EM estimation of a multivariate space-time data fusion model with varying coefficients

Francesco Finazzi and Alessandro Fassò
University of Bergamo - Department of Engineering

January 25, 2013

Abstract

Physical variables can be observed over space and time with different measuring instruments resulting in datasets characterized by different spatial supports. In many cases, the observations are available as point data and pixel data over a regular grid. This paper considers a space-time model for the data-fusion of multiple variables observed as both point data and pixel data. The model is based on latent variables and it is able to handle covariates, missing data and heterotopic spatial locations. The model parameters are estimated by means of the Expectation Maximization algorithm and closed form estimation formulas are derived. The model and the estimation formulas are implemented within the D-STEM software.

1 Introduction

Complex space-time phenomena are usually characterized by multiple physical variables interacting with each other. In order to study and understand the phenomena, the related variables are to be measured over space and time by means of one or more measuring instruments. In general, there is a trade-off between the spatial density of the collected data and their accuracy: the higher the accuracy the lower the spatial density. In many cases, therefore, the data collected using the more accurate measuring instruments are used to calibrate the less accurate but denser data. When the air quality of a region of space is assessed, for instance, the pollutant concentration measured by ground level monitoring stations is used to calibrate the remote sensing observations provided as a grid of pixels over the area of the region (see [1] for more details). In this case, the remote sensing observations are less accurate but, with the exclusion of the missing data, they may cover the entire region. In other cases, the secondary data source is the output of a physical model related to the observed variables, which is also provided as a regular grid of pixels. In all the cases, a data fusion problem is to be solved by taking into account the different spatial support of the data, the different accuracy and their spatio-temporal cross-correlation.

In this paper, a multivariate space-time model for data fusion is considered and the estimation formulas of the model parameters are derived. The model is able to handle multiple variables observed as both point data over a given set of spatial locations and pixel data over a regular grid. Covariates may be considered as space-time varying coefficients to be interacted with the latent variables at the basis of the model. Each observed variable can be characterized by missing data and the sets of spatial locations at which the point data are collected can be disjoint across the variables. Finally, time is considered to be discrete and it is assumed that all the variables are observed synchronously and at fixed temporal intervals. For many applications, the observation at time t can be related to the time average over the temporal interval $(t-1, t]$.

2 The multivariate data fusion model

Suppose that q physical variables are defined over the region of space $\mathcal{D} \subset \mathbb{R}^2$. Each variable, thus, can be measured at each spatial location $\mathbf{s} \in \mathcal{D}$ for each time $t = 1, ..., T$. Moreover, suppose that $\mathcal{D}$ is covered by a regular grid of blocks $\mathcal{B} = \{B_1, ..., B_m\}$ and that the q variables can be observed at each block $B_i \in \mathcal{B}$.

Let $\mathbf{y}(\mathbf{s},t) = (y_1(\mathbf{s},t), ..., y_q(\mathbf{s},t))'$ and $\mathbf{y}(B,t) = (y_1(B,t), ..., y_q(B,t))'$ be the $q-$variate response variables at the generic site $\mathbf{s}$, block $B \ni \mathbf{s}$ and time t. The model equation is given by

$$
\begin{aligned}
\mathbf{y}(\mathbf{s},t) &= \boldsymbol{\alpha}_{BP} \odot \mathbf{x}_{BP}(\mathbf{s},t) \odot \mathbf{w}^B(B,t) + \mathbf{X}_{\boldsymbol{\beta}}(\mathbf{s},t)\,\boldsymbol{\beta}_P + \mathbf{X}_{\mathbf{z}}(\mathbf{s},t)\,\mathbf{z}(t) \\
&\quad + \sum_{j=1}^{c} \boldsymbol{\alpha}_{P,j} \odot \mathbf{x}_{P,j}(\mathbf{s},t) \odot \mathbf{w}^{P,j}(\mathbf{s},t) + \boldsymbol{\varepsilon}^P(\mathbf{s},t) \\
\mathbf{y}(B,t) &= \boldsymbol{\alpha}_B \odot \mathbf{w}^B(B,t) + \boldsymbol{\varepsilon}^B(B,t) \\
\mathbf{z}(t) &= \mathbf{G}\mathbf{z}(t-1) + \boldsymbol{\eta}(t)
\end{aligned}
\tag{1}
$$

where $\odot$ is the Hadamard product. In (1), $\mathbf{w}^B(B,t) = \left(w_1^B(B,t), ..., w_q^B(B,t)\right)'$, $\mathbf{w}^{P,j}(\mathbf{s},t) = \left(w_1^{P,j}(\mathbf{s},t), ..., w_q^{P,j}(\mathbf{s},t)\right)'$, $j = 1, ..., c$ and $\mathbf{z}(t) = (z_1(t), ..., z_p(t))'$ are latent random variables, $\boldsymbol{\varepsilon}^P(\mathbf{s},t) = \left(\varepsilon_1^P(\mathbf{s},t), ..., \varepsilon^P(\mathbf{s},t)\right)'$ and $\boldsymbol{\varepsilon}^B(B,t) = \left(\varepsilon_1^B(B,t), ..., \varepsilon_q^B(B,t)\right)'$ are the random measurement errors while $\boldsymbol{\eta}(t) = \left(\eta_1(t), ..., \eta_p(t)\right)'$ are the random temporal innovations. In particular, the following distributions hold

$$
\begin{aligned}
\mathbf{w}^B(B,t) &\sim N_q(\mathbf{0}, \Gamma_B) \\
\mathbf{w}^{P,j}(\mathbf{s},t) &\sim N_q(\mathbf{0}, \Gamma_{P,j})\,;\ j = 1, ..., c \\
\boldsymbol{\varepsilon}^P(\mathbf{s},t) &\sim N_q\left(\mathbf{0}, \operatorname{diag}\left(\boldsymbol{\sigma}_P^2\right)\right) \\
\boldsymbol{\varepsilon}^B(B,t) &\sim N_q\left(\mathbf{0}, \operatorname{diag}\left(\boldsymbol{\sigma}_B^2\right)\right) \\
\boldsymbol{\eta}(t) &\sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\eta}}) \\
\mathbf{z}(0) &\sim N_p(\boldsymbol{\nu}_0, \boldsymbol{\Sigma}_0)
\end{aligned}
$$

where

$$
\Gamma_B = \mathbf{V}_B \odot \begin{pmatrix} \operatorname{corr}_{\boldsymbol{\theta}_B}\left(w_1^B(B,t), w_1^B(B',t)\right) & \cdots & \operatorname{corr}_{\boldsymbol{\theta}_B}\left(w_1^B(B,t), w_q^B(B',t)\right) \\ \vdots & \ddots & \vdots \\ \operatorname{corr}_{\boldsymbol{\theta}_B}\left(w_q^B(B,t), w_1^B(B',t)\right) & \cdots & \operatorname{corr}_{\boldsymbol{\theta}_B}\left(w_q^B(B,t), w_q^B(B',t)\right) \end{pmatrix}
$$

is a matrix correlation function with $\mathbf{V}_B$ a valid correlation matrix and $\operatorname{corr}_{\boldsymbol{\theta}_B}\left(w_h^B(B,t), w_k^B(B',t)\right)$ the spatial correlation function between block B and block B' and parametrized by the parameter vector $\boldsymbol{\theta}_B$. Similarly,

$$
\Gamma_{P,j} = \mathbf{V}_{P,j} \odot \begin{pmatrix} \operatorname{corr}_{\boldsymbol{\theta}_{P,j}}\left(w_1^{P,j}(\mathbf{s},t), w_1^{P,j}(\mathbf{s}',t)\right) & \cdots & \operatorname{corr}_{\boldsymbol{\theta}_{P,j}}\left(w_1^{P,j}(\mathbf{s},t), w_q^{P,j}(\mathbf{s}',t)\right) \\ \vdots & \ddots & \vdots \\ \operatorname{corr}_{\boldsymbol{\theta}_{P,j}}\left(w_q^{P,j}(\mathbf{s},t), w_1^{P,j}(\mathbf{s}',t)\right) & \cdots & \operatorname{corr}_{\boldsymbol{\theta}_{P,j}}\left(w_q^{P,j}(\mathbf{s},t), w_q^{P,j}(\mathbf{s}',t)\right) \end{pmatrix}
$$

are c matrix correlation functions. The measurement errors are white noise in space and time but each variable retains its own variance so that $\boldsymbol{\sigma}_P^2 = \left(\sigma_{P,1}^2, ..., \sigma_{P,q}^2\right)'$ and $\boldsymbol{\sigma}_B^2 = \left(\sigma_{B,1}^2, ..., \sigma_{B,q}^2\right)'$. Finally, $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$ are valid covariance matrices.

The $q \times 1$ loading vectors

$$\begin{aligned}
\mathbf{x}_{BP}(\mathbf{s},t) &= (x_{BP,1}(\mathbf{s},t),...,\mathbf{x}_{BP,q}(\mathbf{s},t))' \\
\mathbf{x}_{P,j}(\mathbf{s},t) &= (x_{P,j,1}(\mathbf{s},t),...,\mathbf{x}_{P,j,q}(\mathbf{s},t))' ; j = 1,...,c
\end{aligned}$$

the $q \times b$ loading matrix $\mathbf{X}_{\boldsymbol{\beta}}(\mathbf{s},t)$ and the $q \times p$ loading matrix $\mathbf{X}_{\mathbf{z}}(\mathbf{s},t)$ are characterized by known coefficients. In particular

$$\begin{aligned}
\mathbf{X}_{\boldsymbol{\beta}}(\mathbf{s},t) &= \begin{pmatrix} \mathbf{x}_{\boldsymbol{\beta},1}(\mathbf{s},t) & \cdots & \mathbf{0}_{1\times b_q} \\ \vdots & \ddots & \vdots \\ \mathbf{0}_{1\times b_1} & \cdots & \mathbf{x}_{\boldsymbol{\beta},q}(\mathbf{s},t) \end{pmatrix} \\
\mathbf{X}_{\mathbf{z}}(\mathbf{s},t) &= \begin{pmatrix} \mathbf{x}_{\mathbf{z},1}(\mathbf{s},t) & \cdots & \mathbf{0}_{1\times p_q} \\ \vdots & \ddots & \vdots \\ \mathbf{0}_{1\times p_1} & \cdots & \mathbf{x}_{\mathbf{z},q}(\mathbf{s},t) \end{pmatrix}
\end{aligned}$$

where $\mathbf{x}_{\boldsymbol{\beta},i}(\mathbf{s},t)$, $i = 1,...,q$ are $1 \times b_i$ vectors, $b_1 + ... + b_q = b$, $\mathbf{x}_{\mathbf{z},1}(\mathbf{s},t)$ are $1 \times p_i$ vectors, $p_1 + ... + p_q = p$ and $\mathbf{0}_{h\times k}$ is the $h \times k$ matrix of zeros.

Finally $\boldsymbol{\alpha}_{BP} = (\alpha_{BP,1},...,\alpha_{BP,q})'$, $\boldsymbol{\alpha}_{P,j} = (\alpha_{P,1,j},...,\alpha_{P,q,j})'$, $j = 1,...,c$, $\boldsymbol{\alpha}_B = (\alpha_{B,1},...,\alpha_{B,q})'$ and $\boldsymbol{\beta}_P = (\boldsymbol{\beta}'_{P,1},...,\boldsymbol{\beta}'_{P,q})'$, $\boldsymbol{\beta}_{P,i} = (\beta_{i,1},...,\beta_{i,b_i})$, $i = 1,...,q$, $b_1 + ... + b_q = b$ are vectors of unknown scale coefficients. The full model parameter set is

$$\Psi = \left\{\boldsymbol{\alpha}_{BP}, \boldsymbol{\alpha}_{P,1},...,\boldsymbol{\alpha}_{P,c}, \boldsymbol{\alpha}_B, \boldsymbol{\beta}_P, \boldsymbol{\sigma}^2_P, \boldsymbol{\sigma}^2_B, \mathbf{V}_{P,1},...,\mathbf{V}_{P,c}, \boldsymbol{\theta}_{P,1},...,\boldsymbol{\theta}_{P,c}, \mathbf{V}_B, \boldsymbol{\theta}_B, \boldsymbol{\Sigma}_{\boldsymbol{\eta}}, \mathbf{G}, \boldsymbol{\nu}_0\right\}$$

and, as discussed in the next section, it is estimated following the maximum likelihood approach by means of the expectation-maximization (EM) algorithm.

Note that (1) extends both the models developed in [1] and [3] which can be considered as particular instances of (1). Indeed, model (1) is flexible in the sense that it can be based on a subset of the latent variables actually included while the known coefficients may be either time-variant, time-invariant or vectors of one. Moreover, constraints can be imposed on the model parameters. For instance $\mathbf{V}_B$, $\mathbf{V}_{P,j}$, $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$ and $\mathbf{G}$ can be diagonal matrices.

3 Model estimation

Suppose that the q variables are observed at the sets of spatial locations $\mathcal{S}_i = \{\mathbf{s}_{i,1},...\mathbf{s}_{i,n_i}\}$, $i = 1,...,q$ and over the regular grid $\mathcal{B}$. For each time t, thus, the $(n+mq) \times 1 = N \times 1$ observation vector is given by $\mathbf{y}_t = \left(\mathbf{y}_t(\mathcal{S})', \mathbf{y}_t(\mathcal{B})\right)'$ where

$$\begin{aligned}
\mathbf{y}_t(\mathcal{S}) &= \left(y_t(\mathcal{S}_1)', ..., y_t(\mathcal{S}_q)'\right)' \\
&= \left((y_1(\mathbf{s}_{1,1},t),...,y_1(\mathbf{s}_{1,n_1},t)),...,\left(y_q(\mathbf{s}_{q,1},t),...,y_1(\mathbf{s}_{q,n_q},t)\right)\right)'
\end{aligned}$$

$n = n_1 + ... + n_q$, $\mathcal{S} = \uplus_{i=1}^q \mathcal{S}_i$, while

$$\mathbf{y}_t(\mathcal{B}) = ((y_1(B_1,t),...,y_1(B_m,t)),...,(y_q(B_1,t),...,y_1(B_m,t)))'$$

Note that, in general, $\mathcal{S}_i \neq \mathcal{S}_j$, that is, each variable can be observed at a different set of spatial locations. Moreover, even if $\mathcal{S}_i$ and $\mathcal{B}$ are time-invariant, missing data are allowed.

Given the following vectors

$$
\begin{aligned}
\mathbf{x}_{BP,i,t} &= \left(x_{BP}(\mathbf{s}_{i,1},t),...,x_{BP}(\mathbf{s}_{i,n_i},t)\right)' \\
\mathbf{x}_{P,i,j,t} &= \left(x_{P,j}(\mathbf{s}_{i,1},t),...,x_{P,j}(\mathbf{s}_{i,n_i},t)\right)' \\
\mathbf{w}_t^B &= \left(\left(w_1^B\left(B_1,t\right),...,w_1^B\left(B_m,t\right)\right),...,\left(w_q^B\left(B_1,t\right),...,w_q^B\left(B_m,t\right)\right)\right)' \\
\mathbf{w}_{j,t}^P &= \left(\left(w_1^{P,j}(\mathbf{s}_{i,1},t),...,w_1^{P,j}(\mathbf{s}_{i,n_1},t)\right),...,\left(w_q^{P,j}(\mathbf{s}_{q,1},t),...,w_q^{P,j}(\mathbf{s}_{q,n_q},t)\right)\right)' \\
\boldsymbol{\varepsilon}_t &= \left(\left(\varepsilon_1^P(\mathbf{s}_{i,1},t),...,\varepsilon_1^P(\mathbf{s}_{i,n_1},t)\right),...,\left(\varepsilon_q^P(\mathbf{s}_{q,1},t),...,\varepsilon_q^P(\mathbf{s}_{q,n_q},t)\right),\right. \\
&\quad \left.,\left(\varepsilon_1^B(B_1,t),...,\varepsilon_1^B(B_m,t)\right),...,\left(\varepsilon_q^B(B_1,t),...,\varepsilon_q^B(B_m,t)\right)\right)'
\end{aligned}
$$

and the following matrices

$$
\mathbf{X}_{i,t}^{BP} = \begin{pmatrix} \operatorname{diag}\left(\mathbf{0}'_{n_1\times 1},...,\overbrace{\mathbf{x}'_{BP,i,t}}^{i-th\ position},...,\mathbf{0}'_{n_q\times 1}\right)' \\ \mathbf{0}_{(mq)\times n} \end{pmatrix} \qquad i=1,...,q
$$

$$
\mathbf{X}_{i,t}^{B} = \begin{pmatrix} \mathbf{0}_{n\times(mq)} \\ \operatorname{diag}\left(\mathbf{0}'_{m\times 1},...,\overbrace{\mathbf{1}'_{m\times 1}}^{i-th\ position},...,\mathbf{0}'_{m\times 1}\right)' \end{pmatrix} \qquad i=1,...,q
$$

$$
\mathbf{X}_{i,j,t}^{P} = \begin{pmatrix} \operatorname{diag}\left(\left(\mathbf{0}'_{n_1\times 1},...,\overbrace{\mathbf{x}'_{P,i,j,t}}^{i-th\ position},...,\mathbf{0}'_{n_q\times 1}\right)'\right) \\ \mathbf{0}_{(mq)\times n} \end{pmatrix} \qquad \begin{array}{l} i=1,...,q \\ j=1,...,c \end{array}
$$

where $\mathbf{1}_{h\times k}$ is the $h\times k$ matrix of ones, model (1) can be written in this form

$$
\begin{aligned}
\mathbf{y}_t &= \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t \\
\boldsymbol{\mu}_t &= \sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP}\tilde{\mathbf{w}}_t^B + \sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B}\mathbf{w}_t^B + \mathbf{X}_t^{\boldsymbol{\beta}}\boldsymbol{\beta}_P + \mathbf{X}_t^{\mathbf{z}}\mathbf{z}_t + \sum_{j=1}^{c}\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P}\mathbf{w}_{j,t}^P \\
\mathbf{z}_t &= \mathbf{G}\mathbf{z}_{t-1} + \boldsymbol{\eta}_t
\end{aligned}
\tag{2}
$$

where the $n\times 1$ vector $\tilde{\mathbf{w}}_t^B$ is given by $\tilde{\mathbf{w}}_t^B = \mathbf{M}\mathbf{w}_t^B$ and the matrix $\mathbf{M}$ maps $\mathbf{w}_t^B$ over the spatial locations in $\mathcal{S}$ in such a way that, for each $\tilde{w}_i^B\left(B_k,t\right)$, $i=1,...,q$, $k=1,...,n_i$, $B_k \ni \mathbf{s}_{i,k}$. This implies that a given element of $\mathbf{w}_t^B$ can be replicated more than one time in $\tilde{\mathbf{w}}_t^B$ if more than one $\mathbf{s}_{i,k}$ fall into the same block. Finally

$$
\begin{aligned}
\mathbf{X}_t^{\boldsymbol{\beta}} &= \operatorname{blockdiag}\left(\mathbf{X}_{\boldsymbol{\beta},1}(\mathcal{S}_1,t),...,\mathbf{X}_{\boldsymbol{\beta},q}(\mathcal{S}_q,t)\right) \\
\mathbf{X}_t^{\mathbf{z}} &= \operatorname{blockdiag}\left(\mathbf{X}_{\mathbf{z},1}(\mathcal{S}_1,t),...,\mathbf{X}_{\mathbf{z},q}(\mathcal{S}_q,t)\right)
\end{aligned}
$$

where

$$
\mathbf{X}_{\boldsymbol{\beta},i}(\mathcal{S}_i,t) = \begin{pmatrix} \mathbf{x}_{\boldsymbol{\beta},i}(\mathbf{s}_{1,1},t) \\ \vdots \\ \mathbf{x}_{\boldsymbol{\beta},i}(\mathbf{s}_{1,n_i},t) \end{pmatrix} \quad \mathbf{X}_{\mathbf{z},i}(\mathcal{S}_i,t) = \begin{pmatrix} \mathbf{x}_{\mathbf{z},i}(\mathbf{s}_{1,1},t) \\ \vdots \\ \mathbf{x}_{\mathbf{z},i}(\mathbf{s}_{1,n_i},t) \end{pmatrix} \ ; i=1,...,q
$$

Note that (2) is overly complicated with respect to (1) but it allows to write the model in such a way that the model parameters $\alpha_{BP,i}$, $\alpha_{B,i}$ and $\alpha_{P,i,j}$, $i = 1, ..., q$, $j = 1, ..., c$ are explicit as a standard product. This, in turn, allows to write the complete-data likelihood function in an easy manner and to derive closed-form estimation formulas.

3.1 Complete-data likelihood function

The complete-data likelihood $L(\Psi; \mathbf{Y}, \mathbf{Z}, \mathbf{W})$ function can be defined and factorized in the following way

$$
\begin{aligned}
L(\Psi; \mathbf{Y}, \mathbf{Z}, \mathbf{W}) &= \prod_{t=1}^{T} L(\Psi_{\mathbf{y}}; \mathbf{y}_t \mid \mathbf{z}_t, \mathbf{w}_t) \cdot L(\Psi_{\mathbf{z}_0}; \mathbf{z}_0) \cdot \prod_{t=1}^{T} L(\Psi_{\mathbf{z}}; \mathbf{z}_t \mid \mathbf{z}_{t-1}) \\
&\cdot \prod_{t=1}^{T} L\left(\Psi_B; \mathbf{w}_t^B\right) \cdot \prod_{j=1}^{c} \prod_{t=1}^{T} L\left(\Psi_P; \mathbf{w}_{j,t}^P\right)
\end{aligned} \tag{3}
$$

where

$$
\begin{aligned}
\mathbf{Y} &= \{\mathbf{y}_1, ..., \mathbf{y}_T\} \\
\mathbf{W} &= \left\{\mathbf{W}^B, \mathbf{W}_1^P, ..., \mathbf{W}_c^P\right\} \\
\mathbf{W}^B &= \left\{\mathbf{w}_1^B, ..., \mathbf{w}_T^B\right\} \\
\mathbf{W}_j^P &= \left\{\mathbf{w}_{j,1}^P, ..., \mathbf{w}_{j,T}^P\right\};\ j = 1, ..., c \\
\mathbf{Z} &= \{\mathbf{z}_0, \mathbf{z}_1, ..., \mathbf{z}_T\} \\
\mathbf{w}_t &= \left\{\mathbf{w}_t^B, \mathbf{w}_{1,t}^P, ..., \mathbf{w}_{c,t}^P\right\}
\end{aligned}
$$

The statistical distributions involved in (3) are

$$
\begin{aligned}
\mathbf{y}_t \mid \mathbf{z}_t, \mathbf{w}_t &\sim N_N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_\varepsilon) \\
\mathbf{z}_0 &\sim N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \\
\mathbf{z}_t \mid \mathbf{z}_{t-1} &\sim N_p(\mathbf{G}\mathbf{z}_{t-1}, \boldsymbol{\Sigma}_\eta) \\
\mathbf{w}_t^B &\sim N_{mq}(0, \boldsymbol{\Sigma}_B) \\
\mathbf{w}_{j,t}^P &\sim N_n(0, \boldsymbol{\Sigma}_{P,j});\ j = 1, ..., c
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{\Sigma}_\varepsilon &= \text{blockdiag}\left(\sigma_{P,1}^2 \mathbf{I}_{n_1}, ..., \sigma_{P,q}^2 \mathbf{I}_{n_q}, \sigma_{B,1}^2 \mathbf{I}_m, ..., \sigma_{B,q}^2 \mathbf{I}_m\right) \\
\boldsymbol{\Sigma}_B &= \begin{pmatrix} \Gamma_B(\mathbf{H}_{11}) & \cdots & v_{B,(1,q)}\Gamma_B(\mathbf{H}_{1q}) \\ \vdots & \ddots & \vdots \\ v_{B,(q,1)}\Gamma_B(\mathbf{H}_{q1}) & \cdots & \Gamma_B(\mathbf{H}_{qq}) \end{pmatrix} && (4) \\
\boldsymbol{\Sigma}_{P,j} &= \begin{pmatrix} \Gamma_{P,j}(\mathbf{H}_{11}) & \cdots & v_{P,j,(1,q)}\Gamma_{P,j}(\mathbf{H}_{1q}) \\ \vdots & \ddots & \vdots \\ v_{P,j,(q,1)}\Gamma_{P,j}(\mathbf{H}_{q1}) & \cdots & \Gamma_{P,j}(\mathbf{H}_{qq}) \end{pmatrix};\ j = 1, ..., c && (5)
\end{aligned}
$$

and where blockdiag is the block diagonal operator. In (4) and (5), $v_{B,(h,k)}$ is the (h,k) element of the $\mathbf{V}_B$ matrix, $v_{P,j,(h,k)}$ is the (h,k) element of the $\mathbf{V}_{P,j}$ matrix and $\mathbf{H}_{hk} = d(\mathcal{S}_h, \mathcal{S}_k)$ is the distance matrix between the spatial locations in $\mathcal{S}_i$ and the spatial locations in $\mathcal{S}_j$, $h, k = 1, ..., q$. The matrix $\boldsymbol{\Sigma}_0$ has the same dimension of $\boldsymbol{\Sigma}_\eta$ and is supposed to be known.

The complete-data log-likelihood function is given by:

$$
\begin{aligned}
-2l\left(\Psi;\mathbf{Y},\mathbf{Z},\mathbf{W}\right) \;=\;& T\log\left|\mathbf{\Sigma}_{\varepsilon}\right| + \sum_{t=1}^{T}\mathbf{e}_t'\mathbf{\Sigma}_{\varepsilon}^{-1}\mathbf{e}_t \\
& +\log\left|\mathbf{\Sigma}_0\right| + \left(\mathbf{z}_0-\boldsymbol{\mu}_0\right)'\mathbf{\Sigma}_0^{-1}\left(\mathbf{z}_0-\boldsymbol{\mu}_0\right) \\
& +T\log\left|\mathbf{\Sigma}_{\eta}\right| + \sum_{t=1}^{T}\left(\mathbf{z}_t-\mathbf{G}\mathbf{z}_{t-1}\right)'\mathbf{\Sigma}_{\eta}^{-1}\left(\mathbf{z}_t-\mathbf{G}\mathbf{z}_{t-1}\right) \\
& +T\log\left|\mathbf{\Sigma}_{B}\right| + \sum_{t=1}^{T}\left(\mathbf{w}_t^B\right)'\mathbf{\Sigma}_B^{-1}\mathbf{w}_t^B \\
& +\sum_{j=1}^{c}T\log\left|\mathbf{\Sigma}_{P,j}\right| + \sum_{t=1}^{T}\left(\mathbf{w}_{j,t}^P\right)'\mathbf{\Sigma}_{P,j}^{-1}\mathbf{w}_{j,t}^P
\end{aligned}
$$

where $\mathbf{e}_t = \mathbf{y}_t - \boldsymbol{\mu}_t$.

3.2 Missing data

In order to deal with missing data, the following notation is introduced. Given $\mathbf{y}_t$ the observation vector at time t, $\mathbf{y}_t$ is partitioned as $\tilde{\mathbf{y}}_t = \left(\mathbf{y}_t^{(1)},\mathbf{y}_t^{(2)}\right)'$, where $\mathbf{y}_t^{(1)} = \mathbf{L}_t\mathbf{y}_t$ is the sub-vector of non-missing data at time t and $\mathbf{L}_t$ is the appropriate elimination matrix. The vector $\tilde{\mathbf{y}}_t$ is thus a permutation of $\mathbf{y}_t$ and $\mathbf{y}_t = \mathbf{D}_t\tilde{\mathbf{y}}_t$, with $\mathbf{D}_t$ the proper commutation matrix. In the sequel, given $\mathbf{b}_t$ a generic $N\times 1$ vector and $\mathbf{B}_t$ a generic $N\times N$ matrix at time t, $\mathbf{b}_t^{(1)}$ and $\mathbf{B}_t^{(1)}$ will stand for $\mathbf{L}_t\mathbf{b}_t$ and $\mathbf{L}_t\mathbf{B}_t\mathbf{L}_t'$, respectively. On the other hand, if $\mathbf{B}_t$ is a $N\times K$ matrix, than $\mathbf{B}_t^{(1)} = \mathbf{L}_t\mathbf{B}_t$. Finally, $\mathbf{B}_{i,t} = \mathbf{O}_{i,t}\mathbf{B}_t\mathbf{O}_{i,t}'$ is the matrix $\mathbf{B}_t$ the rows and columns of which are restricted to the i-th variable, $i = 1,...,2q$.

The partitioned measurement equation (2) becomes $\mathbf{y}_t^{(l)} = \boldsymbol{\mu}_t^{(l)} + \boldsymbol{\varepsilon}_t^{(l)}, l = 1,2$ and the variance-covariance matrix of the permuted errors is conformably partitioned, namely:

$$
\mathrm{Var}\left[\begin{pmatrix}\boldsymbol{\varepsilon}_t^{(1)} \\ \boldsymbol{\varepsilon}_t^{(2)}\end{pmatrix}\right] = \begin{pmatrix}\mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{12}' & \mathbf{R}_{22}\end{pmatrix}
$$

Since $\mathbf{\Sigma}_{\varepsilon}$ is diagonal, it follows that $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$ are diagonal matrices while $\mathbf{R}_{12} = \mathbf{0}_{(N-U_t)\times U_t}$, with U_t the number of non-missing data at time t.

3.3 EM algorithm

The EM algorithm is considered here in order to estimate the model parameter set Ψ. The algorithm is iterative and it is based on two steps namely the expectation step and the maximization step. The expectation step is defined by the following conditional expectation

$$
\begin{aligned}
Q\left(\Psi,\Psi^{\langle m\rangle}\right) &= E_{\Psi^{\langle m\rangle}}\left[-2l\left(\Psi;\mathbf{Y},\mathbf{Z},\mathbf{W}\right)\mid\mathbf{Y}^{(1)}\right] \\
&= E_{\Psi^{\langle m\rangle}}\left[E_{\Psi^{\langle m\rangle}}\left[-2l\left(\Psi;\mathbf{Y},\mathbf{Z},\mathbf{W}\right)\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right]\mid\mathbf{Y}^{(1)}\right].
\end{aligned}
$$

where $\mathbf{Y}^{(1)} = \left\{\mathbf{y}_1^{(1)},...,\mathbf{y}_T^{(1)}\right\}$. In what follows, $E\left(\cdot\mid\cdot\right)\equiv E_{\Psi^{\langle m\rangle}}\left(\cdot\mid\cdot\right)$, $\mathrm{Var}\left(\cdot\mid\cdot\right)\equiv\mathrm{Var}_{\Psi^{\langle m\rangle}}\left(\cdot\mid\cdot\right)$ and $\mathrm{Cov}\left(\cdot,\cdot\mid\cdot\right)\equiv\mathrm{Cov}_{\Psi^{\langle m\rangle}}\left(\cdot,\cdot\mid\cdot\right)$.

Considering the inner conditional expectation, the following result holds

$$
\begin{aligned}
&E\left[-2l\left(\Psi;\mathbf{Y},\mathbf{Z},\mathbf{W}\right)\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right]=T\log\left|\boldsymbol{\Sigma}_{\varepsilon}\right|\\
&+\operatorname{tr}\left[\boldsymbol{\Sigma}_{\varepsilon}^{-1}\sum_{t=1}^{T}E\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)E\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)'+\operatorname{Var}\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)\right]\\
&+\log\left|\boldsymbol{\Sigma}_0\right|+\operatorname{tr}\left[\boldsymbol{\Sigma}_0^{-1}\left(\mathbf{z}_0-\boldsymbol{\mu}_0\right)\left(\mathbf{z}_0-\boldsymbol{\mu}_0\right)'\right]\\
&+T\log\left|\boldsymbol{\Sigma}_{\eta}\right|+\operatorname{tr}\left[\boldsymbol{\Sigma}_{\eta}^{-1}\sum_{t=1}^{T}\left(\mathbf{z}_t-\mathbf{G}\mathbf{z}_{t-1}\right)\left(\mathbf{z}_t-\mathbf{G}\mathbf{z}_{t-1}\right)'\right]\\
&+T\log\left|\boldsymbol{\Sigma}_{B}\right|+\operatorname{tr}\left[\boldsymbol{\Sigma}_{B}^{-1}\sum_{t=1}^{T}\mathbf{w}_t^{B}\left(\mathbf{w}_t^{B}\right)'\right]\\
&+\sum_{j=1}^{c}T\log\left|\boldsymbol{\Sigma}_{P,j}\right|+\operatorname{tr}\left[\boldsymbol{\Sigma}_{P,j}^{-1}\sum_{t=1}^{T}\mathbf{w}_{k,t}^{P}\left(\mathbf{w}_{k,t}^{P}\right)'\right]
\end{aligned}
\tag{6}
$$

where

$$
E\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)=\mathbf{D}_t\begin{pmatrix}\mathbf{e}_t^{(1)}\\ \mathbf{R}_{21}\mathbf{R}_{11}^{-1}\mathbf{e}_t^{(1)}\end{pmatrix}=\mathbf{D}_t\begin{pmatrix}\mathbf{e}_t^{(1)}\\ \mathbf{0}_{(N-U_t)\times 1}\end{pmatrix}
$$

and

$$
\begin{aligned}
\operatorname{Var}\left[\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right]&=\mathbf{D}_t\begin{pmatrix}\mathbf{0}_{U_t\times U_t}&\mathbf{0}_{U_t\times(N-U_t)}\\ \mathbf{0}_{(N-U_t)\times U_t}&\mathbf{R}_{22}-\mathbf{R}_{21}\mathbf{R}_{11}^{-1}\mathbf{R}_{12}\end{pmatrix}\mathbf{D}_t'\\
&=\mathbf{D}_t\begin{pmatrix}\mathbf{0}_{U_t\times U_t}&\mathbf{0}_{U_t\times(N-U_t)}\\ \mathbf{0}_{(N-U_t)\times U_t}&\mathbf{R}_{22}\end{pmatrix}\mathbf{D}_t'
\end{aligned}
$$

Moreover, $\boldsymbol{\mu}_0\equiv\boldsymbol{\mu}_0^{\langle m\rangle}$, $\boldsymbol{\Sigma}_0\equiv\boldsymbol{\Sigma}_0^{\langle m\rangle}$, $\mathbf{G}\equiv\mathbf{G}^{\langle m\rangle}$, $\boldsymbol{\Sigma}_\eta\equiv\boldsymbol{\Sigma}_\eta^{\langle m\rangle}$, $\boldsymbol{\Sigma}_\varepsilon\equiv\boldsymbol{\Sigma}_\varepsilon^{\langle m\rangle}$, $\boldsymbol{\Sigma}_B\equiv\boldsymbol{\Sigma}_B^{\langle m\rangle}$ and $\boldsymbol{\Sigma}_{P,j}^{\langle m\rangle}$, $j=1,...,c$, that is, vector and matrices are evaluated using the value of the model parameters at the m-th iteration of the EM algorithm.

Applying the outer conditional expectation to the rhs of (6) it follows that

$$
\begin{aligned}
&E\left[E\left[-2l\left(\Psi;\mathbf{Y},\mathbf{Z},\mathbf{W}\right)\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right]\mid\mathbf{Y}^{(1)}\right]=\\
&=T\log\left|\boldsymbol{\Sigma}_{\varepsilon}\right|+\operatorname{tr}\left(\boldsymbol{\Sigma}_{\varepsilon}^{-1}\sum_{t=1}^{T}\boldsymbol{\Omega}_t\right)\\
&+\log\left|\boldsymbol{\Sigma}_0\right|+\operatorname{tr}\left[\boldsymbol{\Sigma}_0^{-1}\left\{\left[E\left(\mathbf{z}_0\mid\mathbf{Y}^{(1)}\right)-\boldsymbol{\mu}_0\right]\left[E\left(\mathbf{z}_0\mid\mathbf{Y}^{(1)}\right)-\boldsymbol{\mu}_0\right]'+\operatorname{Var}\left(\mathbf{z}_0\mid\mathbf{Y}^{(1)}\right)\right\}\right]\\
&+T\log\left|\boldsymbol{\Sigma}_{\eta}\right|+\operatorname{tr}\left[\boldsymbol{\Sigma}_{\eta}^{-1}\left(\mathbf{S}_{11}-\mathbf{S}_{10}\mathbf{G}'-\mathbf{G}\mathbf{S}_{10}'+\mathbf{G}\mathbf{S}_{00}\mathbf{G}'\right)\right]\\
&+T\log\left|\boldsymbol{\Sigma}_{B}\right|+\operatorname{tr}\left[\boldsymbol{\Sigma}_{B}^{-1}\sum_{t=1}^{T}E\left(\mathbf{w}_t^{B}\mid\mathbf{Y}^{(1)}\right)E\left(\mathbf{w}_t^{B}\mid\mathbf{Y}^{(1)}\right)'+\operatorname{Var}\left(\mathbf{w}_t^{B}\mid\mathbf{Y}^{(1)}\right)\right]\\
&+\sum_{j=1}^{c}T\log\left|\boldsymbol{\Sigma}_{P,j}\right|+\operatorname{tr}\left[\boldsymbol{\Sigma}_{P,j}^{-1}\sum_{t=1}^{T}E\left(\mathbf{w}_{j,t}^{P}\mid\mathbf{Y}^{(1)}\right)E\left(\mathbf{w}_{j,t}^{P}\mid\mathbf{Y}^{(1)}\right)'+\operatorname{Var}\left(\mathbf{w}_{j,t}^{P}\mid\mathbf{Y}^{(1)}\right)\right]
\end{aligned}
\tag{7}
$$

where

$$
\begin{aligned}
\boldsymbol{\Omega}_t&=E\left[E\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)E\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)'+\operatorname{Var}\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)\mid\mathbf{Y}^{(1)}\right]\\
&=E\left[E\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)E\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)'\mid\mathbf{Y}^{(1)}\right]+\operatorname{Var}\left(\mathbf{e}_t\mid\mathbf{Y}^{(1)},\mathbf{Z},\mathbf{W}\right)\\
&=\mathbf{D}_t\begin{pmatrix}\boldsymbol{\Omega}_t^{(1)}&\boldsymbol{\Omega}_t^{(1)}\mathbf{R}_{11}^{-1}\mathbf{R}_{21}\\ \mathbf{R}_{12}\mathbf{R}_{11}^{-1}\left(\boldsymbol{\Omega}_t^{(1)}\right)'&\mathbf{R}_{21}\mathbf{R}_{11}^{-1}\boldsymbol{\Omega}_t^{(1)}\mathbf{R}_{11}^{-1}\mathbf{R}_{21}\end{pmatrix}\mathbf{D}_t'+\mathbf{D}_t\begin{pmatrix}\mathbf{0}_{U_t\times U_t}&\mathbf{0}_{U_t\times(N-U_t)}\\ \mathbf{0}_{(N-U_t)\times U_t}&\mathbf{R}_{22}\end{pmatrix}\mathbf{D}_t'\\
&=\mathbf{D}_t\begin{pmatrix}\boldsymbol{\Omega}_t^{(1)}&\mathbf{0}_{U_t\times(N-U_t)}\\ \mathbf{0}_{U_t\times(N-U_t)}&\mathbf{R}_{22}\end{pmatrix}\mathbf{D}_t'
\end{aligned}
\tag{8}
$$

and

$$\mathbf{S}_{11} = \sum_{t=1}^{T} \mathbf{z}_t^T \left(\mathbf{z}_t^T\right)' + \mathbf{P}_t^T$$
$$\mathbf{S}_{10} = \sum_{t=1}^{T} \mathbf{z}_t^T \left(\mathbf{z}_{t-1}^T\right)' + \mathbf{P}_{t,t-1}^T$$
$$\mathbf{S}_{00} = \sum_{t=1}^{T} \mathbf{z}_{t-1}^T \left(\mathbf{z}_{t-1}^T\right)' + \mathbf{P}_{t-1}^T$$

with $\mathbf{z}_t^T = E\left(\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right)$, $\mathbf{P}_t^T = \mathrm{Var}\left(\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right)$ and $\mathbf{P}_{t,t-1}^T = \mathrm{Cov}\left(\mathbf{z}_t, \mathbf{z}_{t-1} \mid \mathbf{Y}^{(1)}\right)$ the output of the Kalman smoother as detailed in [1].

The matrix $\mathbf{\Omega}_t^{(1)}$ in (8) is given by

$$\mathbf{\Omega}_t^{(1)} = E\left(\mathbf{e}_t^{(1)} \mid \mathbf{Y}^{(1)}\right) E\left(\mathbf{e}_t^{(1)} \mid \mathbf{Y}^{(1)}\right)' + \mathrm{Var}\left(\mathbf{e}_t^{(1)} \mid \mathbf{Y}^{(1)}\right) \tag{9}$$

where

$$\begin{aligned}
E\left(\mathbf{e}_t^{(1)} \mid \mathbf{Y}^{(1)}\right) &= E\left(\mathbf{y}_t^{(1)} - \boldsymbol{\mu}_t^{(1)} \mid \mathbf{Y}^{(1)}\right) && (10)\\
&= \mathbf{y}_t^{(1)} - \sum_{i=1}^{q} \alpha_{BP,i} \mathbf{X}_{i,t}^{BP,(1)} E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{Y}^{(1)}\right) && (11)\\
&\quad - \sum_{i=1}^{q} \alpha_{B,i} \mathbf{X}_{i,t}^{B,(1)} E\left(\mathbf{w}_t^B \mid \mathbf{Y}^{(1)}\right) - \mathbf{X}_t^{\boldsymbol{\beta},(1)} \boldsymbol{\beta}_P + \mathbf{X}_t^{\mathbf{z},(1)} E\left(\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right)\\
&\quad - \sum_{j=1}^{c} \sum_{i=1}^{q} \alpha_{P,i,j} \mathbf{X}_{i,j,t}^{P,(1)} E\left(\mathbf{w}_{j,t}^P \mid \mathbf{Y}^{(1)}\right)
\end{aligned}$$

and

$$
\begin{aligned}
&\mathrm{Var}\left[\mathbf{e}_t^{(1)} \mid \mathbf{Y}^{(1)}\right] = \\
&= \mathrm{Var}\left[\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\tilde{\mathbf{w}}_t^{B} + \sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\mathbf{w}_t^{B} + \mathbf{X}_t^{\mathbf{z},(1)}\mathbf{z}_t + \sum_{j=1}^{c}\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P,(1)}\mathbf{w}_{j,t}^{P} \mid \mathbf{Y}^{(1)}\right] \\
&= \left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\right)\mathrm{Var}\left(\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\right)' \\
&+ \left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\right)\mathrm{Var}\left(\mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\right)' \\
&+\mathbf{X}_t^{\mathbf{z},(1)}\mathrm{Var}\left(\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right)\left(\mathbf{X}_t^{\mathbf{z},(1)}\right)' \\
&+ \sum_{j=1}^{c}\sum_{h=1}^{c}\left(\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P,(1)}\right)\mathrm{Cov}\left(\mathbf{w}_{j,t}^{P},\mathbf{w}_{h,t}^{P} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{P,i,h}\mathbf{X}_{i,h,t}^{P,(1)}\right)' + \\
&+ \left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\right)\mathrm{Cov}\left(\tilde{\mathbf{w}}_t^{B},\mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\right)' + \\
&+ \left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\right)\mathrm{Cov}\left(\mathbf{w}_t^{B},\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\right) \\
&+ \left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\right)\mathrm{Cov}\left(\tilde{\mathbf{w}}_t^{B},\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right)\left(\mathbf{X}_t^{\mathbf{z},(1)}\right)' \\
&+\mathbf{X}_t^{\mathbf{z},(1)}\mathrm{Cov}\left(\mathbf{z}_t,\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\right) \\
&+ \sum_{j=1}^{c}\left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\right)\mathrm{Cov}\left(\tilde{\mathbf{w}}_t^{B},\mathbf{w}_{j,t}^{P} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P,(1)}\right)' \\
&+ \sum_{j=1}^{c}\left(\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P,(1)}\right)\mathrm{Cov}\left(\mathbf{w}_{j,t}^{P},\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\right)' \\
&+ \left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\right)\mathrm{Cov}\left(\mathbf{w}_t^{B,(1)},\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right)\left(\mathbf{X}_t^{\mathbf{z},(1)}\right)' \\
&+\mathbf{X}_t^{\mathbf{z},(1)}\mathrm{Cov}\left(\mathbf{z}_t,\mathbf{w}_t^{B,(1)} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\right)' \\
&+ \sum_{j=1}^{c}\left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\right)\mathrm{Cov}\left(\mathbf{w}_t^{B,(1)},\mathbf{w}_{j,t}^{P,(1)} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P,(1)}\right)' \\
&+ \sum_{j=1}^{c}\left(\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P,(1)}\right)\mathrm{Cov}\left(\mathbf{w}_{j,t}^{P,(1)},\mathbf{w}_t^{B,(1)} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\right)' \\
&+ \sum_{j=1}^{c}\mathbf{X}_t^{\mathbf{z},(1)}\mathrm{Cov}\left(\mathbf{z}_t,\mathbf{w}_{j,t}^{P} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P,(1)}\right)' \\
&+ \sum_{j=1}^{c}\left(\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P}\right)\mathrm{Cov}\left(\mathbf{w}_{j,t}^{P},\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right)\left(\mathbf{X}_t^{\mathbf{z}}\right)'.
\end{aligned}
\tag{12}
$$

Given $\mathbf{a}_t$ and $\mathbf{b}_t$ a generic vector at time t, the following identities hold

$$
\begin{aligned}
E\left(\mathbf{a}_t \mid \mathbf{Y}^{(1)},\mathbf{Z}\right) &= E\left(\mathbf{a}_t \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \\
\mathrm{Var}\left(\mathbf{a}_t \mid \mathbf{Y}^{(1)},\mathbf{Z}\right) &= \mathrm{Var}\left(\mathbf{a}_t \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \\
\mathrm{Cov}\left(\mathbf{a}_t,\mathbf{b}_t \mid \mathbf{Y}^{(1)},\mathbf{Z}\right) &= \mathrm{Cov}\left(\mathbf{a}_t,\mathbf{b}_t \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right)
\end{aligned}
$$

as, conditionally on $\mathbf{Z}$, $\mathbf{y}_t^{(1)} \perp \mathbf{y}_{t'}^{(1)}$ for each $t \neq t'$.

The quantities in (10) and (12) that have to be explicitly evaluated are the following

$$\begin{array}{l} E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{Y}^{(1)}\right) \\ E\left(\mathbf{w}_t^B \mid \mathbf{Y}^{(1)}\right) \\ E\left(\mathbf{w}_{j,t}^P \mid \mathbf{Y}^{(1)}\right); j=1,...,c \end{array} \tag{13}$$

$$\begin{aligned} \mathrm{Var}\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{Y}^{(1)}\right) &= \mathrm{Var}\left[E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &\quad +E\left[\mathrm{Var}\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right]. \end{aligned} \tag{14}$$

$$\begin{aligned} \mathrm{Var}\left(\mathbf{w}_t^B \mid \mathbf{Y}^{(1)}\right) &= \mathrm{Var}\left[E\left(\mathbf{w}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &\quad +E\left[\mathrm{Var}\left(\mathbf{w}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right]. \end{aligned} \tag{15}$$

$$\begin{aligned} \mathrm{Cov}\left(\tilde{\mathbf{w}}_t^B,\mathbf{w}_t^B \mid \mathbf{Y}^{(1)}\right) &= \mathrm{Cov}\left[E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right),E\left(\mathbf{w}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &\quad +E\left[\mathrm{Cov}\left(\tilde{\mathbf{w}}_t^B,\mathbf{w}_t^B \mid \mathbf{y}^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right]. \end{aligned} \tag{16}$$

$$\begin{aligned} \mathrm{Cov}\left(\tilde{\mathbf{w}}_t^B,\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right) &= \mathrm{Cov}\left[E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right),E\left(\mathbf{z}_t \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &\quad +E\left[\mathrm{Cov}\left(\tilde{\mathbf{w}}_t^B,\mathbf{z}_t \mid \mathbf{y}^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &= \mathrm{Cov}\left[E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right),\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right]. \end{aligned} \tag{17}$$

$$\begin{aligned} \mathrm{Cov}\left(\tilde{\mathbf{w}}_t^B,\mathbf{w}_{j,t}^P \mid \mathbf{Y}^{(1)}\right) &= \mathrm{Cov}\left[E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right),E\left(\mathbf{w}_{j,t}^P \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &\quad +E\left[\mathrm{Cov}\left(\tilde{\mathbf{w}}_t^B,\mathbf{w}_{j,t}^P \mid \mathbf{y}^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \end{aligned} \tag{18}$$

$$\begin{aligned} \mathrm{Cov}\left(\mathbf{w}_t^B,\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right) &= \mathrm{Cov}\left[E\left(\mathbf{w}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right),E\left(\mathbf{z}_t \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &\quad +E\left[\mathrm{Cov}\left(\mathbf{w}_t^B,\mathbf{z}_t \mid \mathbf{y}^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &= \mathrm{Cov}\left[E\left(\mathbf{w}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right),\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right]. \end{aligned} \tag{19}$$

$$\begin{aligned} \mathrm{Cov}\left(\mathbf{w}_t^B,\mathbf{w}_{j,t}^P \mid \mathbf{Y}^{(1)}\right) &= \mathrm{Cov}\left[E\left(\mathbf{w}_t^B \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right),E\left(\mathbf{w}_{j,t}^P \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &\quad +E\left[\mathrm{Cov}\left(\mathbf{w}_t^B,\mathbf{w}_{j,t}^P \mid \mathbf{y}^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right]. \end{aligned} \tag{20}$$

$$\begin{aligned} \mathrm{Cov}\left(\mathbf{w}_{j,t}^P,\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right) &= \mathrm{Cov}\left[E\left(\mathbf{w}_{j,t}^P \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right),E\left(\mathbf{z}_t \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &\quad +E\left[\mathrm{Cov}\left(\mathbf{w}_{j,t}^P,\mathbf{z}_t \mid \mathbf{y}^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \\ &= \mathrm{Cov}\left[E\left(\mathbf{w}_{j,t}^P \mid \mathbf{y}_t^{(1)},\mathbf{z}_t\right),\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right]. \end{aligned} \tag{21}$$

$$\operatorname{Cov}\left(\mathbf{w}_{j,t}^{P},\mathbf{w}_{h,t}^{P}\mid\mathbf{Y}^{(1)}\right) = \operatorname{Cov}\left[E\left(\mathbf{w}_{j,t}^{P}\mid\mathbf{y}_{t}^{(1)},\mathbf{z}_{t}\right),E\left(\mathbf{w}_{h,t}^{P}\mid\mathbf{y}_{t}^{(1)},\mathbf{z}_{t}\right)\mid\mathbf{Y}^{(1)}\right] + E\left[\operatorname{Cov}\left(\mathbf{w}_{j,t}^{P},\mathbf{w}_{h,t}^{P}\mid\mathbf{y}^{(1)},\mathbf{z}_{t}\right)\mid\mathbf{Y}^{(1)}\right]. \tag{22}$$

for $j,h=1,...,c$.

In order to evaluate (13-22), let

$$\mathbf{H}_t = \begin{bmatrix} \operatorname{Var}(\mathbf{y}_t) & \mathbf{X}_t^{\mathbf{z}}\operatorname{Var}(\mathbf{z}_t) \\ \operatorname{Var}(\mathbf{z}_t)(\mathbf{X}_t^{\mathbf{z}})' & \operatorname{Var}(\mathbf{z}_t) \end{bmatrix}$$

$$\mathbf{H}_t^{(1)} = \begin{bmatrix} \mathbf{L}_t\operatorname{Var}(\mathbf{y}_t)\mathbf{L}_t' & \mathbf{L}_t\mathbf{X}_t^{\mathbf{z}}\operatorname{Var}(\mathbf{z}_t) \\ \operatorname{Var}(\mathbf{z}_t)(\mathbf{X}_t^{\mathbf{z}})'\mathbf{L}_t' & \operatorname{Var}(\mathbf{z}_t) \end{bmatrix}$$

with

$$\begin{aligned}\operatorname{Var}(\mathbf{y}_t) = {} & \left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP}\right)\mathbf{M}\boldsymbol{\Sigma}_B\mathbf{M}'\left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP}\right)' \\ & + \left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B}\right)\boldsymbol{\Sigma}_B\left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B}\right)' \\ & + \mathbf{X}_t^{\mathbf{z}}\operatorname{Var}(\mathbf{z}_t)(\mathbf{X}_t^{\mathbf{z}})' \\ & + \sum_{j=1}^{c}\left(\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P}\right)\boldsymbol{\Sigma}_{P,j}\left(\sum_{i=1}^{q}\alpha_{P,i,j}\mathbf{X}_{i,j,t}^{P}\right)' + \boldsymbol{\Sigma}_{\varepsilon}\end{aligned}$$

and

$$\operatorname{vec}(\operatorname{Var}(\mathbf{z}_t)) = \left(\mathbf{I}_{p^2}-\mathbf{G}\otimes\mathbf{G}\right)^{-1}\operatorname{vec}(\boldsymbol{\Sigma}_{\eta})$$

The conditional expectations in (13) can be evaluated as follows

$$E\left(\tilde{\mathbf{w}}_t^B\mid\mathbf{Y}^{(1)}\right) = E\left[E\left(\tilde{\mathbf{w}}_t^B\mid\mathbf{y}_t^{(1)},\mathbf{z}_t\right)\mid\mathbf{Y}^{(1)}\right]$$

The inner conditional expectation is given by

$$E\left(\tilde{\mathbf{w}}_t^B\mid\mathbf{y}_t^{(1)},\mathbf{z}_t\right) = \operatorname{Cov}\left[\tilde{\mathbf{w}}_t^B,\begin{pmatrix}\mathbf{y}_t^{(1)}\\ \mathbf{z}_t\end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix}\mathbf{y}_t^{(1)}-\mathbf{X}_t^{\boldsymbol{\beta},(1)}\boldsymbol{\beta}_P\\ \mathbf{z}_t\end{pmatrix}$$

while the outer conditional expectation is equal to

$$E\left(\tilde{\mathbf{w}}_t^B\mid\mathbf{Y}^{(1)}\right) = \operatorname{Cov}\left[\tilde{\mathbf{w}}_t^B,\begin{pmatrix}\mathbf{y}_t^{(1)}\\ \mathbf{z}_t\end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix}\mathbf{y}_t^{(1)}-\mathbf{X}_t^{\boldsymbol{\beta},(1)}\boldsymbol{\beta}_P\\ \mathbf{z}_t^{T}\end{pmatrix} \tag{23}$$

where the covariance in (23) is given by

$$\begin{aligned}\operatorname{Cov}\left[\tilde{\mathbf{w}}_t^B,\begin{pmatrix}\mathbf{y}_t^{(1)}\\ \mathbf{z}_t\end{pmatrix}\right] & = \operatorname{Cov}\left[\tilde{\mathbf{w}}_t^B,\begin{pmatrix}\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\tilde{\mathbf{w}}_t^B+\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\mathbf{w}_t^B\\ \mathbf{z}_t\end{pmatrix}\right] \\ & = \begin{pmatrix}\mathbf{M}\boldsymbol{\Sigma}_B\mathbf{M}'\left(\sum_{i=1}^{q}\alpha_{BP,i}\mathbf{X}_{i,t}^{BP,(1)}\right)'+\mathbf{M}\boldsymbol{\Sigma}_B\left(\sum_{i=1}^{q}\alpha_{B,i}\mathbf{X}_{i,t}^{B,(1)}\right)' & \mathbf{0}_{U_t\times p}\end{pmatrix}\end{aligned}$$

The conditional variances in (14) and (15) are given by

$$\begin{aligned}
&\mathrm{Var}\left[E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)}, \mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] = \\
&= \mathrm{Var}\left(\mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix} \mathbf{y}_t^{(1)} - \mathbf{X}_t^{\boldsymbol{\beta},(1)}\boldsymbol{\beta}_P \\ \mathbf{z}_t \end{pmatrix} \mid \mathbf{Y}^{(1)}\right) \\
&= \mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix} \mathbf{0}_{U_t\times U_t} & \mathbf{0}_{U_t\times p} \\ \mathbf{0}_{p\times U_t} & \mathbf{P}_t^T \end{pmatrix}\left(\mathbf{H}_t^{(1)}\right)^{-1}\mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]'
\end{aligned}$$

while the conditional expectation of the variance is

$$E\left[\mathrm{Var}\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)}, \mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] \equiv \mathrm{Var}\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)}, \mathbf{z}_t\right)$$

with

$$\mathrm{Var}\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)}, \mathbf{z}_t\right) = \boldsymbol{\Sigma}_B - \mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]'$$

The conditional covariances in (16-22) are given by

$$\begin{aligned}
&\mathrm{Cov}\left[E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)}, \mathbf{z}_t\right), E\left(\mathbf{w}_t^B \mid \mathbf{y}_t^{(1)}, \mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] = \\
&= \mathrm{Cov}\{\,\mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix} \mathbf{y}_t^{(1)} - \mathbf{X}_t^{\boldsymbol{\beta},(1)}\boldsymbol{\beta}_P \\ \mathbf{z}_t \end{pmatrix}, \\
&\mathrm{Cov}\left[\mathbf{w}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix} \mathbf{y}_t^{(1)} - \mathbf{X}_t^{\boldsymbol{\beta},(1)}\boldsymbol{\beta}_P \\ \mathbf{z}_t \end{pmatrix} \mid \mathbf{Y}^{(1)}\} \\
&= \mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\mathrm{Var}\left[\begin{pmatrix} \mathbf{y}_t^{(1)} - \mathbf{X}_t^{\boldsymbol{\beta},(1)}\boldsymbol{\beta}_P \\ \mathbf{z}_t \end{pmatrix} \mid \mathbf{Y}^{(1)}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\mathrm{Cov}\left[\mathbf{w}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]' \\
&= \mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix} \mathbf{0}_{U_t\times U_t} & \mathbf{0}_{U_t\times p} \\ \mathbf{0}_{p\times U_t} & \mathbf{P}_t^T \end{pmatrix}\left(\mathbf{H}_t^{(1)}\right)^{-1}\mathrm{Cov}\left[\mathbf{w}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]'
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Cov}\left[E\left(\tilde{\mathbf{w}}_t^B \mid \mathbf{y}_t^{(1)}, \mathbf{z}_t\right), \mathbf{z}_t \mid \mathbf{Y}^{(1)}\right] &= \mathrm{Cov}\left\{\mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix} \mathbf{y}_t^{(1)} - \mathbf{X}_t^{\boldsymbol{\beta},(1)}\boldsymbol{\beta}_P \\ \mathbf{z}_t \end{pmatrix}, \mathbf{z}_t \mid \mathbf{Y}^{(1)}\right\} \\
&= \mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\mathrm{Cov}\left[\begin{pmatrix} \mathbf{y}_t^{(1)} - \mathbf{X}_t^{\boldsymbol{\beta},(1)}\boldsymbol{\beta}_P \\ \mathbf{z}_t \end{pmatrix}, \mathbf{z}_t \mid \mathbf{y}^{(1)}\right] \\
&= \mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right]\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix} \mathbf{0}_{U_t\times p} \\ \mathbf{P}_t^T \end{pmatrix}
\end{aligned}$$

The conditional expectations in (16-22) are given by

$$\mathrm{Cov}\left(\tilde{\mathbf{w}}_t^B, \mathbf{w}_t^B \mid \mathbf{y}^{(1)}, \mathbf{z}_t\right) = \mathbf{C}_{12}$$

where

$$\begin{pmatrix} \mathbf{C}_t^{11} & \mathbf{C}_t^{12} \\ \mathbf{C}_t^{21} & \mathbf{C}_t^{22} \end{pmatrix} = \begin{pmatrix} \mathbf{M}\boldsymbol{\Sigma}_B\mathbf{M}' & \mathbf{0}_{n\times qm} \\ \mathbf{0}_{qm\times n} & \boldsymbol{\Sigma}_B \end{pmatrix} - \begin{pmatrix} \mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right] \\ \mathrm{Cov}\left[\mathbf{w}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right] \end{pmatrix}\left(\mathbf{H}_t^{(1)}\right)^{-1}\begin{pmatrix} \mathrm{Cov}\left[\tilde{\mathbf{w}}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right] \\ \mathrm{Cov}\left[\mathbf{w}_t^B, \begin{pmatrix} \mathbf{y}_t^{(1)} \\ \mathbf{z}_t \end{pmatrix}\right] \end{pmatrix}'$$

Finally,

$$E\left[\operatorname{Cov}\left(\tilde{\mathbf{w}}_t^B,\mathbf{w}_t^B \mid \mathbf{y}^{(1)},\mathbf{z}_t\right) \mid \mathbf{Y}^{(1)}\right] = E\left(\mathbf{C}_t^{12} \mid \mathbf{Y}^{(1)}\right) = \mathbf{C}_t^{12}$$

as $\mathbf{C}_t^{12}$ is constant.

3.4 Updating formulas

The updating formulas for the model parameters are obtained from the maximization step of the EM algorithm as

$$\Psi^{\langle m+1\rangle} = \arg\max_{\Psi} Q\left(\Psi,\Psi^{\langle m\rangle}\right)$$

The updating formulas for

$$\{\Psi_{\mathbf{y}},\Psi_{\mathbf{z}}\} = \left\{\boldsymbol{\alpha}_{BP},\boldsymbol{\alpha}_{P,1},...,\boldsymbol{\alpha}_{P,c},\boldsymbol{\alpha}_B,\boldsymbol{\beta}_P,\boldsymbol{\sigma}_P^2,\boldsymbol{\sigma}_B^2,\boldsymbol{\Sigma}_{\boldsymbol{\eta}},\mathbf{G},\boldsymbol{\nu}_0\right\}$$

can be obtained in closed form by solving $\partial Q\left(\Psi,\Psi^{\langle m\rangle}\right)/\partial\left(\Psi_{\mathbf{y}},\Psi_{\mathbf{z}}\right) = \mathbf{0}$ while the model parameters

$$\{\mathbf{V}_{P,1},...,\mathbf{V}_{P,c},\boldsymbol{\theta}_{P,1},...,\boldsymbol{\theta}_{P,c},\mathbf{V}_B,\boldsymbol{\theta}_B\}$$

are estimated through numerical optimization. In particular

$$\left\{\mathbf{V}_B^{\langle m+1\rangle},\boldsymbol{\theta}_B^{\langle m+1\rangle}\right\} = \arg\max_{\mathbf{V}_B,\boldsymbol{\theta}_B} T\log|\boldsymbol{\Sigma}_B| + \operatorname{tr}\left[\boldsymbol{\Sigma}_B^{-1}\sum_{t=1}^{T}\mathbf{w}_t^B\left(\mathbf{w}_t^B\right)'\right]$$

and

$$\left\{\mathbf{V}_{P,j}^{\langle m+1\rangle},\boldsymbol{\theta}_{P,j}^{\langle m+1\rangle}\right\} = \arg\max_{\mathbf{V}_{P,j},\boldsymbol{\theta}_{P,j}} T\log|\boldsymbol{\Sigma}_{P,j}| + \operatorname{tr}\left[\boldsymbol{\Sigma}_{P,j}^{-1}\sum_{t=1}^{T}\mathbf{w}_{k,t}^P\left(\mathbf{w}_{k,t}^P\right)'\right];\ j=1,...,c$$

The closed form updating formulas for $\Psi_{\mathbf{y}} = \left\{\boldsymbol{\alpha}_{BP},\boldsymbol{\alpha}_{P,1},...,\boldsymbol{\alpha}_{P,c},\boldsymbol{\alpha}_B,\boldsymbol{\beta}_P,\boldsymbol{\sigma}_P^2,\boldsymbol{\sigma}_B^2\right\}$ can be derived by noting that the only term of (7) which is function of $\Psi_{\mathbf{y}}$ is $T\log|\boldsymbol{\Sigma}_\varepsilon| + \operatorname{tr}\left(\boldsymbol{\Sigma}_\varepsilon^{-1}\sum_{t=1}^{T}\boldsymbol{\Omega}_t\right)$. Solving

$$\frac{\partial T\log|\boldsymbol{\Sigma}_\varepsilon| + \operatorname{tr}\left(\boldsymbol{\Sigma}_\varepsilon^{-1}\sum_{t=1}^{T}\boldsymbol{\Omega}_t\right)}{\partial\Psi_{\mathbf{y}}} = 0$$

and by considering that

$$\begin{aligned}
\frac{\partial\operatorname{tr}\boldsymbol{\Omega}_t^{(1)}(\lambda)}{\partial\lambda} &= \operatorname{tr}\left[2\frac{\partial E\left(\mathbf{e}_t^{(1)}(\lambda) \mid \mathbf{Y}^{(1)}\right)}{\partial\lambda}E\left(\mathbf{e}_t^{(1)}(\lambda) \mid \mathbf{Y}^{(1)}\right)' + \frac{\partial\operatorname{Var}\left(\mathbf{e}_t^{(1)}(\lambda) \mid \mathbf{Y}^{(1)}\right)}{\partial\lambda}\right] \\
&= \operatorname{tr}\left[2E\left(\mathbf{e}_t^{(1)}(\lambda) \mid \mathbf{Y}^{(1)}\right)\frac{\partial E\left(\mathbf{e}_t^{(1)}(\lambda) \mid \mathbf{Y}^{(1)}\right)'}{\partial\lambda} + \frac{\partial\operatorname{Var}\left(\mathbf{e}_t^{(1)}(\lambda) \mid \mathbf{Y}^{(1)}\right)}{\partial\lambda}\right]
\end{aligned}$$

with λ the generic model parameter, the following updating formulas can be derived

$$\alpha_{BP,r}^{\langle m+1\rangle} = \frac{\begin{bmatrix} \sum_{t=1}^{T} \left[E\left(\mathbf{e}_t^{(1)} \mid \mathbf{Y}^{(1)}\right) + \alpha_{BP,r}^{\langle m\rangle} \mathbf{X}_{r,t}^{BP,(1)} E\left(\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right)\right] E\left(\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right)' \left(\mathbf{X}_{r,t}^{BP,(1)}\right)' + \\ -\frac{1}{2}\mathbf{X}_{r,t}^{BP,(1)} \operatorname{Var}\left(\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right) \left(\sum_{i=1:i\neq r}^{q} \alpha_{BP,i}^{\langle m\rangle} \left(\mathbf{X}_{r,t}^{BP,(1)}\right)' \right) + \\ -\frac{1}{2} \left(\sum_{i=1:i\neq r}^{q} \alpha_{BP,i}^{\langle m\rangle} \mathbf{X}_{r,t}^{BP,(1)} \right) \operatorname{Var}\left(\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right) \left(\mathbf{X}_{r,t}^{BP,(1)}\right)' + \\ -\frac{1}{2}\mathbf{X}_{r,t}^{BP,(1)} \operatorname{Cov}\left(\tilde{\mathbf{w}}_t^{B}, \mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right) \left(\sum_{i=1}^{q} \alpha_{B,i}^{\langle m\rangle} \left(\mathbf{X}_{i,t}^{B,(1)}\right)' \right) + \\ -\frac{1}{2} \left[\left(\sum_{i=1}^{q} \alpha_{B,i}^{\langle m\rangle} \mathbf{X}_{i,t}^{B} \right) \operatorname{Cov}\left(\mathbf{w}_t^{B}, \tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right) \right] \left(\mathbf{X}_{r,t}^{BP,(1)}\right)' + \\ -\frac{1}{2}\mathbf{X}_{r,t}^{BP,(1)} \operatorname{Cov}\left(\tilde{\mathbf{w}}_t^{B}, \mathbf{z}_t \mid \mathbf{Y}^{(1)}\right) \left(\mathbf{X}_t^{\mathbf{z},(1)}\right)' + \\ -\frac{1}{2}\mathbf{X}_t^{\mathbf{z},(1)} \operatorname{Cov}\left(\mathbf{z}_t, \tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right) \left(\mathbf{X}_{r,t}^{BP,(1)}\right)' + \\ -\frac{1}{2}\mathbf{X}_{r,t}^{BP,(1)} \sum_{j=1}^{c} \operatorname{Cov}\left(\tilde{\mathbf{w}}_t^{B}, \mathbf{w}_{j,t}^{P} \mid \mathbf{Y}^{(1)}\right) \left(\sum_{i=1}^{q} \alpha_{P,i,j}^{\langle m\rangle} \left(\mathbf{X}_{i,j,t}^{P,(1)}\right)' \right) + \\ -\frac{1}{2} \left[\sum_{j=1}^{c} \left(\sum_{i=1}^{q} \alpha_{P,i,j}^{\langle m\rangle} \mathbf{X}_{i,j,t}^{P,(1)} \right) \operatorname{Cov}\left(\mathbf{w}_{j,t}^{P}, \tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right) \right] \left(\mathbf{X}_{r,t}^{BP,(1)}\right)' \end{bmatrix}}{\operatorname{tr}\left\{ \sum_{t=1}^{T} \mathbf{X}_{r,t}^{BP,(1)} \left[E\left(\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right) E\left(\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right)' + \operatorname{Var}\left(\tilde{\mathbf{w}}_t^{B} \mid \mathbf{y}^{(1)}\right)\right] \left(\mathbf{X}_{r,t}^{BP,(1)}\right)' \right\}}$$

$$\alpha_{B,r}^{\langle m+1\rangle} = \frac{\operatorname{tr}\begin{bmatrix} \sum_{t=1}^{T} \left(E\left(\mathbf{e}_t^{(1)} \mid \mathbf{Y}^{(1)}\right) + \alpha_{B,r}^{\langle m\rangle} \mathbf{X}_{r,t}^{B,(1)} E\left(\mathbf{w}_t^{B,(1)} \mid \mathbf{Y}^{(1)}\right)\right) E\left(\mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right)' \left(\mathbf{X}_{r,t}^{B,(1)}\right)' \\ -\frac{1}{2}\mathbf{X}_{r,t}^{B,(1)} \operatorname{Var}\left(\mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right) \left(\sum_{i=1:i\neq r}^{q} \alpha_{B,i}^{\langle m\rangle} \left(\mathbf{X}_{r,t}^{B,(1)}\right)' \right) \\ -\frac{1}{2} \left(\sum_{i=1:i\neq r}^{q} \alpha_{B,i}^{\langle m\rangle} \mathbf{X}_{r,t}^{B,(1)} \right) \operatorname{Var}\left(\mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right) \left(\mathbf{X}_{r,t}^{B,(1)}\right)' \\ -\frac{1}{2} \left[\left(\sum_{i=1}^{q} \alpha_{B,i}^{\langle m\rangle} \mathbf{X}_{i,t}^{B} \right) \operatorname{Cov}\left(\mathbf{w}_t^{B}, \tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right) \right] \left(\mathbf{X}_{r,t}^{BP,(1)}\right)' \\ -\frac{1}{2}\mathbf{X}_{r,t}^{BP,(1)} \operatorname{Cov}\left(\tilde{\mathbf{w}}_t^{B}, \mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right) \left(\sum_{i=1}^{q} \alpha_{B,i}^{\langle m\rangle} \left(\mathbf{X}_{i,t}^{B,(1)}\right)' \right) \\ -\frac{1}{2}\mathbf{X}_{r,t}^{B,(1)} \operatorname{Cov}\left(\mathbf{w}_t^{B}, \mathbf{z}_t \mid \mathbf{Y}^{(1)}\right) \left(\mathbf{X}_t^{\mathbf{z},(1)}\right)' \\ -\frac{1}{2}\mathbf{X}_t^{\mathbf{z},(1)} \operatorname{Cov}\left(\mathbf{z}_t, \mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right) \left(\mathbf{X}_{r,t}^{B,(1)}\right)' \\ -\frac{1}{2}\mathbf{X}_{r,t}^{B,(1)} \sum_{j=1}^{c} \operatorname{Cov}\left(\mathbf{w}_t^{B}, \mathbf{w}_{j,t}^{P} \mid \mathbf{Y}^{(1)}\right) \left(\sum_{i=1}^{q} \alpha_{P,i,j}^{\langle m\rangle} \left(\mathbf{X}_{i,j,t}^{P,(1)}\right)' \right) \\ -\frac{1}{2} \left[\sum_{j=1}^{c} \left(\sum_{i=1}^{q} \alpha_{P,i,j}^{\langle m\rangle} \mathbf{X}_{i,j,t}^{P,(1)} \right) \operatorname{Cov}\left(\mathbf{w}_{j,t}^{P}, \mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right) \right] \left(\mathbf{X}_{r,t}^{B,(1)}\right)' \end{bmatrix}}{\operatorname{tr}\left\{ \sum_{t=1}^{T} \mathbf{X}_{r,t}^{B,(1)} \left[E\left(\mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right) E\left(\mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right)' + \operatorname{Var}\left(\mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right)\right] \left(\mathbf{X}_{r,t}^{B,(1)}\right)' \right\}}$$

$$
(\alpha_{G,r,s})^{\langle m+1\rangle} = \frac{\operatorname{tr}\left[\begin{array}{l}
\sum_{t=1}^{T}\left[E\left(\mathbf{e}_t^{(1)} \mid \mathbf{Y}^{(1)}\right)+\alpha_{G,r,s}^{\langle m\rangle}\mathbf{X}_{r,s,t}^{P,(1)}E\left(\mathbf{w}_{s,t}^{P} \mid \mathbf{Y}^{(1)}\right)\right]E\left(\mathbf{w}_{s,t}^{P} \mid \mathbf{Y}^{(1)}\right)'\left(\mathbf{X}_{r,s,t}^{P,(1)}\right)' \\
-\frac{1}{2}\mathbf{X}_{r,s,t}^{P,(1)}\operatorname{Cov}\left[\mathbf{w}_{s,t}^{P},\mathbf{z}_t \mid \mathbf{Y}^{(1)}\right]\left(\mathbf{X}_t^{\mathbf{z},(1)}\right)' \\
-\frac{1}{2}\mathbf{X}_t^{\mathbf{z},(1)}\operatorname{Cov}\left[\mathbf{z}_t,\mathbf{w}_{s,t}^{P} \mid \mathbf{Y}^{(1)}\right]\left(\mathbf{X}_{r,s,t}^{P,(1)}\right)' \\
-\frac{1}{2}\mathbf{X}_{r,s,t}^{P,(1)}\sum_{j=1}^{c}\operatorname{Cov}\left[\mathbf{w}_{s,t}^{P},\mathbf{w}_{j,t}^{P} \mid \mathbf{Y}^{(1)}\right]\left(\sum_{i=1:(i\neq r\ if\ j=s)}^{q}\alpha_{G,i,h}^{\langle m\rangle}\left(\mathbf{X}_{i,j,t}^{P,(1)}\right)'\right) \\
-\frac{1}{2}\sum_{j=1}^{c}\left(\sum_{i=1:(i\neq r\ if\ j=s)}^{q}\alpha_{G,i,k}^{\langle m\rangle}\mathbf{X}_{i,j,t}^{P,(1)}\right)\operatorname{Cov}\left[\mathbf{w}_{j,t}^{P},\mathbf{w}_{s,t}^{P} \mid \mathbf{Y}^{(1)}\right]\left(\mathbf{X}_{r,s,t}^{P,(1)}\right)' \\
-\frac{1}{2}\mathbf{X}_{r,s,t}^{P,(1)}\operatorname{Cov}\left[\left(\mathbf{w}_{s,t}^{P},\mathbf{w}_t^{B} \mid \mathbf{Y}^{(1)}\right)\right]\left(\sum_{i=1}^{q}\alpha_{B,i}^{\langle m\rangle}\left(\mathbf{X}_{i,t}^{B,(1)}\right)'\right) \\
-\frac{1}{2}\left(\sum_{i=1}^{q}\alpha_{B,i}^{\langle m\rangle}\mathbf{X}_{i,t}^{B,(1)}\right)\operatorname{Cov}\left(\mathbf{w}_t^{B},\mathbf{w}_{s,t}^{P} \mid \mathbf{Y}^{(1)}\right)\left(\mathbf{X}_{r,s,t}^{P,(1)}\right)' \\
-\frac{1}{2}\mathbf{X}_{r,s,t}^{P,(1)}\operatorname{Cov}\left(\mathbf{w}_{s,t}^{P},\tilde{\mathbf{w}}_t^{B} \mid \mathbf{Y}^{(1)}\right)\left(\sum_{i=1}^{q}\alpha_{BP,i}^{\langle m\rangle}\cdot\left(\mathbf{X}_{i,t}^{BP,(1)}\right)'\right) \\
-\frac{1}{2}\left(\sum_{i=1}^{q}\alpha_{BP,i}^{\langle m\rangle}\mathbf{X}_{i,t}^{BP,(1)}\right)\operatorname{Cov}\left(\tilde{\mathbf{w}}_t^{B},\mathbf{w}_{s,t}^{P} \mid \mathbf{Y}^{(1)}\right)\left(\mathbf{X}_{r,s,t}^{P,(1)}\right)'
\end{array}\right]}{\operatorname{tr}\left\{\sum_{t=1}^{T}\mathbf{X}_{r,s,t}^{P,(1)}\left[E\left(\mathbf{w}_{s,t}^{P,(1)} \mid \mathbf{Y}^{(1)}\right)E\left(\mathbf{w}_{s,t}^{P} \mid \mathbf{Y}^{(1)}\right)'+\operatorname{Var}\left(\mathbf{w}_{s,t}^{P} \mid \mathbf{Y}^{(1)}\right)\right]\left(\mathbf{X}_{r,s,t}^{P,(1)}\right)'\right\}}
$$

$$
\begin{aligned}
\boldsymbol{\beta}_P^{\langle m+1\rangle} &= \left[\sum_{t=1}^{T}\left(\mathbf{X}_t^{\boldsymbol{\beta},(1)}\right)'\mathbf{X}_t^{\boldsymbol{\beta},(1)}\right]^{-1}\left(\sum_{t=1}^{T}\left(\mathbf{X}_t^{\boldsymbol{\beta},(1)}\right)'\left[E\left(\mathbf{e}_t^{(1)} \mid \mathbf{Y}^{(1)}\right)+\mathbf{X}_t^{\boldsymbol{\beta},(1)}\boldsymbol{\beta}_P^{\langle m\rangle}\right]\right) \\
\left(\sigma_{P,r}^2\right)^{\langle m+1\rangle} &= \frac{1}{n_r T}\operatorname{tr}\sum_{t=1}^{T}\mathbf{O}_{i,t}\boldsymbol{\Omega}_t\mathbf{O}_{i,t} \\
\left(\sigma_{B,r}^2\right)^{\langle m+1\rangle} &= \frac{1}{mT}\operatorname{tr}\sum_{t=1}^{T}\mathbf{O}_{i+q,t}\boldsymbol{\Omega}_t\mathbf{O}_{i+q,t}
\end{aligned}
$$

with $r = 1,...,q$, $s = 1,...,c$ and $\mathbf{e}_t^{(1)} = \mathbf{y}_t^{(1)} - \boldsymbol{\mu}_t^{(1)}$. Finally, the updating formulas for the model parameters $\{\Psi_{\mathbf{z}_0}, \Psi_{\mathbf{z}}\}$ are the following

$$
\begin{aligned}
\boldsymbol{\nu}_0^{\langle m+1\rangle} &= \mathbf{z}_0^T \\
\mathbf{G}^{\langle m+1\rangle} &= \mathbf{S}_{10}\mathbf{S}_{00}^{-1} \\
\boldsymbol{\Sigma}_\eta^{\langle m+1\rangle} &= T^{-1}\left(\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{G}' - \mathbf{G}\mathbf{S}_{10}' + \mathbf{G}\mathbf{S}_{00}\mathbf{G}'\right)
\end{aligned}
$$

Starting from an initial value $\Psi^{\langle 0\rangle}$ for the model parameter set, the EM algorithm is iterated until convergence, that is, until $\left\|\Psi^{\langle m+1\rangle} - \Psi^{\langle m\rangle}\right\| / \left\|\Psi^{\langle m\rangle}\right\| < \delta$, where the norm $\|\cdot\|$ is evaluated with respect to the non-zero and unique elements of Ψ.

3.5 Software implementation

Model (1) is implemented by the D-STEM software available at `http://code.google.com/p/d-stem/`. As discussed in [2], the software is optimized for large datasets and it is based on parallel and distributed computing.

References

[1] Fassò A, Finazzi F (2011) Maximum likelihood estimation of the dynamic coregionalization model with heterotopic data. Environmetrics; 22: 735:748. ISSN 1099-095X

[2] Finazzi F, Fassò A (2012) D-STEM - A statistical software for multivariate space-time environmental data modeling. METMA VI conference. Guimaraes, 12-14 September 2012

[3] Finazzi F, Scott, M, Fassò A (2012) A model based framework for air quality indices and population risk evaluation. With an application to the analysis of Scottish air quality data. Journal of the Royal Statistical Society - Series C - DOI: 10.1111/rssc.12001