Research Article Open Access

S. Foucart, M. Hielsberg, G. L. Mullendore, G. Petrova*, and P. Wojtaszczyk

Optimal Algorithms for Computing Average Temperatures

https://doi.org/10.1515/mcwf-2019-0003
Received November 13, 2018; accepted March 24, 2019

Abstract: A numerical algorithm is presented for computing average global temperature (or other quantities of interest such as average precipitation) from measurements taken at specified locations and times. The algorithm is proven to be in a certain sense optimal. The analysis of the optimal algorithm provides a sharp a priori bound on the error between the computed value and the true average global temperature. This a priori bound involves a computable *compatibility constant* which assesses the quality of the measurements for the chosen model. The optimal algorithm is constructed by solving a convex minimization problem and involves a set of functions selected a priori in relation to the model. It is shown that the solution promotes sparsity and hence utilizes a smaller number of well-chosen data sites than those provided. The algorithm is then applied to canonical data sets and mathematically generic models for the computation of average temperature and average precipitation over given regions and given time intervals. A comparison is provided between the proposed algorithms and existing methods.

1 Introduction

Computing average temperatures is a particular instance of a common task in data processing, namely that of exploiting measurements made on a function f to estimate a quantity $Q(f)$ that depends on f, referred to as Quantity of Interest (QoI) below. In the present situation, $f = f(x, t)$ represents the temperature (in degrees Celsius) as a function of the position x on the surface of the earth and of the time t. We single out temperature as a running example, but other atmospheric features such as humidity or precipitation can be treated equally well. As a QoI, we single out average temperatures over the whole earth or a smaller region R during a whole year or a shorter period Δ, obtained as the normalized integral

$$Q(f) = \frac{1}{\sigma(R)|\Delta|} \int\limits_R \int\limits_\Delta f(x, t)\, d\sigma(x)\, dt, \tag{1}$$

where $\sigma(R)$ is the surface area of the region and $|\Delta|$ is the duration of the time period. In going further, we denote by $\Omega := \{(x, t) : x \quad R,\ t \quad \Delta\}$ the domain of f, so the QoI also reads $Q(f) = |\Omega|^{-1} \int_\Omega f$, where $|\Omega| = \sigma(R)|\Delta|$ denotes the measure of Ω. Other QoIs such as the temperature $Q(f) = f(\xi, \tau)$ at a specific location ξ and a specific time τ can be considered as well.

S. Foucart: Department of Mathematics, Texas A M University, College Station, TX 77840, USA, E-mail: foucart@tamu.edu
M. Hielsberg: Department of Mathematics, Texas A M University, College Station, TX 77840, USA, E-mail: hielsberg@tamu.edu
G. L. Mullendore: Department of Atmospheric Sciences, University of North Dakota, Grand Forks, North Dakota, USA, E-mail: gretchen.mullendore@und.edu
Corresponding Author: G. Petrova: Department of Mathematics, Texas A M University, College Station, TX 77840, USA, E-mail: gpetrova@math.tamu.edu
P. Wojtaszczyk: Institute of Mathematics, Polish Academy of Sciences, 00-656 Warsaw, ul. Sniadeckich 8, Poland, E-mail: wojtaszczyk@impan.pl

Open Access. © 2019 S. Foucart et al., published by De Gruyter. This work is licensed under the Creative Commons Attribution alone 4.0 License.

Temperatures are continuously monitored by numerous weather stations around the globe, providing us with ample data to estimate the QoI. This data takes the form of a vector

$$w = [f(x_1, t_{1,1}), \dots, f(x_1, t_{1,\ {}_1}), \dots, f(x_m, t_{m,1}), \dots f(x_m, t_{m,\ {}_m})] \tag{2}$$

whose entries are obtained by recording the temperature at locations x_j and at times $t_{j,i}$, $j = 1, \dots, m$ and $i = 1, \dots, \ {}_j$. This vector depends linearly on the temperature function f and can be written succinctly as $w = M(f)$, where M represents the measurement mapping associated with the data sites $(x_j, t_{j,i})$, $j = 1, \dots, m$, $i = j, \dots, \ {}_j$.

This data alone is not sufficient to guarantee that a QoI such as annual global temperature can be accurately computed. Indeed, without additional knowledge, the temperature function could, for example, oscillate wildly between the data sites, in which case the computed average temperature $Q(f)$ based on the collected data would be very far from the true average. To forbid such unrealistic scenarios, one imposes additional requirements on the temperature function f, which usually come either from the physical properties of f or from regularity theorems for f as a solution to (system of) partial differential differential equations (PDEs). These additional demands on f are referred to as model class assumptions and are necessary to quantify the uncertainty in our knowledge about f. Unfortunately, in the case of the temperature function, the complexity of the system of PDEs makes regularity theorems hard to exploit. Therefore, the model class assumptions we consider are not built on the smoothness of f, but instead on its approximability. More precisely, our model class is the set of all continuous functions on the sphere (or part of the sphere) that can be approximated, up to accuracy ϵ, by a finite dimensional space V of functions. The choice of this finite dimensional space is essential and should be tailored to the physical properties of f. We believe that the geophysical community has a lot of experience and knowledge about the physical properties of the temperature/precipitation functions and can propose relevant choices of V for these cases. Once the model class is fixed, i.e., the linear space V is selected, the problem of computing $Q(f)$ is well posed since the uncertainty about f is settled, and we can focus on the computation of $Q(f)$. There are, in general, three main issues related to the actual calculation of the QoI:

(i) Is there an optimal accuracy at which one can estimate a QoI?
(ii) Are there procedures that achieve this accuracy?
(iii) Can one implement these procedures?

Such questions form the cornerstone of a mathematical theory called *optimal recovery* (see [18]). Regarding (ii), for instance, the theory guarantees that there is an optimal algorithm for computing an average temperature $Q(f)$ that takes the form of a weighted average of the data, i.e., using the approximation

$$Q(f) \approx \sum_{j=1}^{m} \sum_{i=1}^{j} a^{\star}_{j,i} f(x_j, t_{j,i}). \tag{3}$$

While the existence of such an algorithm is settled, the difficulty lies in the actual construction of this optimal, or even of a near optimal algorithm, since the data points $(x_j, t_{j,i})$ are not equidistributed. Some of the available current methods used to estimate average temperatures also follow that path and take the form of a weighted average, even though they are not explicitly stated as such. In essence, they use local approximation-based methods to calculate the weights, by obtaining approximations to the values of the temperature function at equidistributed points on the sphere (see Section 2.1). However, no error analysis or optimality of the proposed techniques is provided which would give valuable information on the closeness (or not) between the computed quantity and the true average.

Recently, significant advances were made on (i)-(ii)-(iii) in [7], where it was revealed how to explicitly construct optimal algorithms for the estimation of QoIs relative to the model classes mentioned above. The purpose of this note is to present to the geophysical scientists the latest theoretical and algorithmic developments in optimal recovery and inquire their expertise in atmospheric science in an effort to produce the most realistic model classes (specifically, choices of linear spaces V) for average global temperature computation with certifiable performance. More precisely, we present an optimal algorithm for the computation of the average temperature $Q(f)$ based on the measurements $w = M(f)$. The error between the computed QoI and the true average global temperature is proven to be bounded by the product of the error ϵ of the underlying model and a computable *compatibility constant* μ_V which assesses the quality of the measurements for the chosen model (space V). The algorithm finds the weights in (3) by solving a convex minimization problem. The optimization is performed (preferably) during an offline stage, uses a basis for the space V, and requires only the station locations and not the station readings. The obtained solution utilizes only a subset of the set of station locations involved in the optimization. Thus, only the readings at these stations need to be accessed in the online computation stage. This additional efficiency feature of the proposed algorithm reduces the data transmission time, especially in cases when huge databases are stored in a cloud system and are accessible to multiple scientific teams.

Finally, we test the algorithm on canonical data sets for the computation of average temperature and average precipitation over given regions and given time intervals. Note that, to provide a fair comparison between the proposed algorithm and existing methods, we use for some of the experiments the same processed data that have been used for the existing methods . As to the model class, we employ the standard (in the case of functions defined on the sphere) choice of V being the space of spherical harmonics of certain degree or the space of piecewise constant functions. We believe that one can exploit better choices of spaces V (probably well known to atmospheric scientists), and therefore better model classes based on approximibility, which are intrinsically connected to the nature of the temperature or precipitation function f and use our algorithm to dramatically improve on the quality of the obtained results.

2 The algorithm

We now formalize our approach, in particular by clarifying the notion of optimal algorithms. By an algorithm, we simply mean a mapping taking a measurement vector $w = M(f) \quad {}^m$ as input and returning a number $A(w) \quad$ as an output, hopefully close to the true value $Q(f)$. The error made by this approximation is

$$E(w, A) := |Q(f) - A(w)|, \qquad w = M(f). \tag{4}$$

The performance of the algorithm A is then assessed in a worst-case setting, keeping in mind that, besides the data, the only information one has (or presumes) about f is that it satisfies our model class assumptions. The model classes we advocate are built on the approximability properties of f. This usually involves a specific space of functions that we use for approximation, e.g. polynomials, piecewise polynomials, radial basic functions, wavelets, etc., and thus we assume that we have an n-dimensional space V contained in the space $C(\Omega)$ of continuous functions on a domain Ω. *That space encapsulates all our a priori knowledge about the temperature function f.* It would be ideal but clearly unrealistic for f to be an element of V, since we have only partial knowledge about f. So, the best we can do is to use a space V which can rather well approximate the temperature function f. More precisely, if the distance from f to V in uniform norm is

$$\operatorname{dist}(f, V) := \inf_{v \in V} \max_{\omega \in \Omega} |f(\omega) - v(\omega)| \le \epsilon, \tag{5}$$

our model class is the set K of all functions

$$K = K(\epsilon, V) := \{g \quad C(\Omega) : \ \operatorname{dist}(g, V) \le \epsilon\} \tag{6}$$

that can be approximated by elements from V up to accuracy $\epsilon > 0$. We next introduce the error of the algorithm A for the elements in this model class K, that is

$$E(K, A) := \sup_{f \in K} E(M(f), A) \tag{7}$$

as an indicator of the performance of the algorithm A over the class K. A favorable bound on $E(K, A)$ guarantees that $Q(f)$ is computed well. Finally, an optimal algorithm A^* is one that makes $E(K, A)$ as small as possible, i.e.,

$$E^*(K) := E(K, A^*) = \inf_A E(K, A). \tag{8}$$

Note that $E^{\star}(K)$ quantifies the optimal performance relative to the class K. It was shown in [7] that the optimal performance $E^{\star}(K)$ can be precisely evaluated for the model class (6) and that it decouples as

$$E^{\star}(K(\epsilon, V)) = \mu_V \epsilon. \tag{9}$$

The latter formula reveals that, besides the approximation capability of V, another important part is played by a constant μ_V that encapsulates the compatibility of the data sites $(x_j, t_{j,i})$, $j = 1, \dots, m$, $i = 1, \dots, \ _j$, with the space V and the quantity of interest Q. A poor compatibility will result in a large μ_V, which happens for example when $Q(v)$ is large for some $v \quad V$ while all measurements $v(x_j)$ are small, see (13). The roles of μ_V and ϵ are somewhat competing in the choice of a proper linear space V on which to build the algorithm. Indeed, we want a space with both good approximation capability, so that ϵ is small, and good compatibility with the data, so that μ_V is small. Note that enlarging the space V will have the effect of decreasing ϵ while increasing μ_V. Such a tension is similar to the one encountered in statistical learning when confronted with the problem of overfitting the data. Next, once a "favorite" space V is chosen, the results in [7] also put forward the construction of an optimal algorithm $A^{\star}$ for the estimation of real-valued linear QoIs such as (1). The execution of the algorithms does not depend on how well the particular choice of V approximates the function of interest f. All this is factored in the error estimate (9), which can be recast as

$$|Q(f) - A^{\star}(M(f))| \leq \mu_V \mathrm{dist}(f, V), \tag{10}$$

so in the above sense the proposed algorithm is universal.

For easy comparison with other algorithms, we describe our construction in the special framework where there is no time dependence, i.e., when $f = f(x)$ depends only on the position x. Existing algorithms for estimating annual temperatures can indeed be unscrambled as producing beforehand an annual temperature $f(x_j)$ at each data site x_j, and then computing a weighted average of the $f(x_j)$, see e.g [14] for the general overview. We place ourselves in the same setting where our data consist of measurements $f(x_j)$ which have already been averaged over time. Another instance of the time-independent framework is the case where $f_\tau(x) = f(x, \tau)$ represents an instantaneous temperature at a given time τ.

In the time-independent framework, given data sites $(x_1, \dots, x_m)$ and a basis $(v_1, \dots, v_n)$ for the space V, the optimal algorithm $A^{\star}$ for the estimation of a real-valued linear QoI Q first produces a solution

$$a^{\star} := \operatorname{argmin}\left\{ \sum_{j=1}^{m} |a_j| : \ \sum_{j=1}^{m} a_j v_i(x_j) = Q(v_i), \quad i = 1, \dots, n \right\} \tag{11}$$

of an $_1$-minimization problem with variable $a = [a_1, \dots, a_m]^{\top}$, which can be reformulated as a linear optimization problem and solved (usually offline) using standard techniques. Next, for each measurement vector $w \quad ^m$, the algorithm computes (online) the weighted average

$$A^{\star}(w) := \sum_{j=1}^{m} a_j^{\star} w_j. \tag{12}$$

Notice that the optimal weights $a_1^{\star}, \dots, a_m^{\star}$ depend on the sites $x_1, \dots, x_m$ but not on the data $f(x_1), \dots, f(x_m)$ at these sites. Therefore, once the weights are computed (preferably offline), formula (12) can be reused instantly for each data $w_\tau = [f_\tau\ (x_1), \dots, f_\tau\ (x_m)]$ recorded at another time τ' (or other time averages).

The algorithm $A^{\star}$ is optimal in the sense that $E(K(\epsilon, V), A^{\star}) = E^{\star}(K(\epsilon, V))$ for all $\epsilon > 0$. Notice that the knowledge of $\epsilon > 0$ is not necessary to construct the algorithm $A^{\star}$. In addition, the minimization algorithm gives an explicit formula for the compatibility constant μ_V, see [7], namely

$$\mu_V := \mu((x_j)_{j=1}^{m}, V, Q) = 1 + \sup_{v \in V} \frac{|Q(v)|}{\max_{j=1,\dots,m} |v(x_j)|} = 1 + \sum_{j=1}^{m} |a_j^{\star}|. \tag{13}$$

2 1 Novelty of our work and comparison with other methods

Our approach possesses several compelling features which we discuss next. First, we propose a model class based on the approximability of the temperature function, formally expressed by the choice of a space V.

This space explicitly encodes our general knowledge about the temperature field (or other phenomenon we are studying). Next, we suggest an optimal algorithm $A^\star$, see (11)-(12), for the computation of a QoI for the elements from this model class. The algorithm is universal, because it can be executed for any space V. It is reliable because we provide an a priori error estimate for the difference between the computed quantity and the true average. It is efficient since it uses a basis for V and the station locations (not the station readings) to select at the end only a small number of station locations whose readings are used for the computation of the global average temperature.

The current methods for the computation of global average temperature could be roughly divided into two main categories: approximation-based methods and statistical methods. The approximation-based methods use the available station data readings to assign values $(f_k)_{k=1}^M$ of the temperature function at regular grid points $(g_k)_{k=1}^M$, see e.g. [10], [11], and [12]. Once this is done, an average over the newly computed values $(f_k)_{k=1}^M$ is performed. The choice of the grid and the method of assigning these values (usually done via interpolation) implicitly defines the model class assumptions. In [25], twelve methods of interpolation are reviewed and analyzed. Most of them are various versions of the classical inverse distance weighting (IDW) method, see [23]. Its main idea is that at a grid point g_k one assigns the average f_k of temperature readings from stations whose distance to g_k is smaller than a certain fixed threshold, using weights which decay as the distance from g_k gets bigger. Various modifications of this method are used. For example, the reference station method (RSM) proposed in [10] uses all stations whose distance from g_k is at most 1200 km and the weights decay linearly but with additional preference (i.e. bigger weights) given to stations with longer operating times. Those methods produce non-zero weights for all station locations and the respective algorithms use all available readings.

A whole range of statistical methods, often called optimal interpolation, have also been used for average temperature computation. We refer the reader to [6, chap. 4] for the history and basics of these techniques. They are applied to both the computation of the grid values f_k (or the full temperature field f) and the direct calculation of the weights. Usually, in those methods, the weights are computed using the variance and covariance structure of the temperature field (or other phenomenon we are studying). Assumptions about this structure are incorporated in the respective algorithm, reflecting the model assumptions. For example, the method proposed by [26] based on the work of [8] and [15] calculates the weights in (3) directly, see also [24].

3 Experimental results

In this section, we test our method on several situations encountered in atmospheric science: the estimation of average annual global temperature, the estimation of average seasonal regional temperature, and the estimation of total annual global precipitation. We also give a brief description of the implementation of our algorithm.

3 1 Implementation details

Here, for simplicity, we consider the time-independent framework only and discuss the implementation of problem (11) for computing global temperatures. Recall that once the solution $a^\star = [a_1^\star, \dots, a_m^\star]^\top \quad ^m$ to (11), where m is the number of data sites, has been computed offline, the final step for approximating the quantity of interest consists in outputting

$$\sum_{j=1}^{m} a_j^\star w_j$$

for the given data values $w_1, \dots, w_m$.

If we denote by $a := [a_1, \dots, a_m]^\top$ and introduce the vector $s := [s_1, \dots, s_m]^\top \quad {}^m$ of slack variables such that $|a_j| \le s_j$, $j = 1, \dots, m$, finding $a^\star$ is equivalent to solving the minimization problem, see [3]:

$$\underset{a \in \quad {}^m, s \in \quad {}^m}{\text{minimize}} \sum_{j=1}^{m} s_j \qquad \text{subject to } Ba = c, \; -s \le a \le s.$$

Here B is an $n \times m$ matrix, where $n = \dim(V)$, with entries $b_{ij} = v_i(x_j)$, where the v_i's form a basis for the space V, that is $V = \operatorname{span}\{v_1, \dots, v_n\}$. The vector $c := [c_1, \dots, c_n]^\top \quad {}^n$ has entries $c_i = \frac{1}{4\pi} \int_S v_i d\sigma$, $i = 1, \dots, n$, where S is the unit sphere. The above problem can be handled by any off-the-shelf linear solver. We relied on CVX, a MATLAB package for specifying and solving convex programs, see [1].

In the particular case when V is the space of spherical harmonics of degree at most L, see [19], the entries of B and c are computed as follows:

- We pick the usual spherical harmonics $\{Y^k, \quad = 0, \dots, L, k = - \;, \dots, \; \}$ as the basis $\{v_1, \dots, v_n\}$, so that $n = (L+1)^2$. Among the multiple options to construct the spherical harmonics, we relied on Chebfun, a MATLAB package for computing with functions, see [4].
- We collect the data sites $x_1, \dots, x_m$ from one of the appropriate databases mentioned below and we express each data location x_j in polar coordinates (θ_j, ϕ_j).
- We evaluate each spherical harmonic at all the data sites to form the matrix $B \quad {}^{n \times m}$ with entries

$$Y^k(\theta_j, \phi_j).$$

We also evaluate the average of each spherical harmonic to form the vector $c \quad {}^n$ with entries

$$\frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} Y^k(\theta, \phi) \sin(\theta) d\theta d\phi.$$

The numerical evaluations are carried out using Chebfun built-in capabilities.

All experiments can be duplicated by downloading the MATLAB reproducible available on S. Foucart's webpage.

3 2 Average annual global temperatures

Our first and most exhaustive experiment illustrates the application of our algorithm to the estimation of average annual global temperatures. The results are compared with the ones released by agencies such as NOAA National Climatic Data Center and NASA Goddard Institute for Space Studies. In these cases, the time dependence is discarded by considering for each year a function $f = f(x)$ representing an average annual temperature depending only on the position x.

Data provenance: The experiment relies on the following standard data sets:

- Raw Land Data (RLD), obtained by merging monthly land-based station temperatures (GHCNM) from [16] and monthly Antarctic land-based station temperatures (SCAR MET-READER) from [21];
- Processed Land Data (PLD), obtained by processing the RLD using GISTEMP steps 0–2, see [9] and [12];
- Gridded Land Data (GLD), using the PLD to assign temperatures for each grid center in the covering of the globe by 8,000 equal-area cells, see [10];
- Gridded Sea Data (GSD)[1], downloaded from the GISTEMP's website https://data.giss.nasa.gov/pub/gistemp/SBBX.ERSSTv4.gz and obtained following [13] and [20].

Algorithmic details: Our numerical algorithms require the choice of an approximation space V. We consider the linear spaces of spherical harmonics of degrees $L = 3$, $L = 6$, and $L = 9$, which we denote by SH3, SH6,

[1] These data have undergone some processing and gridding steps, too, but we were unable to obtain the raw data used to generate them.

and SH9, respectively. We also consider the linear spaces of piecewise constant functions on two standard partitions of the globe, namely the coarse and fine partitions used in [10], [11], [12], which we denote by PCC and PCF, respectively. We compare the performance of our algorithms with two standard methods provided by GISTEMP and by [20]. Both of these methods implicitly rely on piecewise constant approximation, however, they differ from our optimal algorithm. We compare these algorithms on two data sets:

- Data set 1: This data set consists of the merging of GLD and GSD. This is the standard data set used in GISTEMP.
- Data set 2: This data set consists of the merging of PLD and GSD. This set works more closely with raw data, at least on land. We did not have access to raw sea data.

Observations: The results using PCC and PCF as well as using SH3, SH6 and SH9 on both data sets 1 and 2 are displayed in Figures 1 and 2. In these figures, temperature anomalies in C were computed from the 1951–1980 baseline average. They are compared with those computed using GISTEMP and those reported by NOAA over the 1950–2016 time period. The results call for a number of comments:

- Our results match those reported by NOAA and NASA tightly for data set 1, because they rely on the exact same data set, and more loosely for data set 2, which is closer to raw data;
- Spherical harmonics are surprisingly effective, considering that $\dim(V) = (L+1)^2 = 100$ for $L = 9$ (SH9) while $\dim(V) = 8{,}000$ for the fine grid (PCF);
- Moving from coarse to fine grid improves the approximation capability ϵ and does not severely deteriorate the compatibility constant μ_V (which is also computed by the algorithm), and so produces more accurate results;
- Likewise, for spherical harmonics, increasing L improves ϵ and does not severely deteriorate μ_V. Note that selecting clustered stations, for example stations on land only, would generate a large constant μ_V, since some spherical harmonics appearing in formula (13) have large values over sea regions and small values on land. The increase of μ_V impacts negatively the QoI estimation, as seen from (10).
- The values of μ_V for all the experiments from Figures 1 and 2 are very close to the lower bound of 2 for μ_V (recall that $\mu_V \geq 2$).

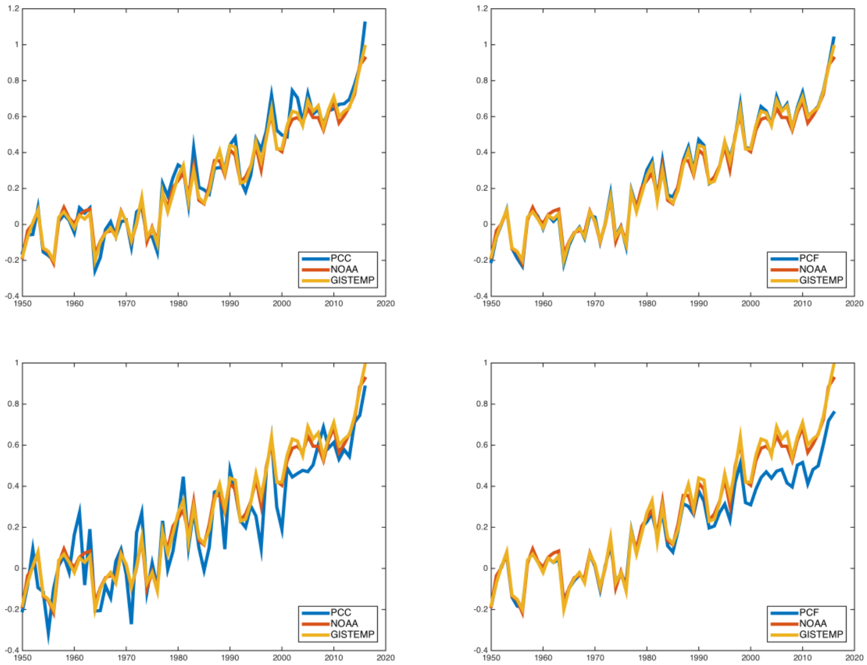

Figure 1: Temperature anomalies in C computed using PCC left column) and PCF right column) with data set 1 top row) and data set 2 bottom row). The values reported by NOAA and by GISTEMP are also shown.

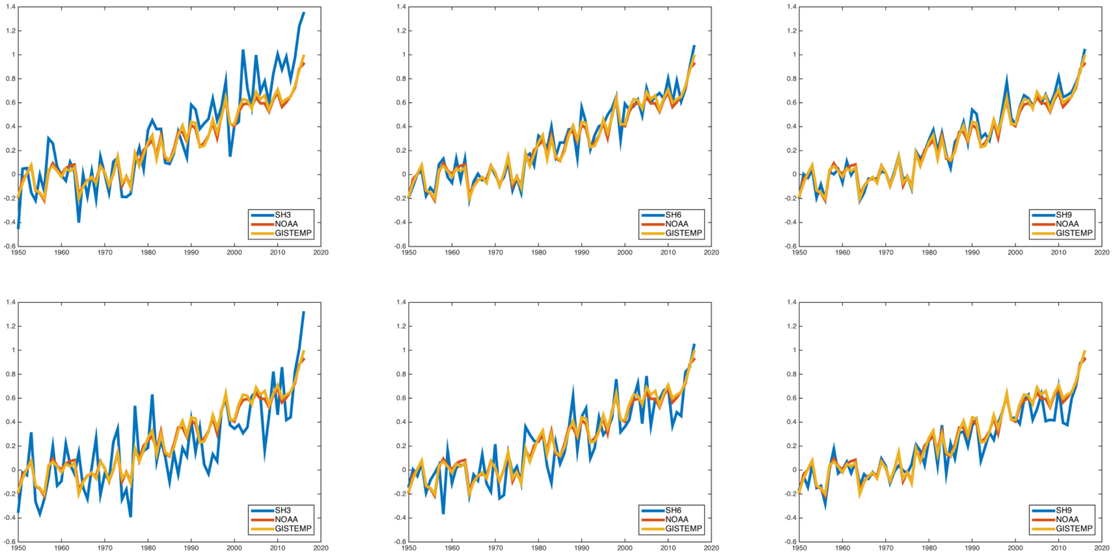

Figure 2: Temperature anomalies in *C* computed using SH3 left column), SH6 middle column) and SH9 right column) with data set 1 top row) and data set 2 bottom row). The values reported by NOAA and by GISTEMP are also shown.

Weather station positioning: According to properties of $_1$-minimizers, the solution $a^\star$ of (11) is sparse, meaning that only $n = \dim(V)$ weights among $a_1^\star, \ldots, a_m^\star$ are nonzero, which implies that only n from the m weather stations are involved in the optimal estimation of average temperatures (recall that the computation of $a^\star$ does not depend on the station readings but only on the positions of the stations). In our spherical harmonics experiments, we have observed that the n stations selected by the method tend to be evenly spread, see Figure 3, and that clustered stations, as in the land-only data set, tend to produce larger values of μ_V, and thus the QoI estimation becomes less reliable in the latter case. When L increases, the value of μ_V does not severely deteriorate as long as there are sufficiently many stations which are reasonably spread around the globe.

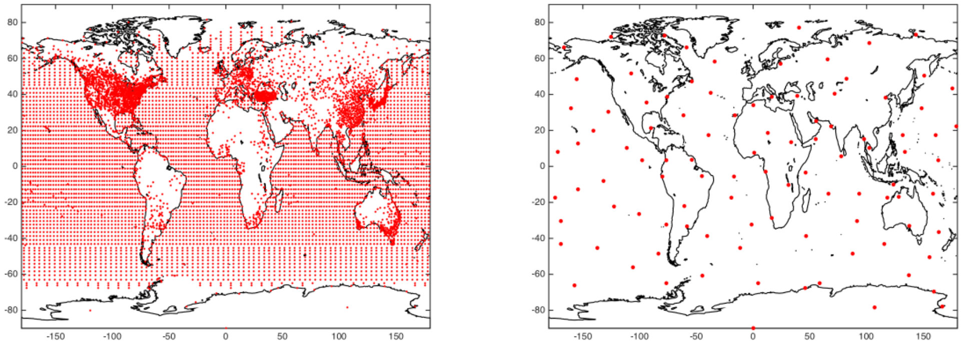

Figure 3: The set of $m = 9{,}187$ locations available in data set 2 left) along with the $n = 100$ selected locations for SH9 right) in 1985. The compatibility constant here is $\mu_V = 2$.

3 3 Average seasonal regional temperatures

We apply our method to estimate the average temperature in the state of Texas, see Figure 4, over two periods of three months (winter and summer) from 2000 to 2016. Time dependence is now incorporated. Thus, the space V consists of functions of two variables: the position x and the time t. We assume a piecewise constant dependence on x and a piecewise linear dependence on t, with breakpoints every week. The data we used underwent a preprocessing step producing average weekly temperatures at each weather station from daily values acquired from [17]. By contrast to the annual global temperature, we do not need to discard a weather station from the record just because one weekly reading is missing. The results in Figure 4 show excellent agreement with the NOAA regional temperatures using all available stations. The volume of data involved in the optimal estimation is reduced by a factor up to 6.

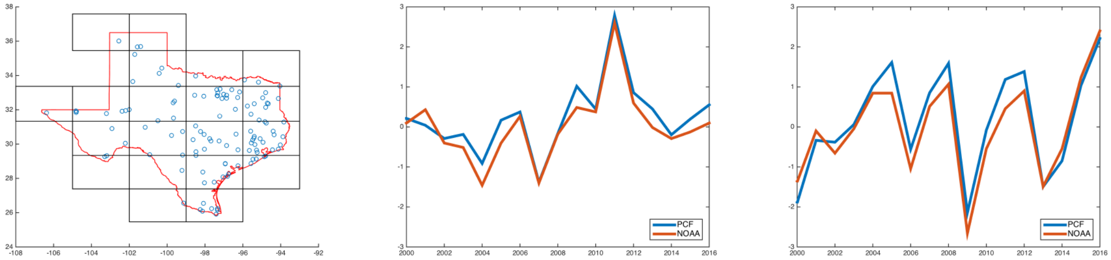

Figure 4: The 115 station located in Texas left) and the temperature anomalies in C for summer middle) and winter right) estimated with the incorporated time-dependence and compared with values reported by NOAA.

3 4 Average annual precipitations for the contiguous US

For our precipitation computations, we use the annual precipitation data for the contiguous US which was extracted from the Global Precipitation Climatology Centre monthly precipitation data set (GPCC) from [22]. Figure 5 shows the results using PCF and compares them with values reported by [2] and downloaded from [5].

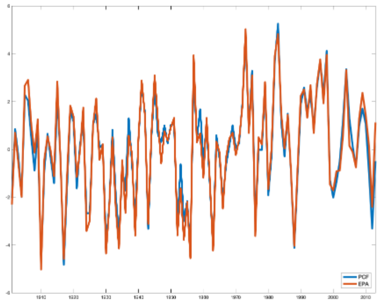

Figure 5: Average annual precipitation anomalies in inches for the contiguous US computed using PCF. The values reported by [2, 5] are also shown.

We observe that PCF performs very well while utilizing only 145 of the data set's 818 grid locations in the contiguous United States, see Figure 6.

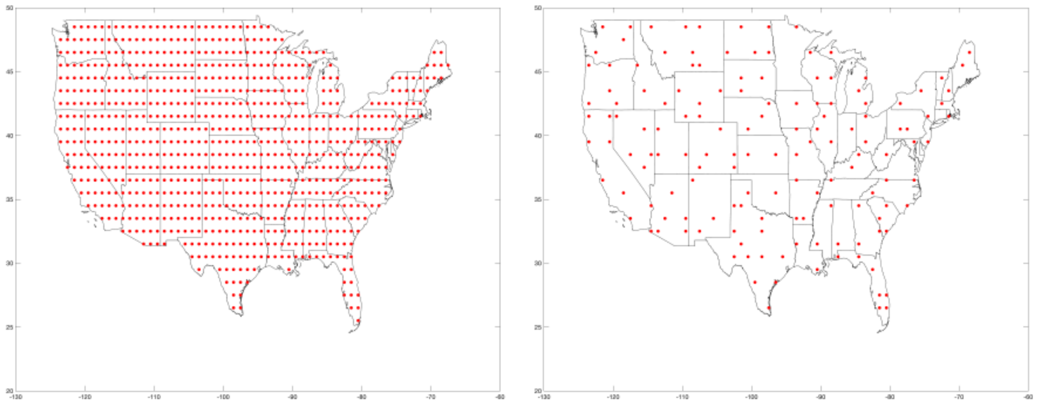

Figure 6: The set of $m = 818$ locations in the contiguous US available in GPCC left) along with the $n = 145$ selected locations for PCF right) in 1985. The compatibility constant here is $\mu_V = 2$.

4 Discussion and conclusion

We have presented a template method to estimate diverse QoIs, with a special emphasis on average temperature. The method, whose parameters are computed offline once and for all, is optimal with respect to the approximation model selected. Ideally, the method should be applied to raw data, where the data manipulation as performed by other methods (such as urban adjustment) should be integrated in the modeling stage, i.e., the choice of the space V. This choice is of course critical for the reliability of the method. It is rather surprising that spherical harmonics provided good results — this would have been conceivable for temperatures in the free troposphere, but on the earth surface they obliterate the important dependence of temperature on latitude, altitude, surface type, etc. We hope that domain experts can chime in and propose relevant choices of V for the model. A good space V should have a small dimension (the space of piecewise constants does not) for at least two reasons: the compatibility constant μ_V will remain small and the estimation formula will feature few nonzero weights. The latter is due to the automatic sparsity of $_1$-minimizers. It does not suggest eliminating numerous weather stations and acquiring less data, as the nonzero weights depend on the QoI, on V, and on the available stations, but it presents an interesting advantage in terms of data transmission.

Acknowledgement: S. F. partially supported by NSF grants DMS-1622134 and DMS-1664803. G. M. partially supported by NSF grant ACI-1450168. G. P. partially supported by NSF grants DMS-1521067 and DMS-1817603. P. W. partially supported by National Science Centre, Poland grant UMO-2016/21/B/ST1/00241.

References

[1] CVX Research, Inc. CVX: MATLAB software for disciplined convex programming, version 2.1. http://cvxr.com/cvx, 2014.
[2] Blunden J., Arndt D. eds.), State of the climate in 2015, B. Am. Meteorol. Soc., 2016, 97 8):S1-S275.
[3] Boyd S., Vandenberghe L., Convex Optimization, Cambridge University Press, 2004.
[4] Chebfun – numerical computing with functions, version 5.7.0. http://www.chebfun.org/.
[5] Climate Change Indicators: US and Global Precipitation. Dataset accessed 2018-04-27 at https://www.epa.gov/sites/production/files/2016-08/precipitation_fig-2.csv.
[6] Daley R., Atmospheric data analysis, Cambridge atmospheric and space science series, 2, 1991.
[7] DeVore R., Foucart S., Petrova G., Wojtaszczyk P., Computing a quantity of interest from observational data, Constructive Approximation, *https://doi.org/10.1007/s00365 018 9433 7*, to appear.

[8] Gandin L., Objective analysis of meteorological fields, Gidrometeoizdat, Leningrad, Translated by Israel Program Scientific Translations, Jerusalem.

[9] GISTEMP Team, 2018: GISS Surface Temperature Analysis GISTEMP). NASA Goddard Institute for Space Studies. Dataset accessed 2018-04-27 at https://data.giss.nasa.gov/gistemp/.

[10] Hansen J., Lebedeff S., Global trends of measured surface air temperature, Journal of Geophysical Research, 1987, 92.D11, 13.

[11] Hansen J., Ruedy R., Glascoe J., Sato M., GISS analysis of surface temperature change, Journal of Geophysical Research: Atmospheres, 1999, 104, no. D24, 30997-31022.

[12] Hansen J., Ruedy R., Sato M., Lo K., Global surface temperature change, Reviews of Geophysics, 2010, 48, no. 4.

[13] Huang B., Banzon V., Freeman E., Lawrimore J., Liu W., Peterson T., Smith T., Thorne P., Woodruff S., Zhang H., Extended Reconstructed Sea Surface Temperature ERSST), version 4, NOAA National Centers for Environmental Information, 2015 doi:10.7289/V5KD1VVF.

[14] Jones P., New M., Parker D., Martin S., Rigor I., Surface air temperature and its changes over the past 150 years. Reviews of Geophysics, 1999, 37,2 173-199.

[15] Kagan R., Averaging of meteorological fields, Gidrometeoizdat, Leningrad in Russian), 1979.

[16] Lawrimore J., Menne M., Gleason B., Williams C., Wuertz D., Vose R., Rennie J., An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3, J. Geophys. Res., 2011, 116, D19121.

[17] Menne M., Durre I., Vose R., Gleason B., Houston T., An overview of the Global Historical Climatology Network-Daily Database, Journal of Atmospheric and Oceanic Technology, 2012, 29, 897-910, doi:10.1175/JTECH-D-11-00103.1.

[18] Micchelli C., Rivlin T., Lectures on optimal recovery. Numerical analysis, Lancaster, 1984), 198521–93, Lecture Notes in Math., 1129, Springer, Berlin.

[19] Müller C., Spherical Harmonics, Lecture Notes in Mathematics, Springer.

[20] NOAA National Centers for Environmental Information, Climate at a Glance. *lobal Time Series*, published February 2018, retrieved on February 28, 2018 from http://www.ncdc.noaa.gov/cag/

[21] REference Antarctic Data for Environmental Research READER), Scientific Committee on Antarctic Research, http://www.antarctica.ac.uk/met/READER, Accessed 2018-04.

[22] Schneider U., Becker A., Finger P., Meyer-Christoffer A., Rudolf B., Ziese M., GPCC Full Data Reanalysis Version 6.0 at 1.0 : Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data, 2011, doi: 10.5676/DWD_GPCC/FD_M_V7_100.

[23] Shepard D, A two-dimensional interpolation function for irregularly-spaced data, Proceedings of the 1968 ACM National Conference, 1968, 517–524, doi:10.1145/800186.810616

[24] Smith T., Reynolds W., Ropelewski C., Optimal averaging of seasonal sea surface temperature and associated confidence intervals 1860-1989), J. Clim., 1994,7, 949-964.

[25] Stahl K., Moore R., Floyer J., Asplin M., McKendry I., Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density, Agricultural and forest meteorology, 2006,139, 224-236.

[26] Vinnikov K., Groisman P. , Lugina K., Empirical data on modern global climate changes temperature and precipitation), J. Clim., 1990, 3, 662-677.