A Bayesian Modelling of Wildfires in Portugal

Giovani L. Silva, Paulo Soares, Susete Marques, M. Inês Dias, M. Manuela Oliveira, and José G. Borges

Abstract In the last decade wildfires became a serious problem in Portugal due to socieconomic and climate change trends. In order to analyse wildfire data, we employ beta regression for modelling the proportion of burned wild area, under a Bayesian perspective. Our main goal is to find out fire risk factors that influence the proportion of area burned and what may make a wild area susceptible or resistant to fire. Then, we analyse wildfire data in Portugal during 1990–1994 through Bayesian normal and beta regression models that use Markov chain Monte Carlo methods for estimating quantities of interest.

1 Introduction

In Portugal, wildfires (related to natural forests and other plant areas) have been increasing in the last years. Fire is indeed an important issue in Mediterranean region affecting namely the ecological and economic aspects of forest areas and causing loss of human life. Many factors have contributed to the increasing number of wildfires, e.g., climate change [7]. Some studies have identified changes in the number of fires, burned area and fire size distribution depending on topographical variables and vegetation type, e.g., in the Spanish region Catalonia [10] and Portugal [13].

Gomes [9] pointed out many causes and consequences of forest fires in Portugal, e.g., currently, rural and forest areas in Portugal are considerably deserted due to

G.L. Silva (✉) • P. Soares
CEAUL and Department Mathematics, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1, 1049-001 Lisbon, Portugal
e-mail: giovani.silva@tecnico.ulisboa.pt; paulo.soares@tecnico.ulisboa.pt

M.I. Dias • M.M. Oliveira
CIMA and Department Mathematics, Universidade de Évora, Évora, Portugal
e-mail: misd@uevora.pt; mmo@uevora.pt

S. Marques • J.G. Borges
FRC and Department Natural Resources, Environment and Territory, Instituto Superior de Agronomia, Universidade de Lisboa, Lisboa, Portugal
e-mail: smarques@isa.ulisboa.pt; joseborges@isa.ulisboa.pt

population migrations from these areas to the main cities, which began in the 1950s. Fernandes et al. [5] proposed a fuel modelling and fire hazard assessment, used to evaluate and compare the fire hazard potential between forest types defined by their composition and structure. They found that potential fire behaviour is primarily driven by stand structure, rather than by cover type.

Marques et al. [13] presented an approach of the characterisation fire occurrence in Portugal, combining the use of geographic information systems (GIS) and generalised linear models (GLM). They emphasised the relationship between ecological and socioeconomic features on the proportion of area burned, recording also the number of fires and fire size for three 5-year periods, including the period 1990–1994. Descriptive statistics indicated variations in the distribution of fires over recent decades, with a significant increase in number of very large fires. Regression models underlined the impact of the forest cover type and the proximity to roads on the proportion of area burned.

For modelling wildfires, GLM [1, 14] have been often adopted, even as that is based on the Gaussian distribution by transforming the response [13]. Ferrari and Cribari-Neto [6] proposed a regression model where the response is beta distributed using a parameterisation of the beta law that is indexed by mean and dispersion parameters. Beta regression can be used for modelling the proportion of area burned that is restricted to the interval (0, 1). The regression parameters of the beta regression model are interpretable in terms of the mean of the response and, when the *logit* link is used, of an odds ratio, unlike the parameters of a linear regression that employs a transformed response [6].

This work proposes to model the proportion of burned area due to wildfires in Portugal, based on beta regression and under a Bayesian perspective (see e.g. [8, 17] for some Bayesian GLMs). The rest of the article is organised as follows. Section 2 succinctly describes the motivation of this work and the different modelling of wildfires. In Sect. 3 we present Bayesian beta regression for modelling the proportion of area burned, taking the use of Markov chain Monte Carlo (MCMC) methods for estimating quantities of interest. Some results of Bayesian beta regression related to the wildfire data analysis in the entire Portuguese mainland between 1990 and 1994, and concluding remarks are done respectively in Sects. 4 and 5, including the identification of the fire risk factors.

2 Motivation and Methods

In Portugal, burned area mapping, obtained by semi-automated classification of high-resolution remote sensing data from Instituto Superior de Agronomia (ISA)—Universidade de Lisboa, identified 35,198 fire perimeters with burned areas equal to or greater than 5 ha in the period 1975–2007 and the corresponding area burned is about 3.8×10^6 ha that is equivalent to nearly 40 % of the country area [13].
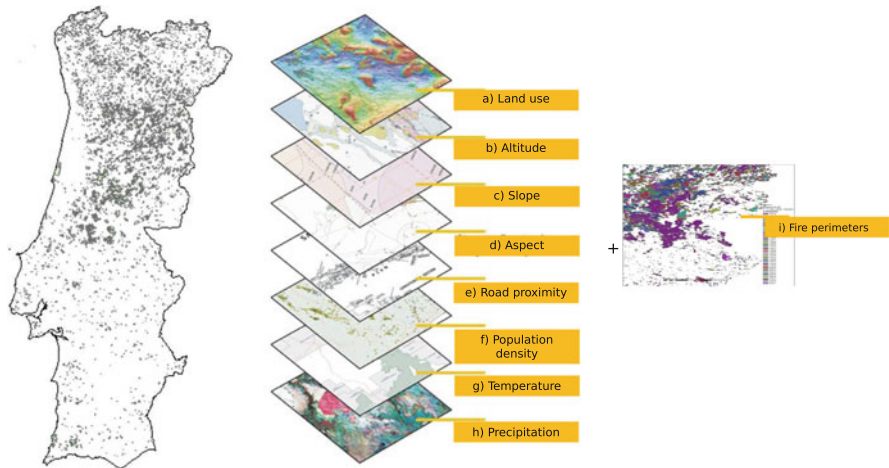
Fig. 1 Fire perimeters between 1990 and 1994 in Portugal (*left*), a zoom over a burned area is shown in the *right*, based on the classes of the covariates: (**a**) land use, (**b**) altitude, (**c**) slope, (**d**) slope orientation (aspect), (**e**) road proximity, (**f**) population density, (**g**) temperature, (**h**) precipitation, (**i**) layer indicating the fire perimeters

In the period 1990–1994, Marques et al. [13] pointed out that: (1) 5706 Portuguese wildfires were recorded and the total burned area extended over 442,745 ha, burning about 4.97 % of the country area, (2) the average area burned per wildfire was 77 ha, (3) 149 wildfires extended over 500 ha, accounting for 44 % of the burned area, (4) none extended over 10,000 ha. Figure 1 (left side) exemplifies the distribution (frequency) of these fires identifying high and critical fire zones that are specially located in the northern and central interior of Portugal.

In order to analyse variations in Portuguese wildfires in 1990–1994, the areas burned were included as map layers in the GIS database according to eight fire features (covariates), which were initially categorised, based on extensive preliminary data analysis and referred in the paper [13], into several classes: altitude (m), slope (%), slope orientation, population (hab/km^2), roads proximity (m), number of days with precipitation greater than 1 mm in the fire season (from May to October), number of days with maximum temperature higher than 25 °C in the fire season, and land cover (Table 1), including the observed proportion of the each land use classes in parentheses. Theses classes were also chosen based on some studies using the same data such as Moreira et al. [15] and Pereira et al. [19].

Figure 1 illustrates the fire perimeters used for constructing burned area data from related map layers. Notice that land cover map used in this study further included a map at the scale 1/25,000 (*Carta de Ocupação do Solo—COS'90*) produced by *Instituto Geográfico Português* using cartographic information from aerial photography mostly dated from 1990 [13], as well as that road proximity

Table 1 Description of the classes of the eight fire features used in the wildfire data

Roads proximity (m)	Population (hab/m^2)	Slope (%)	Altitude (m)	Slope orientation	Precipitation (number of days ≥ 1 mm)[a]	Temperature (number of days $>25\,°C$)[a]
≥ 1000	<25	0–10	< 200	Flat	0–6	0–3
<1000	25–100	10–20	200–400	North	7–13	4–48
	>100	20–30	400–700	East	14–18	49–71
		>30	>700	South	19–22	72–92
				West	23–26	93–112
					≥ 27	≥ 113

Land cover: annual crop (5.3 %); eucalyptus (10.9 %); hardwoods (7 %); hardwoods and softwoods mixed with eucalyptus (HSME) (8.9 %); agro-forestry (5.8 %); permanent crop (3.5 %); shrubs (27.4 %); resinous or softwoods (RS) (18.8 %); softwoods mixed with eucalyptus (SME) (8.6 %); others (e.g. social areas) (3.8 %)
[a]Number of days in the fire season (from May to October)

included trails and was defined (1000 m distance) based on previous work e.g. Catry et al. [3]. Although continuous covariates as temperature and precipitation could be better explored in their natural form, we chose to categorize them because of a matter of simplicity and interpretation for the data collection and the model parameters, respectively.

For the modelling of wildfires, we record the observed proportion of burned area, denoted by r_i that is the burned area out of total area for the ith combination of levels for the covariates in study, $i = 1, \ldots, k$. We propose to model the proportion of burned area from these eight underlying covariates by assuming beta distribution for r_i, i.e.,

- Beta model: $r_i \sim Beta(\mu_i \phi, (1-\mu_i)\phi)$, with mean $E(r_i) = \mu_i$ and variance $Var(r_i) = \frac{\mu_i(1-\mu_i)}{\phi+1}$.

Alternative GLM can model the proportion r_i, for instance:

- Gaussian model: $logit(r_i) \equiv \log(r_i/(1-r_i)) \sim Normal(\mu_i, \sigma^2)$;
- Gamma model: $-\log(r_i) \sim Gamma(\nu, \nu/\mu_i)$, with $E(r_i) = \mu_i$ and $Var(r_i) = \frac{\mu_i^2}{\nu}$.

These two models and other GLM based on transformations of r_i, such as $arcsin(\sqrt{r_i})$ and Box-Cox transformation, are discussed and developed in [1, 14]. Figure 2 displays histograms of the observed proportion of area burned without and with *logit* transformation in Portugal during the period 1990–1994, indicating that transforming response may not be the best way of wildfire modelling, what happens in the proportions close to one in Fig. 2 and notice that the beta model does not transform the response.
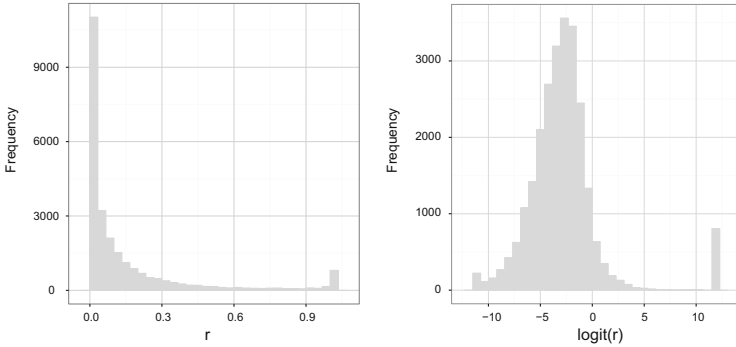
Fig. 2 Histograms of the observed proportion of area burned without (*right*) and with (*left*) *logit* transformation in Portugal during the period 1990–1994

3 Bayesian Beta Regression

Let $r_1, \ldots, r_k$ be random variables, where r_i follows a beta distribution with mean μ_i and unknown precision ϕ, whose probability density function is

$$f(r_i|\mu_i, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu_i\phi)\Gamma((1-\mu_i)\phi)} r_i^{\mu_i\phi-1}(1-r_i)^{(1-\mu_i)\phi-1}, \quad 0 < r_i < 1, \qquad (1)$$

where $\Gamma(\cdot)$ is the gamma function, $0 < \mu_i < 1$ and $\phi > 0$, $i = 1, \ldots, k$. Notice that the parameterisation of the beta distribution (1) was suggested by Ferrari and Cribari-Neto [6] in order to model response variable that is continuous and restricted to the interval (0, 1) and is related to other variables through a regression structure.

The beta regression model is obtained from Eq. (1) by assuming that the mean μ_i can be written as

$$g(\mu_i) = \mathbf{z}_i^T \boldsymbol{\beta} \equiv \eta_i, \qquad (2)$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ is the regression parameter vector associated with the covariate vector $\mathbf{z}_i = (z_{i1}, \ldots, z_{ip})^T$ for the ith observation, $i = 1, \ldots, k$, and $g(\cdot)$ is a *logit* link function $g(\mu) = \log[\mu/(1 - \mu)]$ (for other link functions, see [1, 6, 14]).

3.1 Posterior Distribution

For the likelihood, we can assume different sampling distributions for the proportion r_i, e.g., beta distribution defined in Eq. (1) or normal distribution for the transformed proportion, as referred in Sect. 2. Based on the former distribution with *logit* link

function in Eq. (2), the likelihood function is given by

$$L(\boldsymbol{\beta}, \phi | \mathbf{x}) = \prod_{i=1}^{k} \frac{\Gamma(\phi)}{\Gamma(\mu_i \phi) \, \Gamma((1-\mu_i)\phi)} r_i^{\mu_i \phi - 1} (1-r_i)^{(1-\mu_i)\phi - 1}, \tag{3}$$

where $\mathbf{x} = \{r_i; \mathbf{z}_i, i = 1, \ldots, k\}$ is the data, and $\mu_i = e^{\mathbf{z}_i^T \boldsymbol{\beta}} / (1 + e^{\mathbf{z}_i^T \boldsymbol{\beta}})$, $i = 1, \ldots, k$.

In Bayesian analysis, we also consider information a priori that here consists of assuming independent normal distributions with zero mean and variances v_j^2 for the regression coefficients, $j = 1, \ldots, p$, and inverse gamma distribution with shape a and scale b parameters for the precision parameter ϕ (or gamma distribution with shape a and scale b parameters for the variance σ^2 related to normal regression). In fact, we assigned non-informative prior distribution, i.e., highly dispersed, but proper normal and inverse gamma prior distributions for the model parameters $\boldsymbol{\beta}$ and ϕ (or $1/\sigma^2$), respectively. In that case, one expects that inferential results on the model parameters are not too different from those ones under a frequentist approach.

Assuming a priori independence amongst the model parameters, we can construct the joint posterior density related to the beta regression model (2), which is denoted by

$$\pi(\boldsymbol{\beta}, \phi | \mathbf{x}) \equiv \frac{L(\boldsymbol{\beta}, \phi | \mathbf{x}) \pi_1(\boldsymbol{\beta}) \pi_2(\phi)}{\int \int L(\boldsymbol{\beta}, \phi | \mathbf{x}) \pi_1(\boldsymbol{\beta}) \pi_2(\phi) \, d\boldsymbol{\beta} d\phi}, \tag{4}$$

where $\pi_1(\boldsymbol{\beta})$ and $\pi_2(\phi)$ are the normal and inverse gamma prior distributions of $\boldsymbol{\beta}$ and ϕ, respectively, being the distribution (4) proportional to

$$\prod_{i=1}^{k} \frac{\Gamma(\phi) \, r_i^{\mu_i \phi - 1} (1-r_i)^{(1-\mu_i)\phi - 1}}{\Gamma(\mu_i \phi) \, \Gamma((1-\mu_i)\phi)} \; e^{-\frac{1}{2} \sum_{j=1}^{p} (\beta_j^2 / v_j^2)} \; \phi^{-(a+1)} e^{-b/\phi}. \tag{5}$$

Notice that the mean μ_i is a function of the linear predictor $\eta_i = \mathbf{z}^T \boldsymbol{\beta}$, $i = 1, \ldots, k$.

The joint posterior distribution (5) is awkward to work with, since the marginal posterior distributions of some parameters are not easy to obtain explicitly. These posteriors can be evaluated using MCMC methods (see e.g. [8, 11, 17]). In particular Gibbs sampling that works by iteratively drawing samples for each parameter from the corresponding full conditional distribution, which is friendly implemented in software WinBUGS [12]. Other MCMC method, proposed by Hoffman and Gelman [11], is the No-U-Turn Sampler (NUTS) that is a variant of the Hamiltonian Monte Carlo (HMC), also known as hybrid Monte Carlo. Neal [16] presented HMC method in order to avoid a long time to converge to the posterior distribution as e.g. in Gibbs sampling by using a clever auxiliary variable scheme that transforms the problem of sampling from a posterior distribution into the problem of simulating Hamiltonian dynamics.

3.2 Evaluating and Comparing Models

An important issue in Bayesian data analysis is to choose among postulated sub-models of a statistical model, e.g. the beta regression model (2). Some summary measures of model comparison, such as the posterior mean of Deviance $D(\theta)$, where θ is the model parameter vector, are easily evaluated with MCMC methods. Other two measures of predictive accuracy are Deviance Information Criterion (DIC) and Watanabe-Akaike Information Criterion (WAIC) (see [8, 21]). DIC is here defined as

$$DIC = D(\overline{\theta}) + Var(D(\theta)), \qquad (6)$$

where $\overline{\theta}$ and $Var(D(\theta))$ denote the posterior mean of model parameter θ and the posterior variance of the deviance, respectively, whereas WAIC is defined by

$$WAIC = D(\overline{\theta}) + 2 \sum_{i=1}^{k} Var(D_i(\theta)), \qquad (7)$$

where $Var(D_i(\theta))$ denotes the posterior variance of the ith term of the deviance. DIC and WAIC handle Bayesian models of any degree of complexity, and models with smaller (6) and (7) should be preferred to models with larger ones.

4 Wildfire Data Analysis

For the wildfire data described in Sect. 2, we fitted several regression models based on the response, proportion of the burned area in Portugal during the period 1990–1994, as in Marques et al. [13], but now focusing on the beta regression instead of normal regression, and under a Bayesian perspective. One of the eight covariates presented in Table 1, i.e. slope orientation, was removed from the analysis by not showing any difference among its categories.

Let M_1 and M_3 denote regression beta model (1) with eight covariates showed in Table 1, apart from the covariate slope orientation, whereas M_2 and M_4 represent the corresponding normal models. Table 2 lists these sub-models of the beta and normal model with only main covariate effects, M_1 and M_2, and also with interactions between two covariates. Based on the comparison model measures DIC (6) and WAIC (6), fitted models with interactions had better evaluation than models with only main effects for both beta and normal models. These evaluating values were calculated taking into account the same response $r_i/(1-r_i)$ (so-called *odd*), which generates a sampling log-normal and second-kind beta distributions for normal and beta distributions, respectively. So, normal regression had better fitting than beta

Table 2 Model comparison measures of four fitted regression models for wildfire data

Regression models	WAIC	DIC
M_1: beta regression with only main effects	−10,161.37	−10,165.68
M_2: normal regression with only main effects	−14,593.76	−14,605.28
M_3: beta regression with interactions between two covariates	−10,723.74	−10,728.94
M_4: normal regression with interactions between two covariates	−15,263.72	−15,285.41

regression, and that can namely be associated with the large number of observations ($k = 25{,}388$). However, we chose to select model M_3 in order to illustrate the beta regression model that has not been employed in the analysis of wildfire burned areas, even as it can be considered the natural choice.

For all models showed in Table 2, we assumed prior normal distribution with mean zero and variance 10^4 for the regression parameters and prior inverse gamma and gamma distributions with shape parameter 1 and scale parameter 0.01 for the precision parameter ϕ (beta regression) and the variance σ^2 (normal regression), respectively. That is, highly dispersed, but proper prior distributions. MCMC samples of size 5000 were obtained for all models, after 2500 iterations of burn-in, implemented in software Stan [20]. A study of convergence of the samples was carried out with no worrying features.

For selected beta model M_3, Table 3 displays the model parameter estimates: posterior mean, standard deviation (SD) and 95 % highest posterior density (HPD) credible intervals (CI) for the model parameters. Note that related to the proportion of area burned in Portugal during the period 1990–1994:

1. There is no significant effect of annual and permanent crops in contrast to the other categories of land cover;
2. The land covers with larger likelihood to have wildfires are (in increasing order) agro-forestry, hardwoods, hardwoods and softwoods mixed with eucalyptus (HSME), resinous or softwoods (RS), eucalyptus, softwoods mixed with eucalyptus (SME), and shrubs (the most likelihood).
3. The proportion increases for larger categories of slope and altitude, whereas population and roads proximity display a decreasing effect in the proportion.
4. Because temperature and precipitation had an unexpected negative effect in the proportion, we decided to look for a potential interaction effect between the two covariates. We found significant interaction between temperature and precipitation in model M_3, even as that is not clear for the smaller categories of both covariates.
5. As large the categories of temperature and precipitation as large is the odd of burned area (interaction effect). Notice that the largest category did not have observation enough for confirming that.
6. The estimates in Table 3 also indicates that there is some dispersion in the proportion of area burned (see 95 % HPD credible interval of ϕ).

Table 3 Estimates of the regression parameters and dispersion parameter (ϕ) for model M_3

Parameter	Mean	SD	95 % CI Lower	95 % CI Upper	Parameter	Mean	SD	95 % CI Lower	95 % CI Upper
Roads proximity (β_{17})					Temperature ($\beta_{25}, \ldots, \beta_{29}$)				
<1000	−0.05	0.01	−0.08	−0.02	4–48	−2.63	0.12	−2.87	−2.39
Population (β_{18}, β_{19})					49–71	−2.63	0.13	−2.87	−2.38
25–100	−0.09	0.02	−0.12	−0.05	72–92	−2.67	0.13	−2.92	−2.43
≥100	−0.15	0.02	−0.19	−0.11	93–112	−2.94	0.16	−3.24	−2.63
Slope ($\beta_{14}, \beta_{15}, \beta_{16}$)					≥113	−2.78	0.20	−3.18	−2.41
10–20	0.21	0.02	0.18	0.24	Land cover ($\beta_2, \ldots, \beta_{10}$)				
20–30	0.66	0.02	0.61	0.70	Annual crop	−0.04	0.04	−0.11	0.03
≥30	2.16	0.06	2.05	2.27	Eucalyptus	0.49	0.04	0.42	0.56
Altitude ($\beta_{11}, \beta_{12}, \beta_{13}$)					Hardwoods	0.18	0.04	0.11	0.25
200–400	0.12	0.02	0.08	0.16	HSME	0.32	0.04	0.24	0.39
400–700	0.15	0.02	0.11	0.19	Agro-forestry	0.08	0.04	0.00	0.15
≥700	0.40	0.02	0.35	0.44	Permanent crop	−0.05	0.04	−0.13	0.02
Precipitation ($\beta_{20}, \ldots, \beta_{24}$)					Shrubs	0.61	0.03	0.54	0.68
7–13	−0.58	0.13	−0.83	−0.30	RS	0.48	0.04	0.42	0.55
14–18	0.08	0.18	−0.29	0.43	SME	0.59	0.04	0.51	0.66
19–22	0.43	0.19	0.06	0.80					
23–26	−0.62	0.17	−0.96	−0.29	Intercept (β_1)	0.81	0.13	0.57	1.05
≥27	−2.72	0.96	−4.62	−0.84	ϕ	1.36	0.01	1.34	1.39
(β_{30}) Temperature (4–48) × Precipitation (7–13)						0.70	0.14	0.40	0.95
(β_{31}) Temperature (4–48) × Precipitation (14–18)						0.03	0.19	−0.33	0.40
(β_{32}) Temperature (4–48) × Precipitation (19–22)						−0.11	0.19	−0.46	0.29
(β_{33}) Temperature (4–48) × Precipitation (23–26)						0.62	0.18	0.28	0.98
(β_{34}) Temperature (4–48) × Precipitation (≥27)						2.70	0.96	0.79	4.57
(β_{35}) Temperature (49–71) × Precipitation (7–13)						0.54	0.14	0.27	0.82
(β_{36}) Temperature (49–71) × Precipitation (14–18)						−0.15	0.19	−0.53	0.21
(β_{37}) Temperature (49–71) × Precipitation (19–22)						−0.35	0.19	−0.72	0.04
(β_{38}) Temperature (49–71) × Precipitation (23–26)						0.73	0.18	0.39	1.09
(β_{39}) Temperature (49–71) × Precipitation (≥27)						2.96	0.96	1.09	4.88
(β_{40}) Temperature (72–92) × Precipitation (7–13)						0.47	0.15	0.19	0.75
(β_{41}) Temperature (72–92) × Precipitation (14–18)						−0.04	0.19	−0.42	0.34
(β_{42}) Temperature (72–92) × Precipitation (19–22)						−0.35	0.19	−0.72	0.04
(β_{43}) Temperature (72–92) × Precipitation (23–26)						0.85	0.18	0.49	1.21
(β_{44}) Temperature (72–92) × Precipitation (≥27)						3.41	0.96	1.50	5.30
(β_{45}) Temperature (93–112) × Precipitation (7–13)						0.60	0.18	0.25	0.96
(β_{46}) Temperature (93–112) × Precipitation (14–18)						0.70	0.23	0.27	1.17
(β_{47}) Temperature (93–112) × Precipitation (19–22)						1.38	0.26	0.88	1.89
(β_{48}) Temperature (93–112) × Precipitation (23–26)						3.26	0.23	2.81	3.73
(β_{49}) Temperature (≥113) × Precipitation (7–13)						0.57	0.22	0.10	0.98

5 Concluding Remarks

This analysis of wildfire data in Portugal allow us to figure out the influence of the observed combinations of eight fire risk features on the proportion of burned area. Our results of beta regression are essentially consistent with those ones of normal regression, presented in Marques et al. [13], whose analysis did not include the explanatory variables: slope orientation, precipitation, temperature and the interaction between the last two ones. In fact, our model and conclusions bring improvements on the results reported by them based on a similar data set. So, we also identified changes in the proportion of burned area depending on topographical variables and vegetation type. Pereira et al. [18] pointed out that some variability of the burned area in Portugal is partly due both to the amount of precipitation in the fire season and in the preceding late spring season and to the occurrence of atmospheric circulation patterns that induce extremely hot and dry spells.

In addition, our intuition about interaction between precipitation and temperature was corrected, and we also believe that some latent variables can explain some unobserved heterogeneity in these wildfire data, e.g. spatial extra-variation across fire regions. For instance, Amaral-Turkman et al. [2] proposed a spatio-temporal model to analyse jointly the probability of ignition and fire sizes in Australia and New Zealand. Further research is being developed for capturing the spatio-temporal effects on the proportion of burned area, more proper sampling distributions and link functions. Notice that 4 % of observed burned areas were 0 or 1 being replaced by 10^{-10} and 1×10^{-10}, respectively, for simplicity, We intend to include that issue in future work, as well as to do a full sensitivity analysis of our prior options (see e.g. [8]) and some simulation to clarify the impact of a big data as our wildfires in the results. For the our choice of beta regression instead of normal regression, we also believe that a comprehensive simulation study must be done in order to verify the second choice, as well as the residual analysis for understanding that unexplained situation of the observed proportions close to one (see e.g. Espinheira et al. [4]).

References

1. Amaral-Turkman, M.A., Silva, G.L.: Modelos Lineares Generalizados - da teoria à prática. SPE Edition, Lisbon (2000)
2. Amaral-Turkman, M.A., Turkman, K.F., Le Page, Y., Pereira, J.M.: Hierarchical space-time models for fire ignition and percentage of land burned by wildfires. Environ. Ecol. Stat. **18**, 601–617 (2011)

3. Catry, F., Rego, F., Bação, F., Moreira, F.: Modelling and mapping wildfire ignition risk in Portugal. Int. J. Wildland Fire **18**, 921–931 (2009)
4. Espinheira, P.L., Ferrari, S.L.P., Cribari-Neto, F.: On beta regression residuals. J. Appl. Stat. **35**, 407–419 (2008)
5. Fernandes, P., Luz, A., Loureiro, C., Ferreira-Godinho, P., Botelho, H.: Fuel modelling and fire hazard assessment based on data from Portuguese National Forest Inventory. For. Ecol. Manage. **234S**, S229 (2006)
6. Ferrari, S.L.P., Cribari-Neto, F.: Beta regression for modeling rates and proportions. J. Appl. Stat. **31**, 799–815 (2004)
7. Flannigan, M.D., Amiro, B.D., Logan, K.A., Stocks, B.J., Wotton, B.M.: Forest fires and climate change in the 21st century. Mitig. Adapt. Strat. Glob. Chang. **11**, 847–859 (2005)
8. Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B.: Bayesian Data Analysis, 3rd edn. CRC Press, London (2014)
9. Gomes, J.F.P.: Forest fires in Portugal: how they happen and why they happen. Int. J. Environ. Stud. **63**, 109–119 (2006)
10. González, J.R., Pukkala, T.: Characterization of forest fires in Catalonia (Northeast Spain). Eur. J. For. Res. **126**, 421–429 (2007)
11. Hoffman, M.D., Gelman, A.: The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. arXiv **1111**, 4246 (2011). http://arxiv.org/abs/1111.4246
12. Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D.: WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Stat. Comput. **10**, 325–337 (2000)
13. Marques, S., Borges, J.G., Garcia-Gonzalo, J., Moreira, F., Carreiras, J.M.B., Oliveira, M.M., Cantarinha, A., Botequim, B., Pereira, J.M.C.: Characterization of wildfires in Portugal. Eur. J. For. Res. **130**, 775–784 (2011)
14. McCullagh, P., Nelder, J.A.: Generalized Linear Models, 2nd edn. CRC Press, Boca Raton (1989)
15. Moreira, F., Rego, F.C., Godinho-Ferreira, P.: Temporal (1958–1995) pattern of change in a cultural landscape of northwestern Portugal: implications for fire occurrence. Landsc. Ecol. **16** 557–567 (2001)
16. Neal, R.: Handbook of Markov Chain Monte Carlo, Chap. 5: MCMC Using Hamiltonian Dynamics. CRC Press, Chichester (2011)
17. Paulino, C.D., Amaral-Turkman, M.A., Murteira, B.: Estatística Bayesiana. Fundação Calouste Gulbenkian, Lisboa (2003)
18. Pereira, M.G., Trigo, R.M., da Camara, C.C., Pereira, J.M.C., Leite, S.M.: Synoptic patterns associated with large summer forest fires in Portugal. Agr. Forest. Meteorol. **129**, 11–25 (2005)
19. Pereira, J.M.C., Carreiras, J.M.B., Silva, J.M.N., Vasconcelos, M.J.: Alguns conceitos básicos sobre fogos rurais em Portugal. In: Pereira, J.S., Pereira, J.M.C., Rego, F.C., Silva, J.M.N., Silva, T.P. (eds.) Incêndios Florestais em Portugal, pp. 133–161. ISA Press, Lisboa (2006)
20. Stan Development Team.: Stan: A C++ Library for probability and sampling, Version 2.2. (2014). http://mc-stan.org/
21. Watanabe, S.: Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. **11**, 3571–3591 (2010)