Appendices

Appendix 1: Proof of Equations (3-4).

Let

$$f = \left(\mathbf{x}_t - \mathbf{x}_{t-1}\right)' \left(\mathbf{x}_t - \mathbf{x}_{t-1}\right) + \boldsymbol{\lambda}'\left(\boldsymbol{\tau} - \hat{\boldsymbol{\Psi}}\mathbf{y}_t - \hat{\mathbf{B}}\mathbf{x}_t - \hat{\mathbf{C}}\mathbf{x}_{t-1} - \mathbf{a}_t\right), t \geq 1.$$

Then

$$\frac{\partial f}{\partial \mathbf{x}_t} = 2\mathbf{x}_t' - 2\mathbf{x}_{t-1}' - \boldsymbol{\lambda}'\hat{\mathbf{B}} = \mathbf{0}, \tag{A1}$$

and

$$\frac{\partial f}{\partial \boldsymbol{\lambda}} = \boldsymbol{\tau} - \hat{\boldsymbol{\Psi}}\mathbf{y}_t - \hat{\mathbf{B}}\mathbf{x}_t - \hat{\mathbf{C}}\mathbf{x}_{t-1} - \mathbf{a}_t = \mathbf{0},$$

which leads

$$\boldsymbol{\tau} - \hat{\boldsymbol{\Psi}}\mathbf{y}_t - \hat{\mathbf{B}}\mathbf{x}_{t-1} - \frac{1}{2}\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)\boldsymbol{\lambda} - \hat{\mathbf{C}}\mathbf{x}_{t-1} - \mathbf{a}_t = \mathbf{0}.$$

That is

$$\boldsymbol{\lambda} = 2\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\left(\boldsymbol{\tau} - \hat{\boldsymbol{\Psi}}\mathbf{y}_t - \left(\hat{\mathbf{B}} + \hat{\mathbf{C}}\right)\mathbf{x}_{t-1} - \mathbf{a}_t\right). \tag{A2}$$

Replacing Equation (A2) in Equation (A1) leads

$$\mathbf{x}_t = \left(\mathbf{I}_{m\times m} - \hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\left(\hat{\mathbf{B}} + \hat{\mathbf{C}}\right)\right)\mathbf{x}_{t-1} + \hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\left(\boldsymbol{\tau} - \hat{\boldsymbol{\Psi}}\mathbf{y}_t - \mathbf{a}_t\right).$$

Therefore,

$$\begin{aligned}\mathbf{x}_t - \mathbf{x}_{t-1} = &-\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\left(\hat{\boldsymbol{\Psi}} + \boldsymbol{\Omega}\right)\left(\mathbf{y}_t - \boldsymbol{\tau}\right) + \hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\hat{\boldsymbol{\Psi}}\left(\mathbf{y}_{t-1} - \boldsymbol{\tau}\right) \\ &+ \left(\mathbf{I}_{m\times m} - \hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\left(\hat{\mathbf{B}} + \hat{\mathbf{C}}\right)\right)\left(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}\right).\end{aligned}$$

Appendix 2: Proof of Lemma 1.

Equations (2) and (3) lead the following equations:

$$\mathbf{a}_{t-1} - \mathbf{a}_{t-2} = \boldsymbol{\Omega}\left(\mathbf{y}_{t-1} - \hat{\boldsymbol{\Psi}}\mathbf{y}_{t-2} - \hat{\mathbf{B}}\mathbf{x}_{t-2} - \hat{\mathbf{C}}\mathbf{x}_{t-3} - \mathbf{a}_{t-2}\right) = \boldsymbol{\Omega}\left(\mathbf{y}_{t-1} - \boldsymbol{\tau}\right),$$

for $t > 1$, and

$$
\mathbf{x}_{t-1}-\mathbf{x}_{t-2}=\begin{cases}
-\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\hat{\mathbf{C}}\mathbf{x}_0-\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\left(\hat{\boldsymbol{\Psi}}+\boldsymbol{\Omega}\right)\left(\mathbf{y}_1-\boldsymbol{\tau}\right)-\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\hat{\boldsymbol{\Psi}}\boldsymbol{\tau} & t=1,\\
-\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\hat{\mathbf{C}}\left(\mathbf{x}_{t-2}-\mathbf{x}_{t-3}\right)-\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\left(\hat{\boldsymbol{\Psi}}+\boldsymbol{\Omega}\right)\left(\mathbf{y}_{t-1}-\boldsymbol{\tau}\right) & \\
\quad+\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\hat{\boldsymbol{\Psi}}\left(\mathbf{y}_{t-2}-\boldsymbol{\tau}\right) & t>1.
\end{cases}
$$

Expressing Equation (1) as a difference equation and using the above results provide the iterative formula for $\mathbf{y}_t-\boldsymbol{\tau}$ as follows:

$$
\begin{aligned}
\mathbf{y}_t-\boldsymbol{\tau}=&\left(\mathbf{I}_{n\times n}+\boldsymbol{\Psi}-\mathbf{B}\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\left(\hat{\boldsymbol{\Psi}}+\boldsymbol{\Omega}\right)\right)\left(\mathbf{y}_{t-1}-\boldsymbol{\tau}\right)\\
&+\left(-\boldsymbol{\Psi}+\mathbf{B}\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\hat{\boldsymbol{\Psi}}\right)\left(\mathbf{y}_{t-2}-\boldsymbol{\tau}\right)\\
&+\left(\mathbf{C}-\mathbf{B}\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{-1}\hat{\mathbf{C}}\right)\left(\mathbf{x}_{t-2}-\mathbf{x}_{t-3}\right)\\
&+\left(\boldsymbol{\eta}_t-\boldsymbol{\eta}_{t-1}\right).
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\begin{bmatrix}\mathbf{y}_t-\boldsymbol{\tau}\\ \mathbf{y}_{t-1}-\boldsymbol{\tau}\\ \mathbf{x}_{t-1}-\mathbf{x}_{t-2}\end{bmatrix}&=\mathbf{M}\begin{bmatrix}\mathbf{y}_{t-1}-\boldsymbol{\tau}\\ \mathbf{y}_{t-2}-\boldsymbol{\tau}\\ \mathbf{x}_{t-2}-\mathbf{x}_{t-3}\end{bmatrix}+\begin{bmatrix}\left(1-\mathcal{B}\right)\boldsymbol{\eta}_t\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix}\\
&=\qquad\vdots\\
&=\mathbf{M}^{t-1}\begin{bmatrix}\boldsymbol{\alpha}+\mathbf{B}\mathbf{x}_0-\boldsymbol{\tau}\\ -\boldsymbol{\tau}\\ \mathbf{x}_0\end{bmatrix}+\sum_{j=0}^{t-1}\left(\mathbf{M}\right)^j\begin{bmatrix}\left(1-\mathcal{B}\right)\boldsymbol{\eta}_{t-j}\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix}.
\end{aligned}
$$

Hence,

$$
\mathbf{y}_t-\boldsymbol{\tau}=\boldsymbol{\gamma}_{t-1}+\boldsymbol{\xi}_t,
$$

where,

$$
\boldsymbol{\gamma}_{t-1}=\left[\mathbf{I}\quad\mathbf{0}\quad\mathbf{0}\right]\mathbf{M}^{t-1}\begin{bmatrix}\boldsymbol{\alpha}+\mathbf{B}\mathbf{x}_0-\boldsymbol{\tau}\\ -\boldsymbol{\tau}\\ \mathbf{x}_0\end{bmatrix}=\mathbf{E}_{t-1}\boldsymbol{\gamma}_0+\mathbf{F}_{t-1}\left(-\boldsymbol{\tau}\right)+\mathbf{G}_{t-1}\mathbf{x}_0,
$$

and

$$\boldsymbol{\xi}_t = \begin{bmatrix}\mathbf{I} & \mathbf{0} & \mathbf{0}\end{bmatrix}\sum_{j=0}^{t-1}(\mathbf{M})^j \begin{bmatrix}(1-\mathcal{B})\boldsymbol{\eta}_{t-j} \\ \mathbf{0} \\ \mathbf{0}\end{bmatrix} = (\boldsymbol{\eta}_t - \boldsymbol{\eta}_{t-1}) + \sum_{i=1}^{t-1}\mathbf{E}_i(\boldsymbol{\eta}_{t-i} - \boldsymbol{\eta}_{t-i-1}).$$

Note that $\boldsymbol{\xi}_t = 0, \text{for } t \le 0,$ hence $\boldsymbol{\xi}_t$ can be further expressed as

$$\begin{aligned}\boldsymbol{\xi}_t &= \begin{bmatrix}\mathbf{I}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m}\end{bmatrix}(\mathbf{I} - \mathcal{B}\mathbf{M})^{-1}\begin{bmatrix}\mathbf{I}_{n\times n} \\ \mathbf{0}_{n\times n} \\ \mathbf{0}_{m\times n}\end{bmatrix}(1-\mathcal{B})\boldsymbol{\eta}_t \\ &= [(\mathbf{I}_{(2n+m)\times(2n+m)} - \mathcal{B}\mathbf{M})^{-1}]_{11}(1-\mathcal{B})\boldsymbol{\eta}_t.\end{aligned} \tag{A3}$$

Appendix 3: Proof of Theorem 1.

If $\rho(\mathbf{M}) < 1,$ then

$$\lim_{t\to\infty}\mathbf{M}^t = \mathbf{0}_{(2n+m)\times(2n+m)}.$$

Therefore, it implies

$$\lim_{t\to\infty}\mathbf{E}_t = \lim_{t\to\infty}\mathbf{F}_t = \mathbf{0}_{n\times n}, \lim_{t\to\infty}\mathbf{G}_t = \mathbf{0}_{n\times m}.$$

Hence, in this case, Equation (12) leads

$$\lim_{t\to\infty}\mathrm{E}(\mathbf{y}_t) = \lim_{t\to\infty}\left\{\boldsymbol{\tau} + \mathbf{E}_{t-1}\boldsymbol{\gamma}_0 + \mathbf{F}_{t-1}(-\boldsymbol{\tau}) + \mathbf{G}_{t-1}\mathbf{x}_0\right\} = \boldsymbol{\tau}.$$

Now, since $\max_{1\le i\le n} d_i \le 1$ guarantees $(1-\mathcal{B})\boldsymbol{\eta}_t$ is stationary, the stability conditions depend on $\{E_i\}_{i=0}^{t-1}$. To show the covariance matrix of $\boldsymbol{\xi}_t$ is bounded, we first denote $e_{t,ij}$ as the entry of $\mathbf{E}_t$. According to Horn and Johnson (1990), since $\lim_{t\to\infty}\mathbf{E}_t = \mathbf{0}_{n\times n}$, for arbitrary $\varepsilon > 0,$ there exists a positive integer N_ε such that

$$\sqrt{\sum_{i,j} e_{t,ij}^2} < \varepsilon, \ \forall t > N_\varepsilon,$$

which leads the Euclidean norm $\left(\|\cdot\|_2\right)$ of $\mathbf{E}_t$ to be less than ε for large t.

Define $s_t = \sum_{i=0}^{t}\|\mathbf{E}_i\|_2$. Then

$$s_{n_1} - s_{n_2} = \sum_{i=n_2+1}^{n_1} \|\mathbf{E}_i\|_2 < \varepsilon,\ \forall n_1 > n_2 > N_\varepsilon .$$

This demonstrates that s_t is a Cauchy sequence. Therefore, the filter relating $(1-\mathcal{B})\boldsymbol{\eta}_t$ to $\boldsymbol{\xi}_t$ is stable. According to Reinsel (2003), the process outputs controlled by the dpMEWMA controller satisfy Equations (8) and (9).

Appendix 4: Proof of Equation (17).

From Equation (10), we have

$$\mathbf{I} - \mathcal{B}\mathbf{M} = \begin{bmatrix} \mathbf{I}_{n\times n} - \mathcal{B}\mathbf{M}_{11} & -\mathcal{B}\mathbf{M}_{12} & -\mathcal{B}\mathbf{M}_{13} \\ -\mathcal{B}\mathbf{I}_{n\times n} & \mathbf{I}_{n\times n} & \mathbf{0} \\ -\mathcal{B}\mathbf{M}_{31} & -\mathcal{B}\mathbf{M}_{32} & \mathbf{I}_{m\times m} - \mathcal{B}\mathbf{M}_{33} \end{bmatrix}.$$

By the inverse of block matrix (Graybill, page 19), we have

$$[(\mathbf{I} - \mathcal{B}\mathbf{M})^{-1}]_{11} = \left\{ \mathbf{I}_{n\times n} - \mathcal{B}\mathbf{M}_{11} - \mathcal{B}^2 \begin{bmatrix} \mathbf{M}_{12} & \mathbf{M}_{13} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{n\times n} & \mathbf{0} \\ -\mathcal{B}\mathbf{M}_{32} & \mathbf{I}_{m\times m} - \mathcal{B}\mathbf{M}_{33} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I}_{n\times n} \\ \mathbf{M}_{31} \end{bmatrix} \right\}^{-1},$$

where

$$\begin{bmatrix} \mathbf{I}_{n\times n} & \mathbf{0} \\ -\mathcal{B}\mathbf{M}_{32} & \mathbf{I}_{m\times m} - \mathcal{B}\mathbf{M}_{33} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I}_{n\times n} & \mathbf{0} \\ -\left(\mathbf{I}_{m\times m} - \mathcal{B}\mathbf{M}_{33}\right)^{-1}\left(-\mathcal{B}\mathbf{M}_{32}\right) & \left(\mathbf{I}_{m\times m} - \mathcal{B}\mathbf{M}_{33}\right)^{-1} \end{bmatrix}.$$

Under $\hat{\mathbf{C}} = k\hat{\mathbf{B}},$ we have

$$\left(\mathbf{I}_{m\times m} - \mathcal{B}\mathbf{M}_{33}\right)^{-1} = \mathbf{I}_{m\times m} - \frac{k\mathcal{B}}{1+k\mathcal{B}} \hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{\mathbf{-1}} \hat{\mathbf{B}},$$

and

$$-\left(\mathbf{I}_{m\times m} - \mathcal{B}\mathbf{M}_{33}\right)^{-1}\left(-\mathcal{B}\mathbf{M}_{32}\right) = \frac{\mathcal{B}}{1+k\mathcal{B}} \hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{\mathbf{-1}} \hat{\boldsymbol{\Psi}}.$$

Hence,

$$\begin{aligned}
&\left(1+k\mathcal{B}\right)\left\{ \mathbf{I}_{n\times n} - \mathcal{B}\mathbf{M}_{11} - \mathcal{B}^2 \begin{bmatrix} \mathbf{M}_{12} & \mathbf{M}_{13} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{n\times n} & \mathbf{0} \\ -\mathcal{B}\mathbf{M}_{32} & \mathbf{I}_{m\times m} - \mathcal{B}\mathbf{M}_{33} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I}_{n\times n} \\ \mathbf{M}_{31} \end{bmatrix} \right\} \\
&= \left(1+k\mathcal{B}\right)\mathbf{I}_{n\times n} - \left(1+k\mathcal{B}\right)\mathcal{B}\mathbf{M}_{11} - \mathcal{B}^2\left(1+k\mathcal{B}\right)\left[\mathbf{M}_{12} + \mathbf{M}_{13}\mathbf{M}_{31}\right] \\
&\qquad - \mathcal{B}^3\left[\mathbf{M}_{13}\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{\mathbf{-1}} \hat{\boldsymbol{\Psi}} - k\mathbf{M}_{13}\hat{\mathbf{B}}'\left(\hat{\mathbf{B}}\hat{\mathbf{B}}'\right)^{\mathbf{-1}} \hat{\mathbf{B}}\mathbf{M}_{31}\right].
\end{aligned}$$

From (A3) in Appendix 2, we have

$$\left(\mathbf{I}_{n\times n}-\mathbf{\Phi}_1^*\mathcal{B}-\mathbf{\Phi}_2^*\mathcal{B}^2-\mathbf{\Phi}_3^*\mathcal{B}^3\right)\boldsymbol{\xi}_t=(1+k\mathcal{B})(1-\mathcal{B})\,\boldsymbol{\eta}_t.$$

Appendix 5: The expression of $\sum_{t=1}^{\infty}\left(\boldsymbol{\gamma}_{t-1}'\boldsymbol{\gamma}_{t-1}\right)$.

From Equation (13), we can re-express $\boldsymbol{\gamma}_{t-1}$ and $\sum_{t=1}^{\infty}\left(\boldsymbol{\gamma}_{t-1}'\boldsymbol{\gamma}_{t-1}\right)$ as

$$\boldsymbol{\gamma}_{t-1}=\left[\mathbf{I}_{n\times n}\ \ \mathbf{0}_{n\times n}\ \ \mathbf{0}_{n\times m}\right]\mathbf{M}^{t-1}\begin{bmatrix}\boldsymbol{\gamma}_0\\ -\boldsymbol{\tau}\\ \mathbf{x}_0\end{bmatrix},$$

and

$$\sum_{t=1}^{\infty}\left(\boldsymbol{\gamma}_{t-1}'\boldsymbol{\gamma}_{t-1}\right)=\left[\boldsymbol{\gamma}_0'\ \ -\boldsymbol{\tau}'\ \ \mathbf{x}_0'\right]\left(\sum_{t=1}^{\infty}\left(\mathbf{M}'\right)^{t-1}\begin{bmatrix}\mathbf{I}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m}\\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m}\\ \mathbf{0}_{m\times n} & \mathbf{0}_{m\times n} & \mathbf{0}_{m\times m}\end{bmatrix}\mathbf{M}^{t-1}\right)\begin{bmatrix}\boldsymbol{\gamma}_0\\ -\boldsymbol{\tau}\\ \mathbf{x}_0\end{bmatrix},$$

respectively. To further simplify the above equation, let

$$\mathbf{\Sigma}_{\xi}=\sum_{t=1}^{\infty}\left(\mathbf{M}'\right)^{t}\begin{bmatrix}\mathbf{I}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m}\\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m}\\ \mathbf{0}_{m\times n} & \mathbf{0}_{m\times n} & \mathbf{0}_{m\times m}\end{bmatrix}\mathbf{M}^{t},$$

then

$$\sum_{t=1}^{\infty}\left(\boldsymbol{\gamma}_{t-1}'\boldsymbol{\gamma}_{t-1}\right)=\left[\boldsymbol{\gamma}_0'\ \ -\boldsymbol{\tau}'\ \ \mathbf{x}_0'\right]\mathbf{\Sigma}_{\xi}\begin{bmatrix}\boldsymbol{\gamma}_0\\ -\boldsymbol{\tau}\\ \mathbf{x}_0\end{bmatrix}.$$

Note that $\mathbf{\Sigma}_{\xi}$ can be viewed as the covariance matrix of the time series ς_t, where

$$\varsigma_t=\mathbf{M}'\varsigma_{t-1}+\boldsymbol{\varepsilon}_t,$$

and $\boldsymbol{\varepsilon}_t$ is a vector white noise series with covariance matrix

$$\mathbf{\Sigma}_{\varepsilon}=\begin{bmatrix}\mathbf{I}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m}\\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m}\\ \mathbf{0}_{m\times n} & \mathbf{0}_{m\times n} & \mathbf{0}_{m\times m}\end{bmatrix}.$$

Hence, we have

$$\boldsymbol{\Sigma}_{\xi} = \mathbf{M}'\boldsymbol{\Sigma}_{\xi}\mathbf{M} + \begin{bmatrix} \mathbf{I}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m} \\ \mathbf{0}_{m\times n} & \mathbf{0}_{m\times n} & \mathbf{0}_{m\times m} \end{bmatrix}.$$

Now, by use of procedure in Reinsel (2003, pp 29-30),

$$\boldsymbol{\Sigma}_{\xi} = \mathrm{mat}_{(2n+m)\times(2n+m)}\left(\left(\mathbf{I}_{(2n+m)^2\times(2n+m)^2} - \mathbf{M}'\otimes\mathbf{M}'\right)^{-1}\mathrm{vec}\left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\right)\right),$$

where $\mathrm{vec}(\cdot)$ is the operator stacking column vectors of a $p\times q$ matrix as a $pq\times 1$ vector; $\mathrm{mat}_{p\times q}(\cdot)$ denotes the operator inversely stacking a $pq\times 1$ column vector as a $p\times q$ matrix; $\otimes$ denotes the Kronecker product, that is, $\mathbf{A}\otimes\mathbf{C} = \left(a_{ij}\mathbf{C}\right)_{mp\times nq}$ with $\mathbf{A} = \left(a_{ij}\right)_{m\times n}$ and $\mathbf{C} = \left(c_{ij}\right)_{p\times q}$.

Therefore,

$$\sum_{t=1}^{\infty}\left(\boldsymbol{\gamma}'_{t-1}\boldsymbol{\gamma}_{t-1}\right) = \left[\boldsymbol{\gamma}'_0 \ \ -\boldsymbol{\tau}' \ \ \mathbf{x}'_0\right]\mathrm{mat}_{(2n+m)\times(2n+m)}\left(\left(\mathbf{I}_{(2n+m)^2\times(2n+m)^2} - \mathbf{M}'\otimes\mathbf{M}'\right)^{-1}\mathrm{vec}\left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\right)\right)\begin{bmatrix}\boldsymbol{\gamma}_0 \\ -\boldsymbol{\tau} \\ \mathbf{x}_0\end{bmatrix}.$$

Appendix 6: The expression of υ.

From Equation (17), if $\boldsymbol{\eta}_t$ is a vector $\mathrm{IMA}(1,1)$ process, then $\boldsymbol{\xi}_t$ will follow a stationary vector ARMA(3, 2) process, and the corresponding Yule-Walker equations are

$$\begin{aligned}
\boldsymbol{\Gamma}_0 &= \mathrm{Cov}\left(\boldsymbol{\xi}_t, \boldsymbol{\xi}_t\right) = \boldsymbol{\Gamma}'_1(\boldsymbol{\Phi}^*_1)' + \boldsymbol{\Gamma}'_2(\boldsymbol{\Phi}^*_2)' + \boldsymbol{\Gamma}'_3(\boldsymbol{\Phi}^*_3)' + \mathbf{H}_0, \\
\boldsymbol{\Gamma}_1 &= \mathrm{Cov}\left(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t+1}\right) = \boldsymbol{\Gamma}_0(\boldsymbol{\Phi}^*_1)' + \boldsymbol{\Gamma}'_1(\boldsymbol{\Phi}^*_2)' + \boldsymbol{\Gamma}'_2(\boldsymbol{\Phi}^*_3)' + \mathbf{H}_1, \\
\boldsymbol{\Gamma}_2 &= \mathrm{Cov}\left(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t+2}\right) = \boldsymbol{\Gamma}_1(\boldsymbol{\Phi}^*_1)' + \boldsymbol{\Gamma}_0(\boldsymbol{\Phi}^*_2)' + \boldsymbol{\Gamma}'_1(\boldsymbol{\Phi}^*_3)' + \mathbf{H}_2, \\
\boldsymbol{\Gamma}_3 &= \mathrm{Cov}\left(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t+3}\right) = \boldsymbol{\Gamma}_2(\boldsymbol{\Phi}^*_1)' + \boldsymbol{\Gamma}_1(\boldsymbol{\Phi}^*_2)' + \boldsymbol{\Gamma}_0(\boldsymbol{\Phi}^*_3)'.
\end{aligned}$$

where

$$\mathbf{H}_0 = \mathbf{\Sigma} - \left(\mathbf{\Phi}_1^* - \mathbf{\Theta}_1\right)\mathbf{\Sigma\Theta}_1' - \left(\mathbf{\Phi}_1^*\left(\mathbf{\Phi}_1^* - \mathbf{\Theta}_1\right) + \left(\mathbf{\Phi}_2^* - \mathbf{\Theta}_2\right)\right)\mathbf{\Sigma\Theta}_2' \text{ ;}$$

$$\mathbf{H}_1 = -\mathbf{\Sigma\Theta}_1' - \left(\mathbf{\Phi}_1^* - \mathbf{\Theta}_1\right)\mathbf{\Sigma\Theta}_2' \text{ , and } \mathbf{H}_2 = -\mathbf{\Sigma\Theta}_2',$$

$$\mathbf{\Theta}_1 = \mathbf{\Theta} - k\mathbf{I}_{n\times n}, \text{ and } \mathbf{\Theta}_2 = k\mathbf{\Theta}.$$

Again, by use of procedure in Reinsel (2003, pp 29-30 and 59-60), we obtain

$$\upsilon = \text{trace}\left(\mathbf{\Gamma}_0\right) = \text{trace}\left(\text{mat}_{n\times n}\left(\mathbf{P}_1^{-1}\left(\mathbf{P}_3\mathbf{h}_2 + \mathbf{P}_2\mathbf{h}_1 + \mathbf{h}_0\right)\right)\right),$$

where $\mathbf{h}_i = \text{vec}\left(\mathbf{H}_i\right), i = 1, 2, 3;$ $\mathbf{P}_1 = \mathbf{Q}_1 - \mathbf{Q}_2\mathbf{Q}_3\mathbf{Q}_4;$ $\mathbf{P}_2 = \mathbf{Q}_2\mathbf{Q}_3$, and

$$\mathbf{P}_3 = \mathbf{P}_2\left(\mathbf{\Phi}_3^* \otimes \mathbf{I}_{n\times n}\right)\mathbf{S}_{(n,n)} + \left(\mathbf{\Phi}_2^* \otimes \mathbf{I}_{n\times n} + \mathbf{\Phi}_3^* \otimes \mathbf{\Phi}_1^*\right)\mathbf{S}_{(n,n)},$$

with

$$\mathbf{Q}_1 = \mathbf{I}_{n^2\times n^2} - \left(\mathbf{\Phi}_3^* \otimes \mathbf{\Phi}_3^*\right) - \left(\mathbf{\Phi}_2^* \otimes \mathbf{I}_{n\times n} + \mathbf{\Phi}_3^* \otimes \mathbf{\Phi}_1^*\right)\mathbf{S}_{(n,n)}\left(\mathbf{\Phi}_2^* \otimes \mathbf{I}_{n\times n}\right);$$

$$\begin{aligned}\mathbf{Q}_2 = &\left(\mathbf{\Phi}_1^* \otimes \mathbf{I}_{n\times n} + \mathbf{\Phi}_3^* \otimes \mathbf{\Phi}_2^*\right)\mathbf{S}_{(n,n)} \\ &+ \left(\mathbf{\Phi}_2^* \otimes \mathbf{I}_{n\times n} + \mathbf{\Phi}_3^* \otimes \mathbf{\Phi}_1^*\right)\mathbf{S}_{(n,n)}\left(\mathbf{\Phi}_1^* \otimes \mathbf{I}_{n\times n} + \left(\mathbf{\Phi}_3^* \otimes \mathbf{I}_{n\times n}\right)\mathbf{S}_{(n,n)}\right);\end{aligned}$$

$$\mathbf{Q}_3 = \left(\mathbf{I}_{n^2\times n^2} - \left(\mathbf{\Phi}_2^* \otimes \mathbf{I}_{n\times n}\right)\mathbf{S}_{(n,n)} - \left(\mathbf{\Phi}_3^* \otimes \mathbf{I}_{n\times n}\right)\mathbf{S}_{(n,n)}\left(\mathbf{\Phi}_1^* \otimes \mathbf{I}_{n\times n} + \left(\mathbf{\Phi}_3^* \otimes \mathbf{I}_{n\times n}\right)\mathbf{S}_{(n,n)}\right)\right)^{-1};$$

$$\mathbf{Q}_4 = \mathbf{\Phi}_1^* \otimes \mathbf{I}_{n\times n} + \left(\mathbf{\Phi}_3^* \otimes \mathbf{I}_{n\times n}\right)\mathbf{S}_{(n,n)}\left(\mathbf{\Phi}_2^* \otimes \mathbf{I}_{n\times n}\right),$$

and $\mathbf{S}_{(n,n)} = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\mathbf{E}_{i_1i_2} \otimes \mathbf{E}_{i_1i_2}'$ is the permuted identity matrix, with $\mathbf{E}_{i_1i_2} = \left(e_{ij}\right)_{n\times n}$ and

$$e_{ij} = \begin{cases} 1 & \text{if } i = i_1 \text{ and } j = i_2, \\ 0 & \text{otherwise.} \end{cases}$$

Appendix 7: Proof of Equation (21).

As $\mathbf{B} = \hat{\mathbf{B}}$, $\mathbf{C} = \hat{\mathbf{C}}$, and $\mathbf{\Psi} = \hat{\mathbf{\Psi}}$,

$$\mathbf{M} = \begin{bmatrix} \mathbf{I}_{n\times n} - \mathbf{\Omega} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m} \\ \mathbf{I}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m} \\ -\mathbf{B}'\left(\mathbf{BB}'\right)^{-1}\left(\mathbf{\Psi} + \mathbf{\Omega}\right) & \mathbf{B}'\left(\mathbf{BB}'\right)^{-1}\mathbf{\Psi} & -\mathbf{B}'\left(\mathbf{BB}'\right)^{-1}\mathbf{C} \end{bmatrix}.$$

Hence,

$$\boldsymbol{\gamma}_{t-1}\boldsymbol{\gamma}'_{t-1} = \mathbf{E}_{t-1}\boldsymbol{\gamma}_0\boldsymbol{\gamma}'_0\mathbf{E}'_{t-1},$$

which implies

$$\begin{aligned}
\sum_{t=1}^{N}\operatorname{trace}\left(\boldsymbol{\gamma}_{t-1}\boldsymbol{\gamma}'_{t-1}\right) &= \sum_{t=1}^{N}\operatorname{trace}\left(\mathbf{E}_{t-1}\boldsymbol{\gamma}_0\boldsymbol{\gamma}'_0\mathbf{E}'_{t-1}\right)\\
&= \sum_{t=1}^{N}\boldsymbol{\gamma}'_0\mathbf{E}'_{t-1}\mathbf{E}_{t-1}\boldsymbol{\gamma}_0\\
&= \sum_{t=1}^{N}\boldsymbol{\gamma}'_0\begin{bmatrix}(1-\omega_1)^{2t-2} & 0 & \cdots & 0\\ 0 & \ddots & & \vdots\\ \vdots & & \ddots & 0\\ 0 & \cdots & 0 & (1-\omega_n)^{2t-2}\end{bmatrix}\boldsymbol{\gamma}_0 \qquad \text{(A4)}\\
&= \sum_{t=1}^{N}\sum_{i=1}^{n}(1-\omega_i)^{2t-2}\,\gamma_{i0}^2\\
&= \sum_{i=1}^{n}\frac{1-(1-\omega_i)^{2N}}{\omega_i(2-\omega_i)}\gamma_{i0}^2.
\end{aligned}$$

In addition,

$$\left(\mathbf{I}_{n\times n}-\boldsymbol{\Phi}_1\mathcal{B}\right)\boldsymbol{\xi}_t = \left(\mathbf{I}_{n\times n}-\boldsymbol{\Theta}\mathcal{B}\right)\boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\Phi}_1 = \mathbf{I}_{n\times n}-\boldsymbol{\Omega} = \operatorname{diag}\left(1-\omega_1,\ldots,1-\omega_n\right)$, and then the corresponding Yule-Walker equations are

$$\begin{aligned}
\boldsymbol{\Gamma}_0 &= \operatorname{Cov}\left(\boldsymbol{\xi}_t,\boldsymbol{\xi}_t\right) = \boldsymbol{\Gamma}'_1\boldsymbol{\Phi}'_1+\mathbf{H}_0,\\
\boldsymbol{\Gamma}_1 &= \operatorname{Cov}\left(\boldsymbol{\xi}_t,\boldsymbol{\xi}_{t+1}\right) = \boldsymbol{\Gamma}_0\boldsymbol{\Phi}'_1+\mathbf{H}_1,
\end{aligned}$$

where

$$\mathbf{H}_0 = \boldsymbol{\Sigma}-\left(\boldsymbol{\Phi}_1-\boldsymbol{\Theta}\right)\boldsymbol{\Sigma}\boldsymbol{\Theta}' \text{ and } \mathbf{H}_1 = -\boldsymbol{\Sigma}\boldsymbol{\Theta}'.$$

Setting $\mathbf{g}_i = \operatorname{vec}\left(\boldsymbol{\Gamma}_i\right)$ and $\mathbf{h}_i = \operatorname{vec}\left(\mathbf{H}_i\right)$ leads

$$\mathbf{g}_0 = \left(\mathbf{I}_{n\times n}-\boldsymbol{\Phi}_1\otimes\boldsymbol{\Phi}_1\right)^{-1}\mathbf{h}_0^*,$$

where $\mathbf{h}_0^* = \operatorname{vec}\left(\mathbf{H}'_1\boldsymbol{\Phi}'_1+\mathbf{H}_0\right)$.

Let $\mathbf{g}_0 = \left(g_1^{(0)},\ldots,g_{n^2}^{(0)}\right)$. Since

$$\left(\mathbf{I}_{n^2\times n^2}-\boldsymbol{\Phi}\otimes\boldsymbol{\Phi}\right)^{-1}=\begin{bmatrix} \frac{1}{1-(1-\omega_1)(1-\omega_1)} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \frac{1}{1-(1-\omega_n)(1-\omega_n)} \end{bmatrix},$$

and

$$\mathbf{H}_1'\boldsymbol{\Phi}_1'+\mathbf{H}_0=-\boldsymbol{\Theta}\boldsymbol{\Sigma}\boldsymbol{\Phi}_1'+\left(\boldsymbol{\Sigma}-\left(\boldsymbol{\Phi}_1-\boldsymbol{\Theta}\right)\boldsymbol{\Sigma}\boldsymbol{\Theta}'\right),$$

we can obtain

$$g_{n(i-1)+i}^{(0)}=\frac{\sigma_{ii}-2\left(1-\omega_i\right)\boldsymbol{\theta}_i'\boldsymbol{\sigma}_i+\boldsymbol{\theta}_i'\boldsymbol{\Sigma}\boldsymbol{\theta}_i}{1-\left(1-\omega_i\right)^2}, \tag{A5}$$

where $\boldsymbol{\sigma}_i=\left(\sigma_{i1},\ldots\sigma_{in}\right)'$ and $\boldsymbol{\theta}_i=\left(\theta_{i1},\ldots\theta_{in}\right)'$. Therefore, we can obtain Equation (21) by Equations (A4) and (A5).

Appendix 8: The process I-O Model used in Section 6

For the CMP dataset, according to Fan et al. (2002), a dynamic process is a better and more realistic model to explain the complicated process. However, the results are addressed only under the case of the SISO process. In the following, adopt this dataset, we obtain a prediction model as follows:

$$\hat{\mathbf{y}}_t=\begin{bmatrix}-0.58 & -0.02\\ 0.06 & -0.61\end{bmatrix}\mathbf{y}_{t-1}+\begin{bmatrix}449.5\\2513.5\end{bmatrix}+\begin{bmatrix}31.13 & 107.59 & 34.72 & 57.97\\ 130.64 & -31.65 & 172.58 & 19.80\end{bmatrix}\mathbf{x}_{t-1}$$
$$+\begin{bmatrix}10.53 & 41.85 & 12.95 & 18.62\\ 83.35 & -26.17 & 97.90 & 10.42\end{bmatrix}\mathbf{x}_{t-2}.$$

By using Akaike information criterion (AIC), we have AIC(dynamic)=11.19, AIC(static linear model, Tseng et al., 2002)=12.72, AIC(quadratic without interaction term, Castillo and Yeh, 1988)=13.65. Furthermore, the Ljung-Box Portmanteau statistics for multivariate time series is $Q = 84.69$ when the lag is 20, and the corresponding *p*-value is 0.34. The test supports that the residuals are white noise. Hence, it demonstrates that the dynamic model together with IMA (1, 1) disturbance is more appropriate for the CMP dataset.

Appendix 9: The derivation of the ST controller.

Considering $\mathbf{\Theta} = \mathbf{0}_{n\times n}$, Equation (1) with vector $\mathrm{IMA}(1,1)$ disturbance can be reduced as follows:

$$\mathbf{y}_t = \mathbf{\Psi}\mathbf{y}_{t-1} + \boldsymbol{\alpha} + \mathbf{B}\mathbf{x}_{t-1} + \mathbf{C}\mathbf{x}_{t-2} + (1-B)^{-1}\boldsymbol{\varepsilon}_t,$$

which leads

$$\mathbf{y}_t - \mathbf{y}_{t-1} = \mathbf{\Psi}(\mathbf{y}_{t-1} - \mathbf{y}_{t-2}) + \mathbf{B}(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}) + \mathbf{C}(\mathbf{x}_{t-2} - \mathbf{x}_{t-3}) + \boldsymbol{\varepsilon}_t.$$

In this case, the MMSE controller can be written as follows (Reinsel, 2003):

$$\mathbf{x}_t = \mathbf{x}_{t-1} + (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\left(\boldsymbol{\tau} - \mathbf{y}_t - \mathbf{\Psi}(\mathbf{y}_t - \mathbf{y}_{t-1}) - \mathbf{C}(\mathbf{x}_{t-1} - \mathbf{x}_{t-2})\right), \tag{A6}$$

where $(\mathbf{B}'\mathbf{B})^{-1}$ denotes the generalized inverse of $\mathbf{B}'\mathbf{B}$.

Let the estimated matrix of model parameters be

$$\mathbf{\Xi}_t = \left[\hat{\mathbf{\Psi}}_t \,|\, \hat{\mathbf{B}}_t \,|\, \hat{\mathbf{C}}_t\right]$$

and

$$\boldsymbol{\chi}_t = \left[(\mathbf{y}_t - \mathbf{y}_{t-1})' \,|\, (\mathbf{x}_{t-1} - \mathbf{x}_{t-2})' \,|\, (\mathbf{x}_{t-2} - \mathbf{x}_{t-3})'\right]'.$$

In addition, for $t = 1, 2, \ldots,$ set

$$\mathbf{K}_t = \frac{\mathbf{P}_{t-1}\boldsymbol{\chi}_t}{\lambda + \boldsymbol{\chi}_t'\mathbf{P}_{t-1}\boldsymbol{\chi}_t},$$

$$\mathbf{\Xi}_t^{(i)} = \mathbf{\Xi}_{t-1}^{(i)} + \mathbf{K}_t'\left(y_{i,t} - \tau_i - \mathbf{\Xi}_{t-1}^{(i)}\boldsymbol{\chi}_t\right),$$

and

$$\mathbf{P}_t = \left(\mathbf{I}_{(n+2m)\times(n+2m)} - \mathbf{K}_t \boldsymbol{\chi}_t{}'\right)\mathbf{P}_{t-1} + \frac{\mathbf{I}_{(n+2m)\times(n+2m)}\mathbf{K}_t{}'\mathbf{P}_{t-1}\boldsymbol{\chi}_t}{n+2m},$$

where $\boldsymbol{\Xi}_t^{(i)}$ is the i-th row of $\boldsymbol{\Xi}_t$ and the initial setting for $\mathbf{P}_0$ is very close to an identity matrix in practical applications, that is, $\mathbf{P}_0 \approx \mathbf{I}_{(n+2m)\times(n+2m)}$ (Del Castillo and Yeh, 1998). Then modifying Equation (A6) provides the ST controller as follows.

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \left(\hat{\mathbf{B}}_t{}'\hat{\mathbf{B}}_t\right)^{-1}\hat{\mathbf{B}}_t{}'\left(\boldsymbol{\tau} - \mathbf{y}_t - \hat{\boldsymbol{\Psi}}_t\left(\mathbf{y}_t - \mathbf{y}_{t-1}\right) - \hat{\mathbf{C}}_t\left(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}\right)\right).$$