Bennedsen, Hillebrand, and Koopman
A Multivariate Dynamic Statistical Model of the Global Carbon Budget 1959—2020

A note on the implementation of the linear state space model

March 7, 2025

The linear state space model (with constant $\beta_{1,2,5}$) is implemented in R using the KFAS package (Helske, 2019). The KFAS package implements the Gaussian linear state space model

$$\begin{aligned} y_t &= Z_t\alpha_t + \varepsilon_t, \\ \alpha_{t+1} &= T_t\alpha_t + R_t\eta_t, \end{aligned} \tag{1}$$

where $\varepsilon_t \sim \mathsf{N}(0, H_t)$, $\eta_t \sim \mathsf{N}(0, Q_t)$, and $\alpha_1 \sim \mathsf{N}(a_1, P_1)$.

In our case, the observation vector y_t consists of the entries

$$y_t = \begin{bmatrix} C_t \\ E_t \\ S\_LND_t \\ S\_OCN_t \end{bmatrix}.$$

These are the time series from the GCB file, $t = 1, \ldots, 62$ for 1959 to 2020.

The state vector is defined as

$$\alpha_t = \begin{bmatrix} C_t^* \\ G\_ATM_t^* \\ S\_LND_t^* \\ S\_OCN_t^* \\ X_{1,t} \\ X_{2,t} \\ X_{3,t} \\ X_t^E \\ E_t^* \\ c_{1,t} \\ c_{2,t} \\ \beta_{3,t} \\ \beta_{4,t} \\ \beta_{5,t} \\ \beta_{6,t} \\ \beta_{7,t} \\ \beta_{8,t} \end{bmatrix}.$$

Obviously, this definition of the state vector leads to a non-minimal system. $G\_ATM^*_t = C^*_t - C^*_{t-1}$, S_LND^* and S_OCN^* are linear functions of C^*, and all the constant states from c_1 downward are redundant. These choices are made for convenience: The constant states essentially concentrate these parameters out of the likelihood function, reducing the number of parameters to be estimated by numerical maximization of the log-likelihood. The states G_ATM^* and S_LND^* and S_OCN^* are variables of interest, and including them in the state vector yields predicted, filtered, and smoothed values of their expectations and of their variances as output of the Kalman recursions. We provide a separate note in which we detail the equivalent minimal system.

The vector $\psi \in \mathbb{R}^{12}$ of remaining parameters to be estimated by maximum likelihood is

$$\psi = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \phi_1 \\ \phi_3 \\ \phi_E \\ \sigma^2_{\eta_1} \\ \sigma^2_{\eta_2} \\ \sigma^2_{\eta_3} \\ \sigma^2_E \\ r_{12} \\ r_{13} \\ s_E \end{bmatrix}.$$

Here, $\beta_{1,2}$ are the linear coefficients of concentrations in the land and ocean sinks, $\phi_{1,3,E}$ are the autoregressive coefficients in the deviations processes $X_{1,3}$ in the observation equation and the covariance-stationary driver X^E of changes in emissions. Note that X_2 is white noise and therefore does not have a serial correlation parameter. The variances $\sigma^2_{1,2,3,E}$ are the variances of the innovations in the deviations processes $X_{1,2,3}$ and X^E. The correlations r_{12} and r_{13} allow for covariances of the innovations of the disturbance process in the concentrations equation with that in the land sink equation and that in the ocean sink equation, respectively. Only r_{12} is estimated to be significantly different from zero, and we do not mention r_{13} in the paper; it has remained in the model from an earlier version that had more correlations, which were all estimated to be insignificant. The scaling factor s_E captures an increase in the variance of differences in emissions in 1996. This is inherited from the land-use change emissions series; we discuss this in the supplemental material Section S2.

We employ a transformation scheme to hyperparameters $\theta \in \mathbb{R}^{12}$ that can be estimated unrestricted. The transformations are as follows:

$$\begin{aligned}
\beta_{1,2} &= 10\frac{e^{-\theta_{1,2}}}{1+e^{-\theta_{1,2}}}, \\
\phi_{1,3,E} &= \frac{e^{-\theta_{3,4,5}}}{1+e^{-\theta_{3,4,5}}}, \\
\sigma^2_{1,2,3,E} &= e^{\theta_{6,7,8,9}}, \\
r_{12,13} &= \frac{1-e^{-\theta_{10,11}}}{1+e^{-\theta_{10,11}}}, \\
s_E &= 7\frac{e^{-\theta_{12}}}{1+e^{-\theta_{12}}}.
\end{aligned}$$

The transformation for the β_i ensures that they are between 0 and 10, no matter the numerical value of θ. Similarly, the autoregressive parameters ϕ_i are restricted to lie between 0 and 1, the variances σ^2_i are non-negative, the correlations r_{ij} are between -1 and 1, and the scaling factor s_E is between 0 and

7. None of the estimated parameters reaches these boundaries of the hyperparameter transformation. In calculating the standard errors for ψ from the Hessian of the estimated hyperparameters θ, these transformations have to be taken into account. We use the delta method, which requires computing the Jacobian of the transformations. This is straightforward given the exponential and logistic nature of the transformations.

The parts of the system matrices that pertain to the GCB variables are spelled out in the supplementary material. Here, we provide the full system matrices including the deviations and constant processes to facilitate understanding of the objective function in the code.

The observation matrix Z_t is of dimension 4×17 for $t = 1, \ldots, 62$ and given by

$$Z_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & I1997 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The only time-varying element is the dummy for 1997. In the time index $t = 1, \ldots, 62$ for 1959 through 2020, 1997 is at point $t = 39$. Thus, the entry $Z[2, 15]$ is equal to 1 at $t = 39$ and 0 else.

The next pages detail the state equation. The contemporaneous relations, in particular through C^*, give rise to a "structural" leading matrix B that needs to be inverted in order to arrive at a "reduced-form" vector autoregression that conforms with the standard state equation format. The model specified in Sections 2.4 and 2.5 in the paper can be written in matrix form as

$$\begin{aligned} y_t &= Z_t \alpha_t, \\ B\alpha_{t+1} &= \tilde{T}_t \alpha_t + \tilde{\eta}_t, \end{aligned}$$

see also Section S4 in the supplementary material. This system is written out in the following, where we ignore $X_{4,t}$, the deviations process in the emissions equation, because it is estimated to be essentially zero. In Section S4 in the supplementary material, we only note a constant drift d in emissions. Here, we use the time-varying drift $\Delta \log GDP_t$. Note that we write the autoregressive matrix T_t as time-dependent here (in contrast to the supplementary material), because we write out the whole system, including the constant states. The constant states are coefficients on exogenous data, dummies in the simplest cases but also SOI and GDP. These data are included in the autoregressive matrix, which renders it time-dependent.

$$
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\frac{\beta_1}{C_{1750}} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\frac{\beta_2}{C_{1750}} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
C^*_{t+1} \\ G\_ATM^*_{t+1} \\ S\_LND^*_{t+1} \\ S\_OCN^*_{t+1} \\ X_{1,t+1} \\ X_{2,t+1} \\ X_{3,t+1} \\ X^E_{t+1} \\ E^*_{t+1} \\ c_{1,t+1} \\ c_{2,t+1} \\ \beta_{3,t+1} \\ \beta_{4,t+1} \\ \beta_{5,t+1} \\ \beta_{6,t+1} \\ \beta_{7,t+1} \\ \beta_{8,t+1}
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I1991 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & SOI_{t+1} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & SOI_{t+1} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \phi_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \phi_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi_E & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & \Delta\log GDP_{t+1} & 0 & 0 & I1991 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
C^*_{t} \\ G\_ATM^*_{t} \\ S\_LND^*_{t} \\ S\_OCN^*_{t} \\ X_{1,t} \\ X_{2,t} \\ X_{3,t} \\ X^E_{t} \\ E^*_{t} \\ c_{1,t} \\ c_{2,t} \\ \beta_{3,t} \\ \beta_{4,t} \\ \beta_{5,t} \\ \beta_{6,t} \\ \beta_{7,t} \\ \beta_{8,t}
\end{bmatrix}
$$

$$
+
\begin{bmatrix}
0 \\ 0 \\ 0 \\ 0 \\ \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \\ \kappa_t \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{bmatrix}.
$$

In order to transform this system to the standard linear state space format in Equation (1) that can be passed to the KFAS routines, we need to pre-multiply the equation with B^{-1}:

$$
\begin{aligned}
y_t &= Z_t \alpha_t, \\
\alpha_{t+1} &= B^{-1}\tilde{T}_t \alpha_t + B^{-1}\tilde{\eta}_t, \\
&=: T_t \alpha_t + \eta_t.
\end{aligned}
$$

The inverse of B is given as

$$
B^{-1} = \begin{bmatrix}
\frac{1}{c} & \frac{1}{c} & -\frac{1}{c} & -\frac{1}{c} & 0 & 0 & 0 & 0 & \frac{1}{c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{\beta_1^*+\beta_2^*}{c} & \frac{1}{c} & -\frac{1}{c} & -\frac{1}{c} & 0 & 0 & 0 & 0 & \frac{1}{c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{\beta_1^*}{c} & \frac{\beta_1^*}{c} & \frac{1+\beta_2^*}{c} & -\frac{\beta_1^*}{c} & 0 & 0 & 0 & 0 & \frac{\beta_1^*}{c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{\beta_2^*}{c} & \frac{\beta_2^*}{c} & -\frac{\beta_2^*}{c} & \frac{1+\beta_1^*}{c} & 0 & 0 & 0 & 0 & \frac{\beta_2^*}{c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix},
$$

where $\beta_1^* = \beta_1/C_{1750}$, $\beta_2^* = \beta_2/C_{1750}$, and $c = 1 + \beta_1^* + \beta_2^*$. The product $B^{-1}\tilde{T}_t$ is given as

$$
T_t = \begin{bmatrix}
\frac{1}{c} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{c} & \frac{1}{c} & -\frac{1}{c} & -\frac{1}{c} & -\frac{SOI_{t+1}}{c} & -\frac{SOI_{t+1}}{c} & \frac{\Delta \log GDP_{t+1}}{c} & 0 & \frac{I1991}{c} & \frac{I1991}{c} \\
-\frac{\beta_1^*+\beta_2^*}{c} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{c} & \frac{1}{c} & -\frac{1}{c} & -\frac{1}{c} & -\frac{SOI_{t+1}}{c} & -\frac{SOI_{t+1}}{c} & \frac{\Delta \log GDP_{t+1}}{c} & 0 & \frac{I1991}{c} & \frac{I1991}{c} \\
\frac{\beta_1^*}{c} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\beta_1^*}{c} & \frac{\beta_1^*}{c} & \frac{1+\beta_2^*}{c} & -\frac{\beta_1^*}{c} & \frac{(1+\beta_2^*)SOI_{t+1}}{c} & -\frac{\beta_1^* SOI_{t+1}}{c} & \frac{\beta_1^* \Delta \log GDP_{t+1}}{c} & 0 & \frac{\beta_1^* I1991}{c} & \frac{\beta_1^* I1991}{c} \\
\frac{\beta_2^*}{c} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\beta_2^*}{c} & \frac{\beta_2^*}{c} & -\frac{\beta_2^*}{c} & \frac{1+\beta_1^*}{c} & -\frac{\beta_2^* SOI_{t+1}}{c} & \frac{(1+\beta_1^*)SOI_{t+1}}{c} & \frac{\beta_2^* \Delta \log GDP_{t+1}}{c} & 0 & \frac{\beta_2^* I1991}{c} & \frac{\beta_2^* I1991}{c} \\
0 & 0 & 0 & 0 & \phi_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \phi_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi_E & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & \Delta \log GDP_{t+1} & 0 & 0 & I1991 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}.
$$

With regard to the dummy variables, note that the time stamp of the state equation is

$$\alpha_{t+1} = T_t \alpha_t + \eta_t.$$

The notation in our dummy variables always refers to time t. That is, I1991 means a value of one at time $t = 33$. The effect of the dummy manifests itself then in $t + 1 = 34$ or 1992.

For the error vector $\tilde{\eta}_t$ and the product $\eta_t = B^{-1}\tilde{\eta}_t$, note that

$$B^{-1}\tilde{\eta}_t = \begin{bmatrix} \frac{1}{c} & \frac{1}{c} & -\frac{1}{c} & -\frac{1}{c} & 0 & 0 & 0 & 0 & \frac{1}{c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\beta_1^*+\beta_2^*}{c} & \frac{1}{c} & -\frac{1}{c} & -\frac{1}{c} & 0 & 0 & 0 & 0 & \frac{1}{c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\beta_1^*}{c} & \frac{\beta_1^*}{c} & \frac{1+\beta_2^*}{c} & -\frac{\beta_1^*}{c} & 0 & 0 & 0 & 0 & \frac{\beta_1^*}{c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\beta_2^*}{c} & \frac{\beta_2^*}{c} & -\frac{\beta_2^*}{c} & \frac{1+\beta_1^*}{c} & 0 & 0 & 0 & 0 & \frac{\beta_2^*}{c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \\ \kappa_t \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \\ \kappa_t \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

therefore $\eta_t = \tilde{\eta}_t$, and it is not necessary to define R_t in Equation (1) as B^{-1}.

We do employ the R_t object in KFAS to implement the change in variance in κ_t, see Section S2 of the supplementary material. The entry $R_t[8, 8]$ is equal to one up to (and including) 1995. From 1996 onwards, it is equal to s_E. Therefore, the variance of κ_t is equal to σ_E^2 until 1995, and from 1996 onwards, it is equal to $s_E^2 \sigma_E^2$.

The covariance matrix Q of the state equation is of dimension 17×17 and given as

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\eta_1}^2 & r_{12}\sigma_{\eta_1}\sigma_{\eta_2} & r_{13}\sigma_{\eta_1}\sigma_{\eta_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & r_{12}\sigma_{\eta_1}\sigma_{\eta_2} & \sigma_{\eta_2}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & r_{13}\sigma_{\eta_1}\sigma_{\eta_3} & 0 & \sigma_{\eta_3}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_E^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Most of the diagonal entries are zero, because the state equations of the GCB variables C^*, G_ATM^*, S_LND^*, S_OCN^*, E^* do not feature state equation errors. Instead, they have deviations processes $X_{1,3,E}$, and the innovations in these deviations processes populate the Q-matrix in rows and columns 5 through 8. The constant states in positions 10-17 of the state vector have no innovations by definition.

References

Helske, J., "Package "KFAS": Kalman filter and smoother for exponential family state-space models," Technical Report, CRAN repository, version 1.3.4 2019.