State space models for non-stationary intermittently coupled systems: an application to the North Atlantic oscillation

Philip G. Sansom, Daniel B. Williamson and David B. Stephenson

University of Exeter, UK

Summary. We develop Bayesian state space methods for modelling changes to the mean level or temporal correlation structure of an observed time series due to intermittent coupling with an unobserved process. Novel intervention methods are proposed to model the effect of repeated coupling as a single dynamic process. Latent time varying auto-regressive components are developed to model changes in the temporal correlation structure. Efficient filtering and smoothing methods are derived for the resulting class of models. We propose methods for quantifying the component of variance attributable to an unobserved process, the effect during individual coupling events and the potential for skilful forecasts. The methodology proposed is applied to the study of winter time variability in the dominant pattern of climate variation in the northern hemisphere: the North Atlantic oscillation. Around 70% of the interannual variance in the winter (December–January–February) mean level is attributable to an unobserved process. Skilful forecasts for the winter (December–January–February) mean are possible from the beginning of December.

Keywords: Bayesian filtering and smoothing; Non-linear time series; North Atlantic oscillation; Time varying auto-regressive models
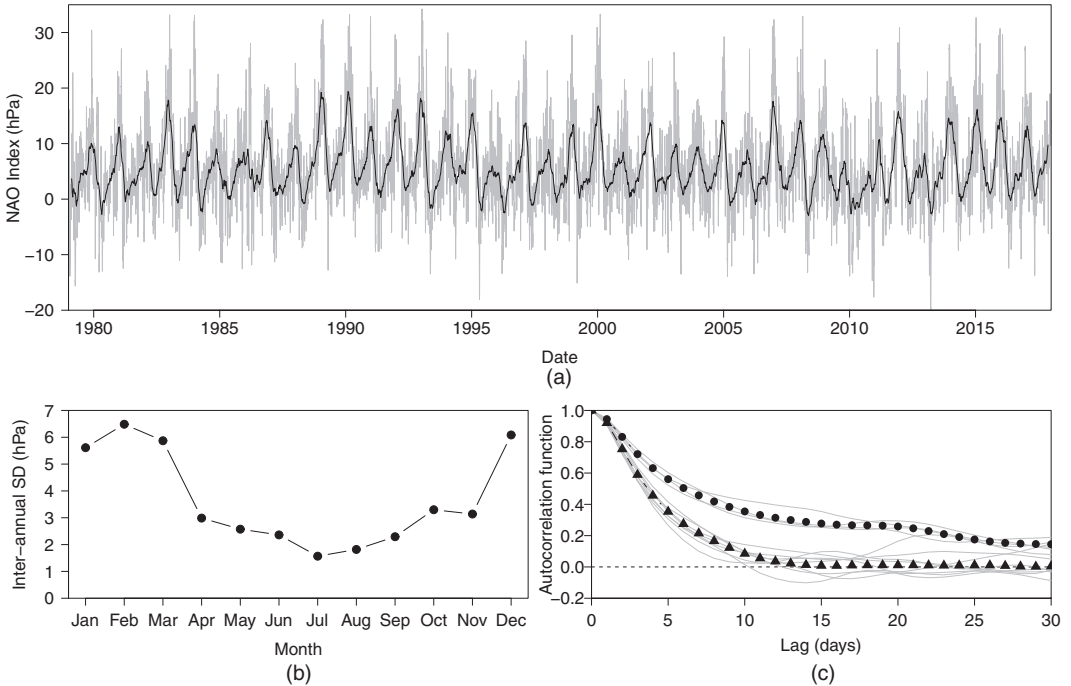
1. Introduction

Intermittently coupled systems can be found in many areas of both the natural and the social sciences. We define an intermittently coupled system as a system which can be modelled by two or more component processes which interact only at certain times. For example, many climate processes are active only during certain times of year; and sea ice and snow cover change the interaction between the surface and the atmosphere (Chapin *et al.*, 2010; Bourassa *et al.*, 2013). Migrating birds and animals mix only at certain times of year, allowing transmission of disease between populations (Olsen *et al.*, 2006; Altizer *et al.*, 2011). Empirical models have been applied to forecasting intermittent demand in production economics and operational research (Croston, 1972; Shenstone and Hyndman, 2005). Interest will often focus on one component of the system, whereas the others may be impossible or impractical to observe or even to identify physically. However, physical reasoning or prior knowledge may support the existence of such components and provide information about their behaviour and their effect on the component of interest. We refer to these secondary processes as intermittently coupled components, and the times at which the processes interact as coupling events. By incorporating this information through careful statistical modelling we can separate the

Fig. 1. NAO: (a) daily time series (———) and 90-day moving average (———) of our daily NAO index; (b) interannual standard deviation of the monthly mean NAO index and (c) the auto-correlation function of the daily NAO index computed for each month of the year (●, December–March; ▲, April–November) (a linear trend and annual and semiannual cycles were estimated by least squares and removed before computing the auto-correlation functions)

effect of intermittently coupled components from the underlying behaviour of the observed system.

The methodology that is developed in this study was motivated by the problem of diagnosing unusual persistence in the dominant mode of climate variability in the northern hemisphere, known as the North Atlantic oscillation (NAO). Because of its effect on European climate, the ability to forecast the NAO is currently a topic of great interest for the development of new climate prediction services (Siegert *et al.*, 2016). A daily time series of NAO observations is shown in Fig. 1. There is a clear annual cycle in the observations. Fig. 1 also indicates that the NAO exhibits greater interannual variability in the extended winter season (December–March) than the extended summer season (April–November). At the same time, the auto-correlation function indicates increased persistence of day-to-day conditions between December and March than between April and November. Increased persistence implies increased predictability. The seasonal contrast in interannual variability and auto-correlation that is visible in Fig. 1 could be caused by a transient shift in the mean, or a change in auto-correlation structure during between December and March. Climate scientists typically fit separate models to different seasons (e.g. Keeley *et al.* (2009) and Franzke and Woollings (2011)). This approach makes it difficult to diagnose whether the apparent change in auto-correlation is the cause of the increased interannual variability, or a symptom of it.

In this study we propose a flexible class of models that are capable of separating variability due to unobservable intermittent components from long-term variability in the observed process itself, accumulated short-term variability and observation errors. We develop tools for diagnosing

whether the intermittent component acts on the mean or the auto-correlation structure of the observed system. If we can learn the state of the intermittent component sufficiently quickly, then it should be possible to make skilful predictions about the remainder of a particular coupled period. Alternatively, the effect of the intermittent component may be similar between coupling events. In that case, it should be possible to make predictions about subsequent coupled periods.

State space models, which are also known as structural time series models, provide a flexible class of models for non-stationary time series (Durbin and Koopman, 2012). By modelling the system in terms of physically meaningful components we can incorporate expert knowledge to help to separate the effects of intermittent components from long-term variability elsewhere in the system. There is an extensive literature on modelling non-stationarity in the mean by state space methods, particularly where the observed process depends linearly on the state parameters and the observation and evolution processes are both normally distributed (Harvey, 1989; West and Harrison, 1997; Durbin and Koopman, 2012). Time varying auto-regressive (TVAR) models generalize classical auto-regressive models to have time varying coefficients, thus capturing changes in the auto-correlation structure (Subba Rao, 1970; Kitagawa and Gersch, 1985; Prado and West, 1997, 2010). In Section 3.1, we propose a class of models containing latent TVAR components that capture changes in short-term temporal dependence while maintaining the interpretability of the mean and unobserved intermittent effects.

Smooth changes in the mean or the temporal dependence structure can be captured by simple random-walk priors on their respective state variables. Rapid changes, such as those that might be expected due to intermittent coupling, often require explicit interventions in the model (Box and Tiao, 1975). Intervention methods were extended to state space models by Harvey and Durbin (1986). Standard intervention approaches (e.g. Harvey (1989), chapter 7.6, West and Harrison (1997), chapter 11, and Durbin and Koopman (2012), chapter 3.2.5) require the introduction of separate intervention and effect variables for each event. The effect is usually assumed to be constant throughout a particular event and independent between events. In the case of intermittent coupling, the underlying cause of each event will usually be the same, although the effect may vary. In Section 3.2, we model the effect of intermittent coupling as a single dynamic process, intermittently identifiable through a series of interventions that determine the timing and duration of the coupling events.

The construction of the NAO time series that is shown in Fig. 1 and analysed in Section 6 is described in Section 2. Following the methodological developments that were outlined above, we discuss efficient posterior inference for the resulting class of models in Section 5. Section 6 contains the results of our study of the NAO. Section 7 concludes with a discussion.

The data that are analysed in the paper and the programs that were used to analyse them can be obtained from

```
https://rss.onlinelibrary.wiley.com/hub/journal/14679876/series-
c-datasets
```

2. The North Atlantic oscillation

The NAO is the name that is given to the difference in surface pressure between the Azores high and the Icelandic low (Walker, 1924). The NAO is important because it affects the strength of the prevailing westerly winds and the position of the storm track, strongly influencing the winter climate of the UK and Europe (Hurrell, 1995). The NAO varies on timescales from a few days to several decades (Hurrell, 1995; Kushnir *et al.*, 2006). Statistical studies have hinted at the potential to predict the NAO on seasonal timescales (Keeley *et al.*, 2009; Franzke and

Woollings, 2011). This potential predictability is often attributed to forcing by slowly varying components of the climate system, including sea surface temperatures, the stratosphere and snow cover (Kushnir *et al.*, 2006). Climate models have recently begun to show significant skill in forecasting the winter NAO one season ahead (Scaife *et al.*, 2014). However, the physical mechanisms behind the predictability remain unclear and the size of the predictable signal appears to be underestimated by the models (Scaife *et al.*, 2014; Eade *et al.*, 2014). Careful statistical modelling may lead to additional insights. If a predictable signal can be extracted from the observations, then it may be possible to identify the source of the forcing effect.

Following Mosedale *et al.* (2006), we construct a simple NAO index as the area-weighted sea level pressure difference between two boxes, one stretching from $20°$ to $55°$ N, the other from $50°$ to $90°$ N, both spanning from $90°$ W to $60°$ E, using pressure data from the ERA-Interim data set (Dee *et al.*, 2011). The resulting daily time series, shown in Fig. 1, spans the period from January 1st, 1979, to December 31st, 2017: a total of $T = 14\,245$ observations.

3. Modelling intermittently coupled systems

In complex systems such as the Earth system, it is reasonable to consider that all components of the system (e.g. mean, seasonality and temporal dependence) may evolve slowly over time. We begin by outlining a general model to capture gradual changes in the underlying components of the observed process. We then propose explicit intervention models to represent rapid transient changes due to intermittent coupling.

3.1. *Latent time varying auto-regressive component models*

Classical auto-regressive models require that we redefine the mean of the observed process, if the mean is non-zero. This makes it difficult to specify physically meaningful models for the time evolution of the mean and the effect of intermittently coupled components. Latent auto-regressive components remove the need to redefine the mean level of the observed time series (Harvey (1989), chapter 2). To allow for possible changes in the mean, seasonal and auto-correlation structure of an observed process, we propose the following latent TVAR component model with observation equation

$$Y_t = \mu_t + \sum_{k=1}^{K} \psi_{kt} + X_t + v_t \qquad v_t \sim N(0, V), \quad k = 1, \ldots, K, \tag{1}$$

and evolution equations

$$\mu_t = \mu_{t-1} + \beta_t + w_{\mu t} \qquad w_{\mu t} \sim N(0, W_\mu), \tag{2a}$$

$$\beta_t = \beta_{t-1} + w_{\beta t} \qquad w_{\beta t} \sim N(0, W_\beta), \tag{2b}$$

$$\psi_{kt} = \psi_{k,t-1} \cos(k\omega) + \psi_{k,t-1}^* \sin(k\omega) + w_{\psi_k t} \qquad w_{\psi_k t} \sim N(0, W_\psi), \quad k = 1, \ldots, K, \tag{2c}$$

$$\psi_{kt}^* = \psi_{k,t-1}^* \cos(k\omega) - \psi_{k,t-1} \sin(k\omega) + w_{\psi_k^* t} \qquad w_{\psi_k^* t} \sim N(0, W_\psi), \quad k = 1, \ldots, K, \tag{2d}$$

$$X_t = \sum_{p=1}^{P} \phi_{pt} X_{t-p} + w_{Xt} \qquad w_{Xt} \sim N(0, W_{Xt}), \tag{2e}$$

$$\phi_{pt} = \phi_{p,t-1} + w_{\phi_p t} \qquad w_{\phi_p t} \sim N(0, W_\phi), \quad p = 1, \ldots, P, \tag{2f}$$

where $\omega = 2\pi/365.25$. The observed process Y_t is modelled as the sum of mean, seasonal and

auto-regressive components. The variable μ_t represents the mean level of the observed process. Any local-in-time systematic trend is captured by the variable β_t. The harmonic components ψ_{kt} and ψ_{kt}^* ($k = 1, \ldots, K$) represent seasonal behaviour. The local trend and seasonal variables are assumed to be time varying, evolving according to independent normal evolution processes $w_{\mu t}$, $w_{\beta t}$, $w_{\psi_k t}$ and $w_{\psi_k^* t}$ ($k = 1, \ldots, K$). The irregular component X_t represents short-term variability in the observed process and is modelled as a latent TVAR process of order P with normal evolution process w_{Xt}. The auto-regressive coefficients ϕ_{pt} are assumed to be time varying, evolving according to independent normal evolution processes $w_{\phi_p t}$ ($p = 1, \ldots, P$). The independent residual v_t represents observation or measurement error.

In the case of the NAO, the variance W_{Xt} of the evolution process w_{Xt} is expected to vary systematically with the solar cycle and is modelled as

$$W_{Xt} = W_X + \sqrt{(a^2 + b^2)} + a\sin(\omega t) + b\cos(\omega t) \qquad W_X > 0. \qquad (3)$$

The other evolution and error variances W_μ, W_β, W_ψ, W_ϕ and V are assumed constant over time. Model (2) is intended to capture gradual changes in the structure of the observed process. Therefore, the evolution variances W_μ, W_β, W_ψ and W_ϕ are expected to be small, in particular $W_\mu, W_\beta, W_\psi, W_\phi \ll W_{Xt}$. The evolution and error variances are assumed unknown and must be inferred from the data. The variance parameters W_X, a and b of the irregular component in model (3) must also be inferred from the data. Expert judgement about the scale of the evolution variances can be incorporated through appropriate prior probability distributions.

The model defined by expressions (1) and (2) is quite general and could be applied to a range of climate, economic or environmental time series. Examination of the sample periodogram for our NAO index showed clear evidence of 6- and 12-month cycles, suggesting a model with $K = 2$ harmonic components. Examination of the sample auto-correlation and partial auto-correlation functions suggest a latent TVAR process with $P = 5$ coefficients (after removing a linear trend, and constant annual and semiannual cycles estimated by least squares).

3.2. Intervention methods for intermittent coupling

The change in the auto-correlation structure of the NAO index in Fig. 1 appears to involve two distinct states, i.e. coupled or not. We model the change from the uncoupled to the coupled state by introducing an intervention variable

$$\lambda_t = \begin{cases} 0 & \text{if } t \notin \tau, \\ 1 & \text{if } t \in \tau \end{cases}$$

where τ is the set of times t where the observed system is believed to be coupled to the unobserved process, e.g. $\tau = \{\text{December, January, February, March}\}$. We assume that the timing and duration of the coupling events is constant between events but not known precisely. We model the intervention λ_t by introducing two hyperparameters α and γ representing the start and duration of the coupled period τ respectively (Fig. 2). In practice, we do not expect an instantaneous change in the behaviour of the system. Therefore, we linearly taper the intervention λ_t from 0 to 1 over a period γ_1 at the start of the coupled period and from 1 to 0 over a period γ_2 at the end of the coupled period. In the absence of stronger beliefs, we assume that the tapering is symmetric (i.e. $\gamma_1 = \gamma_2$) and accounts for a proportion $\rho = (\gamma_1 + \gamma_2)/\gamma$ of the duration γ. The hyperparameters α, γ and ρ are assumed to be unknown and must be inferred from the data.

We consider two alternative models for the effect of intermittent coupling. First, coupling may lead to a transient change in the mean of the observed process; second, coupling may lead to a transient change in the temporal dependence structure of the observed process. If coupling
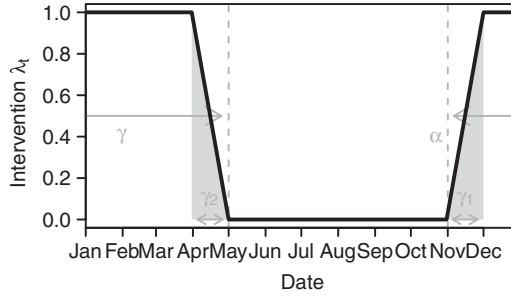
Fig. 2. Example of the form and parameterization of the intervention λ_t: parameters α and γ represent the start and duration of the coupled period, whereas $\rho = (\gamma_1 + \gamma_2)/\gamma$ controls the transition

is believed to induce a change in the mean, then the forecast equation (1) is modified to include the intervention as follows:

$$Y_t = \mu_t + \sum_{k=1}^{K} \psi_{kt} + X_t + \lambda_t \delta_t + v_t \qquad v_t \sim N(0, V). \qquad (4)$$

The effect δ_t is modelled as

$$\delta_t = \varphi \delta_{t-1} + w_{\delta t} \qquad w_{\delta t} \sim N(0, W_\delta). \qquad (5)$$

We interpret the effect δ_t as a change in the mean level, since we expect the effect variance to be small, i.e. $W_\delta \ll W_{X_t}$. However, when $\lambda > 0$, the day-to-day variability of the observed process Y_t will increase slightly, in addition to any systematic change that is captured by W_{X_t} in expression (3).

If coupling is believed to induce a change in the auto-correlation structure, then we modify the forecast equation (1) again,

$$Y_t = \mu_t + \sum_{k=1}^{K} \psi_{kt} + X_t + \lambda_t \sum_{p=1}^{P} \delta_{pt} X_{t-p} + v_t \qquad v_t \sim N(0, V), \qquad (6)$$

and we define P effects δ_{pt}, modelled as

$$\delta_{pt} = \varphi \delta_{p,t-1} + w_{\delta_{pt}} \qquad w_{\delta_p t} \sim N(0, W_\delta), \quad p = 1, \ldots, P, \qquad (7)$$

with common hyperparameters φ and W_δ.

Most of our prior knowledge about coupling events is likely to be about their timing and will be expressed through priors on the hyperparameters α, γ and ρ. Therefore, it is difficult to justify a complex form for the effects δ_t or δ_{pt}. However, a variety of behaviours can be captured depending on the values of the coefficient φ and variance W_δ.

As noted in the previous section, the mean, trend, seasonal and auto-regressive parameters are expected to vary only slowly. Therefore, we can learn their states outside the coupled period and identify the coupling effects δ_t or δ_{pt} ($p = 1, \ldots, P$) when $\lambda_t > 0$. The form and parameterization of the coupling intervention λ_t in Fig. 2 reflect our physical intuition about the likely influence of an unobserved process on the NAO. For other applications, different forms might be appropriate, e.g. no tapering, non-symmetric tapering or non-linear tapering. We recommend keeping $0 \leqslant \lambda_t \leqslant 1$, to make the coupling effect easily interpretable. The only other restriction is that the intervention should be transient, not permanent. Permanent changes can be modelled in the same way, but the effects should be fixed to be identifiable, i.e. $\varphi = 1$ and $W_\delta = 0$.

4. State space form and prior assessment

The model that was proposed in Section 3.1 can be written in state space form as

$$Y_t = f(\boldsymbol{\theta}_t, v_t) \qquad v_t \sim N(0, V),$$
$$\boldsymbol{\theta}_t = g(\boldsymbol{\theta}_{t-1}, \mathbf{w}_t) \qquad \mathbf{w}_t \sim N(\mathbf{0}, \mathbf{W})$$

for $t = 1, \ldots, T$ with $\boldsymbol{\theta}_0 \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\theta}_t = (\mu_t, \beta_t, \psi_{1t}, \psi_{1t}^*, \ldots, \psi_{Kt}, \psi_{Kt}^*, X_t, \ldots, X_{t-P+1}, \phi_{1t}, \ldots, \phi_{Pt})'$ and $\mathbf{w}_t = (w_{\mu t}, w_{\beta t}, w_{\psi_1 t}, w_{\psi_1 t}^*, \ldots, w_{\psi_K t}, w_{\psi_K t}^*, w_{Xt}, w_{\phi_1 t}, \ldots, w_{\phi_P t})'$. The forecast function $f(\boldsymbol{\theta}_t, v_t)$ is given by equation (1). The evolution function $g(\boldsymbol{\theta}_{t-1}, \mathbf{w}_t)$ is given by expression (2). The evolution covariance matrix $\mathbf{W}$ is diagonal with main diagonal $\mathbf{W}_t = (W_\mu, W_\beta, W_\psi, W_\psi, \ldots, W_\psi, W_\psi, W_{Xt}, W_\phi, \ldots, W_\phi)'$. The coupling effect δ_t or effects δ_{pt} ($p = 1, \ldots, P$) can be appended to the state vector $\boldsymbol{\theta}_t$. The evolution process vector $\mathbf{w}_t$ and covariance matrix $\mathbf{W}$ can also be extended to include the coupling effect evolution process $w_{\delta t}$ or processes $w_{\delta_{pt}}$ ($p = 1, \ldots, P$) and variance W_δ respectively.

The prior distribution $\boldsymbol{\theta}_0 \sim N(\mathbf{m}_0, \boldsymbol{\Sigma}_0)$ specifies our beliefs about the state variables at time $t = 0$. We also need to specify priors for the collection of hyperparameters $\boldsymbol{\Phi} = (V, W_\mu, W_\beta, W_\psi, W_\phi, W_X, a, b, \alpha, \gamma, \rho, \varphi, W_\delta)'$.

4.1. Priors for the state variables

Independent normal priors were assigned to each component of the state vector $\boldsymbol{\theta}$ at time $t = 0$. The prior means and variances are listed in Table 1. We use previous studies of the NAO to define informative priors for the mean μ_0 (Hsu and Wallace, 1976), seasonal components $\psi_{1,0}, \psi_{1,0}^*, \psi_{2,0}$ and $\psi_{2,0}^*$ (Chen *et al.*, 2012) and TVAR coefficients $\phi_{1,0}, \ldots, \phi_{5,0}$ (Masala, 2015). The prior on the local trend β_0 is based on our judgement that the NAO mean is very unlikely to experience a local change that is equivalent to more than 1 hPa year^{-1}. The daily NAO in Fig. 1 has a range of approximately 40 hPa. Therefore, the TVAR residuals $X_{-4}, \ldots, X_0$ were assigned independent normal priors with mean 0 hPa and standard deviation 10 hPa, based on a range of four standard deviations.

In Fig. 1, the NAO index has an interannual standard deviation of 5–6 hPa between December and March. Therefore, in the model with a mean intervention, the coupling effect δ_0 was assigned a normal prior with mean 0 hPa and standard deviation 5 hPa. The partial auto-correlation functions (not shown) for December–March and April–November suggest that the change in the auto-correlation structure that is represented by the coefficients $\phi_{1,0}, \ldots, \phi_{Pt}$ is quite small.

Table 1. Prior probability distributions for the state variables θ_0, all normally distributed

Component	Parameter	Mean	Variance
Mean level	μ_0	6.0 hPa	1^2
Local trend	β_0	0.0 hPa year^{-1}	0.002^2
Annual cycle	$\psi_{1,0}$	3.6 hPa	1.0^2
Annual cycle	$\psi_{1,0}^*$	1.0 hPa	1.5^2
Semiannual cycle	$\psi_{2,0}$	1.3 hPa	0.9^2
Semiannual cycle	$\psi_{2,0}^*$	0.7 hPa	1.3^2
Irregular component	$X_{-4}, \ldots, X_0$	0.0 hPa	10^2
TVAR coefficients	$\phi_{1,0}, \ldots, \phi_{5,0}$	$1.8, -1.3, 0.7,$ $-0.3, 0.1$	0.2^2

Therefore, in the model with an intervention on the auto-correlation structure, the coupling effects $\delta_{1,0}, \ldots, \delta_{P,0}$ were assigned normal priors with mean 0.0 hPa and standard deviation 0.2 hPa.

4.2.   Priors on the hyperparameters

The prior distributions that were assigned to the hyperparameters V, W_μ, W_β, W_ϕ and W_X, a and b are listed in Table 2. In the case of the NAO, the variability in the mean and seasonal components will be driven primarily be solar forcing; therefore we assume equal error variances, i.e. $W_\psi = W_\mu$. The observation and evolution variances V, W_μ, W_β and W_ϕ are all expected to be very small, but non-zero. Therefore, boundary avoiding priors were specified in the form of normal distributions on the logarithm of each variance parameter. Simulation studies of the individual components in expression (2) were used to assign priors that reflect the range of variability that we consider plausible for each component. We expect the annual cycle in the day-to-day variance W_{Xt} to peak during the winter season (December–January–February) with an amplitude of up to 5 hPa^2. Corresponding uniform priors were assigned to the amplitude and phase of W_{Xt} and transformed to approximate normal priors for a and b by simulation.

Table 3 lists the priors for the intervention parameters α, γ and ρ and the coupling effect parameters φ and W_δ. Our beliefs about the timing of the intervention λ_t are the same regardless of whether coupling affects the mean or the auto-correlation structure. Vague triangular priors are specified for the beginning α and duration γ of the coupled period. These suggest a coupled

Table 2.   Prior densities for hyperparameters

Component	Parameter	Prior	$\approx 95\%$ interval
Observation variance	$\log(V)$	$N(-10, 3^2)$	$(-16, -4)$
Mean variance	$\log(W_\mu)$	$N(-12, 3^2)$	$(-18, -6)$
Trend variance	$\log(W_\beta)$	$N(-28, 3^2)$	$(-34, -22)$
Irregular variance	$\log(W_X)$	$N(0.0, 1^2)$	$(-2.0, 2.0)$
Irregular variance	a	$N(0.5, 1^2)$	$(-1.5, 2.5)$
Irregular variance	b	$N(2.0, 1^2)$	$(0.0, 4.0)$
Coefficient variance	$\log(W_\phi)$	$N(-18, 3^2)$	$(-24, -12)$

Table 3.   Prior densities for intervention hyperparameters

Component	Parameter	Prior	$\approx 95\%$ interval
Coupling start	$\alpha - 120$	Tri(0, 365, 185)	(40, 325)
Coupling length	γ	Tri(0, 365, 180)	(40, 325)
Tapered proportion	ρ	beta(4, 6)	(0.15, 0.70)
Mean effect variance	$\log(W_{\delta_\mu})$	$N(-8, 4^2)$	$(-16, 0)$
Mean effect coefficient	φ_μ	beta(4, 1)	(0.4, 1.0)
Auto-correlation effect variance	$\log(W_{\delta_\phi})$	$N(-16, 4^2)$	$(-24, -8)$
Auto-correlation effect coefficient	φ_ϕ	beta(45, 1)	(0.9, 1.0)

period with total length around 180 days, beginning around November 1st. A mildly informative prior is specified for the tapering parameter ρ to reflect our physical reasoning that the influence of the unobserved process is unlikely to be constant throughout the coupled period. The coupling coefficients φ_μ and φ_X are expected to be positive and close to but not exceeding 1. The mean coupling effect variance W_{δ_μ} is expected to be greater than the mean variance W_μ, but still small compared with W_{X_t}. Similarly, the auto-correlation coupling effect variance W_{δ_X} is expected to be greater than the coefficient evolution variance W_ϕ.

5. Posterior inference

We want to evaluate the joint posterior of the model components $\theta_1, \ldots, \theta_T$ and the hyper-parameters Φ:

$$\Pr(\theta_{1:T}, \Phi \mid Y_{1:T}) = \Pr(\theta_{1:T} \mid \Phi, Y_{1:T}) \Pr(\Phi \mid Y_{1:T}).$$

If both $f(\theta_t, v_t)$ and $g(\theta_{t-1}, \mathbf{w}_t)$ were linear functions, then, conditionally on Φ, we could sample from the marginal posterior of the state variables $\Pr(\theta_{1:T} \mid \Phi, Y_{1:T})$ by using the well-known *forward filtering–backward sampling* algorithm (Frühwirth-Schnatter, 1994). However, the evolution function $g(\theta_{t-1}, \mathbf{w}_t)$ that is defined by expression (2) is non-linear because of the combination of ϕ_p and X_{t-p} in expression (2e). The form of the observation equation (6) also contains a non-linear combination of δ_{pt} and X_{t-p}. Therefore, we use linear approximations of the observation and state equations:

$$Y_t \approx f(\hat{\theta}_t, \hat{v}_t) + \left.\frac{\partial f}{\partial \theta}\right|_{\hat{\theta}_t, \hat{v}_t} (\theta_t - \hat{\theta}_t) + \left.\frac{\partial f}{\partial v}\right|_{\hat{\theta}_t, \hat{v}_t} (v_t - \hat{v}_t),$$

$$\theta_t \approx g(\hat{\theta}_{t-1}, \hat{\mathbf{w}}_t) + \left.\frac{\partial g}{\partial \theta}\right|_{\hat{\theta}_{t-1}, \hat{\mathbf{w}}_t} (\theta_{t-1} - \hat{\theta}_{t-1}) + \left.\frac{\partial g}{\partial \mathbf{w}}\right|_{\hat{\theta}_{t-1}, \hat{\mathbf{w}}_t} (\mathbf{w}_t - \hat{\mathbf{w}}_t)$$

where $\hat{\theta}_{t-1} = E(\theta_{t-1})$, $\hat{\theta}_t = E(\theta_t)$, $\hat{\mathbf{w}}_t = E(\mathbf{w}_t)$ and $\hat{v}_t = E(v_t)$. This linearization leads to approximate forward filtering–backward sampling recursions, detailed in Appendix A.

In general, we expect the TVAR evolution variance W_ϕ to be small, so the coefficients $\phi_{1t}, \ldots, \phi_{Pt}$ will be only weakly correlated with the other state variables and our uncertainty about the coefficients will decrease rapidly over time. Since the other components of the evolution function $g(\theta_{t-1}, \mathbf{w}_t)$ are linear and the observation errors v_t and joint state evolution process $\mathbf{w}_t$ are normal, forward filtering and backward sampling based on the linear approximation are expected to be very accurate. A simulation study showed that the linear approximation provides excellent filtering and smoothing performance, even when all components of the model evolve much more rapidly than expected (see the on-line supplementary material). The linear approximation sometimes struggles to distinguish the TVAR coefficients $\phi_{1t}, \ldots, \phi_{Pt}$ from the intervention effects $\delta_{1t}, \ldots, \delta_{Pt}$ in the auto-correlation intervention model when both sets of coefficients evolve rapidly. In the case of the NAO, we expect only slow evolution of the TVAR coefficients, and little or no change in the intervention effects. In this scenario, the linearized approximation performs very well.

The marginal posterior of the hyperparameters $\Pr(\Phi \mid Y_{1:T})$ is proportional to

$$\Pr(\Phi \mid Y_{1:T}) \propto \Pr(Y_{1:T} \mid \Phi) \Pr(\Phi).$$

The marginal likelihood $\Pr(Y_{1:T} \mid \Phi)$ can be decomposed as

$$\Pr(Y_{1:T} \mid \Phi) = \Pr(Y_1 \mid \Phi) \prod_{t=2}^{T} \Pr(Y_t \mid Y_{1:t-1}, \Phi).$$

The forward filtering recursions in Appendix A include an expression for the one-step-ahead forecast distribution $\Pr(Y_t \mid Y_{1:t-1}, \boldsymbol{\Phi})$. So the marginal likelihood can be evaluated analytically. Therefore, the joint posterior $\Pr(\boldsymbol{\theta}_{1:T}, \boldsymbol{\Phi} \mid Y_{1:T})$ can be efficiently sampled by combining forward filtering–backward sampling with a Metropolis–Hastings scheme targeting $\Pr(\boldsymbol{\Phi} \mid Y_{1:T})$ as follows.

(a) Let j denote a sample index, at $j = 1$;
    (i) sample starting values $\boldsymbol{\Phi}^{(1)}$;
    (ii) sample $\boldsymbol{\theta}_{1:T}^{(1)} \mid \boldsymbol{\Phi}^{(1)}, Y_{1:T}$ by backward sampling.
(b) For $j = 2, \ldots, J$:
    (i) sample new values $\boldsymbol{\Phi}^*$ from proposal $q(\boldsymbol{\Phi}^* \mid \boldsymbol{\Phi})$;
    (ii) accept $\boldsymbol{\Phi}^*$ with probability

$$\min\left\{ \frac{\Pr(Y_{1:T} \mid \boldsymbol{\Phi}^*)\Pr(\boldsymbol{\Phi}^*)q(\boldsymbol{\Phi} \mid \boldsymbol{\Phi}^*)}{\Pr(Y_{1:T} \mid \boldsymbol{\Phi})\Pr(\boldsymbol{\Phi})q(\boldsymbol{\Phi}^* \mid \boldsymbol{\Phi})}, 1 \right\};$$

    (iii) sample $\boldsymbol{\theta}_{1:T}^{(j)} \mid \boldsymbol{\Phi}^{(j)}, Y_{1:T}$ by backward sampling.

In practice, it is not necessary to perform backward sampling for the state $\boldsymbol{\theta}_{1:T}$ for every sample (j). As with any Markov chain Monte Carlo approach, there is likely to be significant auto-correlation between subsequent samples of the hyperparameters $\boldsymbol{\Phi}^{(j)}$. In the interest of saving storage and computation time, it is sufficient to sample the state $\boldsymbol{\theta}_{1:T}$ for a subset of the $\boldsymbol{\Phi}^{(j)}$.

5.1. Alternative approaches

Conditionally on the TVAR coefficients $\phi_{1t}, \ldots, \phi_{Pt}$, the model defined by expressions (1) and (2) is a normal dynamic linear model. We could split the state vector $\boldsymbol{\theta}_t$ into two parts $\boldsymbol{\theta}_t^* = (\mu_t, \beta_t, \psi_{1t}, \psi_{1t}^*, \ldots, \psi_{Kt}, \psi_{Kt}^*, X_t, \ldots, X_{t-P+1})'$ and $\boldsymbol{\phi}_t^* = (\phi_{1t}, \ldots, \phi_{Pt})'$, and then alternate between forward filtering and backward sampling for each part, conditionally on the other. Gibbs sampling steps could be used to sample the hyperparameters $\boldsymbol{\Phi}$ (West and Harrison (1997), chapter 15). This approach provides exact sampling from the required posterior distribution but has two drawbacks compared with the approximate approach that is proposed here. First, backward sampling must be performed at every iteration, making this approach computationally expensive. Second, Gibbs sampling based on the full conditional distributions of the hyperparameters will tend to mix very slowly, especially for long time series where the data completely overwhelm the prior.

Particle filtering methods provide tools for inference in general non-linear and non-normal state space models (Doucet and Johansen, 2011). However, particle filters are computationally expensive and suffer from problems of 'particle degeneracy', i.e. the state $\boldsymbol{\theta}_t$ will eventually be represented by a single particle at times $t \ll T$. Since we are interested in what happened at all times $t = 1, \ldots, T$, we also require particle smoothing to overcome the degeneracy problem (Godsill *et al.*, 2004; Briers *et al.*, 2010). Particle smoothing is even more computationally expensive, making an alternative approach highly desirable. The problem of efficient inference for unknown hyperparameters also remains an active topic for research in sequential Monte Carlo methods (Chopin *et al.*, 2013).

5.2. Model selection

For some applications, it will be possible to choose between the mean and auto-correlation intervention models on the basis of posterior predictive diagnostics, i.e. whether the model reproduces the observed behaviour. The posterior distributions of the hyperparameters $\boldsymbol{\Phi}$ can

also be useful for choosing between models; for example, is the coupling effect variance W_δ negligible? More formally, we can compare the two intervention models by evaluating the Bayes factor

$$B = \frac{\Pr(Y_{1:T} \mid M_\mu)}{\Pr(Y_{1:T} \mid M_X)} = \frac{\int \Pr(Y_{1:T} \mid \Phi, M_\mu) \Pr(\Phi \mid M_\mu) \, \mathrm{d}\Phi}{\int \Pr(Y_{1:T} \mid \Phi, M_X) \Pr(\Phi \mid M_X) \, \mathrm{d}\Phi} \tag{8}$$

where M_μ is the model including an intervention on the mean, and M_X is the model including an intervention on the temporal dependence structure. The Bayes factor is defined as the ratio of the marginal likelihoods of the competing models (Kass and Raftery, 1995). Values of B greater than 1 indicate support for the mean intervention model M_μ and values of B less than 1 indicate support for the auto-correlation intervention model M_X. The conditional likelihoods $\Pr(Y_{1:T} \mid \Phi, M)$ can be evaluated by using the filtering recursions in Appendix A. The marginal likelihoods $\Pr(Y_{1:T} \mid M)$ can be evaluated by using the posterior samples $\Phi^{(j)} \mid Y_{1:T}, M$ $(j = 1, \ldots, J)$ by bridge sampling (Gronau *et al.*, 2017).

5.3. What is the effect of the coupling?

Given posterior samples $\theta_{1:T}^{(j)} \mid \Phi^{(j)}, Y_{1:T}, M$, we can make inferences about any function of the state variables θ_t for any time period τ of interest, e.g. $\tau = \{\text{December 2009–March 2010}\}$. It is useful to define $\eta_t = \mu_t + \Sigma_k \psi_{kt}$, which we refer to as the systematic component of the observed process. The relative contributions to the variability between coupled periods of the systematic component η_t, the irregular component X_t, the coupling effects δ_t or δ_{pt} $(p = 1, \ldots, P)$ and observation error v_t are of particular interest. The means of the systematic and irregular components during period τ in the jth sample are

$$\bar{\eta}_\tau^{(j)} = \frac{1}{n} \sum_{t \in \tau} \eta_t^{(j)},$$
$$\bar{X}_\tau^{(j)} = \frac{1}{n} \sum_{t \in \tau} X_t^{(j)} \tag{9}$$

where n is the number of time steps in τ. The means of the coupling effects during period τ in the mean and auto-correlation models respectively are

$$\bar{\delta}_{\mu\tau}^{(j)} = \frac{1}{n} \sum_{t \in \tau} \lambda_t \delta_t^{(j)},$$
$$\bar{\delta}_{X\tau}^{(j)} = \frac{1}{n} \sum_{t \in \tau} \lambda_t \sum_{p=1}^{P} \delta_{pt}^{(j)} X_{t-p}^{(j)}. \tag{10}$$

The contribution due to observation error is

$$\bar{v}_\tau^{(j)} = \bar{Y}_\tau - \bar{\eta}_\tau^{(j)} - \bar{\delta}_\tau^{(j)} - \bar{X}_\tau^{(j)}$$

where $\bar{Y}_\tau = \Sigma_{t \in \tau} Y_t / n$. The prior expectations of the irregular component X_t and the coupling effects δ_t or δ_{pt} $(p = 1, \ldots, P)$ during any period τ are 0 by expressions (2e), (5) and (7), i.e. $E(X_t) = 0$ and $E(\delta_t) = E(\delta_{pt}) = 0$ for all t. In general $E(\eta_t) \neq 0$, so for the systematic component η_t it is more useful to consider the anomalies over all similar periods:

$$\bar{\eta}_\tau^{*(j)} = \bar{\eta}_\tau^{(j)} - \frac{1}{|D|} \sum_{t' \in D} \eta_{t'}^{(j)}$$

where $D = \{t \in 1, \ldots, T : d(t) = d(s) \text{ and } s \in \tau\}$ and $d(t)$ is the day of the year at time t. The sample means

$$\bar{\eta}_\tau^* = \frac{1}{J} \sum_j \bar{\eta}_\tau^{*(j)}, \qquad \bar{\delta}_\tau = \frac{1}{J} \sum_j \bar{\delta}_\tau^{(j)}, \qquad \bar{X}_\tau = \frac{1}{J} \sum_j \bar{X}_\tau^{(j)}, \qquad \bar{v}_\tau = \frac{1}{J} \sum_j \bar{v}_\tau^{(j)} \qquad (11)$$

provide a summary of the posterior expected contribution of each component during the period τ. Quantiles can also be computed over the samples to form credible intervals for the contribution of each component.

5.4. Analysis of variance

In a stationary model, elements of the marginal posterior $\Pr(\boldsymbol{\Phi} \mid Y_{1:T}, M)$ would summarize the relative contributions of each model component to the observed variability in the index Y_t. However, since our model is non-stationary, we require an alternative summary of the variance components. In particular, we are interested in the proportion of the interannual variance of the winter (December–January–February) mean of the NAO index explained by each component. Let τ_i be the ith winter period. We propose to perform an analysis of variance of the observed means $\bar{y}_{\tau_i} = (1/N_i) \sum_{t \in \tau_i} y_t$ for each sample j by using the component means $\bar{\eta}_{\tau_i}^{-(j)}$, $\bar{\delta}_{\tau_i}^{(j)}$ and $\bar{X}_{\tau_i}^{(j)}$ defined in expressions (9) and (10) as explanatory variables. The analysis of variance leads to four sums of squares for each sample j, corresponding to the sum of squared deviations explained by the systematic η_t- and irregular X_t-components, the coupling effects δ_t or δ_{pt} ($p = 1, \ldots, P$) and observation errors v_t in each sample trajectory. Posterior summaries over the J samples summarize the overall contributions of each component to the variability between coupled periods.

5.5. Can we make predictions using unobserved components?

Knowledge of the unobserved component through the intervention effect δ_t should provide useful predictability within coupled periods. The model that was proposed in Section 3.2 also allows for dependence between successive coupled periods, so knowledge of the unobserved component during one coupled period may also be useful for predicting the next. The k-step-ahead forecast distribution given data up to time t can be sampled exactly by using the recursions in Appendix A. The correlation between the data and the forecast means provides a simple measure of forecast performance.

6. Results

The Metropolis–Hastings sampler that was outlined in Section 5 was used to draw 1000 samples from each of the joint posteriors $\Pr(\boldsymbol{\theta}_{1:T}, \boldsymbol{\Phi} \mid Y_{1:T}, M_\mu)$ and $\Pr(\boldsymbol{\theta}_{1:T}, \boldsymbol{\Phi} \mid Y_{1:T}, M_X)$. Full details of the sampling design, proposal distributions, diagnostic trace plots and posterior density plots are given in the on-line supplementary material. Both models converge to stable distributions and mix efficiently; however, the burn-in period can be very long depending on the initial values of the hyperparameters $\boldsymbol{\Phi}$.

Despite deliberately vague prior distributions, the posterior distributions of the intervention parameters α and γ are quite sharp for both models. Fig. 3 visualizes the posterior distribution of the intervention λ_t for each model. In the mean intervention model M_μ, an unobserved component acts strongly on the NAO between December and February and into March. There is almost no evidence of coupling between May and October. In the auto-correlation intervention model M_X the situation is reversed. The inverted intervention structure is unexpected but still
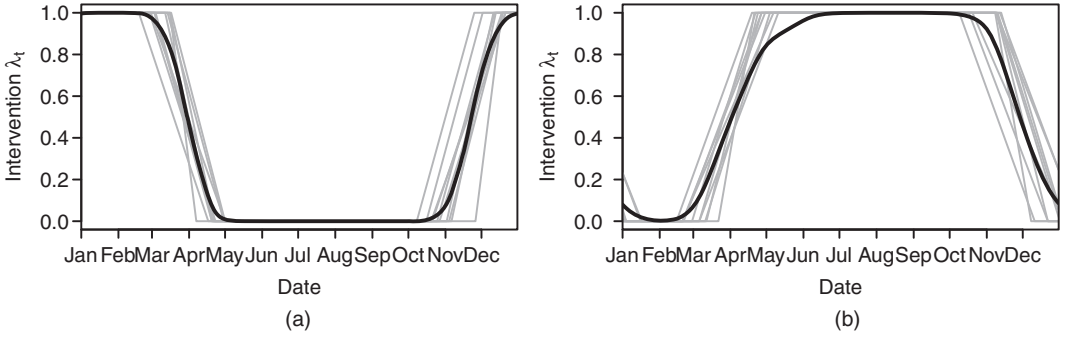
Fig. 3. Posterior of the intervention λ_t (———, random sample of 10 realizations of the intervention $\lambda_t^{(j)}$ based on the posterior samples of α, γ and ρ; ———, pointwise posterior mean over all 1000 realizations of $\lambda_t^{(j)}$): (a) model with an intervention on the mean M_μ; (b) model with an intervention on the auto-correlation structure M_X

consistent with a marked difference in behaviour between the extended winter (December–March) and extended summer (April–November) seasons. Prior sensitivity analysis showed that the posterior distributions of the hyperparameters Φ, including α and γ, are insensitive to the choice of priors in Tables 2 and 3 (see the supplementary material).

To assess the identifiability of the various model components, particularly the coupling effects, we computed correlation matrices for the states $\theta_{1:T}^{(j)} \mid \Phi^{(j)}, \mathbf{Y}_{1:T}$ for each sample j. On average across the 1000 sample covariance matrices, the state variables in both models are all uncorrelated with one another. In particular, the mean intervention effect δ_t is almost completely uncorrelated with the irregular component X_t (90% credible interval of correlation $(-0.02, 0.06)$), and only ever weakly correlated with the mean component μ_t (90% credible interval $(-0.28, 0.20)$). Whereas the auto-correlation intervention effects $\delta_{1t}, \ldots, \delta_{5t}$ are uncorrelated with the other state variables on average, they may be strongly correlated *or* anticorrelated with the mean μ_t and the auto-correlation coefficients $\phi_{1t}, \ldots, \phi_{5t}$. Further investigation showed that these strong associations were the result of the slow rate of change of these parameters, since sampling multiple state trajectories $\theta_{1:T}$ from any single sample of the hyperparameters $\Phi_{(j)}$ produced a similar range of sample correlations.

Posterior predictive diagnostics were used to check the performance of each model in capturing the observed structure of the NAO. In particular, we are interested whether the model can reproduce the seasonal contrast in the interannual variance and auto-correlation structures in Fig. 1. For each sample $\theta_{1:T}^{(j)}, \Phi^{(j)} \mid Y_{1:T}$ from each model we simulate a new sequence of states $\theta_\tau^{*(j)} \mid \Phi^{(j)}, Y_{1:T}$ and observations $Y_\tau^{*(j)}$ for the period $\tau = \{$January 1988–December 2017$\}$. Fig. 4 compares the annual cycle in the interannual standard deviation and the seasonal auto-correlation functions of the observed data Y_τ and the samples $Y_\tau^{*(j)}$ for $j = 1, \ldots, 1000$. The mean intervention model M_μ can reproduce both the interannual variability and the seasonal auto-correlation function. There is a clear difference in the auto-correlation functions simulated between April and November, and between December and March. However, the auto-correlation intervention model M_X cannot reproduce the seasonal auto-correlation behaviour and does not reproduce the interannual variability as well as the mean intervention model M_μ. There is a small difference between the extended summer (April–November) and extended winter (December–March) auto-correlation functions, but much less than observed in the data. The inverted intervention structure in Fig. 3 is an attempt to exploit the extended summer (April–November) period to distinguish the small intervention effects δ_p. Similar checks (which
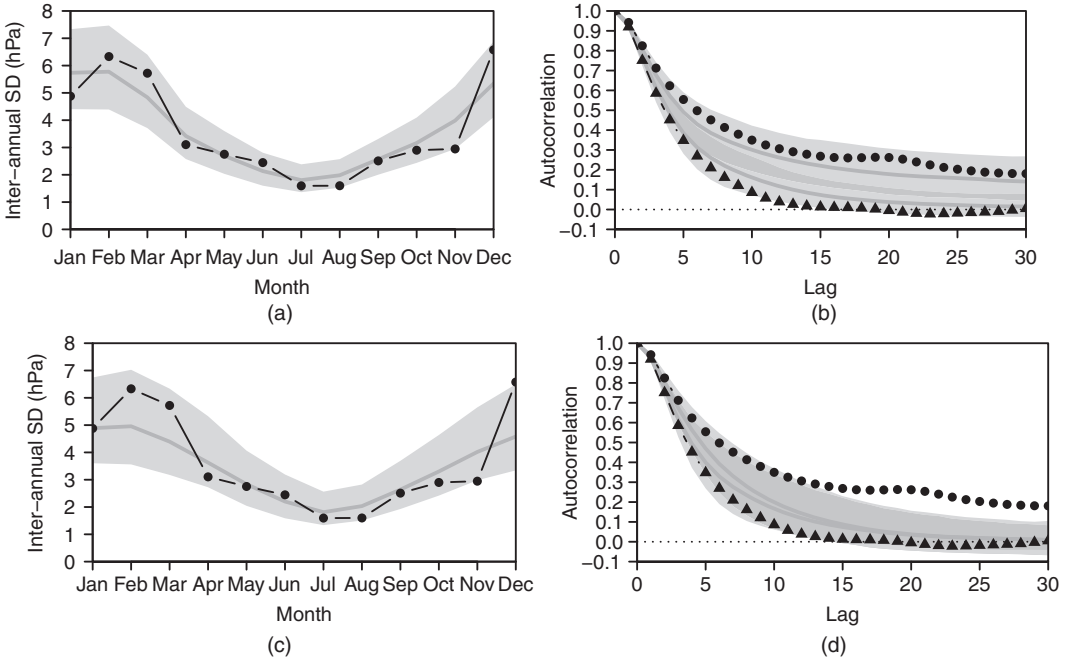
Fig. 4. Posterior predictive checks: (a), (b) model with an intervention on the mean M_μ; (c), (d) model with an intervention on the auto-correlation structure M_χ (●, December–March; ▲, April–November) (before computing the auto-correlation, the mean, a linear trend and annual and semiannual cycles were estimated by least squares and removed; ——, observed statistics; ——, posterior mean; ■, pointwise 90% posterior credible intervals; ■, overlap between credible intervals)

are not shown) suggest that both models can adequately capture the annual cycle in the NAO, indicating that our choice of $K = 2$ harmonics was reasonable.

The posterior predictive checks strongly favour the mean intervention model over the auto-correlation intervention model. The mean intervention can reproduce the observed behaviour; the auto-correlation intervention cannot. The Bayes factor of $B = 1096$ also provides extremely strong evidence in favour of the mean intervention model, i.e. the observed data are over 1000 times more likely to have arisen from the mean intervention model. We conclude that the most likely explanation for the observed behaviour of the NAO index is a transient change in the mean level during the extended winter (December–March) season. The remainder of our analysis focuses on interpreting only the mean intervention model.

Surprisingly for such a complex phenomenon, the mean, trend and seasonal components of the NAO index show very limited evidence of non-stationarity. Fig. 5 shows some posterior trajectories $\theta_{1:T}^{(j)}$ from each component. There is evidence of a fairly constant trend leading to a reduction in the mean level of the NAO of around 0.8 hPa between 1979 and 2017. The posterior distribution of the trend itself suggests that the rate of decrease in the NAO mean peaked around 1993–1994 at around 0.03 hPa year^{-1} ($-0.07, 0.01$), since when the trend has gradually weakened. The amplitudes of the annual and semiannual cycles are almost constant (likewise the phases). The 0.95-quantile of the posterior distribution of the mean evolution standard deviation $\sqrt{W_\mu}$ is 0.005 hPa, so changes in excess of 0.2 hPa year^{-1} to the mean and seasonal components are not ruled out under the random-walk hypothesis. There is no evidence of non-stationarity in the auto-regressive coefficients $\phi_1, \ldots, \phi_5$ which are effectively constant throughout the study period. This suggests that the day-to-day variation in the NAO can be
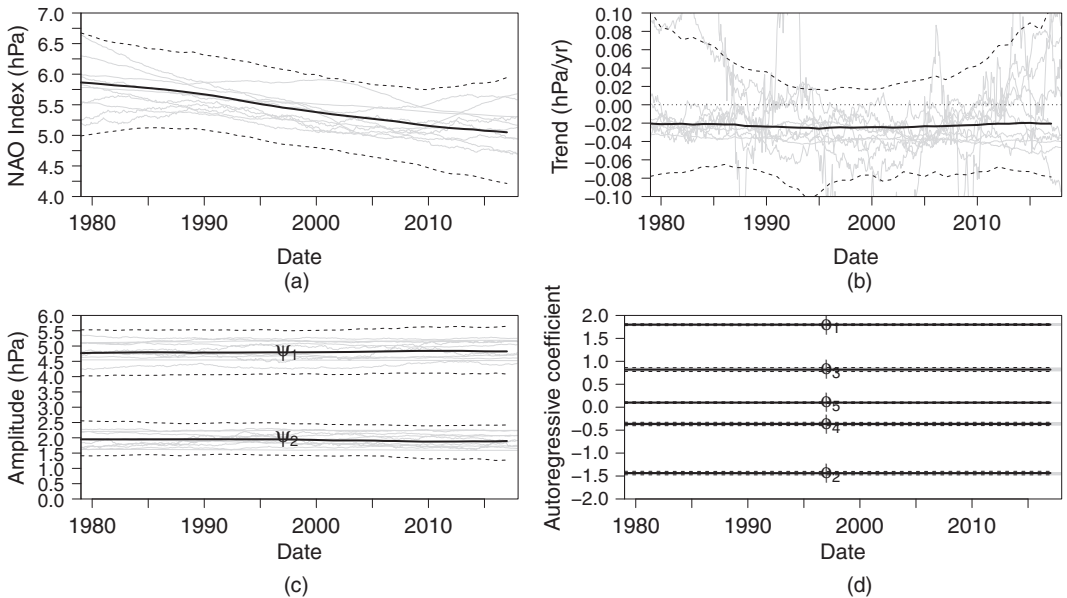
Fig. 5. Posterior distributions of model components (———, pointwise posterior mean; - - - - - -, pointwise 90% credible intervals; ———, random sample of 10 trajectories $\theta_{1:T}^{(j)} | \Phi^{(j)}, Y_{1:T}$): (a) mean μ_t; (b) trend β_t; (c) amplitudes of seasonal harmonics ψ_1 and ψ_2; (d) TVAR coefficients $\phi_{1t}, \ldots, \phi_{5t}$

Table 4. Analysis of variance†

Season	Mean	Coupled component	Irregular component	Error
Winter (December–January–February)	0.00 (0.00,0.05)	0.66 (0.52,0.77)	0.33 (0.23,0.47)	0.00 (0.00,0.00)
Summer (June–July–August)	0.15 (0.06,0.22)	0.00 (0.00,0.00)	0.85 (0.78,0.94)	0.00 (0.00,0.00)

†Values in parentheses indicate 90% credible intervals.

adequately represented by an auto-regressive process rather than a TVAR process. However, this is a useful conclusion given the observed seasonal auto-correlation structure in Fig. 1.

6.1. Quantifying the effect of coupling

The posterior mean estimate of the intervention effect standard deviation $\sqrt{W_{\delta_\mu}}$ is 0.43 (0.33–0.53), indicating a very active process, contributing substantial additional interannual variability during the extended winter season (December–March). Table 4 contains the results of the analysis of variance that was proposed in Section 5 for the mean intervention model M_μ. The effect of coupling δ_t explains around 66% of the observed variation in the winter (December–January–February) means. Accumulated short-term variability that is captured by the TVAR residuals X_t explains around 33% of the interannual variability. Despite the trend that is visible in Fig. 5, the contribution of the mean and seasonal components is negligible. Together they account for a maximum of 5% of the interannual variability in winter (December–January–February). In contrast, the mean and seasonal components account for around 15% of interannual variability
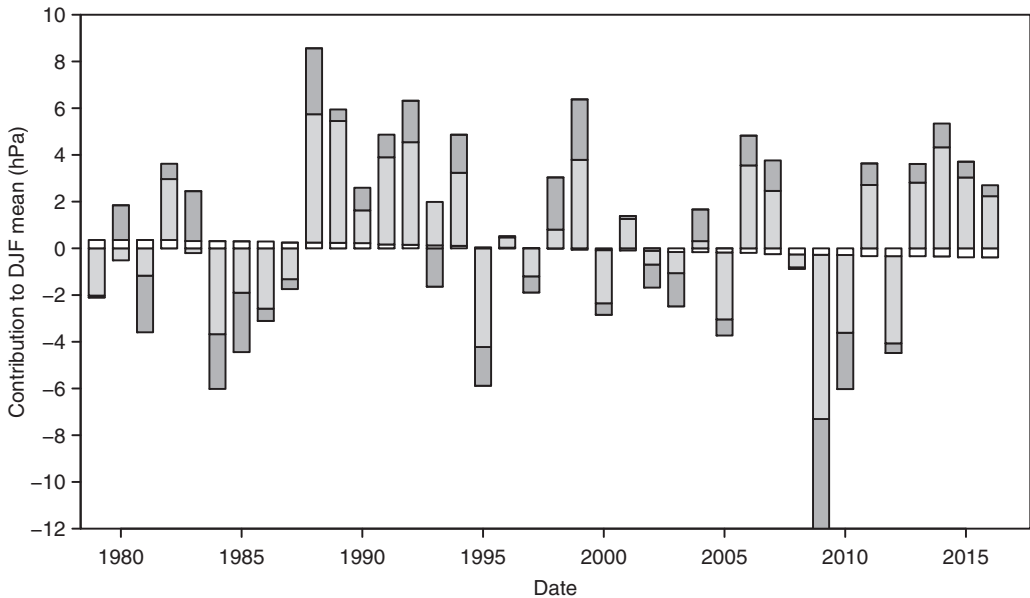
Fig. 6. Contribution of individual model components: posterior mean estimates of the winter (December–January–February) mean levels of the systematic component $\bar{\eta}$ (□), the irregular component $\bar{X}$ (▨), the coupling effect δ (▨) and the observation error $\bar{v}$ (■)

in summer (June–July–August) when coupling has no effect and the day-to-day variability is reduced. The contribution of measurement error is negligible.

Fig. 6 shows the posterior mean contribution of each component to each observed winter (December–January–February) mean level. This is an important and useful advance over existing methods in climate science that estimate only the fraction of total variance explained by each component. The weak negative trend in the mean component μ_t is clearly visible. Both the absolute and the relative contributions of the irregular component X_t and the coupling effect δ_t vary between years, but both components usually have the same sign. This is a product of the limited data that are available to estimate the components during each extended winter (December–March). If the coupling signal cannot be clearly identified during a particular season, then the contribution to the seasonal mean will be split approximately according to the analysis of variance in Table 4 and the two components will have the same sign. The fact that the relative contribution of each component varies widely in Fig. 6 indicates that the model can separate the coupling effect from the noise of the irregular component.

6.2. Forecasting the winter North Atlantic oscillation

The posterior mean estimate of the coupling effect coefficient φ_μ is 0.994 (90% credible interval 0.991–0.997). In terms of interannual variability, this is equivalent to a correlation of around 0.19 (0.05–0.38) between December–January–February means, suggesting limited evidence of persistence and therefore predictability between seasons. However, if we can learn about the coupling effect sufficiently quickly during a specific coupled period, then we can use that knowledge to provide more skilful forecasts for the rest of the period. Fig. 3 suggests that the system is at least partially coupled from the beginning of November until around the middle of April. Using the forecasting recursions in Appendix A, we obtained forecasts beginning each day from
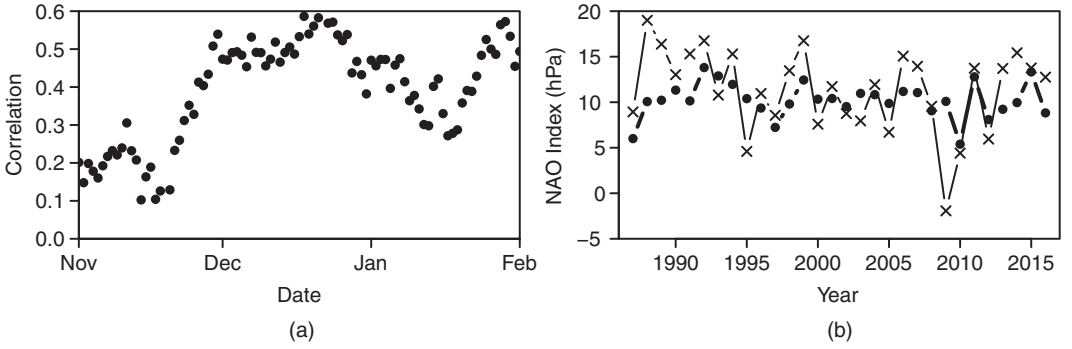
Fig. 7. Predictability of winter (December–January–February) NAO: (a) correlation between the observations and forecasts initialized on each day between November and February, for the mean level over the remainder of the period to February 28th; (b) observations (x) and forecasts (●), for the mean NAO between December 1st and February 28th each year, given data up to November 30th
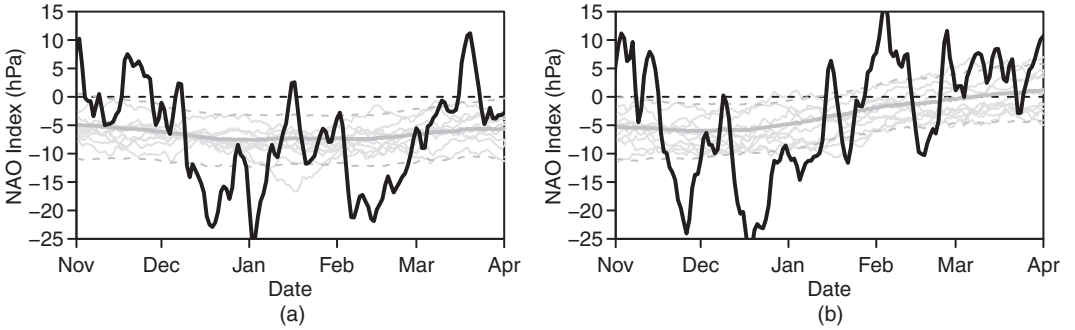
Fig. 8. Deseasonalized observations $Y_t - E(\eta_t \mid Y_{1:T})$ for the winters (December–January–February) of (a) 2009–2010 and (b) 2010–2011: ———, random sample of 10 posterior trajectories for the coupling effect $\delta_t^{(j)} \mid \Phi^{(j)}, Y_{1:T}$; ———, posterior mean for the coupling effect; − − −, pointwise 90% credible interval for the coupling effect

November 1st to February 1st until the end of the fully coupled period on February 28th for every winter between 1987 and 2016. Fig. 7 shows the correlation between the forecast and observed means. By the beginning of December, the correlation approaches 0.5 for the 92-day forecast of the mean NAO to February 28th. This correlation approaches that achieved by computationally expensive numerical weather prediction models (Scaife *et al.*, 2014; Siegert *et al.*, 2016). The correlation increases slightly as more observations are assimilated during December. However, as more observations are assimilated, the forecast period decreases and we are essentially predicting weather noise, so the correlation does not increase further.

Fig. 7 also compares forecasts of the 92-day December–January–February winter mean, initialized on December 1st each year, with the observed mean NAO index for the same periods. The model predicts the 2010, 2011 and 2012 winter seasons with remarkable accuracy and captures the general pattern during the 1990s. However, it fails to predict the extreme winter of 2009–2010. Fig. 8 plots deseasonalized observations of winter 2009–2010 ($Y_t - E(\eta_t \mid Y_{1:T})$). Deseasonalizing the observations leaves only the contributions from the irregular component X_t and the coupling effect δ_t, which represent processes on different timescales. The irregular component X_t captures high frequency fluctuations, whereas the coupling effect δ_t captures any overall departure from the seasonal mean. From the middle of December onwards, the mean

of the deseasonalized data is clearly negative, which the model attributes to the coupling effect δ_t. Since the seasonal forecasts in Fig. 7 were based on information up to November 30th, it is unsurprising that a fairly normal winter was forecast. In contrast, in winter 2010–2011 (Fig. 8), a strong negative signal is visible in November which the model can exploit to forecast skilfully the remainder of the December–January–February season. Winter 2010 also illustrates the time varying nature of the coupling effect δ_t, which starts strongly negative early in the season but weakens from mid-January onwards.

7.   Discussion

In this study we have developed Bayesian state space methods for diagnosing predictability in intermittently coupled systems. Coupling is represented by a transient intervention whose timing and duration are inferred from the data. Interventions to either the mean or temporal dependence structure are considered. The effect of intermittent coupling is modelled as a dynamic process rather than a sequence of constant and independent effects. Latent TVAR components are proposed to capture any inherent non-stationarity in the temporal dependence structure. A linearized approximation is proposed that allows efficient forward filtering and backward sampling for models containing latent TVAR components, without requiring complicated and computationally expensive sequential Monte Carlo methods. In addition, we develop tools for posterior inference in intermittently coupled systems, including evaluating the evidence of a coupling effect, attribution of historical variation in the system and demonstrating potential predictability.

We applied the proposed model and inference methods to diagnose excess winter time variability in the NAO. Existing methods in climate science cannot distinguish between transient changes in the mean or temporal dependence structure. The model strongly points to transient changes in the mean level of the NAO during a period beginning sometime in November and ending around the middle of April. This is an important conclusion given that the excess winter time variability in the NAO is usually characterized by increased temporal dependence. The mean level of the NAO also appears to change on decadal timescales, in addition to a fairly stable annual cycle and the transient changes in winter time. The model can also separate the coupling effect from accumulated day-to-day variability in individual seasons. For the NAO, the two effects actually oppose each other in some seasons.

Like latent auto-regressive components, latent TVAR components improve the interpretability of structural time series models by avoiding the need to redefine the mean level of the observed process. In addition, latent TVAR components permit efficient recursive estimation of the auto-regressive parameters and include standard latent auto-regressive components as a special case when the evolution variance is 0. For the NAO, we found little evidence of changes in the auto-regressive structure throughout the study period, so a standard latent auto-regressive component could be used to represent day-to-day variability. However, the fact that we can confirm that auto-regressive structure is constant on decadal timescales is also a useful conclusion.

The model proposed for intermittently coupled systems differs from standard intervention analysis by modelling the effect of repeated coupling events as a dynamic process, rather than a series of independent events. This allows knowledge that is gained during one coupled period to inform inferences for the next. By modelling the coupling effect as a dynamic process, the effect can also vary within individual coupled periods rather than being assumed constant. Climate scientists usually assume that any coupling effect is constant throughout an arbitrarily defined season. We have shown that the coupling effect on the NAO can vary substantially throughout a single season.

Modelling the effect of coupling as a dynamic process also makes the model robust to minor variations in the timing and duration of the coupled period. However, the specification of a fixed coupling period remains a limitation. Hidden Markov and semi-Markov models are widely used in similar seasonal state switching scenarios to allow for changes in onset and duration (e.g. Carey-Smith *et al.* (2014)). Standard hidden Markov models assume instantaneous switching between states. Although such an assumption may be acceptable for some applications, we do not consider it plausible for the NAO. A completely general alternative would be a reversible jump Markov chain Monte Carlo scheme (Green, 1995). In such a scheme, coupling events could be estimated with varying onset, duration or other parameterized structural changes. However, unless the timing of coupling events varies dramatically, the additional cost and complexity of a reversible jump scheme seems unnecessary. The on-line Bayesian change point methods that were proposed by Fearnhead and Liu (2011) might provide a more efficient approach.

In the methodology that was developed here, we have allowed for non-stationarity in the mean and the temporal dependence structure, but not in the variance. Stochastic volatility models and related auto-regressive conditional heteroscedasticity and generalized auto-regressive conditional heteroscedasticity models have been widely studied and applied, particularly in economics. Masala (2015) applied a generalized auto-regressive conditional heteroscedasticity model to stochastic modelling of the NAO but found that its performance was poor. Efficient filtering and smoothing is possible for time varying observation error variance (West and Harrison (1997), chapter 10.8). However, fully conjugate models that admit analytic filtering and smoothing for time varying state evolution variances are not possible, even in the linear normal case. Of particular interest are changes in the residual TVAR evolution variance W_{Xt} that drives short-term variability in the system. Sequential Monte Carlo methods or further approximations are required to model time varying evolution variances.

Acknowledgements

Appendix A: Forward filtering–backward sampling and forecasting

A.1. Forward filtering

The sequence of posterior distributions $\{\boldsymbol{\theta}_t \mid Y_{1:t}, \boldsymbol{\Phi} : t = 1, \ldots, T\}$ can be approximated as follows.

Before observing y_t, the predictive distributions at time $t-1$ are

$$Y_t \mid Y_{1:t-1}, \boldsymbol{\Phi} \sim N(f_t, Q_t),$$
$$\boldsymbol{\theta}_t \mid Y_{1:t-1}, \boldsymbol{\Phi} \sim N(\mathbf{a}_t, \mathbf{R}_t)$$

with

$$\mathbf{a}_t = g(\mathbf{m}_{t-1}, \mathbf{0}),$$
$$\mathbf{R}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t' + \mathbf{H}_t \mathbf{W}_t \mathbf{H}_t',$$
$$f_t = f(\mathbf{a}_t, 0),$$
$$Q_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + \mathbf{J}_t V_t \mathbf{J}_t',$$

where

$$\mathbf{G}_t = \left.\frac{\partial g}{\partial \boldsymbol{\theta}}\right|_{\hat{\boldsymbol{\theta}}_{t-1}, \hat{\mathbf{w}}_t},$$

$$\mathbf{H}_t = \left.\frac{\partial g}{\partial \mathbf{w}}\right|_{\hat{\boldsymbol{\theta}}_{t-1}, \hat{\mathbf{w}}_t},$$

$$\mathbf{F}_t = \left.\frac{\partial f}{\partial \boldsymbol{\theta}}\right|_{\hat{\boldsymbol{\theta}}_t, \hat{v}_t},$$

$$\mathbf{J}_t = \left.\frac{\partial f}{\partial v}\right|_{\hat{\boldsymbol{\theta}}_t, \hat{v}_t}.$$

After observing Y_t, the posterior distribution of the state vector at time t is

$$\boldsymbol{\theta}_t \mid Y_{1:t}, \boldsymbol{\Phi} \sim N(\mathbf{m}_t, \mathbf{C}_t)$$

with

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t e_t,$$
$$\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t Q_t \mathbf{A}_t'$$

where $e_t = Y_t - f_t$ and $\mathbf{A}_t = \mathbf{R}_t \mathbf{F}_t / Q_t$.

A.2. Backward sampling

The joint posterior $\boldsymbol{\theta}_{1:T} \mid Y_{1:T}, \boldsymbol{\Phi}$ can be sampled recursively as follows.

(a) Sample $\boldsymbol{\theta}_T^{(j)}$ from $\boldsymbol{\theta}_T \mid Y_{1:T}, \boldsymbol{\Phi} \sim N(\mathbf{m}_T, \mathbf{C}_T)$.
(b) For $k = 1, \ldots, T - 1$: sample $\boldsymbol{\theta}_{T-k}^{(j)}$ from $\boldsymbol{\theta}_{T-k}^{(j)} \mid \boldsymbol{\theta}_{T-k+1}^{(j)}, Y_{1:T}, \boldsymbol{\Phi} \sim N\{\mathbf{h}_T(k), \mathbf{H}_T(k)\}$ where

$$\mathbf{h}_T(k) = \mathbf{m}_{T-k} + \mathbf{B}_{T-k}(\boldsymbol{\theta}_{T-k+1}^{(j)} - \mathbf{a}_{T-k+1}),$$
$$\mathbf{H}_T(k) = \mathbf{C}_{T-k} - \mathbf{B}_{T-k}\mathbf{R}_{T-k+1}\mathbf{B}_{T-k}'$$

and $\mathbf{B}_{T-k} = \mathbf{C}_{T-k}\mathbf{G}_{T-k+1}'\mathbf{R}_{T-k+1}^{-1}$. The quantities $\mathbf{m}_t$, $\mathbf{C}_t$, $\mathbf{a}_t$, $\mathbf{R}_t$ and $\mathbf{G}_t$ are obtained from the filtering recursions.

A.3. Forecasting

The k-step-ahead forecast distribution given data up to time t can be sampled sequentially as follows.

(a) Sample $\boldsymbol{\theta}_t^{(j)}$ from $\boldsymbol{\theta}_t \mid Y_{1:t}, M \sim N(\mathbf{m}_t, \mathbf{C}_t)$.
(b) For $i = 1, \ldots, k$:
  (i) sample $\boldsymbol{\theta}_{t+i}^{(j)}$ from $g(\boldsymbol{\theta}_{t+i-1}^{(j)}, \mathbf{w}_{t+i})$;
  (ii) sample $Y_{t+i}^{(j)}$ from $f(\boldsymbol{\theta}_{t+i}^{(j)}, v_{t+i})$.

References

Altizer, S., Bartel, R. and Han, B. A. (2011) Animal migration and infectious disease risk. *Science*, **331**, 296–302.
Bourassa, M. A., Gille, S. T., Bitz, C., Carlson, D., Cerovecki, I., Clayson, C. A., Cronin, M. F., Drennan, W. M., Fairall, C. W., Hoffman, R. N., Magnusdottir, G., Pinker, R. T., Renfrew, I. A., Serreze, M. C., Speer, K., Talley, L. D. and Wick, G. A. (2013) High-latitude ocean and sea ice surface fluxes: challenges for climate research. *Bull. Am. Meteorol. Soc.*, **94**, 403–423.
Box, G. E. P. and Tiao, G. C. (1975) Intervention analysis with applications to economic and environmental problems. *J. Am. Statist. Ass.*, **70**, 70–79.

Briers, M., Doucet, A. and Maskell, S. (2010) Smoothing algorithms for state-space models. *Ann. Inst. Statist. Math.*, **62**, 61–89.

Carey-Smith, T., Sansom, J. and Thomson, P. (2014) A hidden seasonal switching model for multisite daily rainfall. *Wat. Resour. Res.*, **50**, 257–272.

Chapin III, F. S., Sturm, M., Serreze, M. C., McFadden, J. P., Key, J. R., Lloyd, A. H., McGuire, A. D., Rupp, T. S., Lynch, A. H., Schimel, J. P., Beringer, J., Chapman, W. L., Epstein, H. E., Euskirchen, E. S., Hinzman, L. D., Jia, G., Ping, C.-L., Tape, K. D., Thompson, C. D. C., Walker, D. A. and Welker, J. M. (2010) Role of land-surface changes in Arctic summer warming. *Science*, **657**, 9–13.

Chen, G., Qian, C. and Zhang, C. (2012) New insights into annual and semiannual cycles of sea level pressure. *Mnthly Weath. Rev.*, **140**, 1347–1355.

Chopin, N., Jacob, P. E. and Papaspiliopoulos, O. (2013) SMC^2: an efficient algorithm for sequential analysis of state space models. *J. R. Statist. Soc.* B, **75**, 397–426.

Croston, J. D. (1972) Forecasting and stock control for intermittent demands. *Oper. Res. Q.*, **23**, 289–303.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553–597.

Doucet, A. and Johansen, A. M. (2011) A tutorial on particle filtering and smoothing: fifteen years later. In *The Oxford Handbook of Nonlinear Filtering* (eds D. Crisan and B. Rozovskii), ch. 8.2, pp. 656–705. Oxford: Oxford University Press.

Durbin, J. and Koopman, S. J. (2012) *Time Series Analysis by State Space Methods*, 2nd edn. Oxford: Oxford University Press.

Eade, R., Smith, D. M., Scaife, A. A., Wallace, E., Dunstone, N. J., Hermanson, L. and Robinson, N. (2014) Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys. Res. Lett.*, **41**, 5620–5628.

Fearnhead, P. and Liu, Z. (2011) Efficient Bayesian analysis of multiple changepoint models with dependence across segments. *Statist. Comput.*, **21**, 217–229.

Franzke, C. L. E. and Woollings, T. (2011) On the persistence and predictability properties of North Atlantic climate variability. *J. Clim.*, **24**, 466–472.

Frühwirth-Schnatter, S. (1994) Data augmentation and dynamic linear models. *J. Time Ser. Anal.*, **15**, 183–202.

Godsill, S. J., Doucet, A. and West, M. (2004) Monte Carlo smoothing for nonlinear time series. *J. Am. Statist. Ass.*, **99**, 156–168.

Green, P. J. (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**, 711–732.

Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., Leslie, D. S., Forster, J. J., Wagenmakers, E. J. and Steingroever, H. (2017) A tutorial on bridge sampling. *J. Math. Psychol.*, **81**, 80–97.

Harvey, A. C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Harvey, A. C. and Durbin, J. (1986) The effects of seat belt legislation on British road casualties: a case study in structural time series modelling (with discussion). *J. R. Statist. Soc.* A, **149**, 187–227.

Hsu, C.-P. F. and Wallace, J. M. (1976) The global distribution of the annual and semiannual cycles in sea level pressure. *Mnthly Weath. Rev.*, **104**, 1597–1601.

Hurrell, J. W. (1995) Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science*, **269**, 676–679.

Kass, R. E. and Raftery, A. E. (1995) Bayes factors. *J. Am. Statist. Ass.*, **90**, 773–795.

Keeley, S. P. E., Sutton, R. T. and Shaffrey, L. C. (2009) Does the North Atlantic Oscillation show unusual persistence on intraseasonal timescales? *Geophys. Res. Lett.*, **36**, L22706.

Kitagawa, G. and Gersch, W. (1985) A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series. *IEEE Trans. Autom. Control*, **30**, 48–56.

Kushnir, Y., Robinson, W. A., Chang, P. and Robertson, A. W. (2006) The physical basis for predicting Atlantic sector seasonal-to-interannual climate variability. *J. Clim.*, **19**, 5949–5970.

Masala, G. (2015) North Atlantic Oscillation index stochastic modelling. *Int. J. Climtol.*, **35**, 3624–3632.

Mosedale, T. J., Stephenson, D. B., Collins, M. and Mills, T. C. (2006) Granger causality of coupled climate processes: ocean feedback on the North Atlantic Oscillation. *J. Clim.*, **19**, 1182–1194.

Olsen, B., Munster, V. J., Wallensten, A., Waldenström, J., Osterhaus, A. D. M. E. and Fouchier, R. A. M. (2006) Global patterns of Influenza A Virus in wild birds. *Science*, **312**, 384–388.

Prado, R. and West, M. (1997) Exploratory modelling of multiple non-stationary time series: latent process structure and decompositions. In *Modelling Longitudinal and Spatially Correlated Data* (eds T. G. Gregoire, D. R. Brillinger, P. J. Diggle, E. Russek-Cohen, W. G. Warren and R. D. Wolfinger), pp. 349–361. New York: Springer.

Prado, R. and West, M. (2010) *Time Series: Modelling, Computation and Inference*. Boca Raton: Chapman and Hall–CRC.

Scaife, A. A., Arribas, A., Blockey, E., Brookshaw, A., Clark, R. T., Dunstone, N. J., Eade, R., Fereday, D. R., Folland, C. K., Gordon, M., Hermanson, L., Knight, J. R., Lea, D. J., MacLachlan, C., Maidens, A., Martin, M., Peterson, K. A., Smith, D. M., Vellinga, M., Wallace, E., Waters, J. and Williams, A. (2014) Skillful long range prediction of European and North American winters. *Geophys. Res. Lett.*, **5**, 2514–2519.

Shenstone, L. and Hyndman, R. J. (2005) Stochastic models underlying Croston's method for intermittent demand forecasting. *J. Forecast.*, **24**, 389–402.

Siegert, S., Stephenson, D. B., Sansom, P. G., Scaife, A. A., Eade, R. and Arribas, A. (2016) A Bayesian framework for verification and recalibration of ensemble forecasts: how uncertain is NAO predictability? *J. Clim.*, **29**, 995–1012.

Subba Rao, T. (1970) The fitting of non-stationary time-series models with time-dependent parameters. *J. R. Statist. Soc.* B, **32**, 312–322.

Walker, G. T. (1924) Correlation in seasonal variation of weather, IX, a preliminary study of world weather. *Mem. Ind. Meteorol. Dept.*, **24**, 275–332.

West, M. and Harrison, P. J. (1997) *Bayesian Forecasting and Dynamic Models*, 2nd edn. New York: Springer.