Journal of the Korean Statistical Society (2007), 36: 2, pp 209–228

NETWORK DESIGN AND PREPROCESSING FOR MULTI-SCALE SPHERICAL BASIS FUNCTION REPRESENTATION†

HEE-SEOK OH[1] AND DONGHOH KIM[2]

ABSTRACT

Given scattered surface air temperatures observed by a network of weather stations, it is an important problem to estimate the entire temperature field for every location on the globe. Recently, a multi-scale spherical basis function (SBF) representation was proposed by Li (1999) for representing scattered data on the sphere. However, for a successful application of Li (1999)'s method, some practical issues such as network design, bandwidth selection of SBFs and initial coefficients are to be resolved. This paper proposes automatic procedures to design network and to select bandwidths. This paper also considers a preprocessing problem to obtain a stable initial coefficients from scattered data. Experiments with real temperature data demonstrate the promising empirical properties of the proposed approaches.

AMS 2000 subject classifications. Primary 62G05; Secondary 62G08.
Keywords. Bandwidth, multi-resolution, network design, preprocessing, spherical basis function.

1. INTRODUCTION

Suppose that we have, as shown in Figure 1.1, scattered surface air temperatures observed by a network of weather stations x_i satisfying

$$y_i = f(x_i) + \epsilon_i, \qquad i = 1, 2, \ldots, n,$$

Received October 2006; accepted January 2007.
†The work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2005-070-C00021).
[1]Department of Statistics, Seoul National University, Seoul 151-747, Korea (e-mail: heeseok@snu.ac.kr)
[2]Corresponding author. Department of International Management, Hongik University, Jochiwon 339-701, Korea (e-mail: donghohk@hongik.ac.kr)

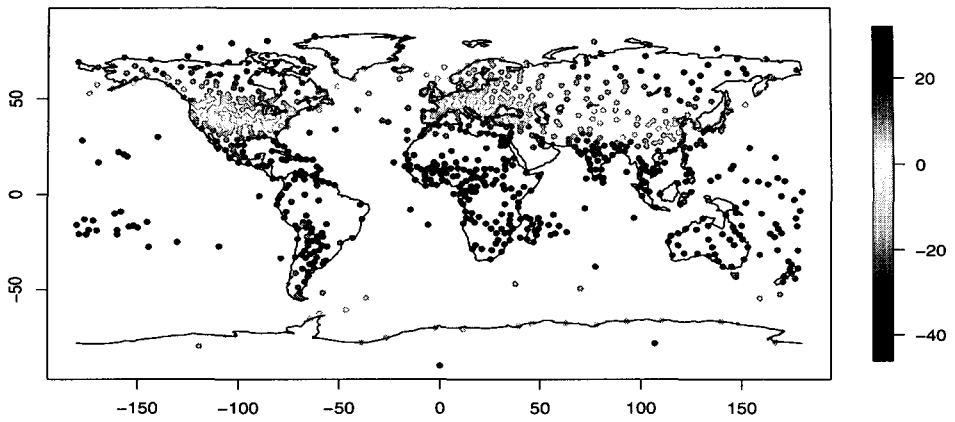

FIGURE 1.1 *Average winter temperatures observed at 913 weather stations during year 1970-1971.*

where ϵ_i denotes the measurement errors. We want to estimate the temperature field $f(x)$ for every location x on the globe. Figure 1.1 shows the average winter temperatures during year 1970-1971. The original large data set was organized by Jones *et al.* (1991). A subset of the data set was used by Luo and Wahba (1998), Li (1999) and Oh and Li (2004) for their analysis.

Being capturing global waves without good localization properties, spherical harmonics are not very efficient in representing such data as shown in Figure 1.1 that have inherent multi-scale characteristics (for example, global patterns of northern hemisphere winter coupled with local anomalies of different sizes in areas such as the Andes, Central Siberia, and Central Canadian Shield). The spherical smoothing spline method tends to produce uniformly smooth results, even though the data have intrinsic multi-scale structure (Oh and Li, 2004). Wavelets with localization properties are particularly effective in representing multi-scale phenomena that comprise activities of different scales at different locations. Moreover, their orthogonality gives rise to multi-scale decompositions that make wavelets a powerful tool for extracting the field's activities at different scales or detecting regional anomalies from global trends.

Several procedures have been proposed for constructing spherical wavelets (SWs). Narcowich and Ward (1996) proposed a construction method of SWs for scattered data on the unit sphere. SWs of Narcowich and Ward (1996) are constructed as the basis functions of the orthogonal complements subspaces pro-

duced by projecting the linear combinations of localized spherical basis functions (SBFs) onto smaller subspaces obtained by successively removing subsets of the SBFs. Since SWs and SW representations of Narcowich and Ward (1996) are designed for scattered data, their method can be easily applied to represent meteorological data which is collected over the surface of the earth via satellites or ground stations with the advantages of spherical wavelets. However, their spherical wavelets always have the same spatial scale that only depends on the pre-determined bandwidth parameter regardless of the resolution level or the distribution of the station. In climate study, it is important to detect local activities from global trends. To achieve this, we need a multi-scale analysis that can extract local patterns of different sizes of anomalies. The SW representation suggested by Narcowich and Ward (1996) may not be effective for this purpose because it relies on spherical wavelets that have a constant size of bandwidth. To solve the above problem, Li (1999) introduced multi-scale SBF representation in which SBFs of different sizes of bandwidths are employed.

In this paper, we focus on some practical issues for implementation of SW methods: (1) a choice of network design and bandwidth of SBFs, and (2) preprocessing of multi-scale analysis. Since appropriately designed network and properly selected bandwidths ensure stable multi-scale SBF representation and meaningful multi-scale interpretation, these are important for successful applications of SW methods. We propose fast automatic procedures for selecting network design and bandwidths of SBFs without considering physical and geographical constraints, which is desirable in applications such as data compression and objective analysis with a large data.

In addition, as pointed out by Oh and Li (2004), for the scattered data, a multi-scale analysis based on the discrete samples of the underline function is inappropriate, because it produces severely biased representations of the function. We consider a preprocessing to obtain the initial coefficients from the scattered spherical data for appropriate multi-resolution analysis.

The rest of the paper is organized as follows. In Section 2, we first review the SBF representation of Narcowich and Ward (1996) and the multi-scale SBF representation of Li (1999). Section 3 presents automatic selection methods of network design and bandwidths of SBFs. In Section 4, a preprocessing of multi-scale analysis is discussed. In Section 5, we apply the proposed methods to a real temperature data. Lastly, we present the conclusion in Section 6.

2. MULTI-SCALE SPHERICAL BASIS FUNCTION REPRESENTATION

In this section, we review the SBF representation proposed by Narcowich and Ward (1996) and Li (1999) for our application.

2.1. Spherical basis function representation

Narcowich and Ward (1996) proposed spherical wavelets for scattered data on the sphere from spherical basis functions (SBFs), $\phi(x, x_i) = \phi(x \cdot x_i)$, where $x \cdot x_i$ is the cosine of the angle between x and x_i. They suggested a SBF representation of a field $f(x)$ in the form of

$$\hat{f}(x) = \sum_{i=1}^{n} \beta_i \phi(x \cdot x_i).$$

For a valid SBF, the matrix $\Phi := [\phi(n_i, n_j)]_{i,j=1}^{n}$ is to be positive matrix. This property can be preserved by the SBFs if in the Legendre series $\phi(x) = \sum_{l=0}^{\infty} g_l \times P_l(x)$, the coefficients g_l are positive. Here $P_l(x)$ denotes the Legendre polynomial of degree l satisfying $\int_{-1}^{1} P_l^2(x)dx = 1$. A simple and useful example of SBF is the generating function of Legendre polynomials

$$\phi(x; \eta) := \frac{1 - \eta^2}{(1 - 2\eta x + \eta^2)^{3/2}} = \sum_{l=0}^{\infty} (2l + 1)\eta^l P_l(x),$$

where $\eta \in (0, 1)$ is a bandwidth parameter. We call this function the Poisson kernel. Figure 2.1 shows the normalized Poisson kernel $\tilde{\phi}(\cos\vartheta; \eta) := \{(1-\eta)^2/(1+\eta)\}\phi(\cos\vartheta; \eta)$ as a function of the angle variable ϑ which satisfies $\tilde{\phi}(\cos(0); \eta) = 1$. As can be seen, the Poisson kernel has a peak at angle 0 (*i.e.* $x = x'$) and decreases monotonically as the angle deviates from 0 to $\pm\pi$. Therefore, one can regard η as a spatial bandwidth parameter of the Poisson SBF representation: small bandwidth when η is large and large bandwidth when η is small. Therefore, a spatially localized representation can be achieved by setting η near unity.

To construct spherical wavelets, Narcowich and Ward (1996) considered the characterization of the loss in a SBF representation as stations are removed progressively from the original network in the representation. An interesting question is how to describe the loss in the SBF representation when some stations are removed. It is this answer that defines spherical wavelets and results in a spherical wavelet representation.

Let $\mathcal{N}_1 := \{x_i\}_{i=1}^{n}$ be the original network of stations, and $\mathcal{N}_2 := \{x_i\}_{i=1}^{m}$, $(m < n)$ be a smaller network obtained by deleting the last $n - m$ stations from

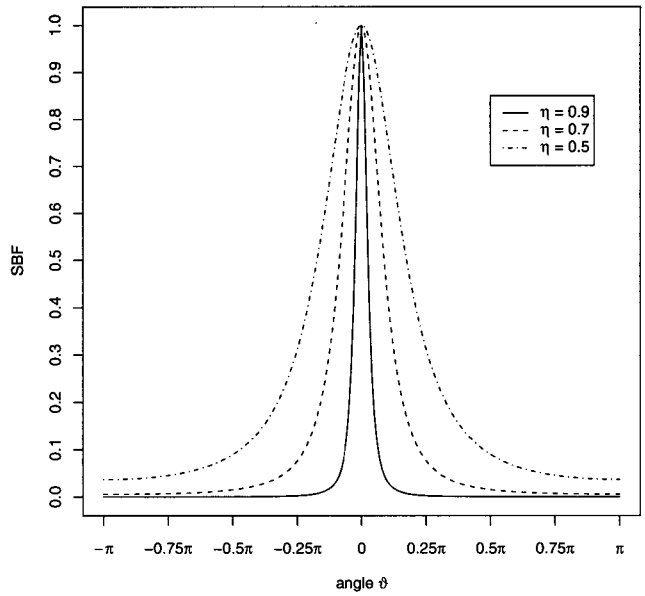

FIGURE 2.1 *Plot of normalized Poisson kernel* $\tilde{\phi}(\cos\vartheta;\eta)$.

$\mathcal{N}_1$. Given the observation from $f(x)$, let

$$f_1(x) := \sum_{i=1}^{n} \beta_{1,i}\phi(x \cdot x_i) = \boldsymbol{\beta}_1^T \boldsymbol{\phi}_1(x) \tag{2.1}$$

be an SBF representation of $f(x)$. If the $\beta_{1,i}$ vary, the equation (2.1) generates a collection of SBFs, $\mathcal{V}_1 := \text{span}\{\phi(x \cdot x_i) : x_i \in \mathcal{N}_1\}$. The SBF representation using the smaller network $\mathcal{N}_2$ is

$$f_2(x) := \sum_{i=1}^{m} \beta_{2,i}\phi(x \cdot x_i) = \boldsymbol{\beta}_2^T \boldsymbol{\phi}_2(x).$$

Since $f_1(x)$ is known throughout the sphere, one can choose $f_2(x)$ to best approximate $f_1(x)$ in terms of least squares, that is, by choosing $\beta_{2,i}$ such that

$$\underset{\beta_{2,i}}{\text{argmin}} \int \left| f_1(x) - \sum_{i=1}^{m} \beta_{2,i}\phi(x \cdot x_i) \right|^2 d\Omega(x), \tag{2.2}$$

where the integral is over the entire sphere Ω. This criterion leads to

$$\boldsymbol{\beta}_2 := \text{vec}\{\beta_{2,i}\}_{i=1}^m = \boldsymbol{A}_2^{-1}\boldsymbol{b}_2,$$

where the matrices $\boldsymbol{A}_2$ and $\boldsymbol{b}_2$ are given by

$$\boldsymbol{A}_2 := [\langle \phi(\cdot x_i), \phi(\cdot x_j) \rangle]_{i,j=1}^m \quad \text{and} \quad \boldsymbol{b}_2 := \text{vec}\{\langle f_1(\cdot), \phi(\cdot x_i) \rangle\}_{i=1}^m. \tag{2.3}$$

The inner product in (2.3) is defined $\langle A(\cdot), B(\cdot) \rangle := \int A(\cdot x) B(x \cdot) d\Omega(x)$ for any square integrable spherical functions A and B. With $\boldsymbol{\beta}_2$ given by (2.2), the resulting field $f_2(x)$ is the projection of $f_1(x)$ on the space $\mathcal{V}_2 := \text{span}\{\phi(x \cdot x_i) : x_i \in \mathcal{N}_2\}$. If the difference (or residual) field is denoted by

$$r_1(x) := f_1(x) - f_2(x),$$

and the collection of $r_1(x)$ is denoted by $\mathcal{W}_1$, then we have $\mathcal{V}_1 = \mathcal{V}_2 \oplus \mathcal{W}_1$. The space $\mathcal{W}_1$ represents the new information in $\mathcal{V}_1$ that is not contained in the smaller space $\mathcal{V}_2$. Any field $f_1(x) \in \mathcal{V}_1$ can be decomposed as

$$f_1(x) = f_2(x) + r_1(x), \tag{2.4}$$

where $f_2(x) \in \mathcal{V}_2$ and $r_1(x) \in \mathcal{W}_1$. Narcowich and Ward (1996) regarded (2.4) as a "multi-resolution" analysis of $f_1(x)$, since $f_1(x)$ stems from a larger (possibly denser) network and thus has a higher spatial resolution than $f_2(x)$. Based on this interpretation, the field $r_1(x)$ may be regarded as the lost high resolution detail when representing $f_1(x)$ by $f_2(x)$.

Since the decomposition is based on the nested networks of stations, it is more meaningful to interpret $r_1(x)$ as local activities of $f_1(x)$ near the deleted stations that cannot be accounted for by the activities of $f_2(x)$. From this interpretation, $r_1(x)$ can be considered as a *local* component and $f_2(x)$ as a *global* component of the field $f_1(x)$. With the nested property of spaces, $\mathcal{V}_1 \supset \mathcal{V}_2 \supset \cdots \supset \mathcal{V}_L$ and $\mathcal{V}_l = \mathcal{V}_{l+1} \oplus \mathcal{W}_l$, $(l = 1, \ldots, L-1)$, the decomposition of (2.4) is generalized by

$$f_1(x) = f_l(x) + r_{l-1}(x) + \cdots + r_1(x), \quad (l = 2, \ldots, L).$$

2.2. Multi-scale spherical basis function representation

To overcome the single scale problem, Li (1999) proposed to employ a set of SBFs with different bandwidths that are adapted to the stations. To explain more precisely, let us define a nested sequence of networks $\mathcal{N}_1 \supset \mathcal{N}_2 \supset \cdots \supset \mathcal{N}_L$, where $\mathcal{N}_l := \{x_i : i = 1, \ldots, n_l\}$ for $l = 1, \ldots, L$ and $n_{L+1} := 0 < n_L < \cdots < n_2 < n_1 := n$. For a given field $f(x)$, a *multi-scale SBF representation* is expressed as

$$\hat{f}(x) = \sum_{l=1}^{L} \sum_{i=n_{l+1}+1}^{n_l} \beta_{l,i} \phi_l(x \cdot x_i). \tag{2.5}$$

All stations in the l-th subnetwork $\mathcal{M}_l = \{x_i : i = n_{l+1} + 1, \ldots, n_l\}$ employ the same SBF $\phi_l(x)$, but different SBFs with different scales (bandwidths) are assigned to the L subnetworks. In this representation, a larger bandwidth is allowed for sparsely located stations and a smaller bandwidth for densely located stations. Furthermore, the nested networks $\mathcal{N}_l$ can be arranged so that the sparseness of stations in $\mathcal{N}_l$ increases with the increase of l. One can choose the multi-scale SBFs so that the bandwidth of $\phi_l(x)$ increases with the sparseness of $\mathcal{N}_l$, and thus the variable l can be truly regarded as a *scale parameter.*

To apply the multi-scale SBF representation in (2.5), the appropriate $\phi_l(x)$ and $\mathcal{N}_l$ should be chosen in advance. Suppose that we choose $\phi_l(x)$ and $\mathcal{N}_l$. The decomposition step of the multi-scale SBF representation works in the same way as the single-scale SBF representation. As before, any field $f_l(x) \in \mathcal{V}_l$ can be decomposed as (2.4) with $f_{l+1}(x) \in \mathcal{V}_{l+1}$ and $r_l(x) \in \mathcal{W}_l$, where $\mathcal{V}_l := \text{span}\{\phi_{l'}(x \cdot x') : x' \in \mathcal{M}_{l'}; l' = l, \ldots, L\}$. However, the fundamental difference between the two decompositions is in the structure of the space $\mathcal{V}_l$. In the decomposition using the multi-scale SBFs, the space $\mathcal{V}_l$ is composed of multi-scale SBFs $\phi_{l'}(x \cdot x')$ with scale l' greater than or equal to l. Since l can be considered as scale parameter for SBFs $\phi_l(x)$ and networks $\mathcal{N}_l$, the space $\mathcal{V}_l$ can be interpreted as having a smaller scale (*i.e.* a higher spatial resolution) than the space $\mathcal{V}_{l+1}$. The space $\mathcal{W}_l$ contains a higher resolution detail of $\mathcal{V}_l$ that cannot be explained by $\mathcal{V}_{l+1}$.

3. NETWORK DESIGN AND BANDWIDTH SELECTION

As mentioned early, a judiciously designed network $\mathcal{N}_L$ is required for a stable multi-scale SBF representation of (2.5). To obtain $\mathcal{N}_L$ we can start with a few sparsely located stations over the area and gradually increase the density by adding more stations so that we can obtain $\mathcal{N}_l = \mathcal{N}_{l+1} \cup \mathcal{M}_l$ sequentially for $l = L-1, \ldots, 1$. This is a bottom-up design approach. In this section, we suggest some schemes to choose the nested networks $\mathcal{N}_l$ systematically for performing the bottom-up design approach. Our network design depends only on the location of data and the type of grid which is predetermined without considering geophysical information. Therefore, this approach is fast and efficient so that it will be desirable to work with a large data set.

3.1. Type of grid

Before explaining the steps of network design, we propose two grid types, and introduce Göttelmann's grid by Göttelmann (1996). Each grid type has a

regular and a reduced grid, respectively. The reduced grid is designed to overcome the regular grid problem of a strong concentration of points near the poles. But since the network is selected by the relation of data-observing sites and a grid type, as will be mentioned later, we can not assure that the reduced grid produces networks that can better represent spherical fields. Let us define $[0, 2\pi]$ as the range of longitude and $[0, \pi]$ as the range of latitude to obtain grid points. Then, by simple transformation, all grid points can be located on $[-180°, 180°]$ as longitude and $[-90°, 90°]$ as latitude.

3.1.1. Rectangular regular grid. We define the index set, for $l \in \mathbb{N}$,

$$K_l^h := \{(i,k) : k = 0, 1, \ldots, 2^l; i = 0, 1, \ldots, 2^l\}. \tag{3.1}$$

The regular grid by the index set in (3.1) is defined

$$\mathcal{T}_l^h := \{(\phi_{i,l}, \theta_{k,l}); (i,k) \in K_l^h\},$$

where grid points of longitude, $\phi_{i,l} = i\pi/2^{l-1}$ and grid points of latitude, $\theta_{k,l} = k\pi/2^l$. In case of $l = 1$, the grid $\mathcal{T}_1^h$ is

$$\mathcal{T}_1^h := \{(0,0), (0, \frac{\pi}{2}), (0,\pi), (\pi, 0), (\pi, \frac{\pi}{2}), (\pi,\pi), (2\pi, 0), (2\pi, \frac{\pi}{2}), (2\pi,\pi)\}.$$

3.1.2. Rectangular reduced grid. Define the index set $K_l^{h,r}$

$$K_l^{h,r} := \{(i,k) : k = 0, 1, \ldots, 2^l; i = 0, 1, \ldots, \frac{2^l - 2^{r_{k,l}^h}}{2^{r_{k,l}^h}} + 1\} \subseteq K_l^h,$$

where the control parameter $r_{k,l}^h$ is given as

$$r_{k,l}^h = \begin{cases} 0, & \frac{\pi}{4} \le \theta_{k,l} \le \frac{3\pi}{4}, \\ \text{round}\{l - \log_2(\pi k)\}, & 0 < \theta_{k,l} < \frac{\pi}{4}, \\ \text{round}\{l - \log_2(\pi(2^l - k))\}, & \frac{3\pi}{4} < \theta_{k,l} < \pi, \\ l - 1, & \theta_{k,l} = 0, \pi. \end{cases}$$

Here round$\{x\}$ denotes the closest integer to x. The reduced grid is given by

$$\mathcal{T}_l^{h,r} := \{(\phi_{i,l}, \theta_{k,l}); (i,k) \in K_l^{h,r}\} \subseteq \mathcal{T}_l^h,$$

where $\phi_{i,l} = i2^{r_{k,l}^h}\pi/2^{l-1}$ are the grid points of longitude and $\theta_{k,l} = k\pi/2^l$ denote the grid points of latitude.

3.1.3. Modified Göttelmann's regular grid. For $l \in \mathbb{N}$, we define the index set

$$K_l^s := \{(i,k) : k = 0,1,\ldots,2^l; i = 0,1,\ldots,2^{l+1}\}, \tag{3.2}$$

where l is the index of level, i is the index of grid point of longitude and k denotes the index of grid point of latitude. The regular grid induced by the index set in (3.2) is defined

$$\mathcal{T}_l^s := \{(\phi_{i,l}, \theta_{k,l}); (i,k) \in K_l^s\},$$

where grid points of longitude, $\phi_{i,l} = i\pi/2^l$ and grid points of latitude, $\theta_{k,l} = k\pi/2^l$. Obviously, the sequence $\{\mathcal{T}_l^s\}_{l\in\mathbb{N}}$ of grids is hierarchical: $\mathcal{T}_l^s \subseteq \mathcal{T}_{l+1}^s$ for all $l \in \mathbb{N}$. For the simplest example, let $l = 1$, the grid $\mathcal{T}_1^s$ is

$$\mathcal{T}_1^s := \{(\phi_{i,1}, \theta_{k,1}); (i,k) \in K_1^s\},$$

where $\phi_{i,1}$ are $0, \pi/2, \pi, 3\pi/2, 2\pi$ and $\theta_{k,1}$ are $0, \pi/2, \pi$ and $K_1^s = \{(i,k) : i = 0,1,2,3,4;\ k = 0,1,2\}$.

3.1.4. Modified Göttelmann's reduced grid. The index set $K_l^{s,r}$ for reduced grid is defined as

$$K_l^{s,r} := \{(i,k) : k = 0,1,\ldots,2^l; i = 0,1,\ldots,\frac{2^{l+1} - 2^{r_{k,l}^s}}{2^{r_{k,l}^s}} + 1\} \subseteq K_l^s,$$

where the control parameter for dropping grid points according to latitude level, $r_{k,l}^s$ is defined to be

$$r_{k,l}^s = \begin{cases} 0, & \frac{\pi}{4} \le \theta_{k,l} \le \frac{3\pi}{4}, \\ \lfloor l - \log_2(\pi k) \rfloor, & 0 < \theta_{k,l} < \frac{\pi}{4}, \\ \lfloor l - \log_2(\pi(2^l - k)) \rfloor, & \frac{3\pi}{4} < \theta_{k,l} < \pi, \\ l - 1, & \theta_{k,l} = 0, \pi. \end{cases}$$

Here $\lfloor x \rfloor$ denotes the largest integer less than or equals to x. The reduced grid is given by

$$\mathcal{T}_l^{s,r} := \{(\phi_{i,l}, \theta_{k,l}); (i,k) \in K_l^{s,r}\} \subseteq \mathcal{T}_l^s,$$

where grid points of longitude, $\phi_{i,l} = i2^{r_{k,l}^s}\pi/2^l$ and grid points of latitude, $\theta_{k,l} = k\pi/2^l$.

3.1.5. Göttelmann's regular grid. Define the index set K_l^G

$$K_l^G := \{(i,k) : k = 0,1,\ldots,2^{l+1}; i = 0,1,\ldots,2^{l+2}\}.$$

By setting grid points of longitude and grid points of latitude

$$\phi_{i,l} = i\frac{\pi}{2^{l+1}} \quad \text{and} \quad \theta_{k,l} = k\frac{\pi}{2^{l+1}},$$

we obtain the grid $\mathcal{T}_l^G := \{(\phi_{i,l}, \theta_{k,l}); (i,k) \in K_l^G\}$. The grid $\mathcal{T}_l^G$ is the same as $\mathcal{T}_{l+1}^s$.

3.1.6. Göttelmann's reduced grid. Let grid points of longitude, $\phi_{i,l} = i2^{r_{k,l}^G}\pi/2^{l+1}$ and grid points of latitude, $\theta_{k,l} = k\pi/2^{l+1}$. The control parameter, $r_{k,l}^G$ is defined to be

$$r_{k,l}^G = \begin{cases} 0, & \frac{\pi}{4} \le \theta_{k,l} \le \frac{3\pi}{4}, \\ \lfloor l + 1 - \log_2(\pi k) \rfloor, & 0 < \theta_{k,l} < \frac{\pi}{4}, \\ \lfloor l + 1 - \log_2(\pi(2^{l+1} - k)) \rfloor, & \frac{3\pi}{4} < \theta_{k,l} < \pi, \\ l - 1, & \theta_{k,l} = 0, \pi. \end{cases}$$

Then, the reduced grid is given by

$$K_l^{G,r} := \{(i,k) : k = 0, 1, \ldots, 2^{l+1}; i = 0, 1, \ldots, \frac{2^{l+2} - 2^{r_{k,l}^G}}{2^{r_{k,l}^G}} + 1\} \subseteq K_l^G.$$
$$\mathcal{T}_l^{G,r} := \{(\phi_{i,l}, \theta_{k,l}); (i,k) \in K_l^{G,r}\} \subseteq \mathcal{T}_l^G.$$

The grid $\mathcal{T}_l^{G,r}$ is similar to the grid $\mathcal{T}_{l+1}^{s,r}$. But the grid points near the two poles are different from $\mathcal{T}_{l+1}^{s,r}$.

Figure 3.1 illustrates the grid points obtained from the above three grid types.

3.2. Network design

Here we propose a fast method for network design by bottom-up design approach. Our network is designed systematically based on the location of data and the type of grid. As the resulting network generated by the proposed method, we expect that in each level (resolution) $\mathcal{N}_l$ and $\mathcal{M}_l$, stations are distributed over the sphere as uniformly as possible, and stations between two levels are not too close so that we can apply SBFs with different bandwidths to these stations. Here are the steps of proposed network design.

1. Obtain the center points c_i of each grid box from the grid $\mathcal{T}_1$ which has the smallest number of grid points. Form a territory (circle) D_i around a given center point. Thus the number of territories should be equal to the number of center points. Then compute the geodesic distance from locations of data

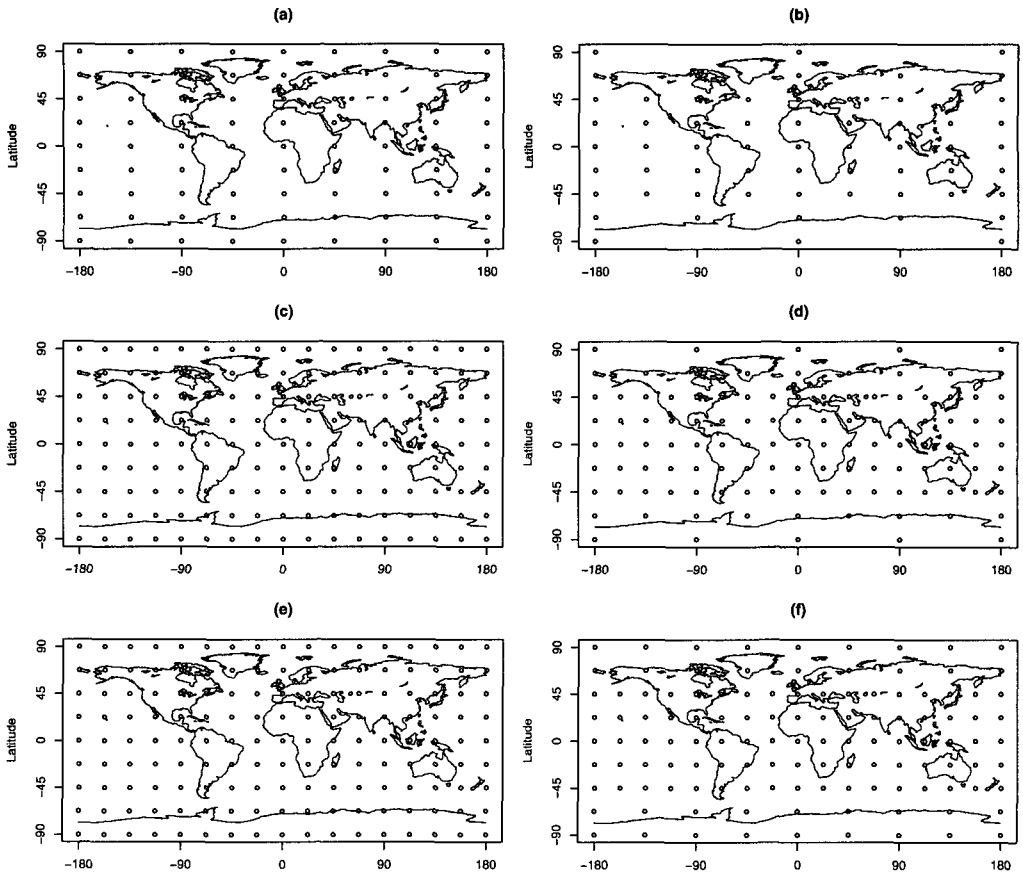

FIGURE 3.1 *(a) Regular grid points of rectangular grid type for $l = 3$; (b) Reduced grid points of rectangular grid type for $l = 3$; (c) Regular grid points of modified Göttelmann's grid type for $l = 3$; (d) Reduced grid points of modified Göttelmann's grid type for $l = 3$; (e) Regular grid points of Göttelmann's grid type for $l = 2$; and (f) Reduced grid points of Göttelmann's grid type for $l = 2$.*

$d_j \in D_i$ within a territory to the center point of the territory, $\{\arccos(c_i \cdot d_j)\}$ and find the location of data which has the minimum distance to the center point. A network with sparsely located stations, $\mathcal{N}_L$ is made up from these locations.

2. From the next grid $\mathcal{T}_2$, compute the center point c_i of each grid box. As with step 1, draw a territory (circle) D_i around a given center point and find the closest station to the center point within its territory, D_i, for all i. With the selected stations, build the set, say $\mathcal{M}^*_{L-1}$. Then compare

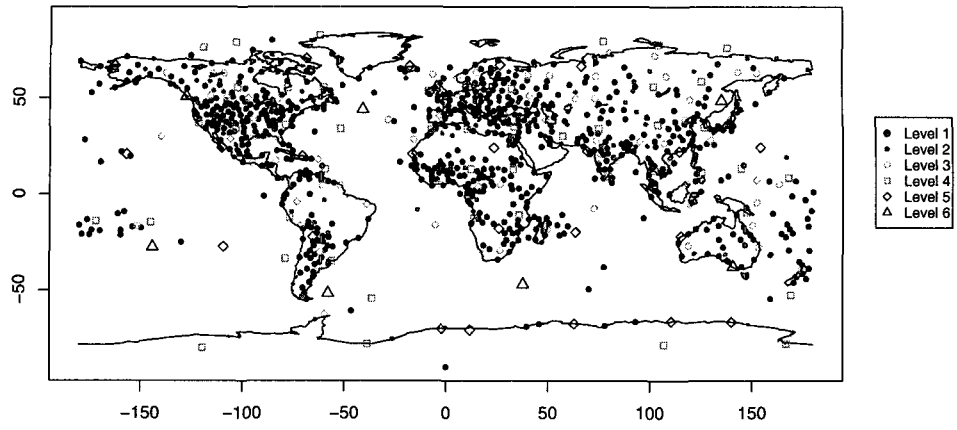

FIGURE 3.2 *Nested network with 6 levels for 913 stations to observe the surface air temperatures of year 1970-1971.*

these stations with stations in $\mathcal{N}_L$. That is, compute the geodesic distance from the stations in $\mathcal{M}_{L-1}^*$ to the stations in $\mathcal{N}_L$. If some distance between two locations is closer than a criterion, the location is deleted from $\mathcal{M}_{L-1}^*$. After the comparison procedure, the network $\mathcal{M}_{L-1}$ is formed. Note that $\mathcal{N}_{L-1} = \mathcal{N}_L \cup \mathcal{M}_{L-1}$.

3. Repeat step 2 for remaining data and grid points from $\mathcal{T}_l$ for $l = 3, 4, \ldots$ until the longitude of grid box from $\mathcal{T}_l$ becomes 5 degrees. Finally, $\mathcal{N}_1 = \mathcal{N}_L \cup \mathcal{M}_{L-1} \cup \cdots \cup \mathcal{M}_1$ is obtained.

For performing the above steps, it is necessary to consider the territory and deletion criterion for each step. The radius of territory should not be too large to reduce the possibilities which select stations located far from the center point of each grid box. Note that when there are no stations in a grid box, and the radius of territory is too large, a station in another grid box can be selected as closest stations to the center point of a grid box.

Figure 3.2 shows a nested network with 6 levels for 913 stations which recorded the surface air temperatures of year 1970-1971. The modified Göttelmann's regular grid is used for this network design.

3.3. The selection of bandwidth

We now consider bandwidths of SBFs to be employed for each network $\mathcal{N}_L$, $\mathcal{M}_{L-1}, \ldots, \mathcal{M}_1$. In view of SBF $\phi_l(x)$, for sparsely located stations, we assign

SBFs with larger bandwidths to cover such area. As the density of stations increases, the bandwidth is decreased in order to represent local area activities of the field efficiently.

First of all, let us look at the scale parameter of 1-d wavelets defined on a real line in order to preview the bandwidth of spherical wavelets and to understand the procedure of the choice of bandwidths to be explained later. From the definition of 1-d wavelets

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(\frac{t - 2^j k}{2^j}\right), \text{ where } \psi \text{ is mother wavelet,}$$

we know that the scale parameter j is decided systematically. That is, as the scale parameter j increases 1 unit every time, the length of support of the wavelet increases twice as much as before. Thus there is a relationship between the scale parameter and the length of the wavelet.

Similarly, we suggest that the bandwidths η_l of SBFs be chosen such that

$$\eta_{L-l} = e^{-\rho_l}, \quad l = 1, 2, \ldots, L-1, \tag{3.3}$$

where $\rho_l = \rho^*/2^l$. The ρ^* can be obtained by $\rho^* = -\log \eta_L$, where η_L is the bandwidth of the coarsest network level L. As mentioned before, the networks are related to a grid. As the level l decreases, a grid box related to the level decreases one fourth in size (both intervals of latitude and longitude decrease by a half). Thus, the area covered by a SBF decreases as the level l decreases. That is the reason we use the bandwidths η_l produced by (3.3). Hence, if the bandwidth η_L of the sparsest level L is decided, all bandwidths can be obtained systematically.

Now let us discuss how we can get the bandwidth η_L. From simple geometry, the surface area covered by surface mass distribution with variance σ^2 over unit sphere Ω is $2\pi(1 - \sqrt{1 - \sigma^2})$. Since the variance of SBF or spherical wavelet from Poisson kernel is $\sigma^2 = ((1 - \eta^2)/(1 + \eta^2))^2$, the surface area covered by SBFs is $2\pi(1 - \sqrt{1 - ((1 - \eta^2)/(1 + \eta^2))^2})$. Note that as the bandwidth of SBF is close to 0, the surface area covered becomes 2π. Since the total surface area of the unit sphere is 4π, in this case, we need two stations to cover the whole sphere Ω. In another extreme case, as the η goes to 1, the surface area covered is close to 0. To cover the whole sphere Ω, we need infinitely many stations. Under the assumption that the stations are distributed equally over the sphere, it can be easily known how many stations are needed in order to cover the whole sphere with fixed η and how large the bandwidth of SBF is needed to cover the whole

sphere when the number of stations are fixed from the following

$$\#\text{ of stations} = n = \frac{2}{1-\sqrt{1-\left(\frac{1-\eta^2}{1+\eta^2}\right)^2}} \quad \text{and} \quad \eta = \sqrt{\frac{1-a_n}{1+a_n}},$$

where $a_n = \sqrt{1-(1-2/n)}$. Stations in the sparsest network $\mathcal{N}_L$ based on any grid scheme are almost distributed equally. Thus, from the above equations, we can decide the bandwidth η_L. For example, we decide SBFs with 6 different bandwidths to be employed at networks with six levels in Figure 3.2. The result is that $\eta_1 = 0.97$ for $\mathcal{M}_1$, $\eta_2 = 0.95$ for $\mathcal{M}_2$, $\eta_3 = 0.90$ for $\mathcal{M}_3$, $\eta_4 = 0.81$ for $\mathcal{M}_4$, $\eta_5 = 0.67$ for $\mathcal{M}_5$ and $\eta_6 = 0.45$ for $\mathcal{N}_6$.

4. PREPROCESSING OF MULTI-SCALE SBF REPRESENTATION

For a given climate field $f(x)$, we have proposed a multi-scale SBF representation of the form

$$f_1(x) = \sum_{(l',i)\in\mathcal{I}_1} \beta_{1,i}\phi_{l'}(x \cdot x_i) = \boldsymbol{\beta}_1^T \boldsymbol{\phi}_1(x), \tag{4.1}$$

where $\mathcal{I}_1 := \{(l',i) : l' = 1,\ldots,L; i = N_{l'+1}+1,\ldots,N_{l'}\}$. The coefficients $\boldsymbol{\beta}_1$ can be obtained in many ways. The simplest example is observations $\boldsymbol{y}$. Once the $f_1(x)$ is obtained, all global fields $f_l(x)$ for $l = 2,\ldots,L$ and detail fields $r_l(x)$ for $l = 1,2,\ldots,L-1$ are decided by the multi-resolution analysis. Thus, whether the $f_1(x)$ has a good performance is very important for a good representation of a climate field. To that end, we consider two approaches, interpolation and approximation method.

We first discuss interpolation methods for obtaining coefficients $\boldsymbol{\beta}_1$. In (4.1), since the interpolation matrix $\boldsymbol{\Phi}_1 = [\phi_{l'}(x_i \cdot x_j)]_{i,j}$ for $l' = 1,\ldots,L$ is invertible, we can obtain $\tilde{\boldsymbol{\beta}}_1 = \text{vec}\{\tilde{\beta}_{1,j}\} = (\boldsymbol{\Phi}_1^T\boldsymbol{\Phi}_1)^{-1}\boldsymbol{\Phi}_1^T\boldsymbol{y}$ by the least squares method. The solution $\tilde{\boldsymbol{\beta}}_1$ gets values passing through all data points at the observation sites. Thus $f_1(x)$ in (4.1) is an interpolation function. As an extension, let us consider the interpolation method by penalized least squares. The penalized least squares solution is $\tilde{\boldsymbol{\beta}}_1^p = \text{vec}\{\tilde{\beta}_{1,i}^p\} = (\boldsymbol{\Phi}_1^T\boldsymbol{\Phi}_1 + \lambda\boldsymbol{Q})^{-1}\boldsymbol{\Phi}_1^T\boldsymbol{y}$. The value λ should be selected appropriately. Note that the penalized least squares becomes the ridge regression method by setting $\boldsymbol{Q} = \boldsymbol{I}$.

The second approach is to obtain the coefficients $\boldsymbol{\beta}_1$ using the approximation matrix $\boldsymbol{\Phi}_1^* = [\Phi_{l'}(x_i \cdot x_j)]_{i=1,\ldots,n,j=1,\ldots,m}$ which consists of $m(< n)$ selected sites

among all observation sites n whereas the interpolation method uses all observation sites for matrix $\mathbf{\Phi}_1$. Thus, the least squares estimator $\hat{\boldsymbol{\beta}}_1 = \text{vec}\{\hat{\beta}_{1,j}\} = (\mathbf{\Phi}_1^{*T}\mathbf{\Phi}_1^*)^{-1}\mathbf{\Phi}_1^{*T}\boldsymbol{y}$ provides an approximation field of $f_1(x)$. But when the selected sites are located with some gaps over the sphere, the coefficients $\hat{\boldsymbol{\beta}}_1$ by least squares estimation may not be stable because some outliers can have profound influence on $\hat{\boldsymbol{\beta}}_1$. Moreover, when some stations are very closely located, the coefficients $\hat{\boldsymbol{\beta}}_1$ may not provide stable estimates due to multicollinearity. Thus, we suggest the ridge regression estimation as a shrinkage estimation method for more stable coefficients, $\hat{\boldsymbol{\beta}}_1^r = \text{vec}\{\hat{\beta}_{1,i}^r\} = (\mathbf{\Phi}_1^{*T}\mathbf{\Phi}_1^* + \lambda \boldsymbol{I})^{-1}\mathbf{\Phi}_1^{*T}\boldsymbol{y}$. The value λ can be decided from the minimization of

$$\text{GCV}(\lambda) = \frac{\text{RSS}}{\left\{\frac{1}{n}\sum_i [1 - h_{ii}(\lambda)]\right\}^2},$$

where the $h_{ii}(\lambda)$ are the diagonal elements of $\boldsymbol{H}(\lambda) = \mathbf{\Phi}_1^*(\mathbf{\Phi}_1^{*T}\mathbf{\Phi}_1^* + \lambda \boldsymbol{I})^{-1}\mathbf{\Phi}_1^{*T}$.

5. APPLICATION TO TEMPERATURE DATA

In this section, we apply the proposed network design, bandwidth selection, and preprocessing to construct multi-scale SBF representations for average winter temperature of year 1970-1971 shown in Figure 1.1. In addition, we perform multi-resolution analysis to obtain global components $f_l(x)$ and local components $r_l(x)$ from a multi-scale SBF representation.

The multi-scale SBF representations for the 1970-1971 average winter global temperature with different network design, bandwidth selections, and preprocessing are displayed in Figures 5.1 and 5.2. Among all combinations, in Figure 5.1, we employ modified Göttelmann's regular grid with six levels for network design, and the bandwidths $\eta_1 = 0.97$, $\eta_2 = 0.95$, $\eta_3 = 0.90$, $\eta_4 = 0.81$, $\eta_5 = 0.67$, and $\eta_6 = 0.45$. As preprocessing, the interpolation method by penalized least squares (top panel) and the approximation approach by least squares (bottom panel) are adapted for obtaining $f_1(x)$. In Figure 5.2, we use the rectangular reduced grid, the bandwidths ($\eta_1 = 0.97$, $\eta_2 = 0.93$, $\eta_3 = 0.88$, $\eta_4 = 0.80$, $\eta_5 = 0.65$, and $\eta_6 = 0.40$), and the interpolation method by least squares (top panel) and the approximation approach by penalized least squares (bottom panel).

As expected, all multi-scale SBF representations capture local activities in some regions such as Siberia as a cold place and North-west of Australia as a hot spot as well as the global trends. However, as you can see, the results are varying with the choice of network design, bandwidth selection, and preprocessing. In

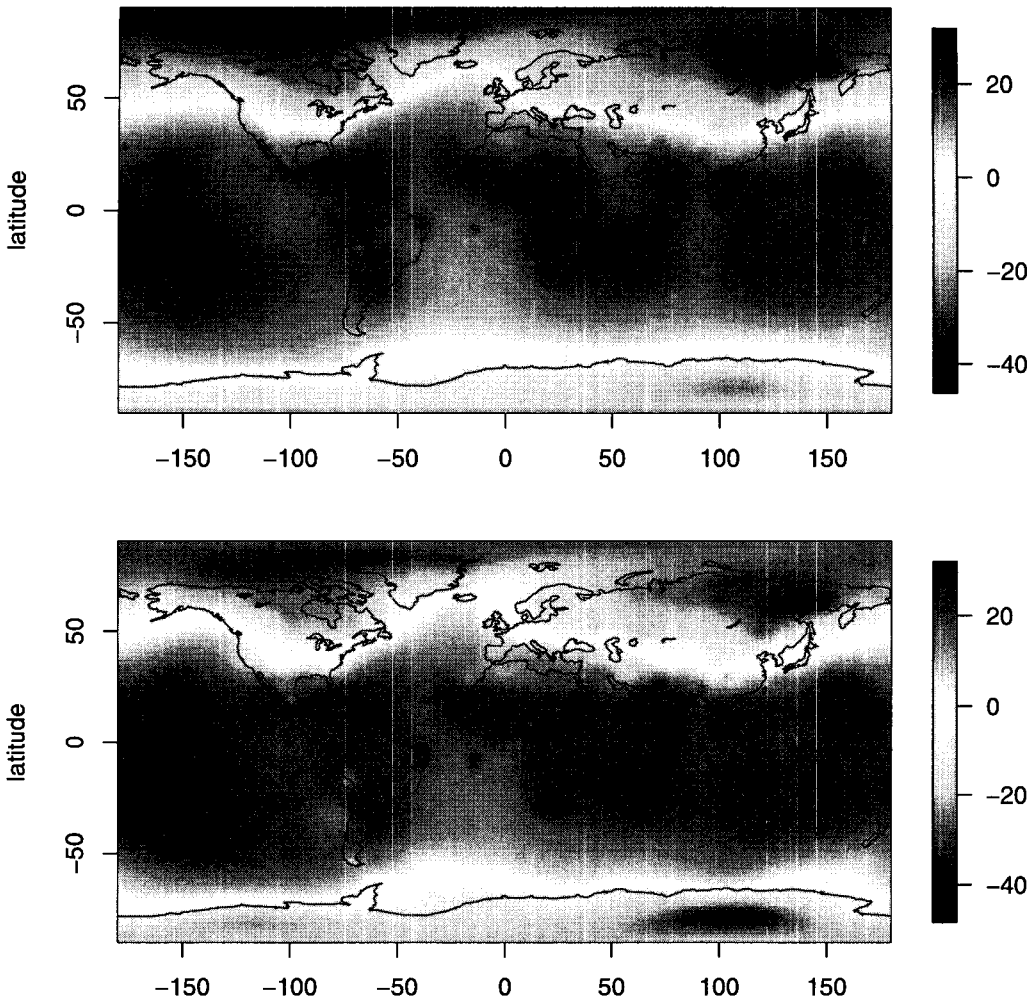

FIGURE 5.1 *Multi-scale SBF representation for 1970-1971 average winter temperature with the network generated by the modified Göttelmann's regular grid. The result in top panel is adapted by an interpolation method by penalized least squares, and the bottom one is implemented by an approximation approach by least squares.*

general, the multi-scale SBF representation with PLS interpolation performs well for detecting local activities from the global trends. Comparing between regular grid and reduced grid, activities near equator are similar but patterns are quite different near the south and north poles.

Figure 5.3 shows global components $f_l(x)$ and local components $r_l(x)$ for

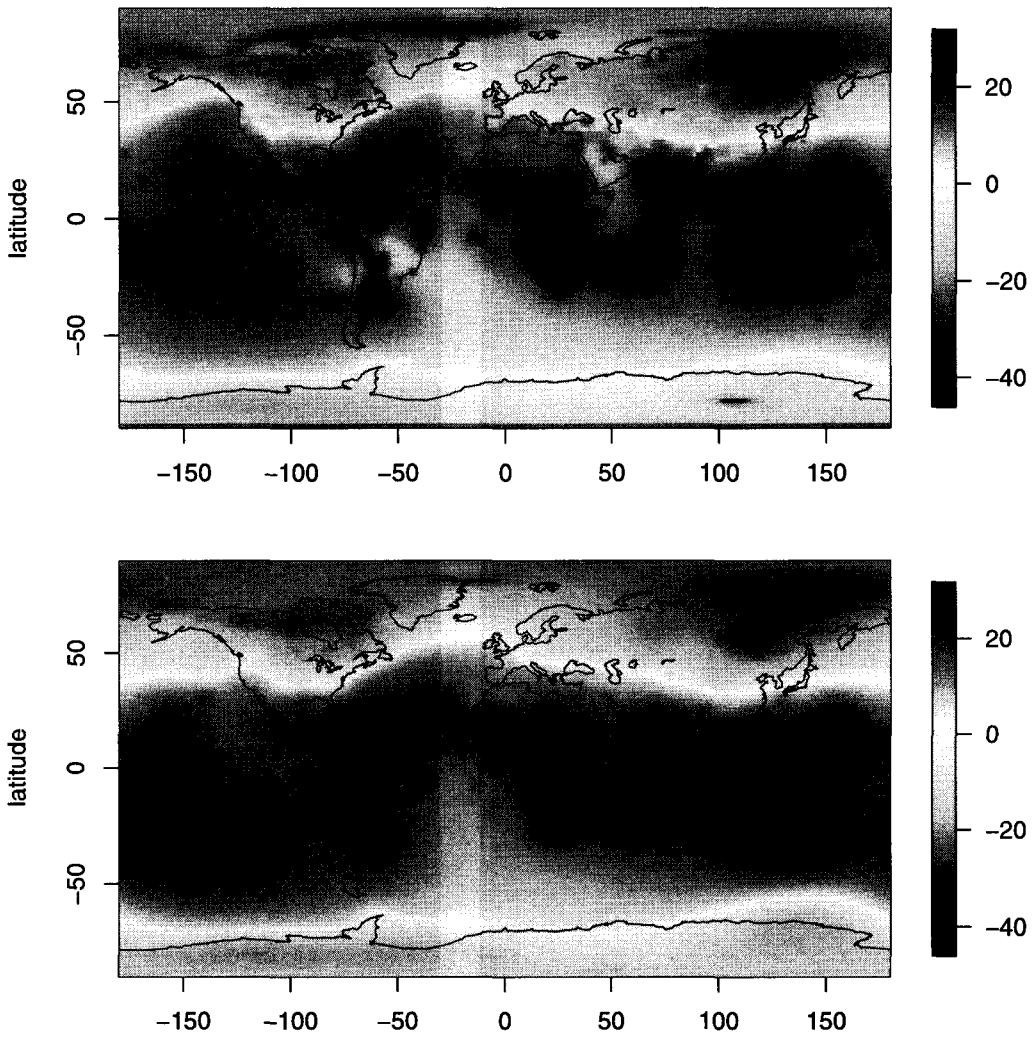

FIGURE 5.2 *Multi-scale SBF representation for 1970-1971 average winter temperature with the network generated by rectangular reduced grid. The result in top panel is adapted by an interpolation method by least squares, and the bottom one is implemented by an approximation approach by penalized least squares.*

the multi-scale SBF representation in the top panel of Figure 5.1 generated by the multi-resolution analysis. As you can see in Figure 5.3, we obtain various representations of a field $f(x)$ according to level of resolution. As the level of resolution l increases, the resulting approximation tends to represent only global

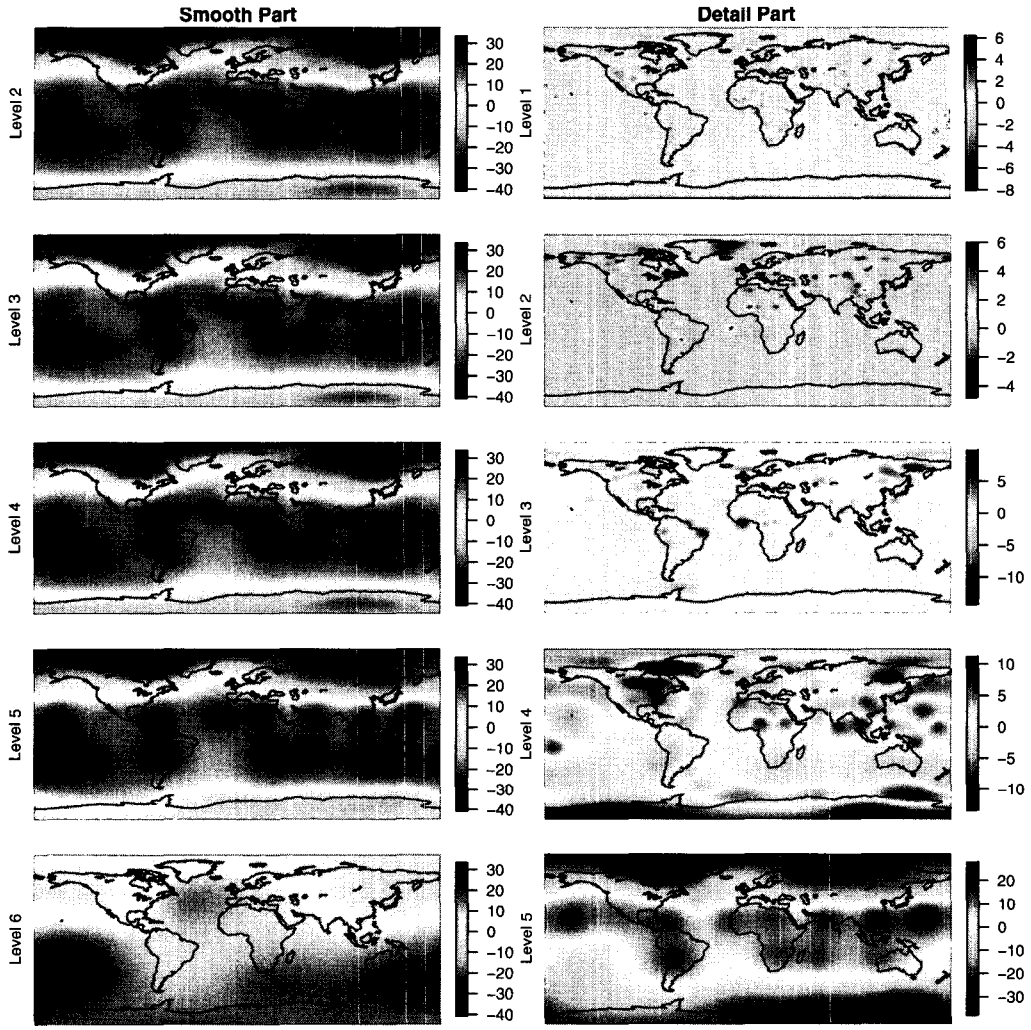

FIGURE 5.3 *Multiresolution analysis of the multi-scale SBF representation in the top panel of Figure 5.1,* $f_1(x)$. *Note that the field* $f_1(x)$ *is decomposed as* $f_1(x) = f_6(x) + r_1(x) + r_2(x) + r_3(x) + r_4(x) + r_5(x)$.

patterns of a field.

Spherical smoothing splines as a traditional approach provide a temperature reconstruction in Figure 5.4. As seen, the resulting fit determined by single smoothing parameter tends to be uniformly smooth and only represents the global trend with ignoring local phenomena. Thus, the traditional approach which depends on single-scale parameter may not be effective for representing various activities of the temperature field.

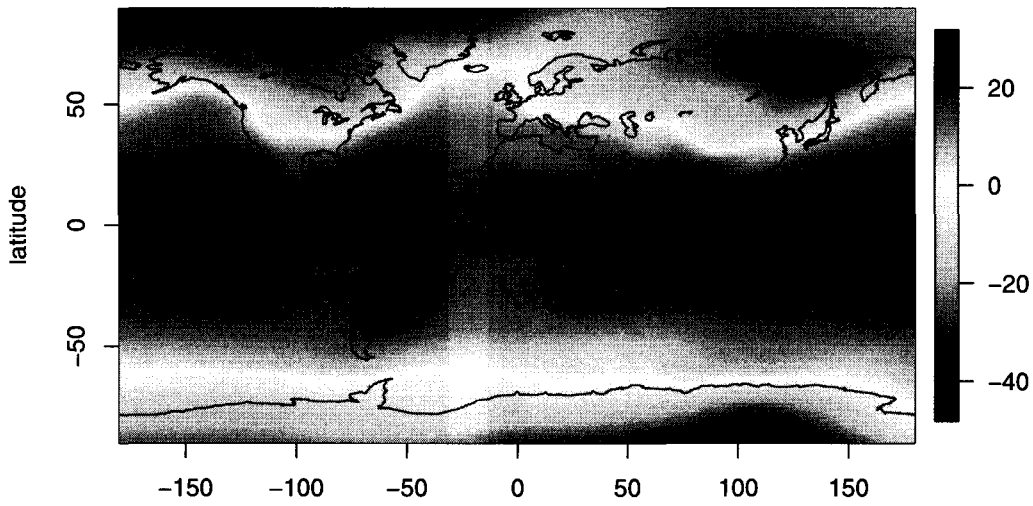

FIGURE 5.4 *Single-scale representation by spherical smoothing splines for 1970-1971 average winter temperature. The result is adapted by thin plate spline method and single smoothing parameter is chosen by generalized cross-validation.*

6. CONCLUSION

To ensure the stability of the multi-scale SBF representation and the corresponding multi-scale analysis, practical implementation of the method requires properly chosen bandwidths of SBFs and networks, and an appropriate preprocessing which obtains initial coefficients from scattered data. In this paper, we have proposed automatic procedures for network design $\mathcal{N}_l$ and bandwidth selection η_l. The proposed method is a kind of non-adaptive approach which does not depend on the response variable. We focus on a systematic way to make the best use of the advantage of a fast algorithm. This paper also discusses statistical estimation methods for initial coefficients from scattered data based on approximation and interpolation approach.

In the experimental study for global surface air temperatures, we have shown that multi-scale SBF representations are very powerful for detecting local activities as well as extracting global trends of temperature fields that cannot be easily detected by the traditional spherical smoothing spline method.

REFERENCES

GÖTTELMANN, J. (1996). "Locally supported wavelets on the sphere", *Preprint*, Johannes Gutenberg University, Mainz.

JONES, P. D., RAPER, S. C. B., CHERRY, B. S. G., GOODESS, C. M., WIGLEY, T. M. L, SANTER, B., KELLY, P. M., BRADLEY, R. S. AND DIAZ, H. F. (1991). An updated global grid point surface air temperature anomaly data set: 1851-1988. Environmental Sciences Division Publication No. 3520, U.S. Department of Energy.

LI, T-H. (1999). "Multiscale representation and analysis of spherical data by spherical wavelets", *SIAM Journal on Scientific Computing*, **21**, 924–953.

LUO, Z.,WAHBA, G. AND JOHNSON, D. R. (1998). "Spatial-temporal analysis of temperature using smoothing spline anova", *Journal of Climate*, **11**, 18–28.

NARCOWICH, F. J. AND WARD, J. D. (1996). "Nonstationary wavelets on the m-sphere for scattered data", *Applied and Computational Harmonic Analysis*, **3**, 324–336.

OH, H-S. AND LI, T-H. (2004). "Estimation of global temperature fields from scattered observations by a spherical-wavelet-based spatially adaptive method", *Journal of the Royal Statistical Society*, Ser. B, **66**, 221–238.