Weekly cycles in daily global fire count time series[1]

Pereira J. M. C.[1], Amaral Turkman M. A.[2], Turkman K. F.[2] and Oom D.[1]

1 Centro de Estatística e Aplicações, Faculdade de Ciências da Universidade de Lisboa
2 Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa

Abstract

Vegetation burning displays spatial and temporal patterns at various scales, affected by many factors including the seasonal occurrence of dry periods and thunderstorms, and by sporadic heat waves and droughts. Regular cycles in vegetation burning are found at annual scale, mostly in seasonally dry areas, and at the daily scale as a consequence of underlying cycles in meteorological variables and land management practices. Since there is no natural forcing with a seven-day period, weekly cycles provide evidence of an anthropogenic fingerprint. Hence one natural approach to search and quantify anthropogenic influence on wildfire patterns is to look for existence of weekly cycles in wildfire incidences. Since biomass burning emissions have important climatic effects such as strong absorption of sunlight, study of weekly cycles in wild-fires may contribute in understanding and explanation of weekly cycles in meteorological variables. There is an ongoing debate if the reported existence of weekly cycles in vegetation burning is statistically significant or not. In order to answer this question we looked at global World daily number of fire activity consisting of 10 years of daily global fire counts observed during 2002-2012. Results clearly show statistically significant 7 day cycle with minimum fire activity on Sundays. Separating the study by land use, it was found that the seven day cycle and smaller fire activities on Sunday was observed on agriculture land and at some extend in forested and range land. Fire activity on settlements, villages and wild land did not show a seven day cycle, confirming the anthropogenic influence on fire activity.

[1]Research partially funded by FCT Fundação para a Ciência e a Tecnologia, Portugal, through the projects - UID/MAT/00006/2013 (AT, KFT, JMCP), UID/AGR/00239/2013 (JMCP and DO) and SFRH/BD/4752/2008 (DO).

1 Introduction

A potent tool for the investigation of the anthropogenic effect on variables is the identification of weekly cycles in them. Weekly cycles over broad geographical areas have been observed in several meteorological and climate variables (e.g. Georgoulias and Kourtidis, 2011).

Since there is no natural forcing with a seven-day period, weekly cycles provide evidence of an anthropogenic fingerprint.

The most plausible hypothesis to explain large- scale weekly cycles is through direct and indirect effects of anthropogenic aerosols, whose emissions also tend to show weekly periodicity at large-scale. Vegetation burning is a large source of greenhouse gases and aerosols and constitutes a major factor controlling variability of atmospheric composition at the inter-annual scale. Since biomass burning emissions have important climatic effects such as strong absorbtion of sunlight, study of weekly cycles in wild-fires may also contribute in understanding and explanation of weekly cycles in meteorological variables. Weekly cycles of global fires were reported by Earl et al.(2015). Also Pereira et al. (2015) found significantly weekly cycles in sub-Saharan Africa cropland burning, with Sunday minima in predominantly Christian regions, and Friday minima in mainly Muslim regions.

Our goal in the present study is to revisit the issue of detecting weekly cycles in vegetation fire activity at global scale, relying on the original MODIS active fire detections at 1 km spatial resolution (Oom and Pereira, 2013). Earl et.al (2015) in their analysis of weekly cycles of global fires, used Monte Carlo methods to test the null hypothesis of no significant weekly cycles in fire activity. This was done based on the range between the days of the week with the minimum and maximum of the total of active fires. To remove the effect of annual and seasonal cycles, they used the daily deviations from the 31 day running mean. They used two ways to construct the data for the Monte Carlo approach. One relied on random reshuffling the original time series and the other on randomly removing 5% of the data retaining the order of the elements of the time series. Our analysis goes much deeper into the problem in that, rather than using bootstrap techniques, we identify seven day cycles trough the analysis of the periodogram of the original time series. Based on peridogram analysis and tests on the significance of jumps observed in the periodogram, we construct harmonic regression models for the mean function of daily fire activity and adjust a linear model for the de-seosanized time series of daily global fire activity. Using a Bayesian analysis this mean function is treated as random and a simultaneous 95% credible band is constructed. Sunday minimum is hence directly investigated by computing the probabilities that the the mean functions of every week day (Monday to Saturday) are inside the credible band corresponding to Sunday. Since these probabilities are considerably small, they indicate that there is a statistical evidence that Sunday fire activities are stochastically less than the fire activities observed in other days of the week. Our findings also clearly show that the estimated mean of Sunday fire activity is uniformly less than the estimated mean fire activity observed on other days of the week.

The data analyzed consist of $n = 3675$ daily global fire counts $N(t)$ from 08-07-2002 to 29-07-2012, corresponding to 525 weeks.

We also spatially disaggregate the analysis of weekly cycles in global fire activity by major global land use type, represented by the villages, settlements, croplands, rangelands, forests, and wildlands (Ellis et al. 2010). Anthrome-specific analysis provides valuable information to elucidate the land management activities, regions, and times of the year most responsible for inducing weekly cycles in global vegetation burning.

In section 2 we introduce the methodology used to identify the existence of a seven day cycle in the time series of the daily global fire counts. We start with a preliminary data analysis of the periodogram of the series, propose a parametric model based on this analysis, estimate the model and test its adequacy. In subsection 2.3 we propose a Bayesian approach to compare the mean functions corresponding to the different week days. In section 3 we perform an anthrome-specific analysis. In Appendix we have extra plots which are referred to in the main text.

2 Methods and Results; global data

2.1 Preliminary data analysis

Log transformed time series of counts together with its autocorrelation function (acf) and partial autocorrelation function (pacf) up to 700 time lags are displayed in Figures 1 and 2.

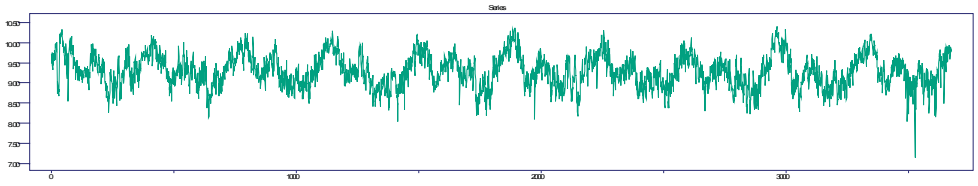

Figure 1: Log-transformed daily fire counts

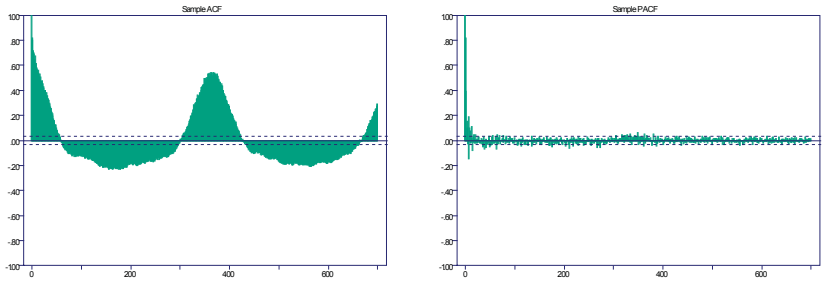

Figure 2: acf and pacf of the series up to lag 700

The plot of the time series $N(t), t = 1, ..., 3675$, as well as the acf-pacf clearly display strong yearly seasonal variations and this can be confirmed by the periodogram of the series given in Figure 3.

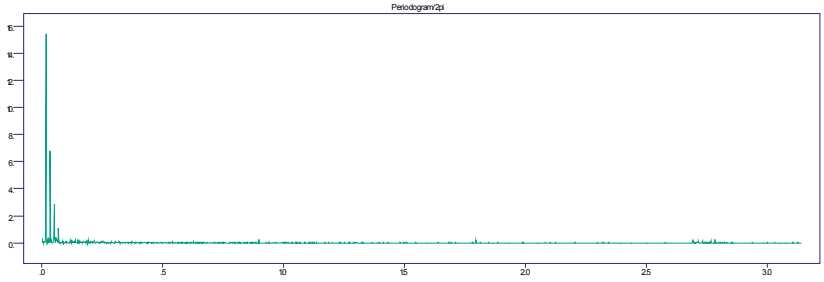

Figure 3: Periodogram of the series

The periodogram shows statistically significant jumps at frequencies ($w = 2\pi j/n$) corresponding to 1 cyle in 365 days ($j = 10$) and its first 3 harmonics corresponding respectively to 1 cyle in 6 months ($j = 20$) 4 months ($j = 30$) and 3 months ($j = 40$). When these harmonics are eliminated by fitting an harmonic regression to the data, the periodogram and its adequately smoothed versions (see Figures 4 and 2.1) of the residual process unmasks further significant jumps at frequencies corresponding respectively to 1 cycle in 7 days ($j = 525$) and 1 cycle at 3 and half days ($j = 1050$).

The significant jumps in the periodogram at frequencies 0.898 ($j = 525$) and 1.795 ($j = 1050$) corresponding to 7 day and 3 and half day cycles are particularly interesting. Using the software ITSM (Brockwell and Davis, 1991) Fisher's spectrum test (Brockwell and Davis, 1991) gives the ratio of maximum periodogram ordinate to mean as 595.3484 and consequently the probability of exceeding this value under the null hypotheses is 0. Fisher's test (see Appendix) is constructed to test the null hypotheses that the time series is a Gaussian white noise against the alternative hypothesis that the series contains an added deterministic periodic component of unspecified frequency. Whittle's test (in

Appendix) can be extended to test the significance of second, third, etc, largest peaks in the periodogram (Nicholls, 1967) estimation of the spectral density function when testing for a jump in the spectrum. See also Murteira et al. (1993).

Application of these tests to the observed peaks at frequencies 0.898 and 1.795, shows that these jumps are significant (see Figure 15 in Appendix), thus are due to the added deterministic periodic components in these particular frequencies.

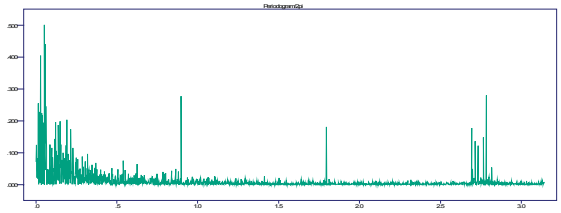

Figure 4: Periodogram of the partially deseasonalized series

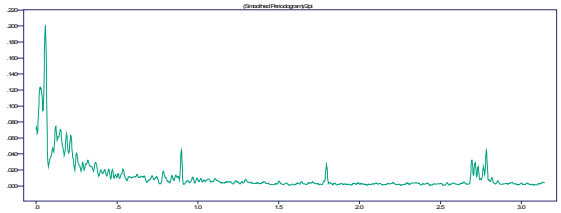

Figure 5: Smoothed periodogram

2.2 The model

The above results based on spectral analysis clearly suggest that the time series is non-stationary with seasonal mean function which can be represent by an harmonic sum

$$S(t) = \sum_{i=1}^{6} c_i \cos \omega_i t + d_i \sin \omega_i t,$$

where ω_i, $i = 1, 2, .., 6$ are the angular frequencies corresponding respectively to 1 cycle per year, 6 months,4 months, 3 months, 7 days and 3 and half days days. Based on these observations, we suggest the model

$$Y_t = M_t + X_t, \tag{1}$$

where $Y_t = \log(N(t))$, $M_t = T_t + S_t$ is the deterministic mean function with linear trend $T_t = a_0 + bt$, and seasonal component

$$\begin{aligned} S_t &= c_1 \cos(2\pi t 10/n) + d_1 \sin(2\pi t 10/n) \\ &= c_2 \cos(2\pi t 20/n) + d_2 \sin(2\pi t 20/n) \\ &= c_3 \cos(2\pi t 30/n) + d_3 \sin(2\pi t 30/n) \\ &= c_4 \cos(2\pi t 40/n) + d_4 \sin(2\pi t 30/n) \\ &= c_5 \cos(2\pi t 525/n) + d_5 \sin(2\pi t 525/n) \\ &= c_6 \cos(2\pi t 1050/n) + d_6 \sin(2\pi t 1050/n) \end{aligned} \tag{2}$$

and $X_t = \phi_0 + \sum_{i=1}^{7} \phi_i + Z_t$ a stationary, zero mean process compatible with a stationary and invertible Gaussian $AR(7)$ process, where, Z_t is a zero mean, uncorrelated (white noise) process.

To study the adequacy of this model, we fitted it to the log transformed data with $n = 3675$ using the software ITSM (Brockwell and Davis, 1991). Figure 6 shows the logarithm of the daily fire counts and the fitted mean function.

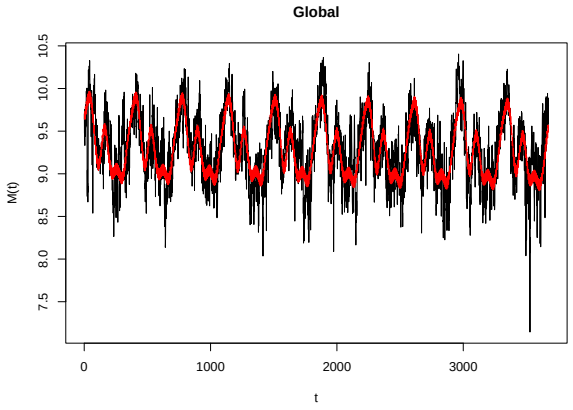

Figure 6: logarithm of the daily fire counts and the fitted mean function

Figure 7 gives the empirical(green) and theoretical(red) acf and pacf of the process X_t, indicating the good fit of the above $AR(7)$ process, whereas Figure 8 gives the acf and the pacf of the residuals obtained from the fitted model $AR(7)$, which is consistent with a zero mean white noise process.

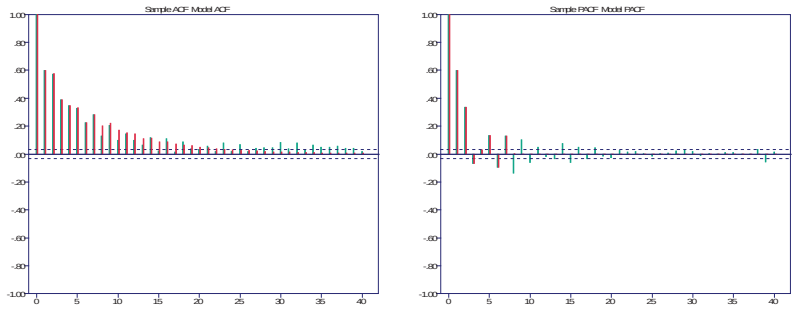

Figure 7: acf and pacf of the sample(green) and model(red) for the process $X_t = Y_t - \hat{M}_t$

Figures 9 and 10, respectively give the Normal qq-plots and the periodogram of the residuals obtained upon fitting the model $AR(7)$.

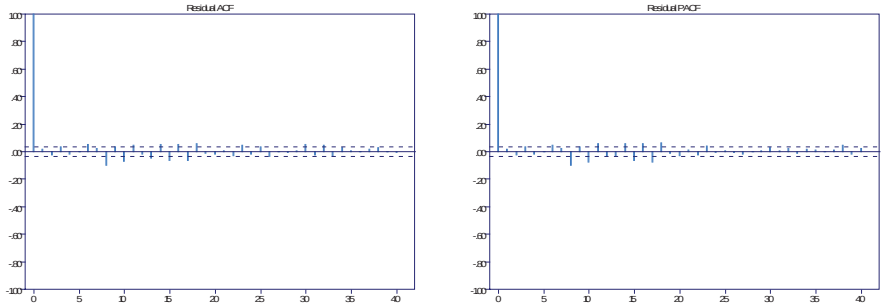

Figure 8: acf and pacf of the residuals upon fitting the model $AR(7)$ model

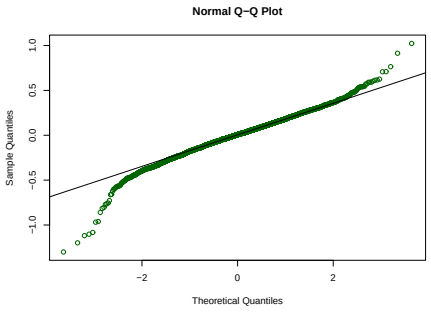

Figure 9: qq-plot of the residuals upon fitting the model $AR(7)$ model

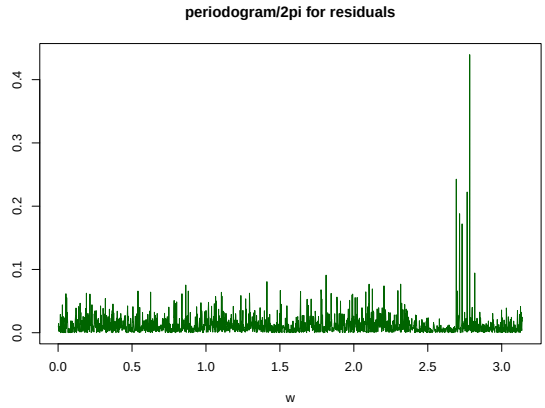

Figure 10: Periodogram of the residuals upon fitting the model $AR(7)$ model

Although the qq-plot of the residuals in Figure 9 shows some problems on the tails, it is clear that the model in (1) gives an adequate fit to the log transformed counts, resulting in residuals consistent with zero mean Gaussian white noise process. A model selection procedure also showed, using a Likelihood Ratio Test (LRT) that all the six harmonics are significant (see table 4). A detailed inspection of the model, also showed that the linear trend was not significant (see table 7 in the Apendix).

2.3 Comparing the mean functions: Bayesian approach

While studying weekly cycles in meteorological, and in general in environmental data, it is of importance to show that with some level of confidence, the observed cycles are not the result of some natural variability. It is therefore legitimate to ask if the observed 7 day and 3 and half day cycles are merely due to some natural variability. However, for this study an explanation can be given as these observed cycles are not caused by simple natural variability. First, there are natural explainable causes for such cyclic movements. Second, the existence of statistically significant jumps in spectrum at frequencies corresponding to these cycles, is a clear sign that the mean of the process has a seasonal component composed of fourier terms at these frequencies.

Once we establish that fire activity indeed displays weekly cycles, it may be interesting to see in which day of the week the minimum fire activity is observed, in particular, one may want to find some evidence that indeed fire activity is minimum on Sundays. One way of answering this question is to compare the mean functions M_{t_i} in (1) with $t_i = i, i+7, i+14, ...$ and $i = 1, ..., 7$, where $i = 1$ refers to Monday and $i = 7$ refers to Sunday. That is compare the mean functions corresponding to the different days of the week.

In order to answer this question, we followed a Bayesian approach. The choice of this methodology is justified by the fact that one can interpret the

mean function M_t as a vector of random parameters and perform the inference based on their posterior distribution.

To fit the model $Y_t = M_t + X_t$, for the logarithm of the counts, with X_t an AR(7) process, we used an integrated nested Laplace approximation (INLA) introduced by Rue et al. (2009), which is implemented in the software R-INLA available in www.r-inla.org. Relatively to the fitting of the model, results were compatible with the ones obtained using the frequentist approach implemented with the software ITSM (Brockwell and Davis, 1991) as described in 2.2. From the analysis of the model, we observe that all the harmonic coefficients are significant, but there is no evidence of a linear trend.

According to the theory behind the INLA approach, the posterior distribution for $M_t, t = 1, ..., 3765$ is approximately multivariate normal with mean $\hat{M}_t, t = 1, ..., 3765$ and posterior covariance matrix S, which can be obtained from R-INLA using `inla.make.lincomb` function. Pointwise $100(1-\gamma)\%$ HPD intervals for each component of the vector M_t can also de obtained using the function `inla.hpdmarginal`. However, rather than obtaining pointwise credible intervals, one should seek for a simultaneous credible band for M_t of size $100(1-\alpha)\%$. For that we follow the suggestion given by Sorbye and Rue (2011).

Defintion: For the random vector $\{M_t, t = 1, ..., n\} = (M_1, ..., M_n)$ a $100(1-\alpha)\%$ simultaneous credible band can be defined as a hyper-rectangular region

$$R = \{\{M_t\} : \bigcap_{i=1}^{n} (M_i \in I_{i,\gamma})\}$$

where $\{I_{i,\gamma}\}$, $i = 1, ..., n$ denotes a set of individual $100(1-\gamma)\%$ HPD credible intervals for each component of the vector M_t, with γ chosen so that $P(\{M_t\} \in R) = 1 - \alpha$.

Since the posterior distribution for the vector M_t is approximately multivariate normal with known mean vector and known covariance matrix, by trial and error, one can obtain a simultaneous $100(1-\alpha)\%$ HPD credible band for the vector $\{M_t, t = 1, ..., n\}$, by computing, for every individual component of the vector, a $100(1-\gamma)\%$ HPD credible interval, varying γ till the overall probability is $1 - \alpha$.

$$P(\{M_t, t = 1, ..., n\} \in R) = 1 - \alpha.$$

Actually this can also be obtained using the function `excursions.simconf` available in the R package `excursions` (Bolin and Lindgren, 2015).

Considering the fact that the linear trend was not significant and considering the cyclic nature of M_t, this was done on year basis, rather than for the full period.

In Figure 11 we display the series of the log counts with the estimated mean function $\hat{M}_t$, and in Figure 12 the estimated mean function for the first year, ie, $\hat{M}_t, t = 1, ..., 364$ together with 95% pointwise HPD intervals and 95% simultaneous credible bands.

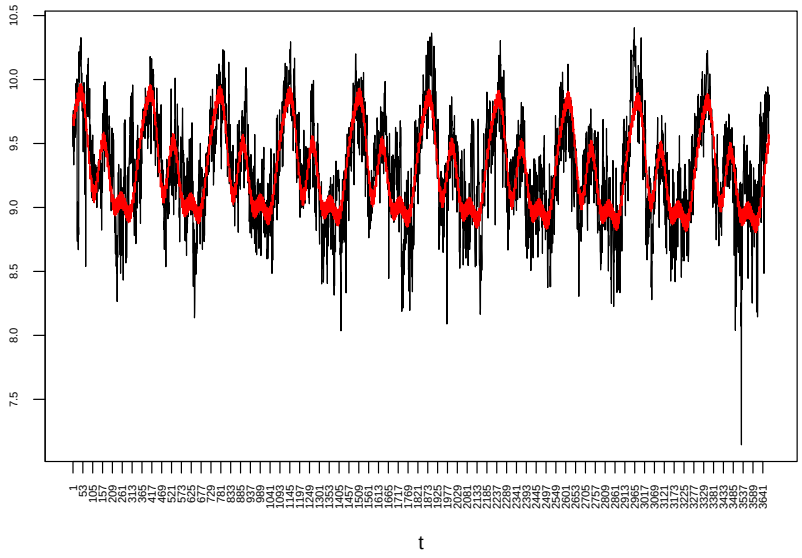

Figure 11: Log transformed fire counts and fitted posterior mean functions $\hat{M}_t$

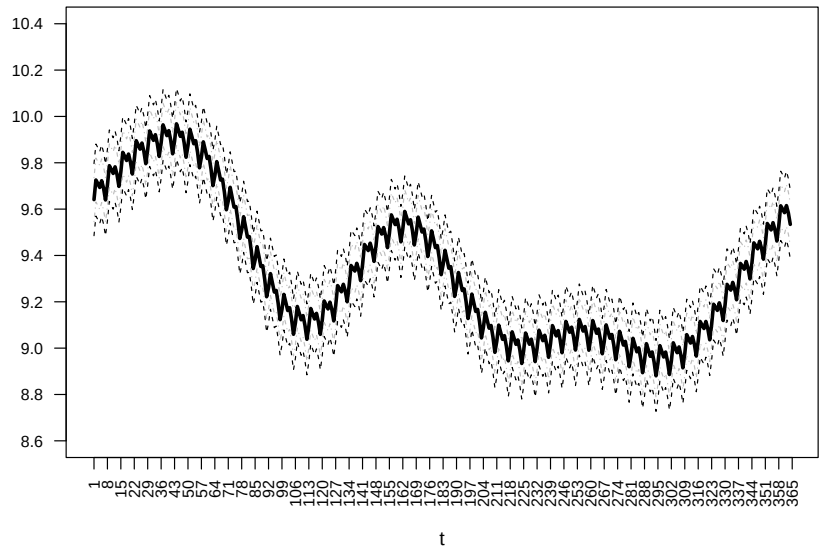

Figure 12: Estimated posterior mean function $\hat{M}_t, t = 1, ..., 364$ together with the 95% pointwise HPD intervals (dotted, grey) and 95% simultaneous credible bands (dashed)

Since the posterior distribution of the mean function M_t is multivariate normal, M_t can be partitioned into $M_{t_i}, i = 1, ..., 7$ (with 1 for Monday, 7 for Sunday), as explained before, to represent the mean functions for the different days of the week. Properties of the multivariate normal distribution imply that the posterior distributions of M_{t_i}, $t_i = i, i + 7, ...$ and $i = 1, ..., 7$ are also approximately multivariate normal with means given by the $\hat{M}_{t_i}$, $t_i = i, i+7, ...$ and $i = 1, ..., 7$ and corresponding covariance matrixes, which are sub-matrices of the overall covariance matrix. Also if we represent the $100(1 - \alpha)\%$ HPD credible bands for M_t by the two series LB_t and UP_t, they can similarly be partitioned into LB_{t_i} and UP_{t_i}, for $i = 1, ..., 7$.

We propose as a methodology to "test" the hypothesis that there is, on average, smaller fire activity on Sundays, by computing the probabilities, for $i = 1, ..., 6$

$$P\left(LB_{t_7} \leq M_{t_i} \leq UB_{t_7}\right),$$

that is, the probabilities that the mean function for every day of the week is inside the credible bands corresponding to Sunday, ie, inside LB_{t_7} and UP_{t_7}. These probabilities can be easily obtained using the function `gaussint()` from the R package `excursions`.

In table 1 we show these probabilities with the associated error in brackets

computed using the first year. For the other years the results are very similar.

Table 1: Probabilities that mean functions for the logarithm of the fire counts for each day of the week are inside the 95% simultaneous credible band for Sunday

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
0.1965	0.0066	0.0712	0.3140	0.2041	0.7019	0.9696
(0.0036)	(0.0006)	(0.0022)	(0.0041)	(0.0033)	(0.0039)	(0.0013)

We also compute the posterior probabilities that the mean function for every other week day $M_{t_i}, i = 1, ..., 7$ is above the estimated mean function for Sunday (the mean of the posterior distribution of the mean vector M_{t_7}). These probabilities are on table 2.

Table 2: Probabilities $P = P(M_{t_i} > \hat{M}_{t_7}$ that mean functions for the logarithm of the fire counts for each day of the week are above the estimated mean function for Sunday

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
P	0.0556	0.7644	0.6303	0.4531	0.7412	0.3511	0.0172
	(0.0012)	(0.0033)	(0.0037)	(0.0037)	(0.0034)	(0.0036)	(0.0007)

Figure 13 is an unfolded reproduction of figure 12, that is, the estimated mean function $\hat{M}_t$ is partitioned into $\hat{M}_{t_i}$, $t_i = i, i+7, i+14, ...$, and $i = 1, ..., 7$ to represent the estimated mean function for the different days of the week. For easiness in reading, only the first year is plotted. Other years have the same behaviour, due to the cyclic nature of the mean function, since the linear trend is not statistically significant.

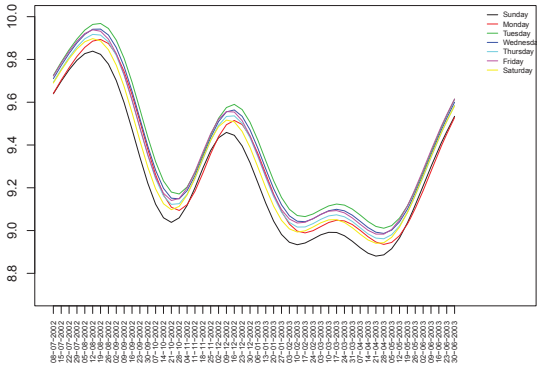

Figure 13: Estimated mean functions $\hat{M}_{t_i}$ of the logarithm of the fire counts for each day of the week

It is clear from that plot that the estimated mean function corresponding to Sunday fire counts is smaller than the other estimated mean functions almost uniformly for every week. Looking in more detail we remark that the estimated mean function for Sunday shows always lower values, except for the period ranging from the second week of November till the first week of December and from the third week of May till the end of June, when it is always very close to the estimated mean function for Monday. The second minimum is basically on Saturday or Monday. This may indicate the existence of a weekend effect on daily fire counts. To understand how much smaller the estimated mean function for Sunday is from the other days of the week, we display in Figure 14 the estimated mean function for Sunday together with the corresponding 95% simultaneous credible band and the estimated mean functions for each of the other days of the week. We observe that the estimated mean function for Saturday is always inside the credible band for Sunday and only during two periods of the year, namely, third week of August till third week of October and second week of December till the second week of April, the estimated mean functions for some of other days of the week are outside the upper bound of the credible band for Sunday. However, they are not outside the upper bound of the credible band for Saturday.

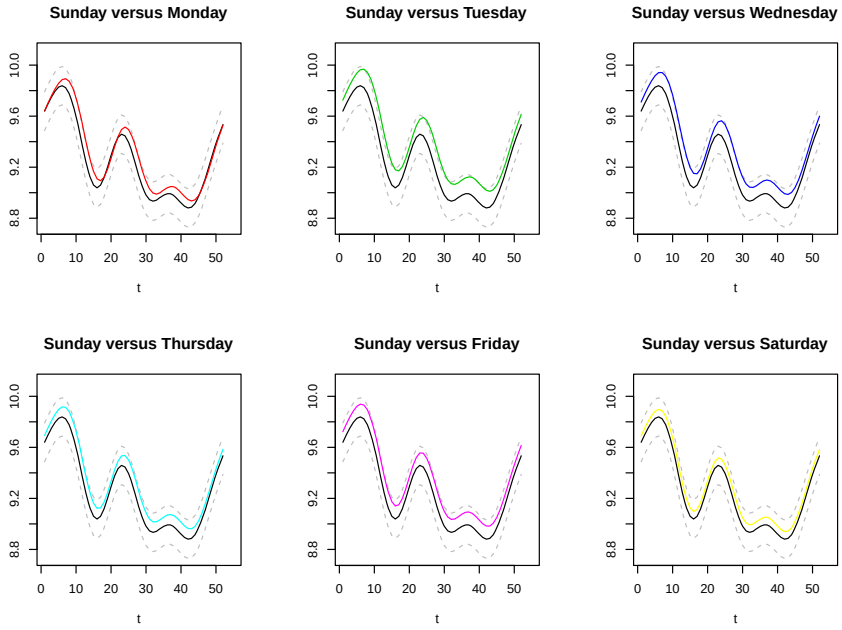

Figure 14: Estimated mean function of the logarithm of the fire counts for Sunday and corresponding 95% simultaneous credible band, with the estimated mean functions $\hat{M}_{t_i}$ of the logarithm of the fire counts for each other day of the week

3 Disaggregation of the global data by anthromes

The same methodology was applied to the data after being spatially disaggregated by anrhromes. A summary statistics for the logarithm of the number of daily fire counts $N_t, t = 1, ..., 3675$, according to the type of anthrome is in table 3.

Table 3: Summary statistics for the logarithm of the number of daily fire counts according to the type of Anthrome

	Croplands	Rangelands	Forested	Wildland	Settlements	Villages
Min.	6.44	5.59	5.78	0	0.69	3.69
1st Qu.	8.11	7.13	7.64	3.20	3.04	5.53
Median	8.41	7.72	8.04	4.38	3.43	5.88
Mean	8.37	7.63	8.05	4.44	3.43	5.90
3rd Qu.	8.66	8.10	8.46	5.71	3.78	6.26
Max.	9.48	9.20	9.56	8.64	6.97	8.23

The plot of the periodograms of the logarithm of fire counts according to the

six type of anthromes are in Figures 16 to 21 on the left in the Appendix. Only for croplands Whitle's test (Figures 16 to 21 on the right side), show strong evidence of significant jumps in the periodogram at frequencies 0.898 ($j = 525$) and 1.795 ($j = 1050$), although some evidence is also observed for rangelands and forested.

Model (1) was applied separately to the six time series of the logarithm of fire counts according to the type of anthrome. Likelihood ratio test (LRT) was used to choose the "best" model by sequentially eliminating the higher frequencies.

Table 4 shows the result of the LRT result test (with corresponding p-value in brackets) comparing the full model (1) and the different models fitted. Since, from one model to the other only the value of k in $S_t = \sum_{i=1}^{k} c_i \cos \omega_i t + d_i \sin \omega_i t$ changes, in the table the first column indicates the value of k for the fitted model. Notice that $k = 6$ corresponds to the model (1) with which the others are compared to, $k = 1$ to the model with only the one year cycle and $k = 0$ a model without the harmonic sum S_t.

Table 4: Likelihood ratio test results

k	Global	Croplands	Rangeland	Forested	Wildland	Settlements	Villages
5	64.26 (0)	75.48 (0)	38.15 (0)	35 (0)	1.06 (0.59)	5.10 (0.08)	11.09 (0.004)
4	-	-	-	-	6.40 (0.17)	7.06 (0.13)	15.68 (0.003)
3	-	-	-	-	7.97 (0.24)	8.31 (0.22)	31.02 (0)
2	-	-	-	-	9.81 (0.28)	10.66 (0.22)	-
1	-	-	-	-	17.45 (0.07)	29.02 (0.001)	-
0	-	-	-	-	59.16 (0)	-	-

As a result of this study we applied the methodology described in subsection 2.3 to the full model for Croplands, Rangelands, Forested and Villages and for Wildlands we used the model with only one year cycle, and for Settlements with the one year and six month cycles. Similar plots to the ones in figures 13 and 14, for the different anthromes are in Appendix (Figures 22 to 32 and 23 to 33).

For the time series corresponding to the different anthromes, table 5 shows the probabilities that the mean function for every day of the week is inside the 95% credible band corresponding to Sunday and table 6 displays the probabilities that the mean function for every day of the week is above the estimated mean function for Sunday.

Table 5: probabilities that the mean function for every day of the week is inside the 95% credible band corresponding to Sunday

Anthrome	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Croplands	0.0122	0.0000	0.0002	0.0227	0.0013	0.3770	0.9750
	(0.0007)	(0.0000)	(0.0000)	(0.0012)	(0.0002)	(0.0040)	(0.0012)
Rangeland	0.0455	0.0005	0.0145	0.1779	0.2051	0.7628	0.9724
	(0.0017)	(0.0002)	(0.0009)	(0.0032)	(0.0033)	(0.0037)	(0.0013)
Forested	0.2684	0.0309	0.1153	0.2946	0.1890	0.6995	0.9673
	(0.0041)	(0.0014)	(0.0027)	(0.0040)	(0.0033)	(0.0039)	(0.0014)
Wildland	0.7001	0.7898	0.8584	0.9026	0.9293	0.9487	0.9554
	(0.0043)	(0.0037)	(0.0032)	(0.0027)	(0.0023)	(0.0020)	(.0020)
Settlements	0.9011	0.8047	0.9164	0.9566	0.9729	0.9653	0.9770
	(0.0027)	(0.0035)	(0.0024)	(0.0018)	(0.0014)	(0.0015)	(0.0012)
Villages	0.7517	0.5565	0.7974	0.9440	0.9384	0.9526	0.9742
	(0.0039)	(0.0045)	(0.0036)	(0.0020)	(0.0021)	(0.0019)	(0.0013)

Table 6: Probabilities that the mean function for every day of the week is above the estimated mean function for Sunday.

Anthrome	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Croplands	0.0367	0.9273	0.8004	0.6606	0.9492	0.5014	0.0010
	(0.0012)	(0.0022)	(0.0032)	(0.0038)	(0.0017)	(0.0039)	(0.0001)
Rangeland	0.0069	0.5622	0.4803	0.2961	0.5133	0.1209	0.0014
	(0.0003)	(0.0038)	(0.0037)	(0.0033)	(0.0038)	(0.0022)	(0.0002)
Forested	0.0034	0.3630	0.3361	0.2981	0.5919	0.1801	0.0013
	(0.0003)	(0.0038)	(0.0037)	(0.0035)	(0.0040)	(0.0029)	(0.0002)
Wildland	0.0001	0.0002	0.0002	0.0005	0.0010	0.0008	0.0012
	(.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)
Settlements	0.0142	0.1173	0.0560	0.0041	0.0146	0.0522	0.0199
	(0.0008)	(0.0025)	(0.0017)	(0.0004)	(0.0007)	(0.0015)	(0.0008)
Villages	0.0150	0.1597	0.0821	0.0172	0.0588	0.0664	0.0167
	(0.0006)	(0.0027)	(0.0019)	(0.0007)	(0.0015)	(0.0017)	(0.0007)

From tables 5 and 6 we see that there is a strong evidence that for croplands, the mean function for Sunday is statistically smaller than the mean function for every other day of the week. This evidence is not so strong for rangelands and forests since the probability that Saturday mean function is inside the 95% credible band or Sunday is relatively high. For the other anthromes there is no statistical evidence of a difference between the mean functions corresponding to the different days of the week.

From the plots in Appendix, it is clear that, for croplands, the estimated mean function corresponding to Sunday fire counts is smaller than the other estimated mean functions almost uniformly for every week. To understand how much smaller the estimated mean function for Sunday in Croplands is from the

other days of the week, we display in Figure 23 the estimated mean function for Sunday together with the 95% simultaneous credible band and the estimated mean functions for each of the other days of the week.

In section ?? of the Appendix, we refer to the periods of the year when the estimated mean function of every specific week day is above the upper simultaneous credible bound corresponding to Sunday, according to the series of Croplands, Rangelands and Forested. For Wildland, Settlements and Villages, the estimated mean functions are always inside the simultaneous credible band for Sunday.

4 Discussion and Conclusions

In order to investigate if the reported existence of weekly cycles in vegetation burning is statistically significant or not, we looked at a time series of the global World daily number of fire activity consisting of 10 years of daily global fire counts observed during 2002-2012. Results clearly show statistically significant 7 day cycle with minimum fire activity on Sundays. Disaggregation of the analysis of weekly cycles in global fire activity by major global land use type, showed that the seven day cycle and a smaller fire activities on Sunday was observed on cropland and at some extend in forested and rangeland. Fire activity on settlements, villages and wild land did not show a seven day cycle, confirming the anthropogenic influence on fire activity.

To arrive at these conclusions we started with a preliminary analysis of the time series of the logarithm of daily global fire counts. The autocorrelation and partial autocorrelation functions showed the existence of strong yearly seasonal variations, confirmed by the periodogram of the series. The periodogram shows the existence of significant jumps at frequencies corresponding to one year cycle, six months cycle, four months cycle and three months cycle. When these harmonics are eliminated by fitting an harmonic regression to the data, the periodogram and its adequately smoothed versions of the residual process unmasks further significant jumps at frequencies corresponding respectively to one cycle in 7 days and one cycle at 3 and half days. Application of Fisher and Whittle's tests to the observed peaks at those frequencies, shows that these jumps are statistically significant.

Based on the results on spectral analysis a non-stationary time series model was fitted to the logarithm of the daily fire counts with a seasonal mean function represented by an harmonic sum with six angular frequencies corresponding to one cycle per year, six months, four months, three months, seven days and three and half days. The error structure was considered to be a stationary, zero mean process compatible with a stationary and invertible Gaussian AR(7) process. A detailed analysis of the fitted model showed an adequate fit. Besides a model selection based on a Likelihood ratio test showed that all the harmonic coefficients were needed in the model.

Once we established that fire activity displays weekly cycles we looked for evidence that fire activity is, on average, minimum on Sundays, by comparing the

mean functions relative to the different days of the week, with the mean function relative to Sundays. This comparison was made using a Bayesian methodology in which the mean function is considered random and the inference is made based on the corresponding posterior distribution. The same model as before was fitted using an Integrated Nested Laplace Approximation. Based on the fact that the posterior distribution of the mean function is approximated by a multivariate normal, a simultaneous 95% credible band for the mean function was constructed. Also based on properties of the multivariate distribution, it was possible to compute, for each week day (Monday to Sunday), the probabilities that the mean functions are inside the credible band corresponding to Sunday. For the global data we observed that these probabilities are low except for Saturday in which case is approximately 0.7.

The above analysis and methodology for the global data was replicated for the six time series corresponding to the different types of land use. Base on the analysis of the periodograms and on Whittle's test, the existence of the seven day cycle was more evident for croplands and in a small extent for rangelands and forested. A Likelihood ratio test confirmed the need for keeping in the models for the mean function the six harmonics. For each of the time series, computation for each week day (Monday to Sunday), of the probabilities that the mean functions are inside the credible band corresponding to Sunday, showed that, for croplands, all the probabilities are very low (below 0.025), except for Saturday when this probability is around 0.38. Probabilities for the time series for Rangelands and for Forested are on the same range as those obtained for the global data, being relatively again high for Saturday (above 0.7). On the other hand for wildlands, settlements and villages, the seven day cycle was not observed through the analysis of the periodogram. Besides, the calculation of the probabilities that the mean functions are inside the credible band corresponding to Sunday were all very high.

REFERENCES

Bolin D. and Lindgren F. (2015). Excursion and contour uncertainty regions for latent Gaussian models, *Journal of the Royal Statistical Society*, Series B, **77**, 85-106.

Brockwell P.J. and Davis R.A. (1991) *Time Series: Theory and Methods.* New York: Springer.

Earl N., Simmonds I and Tapper N. (2015). Weekly cycles of global fires-Association with religion, wealth and culture, and insights into anthropogenic influences on global climate. *Gheophysical Reserach Letters*, **42**, 9579-9589, doi:10.1002/2015GL066383.

Ellis EC, Goldewijk KK, Siebert S, Lightman D, and Ramankutty N (2010) Anthropogenic transformation of the biomes, 1700 to 2000. *Global Ecology and Biogeography*, **19**, 589-606.

Georgoulias AK and Kourtidis KA (2011) On the aerosol weekly cycle spatiotemporal variability over Europe. *Atmos. Chem. Phys.*, **11**, 4611-4632.

Murteira B.F., Muller D. and Turkman K.F. (1993) *Análise de sucessões cronológicas*. Lisboa: McGraw-Hill.

Nicholls D.F. Estimafion of the Spectral Density Function When Testing for a Jump in the Spectrum (1967). *Australian Journal of Statistics*, **9**, 103-108

Oom, D., and Pereira, J.M.C. (2013) Exploratory spatial data analysis of global MODIS active fire data. *International Journal of Earth Observation and Geoinformation*, **21**, 326âĂŞ340.

Rue H., Martino S. and Chopin N. (2009). Approximate Bayesian inference for latent Gaussian models using inte- grated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society*, Series B, **71**, 319-392.

Sørbye S.H. and Rue H. (2011) Simultaneous Credible Bands for Gaussian Latent Models. *Scandinavian Journal of Statistics*, **38**, 712-725.

5 Appendix

5.1 Periodograms and Whittle test on the periodograms

Given a time series Y_t with $t = 1, ..., N$, the periodogram is defined in the interval $[-\pi, \pi]$ as

$$I_N(\omega) = \frac{2}{N} \mid \sum_{t=1}^{N} Y_t \mathrm{e}^{-iwt} \mid^2 .$$

Although defined on the interval $[-\pi, \pi]$ it is usually calculated at the frequencies $\omega_j = \frac{2\pi j}{N}$ for j=0,1,...,[N/2].

To test if the periodogram contains a value substantially larger than the mean value, Fisher constructed a test of the null hypothesis that Y_t is a Gaussian white noise against the alternative hypothesis that Y_t contains an added deterministic periodic component. The test is based on the statistic

$$T_1 = \frac{\max_{j=1,...,q}(I_N(\omega_j)}{\sum_{j=1}^{q} I_N(\omega_j)},$$

where $q = [N/2]$ Under the null hypothesis $P(T_1 > g)$ can be approximated by

$$P(T_1 > g) \approx [N/2](1-g)^{[N/2]-1}.$$

Whittle suggested that Fisher's test could be extended to the significance of the peaks of 2nd, 3rd, ..., order by systematically removing the peaks revealed to be significant and applying the Fisher's test to study the significance of the ones of lower order. Suppose that $I^{\star}_{\max}$ is the peak of second order. Than FisherÂt's test is applied to the statistic

$$T_2 = \frac{I^{\star}_{\max}}{\sum_{j=1}^{q} I_N(\omega_j) - I_{\max}}.$$

We applied Whittle's test to study the significance of the peaks of the periodograms of the different time series under study. The p-values obtained were adjusted using Bonferroni's method of adjusting considering the possibility of finding a high number of significant peaks. Peaks corresponding to the frequencies for $j = 10, 20, 30, 40, 525, 1050$ where considered significant if the adjusted p-value was smaller than 0.01. If in Figures from 15 to 21 the corresponding points $(j, p - value_j)$ are not marked, it means that the peak is not significant. We see, for instance, that these peaks are all significant for the global data and for croplands. For rangelands and forested the peak corresponding to $j = 1050$ is not significant; for wildlands only the peaks corresponding to $j = 10, 20, 30, 40$ are significant; for settlemnts only the peaks for $j = 10, 20$ are significant and for villages, only the peaks for $j = 10, 20, 40$.

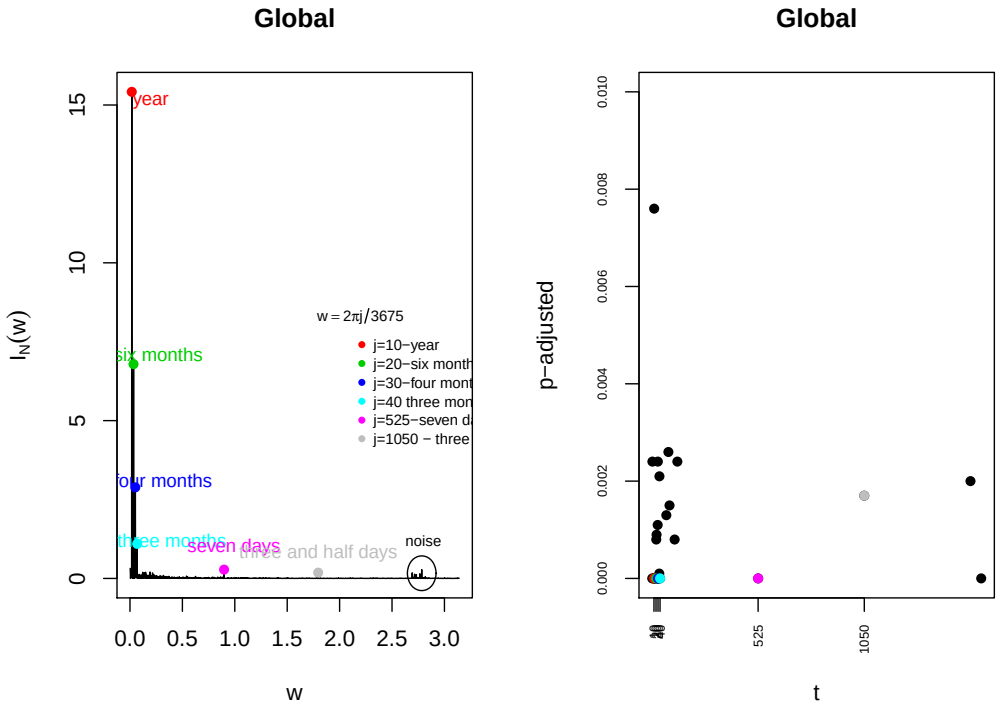

Figure 15: Periodogram and Whittle test for the log daily counts for croplands

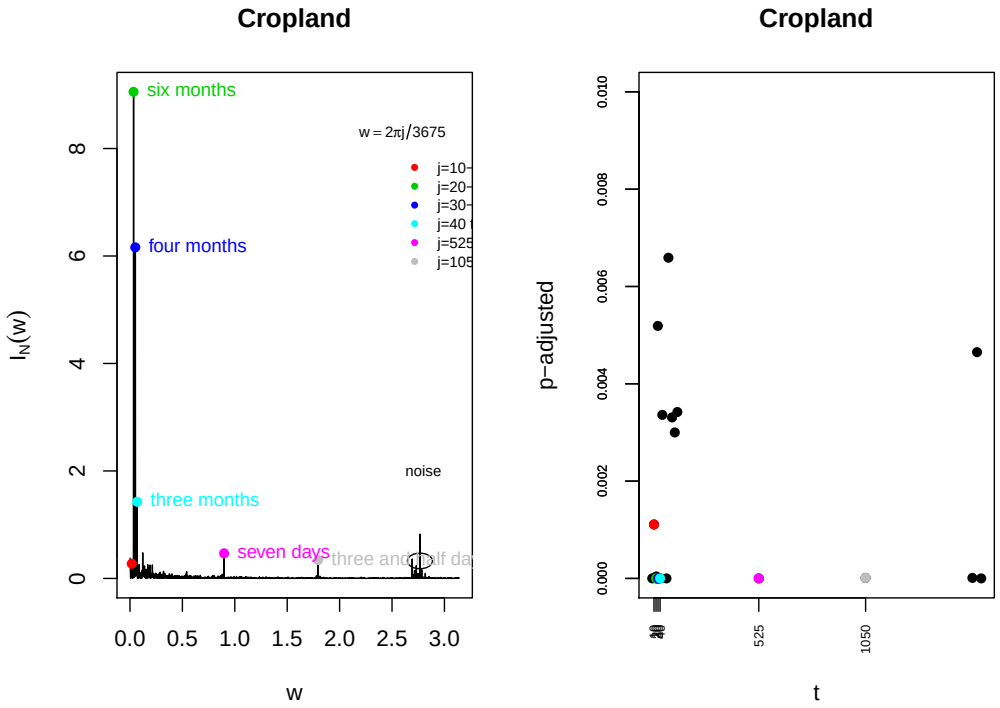

Figure 16: Periodogram and Whittle test for the log daily counts for croplands

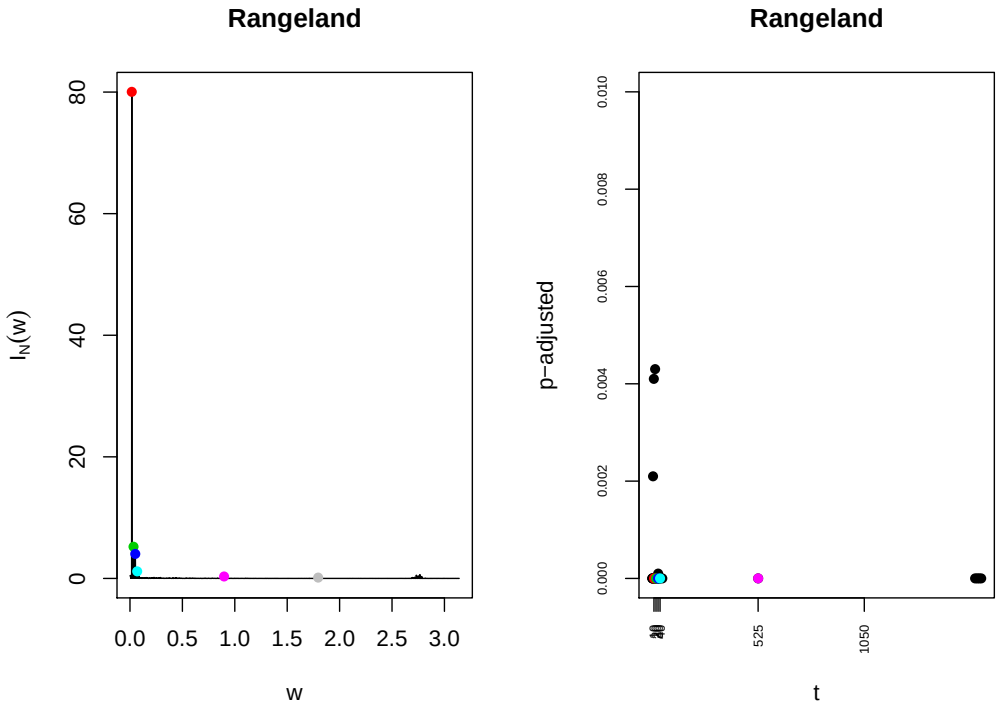

Figure 17: Periodogram and Whittle test for the log daily counts for rangelands

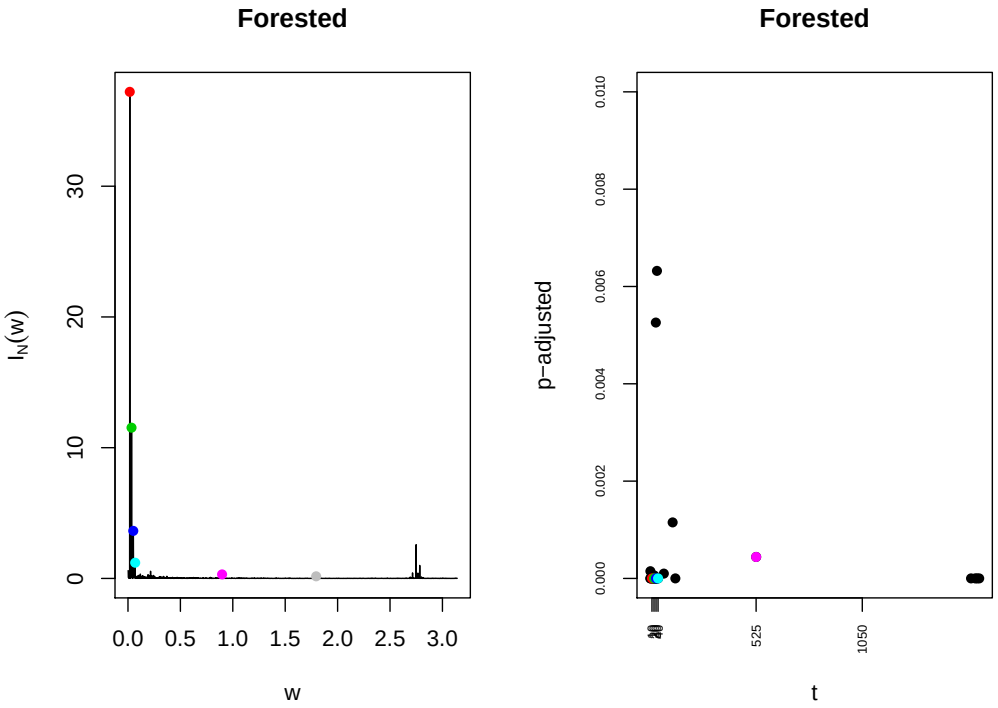

Figure 18: Periodogram and Whittle test for the log daily counts for forests

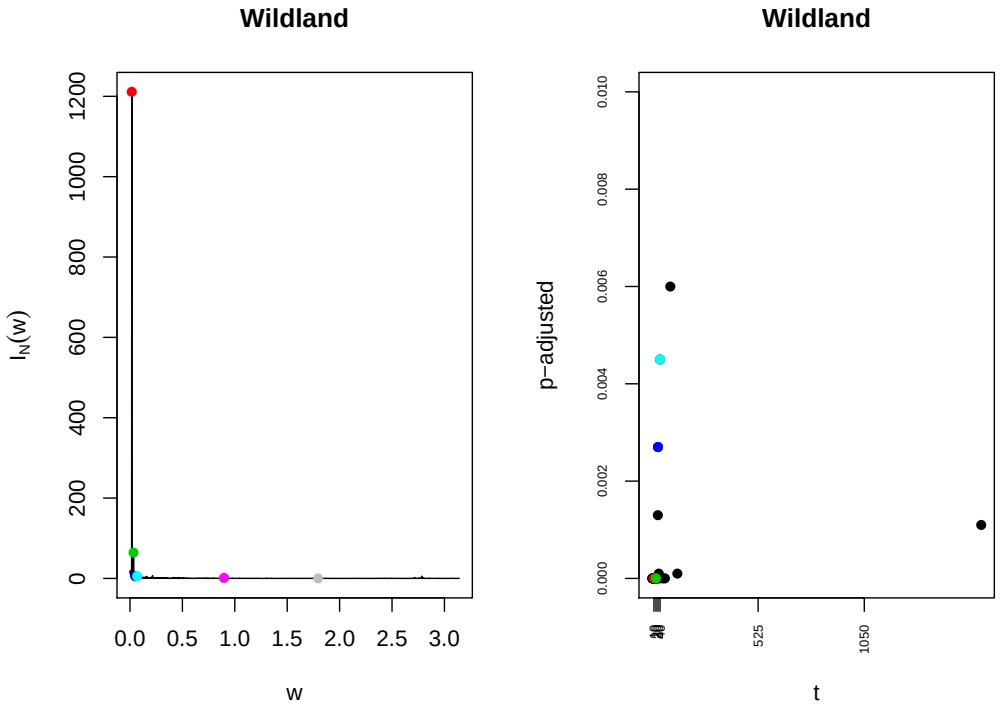

Figure 19: Periodogram and Whittle test for the log daily counts for wildlands

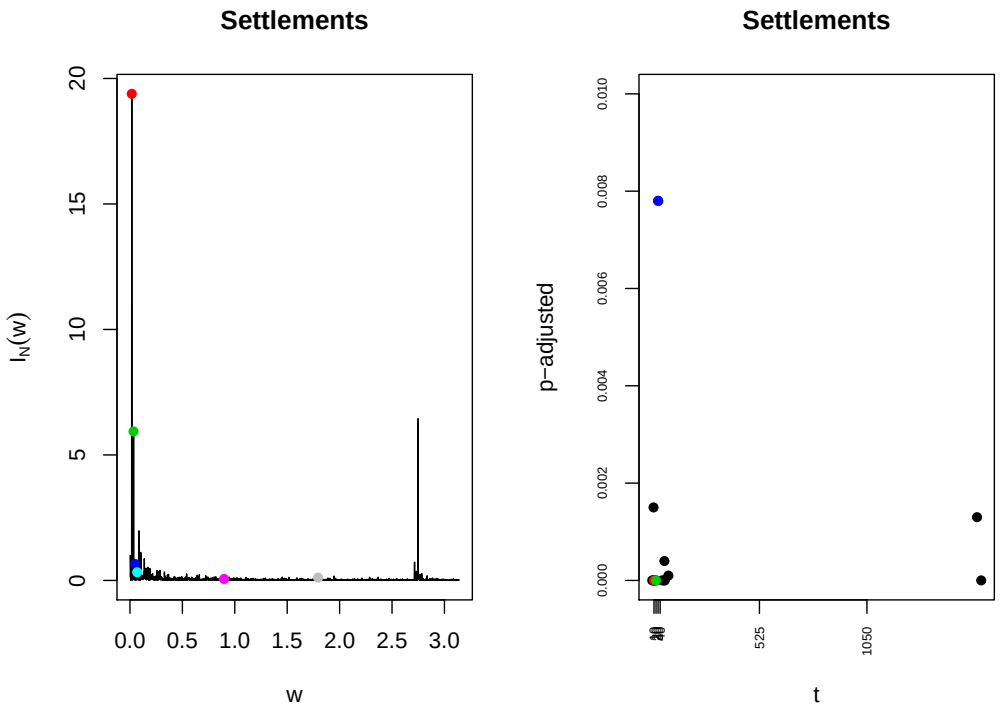

Figure 20: Periodogram and Whittle test for the log daily counts for settlements

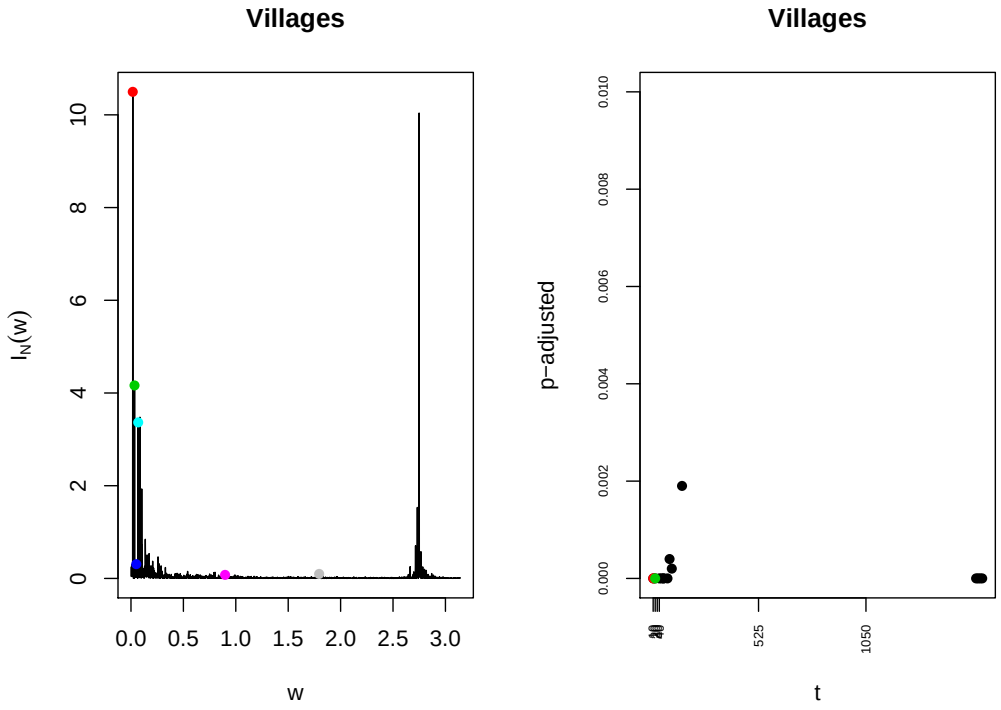

Figure 21: Periodogram and Whittle test for the log daily counts for villages

5.2 Fiited model for the global data

We considered the following mean function

$$M_t = a_0 + bt + S_t,$$

with

$$\begin{aligned} S_t &= c_1 \cos(2\pi t 10/n) + d_1 \sin(2\pi t 10/n) \\ &= c_2 \cos(2\pi t 20/n) + d_2 \sin(2\pi t 20/n) \\ &= c_3 \cos(2\pi t 30/n) + d_3 \sin(2\pi t 30/n) \\ &= c_4 \cos(2\pi t 40/n) + d_4 \sin(2\pi t 30/n) \\ &= c_5 \cos(2\pi t 525/n) + d_5 \sin(2\pi t 525/n) \\ &= c_6 \cos(2\pi t 1050/n) + d_6 \sin(2\pi t 1050/n) \end{aligned} \tag{3}$$

Table 7: Estimated coefficients (frequentist approach) of the mean function relative to the model (1) for the global data.

	Estimate	Std. Error	t-value
a_0	9.3408	0.0280	333.2810
b	-0.0000	0.0000	-1.1328
c_1	0.1865	0.0196	9.5152
d_1	0.2637	0.0197	13.3708
c_2	0.1799	0.0191	9.4055
d_2	0.1186	0.0192	6.1777
c_3	-0.0480	0.0184	-2.6057
d_3	0.1313	0.0185	7.1068
c_4	-0.0069	0.0176	-0.3944
d_4	-0.0865	0.0176	-4.9164
c_5	-0.0424	0.0050	-8.4597
d_5	0.0107	0.0050	2.1443
c_6	-0.0337	0.0029	-11.4430
d_6	-0.0105	0.0029	-3.5731

5.3 Mean functions

Here we display the plots of the estimated mean functions for the different anthromes, using the Bayesian approach.

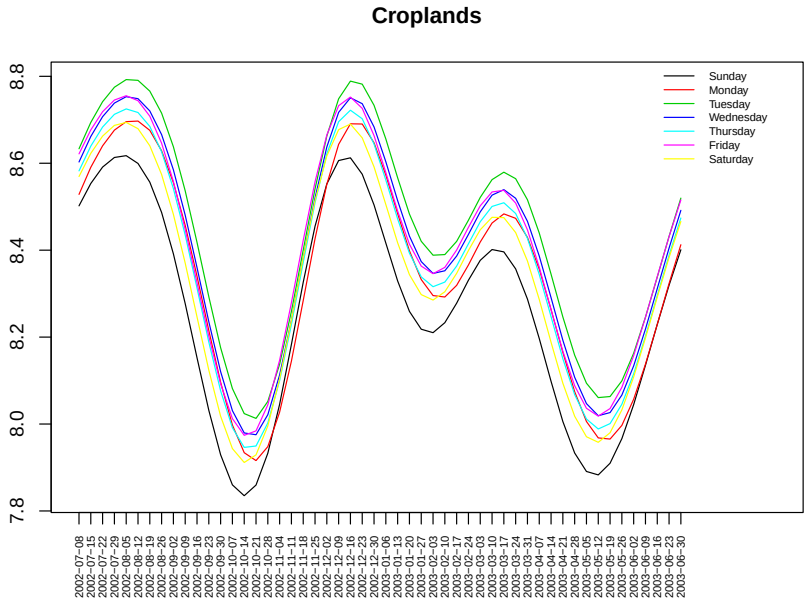

Figure 22: Estimated mean functions for croplands

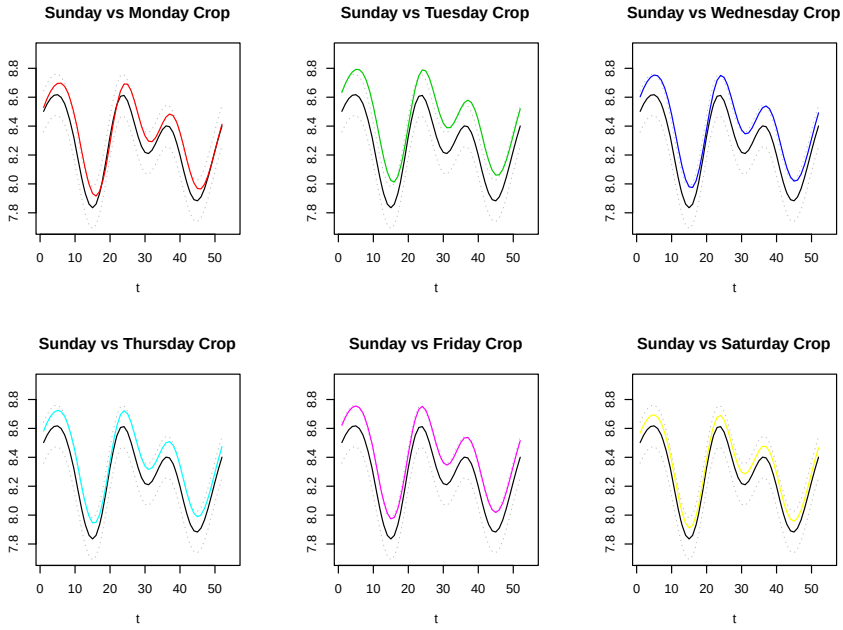

Figure 23: Croplands: comparing estimated mean functions for every day of the week with Sunday mean function

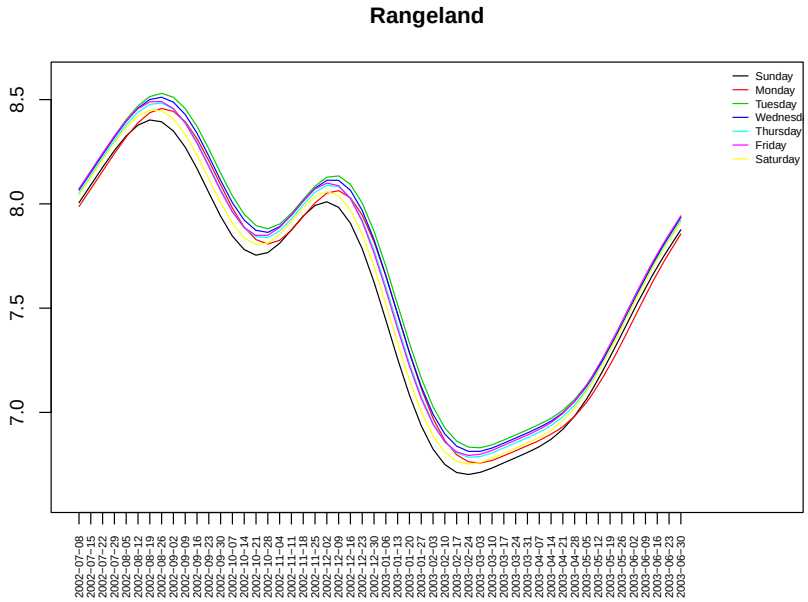

Figure 24: Estimated mean functions for rangelands

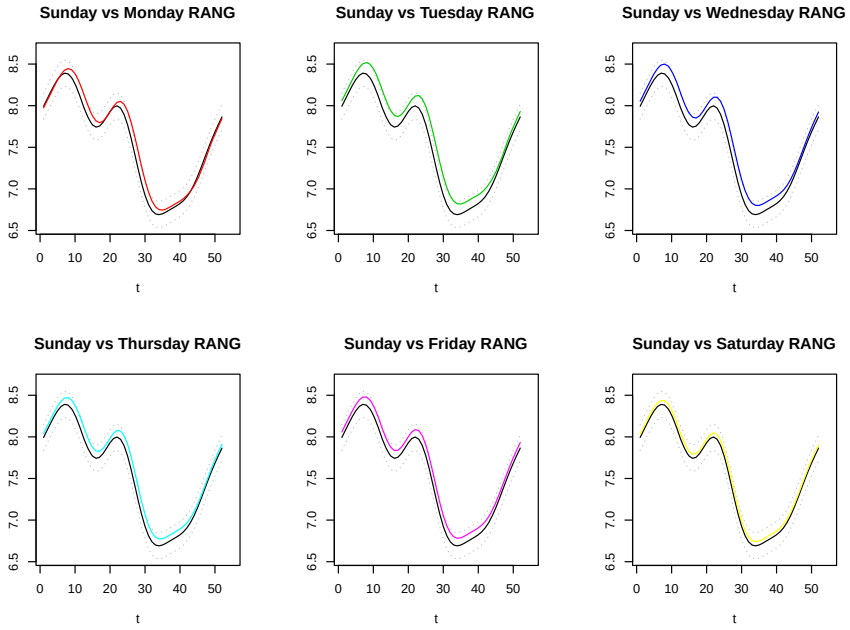

Figure 25: Rangelands: comparing estimated mean functions for every day of the week with Sunday mean function

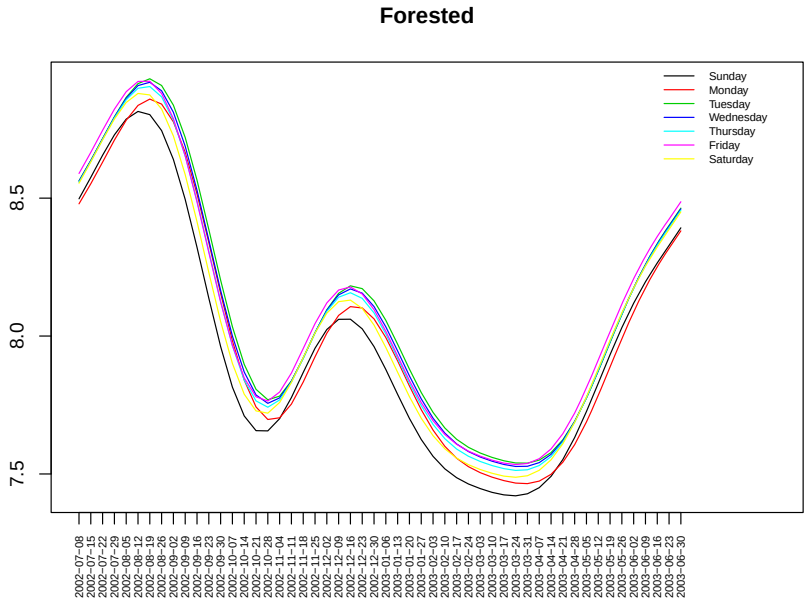

Figure 26: Estimated mean functions for forests

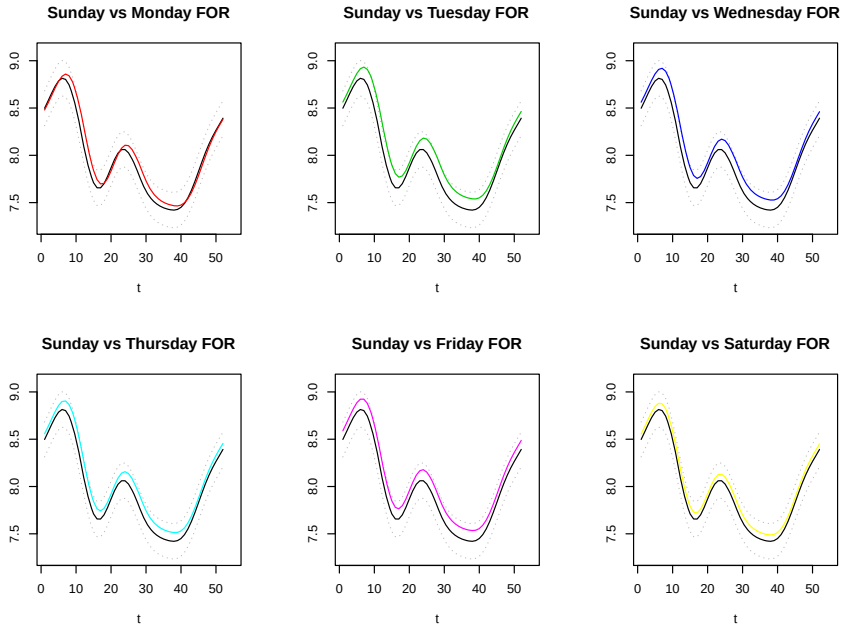

Figure 27: Forests: comparing estimated mean functions for every day of the week with Sunday mean function

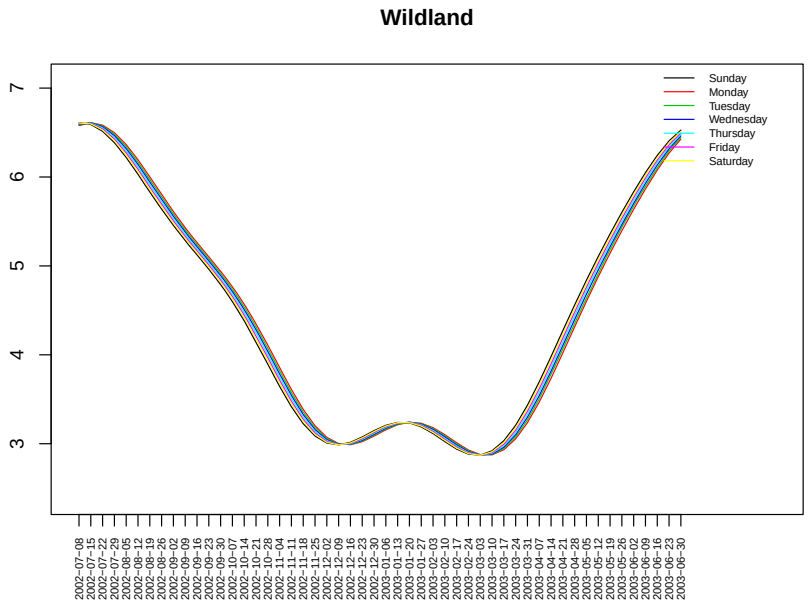

Figure 28: Estimated mean functions for wildlands

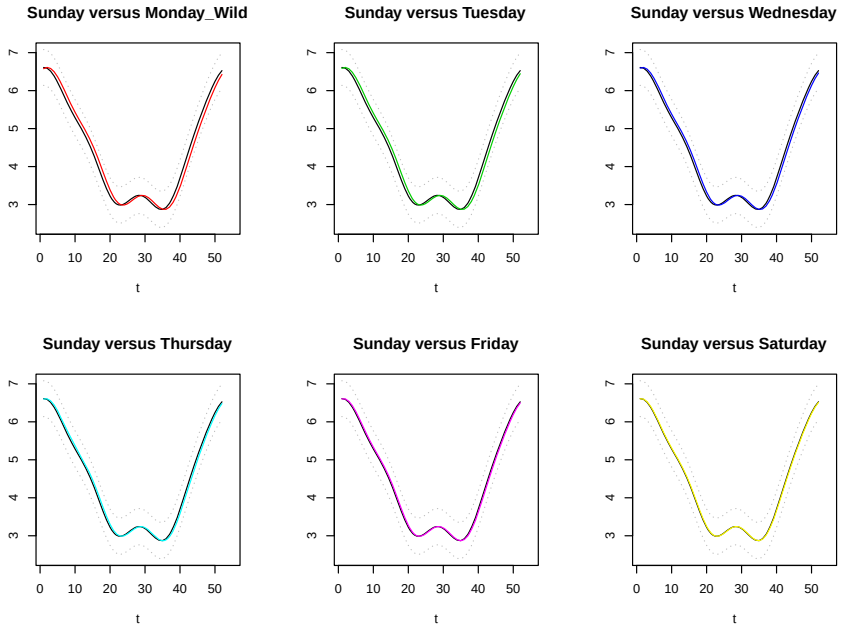

Figure 29: Wildlands: comparing estimated mean functions for every day of the week with Sunday mean function

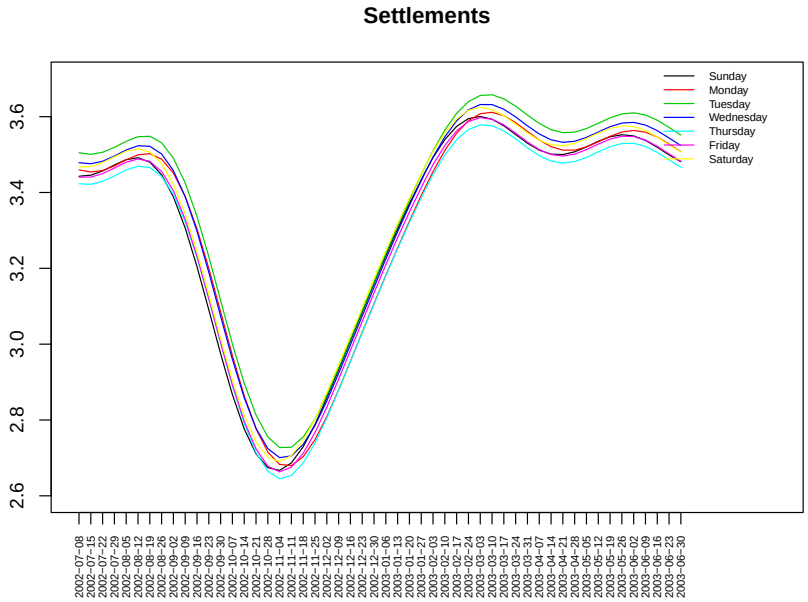

Figure 30: Estimated mean functions for settlements

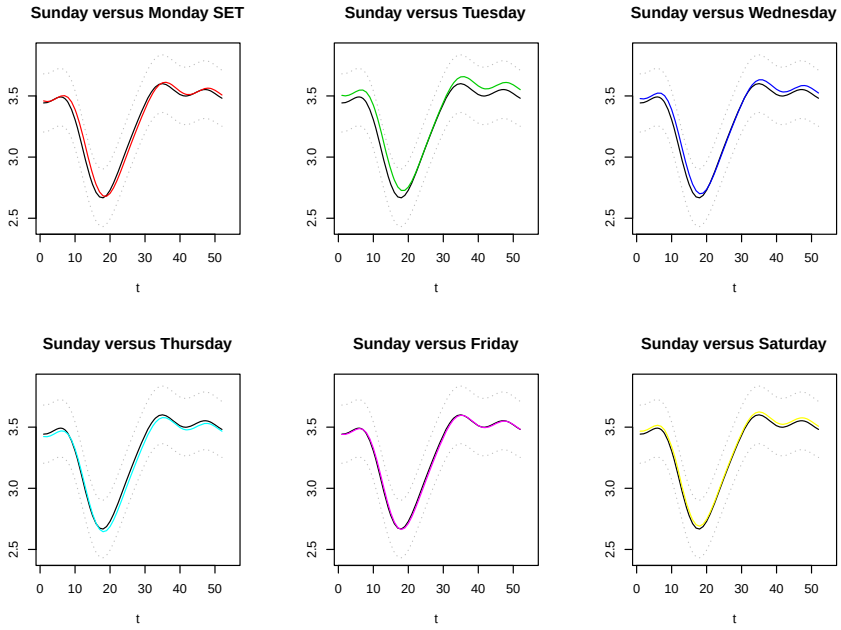

Figure 31: Settlements: comparing estimated mean functions for every day of the week with Sunday mean function

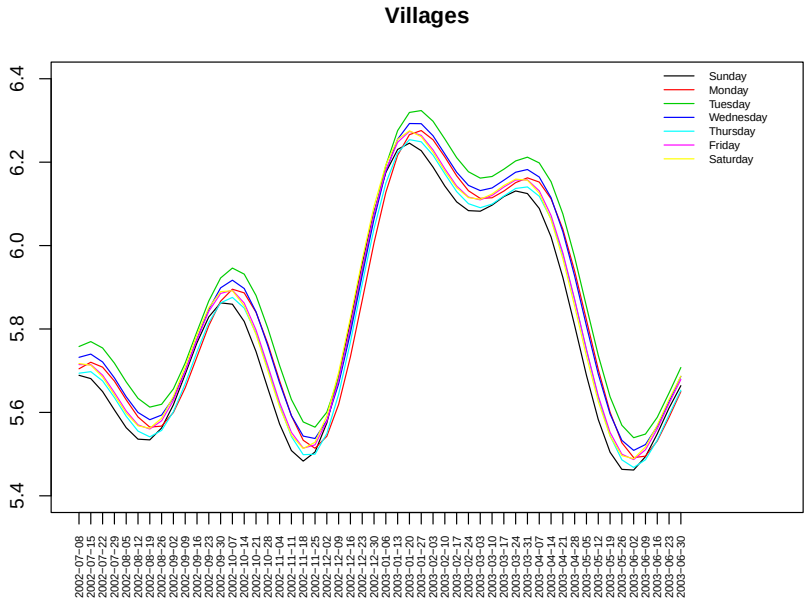

Figure 32: Estimated mean functions for villages

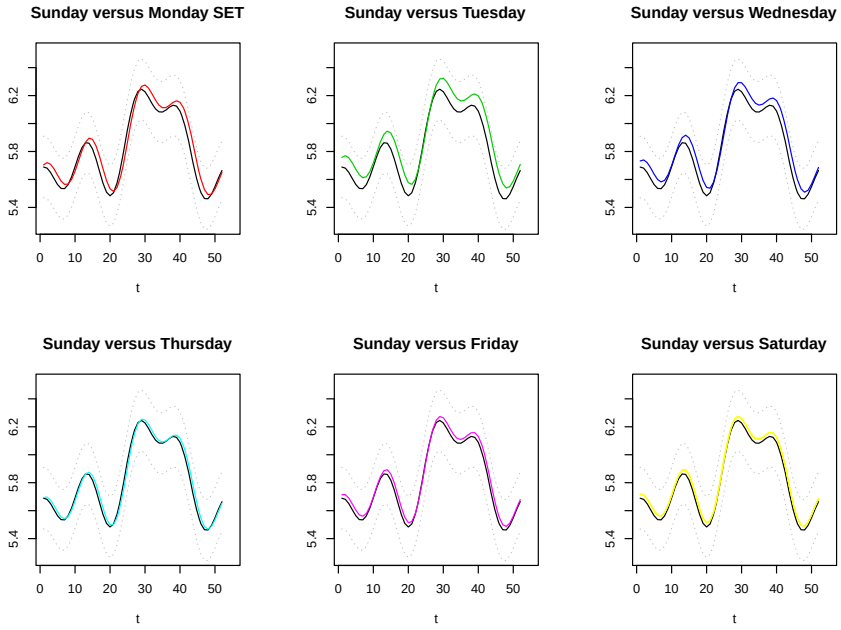

Figure 33: Villages: comparing estimated mean functions for every day of the week with Sunday mean function

5.4 Periods of interest

Here we refer to the periods of the year when the estimated mean function corresponding to each week day is above the 95% upper credible band for Sunday mean function.

5.4.1 Croplands

- Monday is above the upper bound of Sunday mean on

  "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2003-01-06", "2003-01-13", "2003-04-07", "2003-04-14", "2003-04-21"

- Tuesday is above the upper bound of Sunday mean on

  "2002-07-22", "2002-07-29", "2002-08-05", "2002-08-12", "2002-08-19", "2002-08-26", "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2002-10-07", "2002-10-14", "2002-10-21", "2002-12-16", "2002-12-23", "2002-12-30", "2003-01-06", "2003-01-13", "2003-01-20", "2003-01-27", "2003-02-03", "2003-02-10", "2003-03-03", "2003-03-10", "2003-03-17", "2003-03-24", "2003-03-31", "2003-04-07", "2003-04-14", "2003-04-21", "2003-04-28", "2003-05-05", "2003-05-12", "2003-05-19"

- Wednesday is above the upper bound of Sunday mean on

  "2002-08-12", "2002-08-19", "2002-08-26", "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2002-10-07", "2002-10-14", "2002-12-23", "2002-12-30", "2003-01-06", "2003-01-13", "2003-01-20", "2003-01-27", "2003-03-24", "2003-03-31", "2003-04-07", "2003-04-14", "2003-04-21", "2003-04-28", "2003-05-05"

- Thursday is above the upper bound of Sunday mean on

  "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2003-01-06", "2003-04-07", "2003-04-14", "2003-04-21"

- Friday mean is above the upper bound for the Sunday mean on

  "2002-08-12", "2002-08-19", "2002-08-26", "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2002-10-07", "2002-12-23", "2002-12-30", "2003-01-06", "2003-01-13", "2003-01-20", "2003-01-27", "2003-03-24", "2003-03-31", "2003-04-07", "2003-04-14", "2003-04-21", "2003-04-28", "2003-05-05"

- Saturday estimated mean is always below the upper bound of Sunday mean function

- Summary: weeks when there is at least one day of the week with an estimated mean above the upper bound of Sunday mean function

  "2002-07-22", "2002-07-29", "2002-08-05", "2002-08-12", "2002-08-19", "2002-08-26", "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2002-10-07", "2002-10-14", "2002-10-21", "2002-12-16",

"2002-12-23", "2002-12-30", "2003-01-06", "2003-01-13", "2003-01-20", "2003-01-27", "2003-02-03", "2003-02-10", "2003-03-03", "2003-03-10", "2003-03-17", "2003-03-24", "2003-03-31", "2003-04-07", "2003-04-14", "2003-04-21", "2003-04-28", "2003-05-05", "2003-05-12", "2003-05-19".

5.4.2 Rangelands

- Monday is above the upper bound of Sunday mean on

  "2002-12-23", "2002-12-30", "2003-01-06", "2003-01-13", "2003-01-20", "2003-01-27"

- Tuesday is above the upper bound of Sunday mean on

  "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2002-10-07", "2002-10-14", "2002-12-16", "2002-12-23", "2002-12-30", "2003-01-06", "2003-01-13", "2003-01-20", "2003-01-27", "2003-02-03", "2003-02-10"

- Wednesday is above the upper bound of Sunday mean on

  "2002-09-16", "2002-09-23", "2002-09-30", "2002-10-07", "2002-12-16", "2002-12-23", "2002-12-30", "2003-01-06", "2003-01-13", "2003-01-20", "2003-01-27", "2003-02-03"

- Thursday is above the upper bound of Sunday mean on

  "2003-01-06", "2003-01-13"

- Friday mean is always below the upper bound for the Sunday mean

- Saturday estimated mean function is always below the upper bound for the Sunday mean function

- Summary: weeks when there is at least one day of the week outside the upper bound of Sunday mean

  "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2002-10-07", "2002-10-14", "2002-12-16", "2002-12-23", "2002-12-30", "2003-01-06", "2003-01-13", "2003-01-20", "2003-01-27", "2003-02-03", "2003-02-10"

5.4.3 Forested

- Monday is above the upper bound of Sunday mean on

  "2002-09-16", "2002-09-23", "2002-09-30"

- Tuesday is above the upper bound of Sunday mean on

  "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2002-10-07", "2002-10-14"

- Wednesday is above the upper bound of Sunday mean on

  "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30"

- Thursday mean is always below the upper bound for the Sunday mean function

- Friday mean is always below the upper bound for the Sunday mean function

- Saturday mean is always below the upper bound for the Sunday mean function

- Summary: weeks when there is at least one day of the week outside the upper bound of Sunday mean

  "2002-09-02", "2002-09-09", "2002-09-16", "2002-09-23", "2002-09-30", "2002-10-07", "2002-10-14"