

Automatic Forecasts Algorithms

<https://doi.org/10.31978/639-19-010-0.187>

João Rio¹ (joao.rio@ipma.pt)
Pedro Silva¹ (pedro.silva@ipma.pt)
Ilda Novo¹ (ilda.novo@ipma.pt)

¹*Instituto Português do Mar e da Atmosfera*

SUMMARY

The need for forecasts at an increasingly higher number of locations and time frequency requires a fully automated forecast system. Due to model biases, statistical post-processing is required to improve the quality of the forecasts in variables such as temperature, humidity and wind speed. The automatic forecasting algorithms developed at IPMA provide forecasts up to 10 days and use ECMWF and AROME data.

KEY WORDS: automatic weather forecasting; algorithms; ECMWF high-resolution and ensemble system; AROME.

1. INTRODUCTION

The need for high resolution forecasts in both frequency and locations requires the development of algorithms that allow defining the weather and the most significant meteorological variables using a fully automatic procedure. In the system developed at Instituto Português do Mar e da Atmosfera (IPMA), which has been updated twice, in late 2016 and 2017, several parameters and significant weather are available up to 10 days. Apart from the overall picture of the day, in the first three days hourly forecasts are available to the public; after day 3, three-hourly forecasts are available up to day 6. From day 7 to 10 only the daily view is given.

The automatic forecasts are made for over 400 locations in Portugal, ranging from the main cities, to beaches and mountains. To reduce the bias from numerical weather forecasts, statistical post-processing is used. As the forecast errors increase with lead time, using the deterministic forecasts becomes a limited approach and therefore the system developed uses also the forecasts from the ECMWF ensemble prediction system.

One major issue with automatic forecasting are the high uncertainty regarding precipitation events. To take into account this problem and in an effort to help the end-user understand a given weather symbol, in the latest updates made in late 2017, the probability of 3h precipitation is now made available to the public and a complete revision of the weather symbols was made, including both the meaning and graphical display.

2. FORECAST OF HOURLY AND 3 HOURLY PARAMETERS

In the hourly and 3 hourly frequency forecasts, the 2 m temperature and 10 m wind are provided up to day 6. To decrease the systematic error, statistical post-processing is applied to the direct model outputs from ECMWF high resolution (ECMWF-HRES) and the operational AROME Canonical Model Configuration (TERMONIA *et al.*, 2018) run at IPMA. The forecasts from the latter are available up to 48 h.

The statistical methods used are the Kalman Filter and Multivariate Linear Regression model and their coefficients are updated daily. The training periods have been set to 30 and 40 days, respectively. The statistical models are applied to the direct model output (DMO) of the two available models. In the first 48 h, there are four estimates for a given variable; from day 3 there are only two, because AROME forecasts are not available. The final forecast (STA) is then computed as the average value from all the statistical post-processed values available.

In locations without weather stations, the statistical post-processing cannot be done directly as there are no observations. In this case the forecasts for such locations take into account the statistical adjustments in sites with weather stations, as long as they are close by and have similar topography.

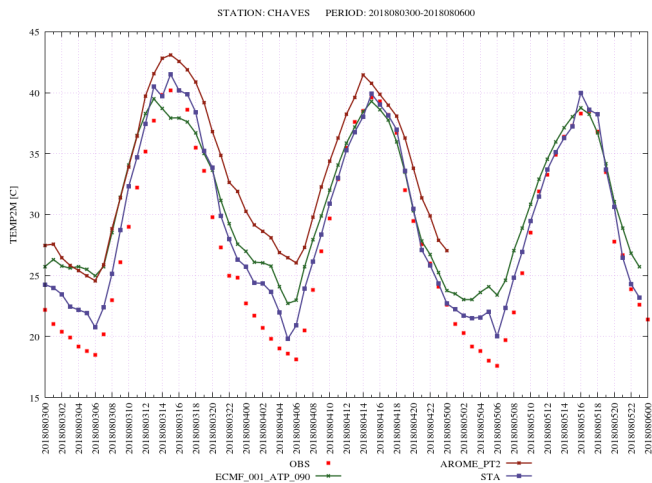


Figure 1. Time series of the 2 m temperature in Chaves, in the period 3rd to 6th August 2018. The values shown are the observations, ECMWF and AROME, as well as the statistical post-processing STA.

Figure 1 shows the hourly timeseries of the 2 m temperature in Chaves, between August the 3rd and the 6th. This is an example of the benefit of the post-processing in an open valley area, where night-time 2 m temperatures drop much more than the models suggest.

To provide the public information about the thermal confort, an algorithm (Rio *et al.*, 2015) based on the formulations of the Wind Chill (OSCZEWSKI and RANDALL, 2005) and the Heat Index (STEADMAN, 1979) was tested and is applied in the whole range of temperatures and wind speeds that occur in Portugal.

3. DAILY EXTREME TEMPERATURES

The forecast of the daily extreme temperatures is done for the period 00-24 h and the procedure done is similar to the one shown in 2. In this case, the DMO variables are the maximum or minimum values of 2 m temperature in 3 or 6 h periods. These values are then processed to get the daily values and this method is applied up to day 10.

However, as the atmosphere is a caotic system, from days 6-7 somewhat excessive variability in forecasts with the same validity time arise. Even though this is inherent to the system and provides valuable information for meteorologists, the public does not understand these differences and hence a system that diminishes this issue was developed.

The system takes into account ECMWF ensemble forecasts (ECMWF-ENS) and the main goal is to limit the STA forecasts to a reasonable distance from the ensemble mean. Figure 2 shows a scheme view of this system which assumes:

- a) up to day 3, the forecast values are the ones computed from the statistical post-processed values;
- b) from day 6 onwards, the forecast values are given by unbiased ensemble mean, with the bias computed from the first 3 days.

As the system is applied to daily values and forecasts are provided with a 3 h frequency in days 4 and 5, the latter values must be adjusted taking into account any changes that may have been required in the transition period.

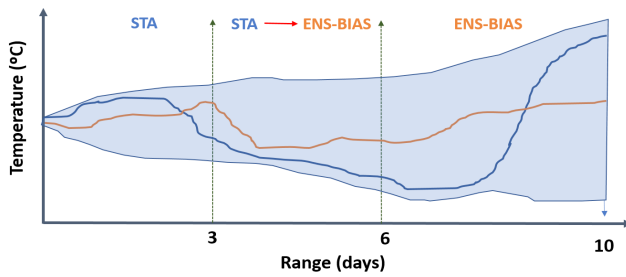


Figure 2. Schematic view of the forecasting system for the 2 m temperature up to 10 days.

Figures 3 and 4 show, respectively, the root mean squared error (RMSE) of minimum and maximum daily temperature computed for 100 locations, in summer 2016. Figures 5 and 6 are similar but show the percentage of correct forecasts (ACCU). A given forecast is considered correct if the forecast is within ± 2 °C of the observed value. In the plots DMO stands for the direct model output from ECMWF, STA stands for the statistical post-processed forecast and STA_NEW is the forecast using ECMWF-ENS as presented above.

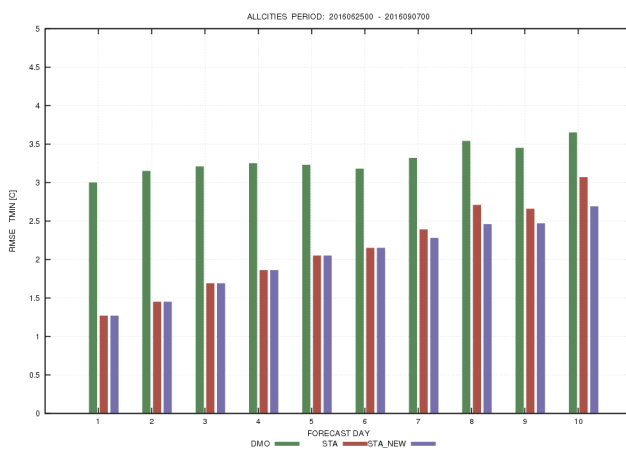


Figure 3. RMSE of the minimum daily 2 m temperature, in summer 2016.

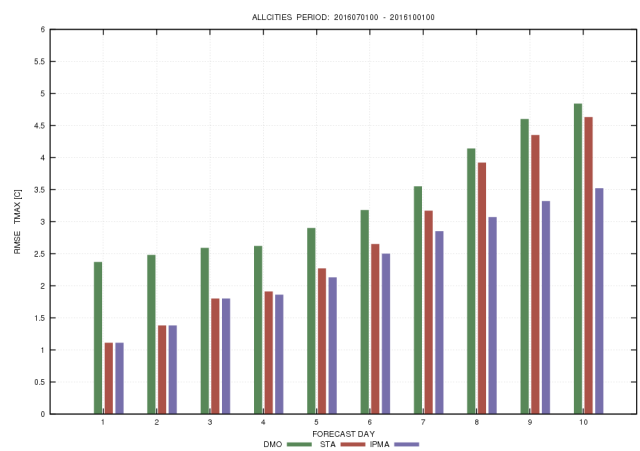


Figure 4. RMSE of the maximum daily 2 m temperature, in summer 2016.

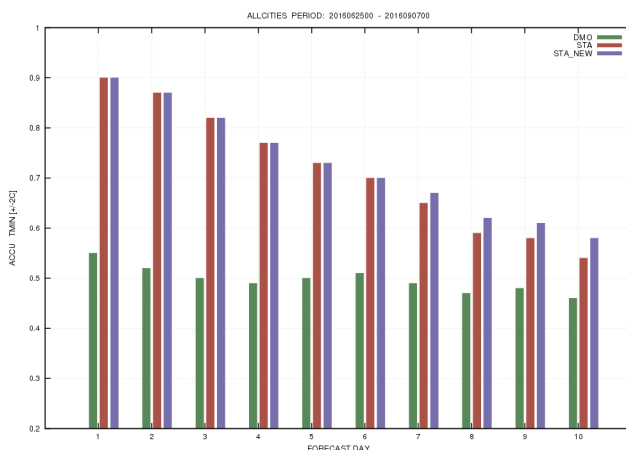


Figure 5. ACCU of the minimum daily 2 m temperature, in summer 2016.

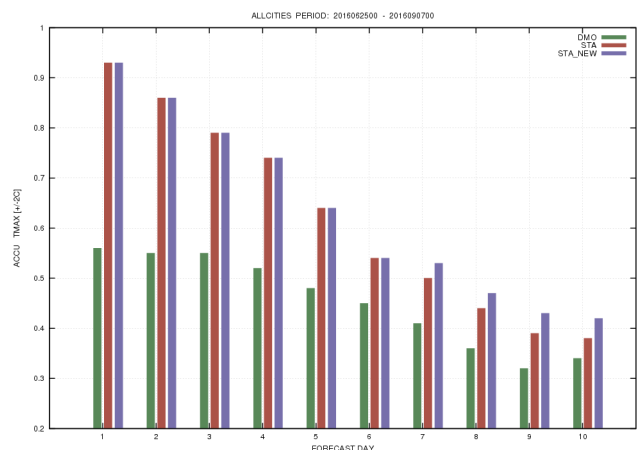


Figure 6. ACCU of the maximum daily 2 m temperature, in summer 2016.

Figures 3 and 4 show the clear improvement of the statistical forecasts when compared to DMO, but this improvement decreases with lead time, as by days 7 to 10 it becomes rather limited. Using the ECMWF-ENS as described above, diminishes the increase of the error in long lead times and so becomes the forecast with the smallest values of RMSE. Additionally, another benefit of this approach is that the variability of the forecasts is much lower than if only DMO or STA were used, which is better understood by the public. In the short-term the percentage of correct forecasts are overall in the 80-95 % range, but this percentage changes considerably as it is a function of the weather conditions, site and time of the year.

4. DEFINING HOURLY AND 3h WEATHER

The high frequency weather is computed in two steps. The first estimate of how the weather will be, is done using exclusively the forecasts from ECMWF-HRES. The variables used include the cloudiness, total precipitation, temperature at the surface and at the 850 hPa level and stability indices.

One of the main problems in defining the weather automatically arises from the precipitation field, as this variable has a high degree of uncertainty. To take this problem into account, the deterministic view is somewhat dropped, as the initial weather estimate is checked against the probability of 3h precipitation from ECMWF-ENS. As an example, if ECMWF-HRES has precipitation over 0.2 mm/3h at a given site but the probability of this amount is below 0.35, then the weather symbol will not have precipitation. This methodology also becomes very useful in reducing the unwanted «jumpiness» in updated forecasts. From late 2017, the probability of precipitation in 3h is made public to help the end-user to identify when it is more likely to have precipitation.

At the present there are 19 weather symbols, ranging from clear skies to rain and thunderstorms. All these weather symbols are available in the hourly/3h forecasts display.

In late 2017 the weather symbols were updated with two main goals: to decrease the difference in symbols that represented similar weather and to revisit the meaning of the symbols. It is worth mentioning that the traditional view of symbols with «rain» and «showers» was dropped, hence now there is only a «rain/showers» symbol. Whether this symbol has a sun or not depends exclusively on the cloudiness of the model forecasts.

5. DAILY OVERVIEW

In the daily view there is information about the overall weather and the wind. As an additional information the probability of precipitation equal or above 1 mm/24h is provided as a guidance to assess how likely is it to have rain/snow. The categorization of the wind speed is usually made using the average speed, but it may not be case if there is high variability during the day. The wind direction is taken as a function of the average wind speed and frequency of each direction.

5.1. Weather symbol: days 1 to 6

The daily weather is computed taking into account all the weather symbols computed in that day, which means a sample of 24 symbols in the first 3 days and 8 in days 4 to 6. The method for choosing the symbol takes into account:

- a) If there is precipitation during that day, then the daily symbol will have precipitation. Otherwise, it will not have precipitation, hence a most severe approach is taken;
- b) In the weather symbols available for precipitation and for no-precipitation, the most frequent approach is taken;
- c) The number of weather symbols available is usually limited to 15, as severe and/or unusual weather symbols are avoided. As an example, in a summer day with clear skies in the morning and occasional thunderstorms in late afternoon, the daily symbol will have the sun, clouds and precipitation, but not the reference to lightning. However, one must be aware that the thunderstorm symbol will be shown in the hourly detail;
- d) In a day with precipitation, whether the daily symbol shows the sun or not depends if the model suggests overcast conditions over most of the day.

5.2. Weather symbol: days 7 to 10

In the medium-range the system assumes that it is meaningless to provide any details regarding the weather at a given site. Therefore, the weather is defined exclusively using the probability of precipitation in 24 h periods, provided by the ECMWF-ENS system, which is also beneficial as this approach provides a weather map with limited spatial variability.

The number of weather symbols available is limited to six, but usually only four are shown. The symbols allowed do not show any severe weather and are usually in the medium range, that is «partly cloudy» and the traditional «showers».

Figure 7 shows the probability of 24 h precipitation equal or above 0.8 and 10 mm/24 h (left/right). Figure 8 shows an example with a forecast for day 8, from the 00 UTC run from 1st November, valid on 8 November 2018. The map, which is not made public, shows a smooth field, with the blue color meaning overcast and rain, the two green colors meaning the partly cloudy and some rain, the grey representing the partly cloudy and the yellow few clouds.

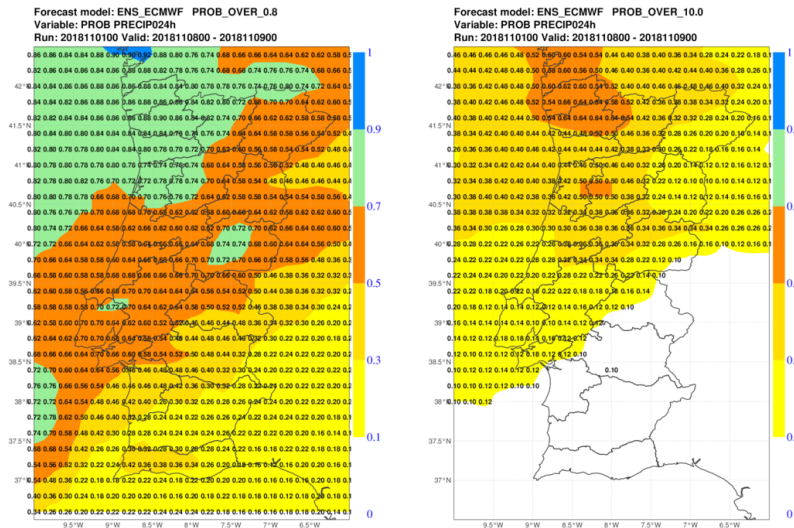
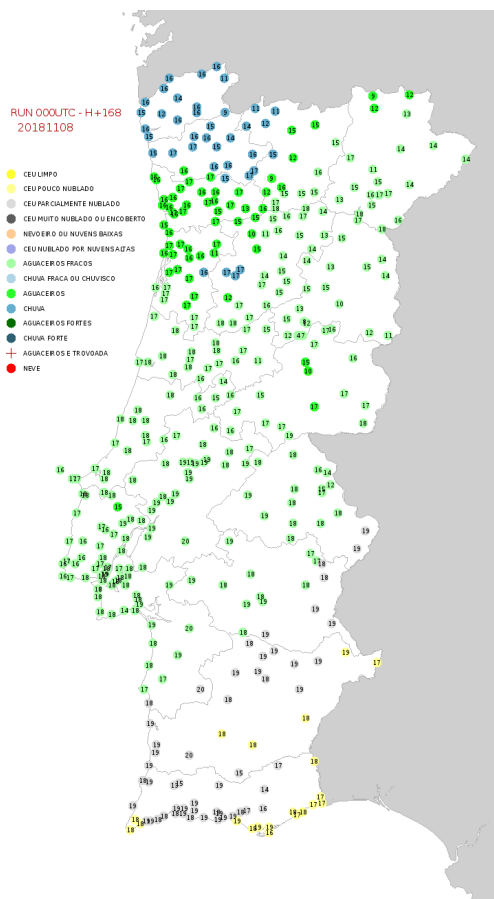


Figure 7. Probability of precipitation equal or above 0.8 and 10 mm/24 h (left/right). Forecast run from 00 UTC November 1st, valid on 8 November 2018.



6. CONCLUSIONS

After the update in late 2017 the automatic forecasting system became more robust and coherent. This is due to the new graphical weather symbols and their meaning, as well as giving the end-user the probability of precipitation in 3h periods.

Also since 2017, the verification of the daily forecast temperatures, made with 1, 4 and 7 days in advance, is available in IPMA’s website. Along with a revised forecast disclaimer, it helps the general public to understand what a forecast is and what it may give the user.

Further developments are being assessed for the near-term, namely using AROME forecasts in the weather symbols and new variables from ECMWF, such as precipitation type and maximum rate and lightning flashes.

Figure 8. Example of daily weather for forecast day 8. Forecast run from 00 UTC November 1st, valid on 8 November 2018.

ACKNOWLEDGMENTS

The authors would like to thank Vanda Costa for her remarks and review of this document.

REFERENCES

STEADMAN, R. G., 1979. The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. *J. Appl. Meteor.*, 18, 861-873.

OSZCZEWSKI, R. and BLUESTEIN, M., 2005. The new wind chill equivalent temperature chart. *Bulletin of the American Meteorological Society*, 86 (10): 1453-1458. Bibcode: 2005BAMS...86.1453O. doi:10.1175/BAMS-86-10-1453.

RIO, J., NOVO, I. e SILVA, P., 2015. Avaliação de uma metodologia de cálculo da Temperatura Aparente (versão 2015). Relatório DivMV 14/2015, Instituto Português do Mar e da Atmosfera.

TERMONIA, P., FISCHER, C., BAZILE, E., BOUYSSSEL, F., BROŽKOVÁ, R., BÉNARD, P., BOCHENEK, B., DEGRAUWE, D., DERKOVÁ, M., EL KHATIB, R., HAMDI, R., MAŠEK, J., POTTIER, P., PRISTOV, N., SEITY, Y., SMOLÍKOVÁ, P., ŠPANIEL, O., TUDOR, M., WANG, Y., WITTMANN, C. and JOLY, A., 2018. The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1, *Geosci. Model Dev.*, 11, 257-281, <https://doi.org/10.5194/gmd-11-257-2018>.