statistical models can be estimated for large spatial datasets, rather than as an approach for mitigating spatial multicollinearity. We consider more general basis expansions of remote coefficients in Section 2.3.

We assume the remote coefficients $\alpha(\boldsymbol{s}, \boldsymbol{r})$ can be well represented by weighted averages of remote coefficients $\alpha(\boldsymbol{s}, \boldsymbol{r}^*)$ at knot locations $\boldsymbol{r}_1^*, \ldots, \boldsymbol{r}_k^* \in \mathcal{D}_Z$, so we make the simplifying approximation that, for some weight function $h(\boldsymbol{r}, \boldsymbol{r}')$ and associated vector $\boldsymbol{h}^*(\boldsymbol{r}) = \left[h\left(\boldsymbol{r}, \boldsymbol{r}_j^*\right)\right]_{j=1}^k \in \mathbb{R}^k$, we can write

$$\alpha(\boldsymbol{s}, \boldsymbol{r}) = \sum_{j=1}^{k} h\left(\boldsymbol{r}, \boldsymbol{r}_j^*\right) \alpha\left(\boldsymbol{s}, \boldsymbol{r}_j^*\right) = \boldsymbol{h}^*(\boldsymbol{r})^T \boldsymbol{\alpha}^*(\boldsymbol{s}), \tag{8}$$

where $\boldsymbol{\alpha}^*(\boldsymbol{s}) = \left[\alpha\left(\boldsymbol{s}, \boldsymbol{r}_j^*\right)\right]_{j=1}^k \in \mathbb{R}^k$. The predictive process approach uses kriging to motivate a choice for the weight vector $\boldsymbol{h}^*(\boldsymbol{r})$, which induces a weight function h. Using Gaussian processes in Section 2.1 to model the remote coefficients implies that $\alpha(\boldsymbol{s}, \boldsymbol{r})$ and $\boldsymbol{\alpha}^*(\boldsymbol{s})$ are jointly normally distributed, yielding the conditional expectation for $\alpha(\boldsymbol{s}, \boldsymbol{r})$

$$E\left[\alpha(\boldsymbol{s}, \boldsymbol{r}) \middle| \boldsymbol{\alpha}^*(\boldsymbol{s})\right] = \boldsymbol{c}^*(\boldsymbol{r})^T R^{*-1} \boldsymbol{\alpha}^*(\boldsymbol{s}) \tag{9}$$

in which $\boldsymbol{c}^*(\boldsymbol{r}) = \left[C_\alpha\left\{(\boldsymbol{s}, \boldsymbol{r}), \left(\boldsymbol{s}, \boldsymbol{r}_j^*\right)\right\}\right]_{j=1}^k \in \mathbb{R}^k$ and $R^* \in \mathbb{R}^{k \times k}$ with entries $R_{ij}^* = C_\alpha\left\{\left(\boldsymbol{s}, \boldsymbol{r}_i^*\right), \left(\boldsymbol{s}, \boldsymbol{r}_j^*\right)\right\}$. Note that the assumption in (6) that C_α is stationary means that $\boldsymbol{c}^*(\boldsymbol{r})$ and R^* do not depend on $\boldsymbol{s}$, despite the term appearing in their definitions. The predictive process approach uses the conditional expectation (9) to define the weight vector $\boldsymbol{h}^*(\boldsymbol{r}) = R^{*-1}\boldsymbol{c}^*(\boldsymbol{r})$ in the approximation (8). Banerjee et al. (2008) show that these types of approximations are reduced rank projections that can capture large-scale spatial structures in data.

Beyond mitigating the statistical issue of multicollinearity, the predictive process approach relates the RESP model to spatially varying coefficient models (1) and also has a scientific interpretation for teleconnection. Using the reduced rank approximation (8) to