Pre-Harvest Sugarcane Burning: a trend-cycle decomposition in spatio-temporal point process

Fernanda C. Valente* Márcio P. Laurini*

December 24, 2019

Abstract

In this paper we propose to analyze the process of change in the pattern of fires in the state of São Paulo, analyzing the efficiency of regulatory changes in the sugarcane production sector. To perform representation and inference procedures for the spatio-temporal point process we adopt a dynamic representation of Log-Gaussian Cox process. This representation is based on the modeling of the intensity function through a decomposition of components of trend, seasonality, cycles, covariates and spatial effects, where the spatial effects change over time through an autoregressive functional structure. The model were estimated using Bayesian inference through the integrated nested Laplace approximations algorithm. We use daily data from MODIS Thermal Anomalies/fires between January 2003 and December 2016 and the results have provided evidences that the trend component reflects the consistent reduction in fire occurrences related to Sugarcane Burning Law (Law 11.241/2002).

Keywords: Spatio-temporal models; Point Process; Log-Gaussian Cox Process; INLA

Resumo

Neste artigo propõe-se analisar as mudanças no padrão das queimadas no estado de São Paulo, analisando a eficiência das mudanças regulatórias na produção de cana-de-açucar. A fim de fazer inferências para o processo pontual espaço-temporal, adotou-se uma representação dinâmica do processo log-Gaussiano de Cox. Essa representação baseia-se na modelagem da função de intensidade através de uma decomposição em componentes de tendência, sazonalidade e ciclo, além de efeitos espaciais, onde tais efeitos espaciais variam no tempo através de uma estrutura funcional autorregressiva. O modelo foi estimado usando inferência Bayesiana através do algoritmo *Integrated Nested Laplace Approximations*. Utilizou-se dados diários do MODIS Thermal Anomalies/fires de janeiro de 2003 a dezembro

*Department of Economics - FEA-RP - Universidade de São Paulo. - Av. dos Bandeirantes 3900, 14040-905, Ribeirão Preto, SP, Brazil. Tel.: +55-16-33290867 - emails - fernanda.valente@usp.br and laurini@fearp.usp.br.

de 2016 e os resultados sugerem que o componente de tendência reflete a redução nas ocorrências de incêndios no estado de São Paulo, relacionada com a Lei da Cana-de-Açucar (Lei 22.241/2002).

Palavras-chave: Modelos espaço-temporais; Processos pontuais; Processo log-Gaussiano de Cox; INLA

1 Introduction

The sugarcane industry has an important role in the world, with the world's largest crop by product according to Food and Agriculture Organization of the United Nations (FAOSTAT, 2017). The three most important countries in sugarcane production are Brazil, India and China (FAOSTAT, 2017). In Brazil, the state of São Paulo is the major producer, representing 55% of the national production in 2017/2018, according to Sugarcane Industry Union (UNICA). After 2003, with the introduction of flex-fuel vehicles in Brazil, the sugarcane production in the country has expanded sharply, increasing from 320 million tons in the harvest of 2002/2003 to 641 million tons in 2017/2018 (UNICA). The production growth can also be related with the increased international interest in alternative energy sources due to environment issues, promoting the use of ethanol to replace fossil fuels to reduce emissions of greenhouse gases. Although ethanol is considered as clean fuel, because it originates from sustainable sources, concerns on its sustainability has raised. Destruction or damage of high-biodiversity areas, degradation of soils and health problems related to pre-harvest burning are some negative impacts from large-scale sugarcane production (GOLDEMBERG; COELHO; GUARDABASSI, 2008).

Sugarcane can be harvest using two different technologies: by hand with necessary pre-harvest field burning or mechanically, with or without pre-harvest burning. Burning of sugarcane residues is responsible for a great amount of pollutant gases in atmosphere which has significant negative effects on human health, as reported in several studies showing the correlation between emissions generated by sugarcane burning and respiratory diseases (e.g., Cançado et al. (2006) and Paraiso and Gouveia (2015)). Due to environmental and economic reasons, the state of São Paulo implemented a legislation (Law N°11.241/2002) to banish the burning practice gradually by 2031. The law establishes two different calendars to adopt mechanical harvest, according to soil slope: the deadline for properties with slope less or equal 12% is 2021, whereas for areas with slope greater then 12% the deadline is 2031. However, in 2007 the "Environmental Protocol" led to acceleration the deadlines, setting 2014 as a target year for properties with slope less or equal 12% and 2017 for non-mechanisable areas.

Aguiar et al. (2011) analyze sugarcane harvest, based on remote sensing images, to evaluate the effectiveness of the "Environmental Protocol" and conclude that the Protocol has been effective with a positive impact on the increase of green harvest, especially on recently expanded sugarcane fields. Capaz, Carvalho and Nogueira (2013) assess the emissions of greenhouse gases (tonCO2eq/ha) in sugarcane harvesting as a function of the simultaneous reduction of previous burning and increase in the use of mechanization in the São Paulo state. Their results show a 39.3% reduction in greenhouse gases emissions between 1990 and 2009. The effects of the Law 11.241/2002 in the sugarcane production sector of Paraná state, second largest national producer of sugarcane, are discussed in Junqueira, Sterchile and Shikida (2011). They analyze the reasons behind the adoption of the mechanized harvest in the sugar and alcohol sector of Paraná state through field research, showing that that the mechanization of the sugarcane harvest in Paraná was due to market pressures allied to standards and costs imputed by São Paulo state, which is one of the reasons of the new technological pattern in the Brazilian

production of sugar and alcohol.

The aforementioned facts show the importance of modelling the spatio-temporal distribution of the fire occurrences and the impacts of regulatory changes on sugarcane burning in the state of São Paulo. Fires can be represented by point processes, which are used to represent occurrences of events in space associated to their spatial coordinates. Moreover, spatial point processes can be associated to temporal instant, which corresponds to spatio-temporal point processes, with applications in several areas like biology, climate, neurology, demography and forestry.

In the modeling context, the Log Gaussian Cox process (LGCP) define a class of useful models to deal with spatio-temporal point process. LGPC, proposed by Møller, Syversveen and Waagepetersen (1998), can be seen as a particular case of Cox processes, where the log-intensity function is a Gaussian random field. Cox processes are inhomogeneous Poisson processes, which the intensity function is a stochastic process. The LGCP formulation is attractive for theoretical and empirical reasons. In special, it allows to add structures in the stochastic intensity function, which permits to control for a general processes of spatial dependence, and also to control for omitted variables with spatial dependence.

Recently Illian et al. (2010) proposed to use integrated nested Laplace approximations (INLA) to fit LGCP models. They construct a Poisson approximation to the true LGCP likelihood based on regular lattices, and counting the number of points in each cell. However, Simpson et al. (2016) show that this approach could be very inefficient since the quality of the likelihood approximation depends on the size of the grid, i.e., it is necessary to construct a much fine grid, which is computationally intense. Therefore, Simpson et al. (2016) propose to use stochastic partial differential equations (SPDE) approach to approximate a Gaussian Field (GF) to a Gaussian Markov Random Fields (GMRF), which is defined by sparse matrices, due to Markov properties. The main result provided by Simpson et al. (2016) is that is not necessary to construct too fine grids in order to approximate LGCP likelihood, allowing computationally effective methods. Moreover, according to Simpson et al. (2016), the LGCP formulation fits naturally within the Bayesian hierarchical modelling framework in combination with Integrated Nested Laplace Approximations (INLA) approach, proposed by Rue, Martino and Chopin (2009), avoiding convergence problems related to Markov Chain Monte Carlo (MCMC) methods.

Another applications of spatio-temporal LGCP to the point process can be found in Pereira et al. (2013), Serra et al. (2014) and Rostami et al. (2017). Serra et al. (2014) analyzed patterns of wildfires incidents in Catalonia from 1994 to 2008 through a spatio-temporal LGCP, estimated using Bayesian inference for GMRF through INLA algorithm. They found statistical evidence that areas closer to humans have more human induced wildfires whereas areas farther have more naturally occurring wildfires. Rostami et al. (2017) propose to model deaths due to substance abuse in Iran between 2005 and 2014 using a LGCP taking SPDE approach and using INLA method to perform Bayesian inference. The considered model incorporated known and unknown influential factors in order to describe spatio-temporal variations in the relative risk of substance abuse mortality. Their results suggest spatial heterogeneity and inequality by gender in deaths related to substance abuse.

In this paper we discuss how to perform inference procedures for spatio-temporal processes using a dynamic representation of a LGCP. This representation is based on the modeling of the intensity function through a decomposition of components in trend, seasonality, cycles, covariates and spatial effects. We also assume that spatial effects are based on a dynamic formulation, which allows it to change over time, based on an autoregressive functional structure. Our model follows the approach introduced by Laurini (2019) which proposes a method for decomposition of trend, cycle and seasonal components in spatio-temporal models, where the spatial component

is based on a continuous projections of spatial covariate functions. Indeed, our proposed model can be viewed as an extension of Laurini (2019) approach to spatial point process, where the dynamics in point process are captured by persistent term and mean reverting components, plus the spatial term, which is time-varying by the autoregressive group structure. We apply this methodology to evaluate the process of change in the pattern of fire occurrences in the state of São Paulo, analyzing the efficiency of regulatory changes in the sugarcane production sector. The results indicate that the trend component reflects the consistent reduction in fire occurrences related to Sugarcane Burning Law (Law 11.241/2002).

This article is organized as follow. Section 2 contains a description of the statistical approach. Section 3 presents the data. Section 4 shows the results with some discussion. Section 5 concludes.

2 Methods

2.1 Log-Gaussian Cox Process

Cox processes are inhomogeneous Poisson processes, which the intensity function is stochastic. Log Gaussian Cox processes, proposed by Møller, Syversveen and Waagepetersen (1998), are a particular case of Cox processes, where the log-intensity function is a Gaussian random field. In order to analyze the process of change in the pattern of fires in the state of São Paulo, we consider a dynamic version of a spatial LGCP where the intensity function is modelled as follows:

$$\begin{aligned}
\log \lambda(s,t) &= \mu_t + s_t + c_t + z(s,t)\beta + \xi(s,t) \\
\mu_t &= \mu_{t-1} + \eta_\mu \\
s_t &= s_{t-1} + s_{t-2} + \ldots + s_{t-m} + \eta_s \\
c_t &= \theta_1 c_{t-1} + \theta_2 c_{t-2} + \eta_c \\
\xi(s,t) &= \Theta\xi(s,t-1) + \omega(s,t) \\
Cov(\omega(s,t)) &= C(h)
\end{aligned} \tag{1}$$

where μ_t is the long term trend, s_t represents the seasonal components, c_t is a cycle component represented by an second-order autoregressive process with complex roots, $z(s,t)$ is a set of covariates observed in the location s and period t, and $\xi(s,t)$ are the spatial random effects represented by the Gaussian process $\omega(s,t)$ continuously projected in space and given by

$$Cov(\omega(s,t)\omega(s',t')) = \begin{cases} 0 & \text{if } t \neq t' \\ \sigma^2 C(h) & \text{if } t = t' \end{cases} \qquad \text{for } s \neq s' \tag{2}$$

where $C(h)$ is a covariance function of the Matérn class, which can be written as

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)}(\kappa||h||)^\nu K_\nu(\kappa||h||) \tag{3}$$

where $= ||s - s'||$ is the Euclidean distance between locations s and s', $\kappa > 0$ is a spatial scale parameter, $\nu > 0$ is the smoothness parameter and K_ν is a modified Bessel function. The marginal variance σ^2 is defined by:

$$\sigma^2 = \frac{\Gamma(\nu)}{4\pi\kappa^{2\nu}\tau^2\Gamma(\nu + \frac{d}{2})} \tag{4}$$

where τ is a scaling parameter and d is the space dimension. Following Lindgren, Rue and Lindström (2011), we adopt a parameterization in terms of $\log\tau$ and $\log\kappa$ for the covariance function:

$$\log\tau = \frac{1}{2}\log\left(\frac{\Gamma(\nu)}{\Gamma(\alpha)(4\pi)^{d/2}}\right) - \log\sigma - \nu\log\rho$$
$$\log\kappa = \frac{\log(8\nu)}{2} - \log\rho \tag{5}$$

where $\rho = \frac{(8\nu)^{1/2}}{\kappa}$. This representation is advantageous since, conditional on the value of ν, it is necessary to estimate only two parameters.

According to Simpson et al. (2016), the LGCP formulation fits naturally within the Bayesian hierarchical modelling framework and are latent Gaussian models, therefore, it may be fitted using the INLA approach of Rue, Martino and Chopin (2009).

2.2 Likelihood approximation

Considering a bounded region $\Omega \in \mathbb{R}^2$, it follows that the likelihood for an LGCP associated with data $Y = \{s_i \in \Omega : i = 1, \ldots, n; t = 1, \ldots, T\}$ is of the form

$$\pi(Y|\lambda) = \exp\left(|\Omega| - \int_\Omega \lambda(s,t)ds\right)\prod_{t=1}^{T}\prod_{i=1}^{n_t}\lambda(s_i,t). \tag{6}$$

Due to the doubly-stochastic property of the intensity function, the likelihood in (6) is analytically intractable. Since the term $\omega(s,t)$ corresponds to a GF with Matérn covariance, it is possible to use the SPDE approach, proposed by Lindgren, Rue and Lindström (2011), to approximate the initial GF to a GMRF, which has very good computational properties due to Markov structure, providing a sparse representation of the spatial effect through a sparse precision matrix (KRAINSKI et al., 2018).

The first main important result, provided by Whittle (1954) and extensively used by Lindgren, Rue and Lindström (2011), is that a GF x(s) with the Matérn covariance function is a stationary solution to the linear fractional SPDE

$$(\kappa - \Delta)^{\alpha/2}x(s) = W(s), \quad s \in \mathbb{R}^d, \quad \alpha = \nu + d/2, \quad \kappa > 0, \quad \nu > 0 \tag{7}$$

where $\Delta = \sum_{i=1}^{d}\frac{\partial^2}{\partial s_i^2}$ is the Laplacian operator and $W(s)$ is a spatial white noise. Therefore, in order to find a GMRF approximation of a GF, we first need to find the stochastic weak solution of SPDE (7).

Lindgren, Rue and Lindström (2011) proposed to use Finite Method Elements (FEM) to construct an approximated solution of SPDE. By proposing FEM, Lindgren, Rue and Lindström (2011) provided a solution for the case of irregular grids, since one of the big advantages of the FEM method is the irregularity, i.e., the domain can be divided into a irregular non-intersecting set of elements. The approximation of SPDE solution is given by

$$x(s,t) \approx \tilde{x}(s,t) = \sum_{j=1}^{n} w_j\varphi_j(s,t) \tag{8}$$

where n is the number of vertices of the triangulation, $\{w_j\}_{j=1}^n$ are the weights with Gaussian distribution and $\{\varphi_j\}_{j=1}^n$ are the basis functions defined for each node on the mesh. Here, the basis functions are chosen to be piecewise linear on each triangle:

$$\varphi_l(s,t) = \begin{cases} 1 & \text{at vertex } l \\ 0 & \text{elsewhere} \end{cases} \tag{9}$$

The stochastic weak solution of (7) is found by requiring

$$\{\langle \phi, (\kappa^2 - \Delta)^{\alpha/2} x \rangle\}_\Omega \stackrel{d}{=} \{\langle \phi, W \rangle\}_\Omega, \tag{10}$$

where $\{\phi_i(s), i = 1, \ldots, m\}$ are test functions and $\stackrel{d}{=}$ denotes equality in distribution. Replacing (8) in (10) gives us

$$\{\langle \phi_i, (\kappa^2 - \Delta)^{\alpha/2} \varphi_j \rangle\}_\Omega \mathbf{w} \stackrel{d}{=} \{\langle \phi_i, W \rangle\}_\Omega, \tag{11}$$

for $i = 1, \ldots, m$, where m is the number of test functions.The finite dimensional solution is obtained by finding the distribution for the Gaussian weights in equation (8) that fulfils (11) for only a specific set of test functions, with $m = n$. Choosing $\alpha = 2$ and $\phi_k = \varphi_k$[1] yields

$$\left(\kappa^2 \{\langle \varphi_i, \varphi_j \rangle\} + \{\langle \varphi_i, -\Delta \varphi_j \rangle\}\right) \mathbf{w} \stackrel{d}{=} \{\langle \varphi_i, W \rangle\}. \tag{12}$$

Define the $n \times n$ matrices, $\mathbf{C}$ and $\mathbf{G}$ as

$$\begin{aligned} C_{ij} &= \langle \varphi_i, \varphi_j \rangle \\ G_{ij} &= \langle \nabla \varphi_i, \nabla \varphi_j \rangle, \end{aligned} \tag{13}$$

then a weak solution to (7) is given by (8), where

$$(\kappa^2 \mathbf{C} + \mathbf{G}) \mathbf{w} \sim N(0, \mathbf{C}) \tag{14}$$

and the precision of the weights, $\mathbf{w}$, is

$$\mathbf{Q}_{\alpha=\mathbf{2}} = (\kappa^2 \mathbf{C} + \mathbf{G})^T \mathbf{C}^{-\mathbf{1}} (\kappa^2 \mathbf{C} + \mathbf{G}). \tag{15}$$

Although G_{ij} and C_{ij} are sparse matrices, $\mathbf{C}^{-1}$ is not sparse. The solution is to replace $C_{ij} = \langle \varphi_i, \varphi_j \rangle$ by the diagonal matrix $C_{ii} = \langle \varphi_i, 1 \rangle$, that yields a Markov approximation. Therefore, $\mathbf{w}$ is a GMRF with precision matrix defined by (15).

Replacing the GF $\omega(s, t)$ by the GMRF approximation $\tilde{\omega}(s, t)$ in equation (1), we obtain

$$\tilde{\xi}(s, t) = \Theta \tilde{\xi}(s, t - 1) + \tilde{\omega}(s, t). \tag{16}$$

Simpson et al. (2016) show that by replacing $\xi(s, t)$ by $\tilde{\xi}(s, t)$ in the intensity function (1) and approximating the integral (6) by a quadrature rule, results that the approximate likelihood consists of $(n + n_t)T$ independent Poisson random variables, where n is the number of vertices and n_t is the number of observed fires.

2.3 INLA method

In this section we present a brief review of the main steps of the INLA method. A detailed discussion can be found in Rue, Martino and Chopin (2009). INLA method provides accurate and efficient approximations on Bayesian hierarchical models that can be represented as a latent Gaussian model (LGN), focuses on models known as Gaussian Markov random field (GMRF). The approach obtains accurate approximations of the marginal posterior distributions of the latent factors, written as

$$\pi(\xi_i | Y) = \int \pi(\xi_i | \theta, Y) \pi(\theta, Y) d\theta \tag{17}$$

[1] When $\phi_k = (\kappa^2 = \Delta)^{1/2} \varphi_k$ for $\alpha = 1$ and $\phi_k = \varphi_k$ for $\alpha = 2$, these two approximations are denoted the *least squares* and the *Galerkin* solution, respectively.

and the marginal posterior distribution of the hyperparameters, given by

$$\pi(\theta_j|Y) = \int \pi(\theta|Y)d\theta_{-j} \tag{18}$$

where θ_{-j} denotes the vector θ without its jth element. In practice, the INLA method can be implemented in three steps. The first step is to find an approximation to the conditional distribution of latent factors:

$$\pi(\xi_i|\theta, Y) \propto \frac{\pi(\xi, \theta, Y)}{\tilde{\pi}(\xi_{-i}|\xi_i, \theta, y)}|_{\xi_{-i}=\xi^*_{-i}(\xi_i,\theta)} =: \tilde{\pi}(\xi|\theta, Y) \tag{19}$$

where ξ_{-i} denotes the latent factors ξ with the ith element omitted, $\tilde{\pi}(\xi_{-i}|\xi, \theta, Y)$ is the Gaussian approximation of $\pi(\xi_{-i}|\xi, \theta, Y)$ and $\xi^*_{-i}(\xi, \theta)$ is its mode. Based on the approximations in (19), the second task is to obtain $\pi(\theta|Y)$ using another Laplace approximation

$$\pi(\theta|Y) \propto \frac{\pi(\xi, \theta, Y)}{\tilde{\pi}(\xi|\theta, Y)}|_{\xi=\xi^*(\theta)} =: \tilde{\pi}(\theta|Y) \tag{20}$$

where $\tilde{\pi}(\xi|\theta, Y)$ is the Gaussian approximation of $\pi(\xi|\theta, Y)$ and $\xi = \xi^*(\theta)$ is its mode. The last step is to obtain the marginal posterior distributions in (17) and (18) by combining the two approximations in the previous steps and integrating out the irrelevant factors.

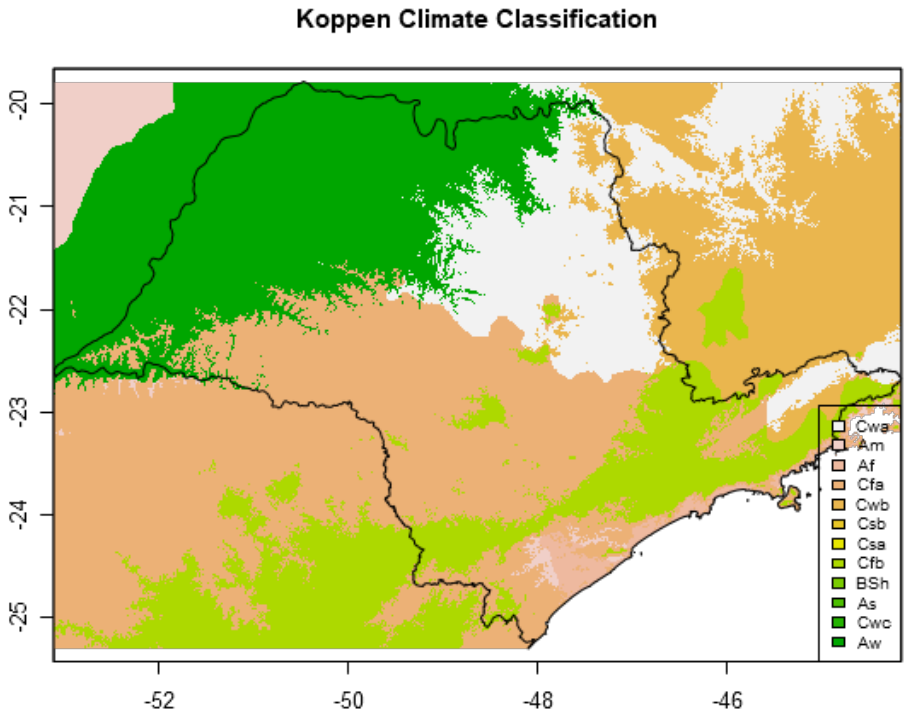
3 Data

In this paper the main goal is to evaluate the effects of Sugarcane Burning Law in fire patterns in the state of São Paulo. To achieve it, we use daily data from MODIS Thermal Anomalies/Fires between January 2003 and December 2016 with confidence detection higher than 50%. The dataset, provided by NASA[2], includes information like fire occurrence (day/night), fire location, the logical criteria for the fire selection and detection confidence. Moreover, in order to facilitate the visualization of the results, we use a quarterly aggregation of the daily data.

As explanatory variables we use Koppen Climate Classification system, following Alvares et al. (2013), and NASA MODIS Land Cover Classification, as a control for possible use of soil in agricultural activities and climate effects. Figures 1 and 2 show a complete classification based on Köppen classification and Land Cover Classification for the state of São Paulo, respectively. The classes of Köppen classification and MODIS land cover classification can be found in Appendix 5.

[2] Available in https://modis.gsfc.nasa.gov/data/dataprod/mod14.php.

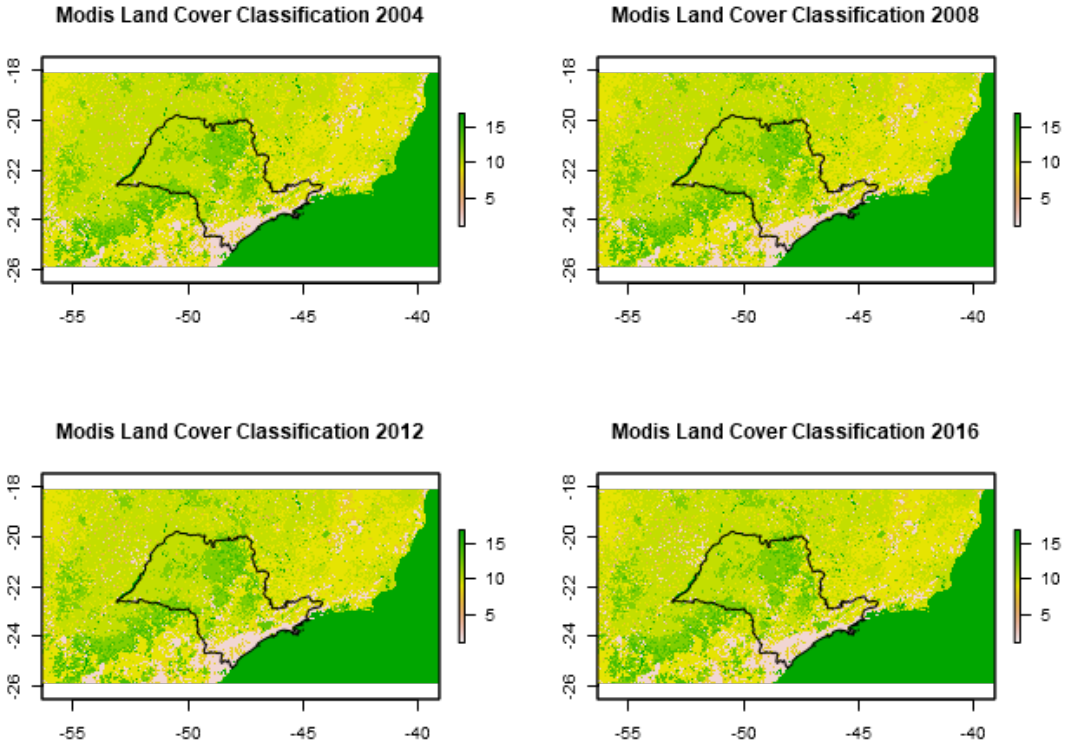
Figure 1: Köppen Climate Classification - São Paulo



Köppen Climate Classification - São Paulo

Cwa: (C) Humid subtropical (w) With dry winter (a) and hot summer
Am: (A) Tropical (m) monsoon
Af: (A) Tropical(f) without dry season
Cfa: (C) Humid subtropical (f) Oceanic climate, without dry season (a) and hot summer
Cwb: (C) Humid subtropical (w) With dry winter (b) and temperate summer
Csb: (C) Humid subtropical (s) With dry summer (b) and temperate summer
Csa: (C) Humid subtropical (s) With dry summer (a) and hot summer
Cfb: (C) Humid subtropical (f) Oceanic climate, without dry season (b) and temperate summer
BSh: (B) Dry (S) Semi-arid (h) low latitude and altitude
As: (A) Tropical (s) with dry summer
Cwc: (c) Humid subtropical (w) With dry winter and (c) short and cool summer
Aw: (A) Tropical (w) with dry winter

Figure 2: MODIS Land Classification - São Paulo



We also include spatially continuous projections of maximum temperature and precipitation for each period as explanatory variables. To construct these data we adopted the methodology proposed by Laurini (2019), based on Brazilian meteorological data provided by the National Institute of Meteorology (INMET) available at http://www.inmet.gov.br. Figures 3 and 4 present maximum temperature (Celsius degrees) and precipitation in Southeast region for the last quarter of 2016, respectively.

Figure 3: Maximum Temperature in Southeast region - 2016/4

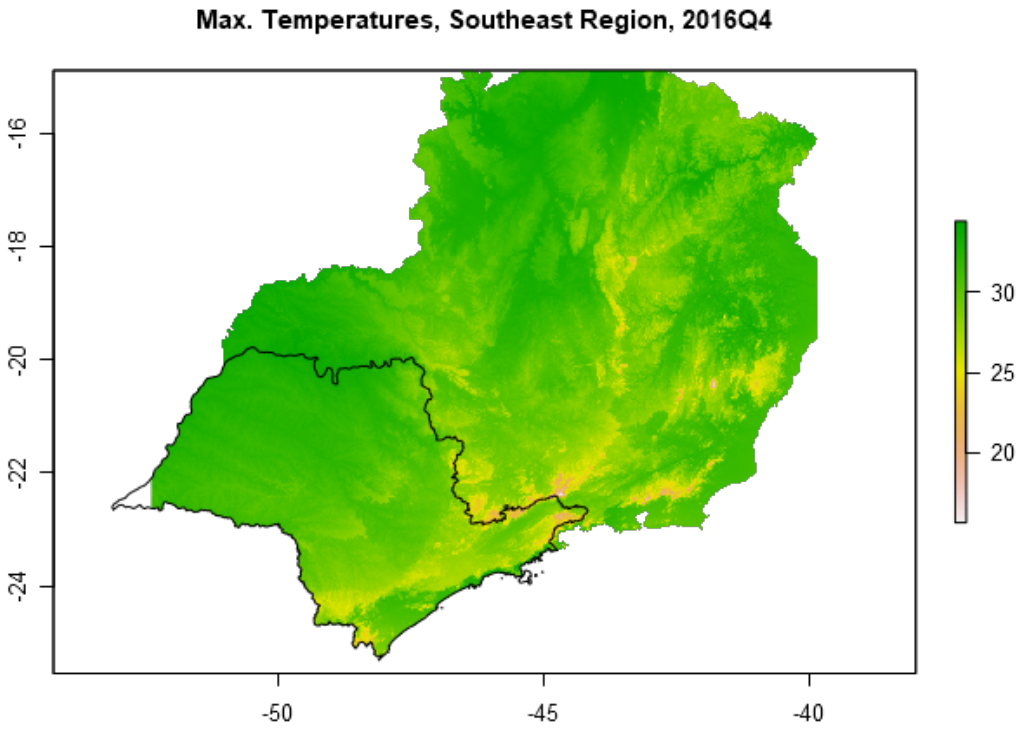
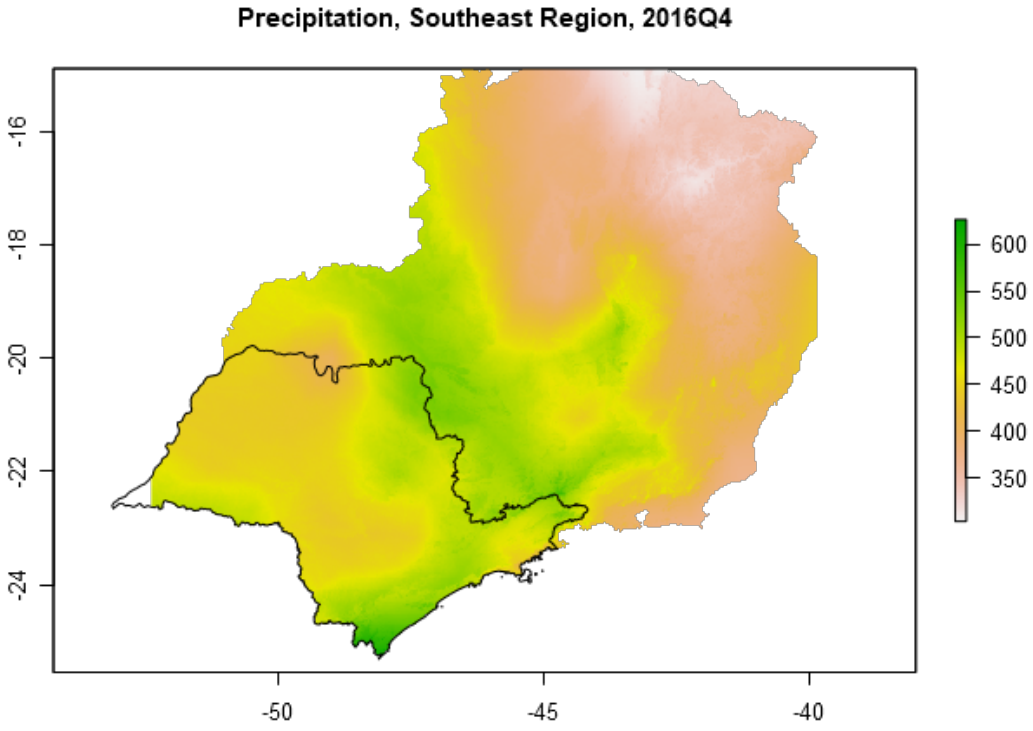
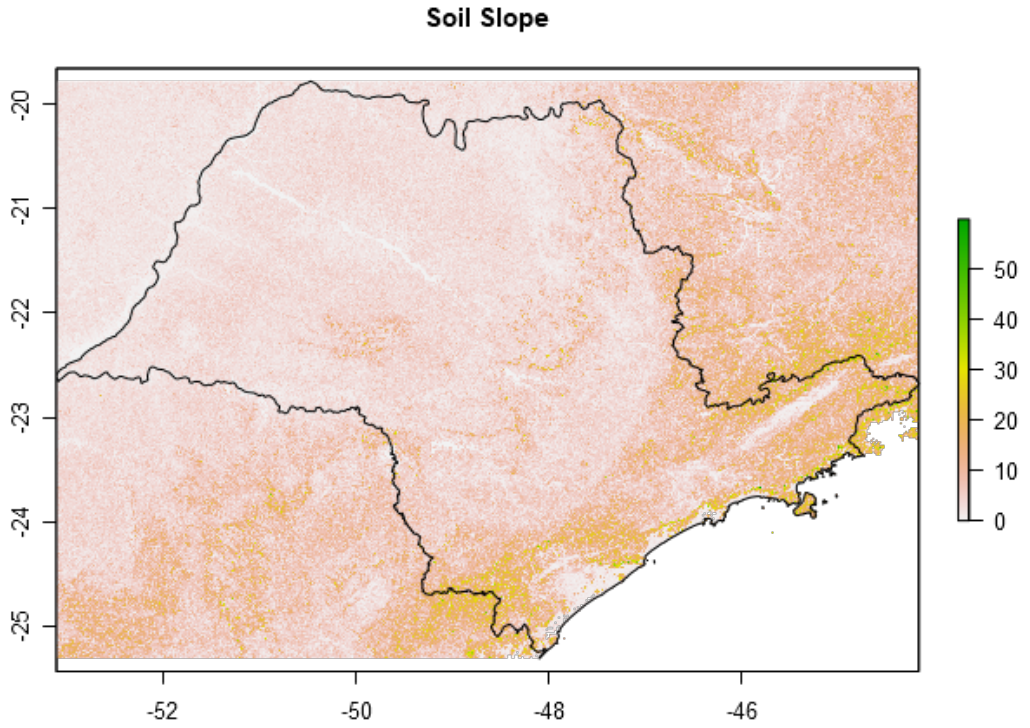


Figure 4: Precipitation in Southeast region - 2016/4



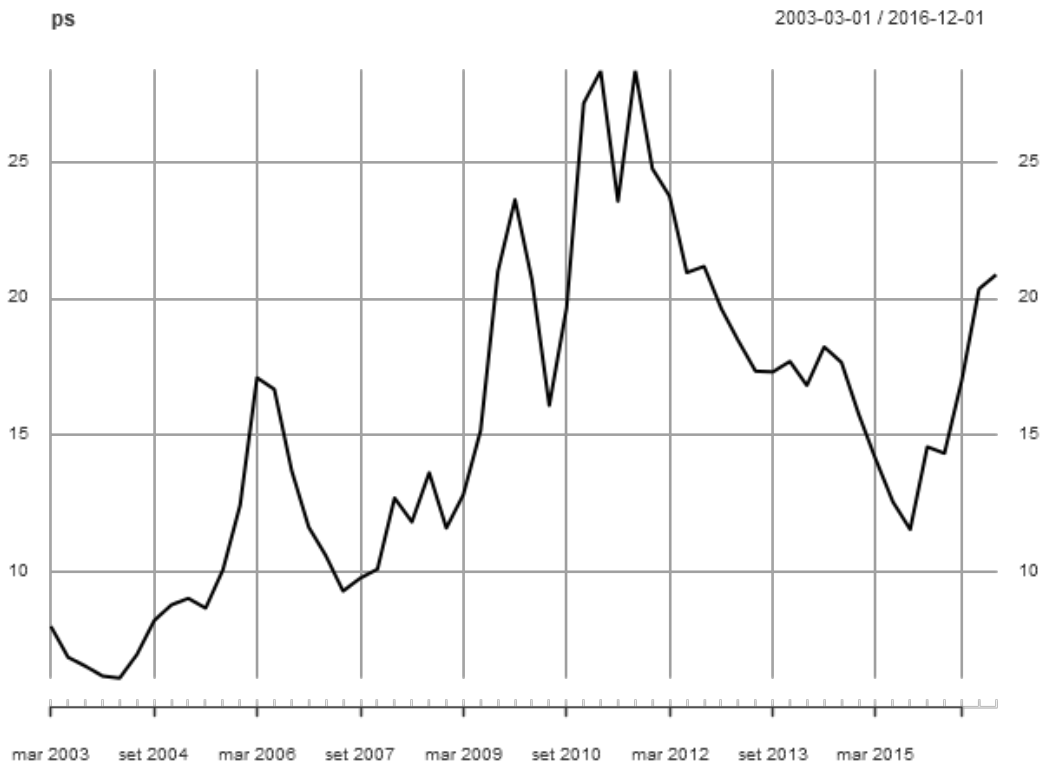
Due to slope restrictions related to mechanical harvesting, we include a soil slope variable, expressed in degrees and presented in Figure 5, from AMBDATA[3]. In our empirical analysis, the soil slope variable is set as a dummy, which is equal 0 in areas with slope less or equal 12% and 1 otherwise. In addition, we also include monthly global future price of sugar (U.S. cents per pound) provided by International Monetary Fund (IMF) and presented in Figure 6.

Figure 5: Soil Slope (degrees) in Southeast region



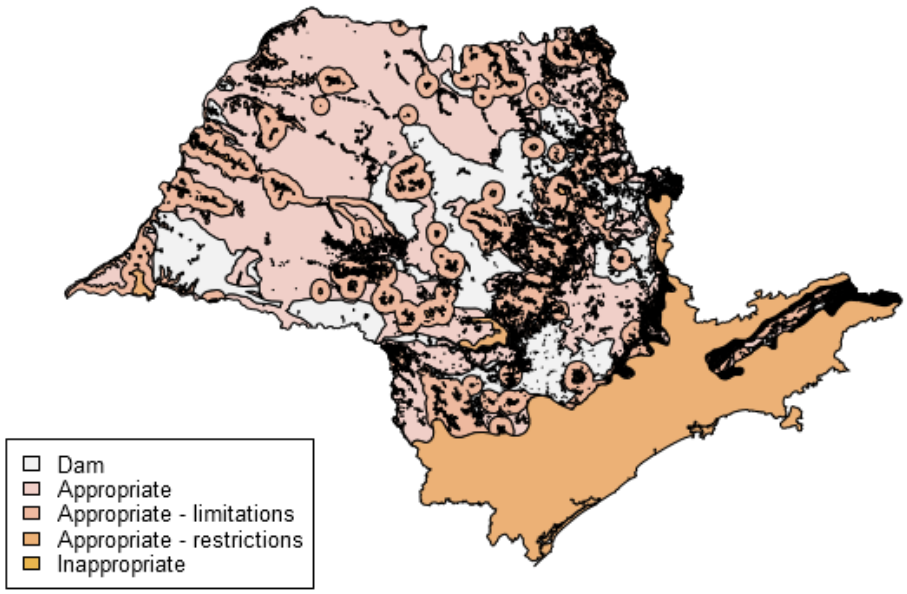
[3] Available in http://www.dpi.inpe.br/Ambdata/index.php

Figure 6: Global Price Sugar N°11 (U.S. Cents per Pound)



Finally, we include as explanatory variable the Sugarcane Agroecological Zoning for the production of ethanol and sugar, provided by the Secretariat for the Environment of the State of São Paulo jointly with the Secretariat of Agriculture of the State of São Paulo. According to Costa et al. (2018), the zoning study evaluate the potential of the land for the production of sugarcane based on some characteristics as the physical, chemical and mineralogical characteristics of the soil, the climate risk and environmental regulations. It was established four areas with different potential of the land for the production of sugarcane: appropriate areas, appropriate areas with environmental limitations, appropriate areas with environmental restrictions and inappropriate areas. Figure 7 present the Sugarcane Agroecological Zoning.

Figure 7: Sugarcane Agroecological Zoning



4 Results

In order to apply the inference procedures discussed in Section 2, the first step is to define a triangulation mesh of the interest region. Figure 8 presents the triangulated mesh adopted in this paper, with 117 triangles covering all the state of São Paulo. The second step is to define a set of knots over time to built a temporal mesh. For the time domain, we define a temporal mesh based on the number of observed quarters, 56. In order to obtain the space-time aggregation, we find to which polygon belongs each data point in the spatial mesh and to which part of the time belongs each data point. Hereafter, we use these both identification index sets to aggregate the data. Table 1 presents the descriptive statistics of the space-time volume unit at each polygon and time knot.

Figure 8: Triangulated Mesh

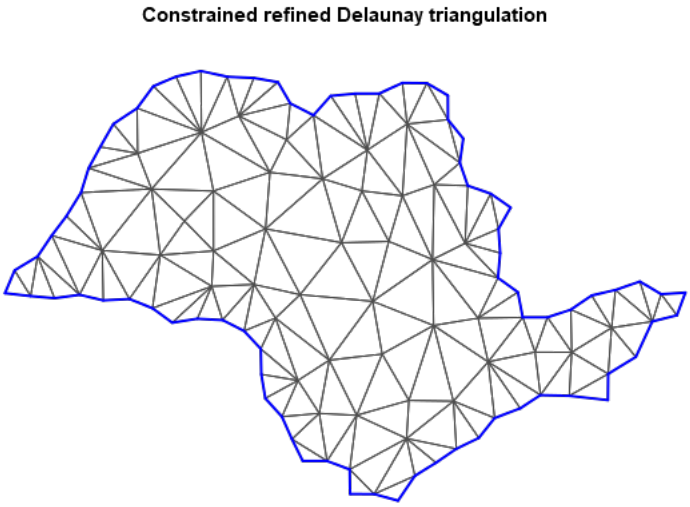


Table 1: Descriptive Statistics - Space-time volume

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Space-time volume	0.011	0.069	0.113	0.185	0.304	0.579

The proposed model follows Laurini (2019) approach, where the specification of the structural decomposition of time series is based on an additive process of trend, seasonality and cycle components. The trend series are modeled as first-order random walk, the seasonality process is modeled as a second-order autoregressive process with restriction that the effects most sum to zero and the cyclic component is represented by an autoregressive process with complex roots, which allows to capture the cyclic patterns. The spatial covariance is defined by a Matérn covariance function, given by the equation 2. As discussed in Section 2 and described in equation 1, the parameters to be estimated are the parameters associated to explanatory variables (β), the precision of the trend component ($1/\eta_\mu$), seasonality component ($1/\eta_s$) and cycle component ($1/\eta_c$). The parameters of the second-order autoregressive process of the cycle as parameterized as partial autocorrelatios (PACF1 and PACF2), whereas the parameters of spatial covariance are represented by $\log \tau$ and $\log \kappa$.

Table 2 presents the results of the estimation of the fixed effects. The results indicate a positive relation between maximum temperatures and the intensity of fires occurrences in São Paulo, and also a positive relation with soil slope, as expected, due to difficulties related to mechanical harvesting and in properties with high slope. In addition, although all regions classified as humid subtropical climate with dry winter and hot summer (Köppen 1), tropical climate without dry season (Köppen 3), humid subtropical oceanic climate, without dry season and hot summer (Köppen 4), humid subtropical climate with dry winter and temperate summer (Köppen 5), humid subtropical oceanic climate, without dry season and temperate summer (Köppen 8) and tropical climate with dry winter (Köppen 12) show a positive relation with fires intensity, the results indicate that under humid subtropical climate with dry winter and hot summer, the intensity of fires occurrences is higher than other regions. The negative relation between sugar price and the intensity of the fires occurrences is related to significant productivity

gains and income generation provided by mechanisation. The Sugarcane Agroecological Zoning presents a negative relation with fires intensity at inappropriate areas, as expected. In summary, the posterior distribution of the fixed effects is consistent with general fire patterns observed in the state of São Paulo.

Table 2: Estimated Parameters - Fixed Effects

	Mean	SD	0.025quant	0.5quant	0.975quant	Mode
Cwa (Köppen)	2.05	0.19	1.69	2.05	2.42	2.05
Af (Köppen)	0.27	0.18	-0.08	0.27	0.63	0.27
Cfa (Köppen)	1.20	0.17	0.86	1.20	1.54	1.20
Cwb (Köppen)	1.54	0.20	1.16	1.54	1.93	1.54
Cfb (Köppen)	0.85	0.18	0.50	0.85	1.19	0.85
Aw (Köppen)	1.53	0.19	1.17	1.53	1.90	1.53
Approp. zone with limitations	0.22	0.07	0.09	0.22	0.35	0.22
Approp. zone with restrictions	0.13	0.07	-0.01	0.13	0.28	0.13
Inappropriate zone	-0.09	0.07	-0.24	-0.09	0.05	-0.09
Land Cover Classification	0.16	0.08	-0.00	0.16	0.33	0.16
Sugar Price	-0.01	0.02	-0.05	-0.01	0.02	-0.01
Soil Slope (dummy)	0.27	0.08	0.11	0.27	0.43	0.27
Max. Temperature	0.02	0.00	0.02	0.02	0.03	0.02
Rainfall	-0.00	0.00	-0.00	-0.00	0.00	-0.00

Table 3 shows the estimated of the random effects. The results indicate a high precision for trend and seasonal components, whereas the precision for the components of cycle are relatively minor. In addition, a group autoregressive persistence equals to 0.70 suggests the high persistence of the spatial random effects.

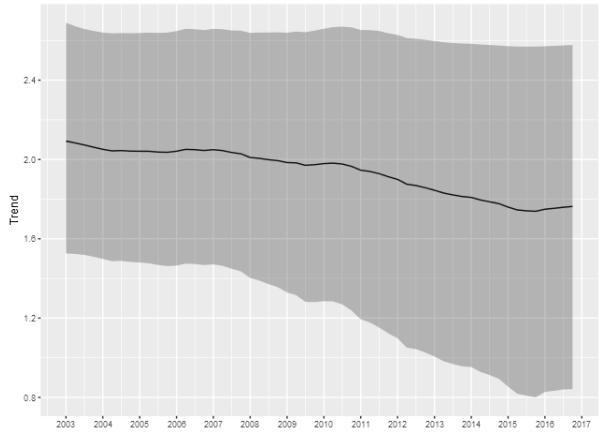
Table 3: Estimated Parameters - Random Effects

	Mean	SD	0.025quant	0.5quant	0.975quant	Mode
Precision for Trend	2379.24	8138.08	60.66	722.71	14788.14	141.72
Precision for Seasonality	26933.38	23395.89	3237.08	20538.46	88612.96	9204.33
Precision for Cycle	7.57	2.41	3.62	7.34	12.96	6.85
PACF1 for Cycle	0.34	0.16	0.01	0.35	0.65	0.35
PACF2 for Cycle	-0.15	0.16	-0.45	-0.16	0.18	-0.17
Log τ	-2.32	0.04	-2.40	-2.32	-2.23	-2.33
Log κ	1.33	0.03	1.26	1.33	1.39	1.33
Group Θ	0.70	0.01	0.67	0.70	0.73	0.71

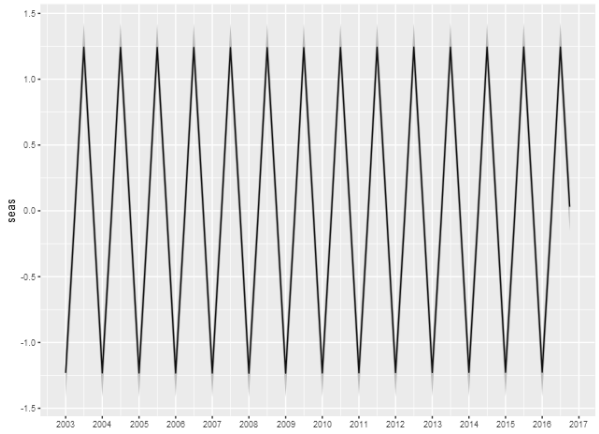
The estimated trend, seasonality and cycle component are shown in Figure 9, which presents the posterior mean of the estimated components with 95% Bayesian credibility interval.

Figure 9: Trend, Seasonal and Cycle decomposition of fire occurrences in São Paulo

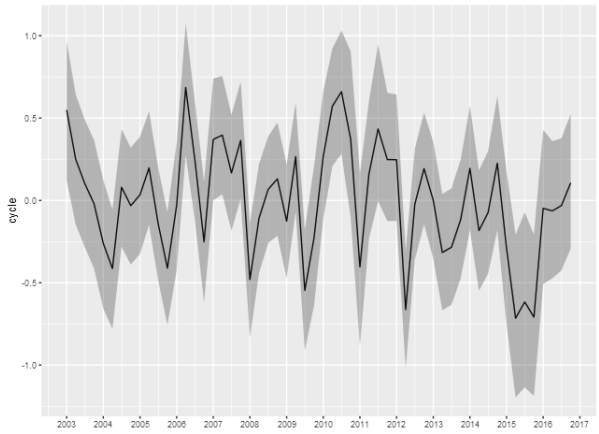
(a) Trend



(b) Seasonal



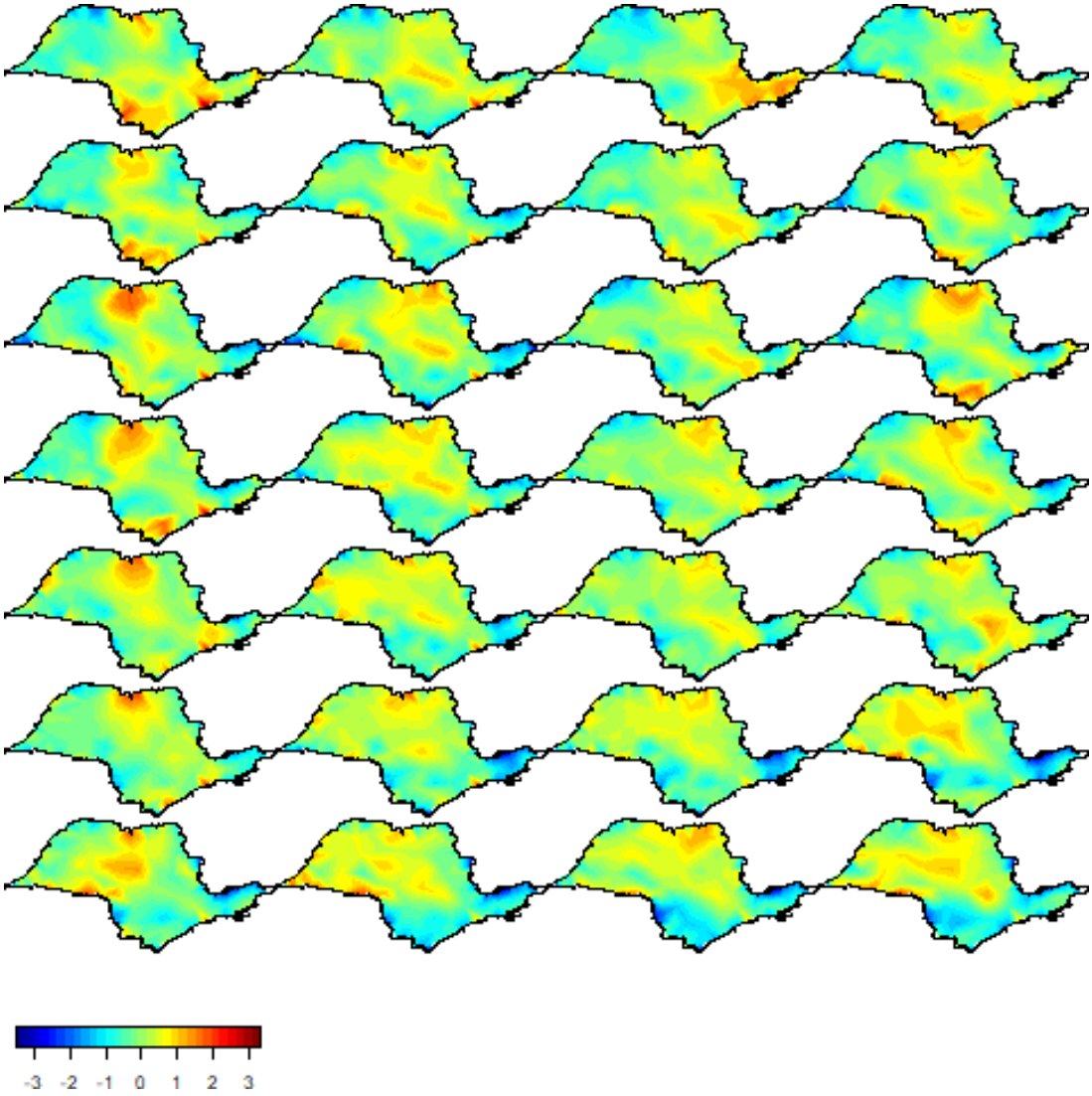
(c) Cycle



The trend component reflects the reduction in fire occurrences related to the sugarcane burning law (Law 11.241/2002), whereas the seasonal component exhibits a very stable behavior over time, as expected. The cycle component corresponds to an AR(2) model with complex roots: 0.392 and -0.150, which is consistent with a cyclic pattern. Plus, this result corresponds to a cycle period of 6.04 quarters.

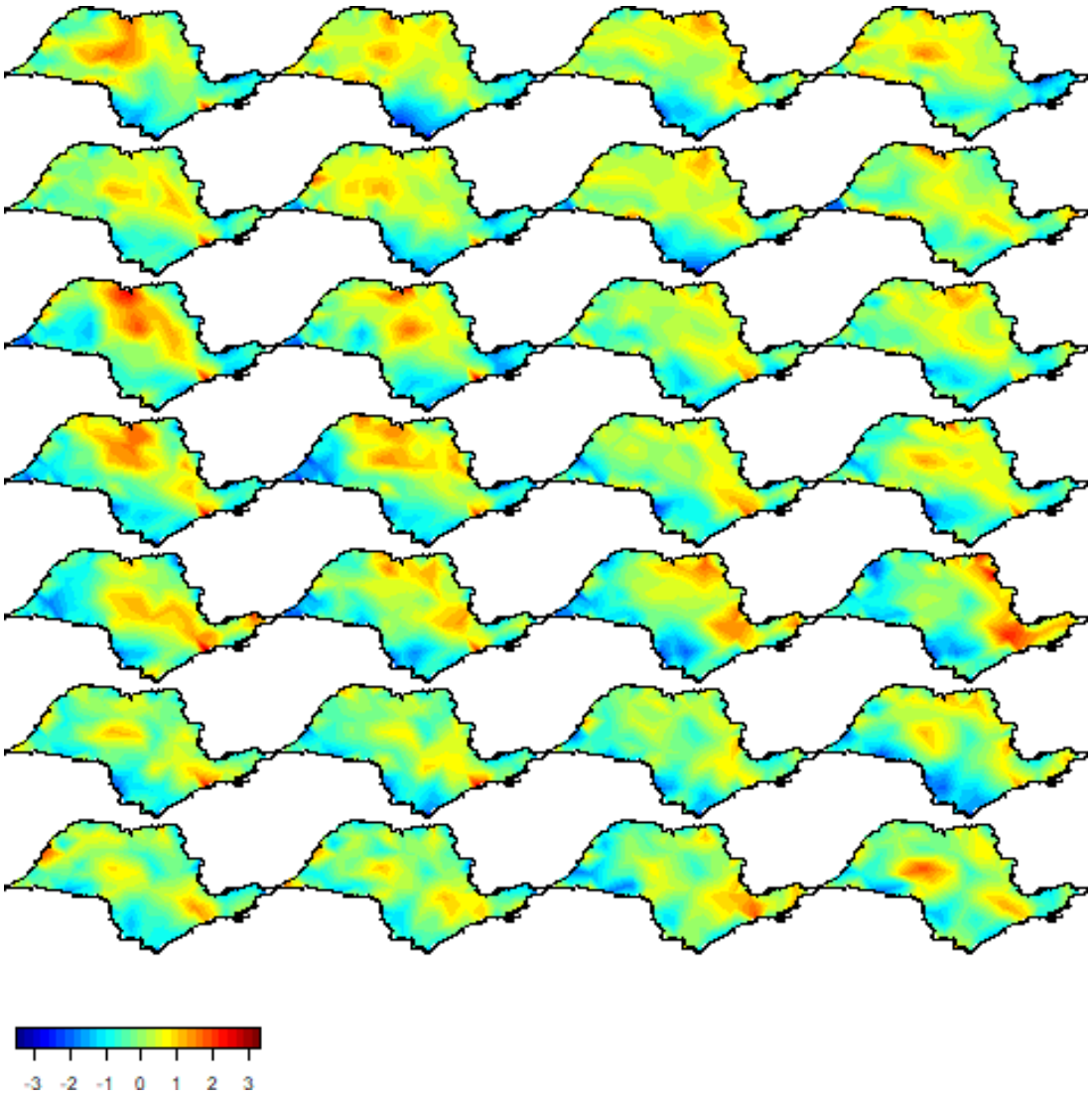
Figures 10 and 11 report the spatial random effects estimated, for 2003-2009 and 2010-2016, respectively. The columns indicate the quarter whereas rows indicate the years. It is possible to observe significant variation in the dynamic evolution of spatial random effects. The intensity of the fires fitted by the model and the observed fires are shown in Figures 12 and 13, for 2003-2009 and 2010-2016 periods, respectively. This adjustment was constructed by adding the estimated posterior mean of the common components of trend, seasonality, cycle and estimated spatial random effects. Again, in the figures, the columns indicate the quarter whereas rows indicate the years, and black dots represent the observed fires in each period. The results observed in these figures suggest that the model has a good fit.

Figure 10: Spatial Random Effects 2003 - 2009



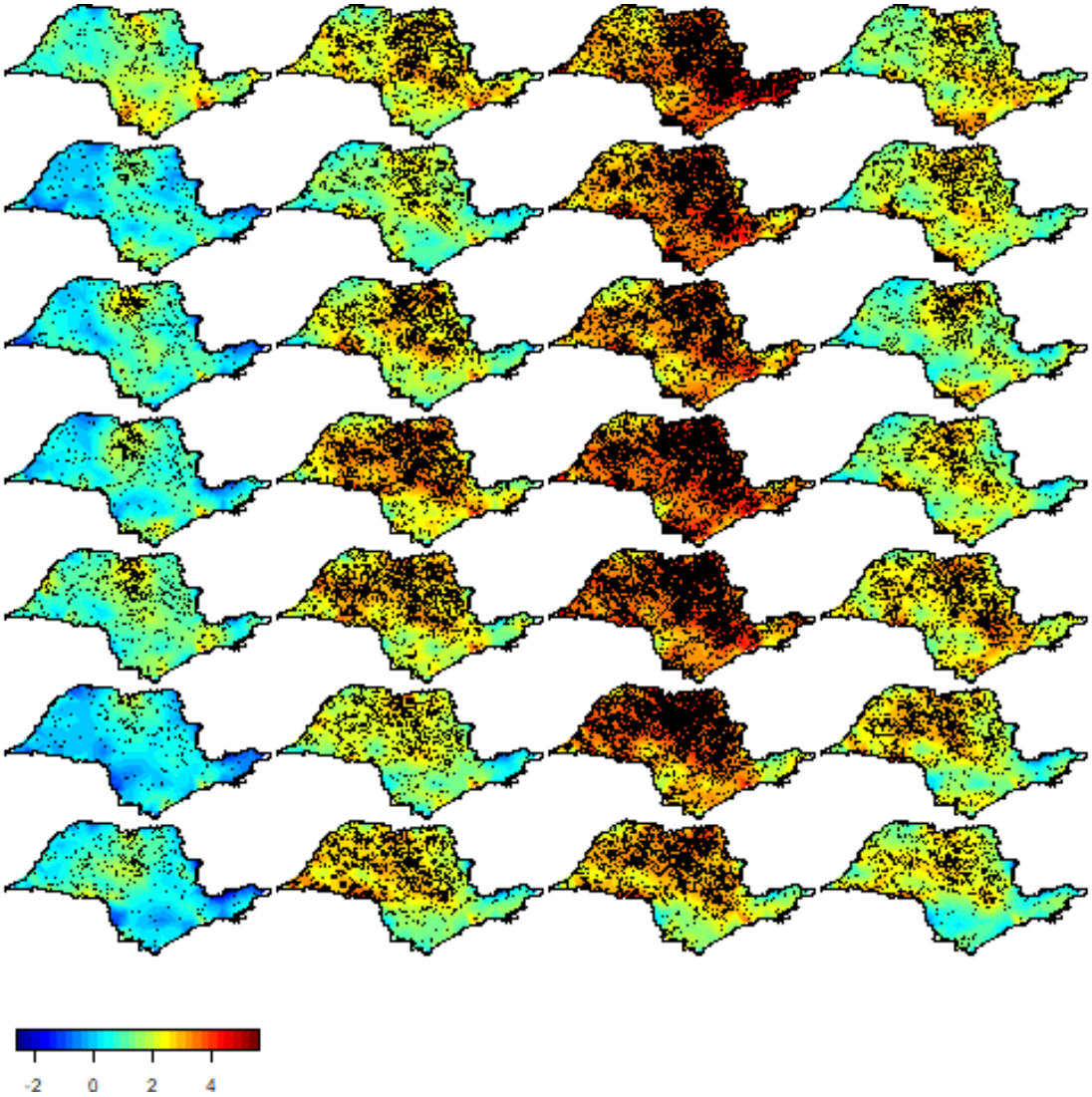
Note: Columns indicate quarter whereas rows indicate the year, from 2003 to 2009.

Figure 11: Spatial Random Effects 2010 - 2016



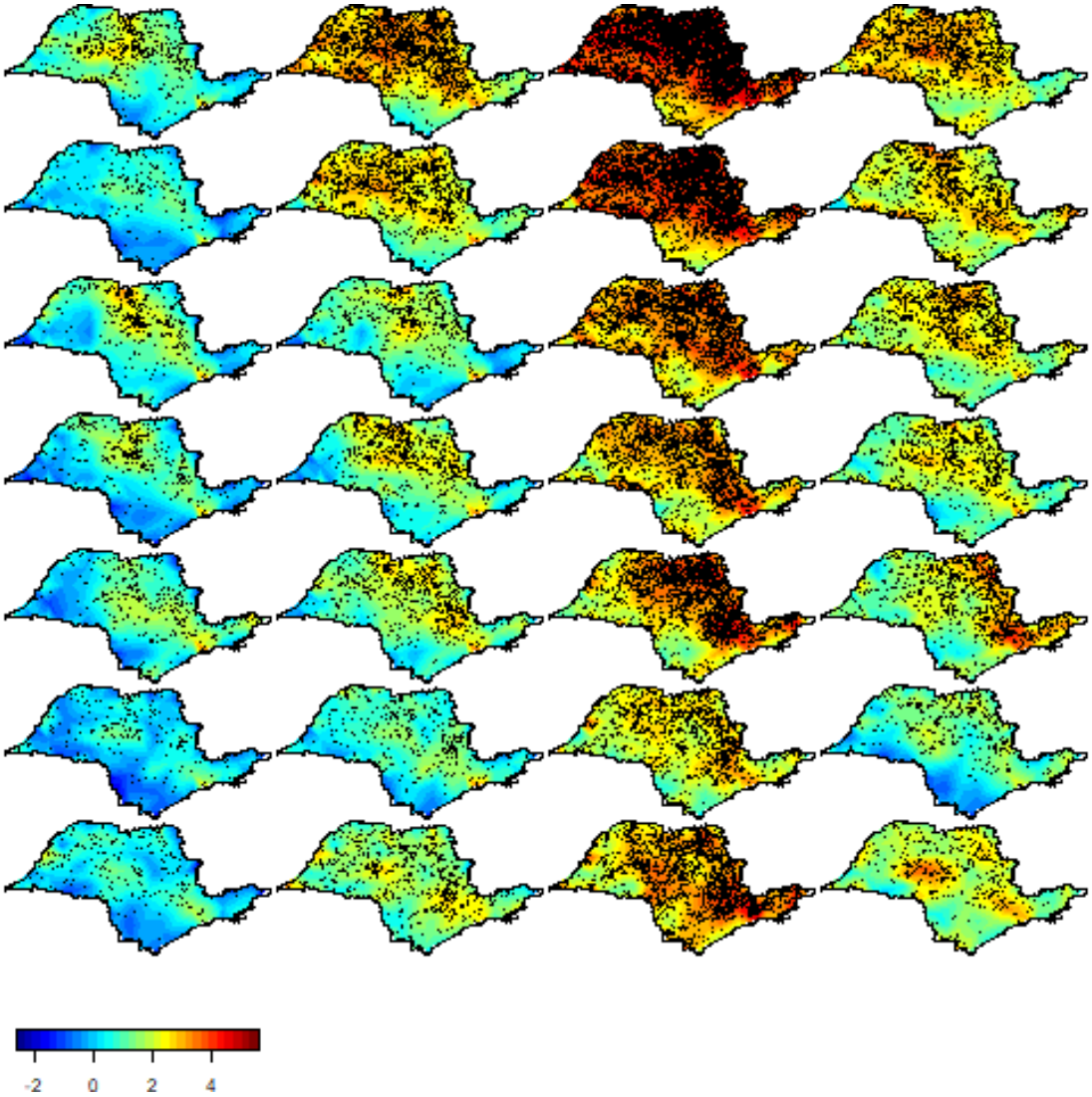
Note: Columns indicate quarter whereas rows indicate the year, from 2010 to 2016.

Figure 12: Fitted Intensity and Fires 2003 - 2009



Note: Columns indicate quarter whereas rows indicate the year, from 2010 to 2016. Black dots indicate observed fires in each period.

Figure 13: Fitted Intensity and Fires 2010 - 2016



Note: Columns indicate quarter whereas rows indicate the year, from 2010 to 2016. Black dots indicate observed fires in each period.

5 Conclusions

In this work we propose to analyze the process of change in the pattern of fires in the state of São Paulo, analyzing the efficiency of regulatory changes in the sugarcane production sector. To perform representation and inference procedures for the spatio-temporal point process we adopt a dynamic representation of Log-Gaussian Cox process. This representation is based on the modeling of the process intensity function through a decomposition of components of trend, seasonality, cycles, covariates and spatial effects, which allows spatial effects to change over time using a autoregressive functional structure. We adopt the SPDE approach proposed by Lindgren, Rue and Lindström (2011) to approximate the initial Gaussian Field to a Gaussian Markov Random Field, which is defined by sparse matrices and fits naturally within Bayesian hierarchical modelling framework, and allows to use Integrated Nested Laplace Approximations approach to inference procedures. We use daily data from MODIS Thermal Anomalies/fires between January 2003 and December 2016. The results provide evidence of the decrease in the trend component of the observed intensity in fires occurrences in the state of São Paulo related to Sugarcane Burning Law (Law 11.241/2002).

Bibliography

AGUIAR, D. A. et al. Remote sensing images in support of environmental protocol: Monitoring the sugarcane harvest in são paulo state, brazil. *Remote Sensing*, Molecular Diversity Preservation International, v. 3, n. 12, p. 2682–2703, 2011.

ALVARES, C. A. et al. Köppen's climate classification map for brazil. *Meteorologische Zeitschrift*, E. Schweizerbart'sche Verlagsbuchhandlung, v. 22, n. 6, p. 711–728, 2013.

CANÇADO, J. E. et al. The impact of sugar cane–burning emissions on the respiratory system of children and the elderly. *Environmental health perspectives*, National Institute of Environmental Health Sciences, v. 114, n. 5, p. 725–729, 2006.

CAPAZ, R. S.; CARVALHO, V. S. B.; NOGUEIRA, L. A. H. Impact of mechanization and previous burning reduction on ghg emissions of sugarcane harvesting operations in brazil. *Applied Energy*, Elsevier, v. 102, p. 220–228, 2013.

COSTA, J. A. F. et al. *Energy Law and Regulation in Brazil*. [S.l.]: Springer, 2018.

GOLDEMBERG, J.; COELHO, S. T.; GUARDABASSI, P. The sustainability of ethanol production from sugarcane. *Energy policy*, Elsevier, v. 36, n. 6, p. 2086–2097, 2008.

ILLIAN, J. et al. Fitting a log gaussian cox process with temporally varying effects–a case study. *Preprint Statistics*, Citeseer, n. 17, 2010.

JUNQUEIRA, C. P.; STERCHILE, S. P. W.; SHIKIDA, P. F. de A. Mudança institucional e o impacto no padrão tecnológico: o caso da mecanização da colheita de cana-de-açúcar no paraná. *Organizações Rurais & Agroindustriais*, v. 11, n. 1, 2011.

KRAINSKI, E. T. et al. *Advanced spatial modeling with stochastic partial differential equations using R and INLA*. [S.l.]: Chapman and Hall/CRC, 2018.

LAURINI, M. A spatio-temporal approach to estimate patterns of climate change. *Environmetrics*, Wiley Online Library, v. 30, n. 1, p. e2542, 2019.

LINDGREN, F.; RUE, H.; LINDSTRÖM, J. An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 73, n. 4, p. 423–498, 2011.

MØLLER, J.; SYVERSVEEN, A. R.; WAAGEPETERSEN, R. P. Log gaussian cox processes. *Scandinavian journal of statistics*, Wiley Online Library, v. 25, n. 3, p. 451–482, 1998.

PARAISO, M. L. d. S.; GOUVEIA, N. Health risks due to pre-harvesting sugarcane burning in são paulo state, brazil. *Revista Brasileira de Epidemiologia*, SciELO Public Health, v. 18, p. 691–701, 2015.

PEREIRA, P. et al. Quantification of annual wildfire risk; a spatio-temporal point process approach. *Statistica*, v. 73, n. 1, p. 55–68, 2013.

ROSTAMI, M. et al. Modeling spatio-temporal variations of substance abuse mortality in iran using a log-gaussian cox point process. *Spatial and spatio-temporal epidemiology*, Elsevier, v. 22, p. 15–25, 2017.

RUE, H.; MARTINO, S.; CHOPIN, N. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, Wiley Online Library, v. 71, n. 2, p. 319–392, 2009.

SERRA, L. et al. Spatio-temporal log-gaussian cox processes for modelling wildfire occurrence: the case of catalonia, 1994–2008. *Environmental and ecological statistics*, Springer, v. 21, n. 3, p. 531–563, 2014.

SIMPSON, D. et al. Going off grid: Computationally efficient inference for log-gaussian cox processes. *Biometrika*, Oxford University Press, v. 103, n. 1, p. 49–70, 2016.

WHITTLE, P. On stationary processes in the plane. *Biometrika*, JSTOR, p. 434–449, 1954.

1 Köppen Climate Classification and MODIS Land Cover Classification

1.1 Köppen Climate Classification

1. Cwa: (C) Humid subtropical (w) With dry winter (a) and hot summer
2. Am: (A) Tropical (m) monsoon
3. Af: (A) Tropical (f) without dry season
4. Cfa: (C) Humid subtropical (f) Oceanic climate, without dry season (a) and hot summer
5. Cwb: (C) Humid subtropical (w) With dry winter (b) and temperate summer
6. Csb: (C) Humid subtropical (s) With dry summer (b) and temperate summer
7. Csa: (C) Humid subtropical (s) With dry summer (a) and hot summer
8. Cfb: (C) Humid subtropical (f) Oceanic climate, without dry season (b) and temperate summer
9. BSh: (B) Dry (S) Semi-arid (h) low latitude and altitude
10. As: (A) Tropical (s) with dry summer
11. Cwc: (C) Humid subtropical (w) With dry winter (c) short and cool summer
12. Aw: (A) Tropical (w) with dry winter

1.2 MODIS Land Cover Classification

1. Evergreen Needleleaf Forests: dominated by evergreen conifer trees (canopy >2m). Tree cover >60%.
2. Evergreen Broadleaf Forests: dominated by evergreen broadleaf and palmate trees (canopy >2m). Tree cover >60%.
3. Deciduous Needleleaf Forests: dominated by deciduous needleleaf (larch) trees (canopy >2m). Tree cover >60%.
4. Deciduous Broadleaf Forests: dominated by deciduous broadleaf trees (canopy >2m). Tree cover >60%.

5. Mixed Forests: dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy >2m).Tree cover >606 Closed Shrublands: dominated by woody perennials (1-2m height) >60% cover.

7. Open Shrublands: dominated by woody perennials (1-2m height) 10-60% cover.

8. Woody Savannas: tree cover 30-60% (canopy >2m). 9 fb13 Savannas: tree cover 10-30% (canopy >2m).

9. Grasslands: dominated by herbaceous annuals (<2m).

10. Permanent Wetlands: permanently inundated lands with 30-60% water cover and >10% vegetated cover.

11. Croplands: at least 60% of area is cultivated cropland.

12. Urban and Built-up Lands: at least 30% impervious surface area including building materials, asphalt and vehicles.

13. Cropland/Natural Vegetation Mosaics: mosaics of small-scale cultivation 40-60% with natural tree, shrub,or herbaceous vegetation.

14. Permanent Snow and Ice: at least 60% of area is covered by snow and ice for at least 10 months of the year.

15. Barren: at least 60% of area is non-vegetated barren (sand, rock, soil) areas with less than 10% vegetation.

16. Water Bodies: at least 60% of area is covered by permanent water bodies.