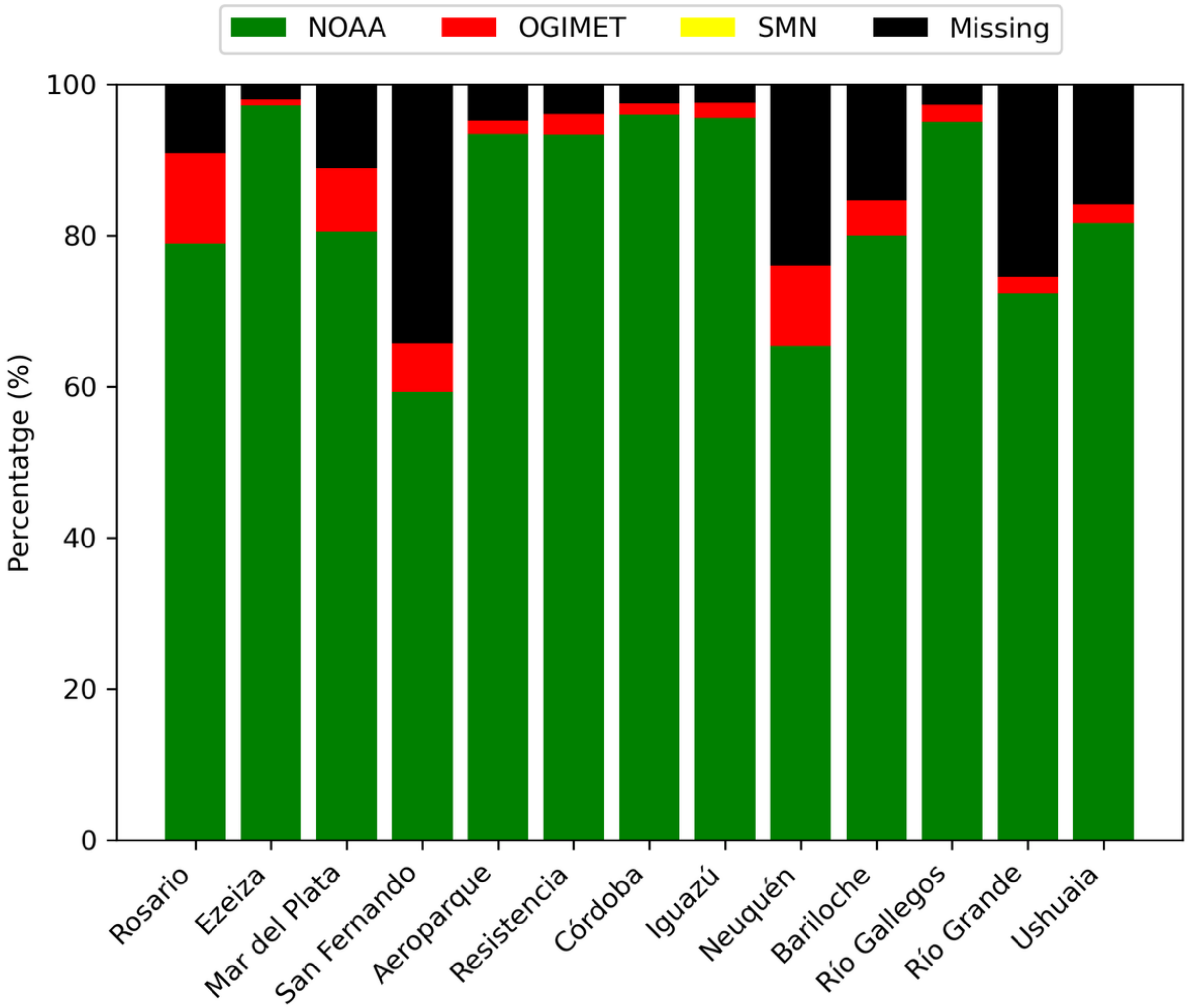

Figure 4

Proportion (%) of the METAR reports provided by NOAA, OGIMET, and SMN sources and of missing data at each airport.

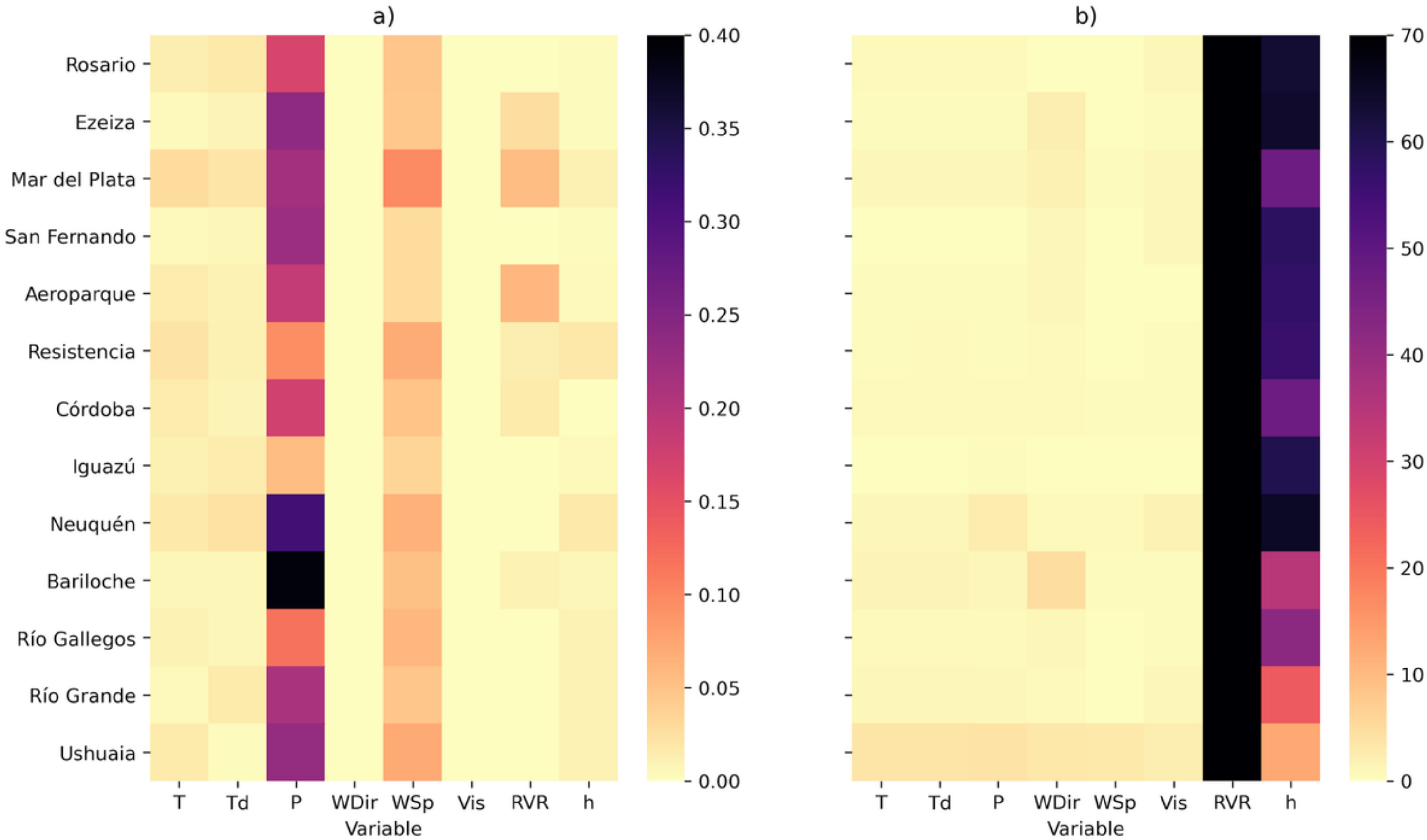

Figure 5

a) Frequency (%) of variable values that do not meet the quality control criteria (described in Section 3.2) relative to the amount of available METARs in each airport. b) Frequency (%) of missing variable values relative to the amount of available METARs at each airport. The ordinate axis lists airports' codes as described in Section 2 and the abscissa axis contains variables: temperature (T), dew point temperature (Td), pressure (P), wind direction (WDir), wind speed (WSp), visibility (Vis), runway visual range (RVR), and cloud base height (h).

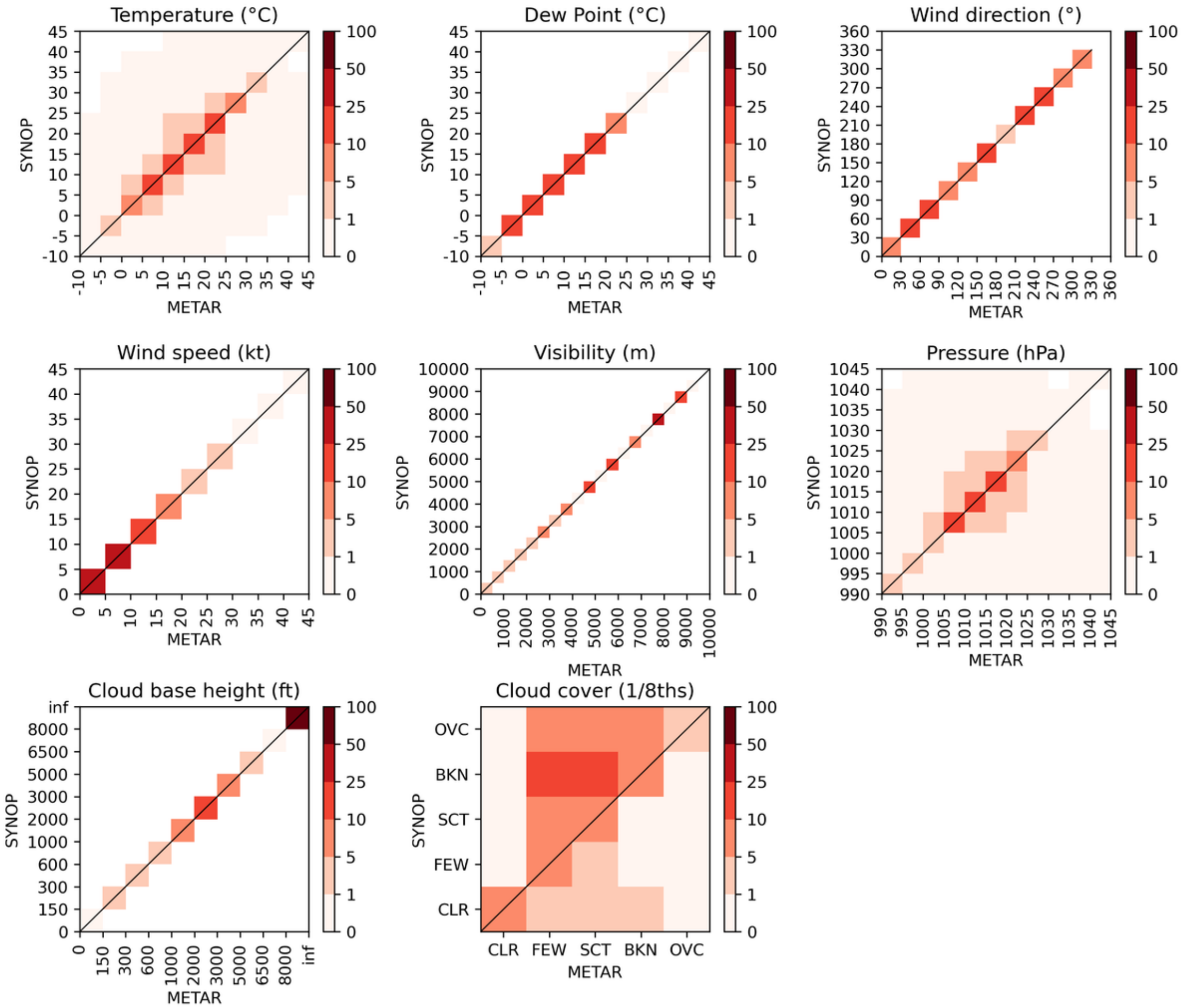

Figure 6

Relative frequency (%) of each variable value reported in METARs and the corresponding value reported in SYNOPs relative to the total amount of hours with simultaneous information available of each variable from both report types.

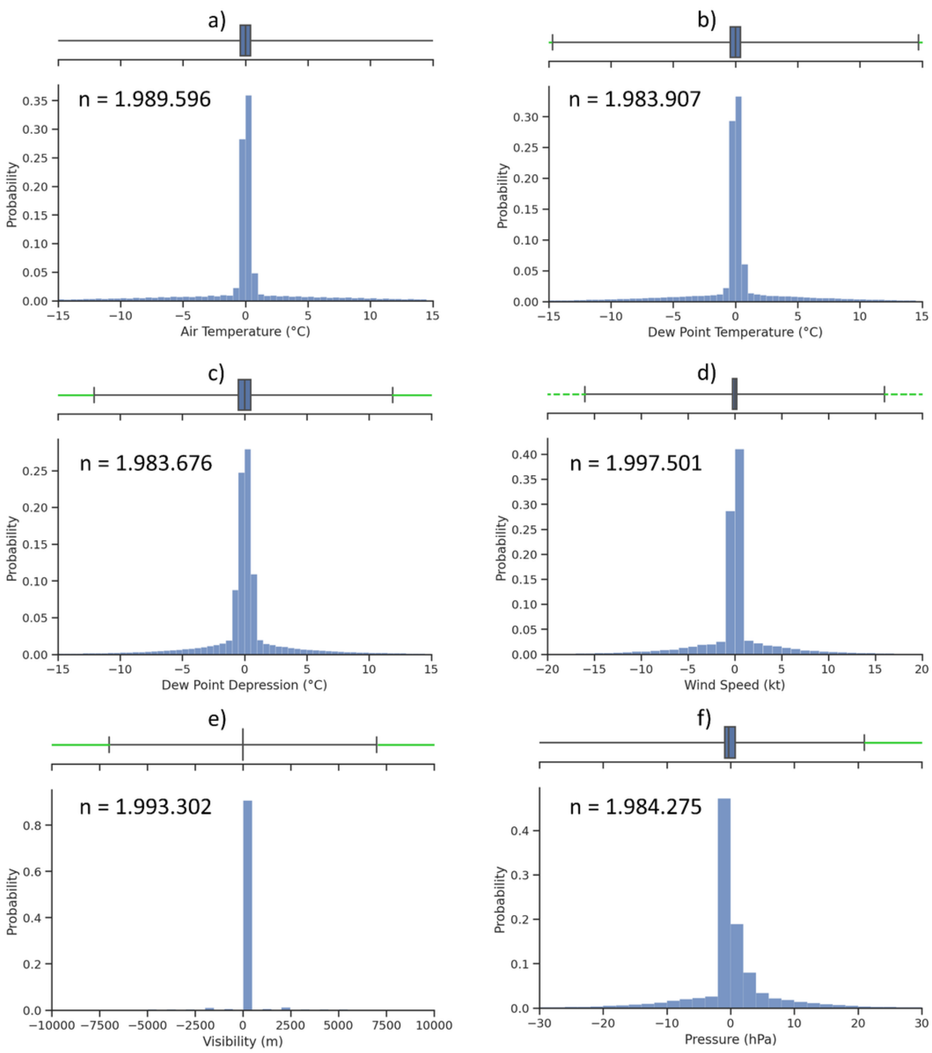

Figure 7

Histogram and Boxplot of the differences between the METAR values and SYNOP values (METAR-SYNOP) for each variable at all airports. The boxplot's box represents the interval between the 25th and 75th percentiles, the gray line inside represents the median (50th), and the brackets outside represent the 0.25th and 99.75th percentiles. The green points outside the brackets are extreme values within the distribution ("outliers"). In the case of e) the entire structure of the boxplot collapses into a single line.

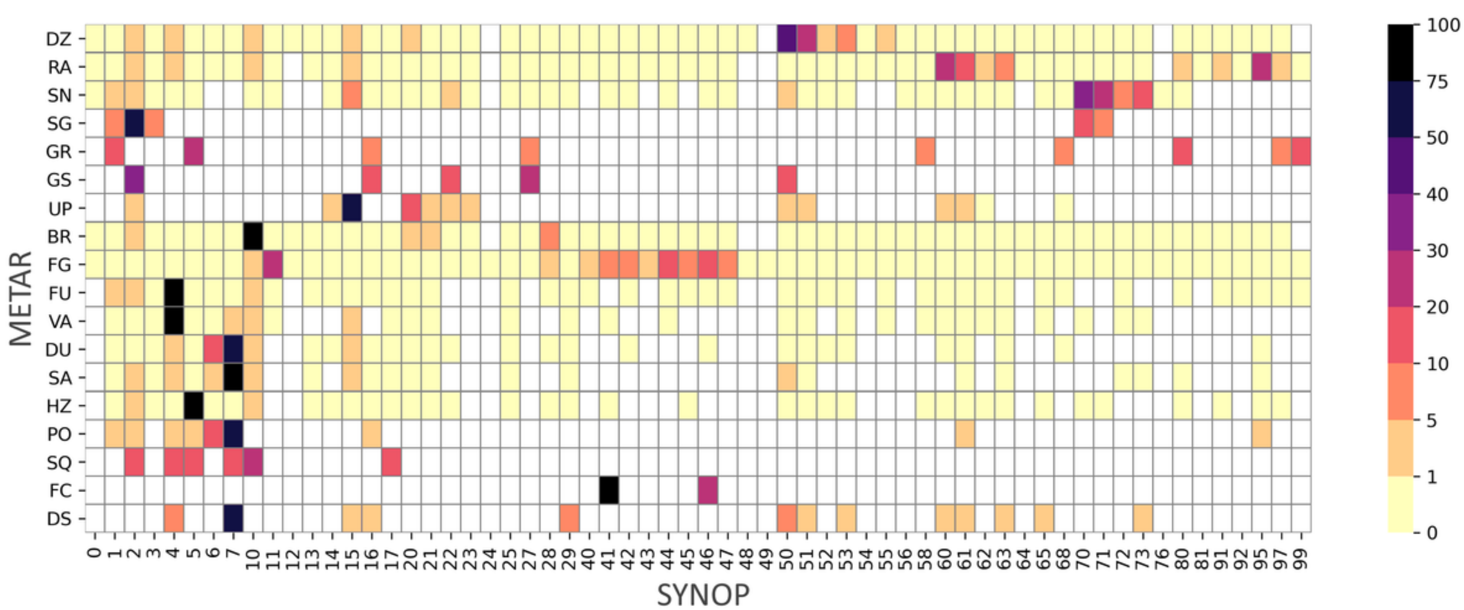

Figure 8

Percentage (%) of hours in which a present weather code was simultaneously reported in the SYNOP (abscissa) and a present weather code in the METAR (ordinate) relative to the total hours with each present weather code was reported alone in METARs (each row adds up to 100%).

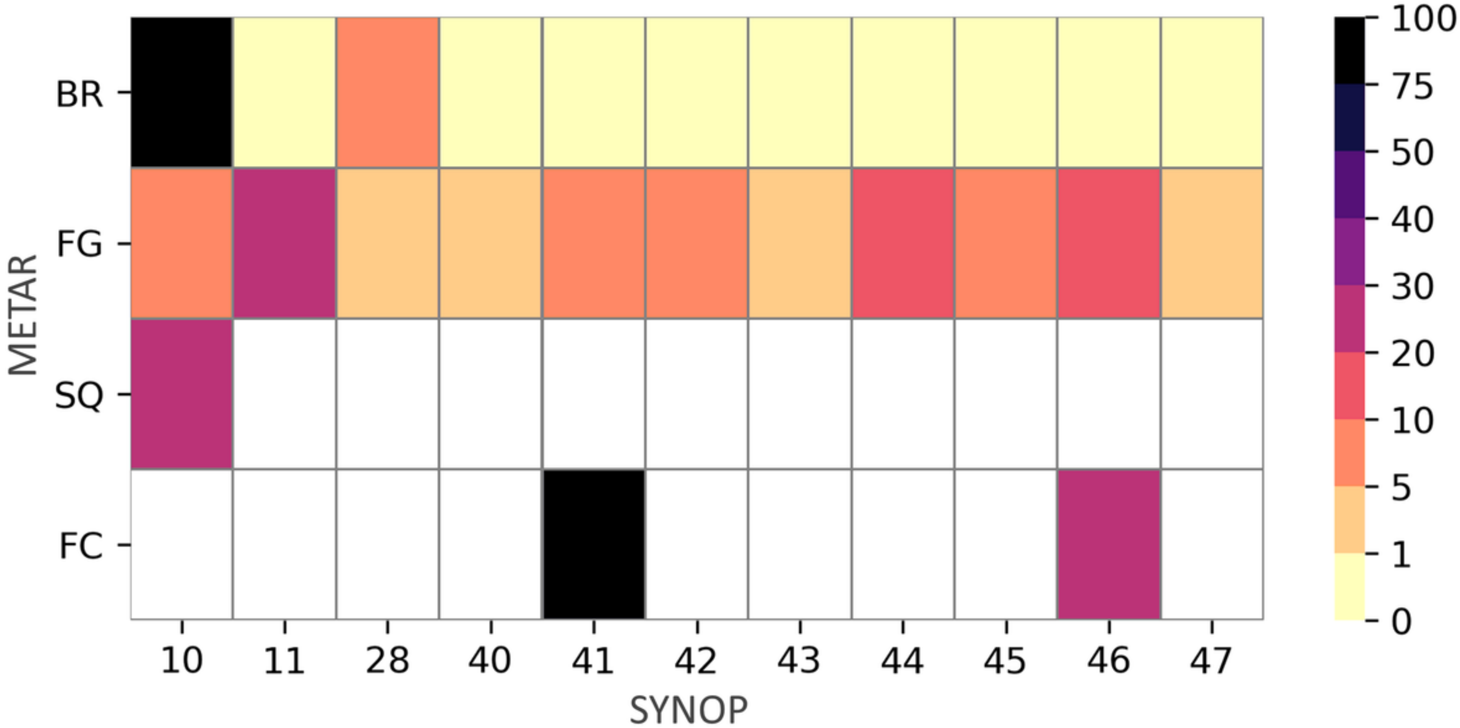

Figure 9

Idem Fig. 8 but reduced to the SYNOP column codes whose frequency coincident with the fog code reported in the METAR ('FG') is greater than or equal to 1% and to the METAR codes whose sum is greater than 10%.

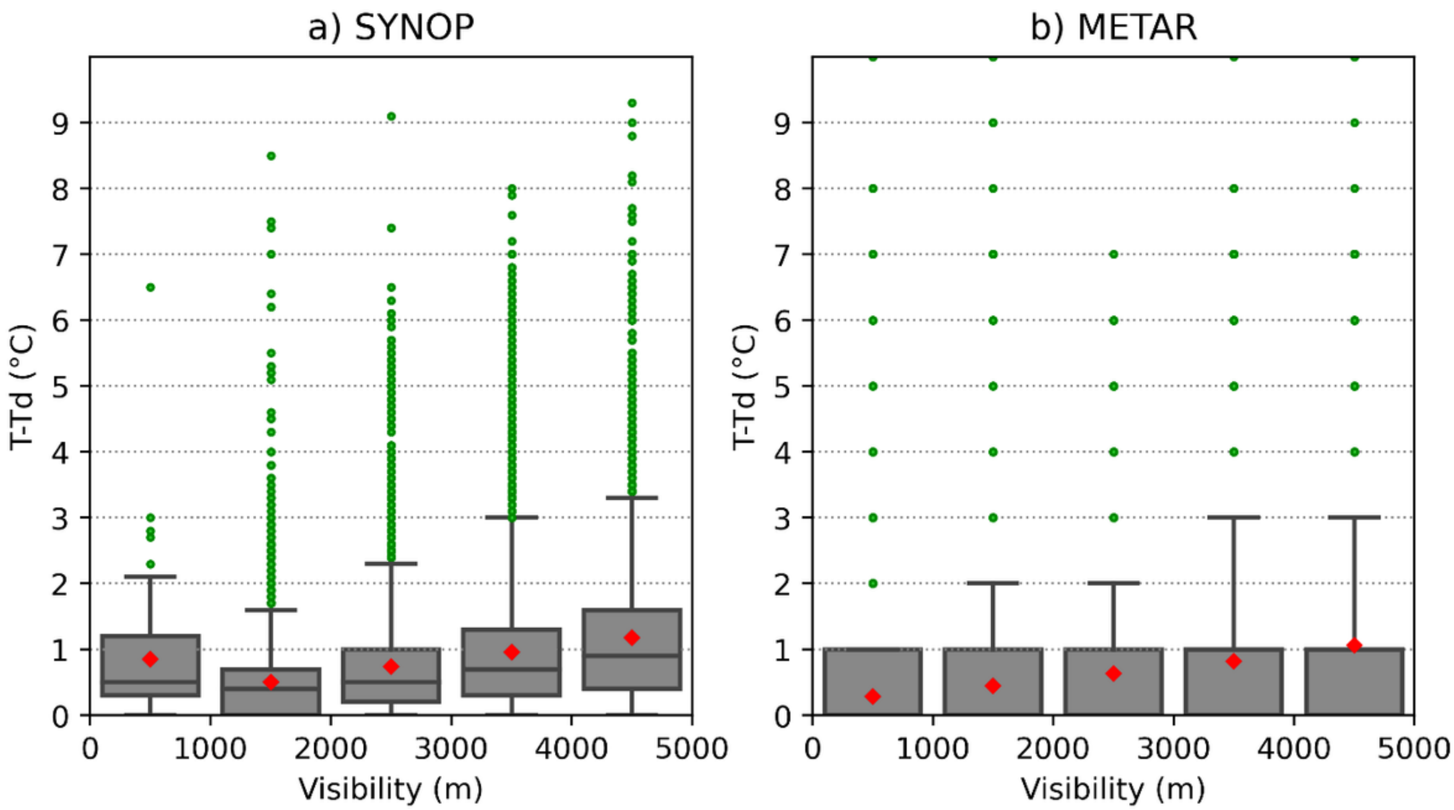

Figure 10

Boxplot of dew point depression distributions split into visibility intervals using fog and mist hours reported in SYNOP (a) and METAR (b) data at all selected airports. The boxplot's box represents the interval between the 25th and 75th percentiles, the gray line inside represents the median (50th), and the brackets outside represent the 0.25th and 99.75th percentiles. The green points outside the brackets are extreme values within the distribution ("outliers") and red diamonds indicate the mean of each distribution.

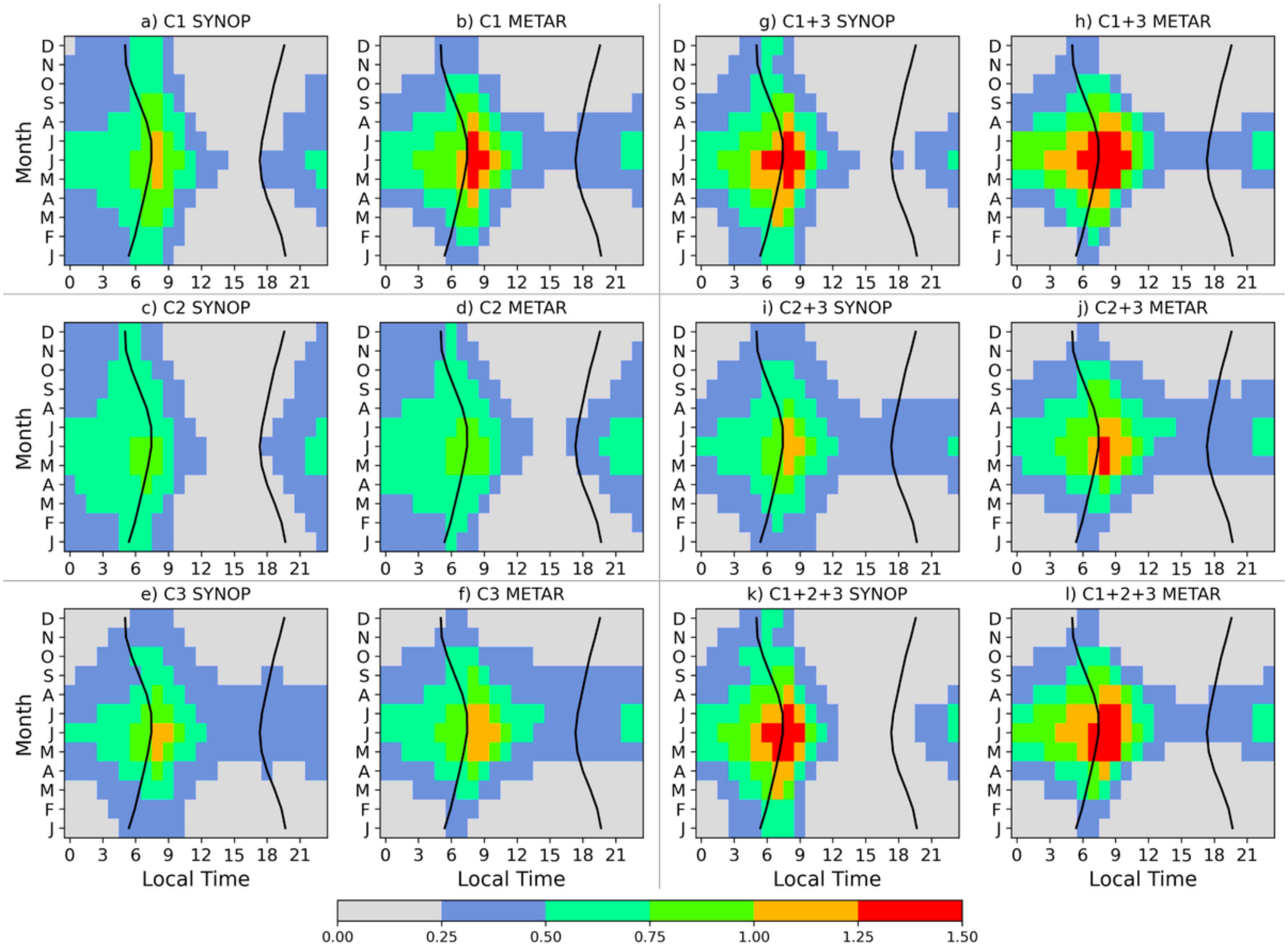

Figure 11

Fog and mist frequency of occurrence (%) using SYNOP (a,c,e,g,i,k) and METAR (b,d,f,h,j,l) data selected using Criterion 1 (a,b), Criterion 2 (c,d), Criterion 3 (e,f), Criterion 1+3 (g,h), Criterion 2+3 (i,j), and Criterion 1+2+3 (k,l) considering all airports. Solid black lines represent the mean sunrise (left) and sunset (right) hours over all airports.

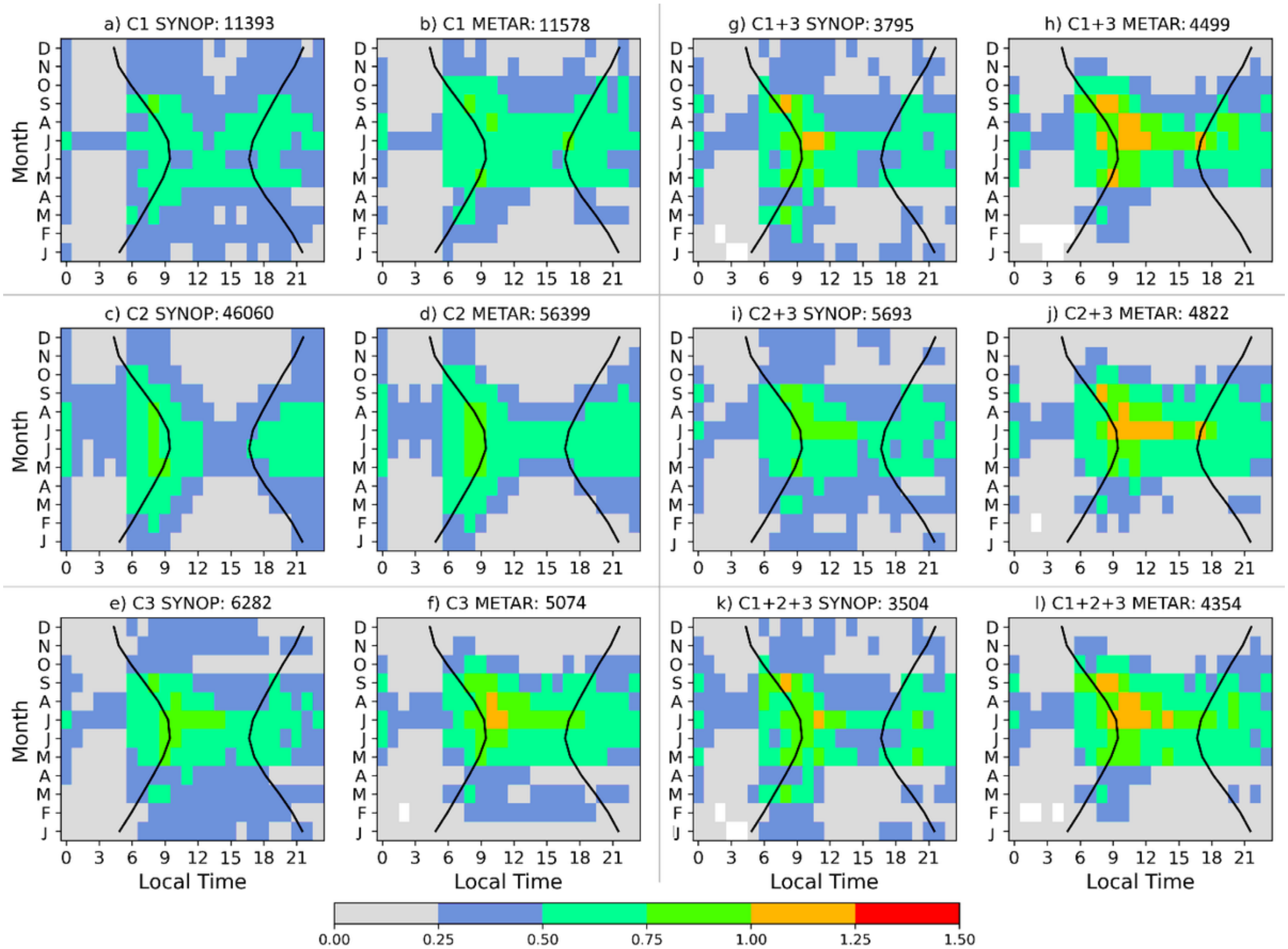

Figure 12

Same as Fig.11 but for Río Grande airport only. Subtitles contain the number of selected hours by each criterion. Solid black lines represent the mean sunrise (left) and sunset (right) hours at Río Grande airport. The lower number of observations between 00 and 06 HOA compared to the resting hours must be taken into account in the analysis.

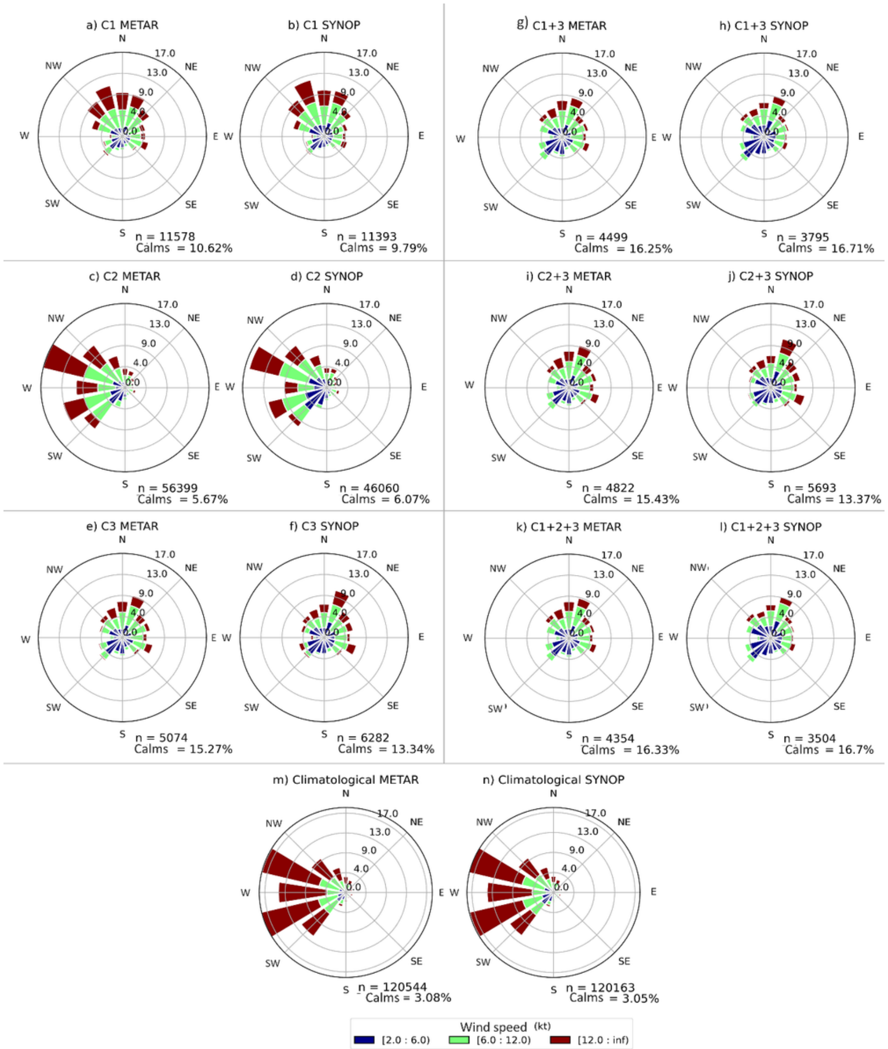

Figure 13

Wind roses at Rio Grande's airport using METAR and SYNOP data selected using Criteria 1 (a, b), 2 (c, d), 3 (e, f), 1+3 (g, h), 2+3 (i, j) and 1+2+3 (k, l). Climatological wind roses using all the available hours are shown in (m, n) subfigures. Wind roses can be interpreted as circular normalized histograms (For example in panel c, 17% of selected hours with Criterion 2 have WNW direction of which 2% have wind speed between 2 and 6 kt, 7% between 6 and 12 kt, and 8% stronger than 12 kt). In the lower right corner of each rose the amount of data used (n) and the percentage of registered calms are specified.