Supplementary Material for "Gaussian process aided function comparison using noisy scattered data"

S.1 Derivation for $c(\boldsymbol{x}, \boldsymbol{x}')$

The predictive mean for $f_1(.)$ given $\mathcal{D}_1$ is as follows:

$$\hat{f}_1(\boldsymbol{x}) = \boldsymbol{r}_1(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(1)},\mathbf{X}^{(1)}} + \sigma_\epsilon^2 \mathbf{I}_{n_1}]^{-1} \boldsymbol{y}^{(1)}.$$

Similarly, the predictive mean for $f_2(.)$ conditioned on $\mathcal{D}_2$ is given by:

$$\hat{f}_2(\boldsymbol{x}) = \boldsymbol{r}_2(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(2)}} + \sigma_\epsilon^2 \mathbf{I}_{n_2}]^{-1} \boldsymbol{y}^{(2)}.$$

Thus $c(\boldsymbol{x}, \boldsymbol{x}') = Cov(\hat{f}_2(\boldsymbol{x}) - \hat{f}_1(\boldsymbol{x}))$ is expressed as follows:

$$\begin{aligned}
&Cov(\hat{f}_2(\boldsymbol{x}) - \hat{f}_1(\boldsymbol{x})) \\
&= Cov(\boldsymbol{r}_2(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(2)}} + \sigma_\epsilon^2 \mathbf{I}_{n_2}]^{-1} \boldsymbol{y}^{(2)} - \boldsymbol{r}_1(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(1)},\mathbf{X}^{(1)}} + \sigma_\epsilon^2 \mathbf{I}_{n_1}]^{-1} \boldsymbol{y}^{(1)}) \\
&= Var(\boldsymbol{r}_2(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(2)}} + \sigma_\epsilon^2 \mathbf{I}_{n_2}]^{-1} \boldsymbol{y}^{(2)}) + Var(\boldsymbol{r}_1(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(1)},\mathbf{X}^{(1)}} + \sigma_\epsilon^2 \mathbf{I}_{n_1}]^{-1} \boldsymbol{y}^{(1)}) \\
&- 2\, Cov(\boldsymbol{r}_2(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(2)}} + \sigma_\epsilon^2 \mathbf{I}_{n_2}]^{-1} \boldsymbol{y}^{(2)}, \boldsymbol{r}_1(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(1)},\mathbf{X}^{(1)}} + \sigma_\epsilon^2 \mathbf{I}_{n_1}]^{-1} \boldsymbol{y}^{(1)}) \\
&= \boldsymbol{r}_2(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(2)}} + \sigma_\epsilon^2 \mathbf{I}_{n_2}]^{-1}\, Var(\boldsymbol{y}^{(2)})\, [\mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(2)}} + \sigma_\epsilon^2 \mathbf{I}_{n_2}]^{-1} \boldsymbol{r}_2(\boldsymbol{x}') \\
&+ \boldsymbol{r}_1(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(1)},\mathbf{X}^{(1)}} + \sigma_\epsilon^2 \mathbf{I}_{n_1}]^{-1}\, Var(\boldsymbol{y}^{(1)})\, [\mathbf{K}_{\mathbf{X}^{(1)},\mathbf{X}^{(1)}} + \sigma_\epsilon^2 \mathbf{I}_{n_1}]^{-1} \boldsymbol{r}_1(\boldsymbol{x}') \\
&- 2\, \boldsymbol{r}_2(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(2)}} + \sigma_\epsilon^2 \mathbf{I}_{n_2}]^{-1}\, Cov(\boldsymbol{y}^{(2)}, \boldsymbol{y}^{(1)})\, [\mathbf{K}_{\mathbf{X}^{(1)},\mathbf{X}^{(1)}} + \sigma_\epsilon^2 \mathbf{I}_{n_1}]^{-1} \boldsymbol{r}_1(\boldsymbol{x}') \\
&= \boldsymbol{r}_2(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(2)}} + \sigma_\epsilon^2 \mathbf{I}_{n_2}]^{-1}\, \boldsymbol{r}_2(\boldsymbol{x}') + \boldsymbol{r}_1(\boldsymbol{x})^\top\, [\mathbf{K}_{\mathbf{X}^{(1)},\mathbf{X}^{(1)}} + \sigma_\epsilon^2 \mathbf{I}_{n_1}]^{-1} \boldsymbol{r}_1(\boldsymbol{x}') \\
&- 2\, \boldsymbol{r}_2(\boldsymbol{x})^\top [\mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(2)}} + \sigma_\epsilon^2 \mathbf{I}_{n_2}]^{-1}\, \mathbf{K}_{\mathbf{X}^{(2)},\mathbf{X}^{(1)}}\, [\mathbf{K}_{\mathbf{X}^{(1)},\mathbf{X}^{(1)}} + \sigma_\epsilon^2 \mathbf{I}_{n_1}]^{-1} \boldsymbol{r}_1(\boldsymbol{x}').
\end{aligned}$$

S.2 Karhunen-Loève expansion of a Gaussian process

Karhunen-Loève expansion provides a framework to decompose any stochastic process as an infinite linear combination of orthogonal basis functions. Since, we are interested in Gaussian processes, we will discuss the KL expansion only for GPs. Let us now consider that $f(\boldsymbol{x})$ is a zero mean Gaussian process with $k(\boldsymbol{x}, \boldsymbol{x}')$ as the covariance function. This process can decomposed as follows:

$$f(\boldsymbol{x}) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\phi_k(\boldsymbol{x})z_k, \tag{1}$$

where $z_k \mid k = 1, \ldots, \infty$ are the uncorrelated standard normal random variables, $\lambda_k \mid k = 1, \ldots, \infty$ are the eigenvalues, and $\phi_k(.) \mid k = 1, \ldots, \infty$ are the basis eigenfunctions. The values of λ_k and $\phi_k(.)$ can be obtained by solving the following integral eigenproblem

$$\int k(\boldsymbol{x}, \boldsymbol{x}')\phi(\boldsymbol{x}')d\boldsymbol{x}' = \lambda\phi(\boldsymbol{x}). \tag{2}$$

In practice, Equation (2) can be solved by discretizing the integral. Let us again assume that we have n data points from the process $f(\cdot)$. Then, we consider the following matrix eigenproblem

$$\mathbf{K}\boldsymbol{u}_k = \lambda_k^{mat}\boldsymbol{u}_k, \tag{3}$$

where $\mathbf{K}$ is again the covariance matrix with entries $\mathbf{K}_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j) \mid i, j = 1 \ldots n$;
λ_k^{mat} are the eigenvalues of the covariance matrix $\mathbf{K}$;
$\boldsymbol{u}_k$ are the normalized unit eigenvectors of the covariance matrix $\mathbf{K}$.
The eigenvalues and eigenfunctions of the integral problem are related to the eigenvalues and eigenvectors of the matrix problem in the following way:

$$\lambda_k \approx \frac{\lambda_k^{mat}}{n}, \tag{4}$$

$$\phi_k(\boldsymbol{x}_j) \approx \sqrt{n}(\boldsymbol{u}_k)_j, \tag{5}$$

where $(\boldsymbol{u}_k)_j$ is the j^{th} component of the eigenvector $\boldsymbol{u}_k$. The above approximation reduces the infinite sum in the KL expansion to a finite sum (truncated KL expansion) as follows:

$$\begin{aligned} f(x_j) &\approx \sum_{k=1}^{n} \sqrt{\frac{\lambda_k^{mat}}{n}}\sqrt{n}(\boldsymbol{u}_k)_j z_k, \\ &= \sum_{k=1}^{n} \sqrt{\lambda_k^{mat}}(\boldsymbol{u}_k)_j z_k. \end{aligned} \tag{6}$$

If λ_k's decay rapidly, this sum can be be truncated further by considering only m largest eigenvalues, where $m < n$. This decomposition can be written compactly in the matrix form. If we consider a vector, $\boldsymbol{f} = (f(\boldsymbol{x}_1), f(\boldsymbol{x}_2), \ldots, f(\boldsymbol{x}_n))^\top$, then it can be decomposed as follows:

$$\boldsymbol{f} = \mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{z}, \tag{7}$$

where $\mathbf{U}$ is the matrix with columns as eigenvectors of covariance matrix $\mathbf{K}$; $\boldsymbol{\Lambda}$ is a diagonal matrix with eigenvalues of $\mathbf{K}$ and $\boldsymbol{z}$ is a vector of length n with uncorrelated standard normal random variables as its components.

S.3 Details of the simulated functions

Piston simulation function

The piston simulation function, as the name suggests, is used to simulate the motion of a piston inside an engine. This function was proposed by Kenett and Zacks (1998). The response is the cycle time in seconds, i.e. the time required to complete one cycle, and is given by:

$$f(x) = 2\pi\sqrt{\frac{M}{k + S^2\frac{P_0V_0}{T_0}\frac{T_a}{V^2}}},$$

where

$$V = \frac{S}{2k}\left(\sqrt{A^2 + 4k\frac{P_0V_0}{T_0}T_a} - A\right),$$

$$A = P_0S + 19.62M - \frac{kV_0}{S},$$

where M is the weight of the piston (kg), k is the coefficient of the spring, S is the piston surface area (m^2) , P_0 is the atmospheric pressure (N/m^2), V_0 is the initial gas volume (m^3), T_0 is the filling gas temperature (K), and T_a is the ambient temperature (K). The number of input variables in this function are seven. We only choose two of them, V_0 and T_0, as input variables. The other variables are fixed at $M = 45$, $S = 0.01$, $k = 2,000$, $P_0 = 100,000$, $T_a = 292$. A perturbation on the function, $g(x)$, is obtained by changing the value of the the spring coefficient from $k = 2,000$ to $k = 2,500$. The range of the function is approximately between $[0.3, 0.7]$, so the value of the noise standard deviation is set at $\sigma_\epsilon = 0.05$. Figure 1 presents $f(x)$ and its perturbation, $g(x)$ along with the noisy datasets.

Borehole simulation function

The borehole function is used to model the flow of water through a borehole (Harper and Gupta, 1983) and has been widely used for computer experiments. See, for example, Morris et al. (1993). The response for this function is the water flow rate in the unit of $m^3/year$, given by:

$$f(x) = \frac{2\pi T_u(H_u - H_l)}{\ln(r/r_w)\left(1 + \frac{2LT_u}{\ln(r/r_w)r_w^2K_w} + \frac{T_u}{T_l}\right)},$$

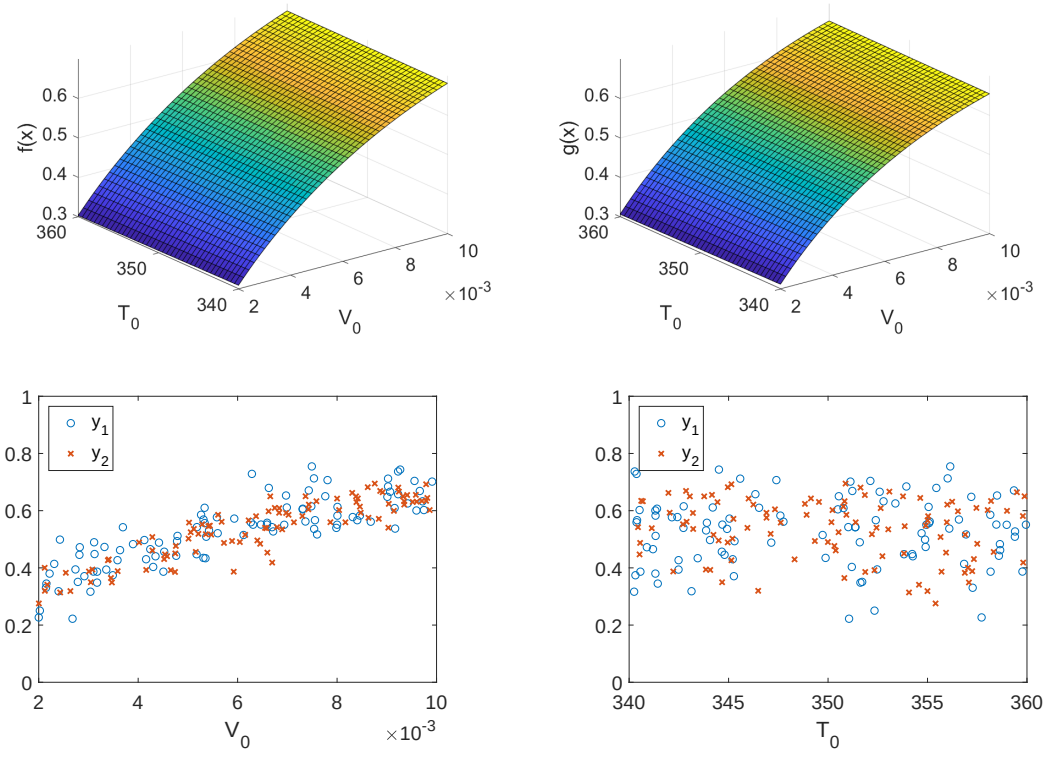

Figure 1: Plots for the piston function. Top left: $f(x)$; Top right: $g(x)$; Bottom left: noisy responses versus V_0; Bottom right: noisy responses versus T_0.

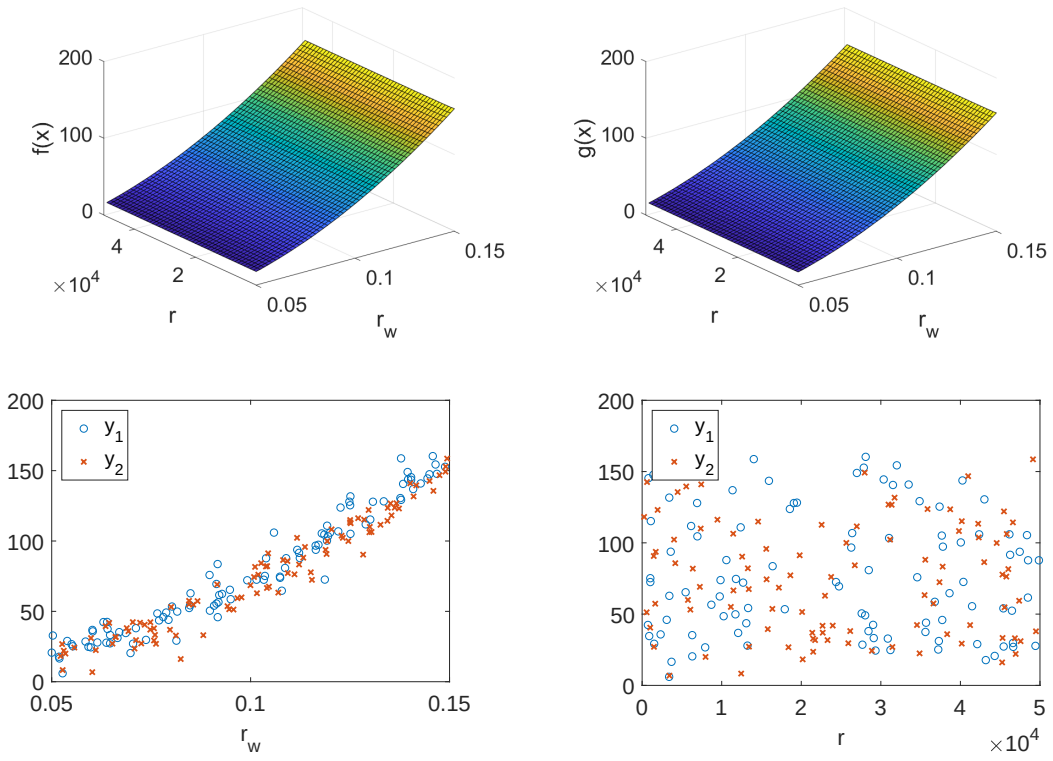

Figure 2: Plots for the borehole function. Top left: $f(x)$; Top right: $g(x)$; Bottom left: noisy responses versus r_w; Bottom right: noisy responses versus r.

where r_w is the radius of the borehole (m), r is the radius of the influence (m), L is the length of the borehole (m), T_u is the transmissivity of the upper aquifer $(m^2/year)$, T_l is the transmissivity of the lower aquifer $(m^2/year)$, H_u is the potentiometric head of the upper aquifer (m), H_l is the potentiometric head of the lower aquifer (m), and K_w is the hydraulic conductivity of the borehole $(m/year)$. The number of input variables for the borehole function is eight. Again, we only consider two input variables (r and r_w) while fixing other variables are fixed at $T_u = 78,000$, $H_u = 1,050$, $T_l = 84$, $H_l = 760$, $L = 1,400$, $K_w = 11,000$. In this simulation study, a perturbation, $g(x)$, is obtained by changing the value of L from 1400 to 1450. The range of this function is approximately between $[0, 150]$, so we set the value of the noise standard deviation at $\sigma_\epsilon = 10$. Figure 2 show the functions and the noisy data plots.

References

Harper, W. and Gupta, S. (1983). *Sensitivity/uncertainty analysis of a borehole scenario comparing Latin Hypercube Sampling and deterministic sensitivity approaches.* BMI/ONWI-516, Office of Nuclear Waste Isolation, Battelle Memorial Institute, Columbus, OH.

Kenett, R. S. and Zacks, S. (1998). *Modern Industrial Statistics: The Design and Control of Quality and Reliability.* Duxbury Press, Pacific Grove, CA.

Morris, M. D., Mitchell, T. J., and Ylvisaker, D. (1993). Bayesian design and analysis of computer experiments: Use of derivatives in surface prediction. *Technometrics*, 35(3):243–255.