Smooth Scalar-on-Image Regression via Spatial Bayesian Variable Selection: Supplement C

By: Jeff Goldsmith, Lei Huang, Ciprian Crainiceanu

C Additional Simulation Results

Here we consider additional scenarios for the two-dimensional simulations described in Section 3 in the main manuscript. We examine two additional cases: first, the coefficient function contains a large, rectangular predictive region on the boundary of the image in which all regression coefficients are equal; second, we examine the impact of registration error on the estimated coefficient image.

C.1 Large, Homogeneous Predictive Region

We construct predictors as in Section 3, using a principal components decomposition of observed 50×50 axial slices to generate simulated images $\boldsymbol{X}_i^S = \sum_{k=1}^{50} c_{ik}\phi_k$. Here $\boldsymbol{\phi} = \{\phi_1, \ldots, \phi_{50}\}$ are eigen-images with accompanying eigenvalues $\boldsymbol{\lambda} = \{\lambda_1, ..., \lambda_{50}\}$, and PC loadings $\boldsymbol{c}_i$ are generated $\boldsymbol{c}_i \sim \text{N}\,[0, \text{diag}(\boldsymbol{\lambda})]$. The coefficient function $\boldsymbol{\beta}$ contains a single rectangular (30×15) predictive region in which the regression coefficients β_l are uniformly equal to 1. As in Section 3, we choose three signal-to-noise ratios $\frac{\sigma_{\boldsymbol{y}}^2}{\sigma_\epsilon^2} = 3, 1, 1/3$.

	$I = 100$			$I = 500$		
$\sigma_{\boldsymbol{y}}^2/\sigma_\epsilon^2 =$	3	1	1/3	3	1	1/3
GMRF - 2D						
MSE_1	0.172	0.273	0.472	0.079	0.117	0.185
MSE_0	0.006	0.019	0.044	0.005	0.007	0.012
Computation	104.4	101.7	104.8	169.5	167.8	165.6
EX. - 2D						
MSE_1	0.462	0.651	0.778	0.370	0.455	0.550
MSE_0	0.026	0.026	0.030	0.003	0.007	0.020
Computation	86.2	92.8	88.9	141.9	143.9	141.5
FPCR - 2D						
MSE_1	0.175	0.256	0.363	0.103	0.133	0.192
MSE_0	0.034	0.040	0.046	0.026	0.029	0.034
Computation	0.4	0.4	0.4	1.8	1.8	1.8

Table C.1: Average mean squared error separated by true predictive and non-predictive location, signal-to-noise ratio, sample size, predictor dimension, and estimation technique ("GMRF" labels the Gaussian MRF prior, "EX" labels the exchangeable prior, and "FPCR" the functional approach). Average computation time (in seconds) to fit the model is also shown.

For each signal-to-noise ratio we generate 500 datasets for both $I = 100$ and $I = 500$ and apply the proposed method, the related scalar-on-image model with an exchangeable prior on the regression coefficients, and the FPCR approach. Tuning parameters are chosen using the first simulated dataset, and the results below may not be fully representative of the performance of the three methods.

Table C.1 shows the MSE taken over all simulated datasets for each combination of sample size and signal-to-noise ratio. The proposed method provides good estimates of the coefficient image in both predictive and non-predictive regions, and performance improves as sample size increases and as signal-to-noise ratio increases. The proposed method substantially outperforms the related approach that uses an exchangeable prior distribution: given the homogeneity of the predictive region, encouraging spatial smoothness improves estimation. The proposed method often outperforms the FPCR approach as well, with the exception of $I = 100$ and $\frac{\sigma_y^2}{\sigma_\epsilon^2} = 1/3$. While these approaches are much more comparable in terms of MSE here than in the bump-and-slab setting considered in Section 3, there remain significant qualitative differences in estimated coefficient images. As Figure C.1 shows, estimated coefficient images for the proposed method contain many locations that are shrunk entirely to zero due to the sparsity constraint. On the other hand, the FPCR method does not impose sparsity, but induces sparsity both through penalization and a basis representation that avoids voxel-by-voxel estimation.

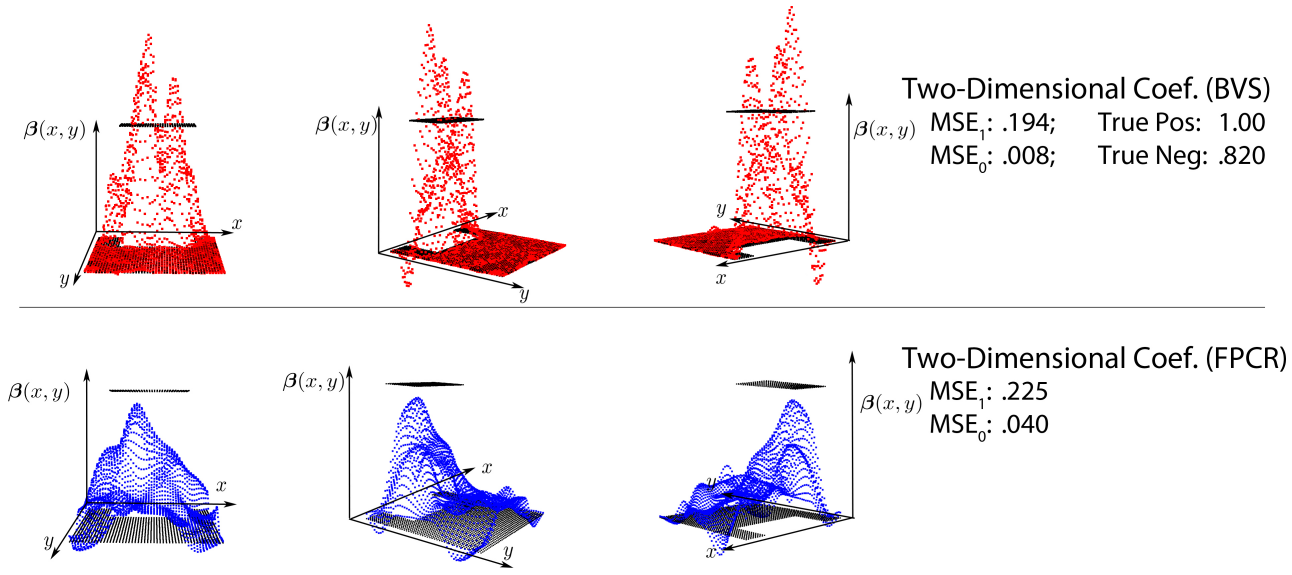

Figure C.1: Plot of typical estimated coefficient images from the simulation with a large, homogeneous predictive region. Estimates from the proposed method (labelled "BVS") and the FPCR method are shown, along with the corresponding MSE_1 and MSE_0. The true coefficient is shown in black.

C.2 Impact of Registration Errors

In this simulation we construct simulated predictors to illustrate the effect of registration error on the proposed procedure. To do so, we modify the previous method for constructing predictors in the following way. From the observed data we extract 56×56 axial slices and perform a principal components decomposition of these images. Simulated images are constructed using $\boldsymbol{X}_i^S = \sum_{k=1}^{50} c_{ik}\phi_k$, where $\boldsymbol{\phi} = \{\phi_1, \ldots, \phi_{50}\}$ are eigen-images with accompanying eigenvalues $\boldsymbol{\lambda} = \{\lambda_1, ..., \lambda_{50}\}$, and PC loadings $\boldsymbol{c}_i$ are generated $\boldsymbol{c}_i \sim \mathrm{N}\left[0, \mathrm{diag}(\boldsymbol{\lambda})\right]$. The coefficient function $\boldsymbol{\beta}$ is defined as in Section 3 using bivariate normal density functions and is applied to the middle 50×50 component of each 56×56 simulated predictor to generate outcomes. As in Section 3, we choose three signal-to-noise ratios $\frac{\sigma_{\boldsymbol{y}}^2}{\sigma_\epsilon^2} = 3, 1, 1/3$.

	$I = 100$			$I = 500$		
$\sigma_{\boldsymbol{y}}^2/\sigma_\epsilon^2 =$	3	1	1/3	3	1	1/3
GMRF – 2D						
MSE_1	2.511	3.053	3.781	1.576	1.712	2.214
MSE_0	0.065	0.082	0.218	0.136	0.110	0.143
Computation	112.8	111.7	111.9	163.5	164.9	164.4
EX. – 2D						
MSE_1	2.731	3.126	3.744	1.977	2.182	2.584
MSE_0	0.233	0.194	0.385	0.110	0.120	0.227
Computation	96.8	95.0	95.3	148.9	147.8	147.2
FPCR – 2D						
MSE_1	2.928	3.753	4.635	1.902	2.151	2.784
MSE_0	0.091	0.058	0.027	0.133	0.123	0.099
Computation	0.4	0.4	0.4	2.1	2.1	2.1

Table C.2: Average mean squared error separated by true predictive and non-predictive location, signal-to-noise ratio, sample size, predictor dimension, and estimation technique ("GMRF" labels the Gaussian MRF prior, "EX" labels the exchangeable prior, and "FPCR" the functional approach). Average computation time (in seconds) to fit the model is also shown.

To simulate registration errors, the "observed" predictors in each simulation are 50×50 images that are off-center components of the complete 56×56 generated images; the predictors are randomly shifted horizontally and vertically by $\{-3, -2, -1, 0, 1, 2, 3\}$ voxels with probability $\{.1, .1, .2, .2, .2, .1, .1\}$. Thus, each observed predictor is shifted from the true generating image by a random amount. We also note that due to the size of predictive regions in the coefficient image, these registration errors are substantial.

Table C.2 shows the MSE taken over all simulated datasets for each combination of sample size and signal-to-noise ratio. A comparison with Table 1 comparison, which provides results for the same simula-

tion design without registration error, unsurprisingly indicates that all methods have higher MSEs in the presence of registration error. Moreover, the results indicate that the proposed method outperforms the competing approaches in terms of MSE for predictive regions. The FPCR has comparable and sometimes lower MSE on non-predictive regions; as Figure C.2 shows, this is due to additional regions declared predictive by the proposed method as well as general oversmoothing by the FPCR method. As expected, performance improves for all methods as signal strength rises and as sample size increases.

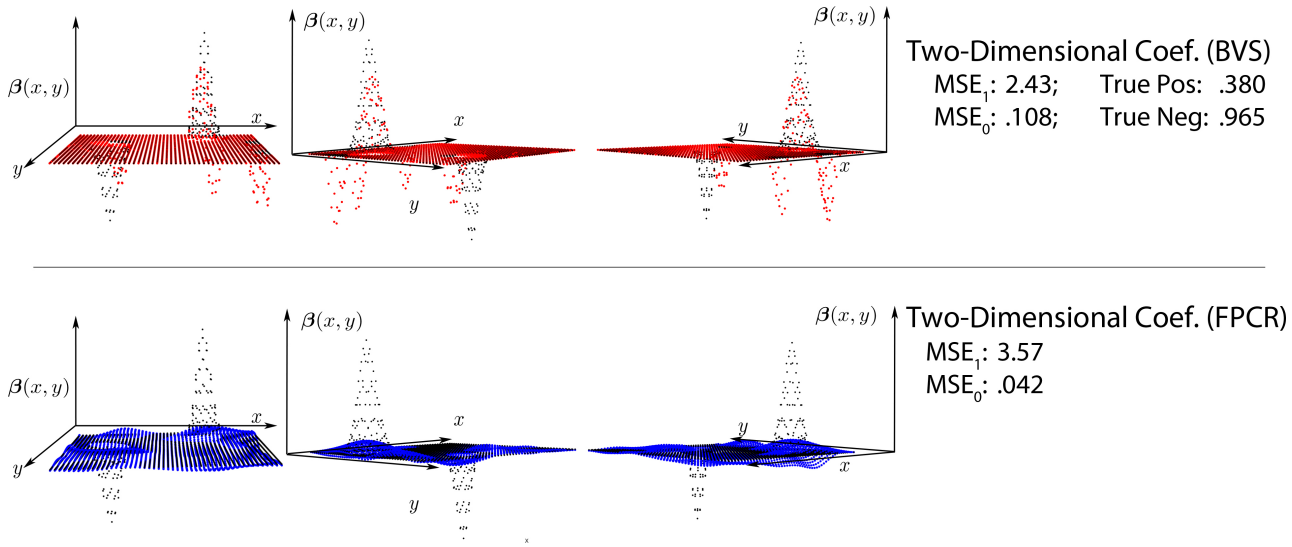

Figure C.2: Plot of typical estimated coefficient images from the simulation with registration errors. Estimates from the proposed method (labelled "BVS") and the FPCR method are shown, along with the corresponding MSE_1 and MSE_0. The true coefficient is shown in black.