Dynamic Harmonic Regression on a Natural Domain:

Modeling Pollution in French Rivers

Joshua R. Christensen

A project submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Masters of Science

Matthew Heiner, Chair
Philip White
Shannon Tass

Department of Statistics

Brigham Young University

Dynamic Harmonic Regression on a Natural Domain: Modeling Pollution in French Rivers

Joshua R. Christensen
Department of Statistics
Masters of Science

BYU Department of Statistics

Abstract

In this project we focus on fitting dynamic harmonic regression (DHR) models to time series of log-nitrate levels in French rivers. The time series display cyclical components with time-varying characteristics which makes them good candidates for DHR models. Our DHR model can accept irregularly sampled data or equally spaced data and models phases and amplitudes directly instead of using Fourier components. Since we model phases and amplitudes directly, our model includes non-Gaussian dynamic linear models. We develop a novel sampling approach to retrieve posterior samples from our model efficiently.

Keywords: dynamic regression, dynamic linear models, slice sampling

Table of Contents

List of Figures

List of Tables

1 Introduction

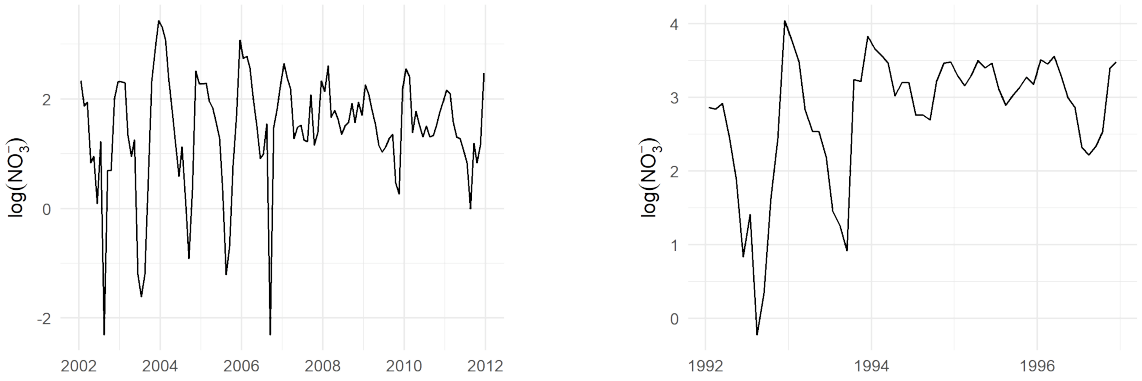
The focus of the project is a data set containing nitrate pollution time series measured at river monitoring stations across France. In this introductory section we will outline the importance of nitrate pollution, the features of our data, and our methods for addressing these data features.

1.1 The Importance of Nitrate Pollution

Nitrate pollution is of interest to hydrology researchers because of its direct contribution to eutrophication and the subsequent creation of hypoxic dead zones (Le Moal et al., 2019; Diaz and Rosenberg, 2008). This process begins with an excess of nutrients available in a water source, often due to pollution from fertilizers. The presence of excess nutrients leads to large blooms of algae. Algae produces oxygen when it is on the surface of water, but begins to consume oxygen when it no longer has sufficient access to sunlight. When too much algae is produced in response to excess nutrients, it prevents sunlight from reaching photosynthesizers deeper underwater, and also leads to some of the algae consuming oxygen as it is pushed away from the surface. At this point oxygen is leaving the water on average as most of the photosynthesis in the system is occurring at the surface. As algae that is pushed away from sunlight begins to die, it is decomposed and metabolized by other organisms, further consuming oxygen and exacerbating the hypoxia of the system. Eventually this cycle leads to the creation of a hypoxic dead zone, where the oxygen content of the water has reached a level too low to support the ecosystem. For a quick review of eutrophication we recommend EarthHow (2022) or U.S. Geological Survey (2019).

Eutrophication already affects approximately two thirds of fresh and estuarine waters worldwide (Dupas et al., 2019). In order to develop and encourage effective interventions to this phenomenon, we must be able to accurately measure the nitrate levels of a given ecosystem. However, while measuring nitrate level at a given point in time is relatively easy, strong seasonal effects make it difficult to extract meaning from isolated measurements (Abbott et al., 2018).

Since a single measurement in time does not provide much information on the underlying pollution state of a given river, it would be useful to have a model that could indicate how the trend and seasonal cycles are likely to behave at a given point in time. This would allow



(a) (b)

Figure 1.1: Examples of two time series showing the log concentration of nitrate in French rivers.

for better monitoring of both the underlying average and the potential peaks of the nitrate levels in the river. Having this clarity would make it easier to identify when rivers are at dangerous levels of nitrate concentration that could lead to eutrophication.

1.2 Data Introduction

Our data set represents time series of nitrate levels from 4,522 monitoring stations across France. Each data point includes an identifier for the station and the day of the measurement. Two examples of time series can be found in Fig. 1.1. Note that we will generally view these concentrations on the log scale which makes the time series less erratic.

One feature of the time series is a strong annual cycle. While the presence of the cycle is clear, the difficulty in accounting for this annual cycle comes from the fact that in many rivers both the underlying average and the amplitude of the cycle change simultaneously through time. This can make it difficult to parse when the underlying trend is moving, when the amplitudes are moving and when we simply have expected cyclical movement or noise causing the shift.

Many of the rivers also display behaviors consistent with system shocks or change points. Fig. 1.1a shows an example where both the level and amplitude seem to change drastically at a specific time point. This could be the result of some intentional or unintentional human intervention, a change in measurement technique or instrument, etc. We will not try to address this data feature in our analysis, but we note it here and will address its effects on model performance in Section 4.1.

The nitrate concentration time series are also sampled irregularly. Some time series have nearly perfect monthly or even weekly data. Others have sporadic yearly data or contain a seemingly random assortment of measurements in a given year. Even among the more consistent time series, variations in sampling frequency are common as policies and funding change over time. This complicates the fitting of many time series models that rely on a

consistent discrete time basis.

We will outline a model that addresses the problems of varying midpoints and amplitudes in the seasonal trends. Our model will also account for the irregular sampling of our data.

1.3 Methods Introduction

Time series exhibiting strong seasonality are often fit using harmonic regression or Fourier regression. In this setup Fourier terms for each desired frequency are included as explanatory variables and then the Fourier coefficients are estimated via regression. The model for a typical harmonic regression including only one frequency and no other explanatory variables is

$$y_t = \beta_0 + a \sin\left(\frac{2\pi t}{\lambda}\right) + b \cos\left(\frac{2\pi t}{\lambda}\right) + \epsilon_t \tag{1.1}$$

where β_0 represents the intercept, a and b are Fourier regression coefficients, λ represents the period and ϵ_t is the error term, which will typically be defined as $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$. From the Fourier coefficients one can easily obtain the corresponding estimated phase and amplitude of the given frequency with simple transformations. We can obtain the phase ϕ by $\phi = \arctan(-a/b)$ and the amplitude α by $\alpha = \sqrt{a^2 + b^2}$. This approach is simple and effective. It does not require any change to the typical regression format where coefficients are unconstrained. When the unconstrained coefficients are transformed to obtain the phase and amplitude the restrictions are imposed implicitly by the ranges of the transformation functions. The main drawback of the Fourier-based harmonic regression is that the parameters in question (the Fourier coefficients) do not have an obvious interpretation in their own right. This is largely inconsequential in the case of typical regression estimated with least-squares, but quickly becomes a problem once Bayesian priors and time-varying parameters enter the model.

Prior specification is an important part of any Bayesian analysis. It would be difficult for a statistician or a hydrologist to specify a meaningful prior on a Fourier coefficient because it has no obvious interpretation when viewed individually and without transformation. Reverse engineering priors from those believed to exist on the phase and amplitude is possible, but introduces an extra layer of abstraction and difficulty in prior specification that is best avoided if possible. Related work in interpretable prior specification for time series has been done by Huerta and West (1999) with applications to ARMA models for periodic or quasiperiodic time series.

Fourier terms also introduce problems when considered in the context of dynamic linear models (DLMs) where we allow regression coefficients to change through time. Since the phase and amplitude in a typical harmonic regression are both transformations of the Fourier coefficients, it is essentially impossible to make modifications to the modeling of one without modifying the other. In the context of DLMs this would mean that if we would like amplitudes to vary with time we would also have to let the phases vary with time. In some

scenarios this is not a problem, but there are many others where it would be reasonable to assume that amplitudes are time-varying while phases remain fixed.

In order to address the shifting amplitudes and trends of the annual cycles in our data we will use a Bayesian dynamic harmonic regression model (DHR). The DHR model extends a typical harmonic regression by allowing the amplitude and phases of the sinusoidal functional bases to vary through time as in a typical DLM. The flexibility of the amplitudes within the DHR framework will allow us to account for the variation in cyclical behavior displayed by the river data we wish to fit.

The DHR model has typically been formulated using Fourier terms (Young et al., 1999; Mindham and Tych, 2019). To avoid the problems mentioned in the preceding paragraphs we have created a model that allows us to directly model phases and amplitudes. This model implies restricted support for our DLM, since amplitudes must be constrained to be positive in order for the model to be identifiable. The most common Markov Chain Monte Carlo (MCMC) scheme for DLMs is the forward filtering, backward sampling (FFBS) algorithm which requires joint multivariate normality in the parameters of the DLM. Since our model requires constrained support we devise our own efficient posterior sampling scheme to accompany our model. We use a novel slice sampling method with a FFBS proposal mechanism to sample efficiently from the posterior distribution.

Another feature of our model is that it allows for data that were sampled at irregular time intervals. This was made possible by returning to the theory of the continuous-time analog to the random walk: the Wiener process or Brownian motion. We will refer to the Wiener process using the notation W_t. The Wiener process is a continuous-time stochastic process that has independent normal increments. This means that $W_t - W_s$ for $s < t$ is independent of all values prior to W_s. Each increment is also distributed normally with mean zero and variance equal to the length of the increment. In other words $W_t - W_s \sim \mathcal{N}(0, t - s)$. Since the discrete-time DLM is based on a random walk, we base our continuous-time DLM on the Wiener process and scale the process variance between adjacent time points by the time elapsed between measurements. This is a natural and intuitive solution that is also simple to implement.

2 Literature Review

In this literature review we will provide a brief overview of literature related to the project. In Section 2.1 we will include a brief description of literature supporting the relevance of the problem to hydrologists. In Section 2.2 we will outline developments in DHR and how they are relevant to the construction of our model, as well as highlight differences between our model and previous models. A brief review of the FFBS algorithm (Frühwirth-Schnatter, 1994; Carter and Kohn, 1994) will then follow in Section 2.3 as it will play a central role in our posterior sampling strategy. Finally, in Section 2.4 we will provide references to other approaches for dealing with irregularly sampled data and constructing DLMs on the continuous-time scale.

2.1 Hydrology

The process of eutrophication and its eventual result of hypoxic dead zones is briefly described in Section 1.1. The impact of dead zones has been increasing in terms of both number of systems affected and geographical area affected at a staggering pace since the 1960s. Hundreds of systems and hundreds of thousands of square kilometers are now affected by dead zones resulting from eutrophication (Diaz and Rosenberg, 2008; Cheng et al., 2020). While the excess nutrients that cause eutrophication and hypoxia are a threat to ecosystems and biodiversity, it is important to note that they also pose a threat to human food and water sources (Abbott et al., 2018).

In fact, the significance of the water pollution described in this project also has very real and current impacts in addition to existential threats. Human health and economic well-being are at risk, with estimates of economic impact ranging from hundreds of billions to trillions of US dollars from nitrate alone (Dupas et al., 2019). The detailed review of eutrophication undertaken by Le Moal et al. (2019) confirms all of the above-mentioned risks, noting that in addition to the more obvious and quantifiable risks, ecosystem degradation nearly always has more far-reaching effects than can be fully accounted for in a simple cost analysis.

2.2 Dynamic Harmonic Regression (DHR)

The typical DHR model is a generalization of typical harmonic regression based on Fourier terms in which the Fourier coefficients are allowed to vary through time. This form of DHR is a subset of typical Gaussian DLMs and can be estimated in the Bayesian paradigm (West, 1995; West and Harrison, 1997; Prado and West, 2010), using spectral decomposition-based loss functions or using more traditional Kalman Filter methods, potentially with smoothing algorithms (Young et al., 1999).

Extensions to DHR or cyclical DLMs have included estimating periods for unknown wavelengths (West, 1995), incorporating change points where the evolution of the state is not smooth (Chappell and Tych, 2012) and irregular sampling adjustments (Mindham and Tych, 2019). DHR has been specifically recommended for some pollution applications by the United Nations Environment Programme or UNEP (Mindham and Tych, 2019) and continues to be used in a variety of complex forecasting and measurement settings (Buckley et al., 2021; Zavala and Messina, 2016).

DHR in all of the referenced cases except for West (1995) is based on Fourier terms, and in the excepted paper both amplitudes and phases are still modeled as time-varying. This will differ from our model which will allow for static phases as the amplitudes vary through time.

One last important note from the DHR literature is that the term dynamic harmonic regression has at least one other meaning within the statistical community: a typical harmonic regression with ARIMA residuals. This is important to note in order to avoid confusion in software implementations and understanding methods for different papers.

2.3 Forward Filtering, Backward Sampling (FFBS)

The FFBS algorithm of Frühwirth-Schnatter (1994) and Carter and Kohn (1994) represents a landmark achievement in efficient sampling from DLM posterior distributions. The FFBS algorithm allows us to efficiently perform block updates for all the time-varying parameters (TVPs) in a DLM at once, provided that the TVPs are jointly multivariate normal. This is done by sampling from a string of conditional distributions which can be found in simple forms using the Markov property.

In describing the algorithm we will use y_t to represent the observation at time t and $\boldsymbol{\theta}_t$ to represent the vector of TVPs at time t. We will also use the colon (:) notation to indicate that a vector contains all indices between the two listed. For example $\boldsymbol{y}_{2:4}$ would represent all observations from time point 2 to time point 4. We will also use T to represent the final time point in our time series.

The forward-filtering portion of the FFBS algorithm is done by the Kalman filter (Kalman, 1960). The Kalman filter recursively finds conditional distributions for each $\boldsymbol{\theta}_t$ given all of the observed data points up through time t. Subsequently only $\boldsymbol{\theta}_T$, the vector of parameters

at the final time point, has a true posterior conditioned on all of the data. We wish to have every $\boldsymbol{\theta}_t$ conditioned on all of the data $\boldsymbol{y}_{1:T}$, rather than just the portion of the data $\boldsymbol{y}_{1:t}$. The backward sampling portion of the algorithm will use the information from the Kalman filter to achieve this full conditioning.

The backward sampling algorithm of the full joint distribution $p(\boldsymbol{\theta}|\boldsymbol{y})$ begins with decomposition into the series of conditionals:

$$p(\boldsymbol{\theta}_{1:T}|\boldsymbol{y}_{1:T}) = p(\boldsymbol{\theta}_T|\boldsymbol{y}_{1:T}) \prod_{T-1}^{1} p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{(t+1):T}, \boldsymbol{y}_{1:T}) \tag{2.1}$$

using conditional probability rules. Then, using the Markov property, each of these conditional probabilities can be simplified further by

$$\begin{aligned} p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{(t+1):T}, \boldsymbol{y}_{1:T}) \propto & p(\boldsymbol{\theta}_t, \boldsymbol{\theta}_{(t+1):T}, \boldsymbol{y}_{(t+1):T}|\boldsymbol{y}_{1:t}) \\ = & p(\boldsymbol{\theta}_t|\boldsymbol{y}_{1:t}) p(\boldsymbol{\theta}_{t+1}|\boldsymbol{\theta}_t, \boldsymbol{y}_{1:t}) \\ & p(\boldsymbol{\theta}_{(t+2):T}|\boldsymbol{\theta}_t, \boldsymbol{\theta}_{t+1}, \boldsymbol{y}_{1:t}) p(\boldsymbol{y}_{(t+1):T}|\boldsymbol{\theta}_{t:T}, \boldsymbol{y}_{1:t}) \\ \propto & p(\boldsymbol{\theta}_t|\boldsymbol{y}_{1:t}) p(\boldsymbol{\theta}_{t+1}|\boldsymbol{\theta}_t, \boldsymbol{y}_{1:t}) \\ = & p(\boldsymbol{\theta}_t, \boldsymbol{\theta}_{t+1}|\boldsymbol{y}_{1:t}) \\ \propto & p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t+1}, \boldsymbol{y}_{1:t}). \end{aligned} \tag{2.2}$$

In moving from the second line to third we can remove the later factors by leveraging conditional independence. For the first of the two removed factors we can remove $\boldsymbol{\theta}_t$ from the conditioning because the Markov property indicates that $\boldsymbol{\theta}_{t+2}$ is conditionally independent of $\boldsymbol{\theta}_t$ given $\boldsymbol{\theta}_{t+1}$. For the second of the two removed factors we can remove $\boldsymbol{\theta}_t$ from the conditioning because each $y_{(t+1):T}$ is conditionally independent $\boldsymbol{\theta}_t$ given $\boldsymbol{\theta}_{(t+1):T}$. Since $\boldsymbol{\theta}_t$ is no longer part of these terms we can drop them using proportionality.

Notice that at the end of Eq. (2.2) we are now only conditioning on the observations up through time t. As a result, we can use the Kalman filter equations to obtain solutions for $p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t+1}, \boldsymbol{y}_{1:t})$. Our backward sampling procedure based on Eq. (2.1) will then begin by sampling $p(\boldsymbol{\theta}_T|\boldsymbol{y}_{1:T})$ which comes directly from the Kalman filter, before subsequently sampling each $p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t+1}, \boldsymbol{y}_{1:t})$ by using the draw from the next time point ahead. As demonstrated in Eq. (2.2), this chain then represents a draw from the desired joint distribution, $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:T})$.

2.4 Continuous-Time DLM

As mentioned in Section 1.2, the sampling of our data is irregular both within and between time series. If our model is to be flexible enough to fit different rivers in the data set without preprocessing or model tuning, we will need to be able to accept irregularly sampled data.

One way to deal with irregular data mentioned in Prado and West (2010) is to simply introduce missing values until there is a consistent discrete time scale for the entire time

series. DLMs include methods for dealing with these missing values, so it essentially turns the continuous time problem into a missing data problem. Unfortunately this only works if there is an obvious choice for the underlying time scale. For our data, where the irregularity is such that in some cases the sampled days seem almost entirely random across a year, we would have to introduce missing values at all other days of the year. For time series that were only sampled a few times a year, this would make the approach impractical.

One solution for developing a true continuous-time DLM is to follow the example of Mindham and Tych (2019) and consider Taylor Series approximations for time intervals not sampled, and then formulate transition matrices and evolution vectors based on the Taylor Series. While this is certainly a workable solution, it complicates the model formulation greatly compared to that of the discrete-time DLM. Another option would be to take the approach of Driver and Voelkle (2018) and return to the underlying stochastic differential equations describing the movement of the parameters through time, and base continuous updates on the solutions to these equations. Again, this solution is logical, but much more mathematically intensive than the discrete-time DLM description. Our proposed method will allow for a continuous-time DLM without requiring computational or conceptual complexity beyond that of the discrete-time DLM.

3 Methods

In this chapter we will describe the methods developed for this project. These will be subdivided into two sections. In Section 3.1 we will give a full specification for our model and highlight its novelties including direct modeling of phase and amplitude and flexibility for irregularly sampled data. Then in Section 3.2 we will describe our posterior sampling scheme and compare it to some potential alternatives in terms of speed and efficiency.

3.1 Model

Our primary goal with the project is to create a model that is both simple to interpret and flexible. The idea is that the model could provide meaningful insights to hydrologists while still managing the complex data features that this data set contains. As mentioned in Section 1.2 our data have clear annual cycles, but the amplitudes and midlines of these periodic cycles vary over time.

Just like the DHR models of Section 2.2, our model is based on the intersection of traditional harmonic regression and DLMs. For a review of harmonic regression see Section 1.3, or for more details see Chapter 3 of Prado and West (2010). We will begin this section with a brief review of the defining equations of a DLM. For a more comprehensive review see Chapter 4 of Prado and West (2010).

In a traditional DLM, models are typically defined by two equations. The first is the state evolution or state innovation equation

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\nu}_t \tag{3.1}$$

where $\boldsymbol{\theta}_t$ represents the vector of time-varying parameters at time t and $\boldsymbol{\nu}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{W}_t)$. The second is the observation model

$$y_t = \boldsymbol{f}_t \boldsymbol{\theta}_t + \epsilon_t \tag{3.2}$$

where $\boldsymbol{f}_t$ is a vector of known covariates or basis function values at time t and $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$. Notice that this model definition assumes normality. In our situation where we wish to model amplitudes directly, we are unable to use this formulation since the amplitudes must remain positive in order for our model to be identifiable.

This has led to our own DHR model formulation. In our model we will include a time-varying intercept term, β_{0t}, which will be unconstrained and can therefore follow the univariate version of the state evolution presented in Eq. (3.1) where we will refer to the univariate $\boldsymbol{W_t}$ as τ^2_{0t}. Our model for the amplitudes will follow a very similar idea, but will include left truncation at zero in order to enforce non-negativity of the amplitudes at every time point. This yields

$$\alpha_{it}|\alpha_{i(t-1)} \sim TN(\alpha_{i(t-1)}, \tau^2_{it}, 0)^1 \quad \text{for } i \in \{1, 2, \ldots, n_\mathrm{f}\} \tag{3.3}$$

where α_{it} represents the amplitude for frequency i, τ^2_{it} is the evolution variance of amplitude i at time t and n_f is the total number of frequencies included in the model. With our time-varying intercept the $\boldsymbol{\theta}$ vector at each time point will include $n_\mathrm{f} + 1$ time-varying parameters. The corresponding $\boldsymbol{f}_t$ for each $\boldsymbol{\theta}_t$ is

$$\boldsymbol{f}_t = \begin{bmatrix} 1 & \cos\left(2\pi[s_t(\xi_1)]\lambda^{-1} + \phi_1\right) & \cdots & \cos\left(2\pi[s_t(\xi_{n_\mathrm{f}})]\lambda^{-1} + \phi_{n_\mathrm{f}}\right) \end{bmatrix} \tag{3.4}$$

where the s_t is a continuous time measure at time index t. The first element of $\boldsymbol{f}_t$ corresponds to the time-varying intercept term (trend) and the rest of the terms represent the cyclical basis functions corresponding to each frequency. Using these values in Eq. (3.2) indicates that each observation y_t is the sum of the time-varying intercept β_{0t} and the seasonal effects for each of the n_f frequencies with period λ, phase ϕ_i, frequency ξ_i, time-varying amplitude α_{it} and white noise term ϵ_t.

Recall that the evolution at each time point for each TVP is governed by τ^2_{it}. For each τ^2_{it} there is a constant underlying process variance, τ^2_i, corresponding to a given TVP (intercept or amplitude) which we multiply by the time gap between time $t-1$ and time t to get τ^2_{it}. This follows the theory of the Wiener process as described in Section 1.3. Note here that t is still used as an index, not an actual measurement of time. If s represents the actual measurement of time then

$$\tau^2_{it} = \tau^2_i \times (s_t - s_{t-1}) \tag{3.5}$$

demonstrates the scaling described. This scaling takes place for each of the amplitudes as well as the intercept and is how we modify each of those DLMs to account for irregularly sampled data.

Note that the τ^2_i parameters are treated as fixed in this model and are used as tuning parameters. Large values for these underlying process variances allow for large movements in the TVPs over time. In some cases this is necessary, but it can also lead to overfitting. Analogously, small values will restrict the movement of the TVPs, protecting against overfitting, but potentially restricting the ability of our model to fit real trends. The values used in our analysis were tuned to each river by qualitatively evaluating whether the

[1] We use the notation $TN(\mu, \sigma, lb)$ to indicate a truncated normal distribution with μ and σ representing the mean and standard deviation of the corresponding normal distribution and lb representing the lower bound for truncation.

posterior evolution of each TVP was smooth. If a given TVP appeared to be overfitting the data, we reduced the value of its corresponding τ_i^2. Note that this less objective process could be easily replaced by estimation via discount factors as suggested in Prado and West (2010). A more elegant solution that we hope will be added to this model in future research would be to estimate the evolution variance directly following the methods and shrinkage priors suggested by Früwirth-Schnatter and Wagner (2010) and Cadonna et al. (2020). This would not only make the estimation of the evolution variance fully Bayesian, it would also make the model more robust to misspecification of the number of frequencies.

The priors in the DLM setting take the form of distributions at an imaginary initial time point from which the rest of the time series originates. For the intercept term this will be denoted

$$\beta_{00} \sim \mathcal{N}(m_0, v_0) \tag{3.6}$$

and for the amplitudes

$$\alpha_{i0} \sim TN(m_i, v_i, 0). \tag{3.7}$$

For two of the rivers we fit with the model we used an annual and a biannual frequency, for the last we only used an annual frequency. The priors for the time-varying intercept and the two amplitudes were as described above with $m_0 = m_1 = m_2 = 0$, $v_0 = 4^2$, $v_1 = 2^2$ and $v_2 = 1.5^2$. We did not have any specific prior information for individual rivers so these priors were chosen because they assign significant mass to a wide range of reasonable values for the intercept and the amplitudes. While most of our focus in this model is on novel specification and sampling for the TVP portion of the model we will also estimate phases and observation variance. The relevant model specifications for these parameters are described below.

The phases, $\boldsymbol{\phi}$, are each individually bounded between 0 and 2π. We used uniform priors over the support since we did not have any strong prior information and the bounded support makes the uniform prior proper in this case. The phases can be easily rescaled to the time scale of the data to provide insight into when the oscillations of a specific frequency begin within a given time period.

The observation variance, σ^2 was modeled using an inverse-gamma prior with shape parameter $\alpha = 2$ and scale parameter $\beta = 0.5$. We note that this prior may appear to overstate our prior knowledge, but the scale of the data for most of the rivers in question indicates that a variance of $\sigma^2 = 1$, which is roughly the 0.9 quantile of our prior distribution, would be quite large. So while the prior is informative, it still allows flexibility within the range of values we reasonably expect σ^2 to take on. The inverse-gamma was chosen since it is semi-conjugate, given our conditionally Gaussian sampling model.

As noted in Section 1.3 one of the great benefits of this model is that only a basic understanding of sinusoidal functions is required to understand how the parameters being estimated relate to real nitrate concentrations.

3.2 MCMC

Now that we have described our model we will detail the MCMC algorithm that will allow us to sample from the posterior. We will break this into a brief description of the quantile slice sampler and then an overview of the full Gibbs sampler used for our model.

3.2.1 Quantile Slice Sampler

The quantile slice sampler described in Heiner et al. (2023) employs the probability integral transform to allow for slice sampling by selecting a pseudo-target, and avoids all other tuning parameters. Let π represent the target distribution, which will generally be a posterior distribution. Let $\hat{\pi}$ represent the pseudo-target distribution with corresponding CDF $\hat{\Pi}$ and inverse CDF $\hat{\Pi}^{-1}$. The quantile slice sampler is based on the following sequence of equalities:

$$\begin{aligned}
\pi(\theta) &= \frac{\pi(\theta)}{\hat{\pi}(\theta)}\hat{\pi}(\theta) \\
&= h(\theta)\hat{\pi}(\theta) \\
&= h\left(\hat{\Pi}^{-1}(\psi)\right)\hat{\pi}\left(\hat{\Pi}^{-1}(\psi)\right)\frac{1}{\hat{\pi}\left(\hat{\Pi}^{-1}(\psi)\right)} \\
&= h\left(\hat{\Pi}^{-1}(\psi)\right)U(\psi;0,1)
\end{aligned} \tag{3.8}$$

where $U(\psi;0,1)$ represents the standard uniform density. Note that we are applying the transformation $\psi = \hat{\Pi}(\theta)$ as we move from line two to line three. On line three we have the result of applying the transformation, including the Jacobian, which then cancels with the second factor. In traditional slice sampling, finding an initial interval on which to slice can be difficult and requires tuning. It is normally achieved by Neal's stepping out procedure described in Neal (2003). In the quantile slice framework we do not have to worry about finding an initial interval that covers the slice at all, because the transformation has restricted our support to the unit interval. We will now briefly describe how to obtain a posterior sample using the quantile slice sampler.

To obtain a posterior draw we first draw the latent variable z uniformly on $[0, h(\theta_0)]$ where θ_0 is the previous draw. This latent variable z defines our slice. We then generate a proposal by drawing ψ_{prop}, from a standard uniform distribution and transforming it back to the θ scale using $\theta_{prop} = \hat{\Pi}^{-1}(\psi_{prop})$. The proposal is rejected if $h(\theta_{prop}) < z$. If the proposal is rejected we use the shrinkage procedure defined in Neal (2003) to shrink the unit interval around $\psi_0 = \hat{\Pi}(\theta_0)$. We perform this shrinkage by replacing the upper bound of the uniform with ψ_{prop} if $\psi_{prop} > \psi_0$ and otherwise replacing the lower bound with ψ_{prop}. We then sample a new ψ_{prop} uniformly using the new bounds and repeat until we accept one of the θ_{prop}'s as the new draw. A more detailed description can be found in Algorithm 3.1. Note that both this description and the algorithm are described on the density scale, but the algorithm should usually be implemented on the log scale for numerical stability.

Algorithm 3.1 Quantile slice sampling.

Input:
θ_0 (previous draw)
f (target density function)
$\hat{\pi}$ (pseudo-target density function)
$\hat{\Pi}$ and $\hat{\Pi}^{-1}$ (CDF and inverse CDF of pseudo-target)
Output:
x_1 (new draw)
Steps:
Draw z from uniform$(0, f(\theta_0)/\hat{\pi}(\theta_0))$
$L \leftarrow 0$
$R \leftarrow 1$
Draw ψ_{prop} from uniform(L, R)
$\theta_{prop} \leftarrow \hat{\Pi}^{-1}(\psi_{prop})$
while $f(\theta_{prop})/\hat{\pi}(\theta_{prop}) < z$ **do**
  if $\psi_{prop} < \hat{\Pi}(\theta_0)$ **then**
    $L \leftarrow \psi_{prop}$
  else
    $R \leftarrow \psi_{prop}$
  end if
  Draw ψ_{prop} from uniform(L, R)
  $\theta_{prop} \leftarrow \hat{\Pi}^{-1}(\psi_{prop})$
end while
$\theta_1 \leftarrow \theta_{prop}$
Return: θ_1

The quantile slice sampler requires that our transformation be a one-to-one mapping between the unit interval and the support of the posterior. In the univariate case the probability integral transform automatically has this property, but the multivariate case requires more care. In the multivariate case we require a multivariate transformation that provides a one-to-one mapping between the support of the posterior and the unit hypercube of the same dimension. Here we take advantage of the fact that the FFBS algorithm has already decomposed our joint posterior into a chain of univariate conditional distributions. We can use the corresponding univariate CDFs to define a transformation that matches these criteria. The details are beyond the scope of this project, but can be found in Heiner et al. (2023).

3.2.2 Gibbs Sampler

Our sampling scheme is based on a Gibbs sampler which samples in turn from $p(\boldsymbol{\phi}|\boldsymbol{\theta}, \sigma^2, \boldsymbol{y})$, $p(\sigma^2|\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{y})$ and $p(\boldsymbol{\theta}|\boldsymbol{\phi}, \sigma^2, \boldsymbol{y})$. The first two steps are relatively simple and will be described first.

To sample the phases given all other parameters we use a multivariate quantile slice sampler. For our pseudo-target we use three independent modular location-scale beta

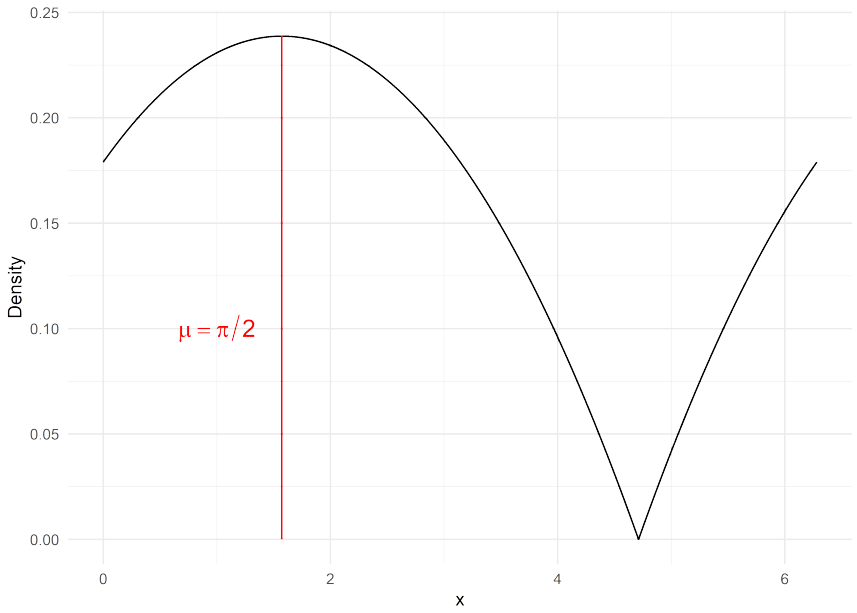


Figure 3.1: Modular location-scale beta distribution.

distributions. The pdf of the location-scale beta is given by

$$f(x) = \sigma^{-1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{x-\mu}{\sigma}\right)^{\alpha-1} \left(1 - \left(\frac{x-\mu}{\sigma}\right)\right)^{\beta-1} \tag{3.9}$$

where α and β are the two shape parameters, μ is the location parameter and σ is the scale parameter. Since our phases parameters $\boldsymbol{\phi}$ have restricted support between 0 and 2π we will use $\sigma = 2\pi$ so that we cover the proper support. We will use $\mu = \phi_0$ so that previously drawn values have more weight in the pseudo-target. Note that shifting the location parameter will modify our support. To solve this problem we impose modular behavior on the location-scale beta, so that values beyond 2π wrap back over zero, and values less than zero wrap back over 2π. A graphical demonstration of the behavior of our pseudo-target can be seen in Fig. 3.1. The quantile slice sampler with the described pseudo-target will behave similarly to a generic slice step, but will favor sampling values close to the previously drawn value.

For the residual variance σ^2 we sample directly from the full conditional. The semi-conjugate inverse-gamma prior combined with our conditionally normal sampling model makes the full conditional for σ^2 an inverse-gamma distribution. More specifically, for each Gibbs scan we draw a value for σ^2 from the specific inverse-gamma distribution with density

$$f(\sigma^2) = \frac{\beta^{\star\alpha^\star}}{\Gamma(\alpha^\star)} \sigma^{2-\alpha^\star-1} e^{-\beta^\star/\sigma^2} \tag{3.10}$$

where $\alpha^\star = \alpha + T/2$ and $\beta^\star = \beta + \sum_{t=1}^{T}(y_t - \boldsymbol{f}_t'\boldsymbol{\theta}_t)^2/2$ and α and β came from our previously defined prior. We now turn our attention to updates for the TVPs.

Each TVP is updated as a separate univariate DLM with all other terms in the mean function being subtracted from the data vector $\boldsymbol{y}$ to create a new vector $\boldsymbol{y}^*$ of partial residuals. These partial residuals are then used in the univariate DLM block update. The intercept $\boldsymbol{\beta}_0$ is an unconstrained TVP so we can use a generic FFBS update (see Section 2.3).

The block update for a time-varying amplitude given all other parameters will combine the FFBS algorithm with the quantile slice sampler. Recall that our target density is no longer a Gaussian DLM because of the truncated normal densities in Eq. (3.3). This means that we cannot use FFBS to sample directly from the posterior. Instead, we will use a multivariate version of the quantile slice sampler, using a transformed Gaussian DLM density as the pseudo-target. This pseudo-target can be easily sampled by using FFBS and then applying a secondary transformation to the resulting draws, which is described below. The multivariate generalization of the shrinking procedure we use is based on the multivariate slice sampler proposed in Neal (2003). In order to use the quantile slice sampler we need a multivariate transformation that describes a one-to-one mapping to the unit hypercube. In the FFBS framework this is easily achieved by subdividing the necessary multivariate $\hat{\Pi}$ into a chain of $\hat{\Pi}_t$ transformations that relate directly to the conditional backward sampling densities described in Eq. (2.2).

The purpose of the secondary transformation to the FFBS draws is to prevent negative proposal values. Negative proposals in any dimension would immediately lead to rejection of the proposed vector since the density of the truncated normal target is zero for negative numbers. This could decrease the efficiency of our sampler by making deterministic rejections based on one dimension, even when other dimensions may have excellent proposals. The transformation applied to the FFBS draw prior to the proposal is

$$f(x) = \begin{cases} x & x \geq c \\ c\, e^{(x/c)-1} & x < c \end{cases} \tag{3.11}$$

which is applied individually to each instance of the time-varying amplitude being sampled. In Fig. 3.2 we see how the transformation replaces negative values with very small values while leaving values above c untouched. This achieves the desired outcome by not biasing larger proposals, while moving negative proposals into an acceptable range. For our sampler we used $c = 0.01$, as this was found to be small enough to minimize impact to positive proposals, but large enough to avoid numerical instability.

Once we have the proposed draw from the transformed FFBS, the proposal is accepted or rejected in a very similar manner to Algorithm 3.1, but with each proposal being vector-valued. Further details of an update for the time-varying amplitudes of one frequency are found in Algorithm 3.2.

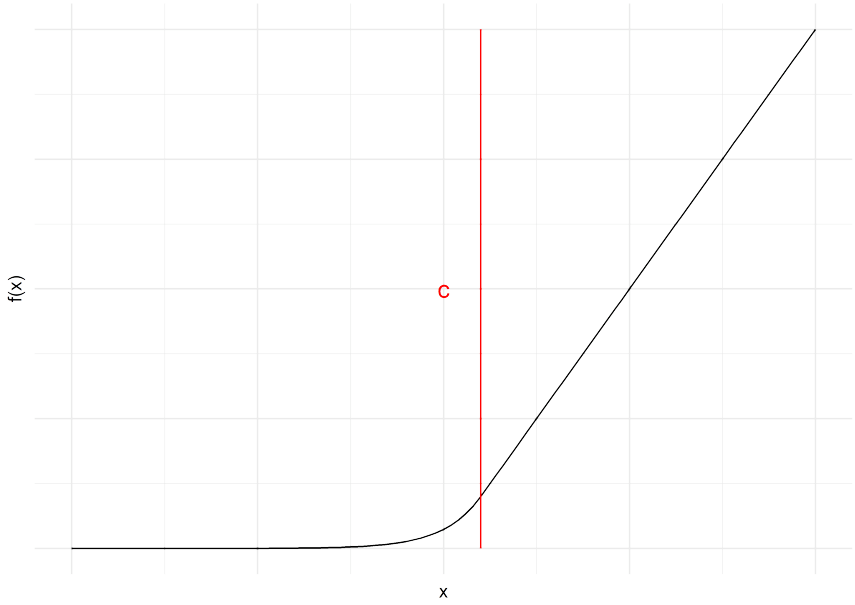


Figure 3.2: Secondary transformation to avoid negative proposals.

Algorithm 3.2 Amplitude block update.

Input:
$\boldsymbol{\alpha}_0$ (previous draw)
f (target density function)
$\hat{\pi}$ (transformed FFBS pseudo-target density)
$\hat{\Pi}$ and $\hat{\Pi}^{-1}$ (multivariate CDF and inverse for transformed FFBS)
Output:
$\boldsymbol{\alpha}_1$ (new draw)
Steps:
Draw y from uniform$(0, f(\boldsymbol{\alpha}_0)/\hat{\pi}(\boldsymbol{\alpha}_0))$
for t in 1:T **do**
    $L_t \leftarrow 0$
    $R_t \leftarrow 1$
end for
Draw $\boldsymbol{\psi}$ from the hyper rectangle defined by bounds $\boldsymbol{L}$ and $\boldsymbol{R}$
$\boldsymbol{\alpha}_{prop} \leftarrow \hat{\Pi}^{-1}(\boldsymbol{u})$
while $f(\boldsymbol{\alpha}_{prop})/\hat{\pi}(\boldsymbol{\alpha}_{prop}) < y$ **do**
    for t in 1:T **do**
        if $\psi_t < \hat{\Pi}_t(\boldsymbol{\alpha}_0)$ **then**
            $L_t \leftarrow \psi_t$
        else
            $R_t \leftarrow \psi_t$
        end if
    end for
    Draw $\boldsymbol{u}$ from the hyper rectangle defined by bounds $\boldsymbol{L}$ and $\boldsymbol{R}$
    $\boldsymbol{\alpha}_{prop} \leftarrow \hat{\Pi}^{-1}(\boldsymbol{\psi})$
end while
$\boldsymbol{\alpha}_1 \leftarrow \boldsymbol{\alpha}_{prop}$
Return: $\boldsymbol{\alpha}_1$

4 Results

In this chapter we will provide details of the model behavior when fit to real data. We will describe the insights obtained by inspecting the data fit and the trajectories of the TVPs. We will also report the performance of our MCMC framework, and compare our quantile slice update for the TVPs to some other samplers that could be used to make this block update.

4.1 Fitting Data from Real Rivers

We tested our model on three different rivers from the data set (Naïades, 2018). These rivers were chosen for their density of data, demonstration of the data features described in Section 1.2, and their quantity of data. For the first time series we investigated, our model was able to fit the data well and quickly. A plot of the data with a line representing the point estimate for the underlying state surrounded by 90% credible bands is displayed in Fig. 4.1.

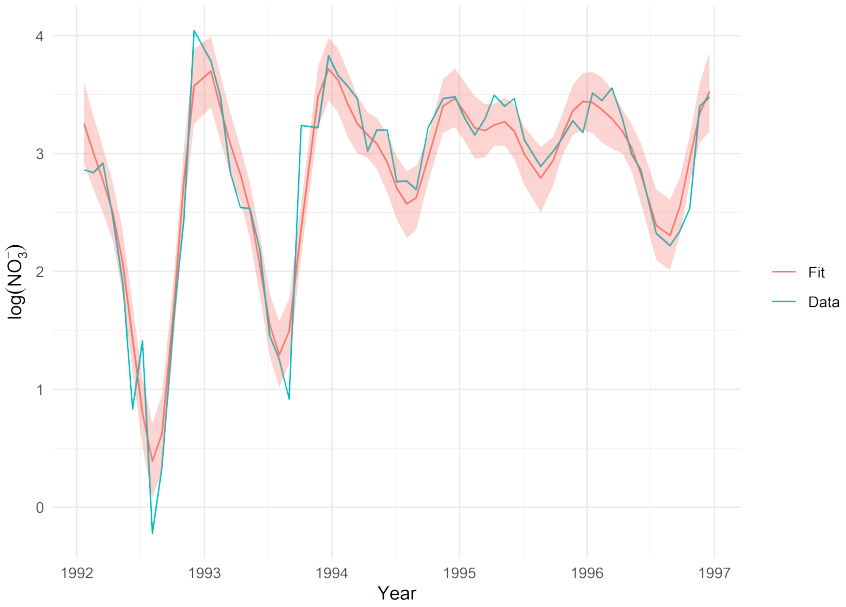
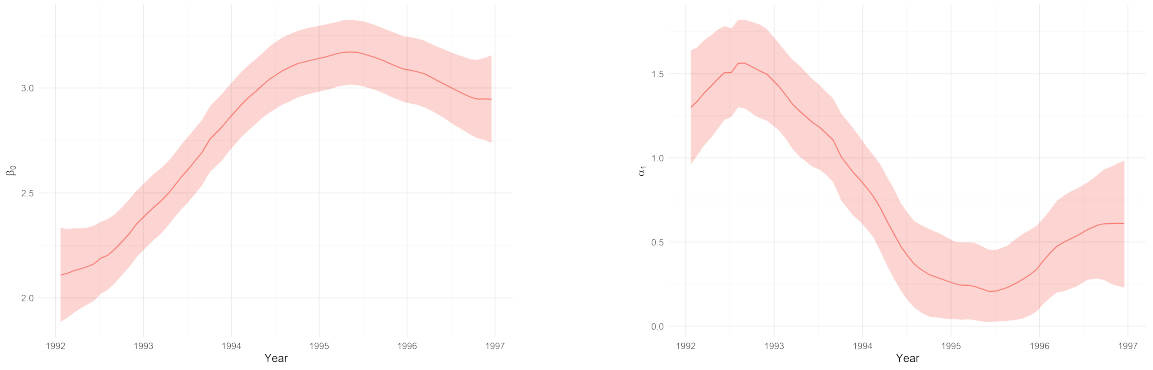


Figure 4.1: Model fit to river 1.

The data from this first river demonstrates very clearly the shifting of the seasonal cycles that our model was designed to capture. From a visual inspection of the river 1 data it appears that the amplitude is large at the beginning of the time series and moves to smaller

values as the time series progresses. We also see movement in the intercept which appears to start at low values and move to higher values over the course of the time series. We can see in Fig. 4.2 that our model has clearly captured and quantified the trends that we identified visually. Also note the smoothness of estimated trajectories of the TVPs. This indicates clear identification of the parameter values without overfitting.



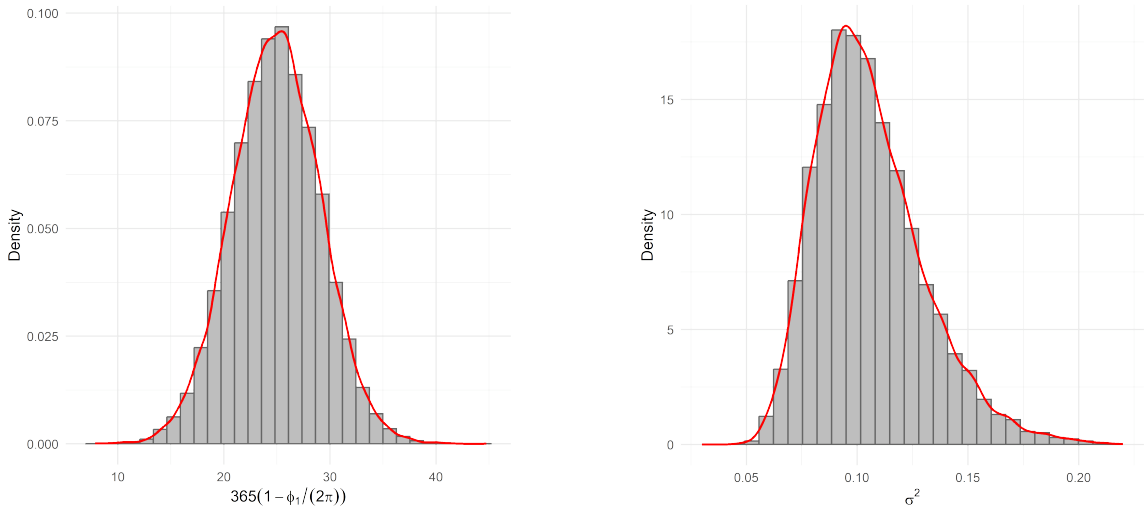
(a) Time-varying trend $\boldsymbol{\beta}_0$.

(b) Time-varying amplitude $\boldsymbol{\alpha}_1$.

Figure 4.2: Trend and amplitude for the primary frequency of river 1.

Posterior distributions for phases and observation variance were also stable. Posterior distributions for the phase shift of the first frequency and the observation variance can be found in Fig. 4.3. Inference on our observation variance can give us an idea of how much noise is left after we fit our model, and its stability can help us identify whether other parameters have also stabilized. The phase shift can show us where oscillations of our seasonal cycles peaked within a given time frame. Since the period for our primary frequency is annual, we can identify when within the year the oscillation is maximized. While phases on the 0 to 2π scale do not obviously provide this information we can simply apply a transformation of the form $365[1 - \phi/(2\pi)]$ to obtain the specific decimal day of the year in which the cycle started. Because we are using the cosine as our periodic basis function, the phase shift also corresponds with the time of the peak. The interpretable form of the phase could be of interest to hydrologists in terms of understanding underlying drivers of the seasonal cycles. In Fig. 4.3a we have reported the phase on the interpretable scale. Our posterior indicates that the annual cycle likely peaks in mid to late January.

We now turn our attention to the other two rivers that we applied our model to. While our model performed fairly well on both of these nitrate concentration time series, they also demonstrate limitations of our model. The fit to river 2 is displayed in Fig. 4.4. We observe that this time series displays many of the same features as river 1, but has a much more sudden shift in underlying level and amplitude. Our DHR model is based on the assumption that the TVPs have smooth trajectories, not jumps. This means that in time series like the nitrate concentration of river 2, the model will attempt to smooth out the sudden change. This will lead to our model overestimating parameters on one side of the



(a) Phase shift in days.

(b) Observation variance σ^2.

Figure 4.3: Estimated marginal posterior densities with histograms for day of the year when the annual seasonal effect is maximized and observation variance for river 1.

jump and underestimating on the other side of the jump, as it spreads out the movement that occurred in the point change. We see this behavior in Fig. 4.4 as the amplitudes appear to become small prematurely and the trend appears to move up prematurely prior to the change point around 2007. While the model does not account for this significant feature of the data from this river, it is still flexible enough to provide reasonably good fit. We suggest the work of Chappell and Tych (2012) to the reader interested in pursuing change-point compatible models related to ours.

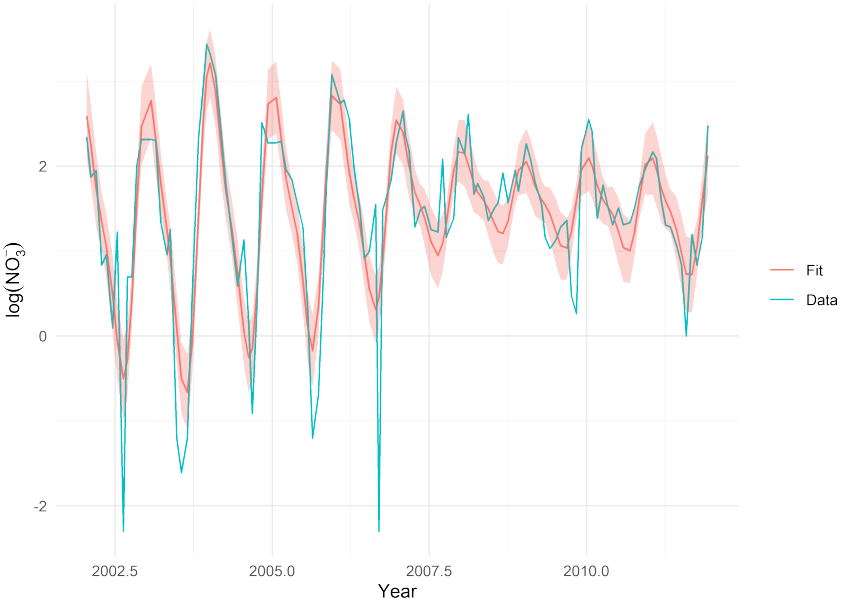


Figure 4.4: Model fit to river 2.

The final river shows a longer time series. This time series also has some drastic shifts, though none that would likely be characterized as a change point. The difficulty with this time series is that over the course of the time series the parameters appear to change at different rates. Since we are using a fixed underlying τ_i^2 parameter for each of the univariate DLMs, we must compromise. We can either overfit in some areas or we can underfit in others. For this analysis we chose values for τ_i^2 that were on the conservative side. As a result we see in Fig. 4.5 that our model has a hard time reaching all the way to the peaks and troughs of our data. Since we move gradually from high to low at first, and then more quickly between low and high, we inevitably miss some of these features when we constrain our model to be more simple.

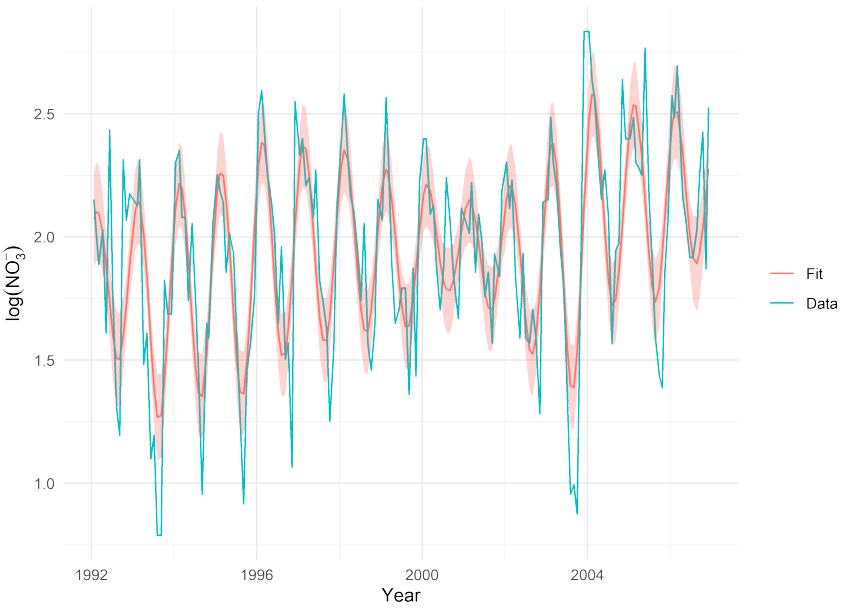


Figure 4.5: Model fit to river 3.

Overall we are quite pleased with the flexibility of our model and its performance in fitting data and providing meaningful insight into the pollution behavior of individual rivers.

4.2 MCMC Sampler Comparison

In this section we compare six different samplers for the amplitude parameters. We compare two variants on our quantile slice method against three easily implemented alternatives, and use the FFBS sampler from the unconstrained Gaussian DLM as a benchmark. Note that we use it as a benchmark for efficiency, but it is not a true competitor since the Gaussian DLM it is sampling does not match our model and is in fact nonidentifiable.

The first of our comparison samplers is the quantile slice sampler. For this first sampler our pseudo-target is a Gaussian DLM with the secondary transform described in Eq. (3.11)

applied. This is the sampler that we suggest using with our model and the one we used to fit the models in Section 4.1. This sampler will be referred to in Table 4.1 and Table 4.2 as QS2 for quantile slice secondary transform. The next sampler is the quantile slice sampler with a Gaussian DLM as the pseudo-target, but no secondary transformation. This sampler will be referred to as QS for quantile slice (Heiner et al., 2023). Next we compare against an independence Metropolis (IM) sampler using a Gaussian DLM to generate proposals as suggested by Fearnhead (2011). We also compare against the multivariate generalization of the slice sampler suggested in Algorithm 8 of Neal (2003) (MS). A newer modification of the slice sampler that is slightly more flexible is the latent slice sampler of Li and Walker (2023) which is expected to perform better in higher dimensional problems than the typical multivariate slice. It will be referred to as LS for latent slice.

Our comparisons will be based on simulation studies designed to replicate some of the difficult data features we see in the nitrate concentration data. Two separate simulations were used. Within each simulation the same priors and tuning parameters were used to generate the data and to fit the model. For each simulation we generated a time series of 120 data points, as this is the length of the river time series we have studied the most throughout this project. In the first simulation, data was generated so that amplitudes were far from zero. This means that the posterior distribution is effectively the same as that of a Gaussian DLM since the truncation at zero has no effect on the parameters when the distribution is already far from zero. We will use this simulation as a control case because of the posterior similarity to a Gaussian DLM. The second simulation data was generated to have amplitudes that were close to zero for most of the time points. This second simulation was designed to show how samplers behaved when the truncated model differed more significantly from the unrestricted model. The second simulation will be our test case to see how the behavior of the samplers changes when the truncation in our model becomes relevant.

Each sampler was used to obtain posterior chains from the generated data set 30 times per simulation. For each run of each sampler, four chains were run with a burn-in period of 1000, and 5000 samples drawn after burn-in. The value for burn-in was chosen by evaluating a conservative estimate for the iteration at which the best few samplers had converged. The four chains were run to be able to better assess convergence.

First we will examine the sampler behavior from the high amplitude simulation. Results from all of the samplers in terms of convergence, effective samples per CPU second and rejections for the slice samplers can be found in Table 4.1. The entries of Table 4.1 and Table 4.2 aggregate all metrics across the four chains and average across the three amplitude parameters. At high amplitudes each of the samplers that used FFBS as a proposal did very well in updating the amplitudes. Performance was nearly identical across the quantile slice samplers and the independence Metropolis sampler, with the independence Metropolis being slightly faster and therefore edging out the quantile slice samplers in effective samples

per CPU second. The actual FFBS algorithm had lower effective samples per CPU second, largely due to the inversion of matrices inherent to the multivariate update of all of the TVPs at once. The benefit of the FFBS is that each sample comes directly from the posterior. Note that FFBS is marked as converging here since amplitudes are far enough from zero that we do not encounter the problem of nonidentifiability. When samples are closer to zero this problem becomes evident as amplitudes flip across zero to their negative value and shift the corresponding phases to account for the change, all without changing the actual model.

Sampler	Converged	n_{eff}/s	time (s)	n_{eff}	Slice Rejections
QS2	Yes	107.8	150	16,167	8
QS	Yes	107.1	151	16,167	8
IM	Yes	114.0	142	16,190	NA
MS	No	0.2	197	48	1,096,501
LS	No	0.2	252	46	1,609,671
FFBS	Yes*	39.9	497	19,818	NA

*updates all amplitudes and intercept at once

Table 4.1: Sampler performance for the time-varying amplitudes in the high amplitude simulation.

We also note that neither the latent slice sampler nor the multivariate slice sampler were able to reach convergence, even in this high amplitude case. Further experimentation with shorter time series indicated that these methods did reach convergence in more reasonable times for shorter time series. For longer time series the movements are too small since the shrinking method for the multivariate slice sampler shrinks in all dimensions at once. In our 120 dimensional problem it only takes one poor proposal to reject and shrink all dimensions. This means if we have 119 excellent proposals and one poor one we may reject the proposal and our shrinkage procedure will then greatly reduce the possibility of movement in the 119 dimensions that had good values in the original proposal.

We now turn our attention to the behavior of the samplers for the low-amplitude simulation. The slice sampler and latent slice sampler once again failed to converge. This is expected since they will have the same problems stemming from the length of the time series, but we have also introduced the extra difficulty of having to reject automatically when any proposal is negative. For amplitudes close to zero this will restrict the movement of these samplers even further. Of greater interest is the difference in behavior for the independence Metropolis sampler and the two quantile slice samplers. We see in Table 4.2 that at low amplitudes all of the samplers were much less efficient. This is expected since the slice samplers will have more rejections before accepting a new sample, and the independence Metropolis will reject new proposals at a higher rate.

At the lower amplitude we also see that independence Metropolis sampler failed to converge. At high amplitudes it was able to outperform the quantile slice samplers because all three were accepting at high rates, but the quantile samplers have more overhead associated with the transformations. However, at low amplitudes the high rejection rate

Sampler	Converged	n_{eff}/s	time (s)	n_{eff}	Slice Rejections
QS2	Yes	3.8	327	1,239	349,634
QS	Yes	2.0	356	686	455,406
IM	No	1.9	123	234	NA
MS	No	0.3	177	47	990,782
LS	No	0.2	227	46	1,629,078
FFBS	No*	40.2	474	19,067	NA

*stabilizes, but is nonidentifiable

Table 4.2: Sampler performance for the time-varying amplitudes in the low amplitude simulation.

causes the independence Metropolis chains to stick in one place for thousands of iterations, making it unable to compete with the quantile slice samplers that are guaranteed to move between iterations. The FFBS still stabilizes, but the non-identifiability of the underlying Gaussian model causes much larger problems at low amplitudes, with amplitudes flipping back and forth across zero and shifting the phases along with them. This leaves the two quantile slice samplers as the only two viable options among the compared samplers. As intended, we were able to prevent some rejections for the quantile slice sampler by using the secondary transformation. By rejecting less, we were able to shrink less per iteration, which means that our samples can explore the posterior space more quickly and achieve higher effective samples sizes.

5 Conclusion

In this master's project we have sought to understand the temporal behavior of nitrate pollution levels in French rivers. As an approach to improve our understanding we created a novel DHR model. Our modeling approach was focused on interpretation and parsimony. Our final model is flexible enough to account for the time-varying cyclical patterns and irregular sampling of our data, while using parameters and distributions tightly tied to physical features of the time series.

We also developed a novel MCMC sampling technique specifically for our model which can obtain samples efficiently from the posterior. We demonstrated by comparison to popular alternatives both the necessity and the effectiveness of our sampler.

We tested our model and sampler through simulation and application to various real nitrate concentration time series. We found that it was flexible and interpretable, and believe that it could be used to improve monitoring of pollutant levels in rivers.

In future work we hope to add more flexibility to our model for the river pollution data. Specific areas of interest for future incorporation into this model are Bayesian shrinkage priors for both variable selection and variance selection (Früwirth-Schnatter and Wagner, 2010; Bitto and Frühwirth-Schnatter, 2019; Cadonna et al., 2020) and change points to account for sudden shocks to the underlying process (Chappell and Tych, 2012). Also of interest would be improving the efficiency and packaging of the software for easier application to thousands of rivers.

References

Abbott, B. W., Moatar, F., Gauthier, O., Fovet, O., Antoine, V., and Ragueneau, O. (2018). "Trends and seasonality of river nutrients in agricultural catchments: 18 years of weekly citizen science in France." *Science of The Total Environment*, 624, 845–858.

Bitto, A. and Frühwirth-Schnatter, S. (2019). "Achieving shrinkage in a time-varying parameter model framework." *Journal of Econometrics*, 210, 1, 75–97.

Buckley, T., Pakrashi, V., and Ghosh, B. (2021). "A dynamic harmonic regression approach for bridge structural health monitoring." *Structural Health Monitoring-an International Journal*, 20, 6, 3150–3181.

Cadonna, A., Fruhwirth-Schnatter, S., and Knaus, P. (2020). "Triple the Gamma–A Unifying Shrinkage Prior for Variance and Variable Selection in Sparse State Space and TVP Models." *Econometrics*, 8, 2, 1–36.

Carter, C. K. and Kohn, R. (1994). "On Gibbs sampling for state space models." *Biometrika*, 81, 3, 541–553.

Chappell, N. A. and Tych, W. (2012). "Identifying step changes in single streamflow and evaporation records due to forest cover change." *Hydrological Processes*, 26, 1, 100–116.

Cheng, F. Y., Van Meter, K. J., Byrnes, D. K., and Basu, N. B. (2020). "Maximizing US nitrate removal through wetland protection and restoration." *Nature*, 588, 7839, 625–630.

Diaz, R. J. and Rosenberg, R. (2008). "Spreading Dead Zones and Consequences for Marine Ecosystems." *Science*, 321, 5891, 926–929.

Driver, C. C. and Voelkle, M. C. (2018). "Hierarchical Bayesian continuous time dynamic modeling." *Psychological Methods*, 23, 4, 774–799.

Dupas, R., Minaudo, C., and Abbott, B. W. (2019). "Stability of spatial patterns in water chemistry across temperate ecoregions." *Environmental Research Letters*, 14, 7, 74015.

EarthHow (2022). "How does eutrophication work? causes, process and examples." Accessed March 2023, url: `https://earthhow.com/eutrophication-causes-process-examples/`.

Fearnhead, P. (2011). "MCMC for State-Space Models." In *Handbook of Markov Chain Monte Carlo*, 513–529. Boca Raton: Chapman & Hall/CRC.

Früwirth-Schnatter, S. and Wagner, H. (2010). "Stochastic model specification search for Gaussian and partial non-Gaussian state space models." *Journal of Econometrics*, 154, 1, 85–100.

Frühwirth-Schnatter, S. (1994). "Data augmentation and dynamic linear models." *Journal of Time Series Analysis*, 15, 2, 183–202.

Heiner, M., Dahl, D., Johnson, S., and Christensen, J. (2023). "Quantile Slice Sampling with Transformations to Approximate Targets." [Manuscript in preparation].

Huerta, G. and West, M. (1999). "Priors and component structures in autoregressive time series models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61, 4, 881–899.

Kalman, R. E. (1960). "A New Approach to Linear Filtering and Prediction Problems." *Journal of Basic Engineering*, 82, 1, 35–45.

Le Moal, M., Lefebvre, A., Ménesguen, A., Pannard, A., Levain, A., Gascuel-Odoux, C., Étrillard, C., Moatar, F., Pinay, G., Souchu, P., and Souchon, Y. (2019). "Eutrophication: A new wine in an old bottle?" *Science of The Total Environment*, 651, 1–11.

Li, Y. and Walker, S. G. (2023). "A latent slice sampling algorithm." *Computational Statistics & Data Analysis*, 179, 107652.

Mindham, D. A. and Tych, W. (2019). "Dynamic harmonic regression and irregular sampling; avoiding pre-processing and minimising modelling assumptions." *Environmental Modelling and Software*, 121.

Naïades (2018). "Physicochemistry data for Whole France." Data retrieved November 2018, url: `http://www.naiades.eaufrance.fr/france-entiere#/`.

Neal, R. M. (2003). "Slice Sampling." *The Annals of Statistics*, 31, 3, 705–741.

Prado, R. and West, M. (2010). *Time series: modeling, computation, and inference*. Chapman and Hall/CRC Press.

U.S. Geological Survey (2019). "Nutrients and eutrophication active." Accessed March, 2023, url: `https://www.usgs.gov/mission-areas/water-resources/science/nutrients-and-eutrophication#overview`.

West, M. (1995). "Bayesian Inference in Cyclical Component Dynamic Linear Models." *Journal of the American Statistical Association*, 90, 432, 1301–1312.

West, M. and Harrison, J. (1997). *Bayesian forecasting and dynamic models*. 2nd ed. Springer, New York.

Young, P. C., Pedregal, D. J., and Tych, W. (1999). "Dynamic Harmonic Regression." *Journal of Forecasting*, 18, 6, 369–394.

Zavala, A. J. and Messina, A. R. (2016). "Dynamic harmonic regression approach to wind power generation forecasting." In *2016 IEEE PES Transmission & Distribution Conference and Exposition-Latin America (PES T&D-LA)*, 1–6. IEEE.