Comparing the Velocity of Climate Change and the Velocity of Phenology Change

Diana Gerardo[1]

[1] *University of California, Santa Cruz*

September 2019

The goal of this paper is to quantify spatial temporal trends for the velocity dataset calculated from TerraClimate and Vipphen EVI2 datasets. We specifically want to see if there is a correlation between the velocity of cumulative vegetation index change (V_CVI) and the velocity of mean annual temperature change (V_MAT) between the years 1981 - 2014. To check if the relationship is stable we use moving windows analysis. We construct a univariate Bayesian spatial regression model for each time period and perform posterior and predictive inferences.

1 Introduction

Global climate change has been continuing to affect terrestrial ecosystems worldwide. Vegetation can adjust quickly to a slow rate of climate change through changes in its phenology or physiology and/or changes in community composition (Huang et al., 2017). However, with rapid climate change, a recent topic of interest is to explore how the rate of climate change prevents vegetation from adapting to its climate by decreasing its ability to stabilize its productivity change rate.

The goal of this report is to explore the relationship between the velocity of cumulative vegetation index (V_CVI) change and the velocity of mean annual temperature (V_MAT).

1.1 Velocities of Change

Global temperature data was obtained from the dataset TerraClimate (J.T. Abatzoglou, 2018). A mean annual temperature (MAT) value is computed for each year from 1981 to 2014. MAT is defined to be the average of mean monthly temperatures in a given year. From an unpublished data obtained through a personal communication with Yiluan Song (Environmental Sciences UCSC), V_MAT change was calculated repeatedly in a 10 year window. In section only, let h denote distance between location points and time period denoted as t. The the velocity was calculated as

$$\text{V\_MAT} = \frac{dh}{dt} = \frac{d\text{MAT}}{dt} \div \frac{d\text{MAT}}{dh}$$

where $\frac{d\text{MAT}}{dt}$ is the temporal gradient in degree Celsius per year and $\frac{d\text{MAT}}{dh}$ is the spatial gradient in degree Celsius per km.

The 2-band enhanced vegetation index (EVI2) data was obtained from the dataset Vipphen EVI2 (Didan and Barreto, 2016b). EVI2 quantifies healthy green vegetation. The annual EVI2 cycle follows a bell shaped curve with low values during the end of winter, high values during spring and early summer, decreasing values during end of summer and fall, and finally low values again during the start of winter. The area under the annual bell shaped curve is the cumulative vegetation index (CVI) (Didan and Barreto, 2016a). Similarly, V_CVI change was calculated as

$$\text{V\_CVI} = \frac{dh}{dt} = \frac{d\text{CVI}}{dt} \div \frac{d\text{CVI}}{dh}$$

repeatedly in a 10-year window.

1.2 Velocity Dataset

The raw velocity dataset consists of the variables longitude, latitude, time period, V_MAT, and V_CVI. V_MAT and V_CVI were decomposed into southern-northern and western-eastern directions, and we use velocities in the southern-northern direction for our analysis. We re-projected the raw velocity data from WGS84 (lat, lon coordinates) into a Lambert Azimuthal Equal Area projection centering on the north pole, in order to

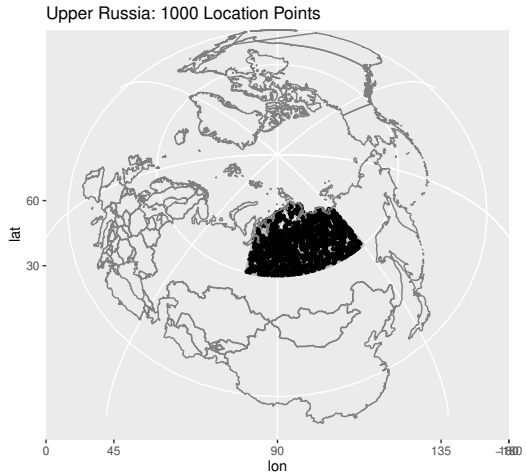

Figure 1: *Area and locations of focus.*

preserve the distances. Here we choose distances in km. Each time period corresponds to a 10-year window, with $t = 1$ corresponding to 1981-1990, $t = 2$ to 1982-1992, and so on, up to $t = 25$ corresponding to 2005-2014.

To explore the relationship between V_CVI and V_MAT, we construct a univariate Bayesian spatial regression model as described in section 3. This model incorporates a regression component to assess the relationship between V_CVI and V_MAT, an underlying process with a Gaussian process spatial structure and a non-spatial Gaussian error component (S. Banerjee and Gelfand, 2015).

Incorporating a moving window analysis of our Bayesian spatial model assesses the stability of the relationship between V_MAT and V_CVI over time and also that of the underlying spatial structure. A common assumption is that the parameters are constant with respect to time. Thus, any change in the parameter estimates at a point in time will be captured because a moving window was incorporated.

We focus on studying a small upper region of Russia. The resolution of the velocity dataset is in 0.5×0.5 degree pixels and in our region of focus there are 2825 location points (excluding oceanic/water points) per time period. So, data has been collected for the same 2825 locations over 25 time periods.

Obtaining posterior inferences from the univariate spatial regression model requires the implementation of Markov chain Monte Carlo methods. These are iterative methods that need to be run until convergence is achieved, and are not practically feasible for settings that involve a large number of spatial locations. Therefore, we choose only 1000 locations randomly, shown in Figure 1, from the total 2825 locations to run the model. These 1000 locations are the same across all time periods. The remaining 1825 locations become part of our test set for posterior predictive checking.

2 Semi-Variograms

The first stage in spatial exploratory data analysis is estimating semi-variograms. Semi-variograms depict spatial autocorrelation of the measured sample points. Let $Y_t(s)$ denote the V_CVI at time period t and location s, for $t = 1, \ldots, 25$ and $s \in \{s_1, \ldots, s_n\}$ with $n = 1000$. We define a linear regression model as follows

$$Y_t(s) = X_t(s)\beta_t + \varepsilon_t(s), \qquad \varepsilon_t(s) \sim N\left(0, \tau_t^2\right)$$

where $X_t(s) = (1, x_t(s))$ and $x_t(s)$ denotes the V_MAT at time t and location s. $\beta_t = (\beta_{t,0}, \beta_{t,1})'$ is the vector of regression coefficients and the error term $\varepsilon_t(s)$. This model assumes that the errors are independent across locations and over time.

The model above does not include any spatial structure, then, any spatial pattern should be in the residuals, which are computed as the observed minus the predicted values, i.e., $\hat{\varepsilon}_t(s) = Y_t(s) - \widehat{Y}_t(s)$. So we conduct the variogram analysis on the residuals. By definition the variogram is given by

$$\begin{aligned} 2\gamma(d) &= E\left\{[Z(s) - Z(s+d)]^2\right\} \\ &= Var\left[Z(s) - Z(s+d)\right] \end{aligned}$$

where $Z(\cdot)$ denotes a spatial process and d is a distance. In our case we take $Z(s) = \hat{\varepsilon}_t(s)$. Plotting the variogram provides the differences in pairs of residual data points, separated by a distance d.

Next, we obtain estimates of the variogram using a semi-variogram model $\gamma(d)$. The most common form of empirical variogram behaviour is that the correlation between $Z(s_i)$ and $Z(s_j)$ decreases as the distance increases (Diggle and Ribeiro, 2007). Therefore, we choose a semi-variogram model that behaves in the same way. We also choose a semi-variogram model that has flexibility with the degrees of smoothness in the underlying spatial process $Z(s)$. The semi-variogram that satisfies these conditions is the Matérn semi-variogram.

2.1 The Matérn Semi-Variogram

The Matérn semi-variogram implemented in the `gstat` R package follows the parameterization in Stein, 2015

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2\left[1 - \frac{(rd)^\kappa}{2^{\kappa-1}\Gamma(\kappa)} \cdot K_\kappa(rd)\right] & \text{if } d > 0 \\ 0 & \text{if } d = 0. \end{cases}$$

The function $\Gamma(\cdot)$ is the gamma function while K_κ is the modified Bessel function of order κ. Here κ controls the smoothness of the realized random field while r is the range parameter. The range is the distance after which the variogram levels off. Note that the range can be parameterized as $r = \frac{3}{\phi}$ where ϕ is the spatial decay parameter which is the parameterization used in the implementation of the Bayesian univariate

Table 1: *SSErr Criterion [note: N/C refers to no convergence].*

SSErr From Matérn Variogram				
t	$\kappa = 0.5$	$\kappa = 1.0$	$\kappa = 1.5$	$\kappa = 2.5$
1	10.5637	10.3642	10.234	10.0747
2	1.14994	1.67276	2.07572	2.61269
3	1.61087	2.27137	2.75322	3.34046
4	28.3998	30.1363	31.2465	32.5921
5	10369.5	10205.5	10124.2	10026.1
6	41470.3	40963.5	40623.7	N/C
7	5632.01	5613.09	5609.4	N/C
8	343.105	333.860	328.57	322.547
9	49.6714	52.8556	55.9717	60.5163
10	202.976	225.578	245.396	273.605
11	583343	574416	571664	569140
12	11037.9	10248.6	9919.31	9657.78
13	329986	309265	298509	287648
14	474369	460315	454110	447901
15	1221530	1234760	1250510	1270150
16	541.517	552.573	577.447	619.253
17	916.644	848.718	813.27	779.275
18	32.1169	30.8878	30.3247	29.8926
19	3.79932	3.95603	4.13945	4.42906
20	10.3706	8.24809	7.7688	7.77173
21	2054.79	2033.66	2055.51	2081.08
22	34.7739	36.785	39.5383	43.5523
23	37.5456	44.0467	48.1282	52.9145
24	15.9212	17.4703	18.2105	18.9063
25	12.1909	13.0214	13.6515	14.5061

Table 2: *Parameter Estimates.*

Parameter Estimates				
t	κ	τ^2	σ^2	$\phi = 3/r$
1	0.5	39.414	59.6499	0.02994
2	0.5	39.3892	37.8421	0.01524
3	0.5	40.3303	28.8451	0.00982
4	0.5	47.8407	43.4009	0.02052
5	2.5	278.953	587.94	0.10232
6	1.5	1166.54	243.182	0.01536
7	1.5	0.00000	719.486	0.11680
8	2.5	407.138	72.0004	0.02611
9	0.5	310.343	105.561	0.01315
10	0.5	427.093	252.916	0.01742
11	2.5	9163.26	3562.11	0.02561
12	2.5	2392.84	704.767	0.02628
13	2.5	8553.14	2671.6	0.02507
14	2.5	9485.45	3236.89	0.02420
15	0.5	11514.2	7619.89	0.01109
16	0.5	590.031	352.785	0.01300
17	2.5	527.491	88.9685	0.01904
18	2.5	164.279	58.8051	0.04516
19	0.5	62.5098	42.4782	0.01748
20	1.5	107.167	51.1709	0.01630
21	1.0	13.4181	466.786	0.05235
22	0.5	173.146	110.739	0.00785
23	0.5	136.423	78.9171	0.01522
24	0.5	124.18	33.5084	0.00754
25	0.5	62.8576	89.4008	0.02959

hierarchical spatial model described later. The parameter τ^2 denotes the nugget which represents the small scale variability of the data. Parameter σ^2 denotes the partial sill which represents the variance of a spatially autocorrelated process without any nugget effect. The partial sill is also the difference between the sill and nugget. The sill is the total variance where the variogram appears to level off which is the sum of the partial sill and nugget ($\sigma^2 + \tau^2$), as shown in Figures 2, 3, 4, and 5.

Note that the Matérn semi-variogram at $\kappa = 0.5$ reduces to the exponential semi-variogram given by

$$\gamma(d) = \begin{cases} \tau^2 + \sigma^2 \left[1 - \exp\left(-rd\right)\right] & \text{if } d > 0 \\ 0 & \text{if } d = 0 \end{cases}$$

and as $\kappa \to \infty$ the Matérn reduces to a Gaussian semi-variogram. Also, κ and ν are both used interchangeably throughout this report, but both denote the smoothness parameter.

2.2 Variogram Attributes

We use the `gstat` `R` package to obtain least squares estimates (LSE) of the partial sill, range, and nugget parameters of the Matérn semi-variogram model. In Figure 1, we perform a sensitivity analysis with smoothness parameter κ equal to 0.5, 1.0, 1.5, and 2.5 for each time period. We choose the best fitted variogram per time period based on the SSErr (sum of squared error) criterion, that is, the variogram with the smallest SSErr value (highlighted in grey). The best fitted parameter estimates of the partial sill, nugget, spatial decay, and smoothness parameters for that time period will help guide in specifying the initial values and priors for the univariate spatial regression model.

Table 2 shows the parameter estimates for the best fitted Matérn variogram, with the exception of the first time period. For $t = 1$, we choose the exponential variogram since there is no large difference between the SSErr values.

The SSErr is a measure of the variance of the observed data from the true mean of the data. For time periods t = 5, 6, 7, 11, 12, 13, 14, 15, and 21 in Table 1, the SSErr value is very large indicating large data variability for that time period.

Looking at their corresponding variograms in Figures 2, 3, 4, and 5, the sample variogram does not follow common empirical variogram behaviour as described before and the Matérn curve appears flat as distance increases. Thus, if there is a strong

spatial correlation, the variogram function will be increasing, usually reaching a saturation point. If the Matérn function shows steady behavior after a given distance d_0, that indicates a potential absence of spatial correlation between pairs of points that have distances greater than d_0 for that time t.

The variograms for all time periods can be seen on the following pages 5 - 8 of this report. In Figure 2 at time period 2, the exponential curve stabilizes at about 500 km, indicating that there is no significant spatial correlation across points that are at distances above 500 km. Other time periods such as period 5 and 6 do not show a clear spatial relationship, with variogram values that do not have a clear upward trend that stabilizes after a given distance value.

For time period 7 in Figure 3, the Matérn curve stabilizes immediately at about 250 km. With all points given after 250 km, this indicates that there is no significant spatial correlation across all points at time period 7. In Figure 5 at time period 22 and 24, the exponential curve does not reach sill until at about 1250 km, indicating there is a significant spatial correlation across points that are distances below 1250 km.

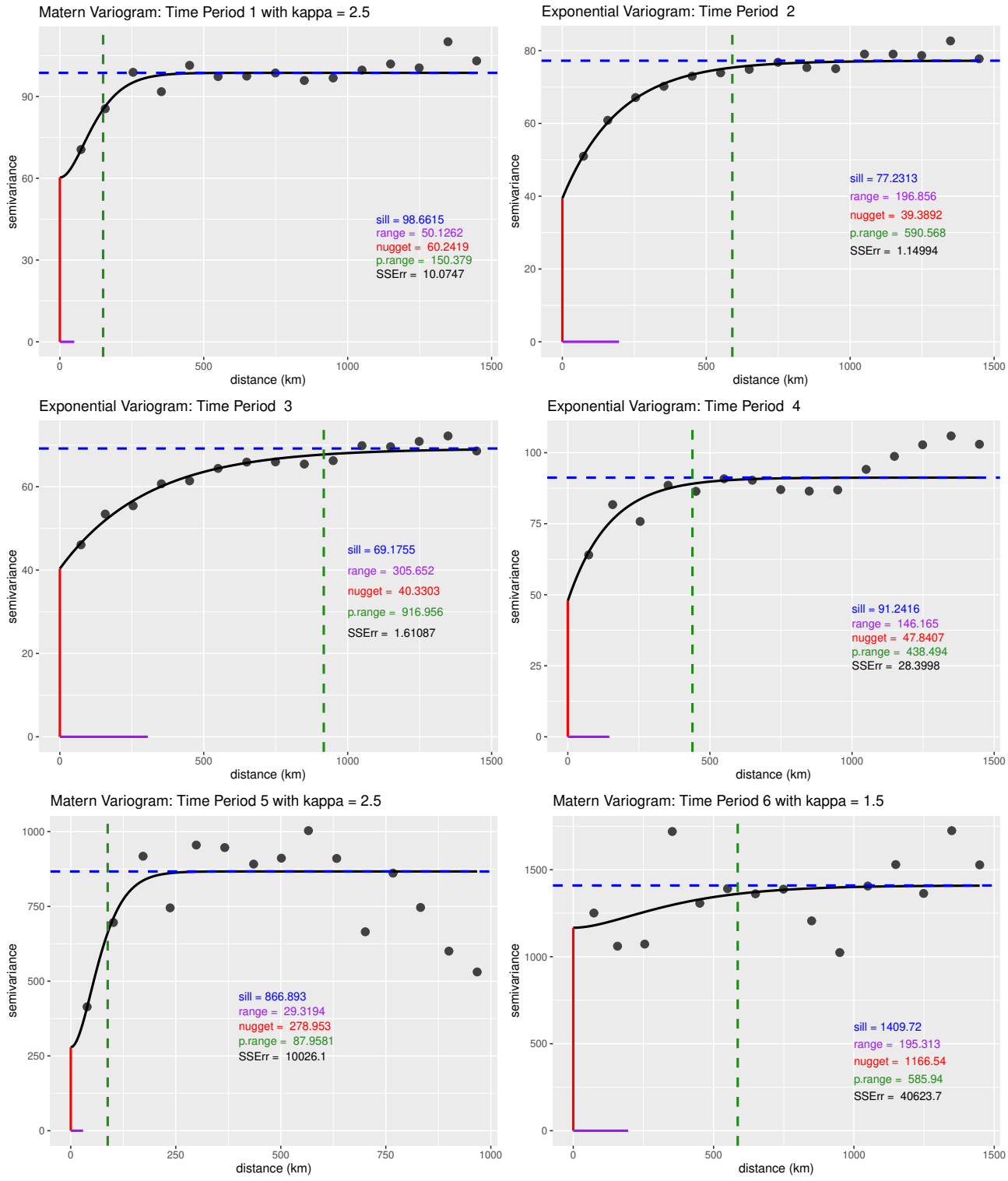

Figure 2: *Variograms for Time Periods 1 - 6.*

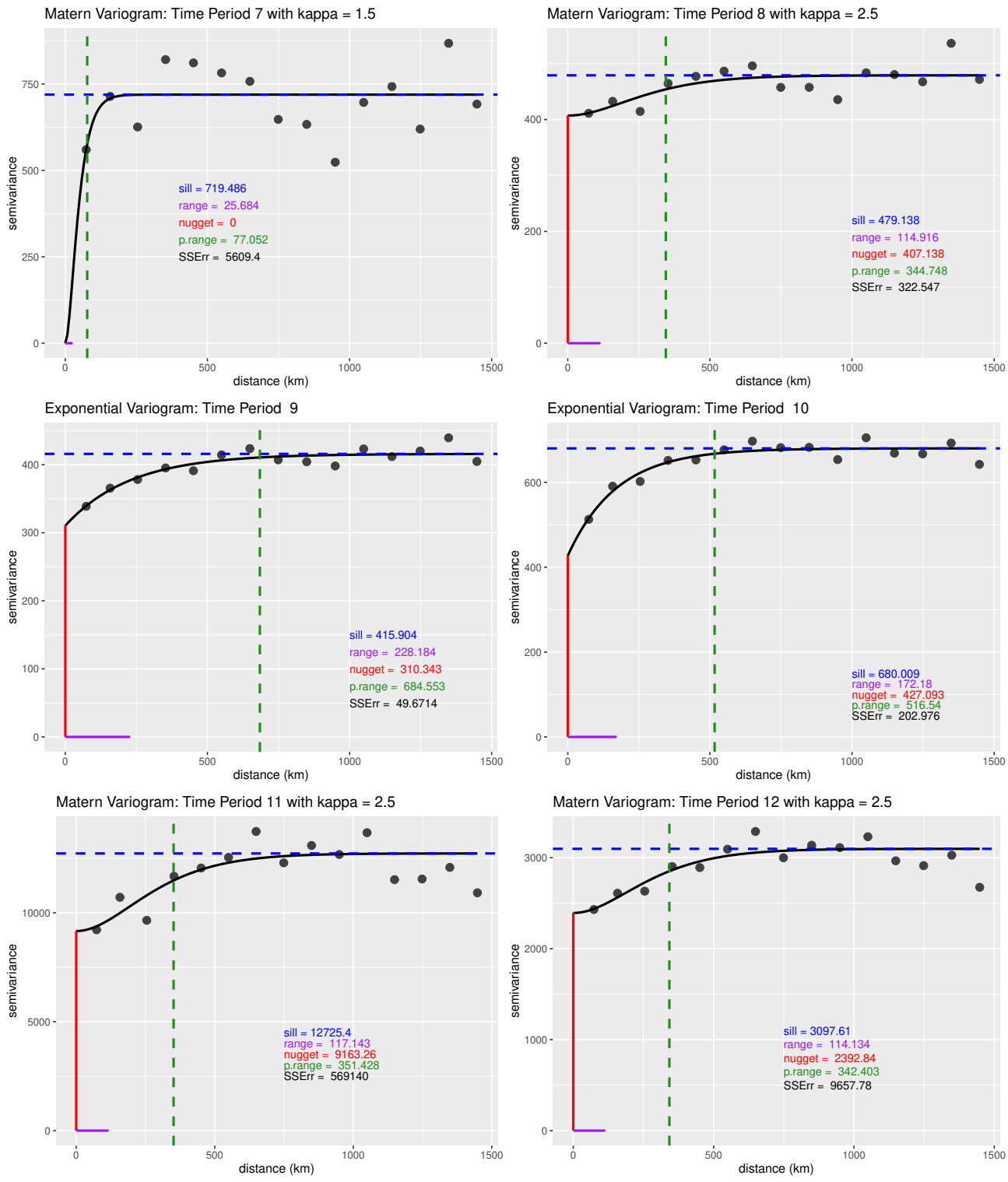

Figure 3: *Variograms for Time Periods 7 - 12.*

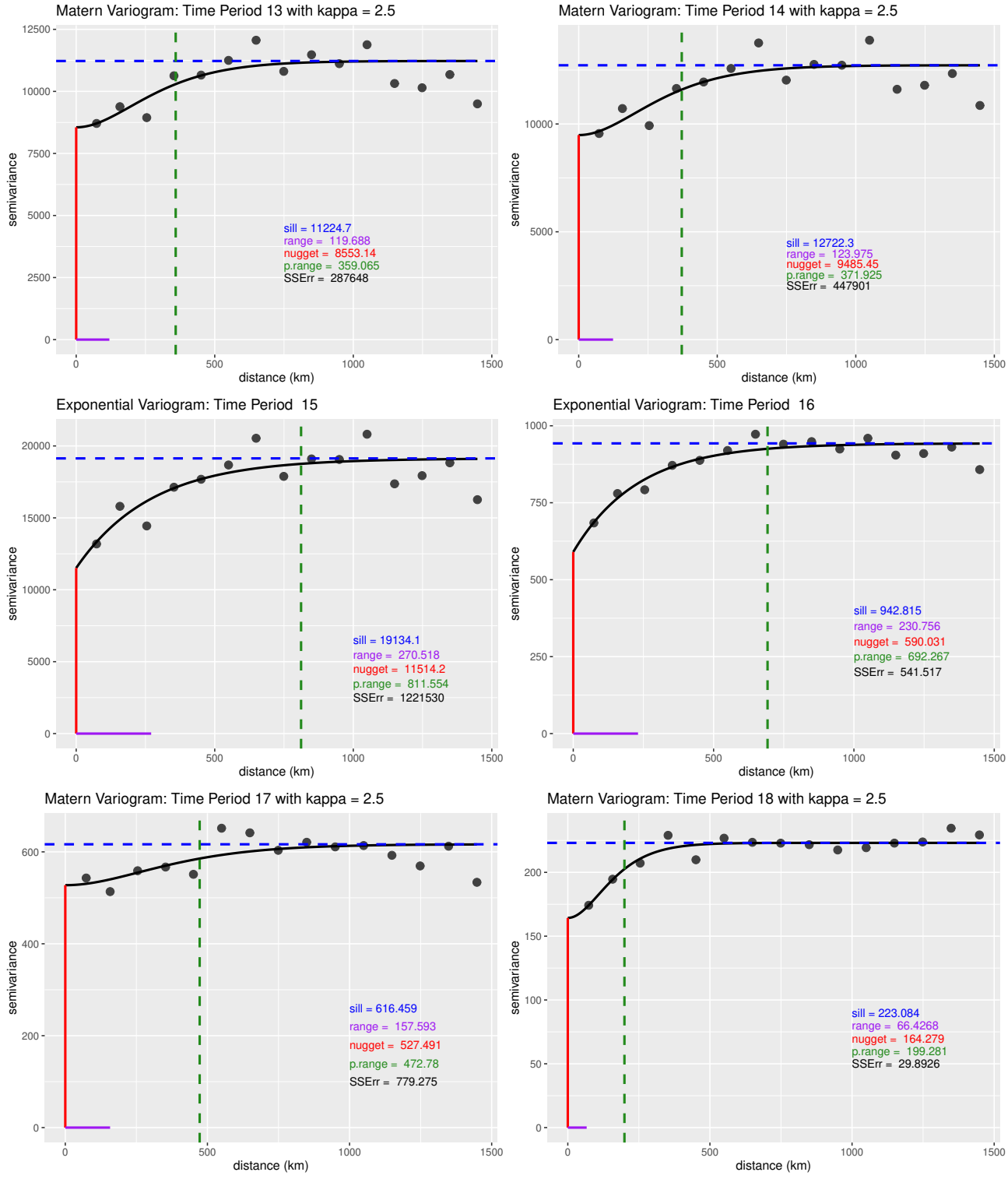

Figure 4: *Variograms for Time Periods 13 - 18.*

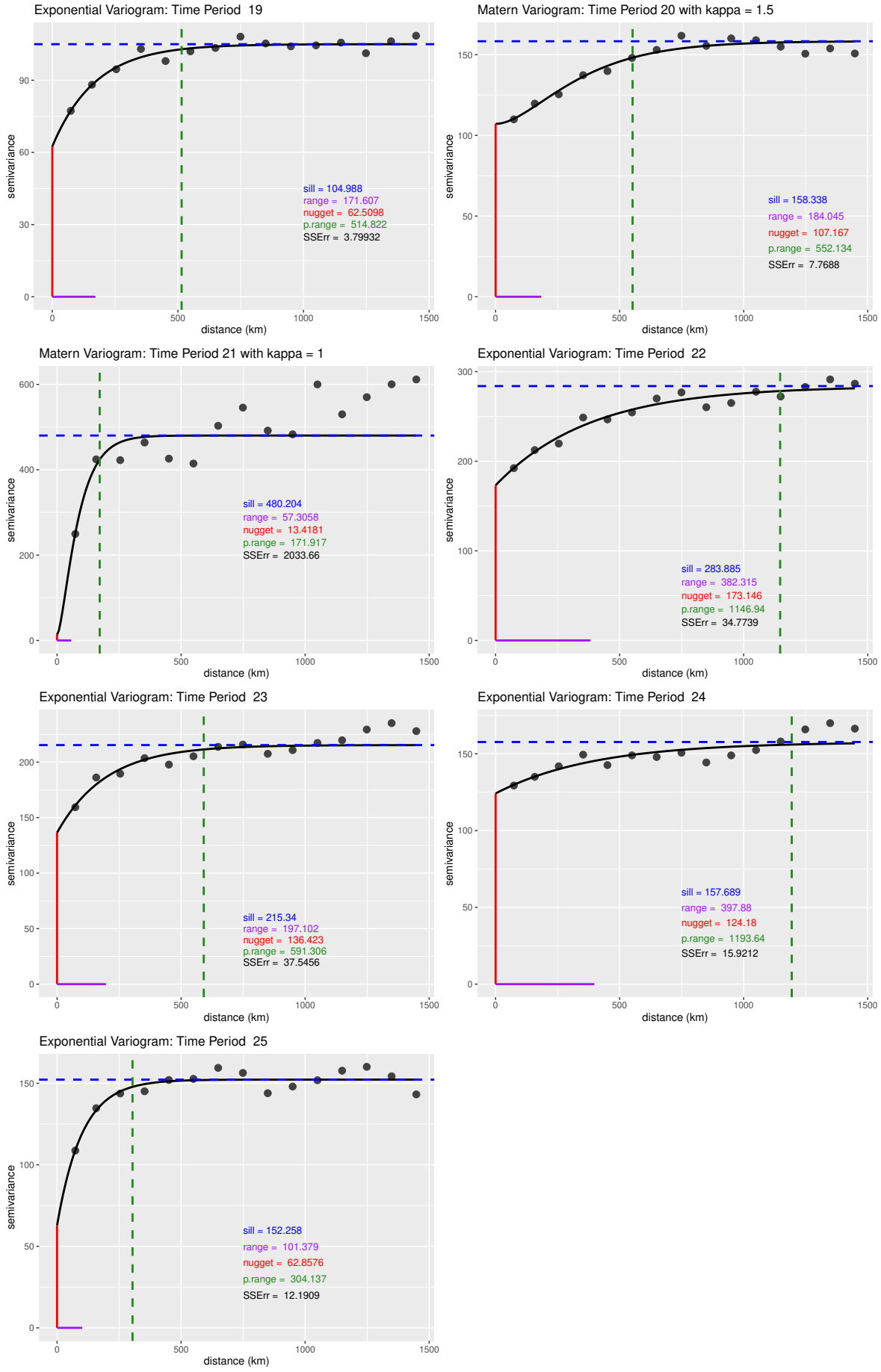

Figure 5: *Variograms for Time Periods 19 - 25.*

3 Univariate Bayesian Spatial Regression Model

For spatially indexed V_CVI denoted as $Y_t(s)$ at each time period t and location s for $t = 1, \ldots, 25$ and $s \in \{s_1, \ldots, s_{1000}\}$, we now consider the model

$$Y_t(s) = X_t(s)\beta_t + w_t(s) + \varepsilon_t(s), \quad (1)$$

where $X_t(s)$ and β_t are defined as before. Note that the model in (1) has an additional process $w_t(s)$, which is assumed to be a zero mean stationary Gaussian spatial process with parameters $\theta_t = \{\sigma_t^2, \phi_t, \nu_t\}$. The $w_t(s)$'s provide local adjustment (with structured dependence) to the mean and capture the effect of unmeasured or unobserved regressors with spatial patterns. Here we assume $\varepsilon_t(s)$ is a Gaussian white noise process with variance τ_t^2, intended to capture measurement error (S. Banerjee and Gelfand, 2015; Finley, Banerjee, and Carlin, 2007). This τ_t^2 is also called the nugget.

The model in (1) can also be written as

$$Y_t(s) = \mu_t(s) + \varepsilon_t(s), \qquad \varepsilon_t(s) \sim N(0, \tau_t^2)$$
$$\mu_t(s) = X_t(s)\beta_t + w_t(s), \quad w_t(s) \sim GP(0, C(d; \theta_t))$$

Customarily, one assumes stationarity, which means that $C(d;\theta_t) = C(s_i - s_j;\theta_t)$ is a function of the distance d between sites s_i and s_j. Isotropy goes further and specifies $C(d;\theta_t) = C(\|s_i - s_j\|;\theta_t)$, where d is the Euclidean distance between the sites s_i and s_j (Finley, Banerjee, and Carlin, 2007). $GP(0, C(d, \theta_t))$ denotes a spatial Gaussian process with the Matérn covariance function $C(d;\theta_t)$. We specify $C(d;\theta_t) = \sigma_t^2 \rho(\cdot;\phi_t,\nu_t)$, where $\rho(d;\phi_t,\nu_t)$ is the correlation function of choice which in this case is the Matérn with ϕ_t controlling the correlation decay and ν_t controlling the smoothness. Thus, our likelihood function is $Y_t(s) \sim N(X_t(s)\beta_t + w_t(s), \tau_t^2 + \sigma_t^2)$.

Given a prior on the model parameters $\pi(\beta_t, \theta_t)$, we have the posterior

$$\pi(\beta_t, \theta_t | Y_t, X_t) \propto L(Y_t | X_t, \beta_t, \theta_t)\, \pi(\beta_t, \theta_t).$$

For each of the 25 models, we assume the following priors

$$\pi(\beta_t) \sim N_2(0, 1000),$$
$$\pi(\sigma_t^2) \sim IG(3, 2),$$
$$\pi(\tau_t^2) \sim IG(2, 2),$$
$$\pi(\phi_t) \sim Unif(0.001, 1).$$

To sample from the posterior we use functions from the `spBayes` package such as `spLM` and `spRecover`. We run the MCMC for 5,000 iterations and discard the first 1000. In this package the Matérn covariance function is parameterized as,

$$C(d;\theta) = \begin{cases} \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\phi d)^{\nu} K_{\nu}(\phi d) & \text{if } d > 0 \\ \tau^2 + \sigma^2 & \text{if } d = 0 \end{cases}$$

Here, we assume ϕ_t lies in the interval between 0.001 to 1 based on our parameter estimates on Table 2. For the variance parameters σ_t^2 and τ_t^2, we assume weakly informative inverse-Gamma (IG) priors to allow for large variability. For the vector of regression coefficients β_t we assume a non-informative multivariate normal prior since we do not have the information of whether or not correlation is present.

We used a predictive process model with 108×2 matrix of knots (locations) using a grid over the extent of the domain as shown in Figure 6. This allows the predictive process to obtain the spatial effects $w_t(s)$. If this set of knots is not specified the package will automatically use the observed locations which in this case would be computationally unfeasible given our large number of locations. In general, the user chooses the number of knots $s^* = \{s_1^*, \ldots, s_h^*\}$ less than the observed number of locations, and defines the process

$$\tilde{w}_t(s) = \mathrm{E}[w_t(s) | w_t(s_i^*), i = 1, 2, \ldots, h].$$

Thus, $\tilde{w}_t(s)$ is called the predictive process. Replacing $w_t(s)$ with $\tilde{w}_t(s)$ in our univariate spatial model, described early in Section 3, yields the predictive process counterpart of the univariate spatial regression model (Finley, Banerjee, and Carlin, 2007).

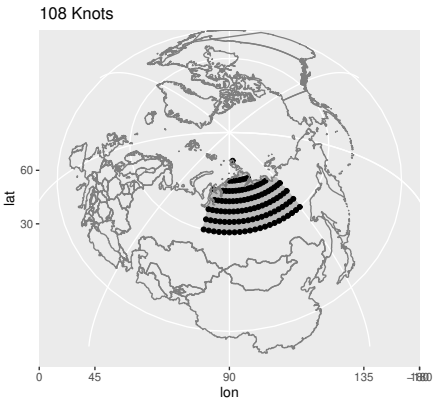

Figure 6: *108 knots used for the predictive spatial process.*

3.1 Overall Posterior Results

For each model run we fix the number of iterations to be 5,000, we fix the burn-in to 1000, and we fix the smoothness parameter ν_t with its corresponding estimate in Table 2. Thus, we have obtained all posterior parameter results across 25 time periods. Among the 25 time periods only 10 time periods did not achieve MCMC convergence. These time periods were at t = 5, 6, 7, 11, 12, 13, 14, 15, 21, and 25.

When assessing convergence, we check to see if the trace plots for the posterior parameters exhibit stable rapid up-and-down variation with no long-term trends or drifts. Thus any unstable differences within the MCMC, for example increasing or decreasing trends, indicates non-convergence.

Figure 7 shows the posterior means and corresponding 95% posterior intervals for the two regression coefficients, the baseline $\beta_{0,t}$ and the coefficient that describes the relationship between V_MAT and V_CVI $\beta_{1,t}$ for each time period t.

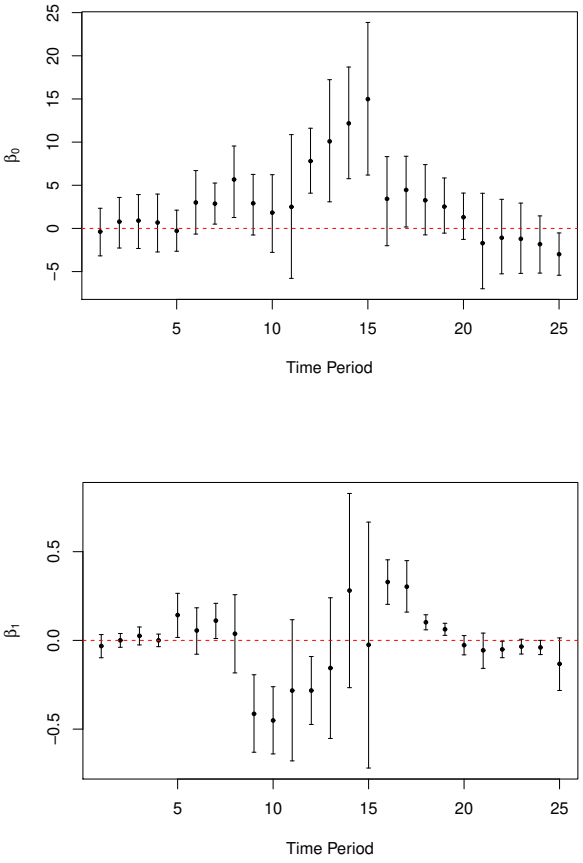

Figure 7: *Posterior means (dots) and 95% credible intervals for the regression coefficients.*

We see that the coefficients vary over different time periods, with a general increase in the baseline level and its associated variability for periods between $t = 6$ and $t = 18$. We also see that the 95% posterior intervals for $\beta_{t,1}$ do not contain zero for time periods $t = 9, 10,$ and 12, in which the linear relationship between the two variables V_MAT and V_CVI is negative, and $t = 5, 7, 16, 17, 18$ and 19, in which the linear relationship between these two variables is positive.

A negative relationship between the two variables V_MAT and V_CVI implies that in those time periods in upper Russia as V_MAT change increased, the V_CVI change decreased meaning vegetation productivity was unable change at the same rate. A positive relationship between the two variables V_MAT and V_CVI implies that in those time periods in upper Russia as V_MAT change increased, the V_CVI change increased meaning vegetation productivity was able to maintain the same rate of change.

The plots in Figure 8 (previous page) show the posterior means and corresponding 95% posterior intervals for the parameters σ_t^2, τ_t^2 and ϕ_t. In these plots we see that there is a marked increase in the values of τ_t^2 and ϕ_t for periods $t = 5, 6, 7$ and $t = 11, 12, 13, 14, 15$. Similarly, marked increase values of σ_t^2 but for periods $t = 5, 6, 7, 11, 14$. Note that for all of these periods MCMC convergence was not achieved within 5,000 iterations so further exploration is needed to determine if these results are reliable. We need to run the model for a significantly larger number of iterations.

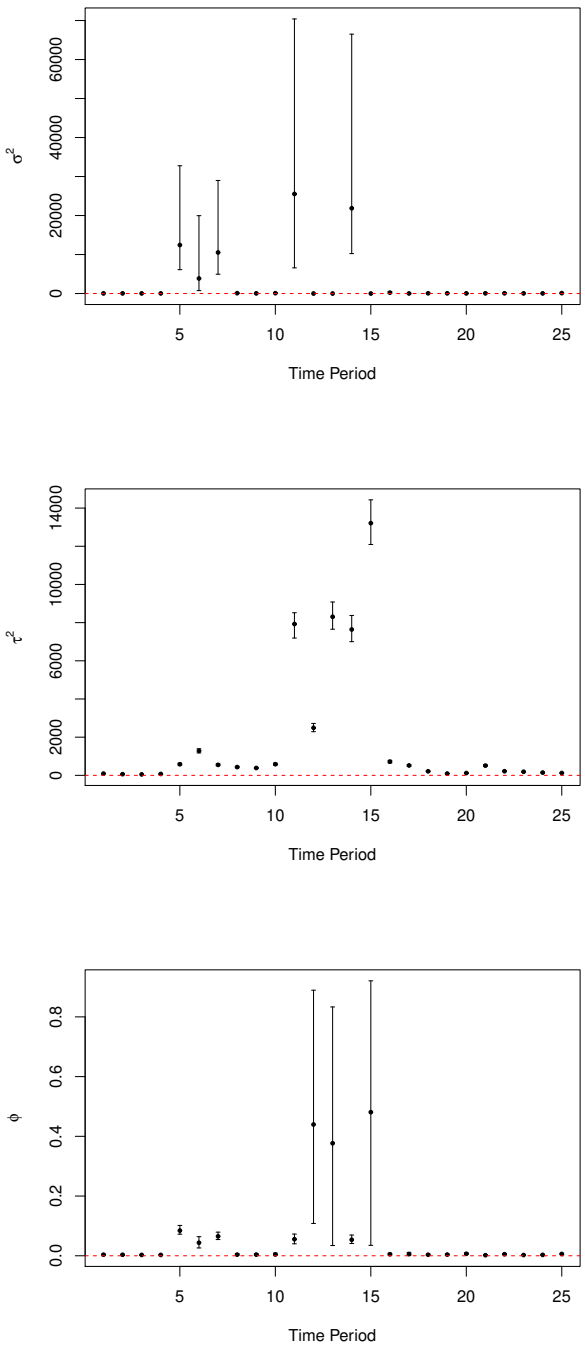

Figure 8: *Posterior means (dots) and 95% credible intervals for the Matérn parameters.*

In the following sections we will take a closer look at the results obtained for few of the time periods, specifically at t= 1, 5, 9, and 15.

3.2 Posterior and Predictive Results

In this section we will look at the results from the univariate Bayesian spatial regression models corresponding to time periods 1, 5, 9, and 15. For time periods 1, 9, MCMC convergence was achieved for all the corresponding models' parameters. Looking back at Figure 7, at t=1 the regression coefficient for V_MAT ($\beta_{1,1}$) was not significantly different from zero whereas $\beta_{t,1}$ was significantly different from zero at t=9. For time periods 5 and 15 MCMC convergence was not achieved for some parameters. At t=5, only parameter σ_5^2 displayed an unstable trace plot with slow mixing and random high spikes in increase. At t=15, the traces for parameter τ_{15}^2 showed slow mixing and random high spikes in increase, and parameter ϕ_{15} showed long term trends with slow up-and-down variation covering the entire interval (0.001,1).

3.2.1 Time Period 1

Time period 1 refers to the years 1981 - 1990. Using the same model structure as described earlier in section 3, we run the model using `spBayes` package and obtain the following 95% posterior credible intervals shown in Table 3. Here $\beta_{1,1}$ is not significantly different from zero as we also saw in Figure 7. This implies that there is no clear linear relationship between the variables V_CVI and V_MAT.

Table 3: *Posterior Estimates at t=1.*

	95% Posterior Credible Interval		
	50%	2.5%	97.5%
$\beta_{1,0}$	-0.371639	-3.177611	2.335278
$\beta_{1,1}$	-0.031697	-0.097747	0.032772
σ_1^2	35.09923	17.73714	76.30841
τ_1^2	92.11599	83.76416	101.2110
ϕ_1	0.003588	0.001745	0.005796

Figure 9 displays estimates of the spatial effects $w_1(s)$. These spatial effects suggest how processes other than V _MAT might structure distribution patterns in V_CVI.

Next we perform a posterior predictive analysis on a test set to check the predictive accuracy and power of the model. As stated in section 1.2, we have a test set of velocity data for 1825 locations and these locations are different from those used to train the model. From the test set, we sample 1000 random location points and use that corresponding velocity data to validate our model.

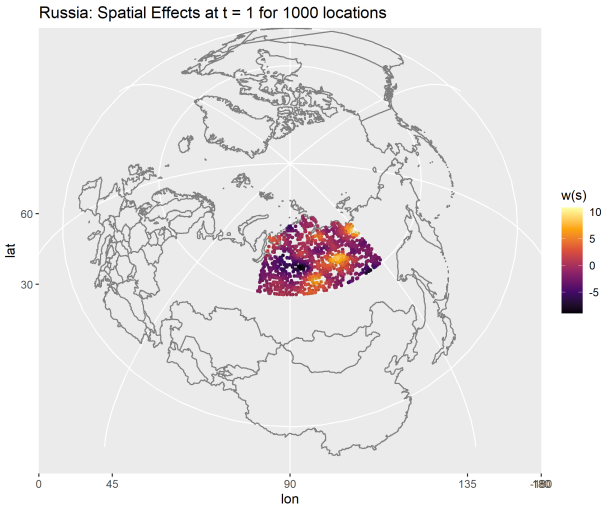

Figure 9: *Time Period 1: Spatial Effects.*

Let $Y_1^*(s^*)$ denote the density of a future V_CVI with $X_1^*(s^*) = (1, x_t(s^*))$ at new location s^*. We obtain 5000 draws from posterior predictive distributions,

$$Y_1^*(s^*) \sim N\left(X_1^*(s^*)\beta_1^* + w_1^*(s^*), \tilde{\tau}_1^2 + \tilde{\sigma}_1^2\right) \qquad (2)$$

for each of the new locations s^* using `spPredict`. Note that `spBayes` package implements the inclusion of the spatial effects for the predictions.

Figure 10, on the following page, shows the observed test set and the predicted test set obtained from the model. The predicted results more or less capture the general structure of the test set. Specifically the forward diagonal of lighter orange locations. The model does not perform well in terms of capturing the locations with large negative V_CVI values. An explanation for this is that perhaps the implementation of a Gaussian process on $w_t(s)$ is the component that is not capturing this variability, or there could be other possible predictors other than V_MAT driving the response variable V_CVI.

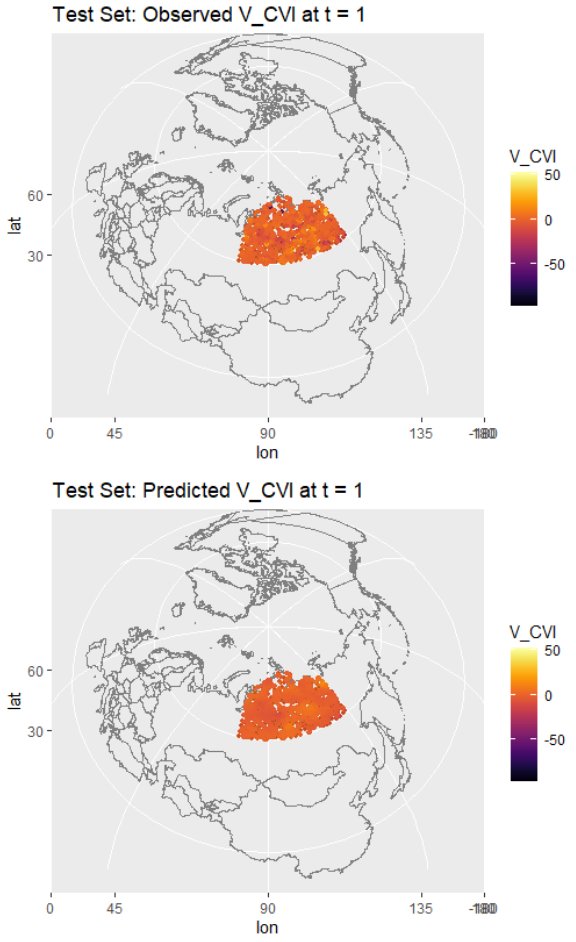

Figure 10: *Time Period 1: Observed and Predictive Test Set.*

Figure 11 summarizes the posterior predictive distributions in terms of the mean and 95% credible intervals for each of the test locations and compares these predictions with the actual observed values. We are looking to see if the observed response values are consistent with the corresponding predictive distributions; any predictive interval that does not cross of the blue line are possible outliers. Finally, I calculate the proportion of observations that fall inside the 95% prediction interval and I get a 97% accuracy rate which indicates the good fit model and thus our inference on $\beta_{1,1}$ is reliable. Once again, we see that for this time period the model does a reasonable job at predicting the new test locations, but has difficulties with some outlier observations.

3.2.2 Time Period 5

For time period 5 (1985-1994) we run the model and obtain the 95% posterior credible intervals shown on Table 4. Here $\beta_{5,1}$ is not significantly different from zero implying no clear correlation between V_CVI and V_MAT. However, for time period 5, MCMC convergence was not achieved for parameter σ_5^2. The trace for σ_5^2 across 4000 iteration displayed an unstable trace plot with slow mixing and high spikes of increase at ran-

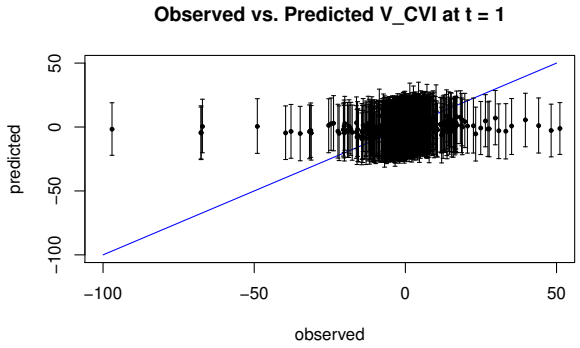

Figure 11: *Obs Test vs Pred Test at t=1.*

dom iterations. Parameters τ_5^2 and ϕ_5 achieved MCMC convergence with traces showing a stable up-and-down variation but the occurrence was not as rapid. Thus, for the results to be reliable, perhaps the model needed to run for a significantly larger number of iterations.

Table 4: *Posterior Estimates at t=5.*

	95% Posterior Credible Interval		
	50%	2.5%	97.5%
$\beta_{5,0}$	-0.27612	-2.64158	2.11788
$\beta_{5,1}$	0.14274	0.01681	0.26529
σ_5^2	12436.6	6118.74	32757.0
τ_5^2	581.113	536.694	640.327
ϕ_5	0.08426	0.07232	0.10162

Figure 12 are the plotted spatial effects $w_5(s)$. Here the $w_5(s)$ suggests no distribution patterns in V_CVI.

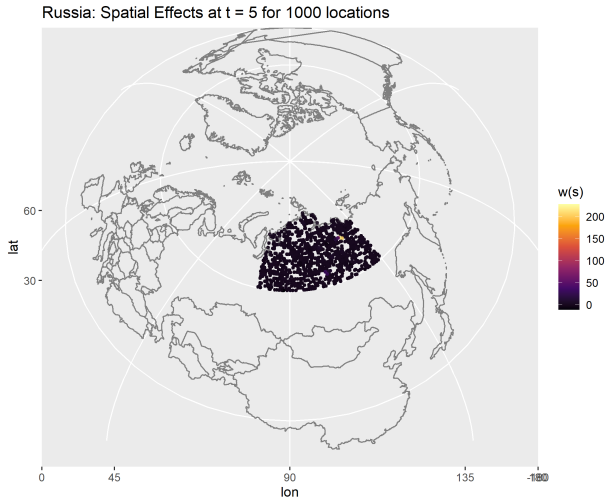

Figure 12: *Time Period 5: Spatial Effects.*

In Figure 13 shows the observed test set and predicted test set obtained from the model. The predicted results captured the structure of the observed test set such that the entire region follows that same spatial pattern as the spatial effects implied. Further analysis is needed to understand this spatial behavior such as using quantiles to bin the V_CVI values.

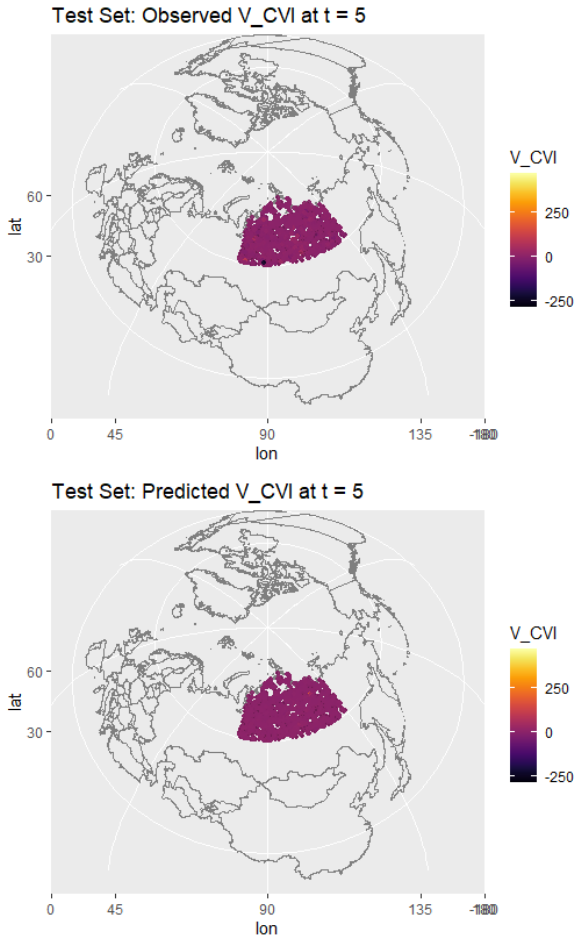

Figure 13: *Time Period 5: Observed and Predictive Test Set.*

Figure 14, summarizes the posterior predictive distribution in terms of the mean and 95% credible intervals for each of the test locations and compares these predictions with the actual observed values. We look to see if the observed response values are consistent with the corresponding predictive distributions; any predictive interval that does not cross the blue line are possible outliers. Thus, in Figure 14 we see poor predictive power for several outlier observations despite their corresponding huge 95% predictive credible intervals crossing the blue line.

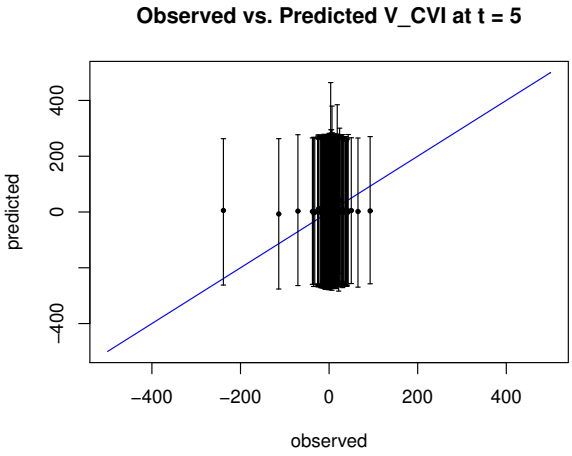

Figure 14: *Obs Test vs Pred Test at t=5.*

3.2.3 Time Period 9

Time period 9 refers to the years 1989 - 1998. Using the same model structure as descried in section 3, we run the model and obtain the following 95% posterior credible intervals shown in Table 5. Here $\beta_{9,1}$ is significant different from zero as we also saw in Figure 7. This implies that there is a negative linear relationship between the variables V_CVI and V_MAT. A negative relationship between the two variables V_MAT and V_CVI implies that in those time periods in upper Russia as V_MAT change increased, the V_CVI change decreased meaning vegetation productivity was unable change at the same rate.

Table 5: *Posterior Estimates at t=9.*

	95% Posterior Credible Interval		
	50%	2.5%	97.5%
$\beta_{9,0}$	2.91160	-0.76008	6.25868
$\beta_{9,1}$	-0.41400	-0.63025	-0.19360
σ_9^2	44.8000	13.1097	106.747
τ_9^2	385.013	354.797	421.116
ϕ_9	0.00385	0.00164	0.00709

Figure 15, on the following page, displays estimates of the spatial effects $w_9(s)$. These spatial effects suggest how the Gaussian process might structure distribution patterns in V_CVI.

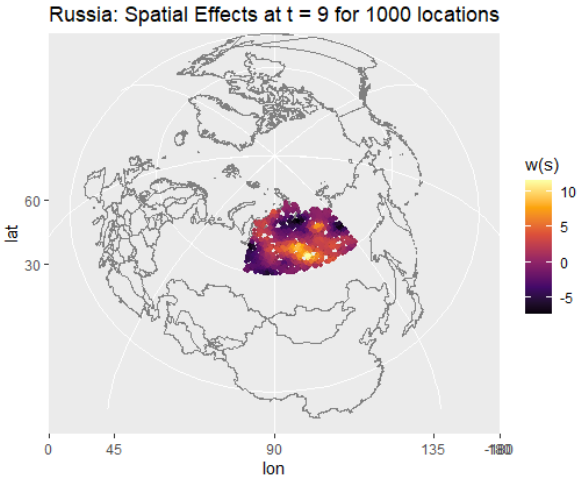

Figure 15: *Time Period 9: Spatial Effects.*

As done with previous time periods, we perform a posterior predictive analysis on the test set of 1000 random location points.

Figure 16 shows the observed test set and the predicted test set obtained from the predictive model. The predicted results captures the general structure of the test set. The predicted test set faintly captures location with high V_CVI values as in the observed test set. The model does not perform well in terms of capturing the locations with large negative V_CVI values. As stated before, an explanation for this is that perhaps the implementation of a Gaussian process on $w_t(s)$ is the component that is not capturing this variability, or there could be other possible predictors other than V_MAT driving the response V_CVI.

Figure 17 summarizes the posterior predictive distribution in terms of the mean and 95% credible intervals for each of the test locations and compares these to the predictions with the actual observed values. Again, we look to see if the observed V_CVI values are consistent with the corresponding predictive V_CVI values. Finally, I calculate the proportion of observations that fall inside the 95% prediction interval and get a 95.7% accuracy rate which indicates a relatively good model fit. However, as mentioned before the model does not adequately predict large (positive or negative) observed V_CVI values. Once again, our inference for this time period is that there is a negative correlation between V_CVI and V_MAT.

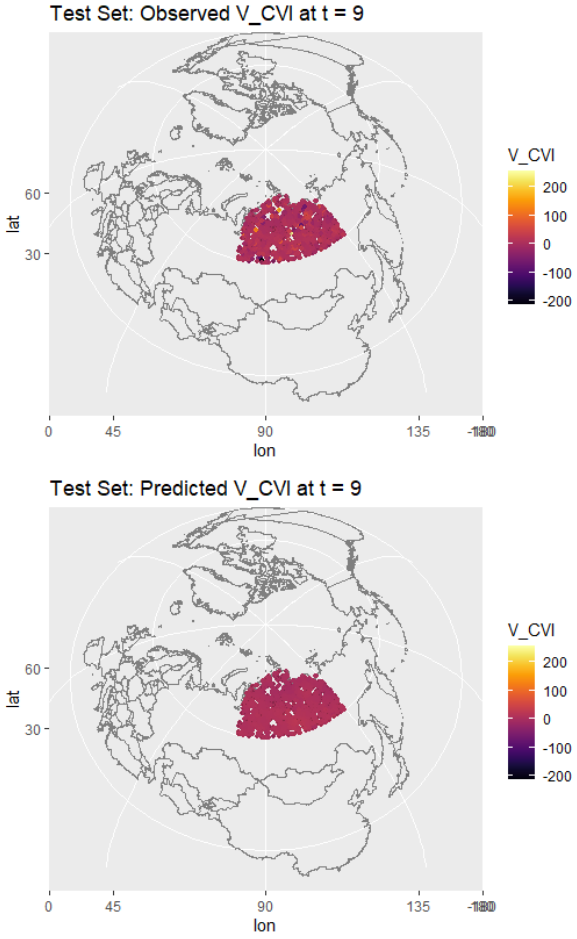

Figure 16: *Time Period 9: Observed and Predictive Test Set.*

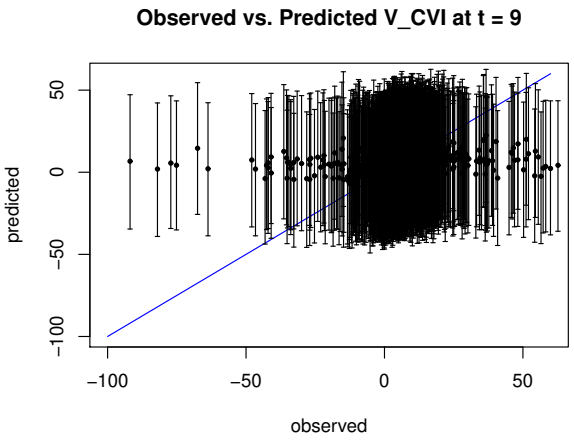

Figure 17: *Obs. Test vs Pred. Test at t=9.*

3.2.4 Time Period 15

Time period 15 refers to the years 1995 - 2004. From section 3.2, parameters σ^2_{15} and ϕ_{15} did not achieve MCMC convergence. At t=15, the traces for parameter σ^2_{15} showed slow mixing and random high spikes in increase across iterations. Parameter ϕ_{15} showed long term trends with slow up-and-down variation covering the entire interval (0.001,1). This indicates that the interval specification for the prior ϕ_{15} maybe incorrect which have affected the iteration values of σ^2_{15} since it is dependent on ϕ. We followed carefully the exact procedure in S. Banerjee and Gelfand, 2015 where the initial estimated value for ϕ_{15} is about 0.01 from table 2, well within the uniform prior specification of the interval 0.001 to 1.

In table 6, are the model parameter posterior means and 95% credible intervals. Here $\beta_{15,1}$ is not significantly different from zero implying no clear correlation between the response variable V_CVI and the explanatory variable V_MAT. However, the 95% credible interval band is very large suggesting large amount of uncertainty and thus our inference on $\beta_{15,1}$ is unreliable.

Table 6: *Posterior Estimates at t=15.*

	Posterior		
	50%	2.5%	97.5%
$\beta_{15,0}$	14.9816	6.18612	23.8650
$\beta_{15,1}$	-0.02496	-0.71943	0.66683
σ^2_{15}	0.69224	0.28783	2.42618
τ^2_{15}	13210.9	12093.7	14429.6
ϕ_{15}	0.48092	0.03497	0.92069

Figure 18 are the plotted spatial effects $w_{15}(s)$. Here $w_{15}(s)$ suggest no spatial distribution pattern in V_CVI.

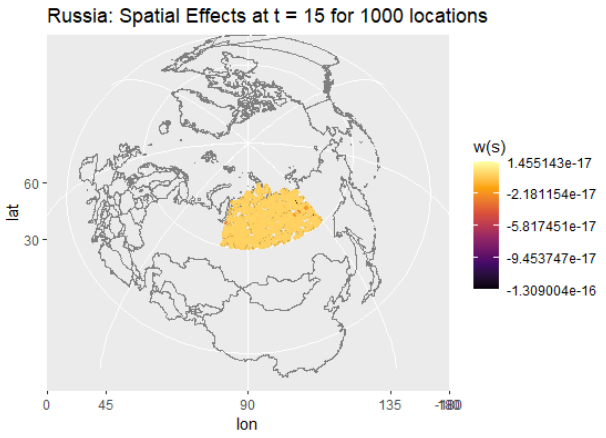

Figure 18: *Time Period 15: Spatial Effects.*

Figure 19 shows the observed test set and the predicted test set of V_CVI. We look to see if the predicted test set capture the same variability as the observed test values of V_CVI. Here we see that the predicted test set did not adequately capture variability from the observed test set. The entire region from the predicted test set is a lighter purple in comparison to the observed values. This implies the model has poor predictive power.

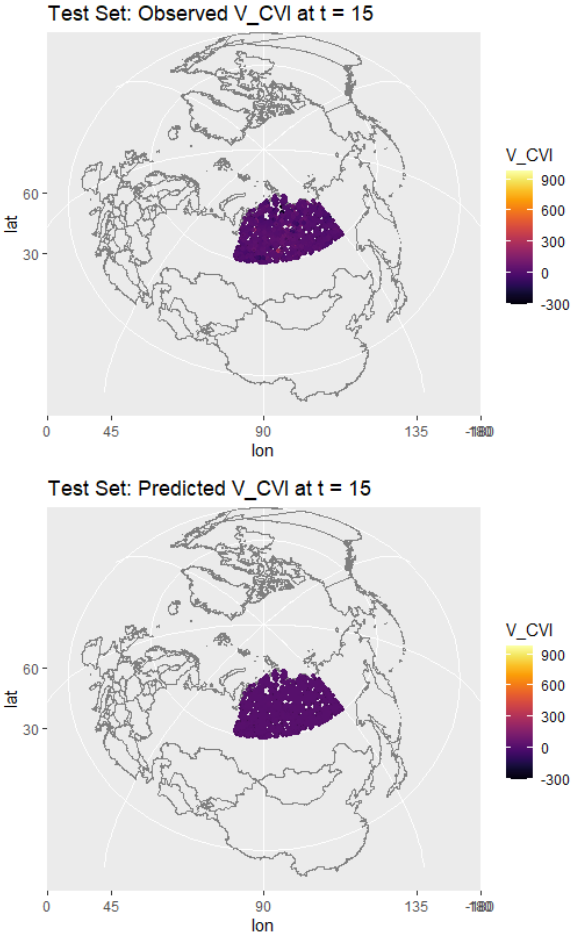

Figure 19: *Time Period 15: Observed and Predictive Test Set.*

Figure 20 summarizes the posterior predictive distribution in terms of the mean and 95% credible intervals for each of the test locations and compares these to the predictions with the actual observed values. We check to see if the observed V_CVI values are consistent with the corresponding predictive V_CVI values. Thus, in Figure 20, we also see poor predictive values due to the many outliers despite their corresponding huge 95% predictive credible intervals crossing the blue line. This reinforces that any inference on $\beta_{15,1}$ is unreliable due to large uncertainty in the data.

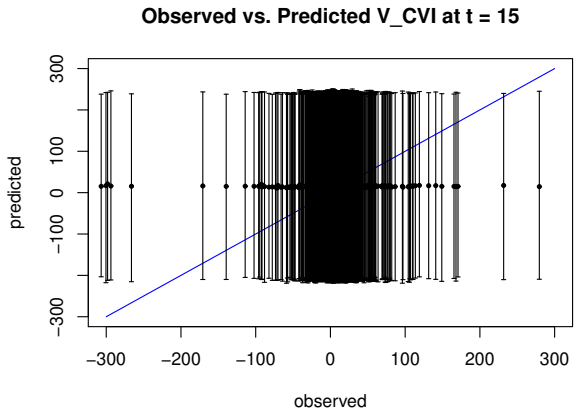

Figure 20: *Obs Test vs Pred Test at t=15.*

4 Conclusion

In this work we find that for those time periods in which MCMC convergence was achieved reliable inferences on $\beta_{t,1}$ can be made. Specifically, for time periods t= 9, 10, and 12 the linear relationship between the two variables V_CVI and V_MAT was negative. A negative relationship between the two variables implies that in those time periods in upper Russia, as V_MAT increased, the V_CVI decreased; meaning vegetation productivity was unable change at the same rate.

For t= 16, 17, 18 and 19, the linear relationship between V_CVI and V_MAT is positive. This implies that in those time periods, in upper Russia, as V_MAT change increased, the V_MAT change increased; meaning vegetation productivity was able to maintain the same rate of change.

Among the 25 time periods 10 time periods did not achieve MCMC convergence. These time periods were t = 5, 6, 7, 11, 12, 13, 14, 15, 21, and 25. In this report we specifically looked at t = 5 and 15 in which there was no clear spatial structure in the response V_CVI. It is for these time periods where we can not make reliable inferences on whether there is a clear linear relationship between V_CVI and V_MAT.

One way to potentially resolve some convergence issues is to increase the run time of the model. In this report we fixed the number of iterations to be 5,000 and discarded the first 1,000 iterations. For t=5 we discussed how the trace of τ_5^2 across 4,000 iterations displayed slow mixing with high increase in value at a few random iterations. If we had run the model say for 40,000 iterations then we could have thinned trace and removed the few iterations where there was a high increase in values. Future work should also include studies to assess the sensitivity of the posterior results with respect to the prior distributions. Users should caution that computational time increases depending on sample size and number of knots. Increasing the number of knots increases the computational time per iteration exponentially while increasing the sample size increases computational time per iteration at a slower rate in comparison.

Another future direction is to implement a new Bayesian model that includes a spatial structure on the regression coefficients, rather than on a separate spatial process. Such spatial structure can be modeled via convolution processes with Gaussian Markov random field (GMRF) priors. Obtaining posterior inference in this type of model would require custom-made MCMC implementation (Higdon, 2006).

Bibliography

Didan, K. and A. Barreto (2016a). *Multi-Sensor Vegetation Index and Phenology Earth Science Data Records Algorithm Theoretical Basis Document and User Guide Version 4.0*. URL: https://lpdaac.usgs.gov/products/vipphen_evi2v004/.

– (2016b). *NASA MEaSUREs Vegetation Index and Phenology (VIP) Phenology EVI2 Yearly Global 0.05Deg CMG*. NASA EOSDIS Land Processes DAAC. DOI: 10.5067/MEaSUREs/VIP/VIPPHEN_EVI2.004. URL: https://doi.org/10.5067/MEaSUREs/VIP/VIPPHEN_EVI2.004.

Diggle, Peter J. and Paulo J. Ribeiro (Mar. 2007). *Model-based Geostatistics*. English. Springer Series in Statistics. Springer. ISBN: 0387329072 978-0387329079.

Finley, Andrew O., Sudipto Banerjee, and Bradley P. Carlin (2007). "spBayes: An R Package for Univariate and Multivariate Hierarchical Point-Referenced Spatial Models". In: *Journal of Statistical Software* 19.4, pp. 1–24. URL: http://www.jstatsoft.org/v19/i04/.

Higdon, Dave (Oct. 2006). "A Primer on Space-Time Modeling from a Bayesian Perspective". In: DOI: 10.1201/9781420011050.ch6.

Huang, Mengtian et al. (Oct. 2017). "Velocity of change in vegetation productivity over northern high latitudes". In: *Nature Ecology and Evolution* 1. DOI: 10.1038/s41559-017-0328-y.

J.T. Abatzoglou S.Z. Dobrowski, S.A. Parks K.C. Hegewisch (2018). *Terraclimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015*. Scientific Data Journal. DOI: 10.1038/sdata.2017.191.

S. Banerjee, B.P. Carlin and A.E. Gelfand (2015). *Hierarchical Modeling and Analysis for Spatial Data*.

Stein, Michael (Feb. 2015). *Interpolation of Spatial Data: Some Theory For Kriging*. DOI: 10.1007/978-1-4612-1494-6.