



# Diurnal temperature cycle deduced from extreme daily temperatures and impact over a surface reanalysis system

F. Besson<sup>1</sup>, E. Bazile<sup>2</sup>, C. Soci<sup>2</sup>, J.-M. Soubeyroux<sup>1</sup>, G. Ouzeau<sup>1</sup>, and M. Perrin<sup>3</sup>

<sup>1</sup>Meteo-France, DCSC/AVH, Toulouse, France

<sup>2</sup>Meteo-France, CNRM/GAME, Toulouse, France

<sup>3</sup>University of Toulouse III – Paul Sabatier, Toulouse, France

Correspondence to: F. Besson (francois.besson@meteo.fr)

Received: 14 January 2015 – Revised: 29 May 2015 – Accepted: 09 June 2015 – Published: 30 June 2015

**Abstract.** Due to the evolution of the observation network, hourly 2 m temperature analysis performed by reanalysis systems shows temporal inhomogeneities. The observation network gap is less present for extreme daily temperature observations. In order to reduce inhomogeneities and enable a climatological use of temperature analysis, information from extreme temperatures could be useful. In this study, the diurnal temperature cycle has been reconstructed for stations which only record extreme temperatures. These new “pseudo” hourly temperature observations are then provided to the analysis system. Two methods have been used to deduce hourly temperatures from extremes and compared to real observations. The results have shown that using those new pseudo-observations as an input for two different reanalysis systems enables reducing the bias in temperature analysis.

## 1 Introduction

Observation network heterogeneity (both in time and space) has to be addressed in order to produce unbiased climatological indexes (Peterson and Vose, 1997). Focusing on France, a big gap in the hourly 2 m temperature observation network occurred in the early 1990s related to the automatic station deployment. From the end of the 1950s to the end of the 1980s, the number of hourly temperature observations available ranges from about 200 to about 500, and then it increases and reaches 2000 in the present day; the number has increased by a factor of 10 in 55 years.

As a consequence the hourly 2 m temperature analysis, performed by SAFRAN (Durand et al., 1993, 1999), and using those hourly temperatures, shows inhomogeneities (Vidal et al., 2009; Soubeyroux et al., 2011). The change in observation network density is not so steep for the extreme daily temperatures (6-fold increase in 55 years, evolution quite linear). So we have tried to extract the information coming from extreme temperatures to deduce hourly data and reduce the gap.

Thus, the diurnal temperature cycle has been reconstructed and used as pseudo-observations for stations recording only extreme temperatures employing two methods. The first one only uses extreme temperatures and geographical characteristics of the station, whereas the second one needs extreme temperatures and hourly observation from the neighbourhood. Those two methods will be briefly described in Sect. 2 and compared in Sect. 3.1. Comparison of reanalyses performed with pseudo-observations using respectively SAFRAN and MESCAN (Soci et al., 2013) analysis systems are shown in Sect. 3.

## 2 Methods

### 2.1 Temperature temporal downscaling methods

#### 2.1.1 Astro method

This method is described in Reicosky et al. (1989) as WCALC. We have called it the astro (astronomical) method because we deduce the hourly temperature at the station from

extreme temperatures and sunrise/sunset hours. The temperature diurnal evolution is subdivided into three segments: from midnight to sunrise +2 h, for daylight hours and from sunset to midnight. For each of these segments coefficients are computed depending on extreme temperatures and sunset/sunrise hours (of the day and/or previous/following days).

The advantage of this method is that very few data are needed and it can be applied whatever the density of the network. However, it is not suitable for days with atypical diurnal cycle (Perrin et al., 2013).

### 2.1.2 Alpha method

$$T(h) = T_{\min} + \alpha(h)(T_{\max} - T_{\min}) \quad (1)$$

We assume that hourly temperature can be defined as in Eq. (1) with  $T(h)$  being an hourly temperature,  $T_{\min}$  the minimum temperature,  $T_{\max}$  the maximum temperature and  $\alpha(h)$  an hourly coefficient varying from 0 to 1. This equation can be applied for all station with hourly data.

The aim of this method is to determine the  $\alpha(h)$  coefficient for stations without hourly data. This can be done by using average values of  $\alpha(h)$  available for stations with hourly data in the neighbourhood. Thus with this  $\alpha(h)$  value we can apply Eq. (1) and get a pseudo hourly observation of temperature for a station without hourly observation.

A previous study (Perrin et al., 2013) of the spatial correlation of the  $\alpha(h)$  showed that a criterion is to be used in order to select stations in the neighbourhood. Thus we only keep in the neighbourhood stations with a difference of elevation lower than 200 m.

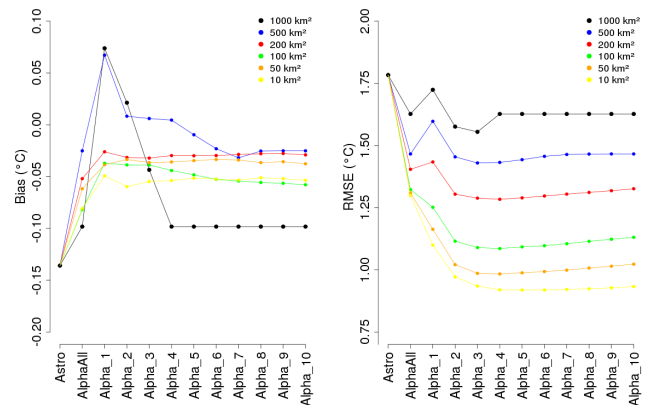
Furthermore we can vary the number of stations selected in the neighbourhood. In our study several experiments are performed:

- one experiment (called Alpha\_All) with all the stations filling the elevation criterion;
- experiments with the  $N$  closest stations filling the elevation criterion (called Alpha\_N), with  $N$  varying from 1 to 10.

Thanks to this method we have quite similar quality even if the diurnal cycle is atypical (Perrin et al., 2013). The inconvenience is that, unlike the astro method, we need a network with hourly observations to apply this method.

### 2.1.3 Reanalysis systems

Two different reanalysis systems have been used: SAFRAN and MESCAN. Both use a guess and observations to perform an optimal interpolation. MESCAN does it on a regular grid of 5.5 km, whereas SAFRAN does it on 615 climatologically homogeneous area with a 300 m vertical gradient. In our study SAFRAN uses as a guess ECMWF's operational archives projected on its analysis points (Quintana-Seguí et al., 2008), and MESCAN uses a HIRLAM guess downscaled from 22 to 5.5 km.



**Figure 1.** Bias (left) and RMSE (right) of pseudo-observations for hourly temperatures using different methods (x axis) and different network densities (colours). If we only have  $N$  hourly data usable for alpha method, the pseudo-observation for Alpha\_M (with  $M > N$ ) is set to the Alpha\_N's value.

## 3 Results

### 3.1 Comparison of temporal downscaling methods

The comparison of the astro and alpha methods is performed over a 12-month period from August 2009 to July 2010 every 6 h, at 318 independent hourly observations stations over France. We performed six experiments with different densities (keeping one observation per 10, 50, 100, 200, 500, 1000 km<sup>2</sup>) for hourly observations available to deduce pseudo-observations at the 318 stations. Thus we can evaluate the astro and alpha methods over different network densities.

In Fig. 1 we notice that, whatever the network density used, alpha methods perform better than the astro method (lowest bias and root-mean-square error (RMSE) values for alpha methods).

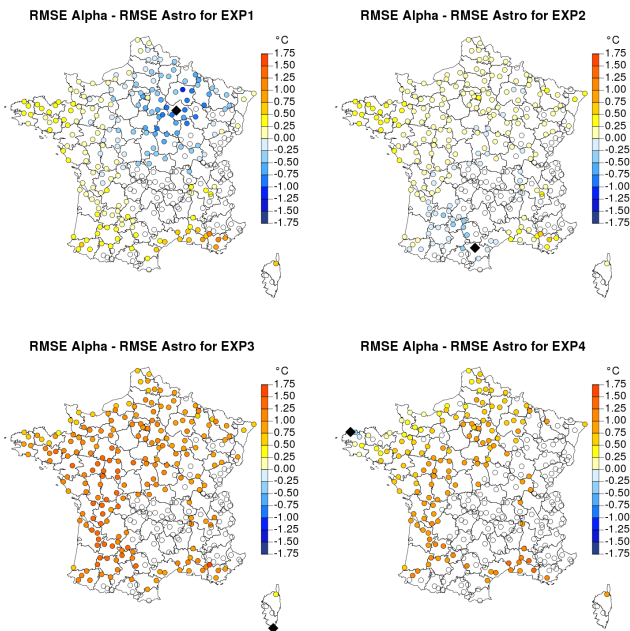
Focusing on alpha methods in Fig. 1, we observe that

- all of them improve when the network density increases (fairly constant low bias and increasing low RMSE);
- using only one neighbour degrades the pseudo-observations compared to the use of two or more neighbours;
- with a high-density network ( $< 200$  km<sup>2</sup>) the Alpha\_all experiment gets the worst results because we use information coming from the whole area, whereas when reducing the number of neighbours we keep local features.

Then we did four extreme experiments using only one hourly observation to deduce pseudo-observation with alpha method. Evaluation is done over the 318 independent stations. For all experiments, bias and RMSE values over the whole domain are better for astro (or quite similar to alpha

**Table 1.** Bias and RMSE of pseudo-observation for the four extreme experiments for astro and alpha methods.

	Exp. 1		Exp. 2		Exp. 3		Exp. 4	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Astro	-0.18	1.76	-0.18	1.76	-0.18	1.76	-0.18	1.76
Alpha	0.27	1.78	-0.15	1.87	-0.36	2.8	-0.19	2.5



**Figure 2.** Comparison of the RMSE between alpha and astro for the four extreme experiments over 318 independent stations (circle). RMSE alpha–RMSE astro is represented by colours of the circles’ background: cold (warm) colours show that alpha is better (worse) than astro. White circle for station where the alpha method cannot be applied (because of the elevation criterion). The black diamond is the location of the hourly observation used to deduce pseudo-observation.

method); see Table 1. Looking at the geographical repartition of the differences of RMSE between the two methods (see Fig. 2), we notice that for Exp. 1 and Exp. 2 where hourly observations are located on a plain far from the sea, close to the station, alpha is better than astro, but the further away the better is astro. Whereas for Exp. 3 and Exp. 4 with observation location along the coast (with specific weather type), astro is almost always the best.

### 3.2 Impact over reanalysis systems

We performed the reanalysis over France and in the period from October 2009 to June 2010. We did it using real observations and then using pseudo-observations from three selected methods: Alpha\_All, Alpha\_5 and astro. We added a criterion for alpha methods regarding the selection in the

**Table 2.** Bias and RMSE of pseudo-observation at 1555 observation points.

	Pseudo-obs.		
	Astro	Alpha_All	Alpha_5
Bias	-0.09	0	0
RMSE	1.73	0.9	0.85

neighbourhood: to be selected, an observation has to be less than 100 km from the pseudo-observation location.

Results are finally interpolated over an irregular grid built with 1555 observations’ location. For SAFRAN, as the temperature analysis is done in a homogeneous area in increments of 300 m of elevation, we can do a linear interpolation between those levels, whereas for MESCAN we determine the gradient using the 4 closest model grid points.

Table 2 shows characteristics of statistic errors for pseudo-observations used as input for analysis systems; results are fairly similar to those obtained in Sect. 3.1. We just notice that the new criterion on the maximal distance improves the results of Alpha\_All experiments. In Table 3, statistics for SAFRAN and MESCAN using those pseudo-observations and the real observations, and additional statistics for MESCAN’s guess, are presented. Both reanalysis systems have low bias values; they slightly increase RMSE values, but experiment performance is sorted in the same way before and after the reanalysis systems: alpha methods give better results compared with astro.

Using Alpha\_5 or Alpha\_All pseudo-observations gives results very close to the experiment with real observations. Furthermore bias and RMSE coming from the guess are reduced whatever the pseudo-observations used.

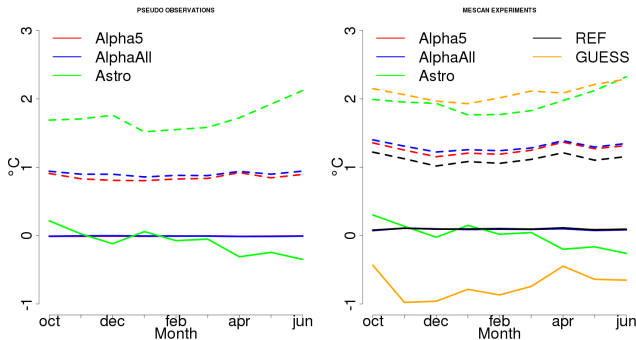
Looking at error statistics month by month (see Fig. 3), we notice that bias and RMSE for the astro method (for both input pseudo-observation and output of MESCAN) is monthly dependent. This is not the case for alpha methods, which show rather stable statistics. Results using SAFRAN are similar (not shown here).

## 4 Conclusion and perspectives

In order to deal with variations in the network’s density of hourly 2 m temperature observations during the reanalysis period, two methods have been tested to produce pseudo-

**Table 3.** Bias and RMSE for MESCAN and SAFRAN (using different data) and MESCAN's guess at 1555 observation points.

	MESCAN				Guess	SAFRAN			
	Astro	Alpha_All	Alpha_5	Real obs.		Astro	AlphaAll	Alpha_5	Real obs.
Bias	0.01	0.09	0.09	0.09	-0.72	-0.1	-0.06	-0.05	0
RMSE	1.96	1.30	1.26	1.12	2.09	1.85	1.15	1.06	0.93

**Figure 3.** Bias (plain lines) and RMSE (dashed lines) for pseudo-observations (left) and MESCAN experiments (REF only uses real observations) (right).

observations. The first method (astro) is the simplest to use and only needs extreme temperatures and geographical data. The second one (alpha) needs hourly data from stations in the neighbourhood. Even if both methods have low bias, alpha method has the best results. The impact of the hourly observation network density available to use the alpha method has been tested. Even in a sparse network (one observation per 1000 km<sup>2</sup> over France), results with alpha remain slightly better. However in extreme experiments with only one hourly observation over the whole country, best results are obtained with the astro method (especially when the weather type of the local observation is particular). Furthermore with a high-density network, using two neighbours for the alpha method leads to satisfying results.

Using the pseudo-observations as an input for reanalysis systems gives results in accordance; i.e. results for alpha methods are better than astro. In addition performances of reanalysis systems with pseudo-observations with respect to performances with real observations are close; even if there is a slight deterioration, the bias coming from the guess is still reduced (even if for astro the bias is monthly dependent). The results of this work show that pseudo-observations can be useful in a reanalysis system. But as we introduce observation errors using pseudo-observation in a reanalysis system, we should use specific error statistics for pseudo-observations in the optimal interpolation.

Finally we will work on the extension over a longer period to be able to compute climatological indexes and compare them to the ones deduced from long-term series.

**Acknowledgements.** The authors would like to thank the reviewers for their comments and suggestions, which improved the quality of this manuscript. The research leading to these results has received funding from the European Union's Seventh Framework Program (FP/2007-2013) under grant agreement no. 242093 and agreement no. 607193.

Edited by: A. K. Kaiser-Weiss

Reviewed by: two anonymous referees

## References

- Durand, Y., Brun, E., Merindol, L., Guyomarc'h, G., Lesaffre, B., and Martin, E.: A meteorological estimation of relevant parameters for snow models, *Ann. Glaciol.*, 18, 65–71, 1993.
- Durand, Y., Giraud, G., Brun, E., Merindol, L., and Martin, E.: A computer-based system simulating snowpack structure as a tool for regional avalanche forecasting, *J. Glaciol.*, 45, 469–484, 1999.
- Perrin, M., Besson, F., Soubeyroux, J. M., and Bazile, E.: Reconstitution du cycle quotidien de température pour le futur système d'analyse atmosphérique MESCAN. Rapport M1/SID – University of Toulouse III – Paul Sabatier, Toulouse, France, 2013.
- Peterson, T. and Vose, R.: An overview of the global historical climatology network temperature database, *B. Am. Meteor. Soc.*, 78, 2837–2849, doi:10.1175/1520-0477(1997)078<2837:A00TGH>2.0.CO;2, 1997.
- Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., and Morel, S.: Analysis of near surface atmospheric variables: validation of the SAFRAN analysis over France, *J. Appl. Meteorol. Climatol.*, 47, 92–107, doi:10.1175/2007JAMC1636.1, 2008.
- Reicosky, D. C., Winkelman, L. J., Baker, J. M., and Baker, D. G.: Accuracy of hourly air temperatures calculated from daily minima and maxima, *Agr. Forest Meteorol.*, 46, 193–209, 1989.
- Soci, C., Bazile, E., Besson, F., Landelius, T., Mahfouf, J.F., Martin, E., and Durand, Y.: EURO4M Project – REPORT D 2.6 Report describing the new system in D2.5, 2013.
- Soubeyroux, J. M., Schneider, M., and Besson, F.: Faisabilité de construction d'un indicateur thermique régionalisé sur la France à partir de la réanalyse SAFRAN, DCLIM/AVH, Météo-France: Toulouse, France, 2011.
- Vidal, J. P., Martin, E., Franchisteguy, L., Baillon, M., and Soubeyroux, J. M.: A 50-year high-resolution atmospheric reanalysis over France with the Safran system, *Int. J. Climatol.*, 30, 1627–1644, doi:10.1002/joc.2003, 2009.