# Estimating daily climatological normals in a changing climate


Alix Rigal[1], Jean-Marc Azaïs[2], Aurélien Ribes[1]

[1] *CNRM, UMR3589, Météo-France and CNRS, Toulouse, France*

[2] *Institut de Mathématiques, Université Paul Sabatier (member of the federal University "Université de Toulouse"), 31062 Toulouse, cedex 09 France*

*email : alix.rigal@meteo.fr*


## Introduction

Climatological normals are widely used baselines for the description and the characterisation of a given meteorological situation. The World Meteorological Organisation (WMO) standard recommends estimating climatological normals as the average of observations over a 30-year period. This approach may lead to strongly biased normals in a changing climate. Here we propose a new method with which to estimate daily climatological normals in a non-stationary climate. Our statistical framework relies on the assumption that the response to climate change is smooth over time, and on a decomposition of the response inspired by the pattern scaling assumption. Estimation is carried out using smoothing splines techniques, with a careful examination of the selection of smoothing parameters. The new method is compared, in a predictive sense and in a perfect model framework, to previously proposed alternatives such as the WMO standard (reset either on a decadal or annual basis), averages over shorter periods, and hinge fits.

## Existing methods/background

- **World Meteorological Organisation (WMO)**

The WMO recommendation is to calculate climatological normals as a simple average over a 30-year period reset each decade:

$$WMO(D+k) = \frac{1}{30}\sum_{i=D-29}^{D} T_i, \qquad (1)$$

where $D+k$ is the current year, $D$ is the current decade (e.g 2010), $k \in [\![1,10]\!]$ denotes the year within the decade, $T_i$ is the mean temperature (or any other meteorological variable) of year $i$.

- **WMO reset**

As a first very simple alternative[4], the same calculation can be made and updated every year, leading to:

$$WMO(y) = \frac{1}{30}\sum_{i=y-30}^{y-1} T_i, \qquad (2)$$

where $y$ is the current year.

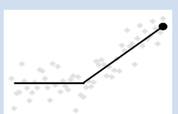
- **OCN - Optimal Climate Normals**

Averaging over a 30-year period was sub-optimal[2,5,6], a 15-year window was found to be a good compromise for temperature normals. In the following it will be updated every 10 years and used as a benchmark for other operational normals based on shorter averaging.

$$OCN(D+k) = \frac{1}{15}\sum_{i=D-14}^{D} T_i, \qquad (3)$$

with $k \in [\![1,10]\!]$.

- **Hinge fit**



Break-point model allowing for a trend in the estimation[3,6]
1975 was an appropriate choice for the continental US[3,6]

$$Hinge(D+k) = \beta_0 + \beta_1 I_{1975}(D+k), \qquad (4)$$

where $I(x)=0$ if $x \leq 1975$ and $I(x) = x - 1975$ if $x > 1975$.
The coefficients $\beta_0$ and $\beta_1$ are estimated from the full observational record available up to year $D$ using simple linear regression.
Again, this type of estimate could be updated each decade or year in which case it will be referred to **hinge fit reset**.

- **From annual to daily normals**

All of the techniques listed above can be used to derive daily normals. This requires an additional procedure and consists of an expansion in a Fourier basis.

## Proposed method - spline model

Let $T_{y,d}$ be the mean temperature of day $d$ in year $y$. Our statistical model assumes that the following decomposition holds:

$$T_{d,y} = f(d) + g(y)h(d) + \varepsilon_{d,y}, \qquad d \in [\![1,365]\!], y \in [\![1,n]\!], \qquad (5)$$

where:

- $f(), g(), h()$ are smooth functions,
- $f(), h()$ are periodic functions with period 365,
- $\varepsilon$ is assumed to be Gaussian white noise with unknown variance $\sigma^2$.

The solution of such a statistical model is achieved by smoothing splines[1].

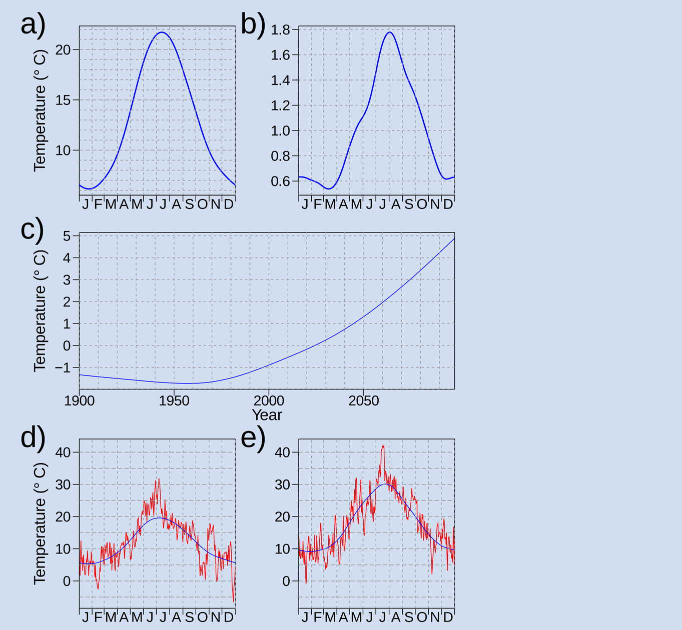



Figure (right): Decomposition of a time series (Paris) by the spline model (5). a) represents the reference seasonal cycle $f$ with *df=11*, b) illustrates the seasonal drift $h$ with *df=10*, and c) represents the annual trend $g$ with *df=10*. The plots d) and e) show the estimation of the annual cycle in 1900 and 2030 respectively. Raw data are shown in red, while the fit of model (1) is in blue.

## Selecting degrees of freedom

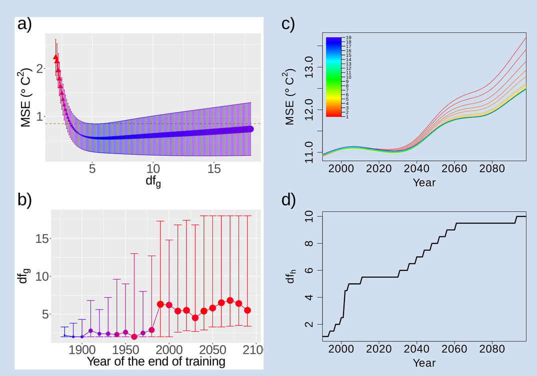
The determination of the different degrees of freedom ($df$) is performed using a variant of cross-validation methods adapted to a prediction context (i.e. the $df$ minimizing the $MSE = \sum(prediction - data)^2$).
The three coefficients: $df_f$, $df_g$, $df_h$ are estimated sequentially (computationally more affordable).

**Degree of freedom of the reference cycle $f()$:**
- Training sample: 1900-1930 , 1931-1940 as a testing sample (stationary climate).
- The $df$ achieving the smallest MSE is taken.
- High signal-to-noise ratio ⇒ well defined.

**Degree of freedom of the annual trend $g()$:**
- We calculate each decade $D$ on annual mean temperatures.
- Training sample: data prior to $D$ and testing sample: D.
- Stability after 1990 (all locations).

**Degree of freedom of the delta cycle $h()$**
- $df_f$ and $df_g$ are determined.
- For a given year $y$

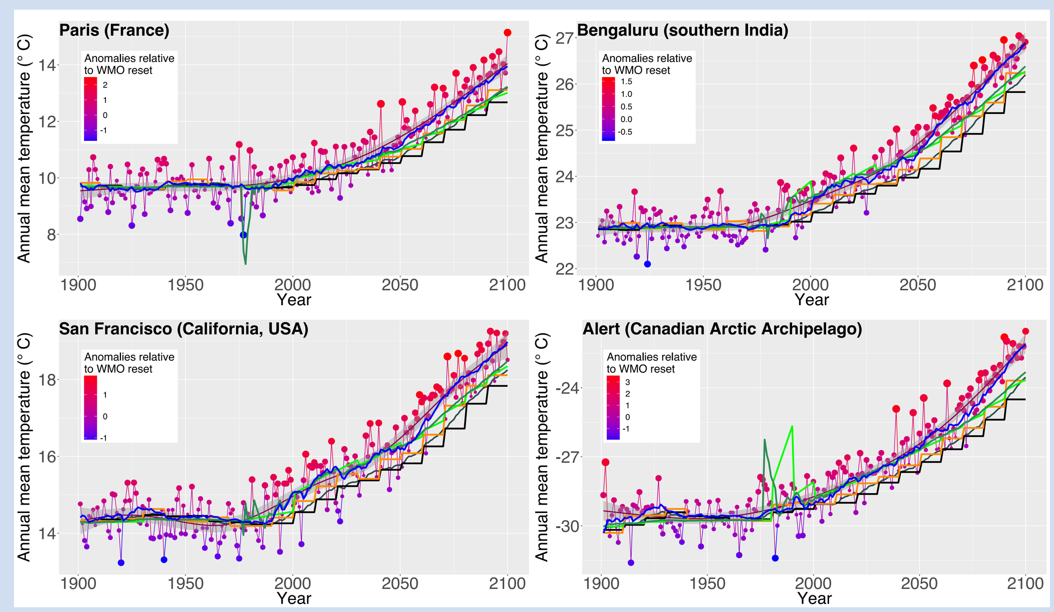
Training sample: days prior $y$ and testing sample: days of year $y$.




a) MSE for the decade 2050 Paris (France). b) Selected $df_g$.
c) MSE of daily normals as a function of time for different values of $df_h$. d) Selected $df_h$.

## Results on annual mean temperature(CNRM-CM5)

Annual mean temperatures and estimated normals. Temperature normals on an RCP8.5 scenario of the CNRM-CM5 model.

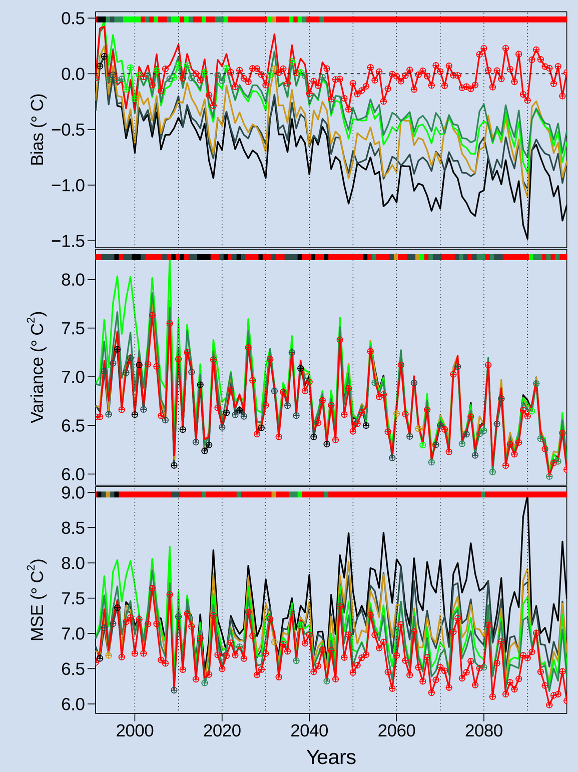



Time-series of annual mean temperature (points) at four different locations (panels). Climatological normals are estimated using 6 different techniques: WMO standard (black line), WMO reset (grey), OCN (yellow), hinge fit (light green), hinge fit reset (dark green) and spline model (blue). A smoothing spline of the entire time-series (1900-2100; purple line) can be considered as a reference. Anomalies with respect to the WMO reset.

## Results for CMIP 5

Daily scores in San Francisco, of RCP8.5 simulation from the CMIP5 archive.

The year in the $x$-axis denotes the end of the training period; prediction is made for the following year. The coloured line (top of each panels) indicates which method performs best, for a given criterion and a given year.




Six techniques for estimating daily normals, namely WMO (black), WMO reset (grey), OCN (yellow), hinge fit (light green), hinge fit reset (dark green) and spline model (red).

## Conclusion

Results show that our technique outperforms all alternatives considered and confirm that previously proposed techniques are substantially biased (from ∼ 0.1° to ∼ 1°). The proposed estimation algorithm is very fast due to a two-step (as opposed to simultaneous) procedure. The main challenge is the tuning of the smoothing parameters which is done using an extension of cross-validation specifically designed for prediction.

**Application**

- Climate monitoring, e.g. qualify whether a year or season is warmer or colder than expected.
- Producing climate change corrected times-series.
- Providing a refined description of on-going climate change with respect to the annual cycle, i.e. beyond the annual mean warming.

## References


1. Azaïs and Ribes, 2016. *Journal of Multivariate Analysis.*
2. Huang et al., 1996. *Journal of Climate*
3. Livezey et al., 2007. *Journal of Applied Meteorology and Climatology*
4. Scherrer et al., 2006. *International Journal of Climatology*
5. Wilks, 2013. *Journal of Applied Meteorology and Climatology*
6. Wilks et al., 2013. *Journal of Applied Meteorology and Climatology*