Designing Combined Physical and Computer Experiments to Maximize Prediction Accuracy

Erin R. Leatherman, Angela M. Dean, Thomas J. Santner

Results from Corrected Software

The code used to construct the local IMSPE-optimal designs in the paper contained an error and was corrected on February 1, 2021. Corrected code is posted on the lead author's website (link). Many thanks to Dr. Selden Crary for alerting us to this error.

Using the corrected software, the design recommendations that are made in Section 5 of the manuscript remain. Additionally, the $D^P_{.5,-}$ designs appear to perform nearly as well as the $D^P_{.25,-}$ designs. Figures 1–4 show updated heatmaps of the *90th percentile* of the standardized EMSPE values that are colored replacements for Figures 6 – 9 in the manuscript. Figures 5–8 show updated heatmaps of the *average* of the standardized EMSPE values.

The conclusions from Section 6 of the manuscript remain unchanged. Figures 9 and 10 show updated boxplots of EMSPE values for various "true" values of the calibration parameter. These plots replace manuscript Figures 10 and 11, respectively.

Figure 1: $(n_p, d_x, n_s, d_x + d_t) = (10, 2, 15, 3)$: A heatmap of the 90^{th} percentile of the standardized EMSPE values for the 14 combined designs listed in Table 3 of the manuscript and the 18 test-bed surface types listed in Table 5 of the manuscript. Test bed surfaces 1–3, 4–6, 7–9, 10–12, 13–15 and 16–18 use $S^{Krig}_{.25}$, $S^{Krig}_{.5}$, S^{SL}_{E}, S^{SL}_{U}, S^{MXB}_{edge}, and S^{MXB}_{mid}, respectively, as $y^s(\boldsymbol{x}, t)$. Within each grouping of three simulator surfaces, the $\zeta_{\text{test}}(\boldsymbol{x})$ to be estimated is the sum of $y^s_{\text{test}}(\boldsymbol{x}, 0.25)$ and $\delta_{\text{test}}(\boldsymbol{x})$ which is: $\equiv 0$, $B^{Krig}_{.5}$, $B^{Krig}_{.75}$, in order. Blue boxes indicate that the design has a 90^{th} percentile EMSPE value lower than the mean EMSPE value for the group of designs for that particular surface.

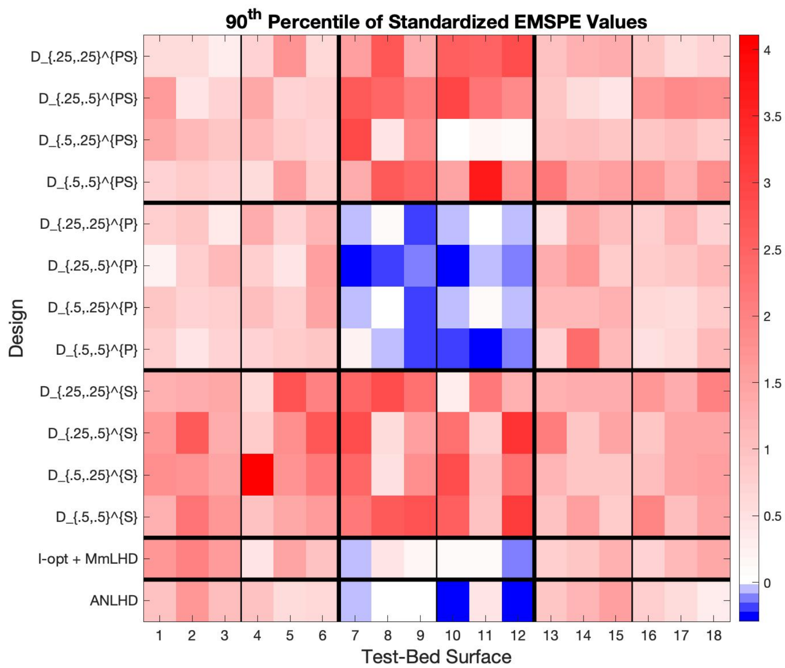

Figure 2: $(n_p, d_x, n_s, d_x + d_t) = (20, 4, 25, 5)$: A heatmap of the 90^{th} percentile of the standardized EMSPE values for the 14 combined designs listed in Table 3 of the manuscript and the 18 test-bed surface types listed in Table 5 of the manuscript. Test bed surfaces 1–3, 4–6, 7–9, 10–12, 13–15 and 16–18 use $S^{Krig}_{.25}$, $S^{Krig}_{.5}$, S^{SL}_{E}, S^{SL}_{U}, S^{MXB}_{edge}, and S^{MXB}_{mid}, respectively, as $y^s(\boldsymbol{x}, t)$. Within each grouping of three simulator surfaces, the $\zeta_{\text{test}}(\boldsymbol{x})$ to be estimated is the sum of $y^s_{\text{test}}(\boldsymbol{x}, 0.25)$ and $\delta_{\text{test}}(\boldsymbol{x})$ which is: $\equiv 0$, $B^{Krig}_{.5}$, $B^{Krig}_{.75}$, in order. Blue boxes indicate that the design has a 90^{th} percentile EMSPE value lower than the mean EMSPE value for the group of designs for that particular surface.

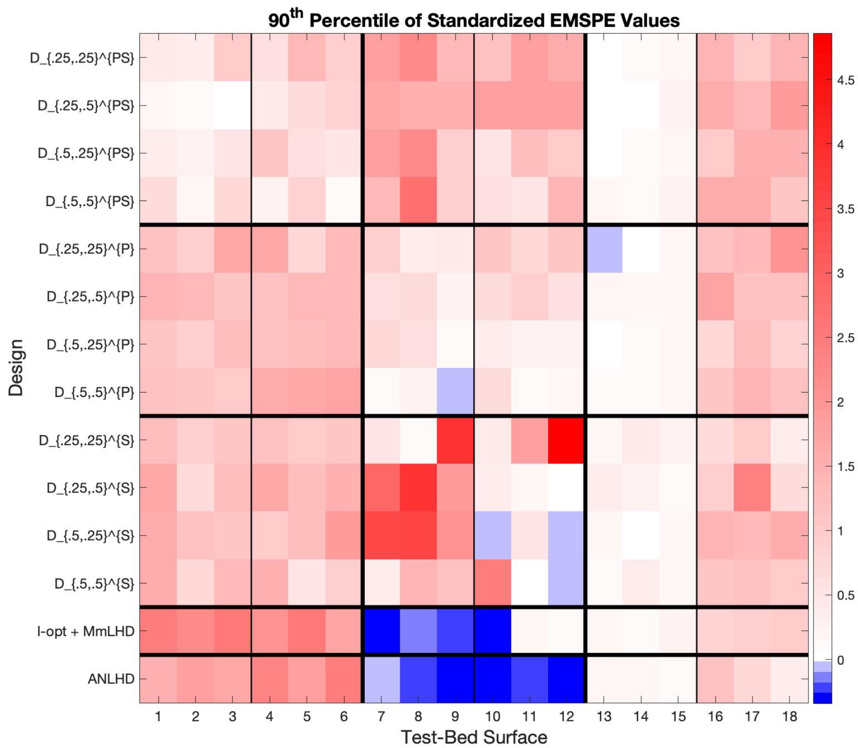

Figure 3: $(n_p, d_x, n_s, d_x + d_t) = (30, 3, 50, 5)$: A heatmap of the 90^{th} percentile of the standardized EMSPE values for the 14 combined designs listed in Table 3 of the manuscript and the 18 test-bed surface types listed in Table 5 of the manuscript. Test bed surfaces 1–3, 4–6, 7–9, 10–12, 13–15 and 16–18 use $S^{Krig}_{.25}$, $S^{Krig}_{.5}$, S^{SL}_{E}, S^{SL}_{U}, S^{MXB}_{edge}, and S^{MXB}_{mid}, respectively, as $y^s(\boldsymbol{x}, \boldsymbol{t})$. Within each grouping of three simulator surfaces, the $\zeta_{\text{test}}(\boldsymbol{x})$ to be estimated is the sum of $y^s_{\text{test}}(\boldsymbol{x}, 0.25 \times \mathbf{1}_2)$ and $\delta_{\text{test}}(\boldsymbol{x})$ which is: $\equiv 0$, $B^{Krig}_{.5}$, $B^{Krig}_{.75}$, in order. Blue boxes indicate that the design has a 90^{th} percentile EMSPE value lower than the mean EMSPE value for the group of designs for that particular surface.

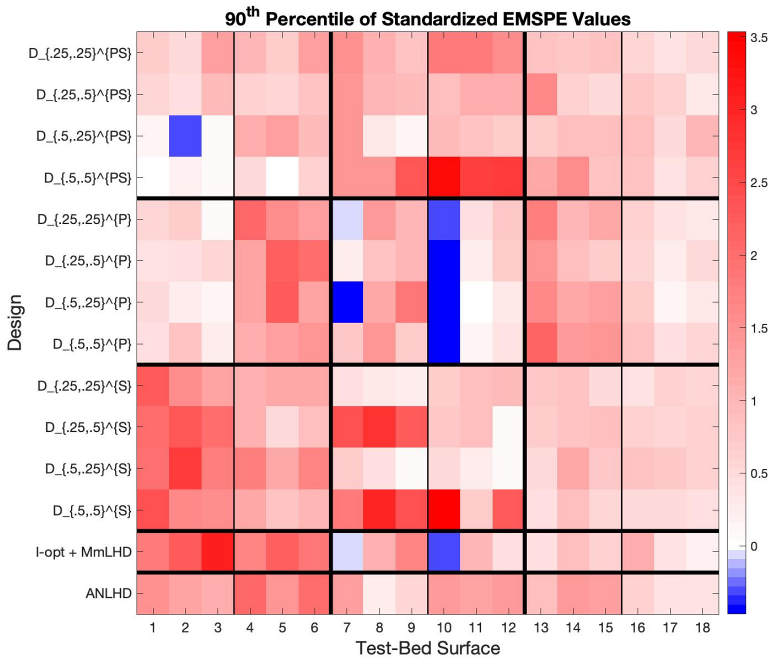

Figure 4: $(n_p, d_x, n_s, d_x + d_t) = (40, 4, 50, 5)$: A heatmap of the 90^{th} percentile of the standardized EMSPE values for the 14 combined designs listed in Table 3 of the manuscript and the 18 test-bed surface types listed in Table 5 of the manuscript. Test bed surfaces 1–3, 4–6, 7–9, 10–12, 13–15 and 16–18 use $S^{Krig}_{.25}$, $S^{Krig}_{.5}$, S^{SL}_{E}, S^{SL}_{U}, S^{MXB}_{edge}, and S^{MXB}_{mid}, respectively, as $y^s(\boldsymbol{x}, t)$. Within each grouping of three simulator surfaces, the $\zeta_{\text{test}}(\boldsymbol{x})$ to be estimated is the sum of $y^s_{\text{test}}(\boldsymbol{x}, 0.25)$ and $\delta_{\text{test}}(\boldsymbol{x})$ which is: $\equiv 0$, $B^{Krig}_{.5}$, $B^{Krig}_{.75}$, in order. Blue boxes indicate that the design has a 90^{th} percentile EMSPE value lower than the mean EMSPE value for the group of designs for that particular surface.

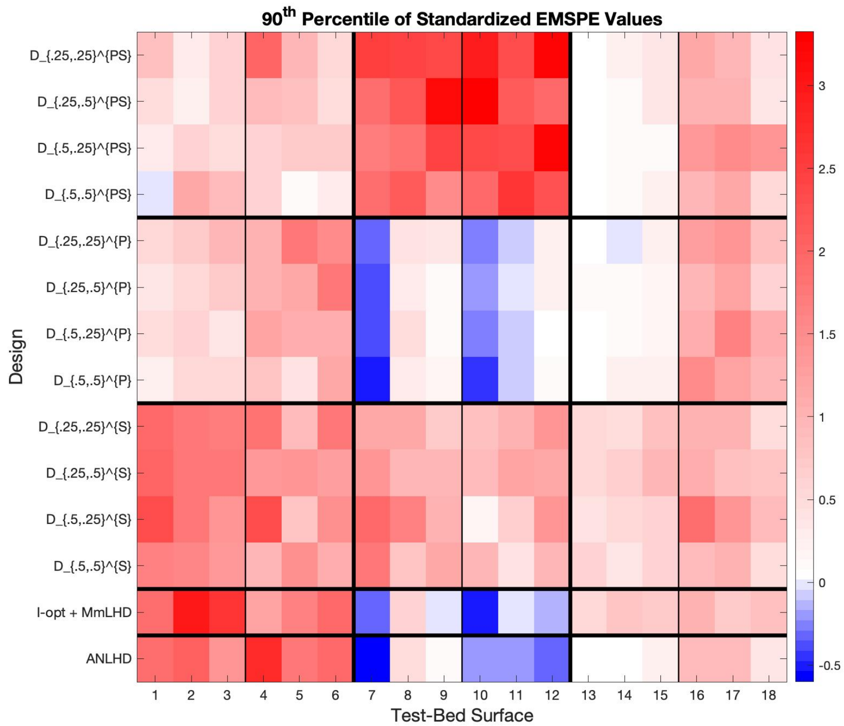

Figure 5: $(n_p, d_x, n_s, d_x + d_t) = (10, 2, 15, 3)$: A heatmap of the mean of the standardized EMSPE values for the 14 combined designs listed in Table 3 of the manuscript and the 18 test-bed surface types listed in Table 5 of the manuscript. Test bed surfaces 1–3, 4–6, 7–9, 10–12, 13–15 and 16–18 use $S^{Krig}_{.25}$, $S^{Krig}_{.5}$, S^{SL}_{E}, S^{SL}_{U}, S^{MXB}_{edge}, and S^{MXB}_{mid}, respectively, as $y^s(\boldsymbol{x}, t)$. Within each grouping of three simulator surfaces, the $\zeta_{\text{test}}(\boldsymbol{x})$ to be estimated is the sum of $y^s_{\text{test}}(\boldsymbol{x}, 0.25)$ and $\delta_{\text{test}}(\boldsymbol{x})$ which is: $\equiv 0$, $B^{Krig}_{.5}$, $B^{Krig}_{.75}$, in order. Blue boxes indicate that the design has a mean EMSPE value lower than the mean EMSPE value for the group of designs for that particular surface.

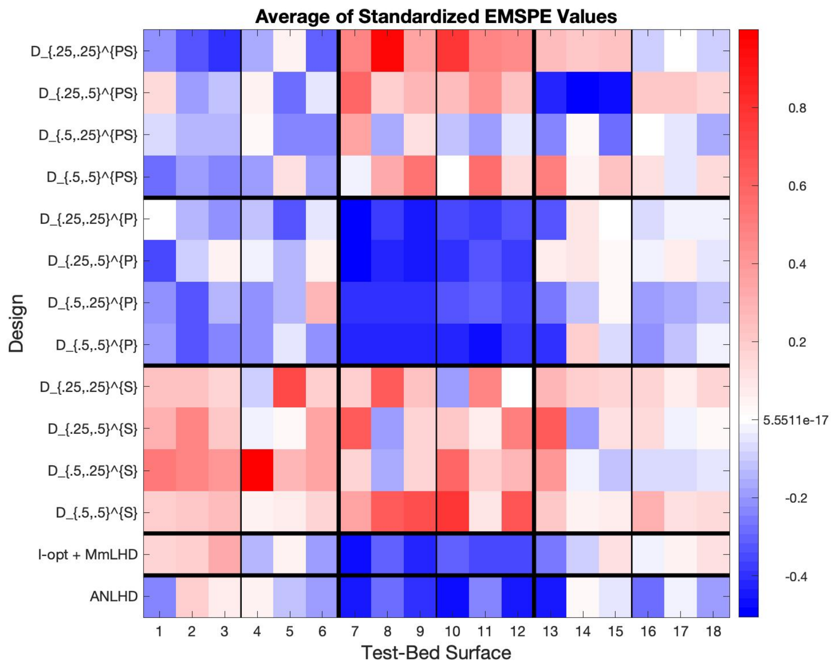

Figure 6: $(n_p, d_x, n_s, d_x + d_t) = (20, 4, 25, 5)$: A heatmap of the mean of the standardized EMSPE values for the 14 combined designs listed in Table 3 of the manuscript and the 18 test-bed surface types listed in Table 5 of the manuscript. Test bed surfaces 1–3, 4–6, 7–9, 10–12, 13–15 and 16–18 use $S^{Krig}_{.25}$, $S^{Krig}_{.5}$, S^{SL}_{E}, S^{SL}_{U}, S^{MXB}_{edge}, and S^{MXB}_{mid}, respectively, as $y^s(\boldsymbol{x}, t)$. Within each grouping of three simulator surfaces, the $\zeta_{\text{test}}(\boldsymbol{x})$ to be estimated is the sum of $y^s_{\text{test}}(\boldsymbol{x}, 0.25)$ and $\delta_{\text{test}}(\boldsymbol{x})$ which is: $\equiv 0$, $B^{Krig}_{.5}$, $B^{Krig}_{.75}$, in order. Blue boxes indicate that the design has a mean EMSPE value lower than the mean EMSPE value for the group of designs for that particular surface.

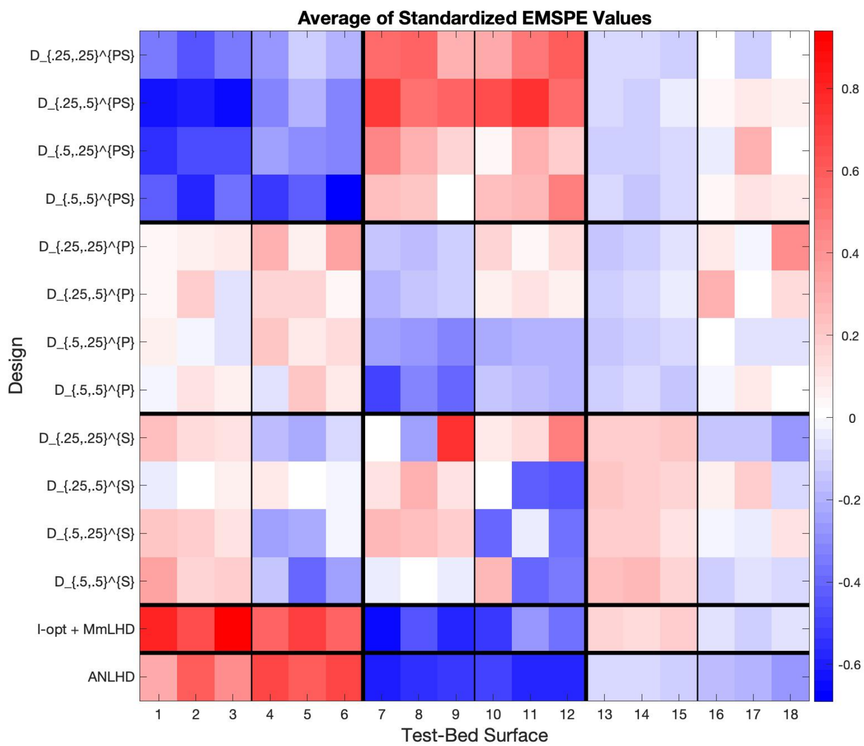

Figure 7: $(n_p, d_x, n_s, d_x + d_t) = (30, 3, 50, 5)$: A heatmap of the mean of the standardized EMSPE values for the 14 combined designs listed in Table 3 of the manuscript and the 18 test-bed surface types listed in Table 5 of the manuscript. Test bed surfaces 1–3, 4–6, 7–9, 10–12, 13–15 and 16–18 use $S^{Krig}_{.25}$, $S^{Krig}_{.5}$, S^{SL}_{E}, S^{SL}_{U}, S^{MXB}_{edge}, and S^{MXB}_{mid}, respectively, as $y^s(\boldsymbol{x}, \boldsymbol{t})$. Within each grouping of three simulator surfaces, the $\zeta_{\text{test}}(\boldsymbol{x})$ to be estimated is the sum of $y^s_{\text{test}}(\boldsymbol{x}, 0.25 \times \mathbf{1}_2)$ and $\delta_{\text{test}}(\boldsymbol{x})$ which is: $\equiv 0$, $B^{Krig}_{.5}$, $B^{Krig}_{.75}$, in order. Blue boxes indicate that the design has a mean EMSPE value lower than the mean EMSPE value for the group of designs for that particular surface.

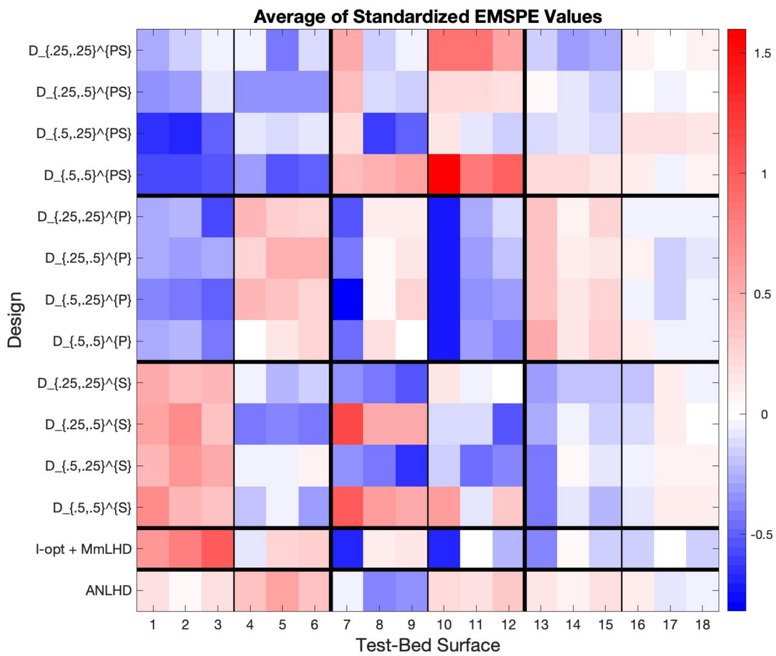

Figure 8: $(n_p, d_x, n_s, d_x + d_t) = (40, 4, 50, 5)$: A heatmap of the mean of the standardized EMSPE values for the 14 combined designs listed in Table 3 of the manuscript and the 18 test-bed surface types listed in Table 5 of the manuscript. Test bed surfaces 1–3, 4–6, 7–9, 10–12, 13–15 and 16–18 use $S^{Krig}_{.25}$, $S^{Krig}_{.5}$, S^{SL}_{E}, S^{SL}_{U}, S^{MXB}_{edge}, and S^{MXB}_{mid}, respectively, as $y^s(\boldsymbol{x}, t)$. Within each grouping of three simulator surfaces, the $\zeta_{\text{test}}(\boldsymbol{x})$ to be estimated is the sum of $y^s_{\text{test}}(\boldsymbol{x}, 0.25)$ and $\delta_{\text{test}}(\boldsymbol{x})$ which is: $\equiv 0$, $B^{Krig}_{.5}$, $B^{Krig}_{.75}$, in order. Blue boxes indicate that the design has a mean EMSPE value lower than the mean EMSPE value for the group of designs for that particular surface.

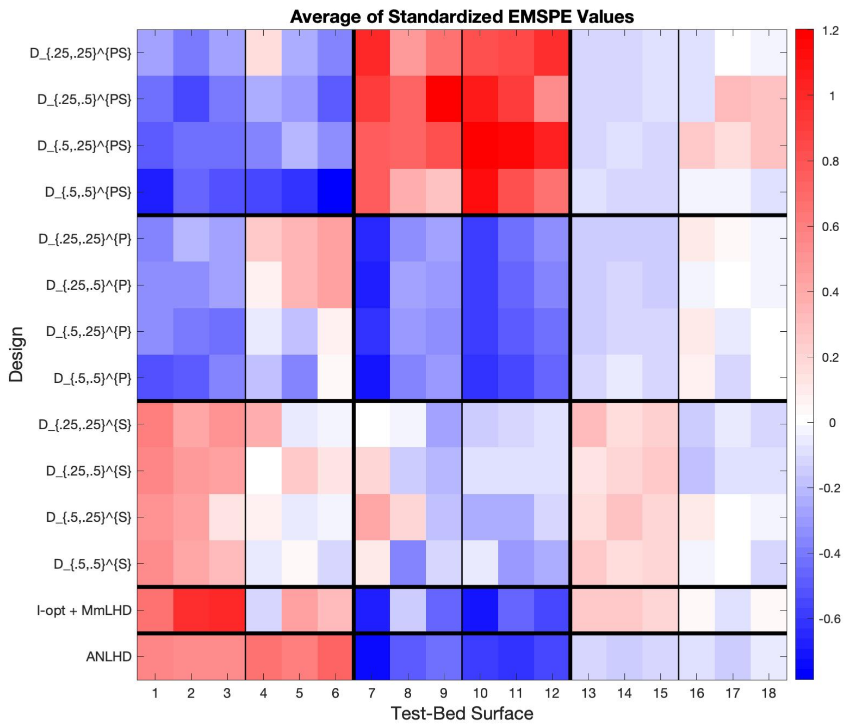

Figure 9: $(n_p, d_x, n_s, d_x + d_t) = (10, 2, 15, 3)$: Boxplots of (non-standardized) EMSPE values when predicting 30 realizations of the surface $\zeta_{\text{test}}(\boldsymbol{x}) = y^s_{\text{test}}(\boldsymbol{x}, \theta) + \delta_{\text{test}}(\boldsymbol{x})$ for $\theta \in \{0.125, 0.25, \ldots, 0.875\}$ when $y^s_{\text{test}}(\boldsymbol{x}, t)$ is $S^{Krig}_{.25}$ and $\delta_{\text{test}}(\boldsymbol{x})$ is $B^{Krig}_{.5}$. Panels (from left to right) correspond to designs $D^{PS}_{.25,.5}$, $D^{P}_{.25,.5}$, and $D^{S}_{.25,.5}$, which were constructed under the assumption $\theta = 0.5$.

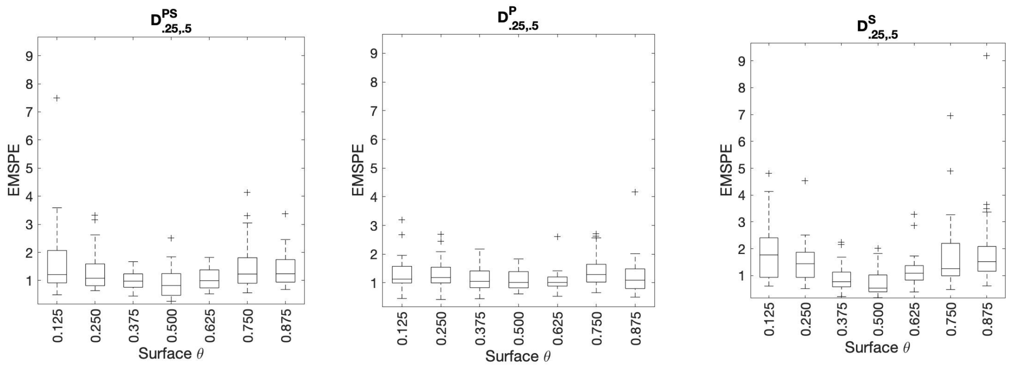

Figure 10: $(n_p, d_x, n_s, d_x + d_t) = (40, 4, 50, 5)$: Boxplots of (non-standardized) EMSPE values when predicting 30 realizations of the surface $\zeta_{\text{test}}(\boldsymbol{x}) = y^s_{\text{test}}(\boldsymbol{x}, \theta) + \delta_{\text{test}}(\boldsymbol{x})$ for $\theta \in \{0.125, 0.25, \ldots, 0.875\}$ when $y^s_{\text{test}}(\boldsymbol{x}, t)$ is $S^{Krig}_{.25}$ and $\delta_{\text{test}}(\boldsymbol{x})$ is $B^{Krig}_{.5}$. Panels (from left to right) correspond to designs $D^{PS}_{.25,.5}$, $D^{P}_{.25,.5}$, and $D^{S}_{.25,.5}$, which were constructed under the assumption $\theta = 0.5$.