Supplementary Materials: Out-of-sample prediction in multidimensional P-spline models

Alba Carballo [1,†], María Durbán [1,†] and Dae-Jin Lee [2,†,*]

Proof of Theorem 1

Differenciating Eq. (12) in the main text with respect to $\boldsymbol{\theta}_+$ leads to

$$\frac{\partial S}{\partial \boldsymbol{\theta}_+} = -2\boldsymbol{B}_+'\tilde{\boldsymbol{R}}_+^{-1}(\boldsymbol{y}_+ - \boldsymbol{B}_+\boldsymbol{\theta}_+) + 2(\lambda_z \boldsymbol{P}_+^z + \lambda_x \boldsymbol{P}_+^{x_+}) = 0$$

i.e., the penalized least squares solution is given by:

$$\hat{\boldsymbol{\theta}}_+ = (\boldsymbol{B}_+'\tilde{\boldsymbol{R}}_+^{-1}\boldsymbol{B}_+ + \lambda_z \boldsymbol{P}_+^z + \lambda_x \boldsymbol{P}_+^{x_+})^{-1}\boldsymbol{B}_+'\tilde{\boldsymbol{R}}_+^{-1}\boldsymbol{y}_+. \tag{1}$$

Let us define $\boldsymbol{C} = (\boldsymbol{B}_+'\tilde{\boldsymbol{R}}_+^{-1}\boldsymbol{B}_+ + \lambda_z \boldsymbol{P}_+^z + \lambda_x \boldsymbol{P}_+^{x_+})$ and $\boldsymbol{C}^{-1} = \begin{bmatrix} \boldsymbol{C}^{11} & \boldsymbol{C}^{12} \\ \boldsymbol{C}^{21} & \boldsymbol{C}^{22} \end{bmatrix}$, with this notation and since $\tilde{\boldsymbol{R}}_+^{-1} = \tilde{\boldsymbol{R}}_{x_+}^{-1} \otimes \tilde{\boldsymbol{R}}_z^{-1} = \text{blockdiag}(\boldsymbol{I}, \boldsymbol{O})$, with $\boldsymbol{I}$ an identity matrix of dimension $n_x n_z \times n_x n_z$ and $\boldsymbol{O}$ a null matrix of dimension $n_{x_p} n_z \times n_{x_p} n_z$, equation (1) can be rewritten as

$$\boldsymbol{\theta}_+ = \boldsymbol{C}^{-1}\boldsymbol{B}_+'\tilde{\boldsymbol{R}}_+^{-1}\boldsymbol{y}_+ = \begin{bmatrix} \boldsymbol{C}^{11}\boldsymbol{B}'\boldsymbol{y} \\ \boldsymbol{C}^{21}\boldsymbol{B}'\boldsymbol{y} \end{bmatrix}. \tag{2}$$

If $\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix}$, by Theorem 8.5.11 given in [1] we have that:

$$\boldsymbol{C}^{-1} = \begin{bmatrix} \boldsymbol{K}^{-1} & -\boldsymbol{K}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22}^{-1} \\ -\boldsymbol{C}_{22}^{-1}\boldsymbol{C}_{21}\boldsymbol{K}^{-1} & \boldsymbol{C}_{22}^{-1} + \boldsymbol{C}_{22}^{-1}\boldsymbol{C}_{21}\boldsymbol{K}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22}^{-1} \end{bmatrix},$$

with $\boldsymbol{K} = \boldsymbol{C}_{11} - \boldsymbol{C}_{12}\boldsymbol{C}_{22}^{-1}\boldsymbol{C}_{21}$. Therefore:

$$\boldsymbol{C}^{11} = \boldsymbol{K}^{-1} = \left(\boldsymbol{B}'\boldsymbol{B} + \lambda_x \boldsymbol{P}_{+11}^{x_+} + \lambda_z \boldsymbol{P}_{+11}^z - \lambda_x^2 \boldsymbol{P}_{+12}^{x_+}(\lambda_x \boldsymbol{P}_{+22}^{x_+} + \lambda_z \boldsymbol{P}_{+22}^z)^{-1}\boldsymbol{P}_{+21}^{x_+}\right)^{-1}$$

and

$$\begin{aligned} \boldsymbol{C}^{21} &= -\boldsymbol{C}_{22}^{-1}\boldsymbol{C}_{21}\boldsymbol{K}^{-1} \\ &= -(\lambda_x \boldsymbol{P}_{+22}^{x_+} + \lambda_z \boldsymbol{P}_{+22}^z)^{-1}\lambda_x \boldsymbol{P}_{+21}^{x_+}\boldsymbol{C}^{11} \end{aligned}$$

and by equation (2) the coefficients for the fit and for the prediction are given by equations (18) and (19) in the main text, respectively, as we wanted to show.

Proof Colloraly 1

Notice that if $\boldsymbol{P}_{+22}^z = \boldsymbol{O}$, by Eq. (18), the coefficients that give the fit are:

$$\hat{\boldsymbol{\theta}}_{+1,\dots,c} = \left(\boldsymbol{B}'\boldsymbol{B} + \lambda_x(\boldsymbol{D}_x'\boldsymbol{D}_x \otimes \boldsymbol{I}_{c_z}) + \lambda_z(\boldsymbol{I}_{c_x} \otimes \boldsymbol{D}_z'\boldsymbol{D}_z)\right)^{-1}\boldsymbol{B}'\boldsymbol{y},$$

i.e., the same as the coefficients we obtain only fitting the data without a prediction. Let us see which are the coefficients that determine the forecast when the penalty orders are two or three.

- Differences of order 2.

Suppose a difference matrix with second order penalty $\boldsymbol{D}_{\boldsymbol{x}_+}$ of dimensions $(c_{\boldsymbol{x}_+} - 2) \times c_{\boldsymbol{x}_+}$:

$$
\boldsymbol{D}_{\boldsymbol{x}_+} = \begin{bmatrix} \boldsymbol{D} & \boldsymbol{O} \\ \boldsymbol{D}_{\boldsymbol{x}(1)} & \boldsymbol{D}_{\boldsymbol{x}(2)} \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix},
$$

where $\boldsymbol{D}_{\boldsymbol{x}(2)}$:

$$
\boldsymbol{D}_{\boldsymbol{x}(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ -2 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix},
$$

of dimension $c_{\boldsymbol{x}_p} \times c_{\boldsymbol{x}_p}$, where $c_{\boldsymbol{x}_p}$ is the number of columns of $\boldsymbol{B}_{\boldsymbol{x}(2)}$. Therefore, $(\boldsymbol{D}_{\boldsymbol{x}(2)} \otimes \boldsymbol{I}_{c_z})^{-1}$ has the form

$$
\begin{bmatrix}
1 & & & & & & & & & & & & \\
 & \ddots & & & & & & & & & & & \\
 & & 1 & & & & & & & & & & \\
2 & & & 1 & & & & & & & & & \\
 & \ddots & & & \ddots & & & & & & & & \\
 & & 2 & & & 1 & & & & & & & \\
3 & & & 2 & & & 1 & & & & & & \\
 & \ddots & & & \ddots & & & \ddots & & & & & \\
 & & 3 & & & 2 & & & 1 & & & & \\
4 & & & 3 & & & 2 & & & 1 & & & \\
 & \ddots & & & \ddots & & & \ddots & & & \ddots & & \\
 & & 4 & & & 3 & & & 2 & & & 1 & \\
 & \vdots & & & \vdots & & & \vdots & & & & \vdots & \ddots
\end{bmatrix}
$$

of dimension $(c_{\boldsymbol{x}_p} \cdot c_{\boldsymbol{z}}) \times (c_{\boldsymbol{x}_p} \cdot c_{\boldsymbol{z}})$, each block has dimension $c_{\boldsymbol{z}} \times c_{\boldsymbol{z}}$. Moreover,

$$
\boldsymbol{D}_{\boldsymbol{x}(1)} \otimes \boldsymbol{I}_{c_z} = \begin{bmatrix}
0 & & & 0 & & & & 1 & & & -2 & & \\
 & \ddots & & & \ddots & & \dots & & \ddots & & & \ddots & \\
 & & 0 & & & 0 & & & & 1 & & & -2 \\
0 & & & 0 & & & & 0 & & & 1 & & \\
 & \ddots & & & \ddots & & \dots & & \ddots & & & \ddots & \\
 & & 0 & & & 0 & & & & 0 & & & 1 \\
0 & & & 0 & & & & 0 & & & 0 & & \\
 & \ddots & & & \ddots & & \dots & & \ddots & & & \ddots & \\
 & & 0 & & & 0 & & & & 0 & & & 0 \\
 & \vdots & & & \vdots & & \dots & & \vdots & & & \vdots &
\end{bmatrix},
$$

i.e., is a matrix of dimension $(c_{x_p} \cdot c_z) \times (c_x \cdot c_z)$ with just three blocks of dimension $c_z \times c_z$ that are not blocks of zeros. Therefore,

$$(\boldsymbol{D}_{\boldsymbol{x}(2)} \otimes \boldsymbol{I}_{c_z})^{-1}(\boldsymbol{D}_{\boldsymbol{x}(1)} \otimes \boldsymbol{I}_{c_z}) = \begin{bmatrix} 0 & & 0 & & & 1 & & -2 & \\ & \ddots & & \ddots & \cdots & & \ddots & & \ddots \\ & 0 & & 0 & & & 1 & & -2 \\ 0 & & 0 & & & 2 & & -3 & \\ & \ddots & & \ddots & \cdots & & \ddots & & \ddots \\ & 0 & & 0 & & & 2 & & -3 \\ 0 & & 0 & & & 3 & & -4 & \\ & \ddots & & \ddots & \cdots & & \ddots & & \ddots \\ & 0 & & 0 & & & 3 & & -4 \\ 0 & & 0 & & & 4 & & -5 & \\ & \ddots & & \ddots & \cdots & & \ddots & & \ddots \\ & 0 & & 0 & & & 4 & & -5 \\ & \vdots & & \vdots & \cdots & \vdots & & \vdots & \end{bmatrix},$$

with dimension $(c_{x_p} \cdot c_z) \times (c_x \cdot c_z)$. Hence, considering the matrix of coefficients that give the fit, $\hat{\boldsymbol{\Theta}}$, and the matrix of coefficients that give the forecast, $\hat{\boldsymbol{\Theta}}_p$, each row $j = 1, ..., c_z$, of the additional matrix of coefficients is a linear combination of two old coefficients of that row:

$$\hat{\boldsymbol{\Theta}}_{j\ \cdot} = \hat{\theta}_{j\ c_x} \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} + (\hat{\theta}_{j\ c_x} - \hat{\theta}_{j\ c_x-1}) \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \end{bmatrix}.$$

- Differences of order 3.
  Suppose a difference matrix with third order penalty, $\boldsymbol{D}_{\boldsymbol{x}_+}$ of dimensions $(c_{\boldsymbol{x}_+} - 3) \times c_{\boldsymbol{x}_+}$:

$$\boldsymbol{D}_+ = \begin{bmatrix} -1 & 3 & -3 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & -1 & 3 & -3 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 3 & -3 & 1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 3 & -3 & 1 \end{bmatrix}.$$

In this case, $\boldsymbol{D}_{\boldsymbol{x}(2)}$ is:

$$\boldsymbol{D}_{\boldsymbol{x}(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 3 & -3 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 3 & -3 & 1 \end{bmatrix}.$$

$(\boldsymbol{D}_{\boldsymbol{x}(2)} \otimes \boldsymbol{I}_{c_z})^{-1}$ and $\boldsymbol{D}_{\boldsymbol{x}(1)} \otimes \boldsymbol{I}_{c_z}$ are:

$$(\boldsymbol{D}_{\boldsymbol{x}(2)} \otimes \boldsymbol{I}_{c_z})^{-1} = \begin{bmatrix}
1 & & & & & & & & & & & & \\
& \ddots & & & & & & & & & & & \\
& & 1 & & & & & & & & & & \\
3 & & & 1 & & & & & & & & & \\
& \ddots & & & \ddots & & & & & & & & \\
& & 3 & & & 1 & & & & & & & \\
6 & & & 3 & & & 1 & & & & & & \\
& \ddots & & & \ddots & & & \ddots & & & & & \\
& & 6 & & & 3 & & & 1 & & & & \\
10 & & & 6 & & & 3 & & & 1 & & & \\
& \ddots & & & \ddots & & & \ddots & & & \ddots & & \\
& & 10 & & & 6 & & & 3 & & & 1 & \\
& \vdots & & & \vdots & & & \vdots & & & & \vdots & \ddots
\end{bmatrix},$$

$$\boldsymbol{D}_{\boldsymbol{x}(1)} \otimes \boldsymbol{I}_{c_z} = \begin{bmatrix}
0 & & & & -1 & & & 3 & & & -3 & & \\
& \ddots & & \cdots & & \ddots & & & \ddots & & & \ddots & \\
& & 0 & & & & -1 & & & 3 & & & -3 \\
0 & & & & 0 & & & -1 & & & 3 & & \\
& \ddots & & \cdots & & \ddots & & & \ddots & & & \ddots & \\
& & 0 & & & & 0 & & & -1 & & & 3 \\
0 & & & & 0 & & & 0 & & & -1 & & \\
& \ddots & & \cdots & & \ddots & & & \ddots & & & \ddots & \\
& & 0 & & & & 0 & & & 0 & & & -1 \\
& \vdots & & \cdots & & \vdots & & & \vdots & & & \vdots &
\end{bmatrix}.$$

Then,

$$(\boldsymbol{D}_{\boldsymbol{x}(2)} \otimes \boldsymbol{I}_{c_z})^{-1}(\boldsymbol{D}_{\boldsymbol{x}(1)} \otimes \boldsymbol{I}_{c_z}) = \begin{bmatrix}
0 & & & & -1 & & & 3 & & & -3 & & \\
& \ddots & & \cdots & & \ddots & & & \ddots & & & \ddots & \\
& & 0 & & & & -1 & & & 3 & & & -3 \\
0 & & & & -3 & & & 8 & & & -6 & & \\
& \ddots & & \cdots & & \ddots & & & \ddots & & & \ddots & \\
& & 0 & & & & -3 & & & 8 & & & -6 \\
0 & & & & -6 & & & 15 & & & -10 & & \\
& \ddots & & \cdots & & \ddots & & & \ddots & & & \ddots & \\
& & 0 & & & & -6 & & & 15 & & & -10 \\
0 & & & & -10 & & & 24 & & & -15 & & \\
& \ddots & & \cdots & & \ddots & & & \ddots & & & \ddots & \\
& & 0 & & & & -10 & & & 24 & & & -15 \\
& \vdots & & \cdots & & \vdots & & & \vdots & & & \vdots &
\end{bmatrix}.$$

Therefore, each row, $j = 1, ..., c_z$, of the additional matrix of coefficients is a linear combination of three old coefficients of that row:

$$\hat{\boldsymbol{\Theta}}_{p_{j\,\cdot}} = \hat{\theta}_{j\,c_x} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} + \frac{3\hat{\theta}_{j\,c_x} - 4\hat{\theta}_{j\,c_x-1} + \hat{\theta}_{j\,c_x-2}}{2} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ \vdots \end{bmatrix} + \frac{\hat{\theta}_{j\,c_x} - 2\hat{\theta}_{j\,c_x-1} + \hat{\theta}_{j\,c_x-2}}{2} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ \vdots \end{bmatrix}^2 .$$

Proof of Theorem 2

Given the extended transformation matrix for the random part $\boldsymbol{\Omega}_{+r}$ and the extended penalty matrix $\boldsymbol{P}_+$ defined in the main text in Eq. (45) and (43), respectively, $\boldsymbol{F}_+^*$ is:

$$\boldsymbol{F}_+^* = \boldsymbol{\Omega}_{r+}^{*\prime} \boldsymbol{P}_+ \boldsymbol{\Omega}_{r+}^* = \begin{bmatrix} 0 & \boldsymbol{U}_{z_+r}^{*\prime} & \cdots & & & \\ \vdots & & \boldsymbol{U}_{x_+r}^{*\prime} & & & \\ & & & \boldsymbol{u}_{x_+f}^{*(2)\prime} \otimes \boldsymbol{U}_{z_+r}^{*\prime} & & \\ & & & \boldsymbol{U}_{x_+r}^{*\prime} \otimes \boldsymbol{u}_{z_+f}^{*(2)\prime} & & \\ & & & \boldsymbol{U}_{x_+r}^{*\prime} \otimes \boldsymbol{U}_{z_+r}^{*\prime} & & \end{bmatrix} \begin{bmatrix} 0 & \cdots & & \\ \vdots & \lambda_z \boldsymbol{D}_{z_+}' \boldsymbol{D}_{z_+} & & \\ & & \lambda_x \boldsymbol{D}_{x_+}' \boldsymbol{D}_{x_+} & \\ & & & \lambda_3 \boldsymbol{D}_{x_+}' \boldsymbol{D}_{x_+} \otimes \boldsymbol{I}_{c_{z_+}} + \lambda_4 \boldsymbol{I}_{c_{x_+}} \otimes \boldsymbol{D}_{z_+}' \boldsymbol{D}_{z_+} \end{bmatrix}$$

$$\begin{bmatrix} 0 & \cdots & & \\ \boldsymbol{U}_{zr} & & & \\ \vdots & \boldsymbol{U}_{x_+r}^* & & \\ & & \boldsymbol{u}_{x_+f}^{*(2)} \otimes \boldsymbol{U}_{z_+r}^* \mid \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{u}_{z_+f}^{*(2)} \mid \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{U}_{z_+r}^* \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_z \boldsymbol{U}_{z_+r}^{*\prime} \boldsymbol{D}_{z_+}' \boldsymbol{D}_{z_+} \boldsymbol{U}_{z_+r}^* & & \\ & \lambda_x \boldsymbol{U}_{x_+r}^{*\prime} \boldsymbol{D}_{x_+}' \boldsymbol{D}_{x_+} \boldsymbol{U}_{x_+r}^* & \\ & & \boldsymbol{F}_+^{(1,2)} \end{bmatrix},$$

where $\boldsymbol{F}_+^{(1,2)} = \begin{bmatrix} \boldsymbol{F}_{+11}^{(1,2)} & \boldsymbol{O} & \boldsymbol{F}_{+13}^{(1,2)} \\ \boldsymbol{O} & \boldsymbol{F}_{+22}^{(1,2)} & \boldsymbol{O} \\ \boldsymbol{F}_{+13}^{(1,2)\prime} & \boldsymbol{O} & \boldsymbol{F}_{+33}^{(1,2)} \end{bmatrix}$, with

$$\begin{aligned}
\boldsymbol{F}_{+11}^{(1,2)} &= \tau_x \boldsymbol{u}_{x_+f}^{*(2)\prime} \boldsymbol{D}_{x_+}' \boldsymbol{D}_{x_+} \boldsymbol{u}_{x_+f}^{*(2)} \otimes \boldsymbol{U}_{zr}^{*\prime} \boldsymbol{U}_{z_+r}^* + \tau_z \boldsymbol{u}_{x_+f}^{*(2)\prime} \boldsymbol{u}_{x_+f}^{*(2)} \otimes \boldsymbol{U}_{z_+r}^{*\prime} \boldsymbol{D}_{z_+}' \boldsymbol{D}_{z_+} \boldsymbol{U}_{z_+r}^* \\
\boldsymbol{F}_{+13}^{(1,2)} &= \tau_z \boldsymbol{u}_{x_+f}^{*(2)\prime} \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{U}_{z_+r}^* \boldsymbol{D}_{z_+}' \boldsymbol{D}_{z_+} \boldsymbol{U}_{z_+r}^* + \tau_x \boldsymbol{u}_{x_+f}^{*(2)\prime} \boldsymbol{D}_{x_+}' \boldsymbol{D}_{x_+} \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{U}_{z_+r}^{*\prime} \boldsymbol{U}_{z_+r}^* \\
\boldsymbol{F}_{+22}^{(1,2)} &= \tau_x \boldsymbol{U}_{x_+r}^{*\prime} \boldsymbol{D}_{x_+}' \boldsymbol{D}_{x_+} \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{u}_{z_+f}^{(*2)\prime} \boldsymbol{u}_{z_+f}^{*(2)} + \tau_z \boldsymbol{U}_{x_+r}^{*\prime} \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{u}_{z_+f}^{*(2)\prime} \boldsymbol{D}_{z_+}' \boldsymbol{D}_{z_+} \boldsymbol{u}_{z_+f}^{*(2)} \\
\boldsymbol{F}_{+23}^{(1,2)} &= \tau_x \boldsymbol{U}_{x_+r}^{*\prime} \boldsymbol{D}_{x_+}' \boldsymbol{D}_{x_+} \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{u}_{z_+f}^{*(2)\prime} \boldsymbol{U}_{z_+r}^* + \tau_z \boldsymbol{U}_{x_+r}^{*\prime} \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{u}_{z_+f}^{*(2)\prime} \boldsymbol{D}_{z_+}' \boldsymbol{D}_{z_+} \boldsymbol{U}_{z_+r}^* \\
\boldsymbol{F}_{+33}^{(1,2)} &= \tau_z \boldsymbol{U}_{x_+r}^{*\prime} \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{U}_{z_+r}^{*\prime} \boldsymbol{D}_{z_+}' \boldsymbol{D}_{z_+} \boldsymbol{U}_{z_+r}^* + \tau_x \boldsymbol{U}_{x_+r}^{*\prime} \boldsymbol{D}_{x_+}' \boldsymbol{D}_{x_+} \boldsymbol{U}_{x_+r}^* \otimes \boldsymbol{U}_{z_+r}^{*\prime} \boldsymbol{U}_{z_+r}^*
\end{aligned}$$

using $\boldsymbol{u}_{if}^{*(2)\prime} \boldsymbol{D}_i' = \boldsymbol{O}$ for $i = \boldsymbol{z}_+, \boldsymbol{x}_+$, we obtain the extended mixed model penalty $\boldsymbol{F}_+^*$ in Eq. (49) in the main text.

Simulation results for Scenario 1

- $n_{z_p} = 0, n_{x_p} = 10$

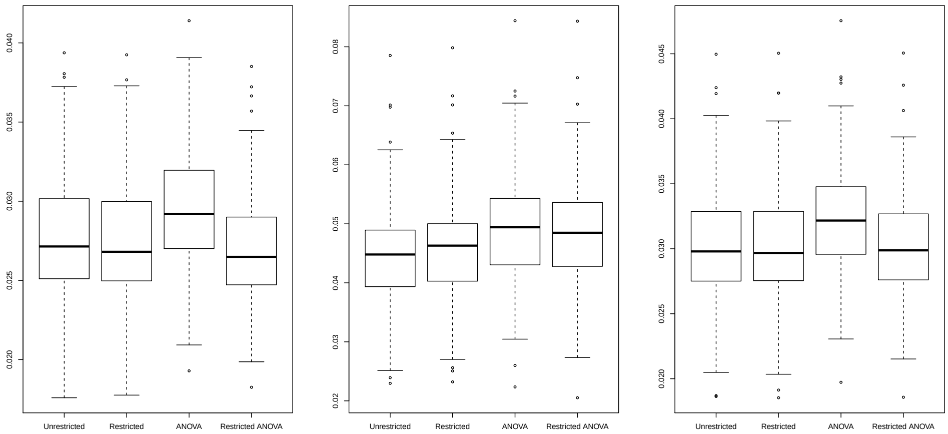

Figure S1. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 0$ and $n_{x_p} = 10$.

- $n_{z_p} = 0, n_{x_p} = 15$

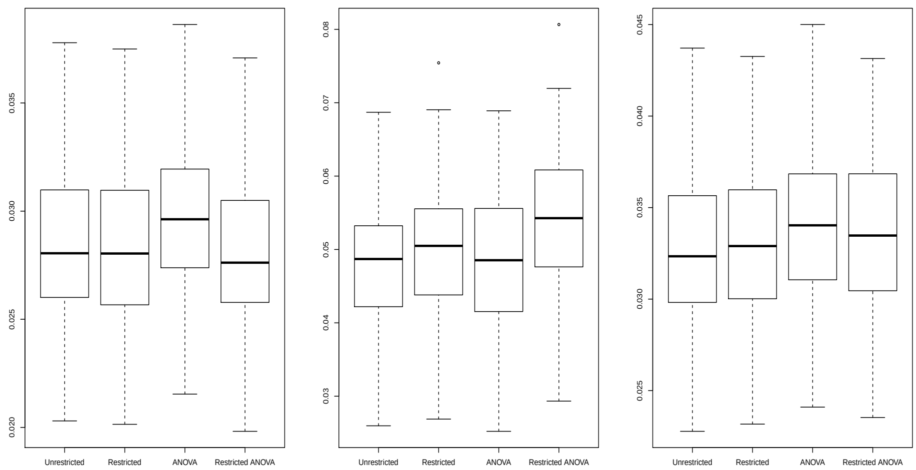

Figure S2. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 0$ and $n_{x_p} = 15$.

- $n_{z_p} = 0, n_{x_p} = 20$

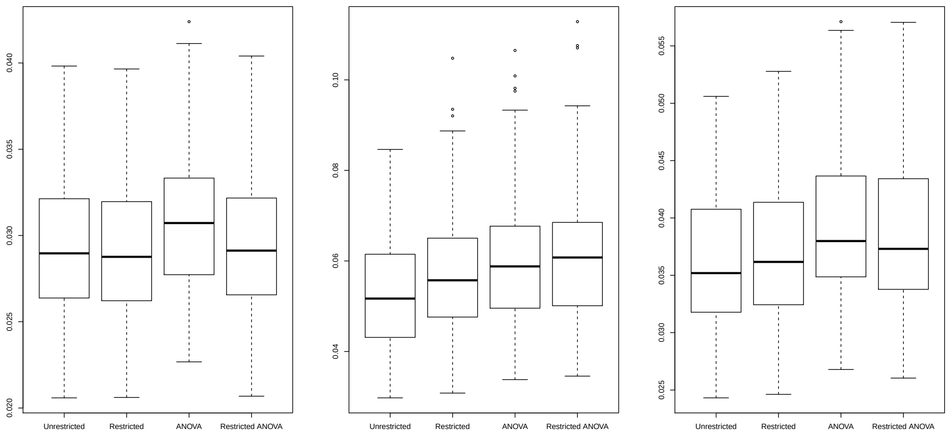

Figure S3. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 0$ and $n_{x_p} = 20$.

- $n_{z_p} = 10, n_{x_p} = 5$

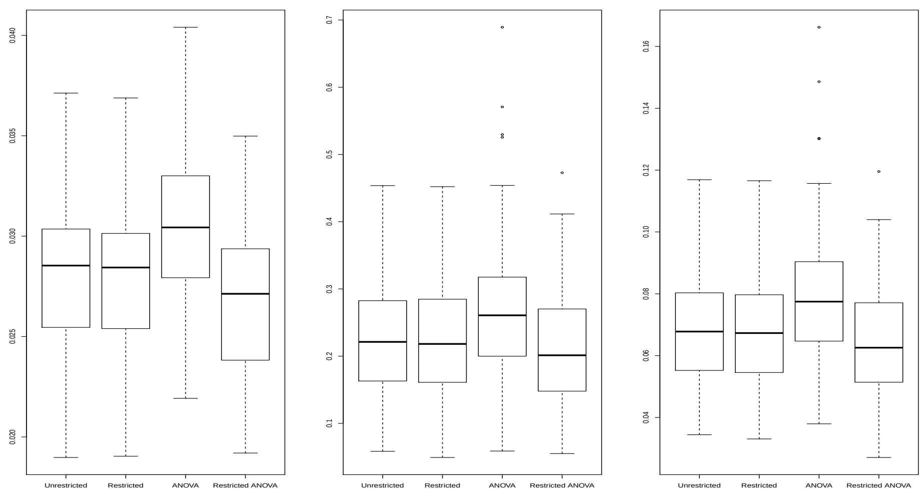

Figure S4. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 10, n_{x_p} = 5$.

- $n_{z_p} = n_{x_p} = 10$

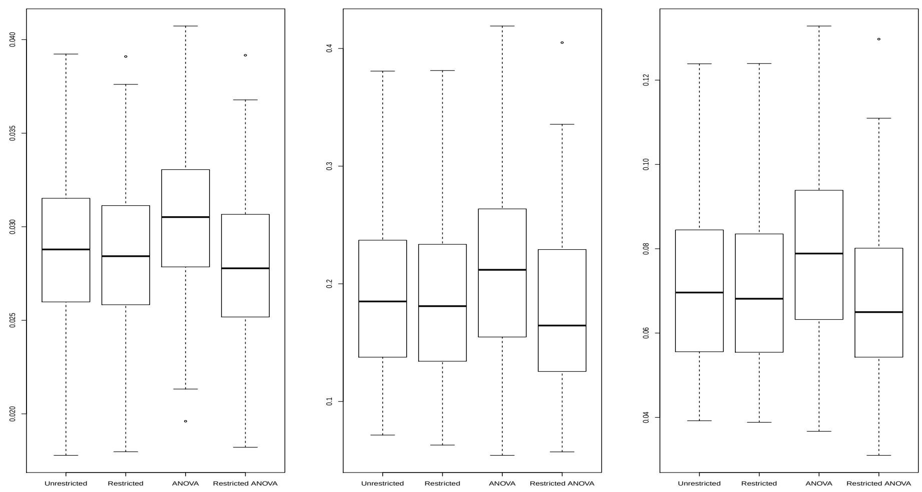

Figure S5. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 10$, $n_{x_p} = 10$.

- $n_{z_p} = 10, n_{x_p} = 15$

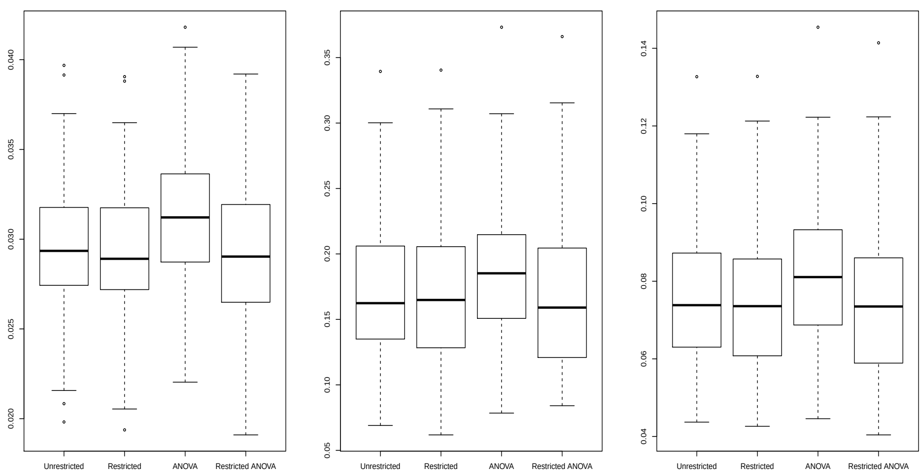

Figure S6. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 10$, $n_{x_p} = 15$.

- $n_{z_p} = 10, n_{x_p} = 20$

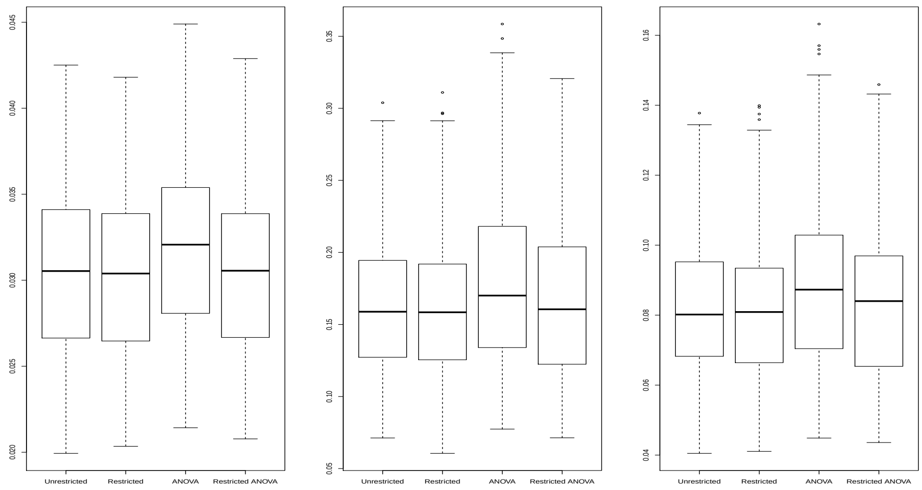

Figure S7. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 10$, $n_{x_p} = 30$.

- $n_{z_p} = 20$, $n_{x_p} = 5$

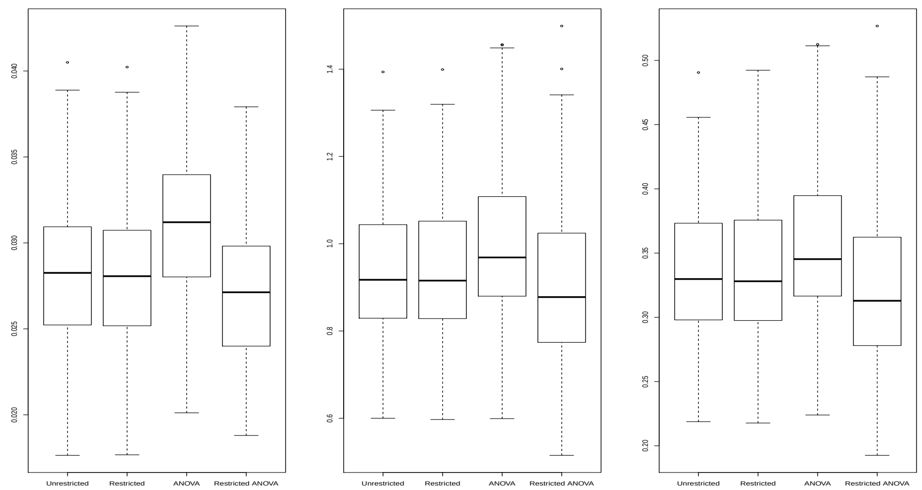

Figure S8. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 20$, $n_{x_p} = 5$.

- $n_{z_p} = 20$, $n_{x_p} = 10$

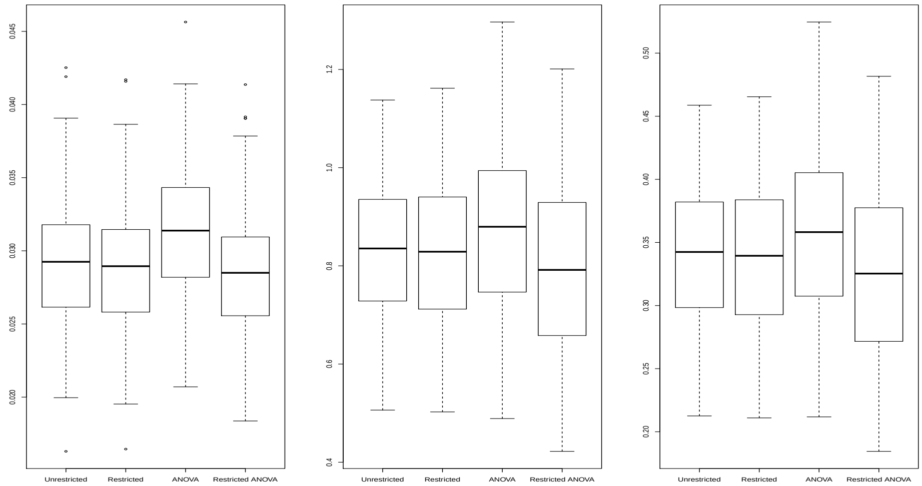

Figure S9. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 20$, $n_{x_p} = 10$.

- $n_{z_p} = 20$, $n_{x_p} = 15$

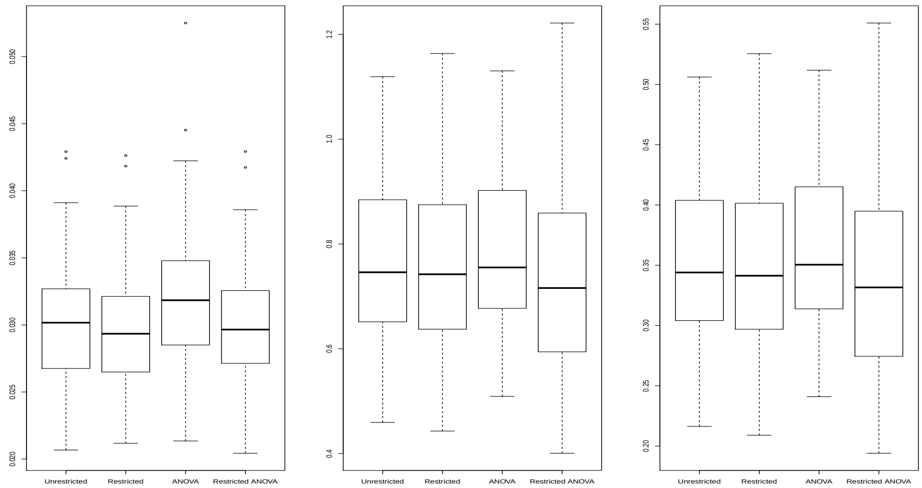

Figure S10. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 20$, $n_{x_p} = 15$.

- $n_{z_p} = 20$, $n_{x_p} = 20$

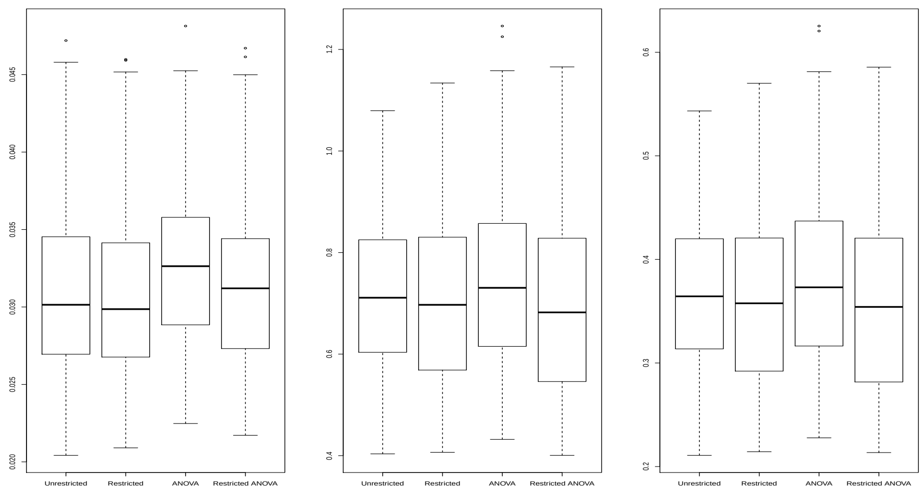

Figure S11. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 1 and $n_{z_p} = 20$, $n_{x_p} = 20$.

Simulation results for Scenario 2

- $n_{z_p} = 0$, $n_{x_p} = 10$

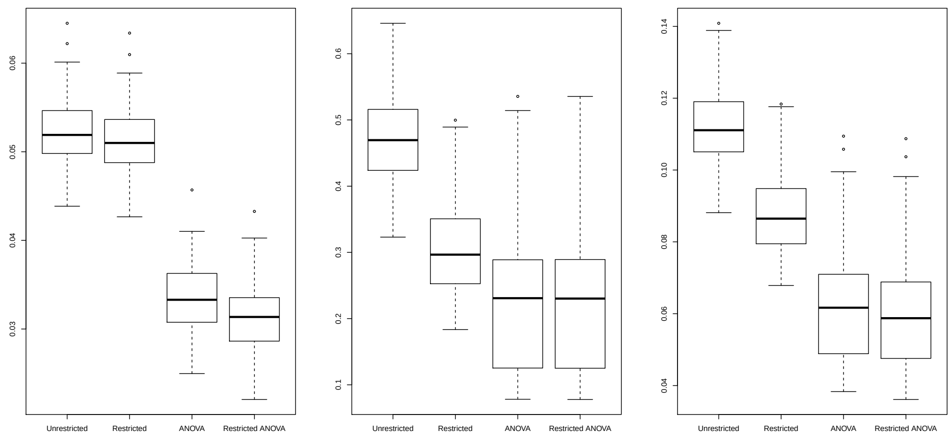

Figure S12. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = 0$ and $n_{x_p} = 10$.

- $n_{z_p} = 0$, $n_{x_p} = 15$

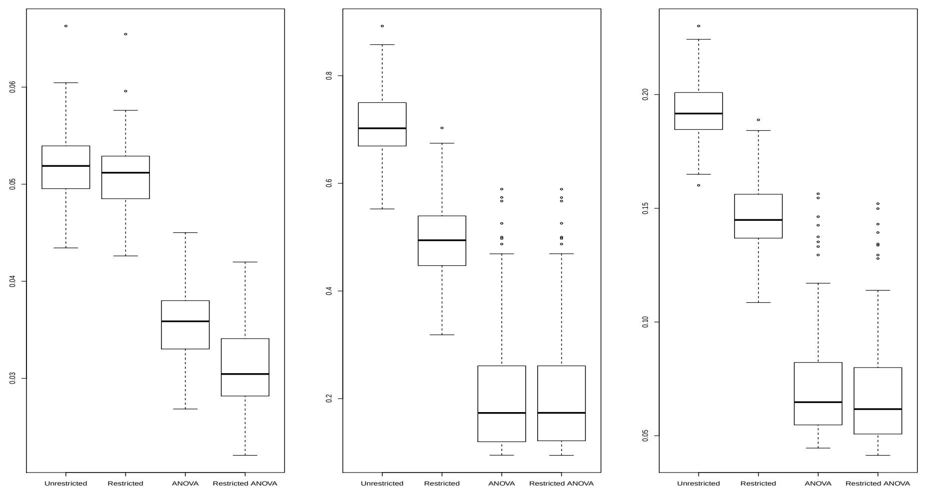

Figure S13. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = 0$ and $n_{x_p} = 15$.

- $n_{z_p} = 0$, $n_{x_p} = 20$

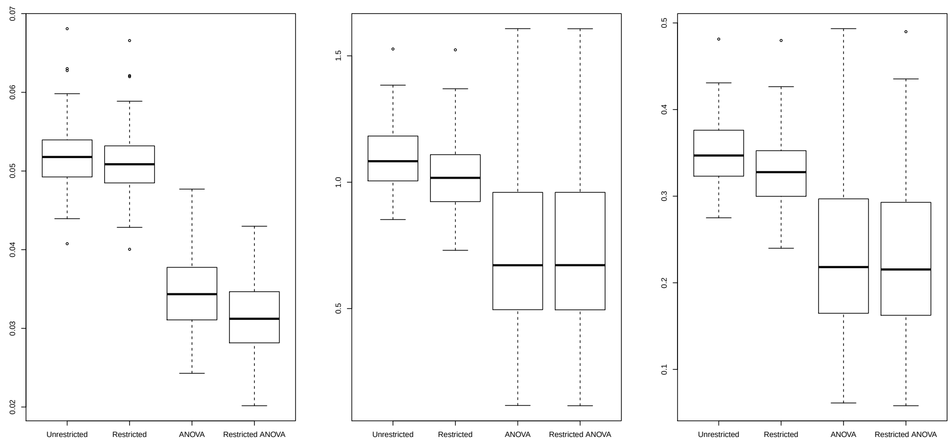

Figure S14. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = 0$ and $n_{x_p} = 20$.

- $n_{z_p} = 10$, $n_{x_p} = 5$

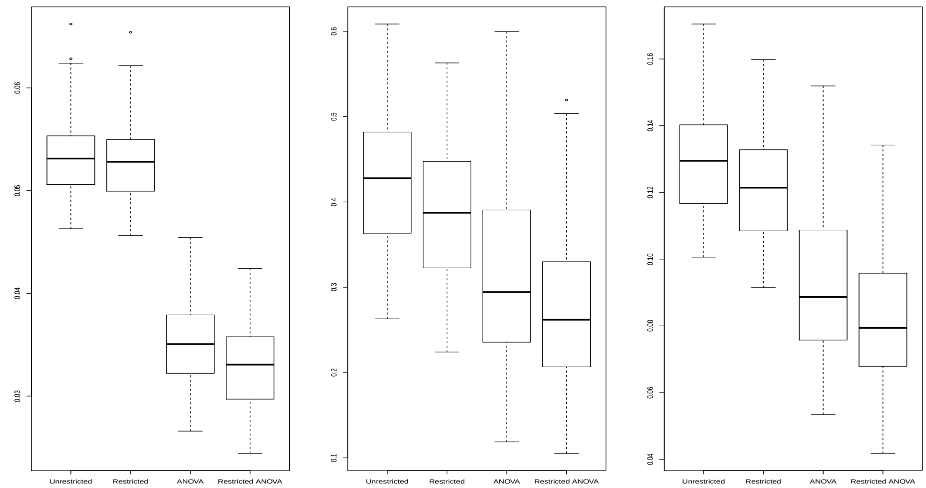

Figure S15. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = 10$ and $n_{x_p} = 5$.

- $n_{z_p} = n_{x_p} = 10$

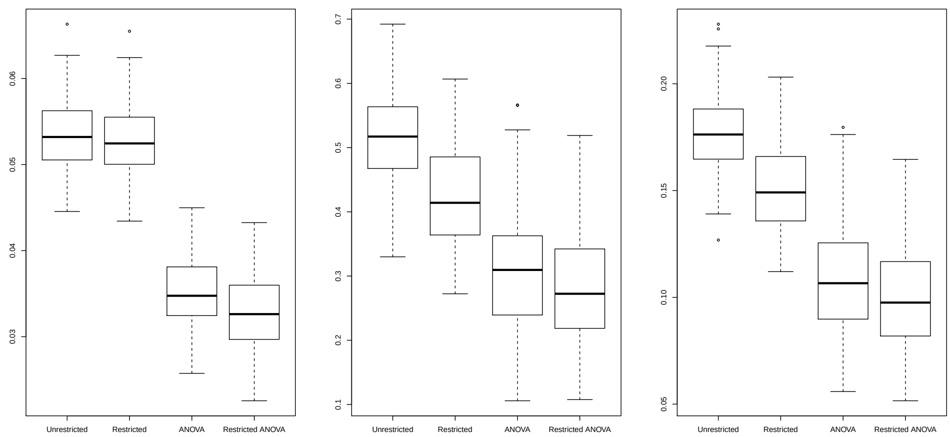

Figure S16. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = n_{x_p} = 10$.

- $n_{z_p} = 10$, $n_{x_p} = 15$

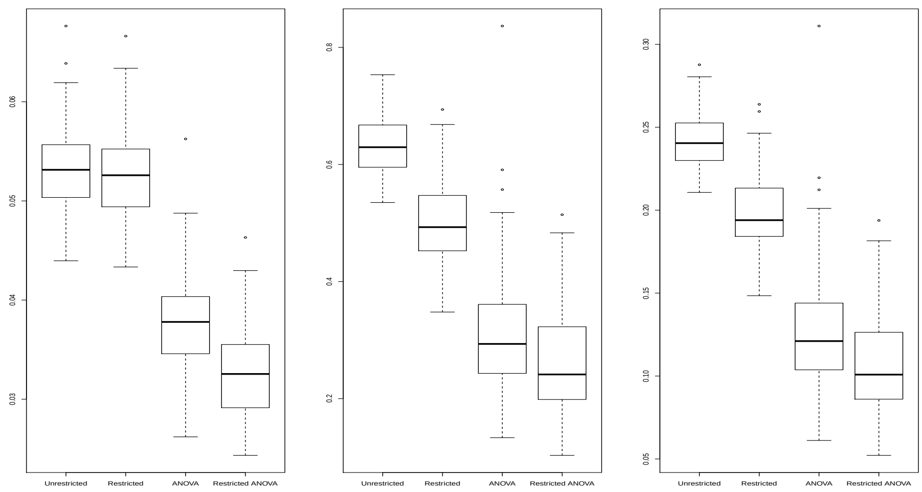

Figure S17. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = 10$ and $n_{x_p} = 15$.

- $n_{z_p} = 10, n_{x_p} = 20$

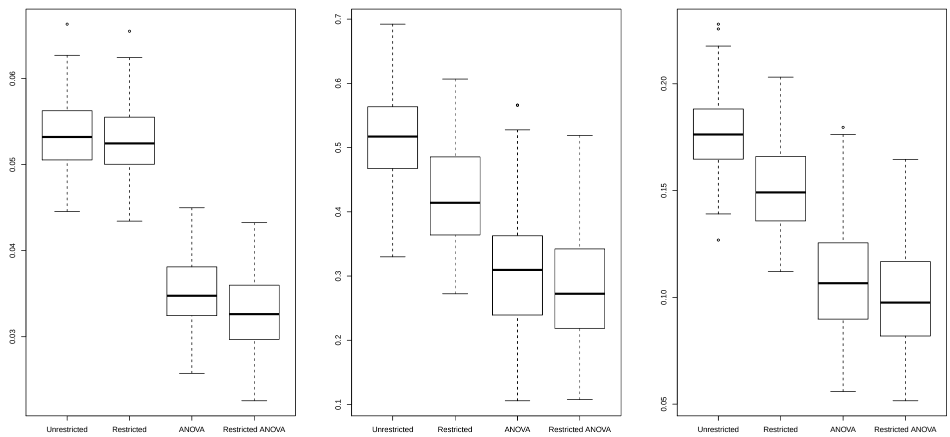

Figure S18. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = 10$ and $n_{x_p} = 20$.

- $n_{z_p} = 20, n_{x_p} = 5$

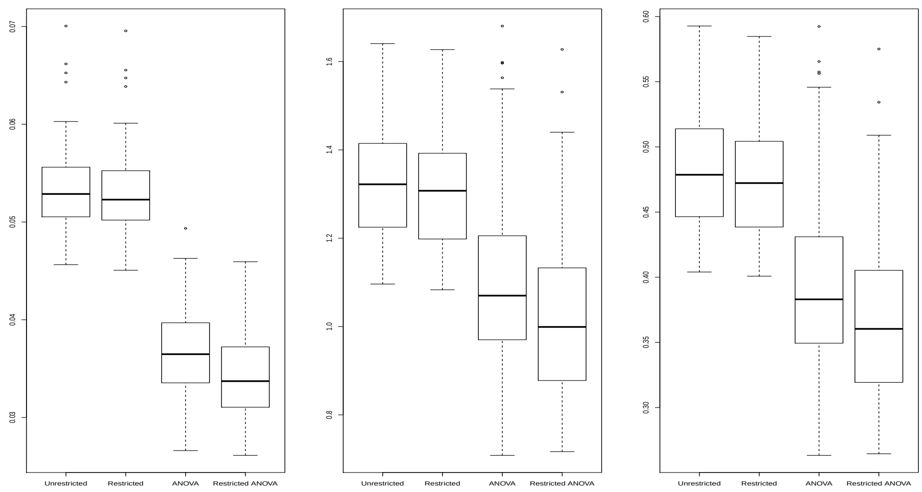

Figure S19. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = 20$ and $n_{x_p} = 5$.

- $n_{z_p} = 20$, $n_{x_p} = 10$

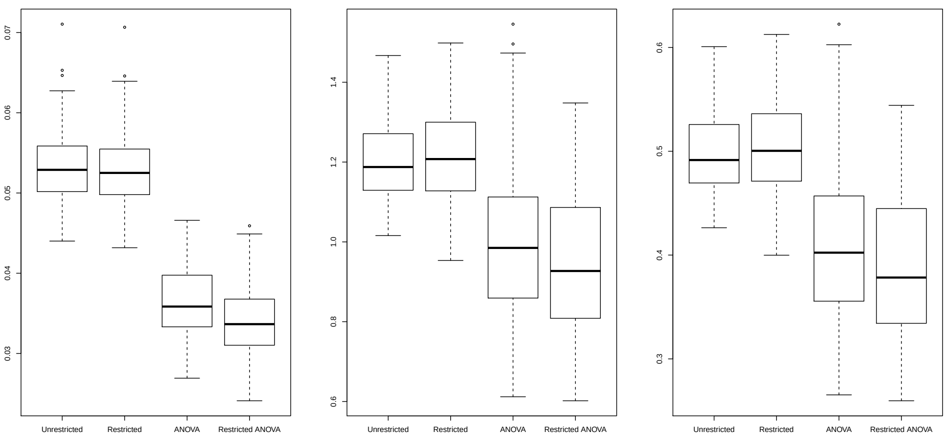

Figure S20. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = 20$ and $n_{x_p} = 10$.

- $n_{z_p} = 20$, $n_{x_p} = 15$

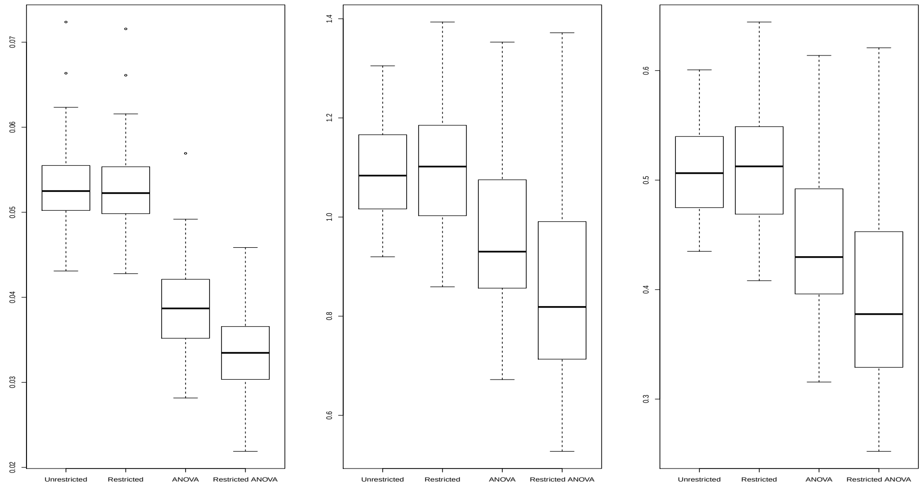

Figure S21. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = 20$ and $n_{x_p} = 15$.

- $n_{z_p} = n_{x_p} = 20$

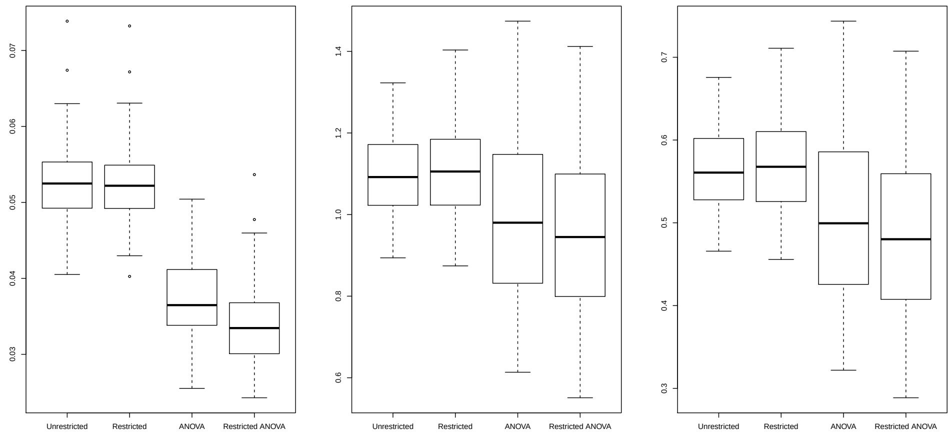

Figure S22. MAE in the fit (left panel), in the forecast (middle panel) and in total (right panel) of smooth models in scenario 2 and $n_{z_p} = n_{x_p} = 20$.

References

1. Harville, D. *Matrix Algebra from a Statistician's Perspective*; Springer, 2000.