**Nonlinear Processes
in Geophysics**

External forcing of earthquake swarms at Alpine regions: example from a seismic meteorological network at Mt. Hochstaufen SE-Bavaria

V. Svejdar[1], **H. Küchenhoff**[1], **L. Fahrmeir**[1], **and J. Wassermann**[2]

[1]Department of Statistics, Ludwig-Maximilians Univ. Munich, Germany
[2]Department of Earth- and Environmental Sciences, Geophysical Observatory, Ludwig-Maximilians Univ. Munich, Germany

Abstract. In the last few years, it has been shown that above-average rainfall and the following diffusion of excess water into subsurface structures is able to trigger earthquake swarms in the uppermost brittle portion of the Earth's crust. However, there is still an ongoing debate on whether the crust already needs to be in a critical-to-failure state or whether it is sufficient that water is transported rapidly within channels and veins of karst or similar geological formations to the underlying, earthquake-generating layers. Also unknown is the role of other forcing mechanisms, possible co-variables and probably necessary tectonic loading in the triggering process of earthquakes. Because of these problems, we do not use an explicit physical model but instead analyze the meteorological and geophysical data via sophisticated statistical models. We are interested in the influence of a more complete set of possible forcing parameters, including the influence of synthetic earth tides, on the occurrence of earthquake swarms. In this context, regression models are the adequate tool, since the calculation of simple correlations can be confounded by the other variables. Since our outcome variable (the number of quakes) is a count, we use Poisson regression models that include the plausible assumption of a Poisson distribution for the counts. For this study, we use nearly continuous recordings of a seismic and meteorological network in the years 2002–2008 at Mt. Hochstaufen in SE-Bavaria. Our non-linear regression model reveals correlations between external forces and the triggering of earthquakes. In addition to the still dominant influence of rainfall, theoretical estimated tidal tilt show some weak influence on the swarm generation. However, the influence of the modeled trend functions shows that rain is by far not the most important forcing mechanism present in the data.

1 Introduction

That above-average rates of rainfall may trigger earthquakes had been speculated since the start of detailed seismic observations in alpine areas (e.g. Schmedes, 1979; Roth et al., 1992; Schwarzmann et al., 1996). Several independent studies carried out at Mt. Hochstaufen (SE-Bavaria, Germany), Central Swiss and Mt. Hood (Oregon, USA) showed that the triggering or forcing of earthquake swarm occurrence is real and might be described by a simple 1-D diffusion of rain into the subsurface structure (Roth et al., 1992; Saar and Manga, 2003; Kraft et al., 2006a; Husen et al., 2007). In this model, the effect of rain on a critical, i.e. close-to-failure crust is modeled by an increase of pore pressure, which in turn leads to a reduction of normal stress acting on pre-existing faults. Hainzl et al. (2006) combined the model of diffusivity with a model of rate-state friction (Dietrich, 1994) in order to quantify the effect of pore pressure changes on seismicity. They concluded that even small stress changes in the range of 100 Pa are sufficient for earthquake triggering if the crust is in a critical state. Husen et al. (2007) further concluded that the earthquakes close to Riemenstalden and Muothal region (Central Swiss) are located on prestressed fault systems and would have happened regardless at a future date. Above-average rainfall is seen as the final catalyst in earthquake triggering. Taking into account that the upper layer at both areas is dominated by karst, Miller (2008) argued that a pure diffusion process might not be the appropriate model to describe the water flow in the upper portion of this prominent geological situation. Consequently, he concluded that the excess pressure due to heavy rainfall is much higher in comparison with a pure diffusion model and therefore earthquakes are produced even if the stress condition of the crust is not close to critical. In all these studies, the question remains whether rainfall is the only acting external forcing mechanism and whether the crust needs to be in an already critical condition. Even when the simple 1-D diffusion equation

describes the time-depth dependency of the earthquake front in some cases fairly well, there might exist different models or details of the triggering and loading process hidden in the data, which describe the time-depth distribution of earthquakes at least equally well. Poisson models with the number of earthquakes have been applied successfully, see e.g. Guillas et al. (2010). These models enable us to find complex relationships between possible triggers and earthquake occurrence. In particular, lagged and nonlinear effects are part of our model.

2 Data

Following Kraft et al. (2006a, b) and Hainzl et al. (2006), we restrict our analysis to the Mt. Hochstaufen area, a 1770 m high mountain massif at the south-east Bavarian, northern alpine front (see Kraft et al., 2006b for a detailed geological description). This area has been known for its seismic activity since 1390 and has therefore been monitored starting in the 1970ies by the geophysical observatory in Fürstenfeldbruck with a changing number of seismometers. Since 1980, the swarm character of these earthquakes, which occur mainly in the summer months, has been investigated using digital mobile seismic equipment and one permanent analog ink recorder in Bad Reichenhall. From 2001 on, the area has been monitored by six digital short period seismic stations with permanent and continuous data transmission to the data center in Fürstenfeldbruck (Fig. 1). During summer, this permanent network is extended by additional mobile stations (two–six) in order to decrease the magnitude of completeness (M_c) and increase the location accuracy. Additionally, meteorological stations (rain gauge, air pressure, air temperature, humidity) have been co-located at three sites since mid-2004 (RJOB, RNON, RMOA, see Fig. 1).

After the construction of the permanent surveillance network, seven distinct earthquake swarms occurred at Mt. Hochstaufen (March and August 2002; May, July, August 2005; September 2007, April 2008). The maximum magnitude was reached in 2007, when an earthquake $M_l = 3.4$ happened and was felt in a wide area. The majority of the earthquakes, however, are small in magnitude ($<1.5\ M_l$) and thus not felt. The largest earthquake swarm by number of events was recorded in August 2002 after a period of extraordinary heavy rainfall with more than 600 events/month. None of these earthquakes were strong enough to be felt (Kraft et al., 2006a). As we will focus on the daily or hourly event rate in the period of 2002–2008, we must first estimate the magnitude of completeness (M_c) for the entire, heterogenous catalogue. We will follow the approach by Woessner and Wiemer (2008) and first calculate M_c by the method of the maximum curvature. Using this value as a first rough approximation, we than apply the "goodness of fit test" method (Woessner and Wiemer, 2008) in an interval centered at this first estimate. In order to test the stability of the M_c value and because the swarm of 2002 was observed with the densest
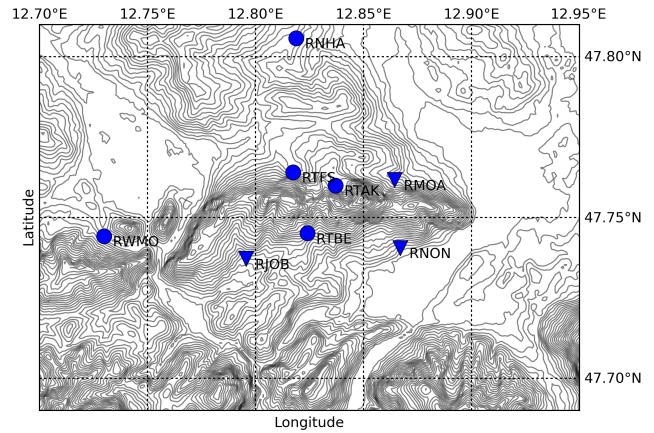

Fig. 1. Seismic and meteorological network installed at Mt. Hochstaufen. The stations shown were recording during most of the time period analyzed in this paper. The circles represent seismic (3C 1 Hz seismic sensors) stations, the triangles mark seismic and additionally installed meteorological stations (rain gauge, air temperature, ambient pressure, humidity).

seismic network deployed so far, we exclude this data in a second estimation of M_c from the catalogue. Both catalogues result in a magnitude of completeness of $M_c = 0.2$. Using this M_c estimate, all earthquakes smaller than this threshold are removed from further analysis.

The error in the hypocenter determination also changed several times during the observation period. At its best in 2002, it ranges roughly 200–300 m in horizontal direction and approximately 500 m in the vertical axes for larger and smaller events, respectively (Kraft et al., 2006b). For the period 2003–2008 and $M_c = 0.2$, the error in horizontal direction is around 400 m while the error in the vertical axis is about 800 m. As the terrain shows strong 3-D topography, shallow hypocenters and small epicentral distances to the network, a pseudo 3-D velocity model is used for the hypocenter determination (Kraft et al., 2006a). While the 3-D model includes the topography, it is using a homogeneous velocity-depth function, which still give rise to questions on errors of estimated locations. Figure 2 gives the hypocenters of all earthquakes between 2002–2008, which were recorded with more than three stations and both P and S phases, respectively.

It becomes apparent that the main seismicity in the area of Bad Reichenhall is connected to the eastern section of the Hochstaufen massif in a depth range between −4 km below sea level up to the surface. This earthquake data (location, source time and magnitude) form, together with rain gauge and air temperature measurements as well as additional theoretical earth tide calculations, the input data for the following statistical time series modeling. For the detailed statistical analysis, we divide the data set into daily (number of earthquakes per day; cumulative rain per day) and hourly sampled time traces.
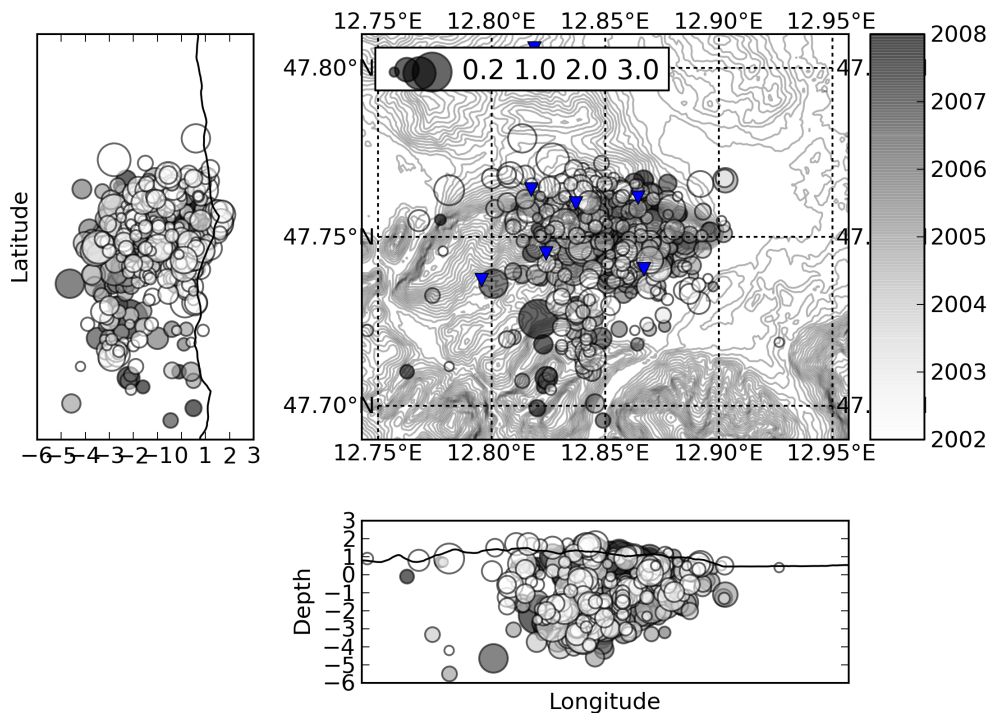
Fig. 2. Hypocenters of seismic events in the time range 2002–2008 and $M_l \geq 0.2$. The swarms are represented on a grayscale: white (2002), light-grey (2005), grey (2007) and black (2008) circles. The size of the circles is proportional to local magnitude estimations.

3 Modeling

As the first step of our analysis, we try to reproduce the time-depth behavior of seismicity following above-average rain fall events in 2003–2008 using the approach described in Kraft et al. (2006a) for the swarm data of 2002. In doing so, we first estimate the time of occurrence of exceptional rain events and check whether there was a following seismic swarm. Already the selection of a simple "above-average" rain threshold produces a large uncertainty as it is sometimes impossible to define a sharp onset of "over-average" rainfall. For the analysis presented here, we chose the threshold of $65 \, \mathrm{mm} \, \mathrm{day}^{-1}$ for practical rather than proper physical reasons. Using this threshold condition, we are able to identify seven periods since 2001 (March 2002, August 2002, May 2005, July 2005, August 2005, September 2007 and April 2008; see Fig. 3), which were also selected by a human interpreter when identifying seismic swarms. Other promising variables may exist in the volume influx or the cumulative rain amount. Without detailed knowledge of the influx/outflux budget of the hydrological system at Mt. Hochstaufen, however, it seems not feasible to use them without adding more complexity to the model. Nonetheless, the definition of swarm is also sometimes difficult and biased by the interpreter. In Fig. 3, the threshold of $65 \, \mathrm{mm} \, \mathrm{day}^{-1}$ rain was apparently exceeded in 2004 and 2006 without being followed by a seismic swarm. This

might already point to the existence of a loading mechanism that transforms the crust from stable to critical conditions. Nevertheless, using the day exceeding the threshold as origin time, we apply the simple 1-D diffusion equation introduced by Shapiro et al. (1997) for modeling the seismic pattern:

$$z = \sqrt{(4\pi D t)} \qquad (1)$$

with z representing the depth of the diffusive water front, D diffusivity and t time from onset of hydraulic pressure disturbance at the surface. In Fig. 4, the time vs. depth plot for all available swarm data with a magnitude exceeding $M_c = 0.2$ are shown together with their corresponding local magnitude estimate. In addition, the theoretical time depth relationship of the hydraulic pressure front according to Eq. (1) is also shown whenever the chosen rain rate threshold of $65 \, \mathrm{mm} \, \mathrm{day}^{-1}$ was exceeded. To keep the analysis simple, we use the diffusivity estimates by Kraft et al. (2006a) for the earthquake swarms of 2002. Overall, this oversimplified 1-D diffusivity model seems not to give reasonable fits for most of the swarm data. Looking more closely at Fig. 4, several problems become apparent when applying the model to the data. First of all, it is not clear what "above-average" rainfall quantitatively means. For example, if the rain is distributed over several days, how should the threshold and the date of exceedance be defined? What is the "effective" location or depth of the pressure disturbance? In
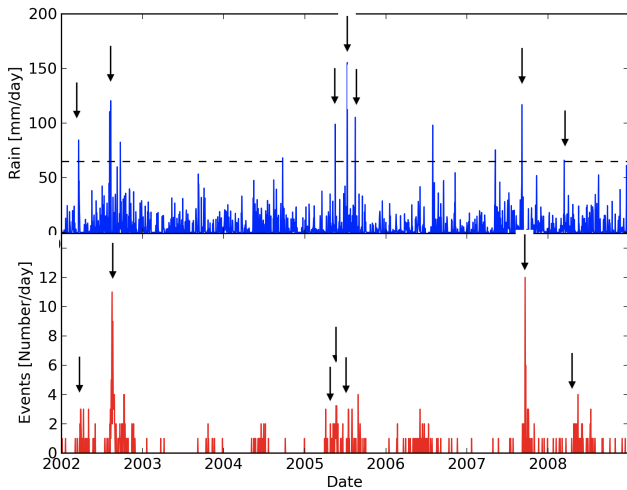
Fig. 3. Rain rate and earthquake rate plots for the period of analysis. The threshold of $65\,\mathrm{mm\,day^{-1}}$ is also shown as a dashed line. The arrows mark the reported occurrence of earthquake swarms.

Fig. 4, we have chosen an "average" free surface of the area at 500 m a.s.l. Clearly a 3-D model of the hydraulic system of the Mt. Hochstaufen massif is needed but with the present knowledge of the hydrological system this is not feasible. In addition, the earthquake data do not show any backfront of seismicity, which is normally observed at hydraulic fracturing experiments when the fluid influx has ended (Shapiro and Dinske, 2007). All these problems lead us not to use even more complicated but less justified physical models but to instead apply a suitable statistical analysis in order to detect and compute the influence of possible external forcing mechanisms. We additionally expand our observed external parameters by including strain or tilt induced by tides and air temperature into the analysis. The statistical multi-component analysis will thus lead to a better understanding of the observed seismicity pattern and will also guide future measurements in this area.

4 Statistical modeling by Poisson regression based on daily values

In our study we want to apply an alternative model in order to prove the possible influence of rainfall and other external forcing parameters on the occurrence of shallow earthquakes. For a first data set, using daily values of the different input variables, the response variable is defined as

$$Y_{jt} = \text{Number of quakes in depth category } j \text{ at day } t.$$

The number of quakes are observed within the four depth categories shown in Table 1. In this case, the depth is given relative to Earth's surface. While the chosen interval bounds are set without direct geological evidence, they reflect the

Table 1. Allocated depth categories of the quakes.

Depth Category j	$j=1$	$j=2$	$j=3$	$j=4$
Depth in km	$]0,-1]$	$]-1,-2]$	$]-2,-4]$	$]-4,-\infty[$
Number of quakes	162	56	144	14

uncertainty in estimating the source depth by this sparse network (Kraft et al., 2006b). The chosen interval bounds were varied several times with different interval bounds and different numbers of depth categories (from 3 to 6), which did not affect the direction of our results in a massive way.

Since the influence of rain possibly lags behind, we define

$x_{t-i} = i$-days lagged amount of rain in mm at day t

for $i = 0, ..., 20$. The number of lagged days included in the model was chosen arbitrarily. Longer lags till $i = 30$ were examined in further studies without noteworthy improvement of the model. Furthermore, the influence of another external covariate $b_t =$ temperature at day t as well as the influence of tide induced variables z_{lt}, defining gravity, NS and EW tilt, NS and EW strain, are examined. Here, we have to emphasize that the tide variables are computed using a purely theoretical relationship by using the SPOTL program package (Agnew, 1996). Real in situ measured tilt and strain might strongly deviate from these values due to the present local geology complexities.

As the response variable is a count, an appropriate modeling is given by the Poisson regression. It is defined by

$$Y_{jt}|x_{t-i},b_t,z_{lt} \sim \text{Poisson}(\exp(\eta_{jt}))$$
$$E(Y_{jt}|x_{t-i},b_t,z_{lt}) = \exp(\eta_{jt})$$

where η_{jt} is the predictor containing the explanatory variables in depth category j at day t.

As already mentioned in the abstract, the use of regression models has advantages over the analysis of simple correlations: a correlation can be confounded by other variables. In a regression model, we are able to analyze correlations of interest of two variables adjusted for the interfering effect of these confounders. We furthermore chose Poisson regression in order to avoid negative expectations. The predictor is linked with the exponential function, which ensures positive expectations (which is adequate for count data) and is the standard option in the context of statistical modeling. This model implies a multiplicative linkage of the covariates. The fact that strong rainfall can have a larger effect, if the general expectation of earthquakes is higher, seems to be reasonable and fits into the concept of criticality of the monitored crustal segment.

Except the instant influence of the rain x_t, we also take the $i = 1, ..., 20$ lagged rain days into the predictor. The predictor in depth category j is given by
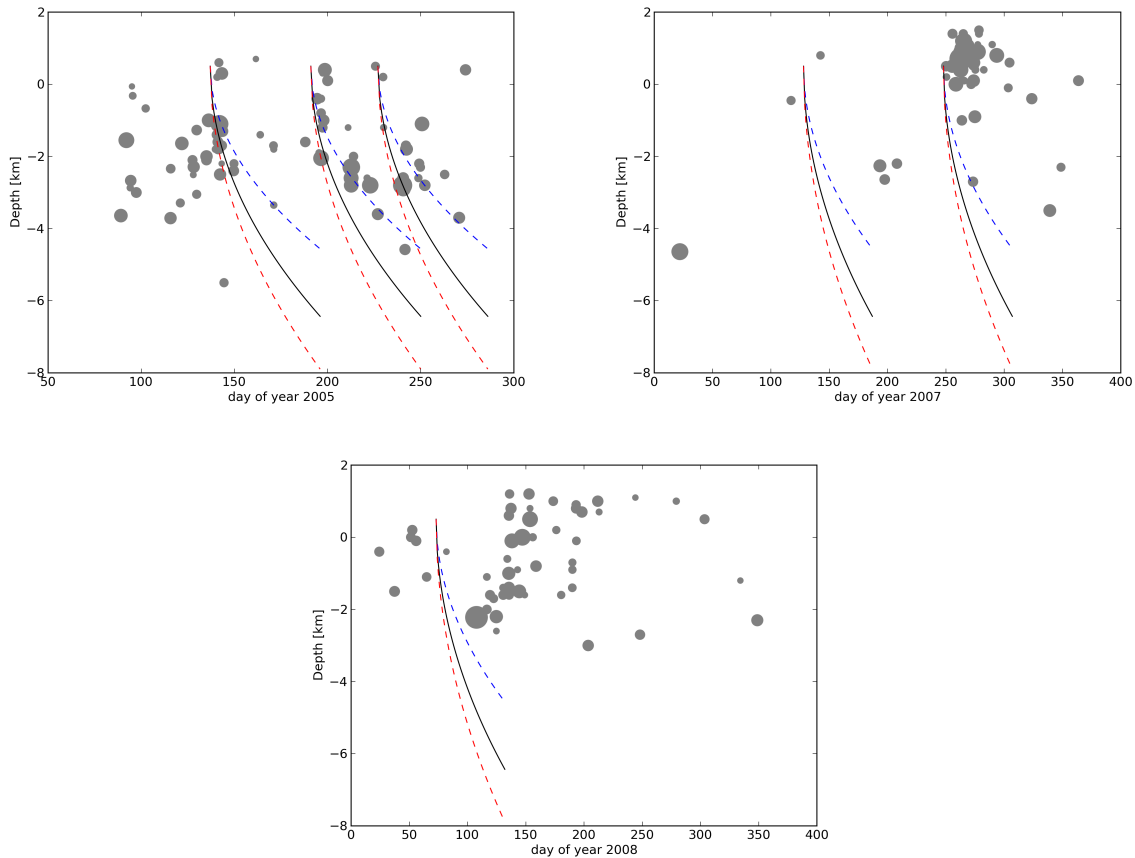
Fig. 4. Time vs. depth relationship of swarm earthquakes in the years 2005, 2007 and 2008. The origin time is set to the corresponding day when the rain rate exceeded the threshold of 65 mm day^{-1}. As reference depth an average value of 500 m a.s.l. was chosen. For comparison, three different diffusivity values are used to show the fit of Eq. (1) ($D = 0.4$; 1.1 – dashed; $D = 0.75$ – solid). The diffusivity values are taken from Kraft et al. (2006a). The size of the circles is proportional to local magnitude estimations and are the same as in Fig. 2.

$$\eta_{jt} = \sum_{i=0}^{20} x_{t-i} \cdot \beta_{ij} + \zeta_j + \kappa \cdot b_t + \sum_{l=1}^{L} z_{lt} \cdot \lambda_l + f(t) \qquad (2)$$

where the β_{ij} are the coefficients for rain lag i in depth category j and ζ_j define intercepts for depth category j. κ is the coefficient describing the instant influence of the temperature. In case of the hourly computed five tide induced variables (gravity; NS and EW tilt; NS and EW strain), the daily maximum of the absolute values was used for each of the five tide variables. In order to evaluate which of the five tide variables should be taken into the model, selection methods based on Akaike's information criterion (AIC, see e.g. Akaike, 1973) were used. Via backward and forward selection (see e.g. Hastie et al., 2009), relevant tide variables were preselected and then put into the model. λ_l describe the parameters of the tide effects. In order to cope with detected correlations of the deviance residuals in the daily data set, generalized estimating equations (Fahrmeir and Tutz, 2001) were calculated with the statistical program R (http://cran.r-project.org/, package gee). Second degree B-Splines (Fahrmeir and Tutz, 2001) with six inner knots

model the trend $f(t)$ in the data in a flexible nonlinear way. We thereby adjust for the effect of not observable covariates, which leads to increased trust in the detected significances in our modeling.

Due to interpretation and collinearity problems of the 84 rain lag depth interaction parameters β_{ij}, it seems to be reasonable to make some constraints on the interaction parameters. We make the following assumptions:

1. within a fixed depth category the effect of adjacent rain lags should be similar,

2. for a fixed rain lag the effect of adjacent depth categories should be similar.

This can be realized by the polynomial distributed lags (PDL) approach (Schwartz, 2000) where it is assumed that the effect of the rain with lag i in depth category j can be described as a two-dimensional low order polynomial. A quadratic polynomial is given by

Table 2. Lags of prime importance on the number of quakes (Poisson regression based on daily values.)

Depth Category j	$j = 1$	$j = 2$	$j = 3$	$j = 4$
Depth in km	]0, −1]	]−1, −2]	]−2, −4]	]−4, −∞[
Number of quakes	162	56	144	14
$i_{\max(j)}$	Lag13	Lag13	Lag13	Lag13
$\hat{\beta}_{i_{\max(j)}}$	0.014	0.010	0.007	0.005

$$\beta_{ij} = \alpha + \delta_1 \cdot i + \gamma_1 \cdot j + \delta_2 \cdot i^2 + \gamma_2 \cdot j^2 + \nu \cdot i \cdot j \qquad (3)$$

for $i = 0, ..., 20$ and $j = 1, ..., 4$.

Due to the linearity in the parameters $\alpha, \delta_1, \gamma_1, \delta_2, \gamma_2, \nu$, the part describing the rain-depth interaction in the predictor can be written as:

$$\sum_i x_{t-i} \cdot \beta_{ij} = \quad \alpha \cdot \underbrace{\sum_i x_{t-i}}_{w_{\alpha t}} + \delta_1 \cdot \underbrace{\sum_i i \cdot x_{t-i}}_{w_{\delta_1 t}} +$$

$$+ \gamma_1 \cdot j \cdot \underbrace{\sum_i x_{t-i}}_{w_{\gamma_1 t}} + \delta_2 \cdot \underbrace{\sum_i i^2 \cdot x_{t-i}}_{w_{\delta_2 t}} +$$

$$+ \gamma_2 \cdot j^2 \cdot \underbrace{\sum_i x_{t-i}}_{w_{\gamma_2 t}} + \nu \cdot j \cdot \underbrace{\sum_i i \cdot x_{t-i}}_{w_{\nu t}}.$$

After construction of the new design variables $\boldsymbol{w}_\alpha$, ..., $\boldsymbol{w}_\nu$ a known generalized linear model (Poisson regression) can be calculated and the estimates $\hat{\beta}_{ij}$ can be determined via Eq. (3) out of the estimates $\hat{\alpha}, ..., \hat{\nu}$. With

$$\boldsymbol{u}_{ij} = (1 \quad i \quad j \quad i^2 \quad j^2 \quad i \cdot j)' \quad \text{and}$$

$$\mathbf{V} = \text{estimated covariance matrix of } \boldsymbol{\theta} = (\alpha, ..., \nu)'$$

the standard error (se) of $\hat{\beta}_{ij}$ can be calculated as

$$\text{se}(\hat{\beta}_{ij}) = \sqrt{\boldsymbol{u}'_{ij} \mathbf{V} \boldsymbol{u}_{ij}}.$$

Furthermore, a statistical test for the possible rain effect can be performed.

5 Results of the Poisson regression based on daily values

In Fig. 5, the number of quakes on each of the $n = 2100$ days for each depth category is shown. As expected, the least quakes are counted in the deepest category 4.

The amount of rain in millimeters as well as temperature for the $n = 2100$ days can be seen in Fig. 6. The temperature is only available in the second part of the data set. Exemplary Tilt North South can be seen in the lowermost part of Fig. 6.

Table 3. External coefficients besides rain-depth coefficients in the Poisson regression based on daily values (Backward Selection). Temperature was not available at the beginning of the observation.

	Estimate	P-Value
Temperature	0.079	0.0195
Tide induced tilt North South	0.028	0.2396
Tide induced strain North South	−0.099	0.2520
Tide induced strain East West	0.047	0.5536

In Fig. 7, the course of the estimated distributed lag coefficients $\hat{\beta}_{ij}$ can be seen for every depth category. In every depth category, the maximum effect of rain on the number of quakes is on lag 13.

The coefficient plot also shows a decrease of the rain effect with increasing depth. The larger the distance of a depth class from the earth's surface (depth category 4), the smaller and less significant the respective coefficients. The strength of the maximum effect in each depth class decreases from 0.014 (depth category 1) to 0.005 (depth category 4), see Table 2. This corresponds to an increase of the expected number of earthquakes of the factor $\exp(10 \times 0.014) = 1.15$ and $\exp(10 \times 0.005) = 1.05$, respectively per 10 mm additional rain.

5.1 Accuracy of the maximum

We address the question of the accuracy of the determination of the maximum in each depth category. Focusing on the polynomial form of the rain-depth coefficients

$$\hat{\beta}_{ij} = \hat{\alpha} + \hat{\delta}_1 \cdot i + \hat{\gamma}_1 \cdot j + \hat{\delta}_2 \cdot i^2 + \hat{\gamma}_2 \cdot j^2 + \hat{\nu} \cdot i \cdot j$$

the maximal lag in every depth category j can be obtained after differentiating with respect to i and equating to zero via:

$$i_{\max(j)} = \frac{-\hat{\delta}_1 - \hat{\nu} \cdot j}{2 \cdot \hat{\delta}_2}.$$

The estimated variance of the maximum in each depth category can then be calculated as:

$$\widehat{\text{Var}(i_{\max(j)})} = \qquad\qquad\qquad (4)$$

$$\left(-\frac{1}{2\hat{\delta}_2}, \ \frac{\hat{\delta}_1 + \hat{\nu} \cdot j}{2\hat{\delta}_2^2}, \ -\frac{j}{2\hat{\delta}_2} \right) \cdot \widehat{\text{Var}(\hat{\delta}_1, \hat{\delta}_2, \hat{\nu})} \cdot \begin{pmatrix} -\frac{1}{2\hat{\delta}_2} \\ \frac{\hat{\delta}_1 + \hat{\nu} \cdot j}{2\hat{\delta}_2^2} \\ -\frac{j}{2\hat{\delta}_2} \end{pmatrix}$$

For the Poisson model based on daily values, the estimated standard deviations (square root of Eq. 4) of the maximum in each depth categories can be seen in Table 4. The standard deviations of the maxima alternate poorly for the four depth categories and the maxima can be assumed to be specified relatively exact with standard deviations of about 0.7.
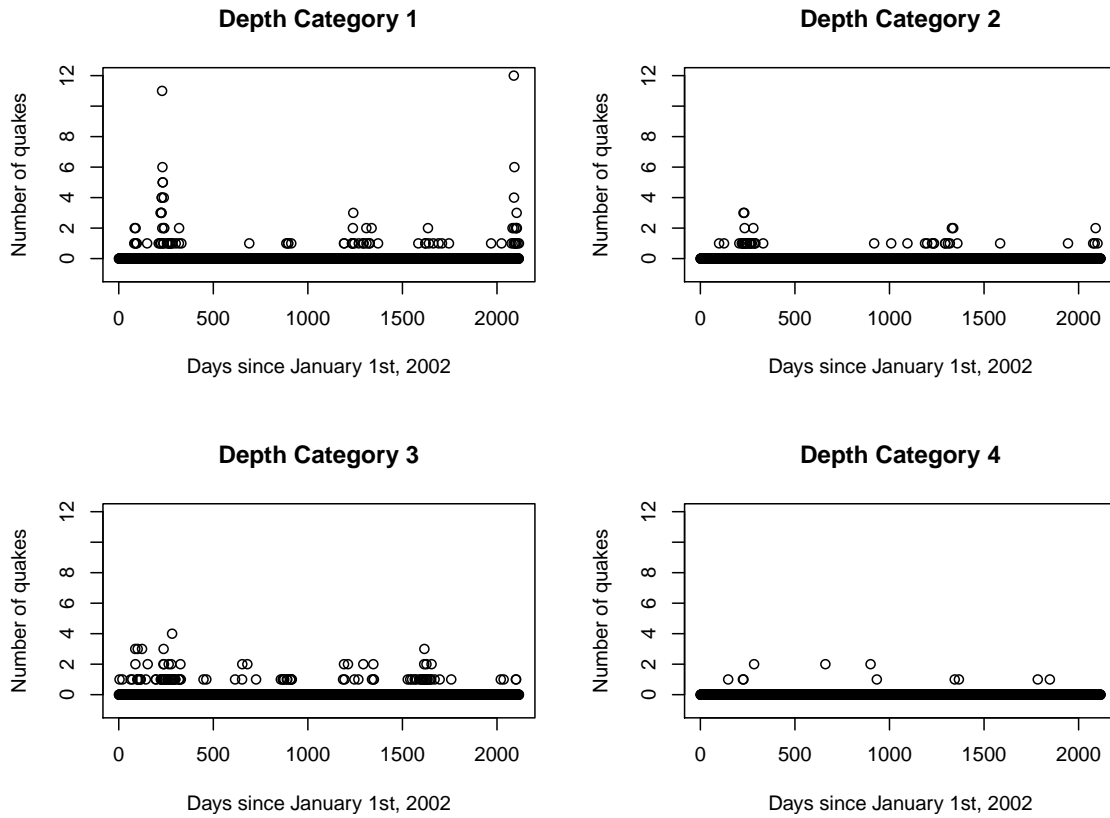
Fig. 5. Daily counts of quakes in the Hochstaufen area in the sample period of 2100 days by depth categories.

As a sensitivity analysis we fitted our model with alternative depth categories. We tried less (three) and more (up to six) depth categories. When testing these alternative depth categories, the maximum lag effect can still be found at lag 13 and lag 14, respectively. The results for the other parameters remain similar.

5.2 Further covariates and prediction

The influence of tides and temperature can be found in Table 3. Three variables were preselected by a backward selection starting with a model with all possible covariates using the AIC criterion. Nevertheless, none of the tide components shows statistical significance (statistical significance ≡ p-Value < 0.05). The influence of the temperature is significant with the coefficient $\hat{\kappa} = 0.079$. This is surprising as the temperature effect is modeled as acting instantly and not with a possible time delay. This delayed temperature effect can be expected when a thermo-elastic strain component is used in the model (Ben-Zion and Leary, 1986). The estimated component of the trend is shown in Fig. 8. The trend contains the information not explained by the available covariates. Figure 9 shows a comparison of the observed number of quakes in each depth category j (black dots) with the predicted number of quakes (red dots) in each depth category.

Table 4. Accuracy of the maximum for the Poisson regression based on daily values.

Depth Category j	$j = 1$	$j = 2$	$j = 3$	$j = 4$
$\dfrac{i_{\max(j)}}{\sqrt{\widehat{\mathrm{Var}(i_{\max(j)})}}}$	Lag 13 0.680	Lag 13 0.661	Lag 13 0.653	Lag 13 0.655

The prediction was done by a 2-fold cross validation (Hastie et al., 2009). One can see that the prediction is promising in depth categories 1 and 3, the categories with large numbers of quakes (see Table 1). The days where the number of quakes increase can be detected in depth category 1 and 3, whereat a certain underprediction can be seen. In depth categories 2 and 4, where we count only 56 and 14 quakes, respectively, the prediction of the earthquakes cannot function reliably enough.

6 Results of the logistic regression based on hourly values

For a shorter time period, the data is analyzed on an hourly basis. This restriction is caused by the (late) installation
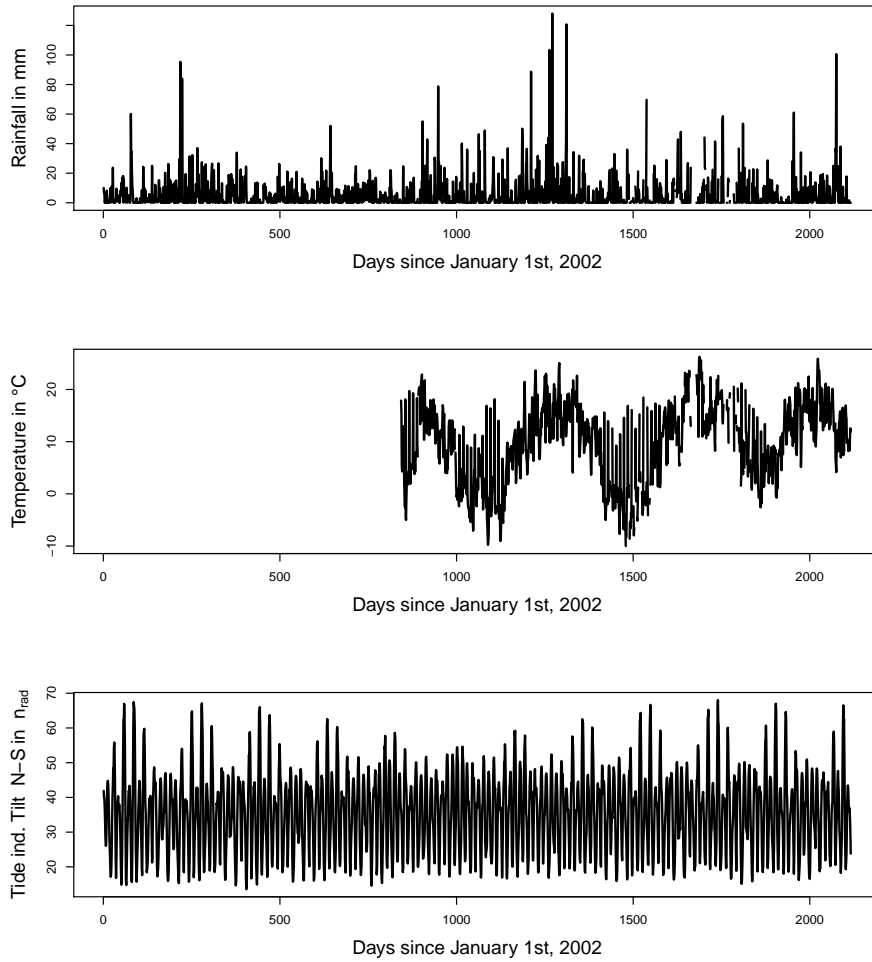
Fig. 6. Daily data in the Hochstaufen area in the sample period of 2100 days. Temperature was not available at the beginning of the observation.

of meteorological stations around Mt. Hochstaufen in mid 2004. For the time period 2001–2003, we only have access to daily rain values from meteorological stations of the national weather service. Therefore, our response variable converts in this section to

Y_{jt} = Number of quakes in depth category j at hour t

with

$x_{t-i} = i$-days lagged amount of rain in mm at hour t.

As the number of quakes in each depth category is rarely >1 (only in 9 h from 21 April 2004–27 February 2007) the Poisson model is changed to a logistic regression which simply distinguishes if there a quake occured or not. As no correlations in the deviance residuals could be detected, a generalized linear model was calculated. A nonlinear trend $f(t)$ was added to the model so that the generalized additive model *gam* was used (http://cran.r-project.org/, package gam).

**Table 5.** Lags of importance on the number of quakes (logistic regression based on hourly values).

Depth Category j	$j = 1$	$j = 2$	$j = 3$	$j = 4$
Depth in km	$]0, -1]$	$]-1, -2]$	$]-2, -4]$	$]-4, -\infty[$
Number of quakes	36	18	57	7
$i_{\max(j)}$	Lag 10	Lag 11	Lag 12	Lag 13
$\hat{\beta}_{i_{\max(j)}}$	0.010	0.006	0.004	0.003

In Fig. 10, we can see that the rain depth interaction coefficients behave similarly to the coefficients based on daily values. The closer a depth category is to the Earth's surface (depth category 1), the larger the coefficients. In depth category 1, the lagged rain of prime importance is the rain 10 days before a quake (see Table 5) while in depth category 4 the lagged rain of prime importance is the lagged rain 13 days before a quake.
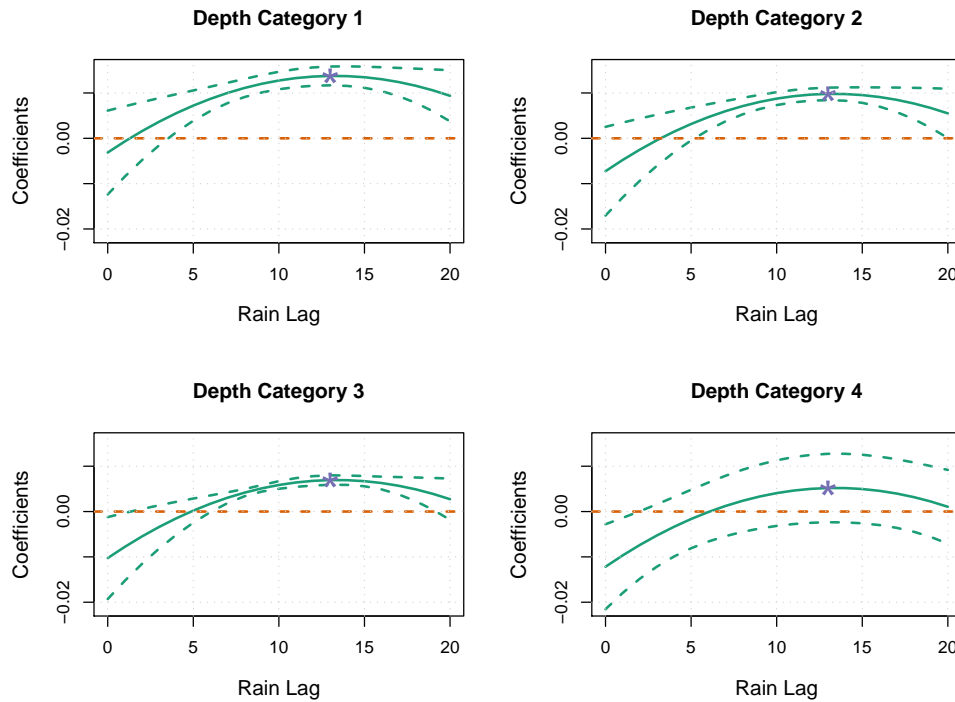
Fig. 7. Estimated rain depth coefficients $\hat{\beta}_{ij}$ with confidence intervals (dashed lines) for every depth category j (Poisson regression based on daily values). * indicates $\hat{\beta}_{i_{\max(j)}}$ for each depth category j.
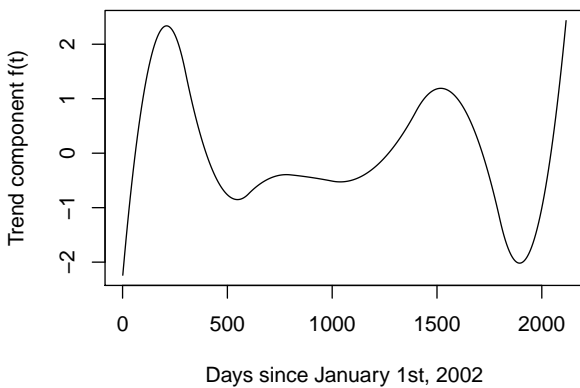
Fig. 8. Estimated trend component ($f(t)$) of the daily data set.

The accuracy of the maximum can be seen in Table 6. The maxima are not as exact as in the daily data set.

The tide variables (tide induced gravity, tilt and strain changes) to include in our model, according to the backward selection based on the AIC criterion, include the variables shown in Table 7. As a sensitivity analysis, we performed a forward selection using the AIC, which gives the preselected variables presented in Table 8. It becomes apparent that, depending on the selection criterion, we receive different tide variables to include into our model. However, when including the preselected variables in the full hourly data model

Table 6. Accuracy of the maximum for the logistic regression based on hourly values.

| Depth Category j	$j=1$	$j=2$	$j=3$	$j=4$
$i_{\max(j)}$	Lag 10	Lag 11	Lag 12	Lag 13
$\sqrt{\mathrm{Var}(\widehat{i_{\max(j)}})}$	1.288	0.934	1.358	2.139

Table 7. External coefficients besides rain–depth coefficients in the logistic regression based on hourly values (Backward Selection).

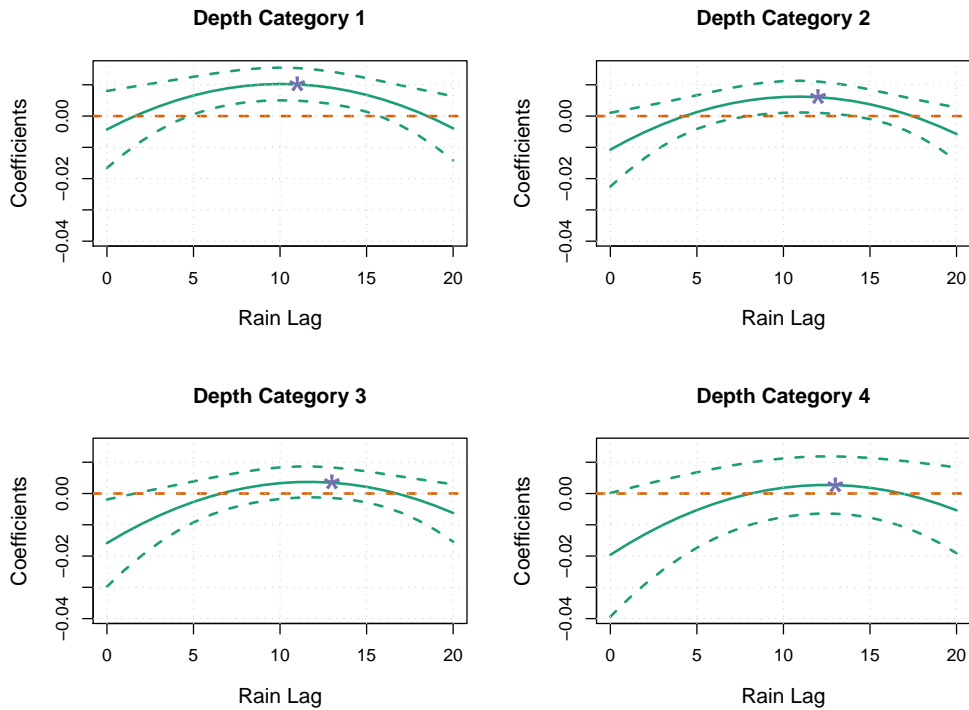|	Estimate	P-Value
Temperature	0.031	0.2326
Tide gravity changes	−0.006	0.3648
Tide induced tilt North South	0.007	0.5680
Tide induced tilt East West	−0.011	0.0018
Tide induced strain North South	0.033	0.5276

with all other covariates, only Tilt East West shows statistical significance (p-Value < 0.05) in both procedures.

7 Conclusions

The application of the established 1-D diffusion model (i.e. depth only) assumption for the earthquake front vs. rain
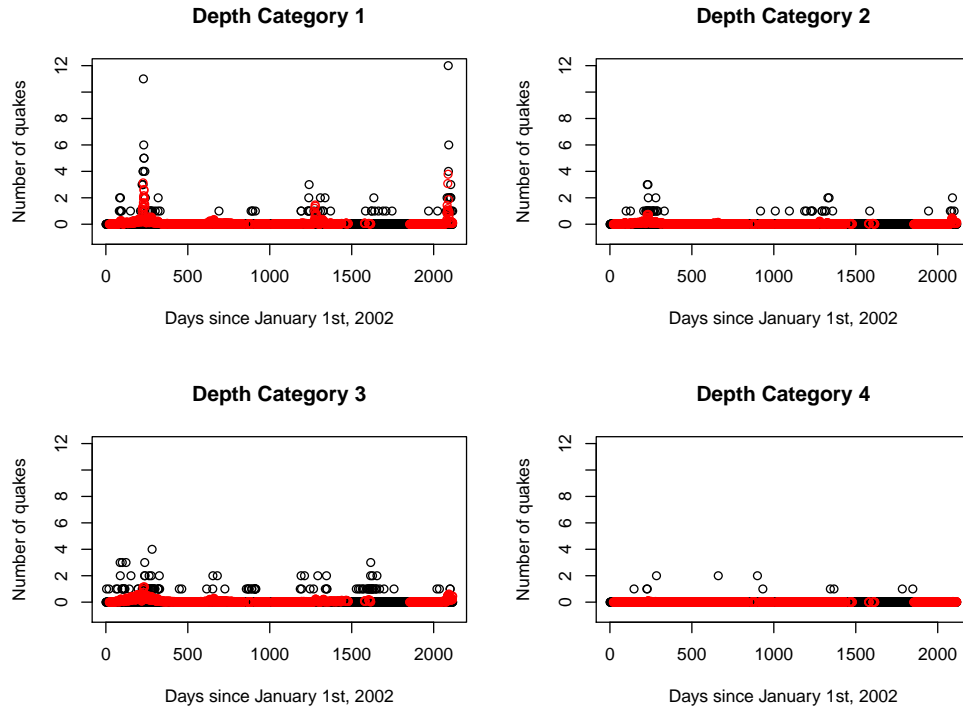
**Fig. 9.** Predicted (red dots) versus observed (black dots) number of quakes in each depth category j.

Fig. 10. Estimated rain depth coefficients $\hat{\beta}_{ij}$ with confidence intervals (dashed lines) for every depth category j (logistic regression based on hourly values). * indicates $\hat{\beta}_{i_{\max(j)}}$ for each depth category j.

Table 8. External coefficients besides rain–depth coefficients in the logistic regression based on hourly values (Forward Selection).

	Estimate	P-Value
Temperature	0.030	0.2527
Tide induced tilt East West	−0.011	0.0014
Tide induced strain East West	−0.013	0.2841

injection point shows profound differences with the observed maximum depth-time distribution for the years 2003–2008 at Mt. Hochstaufen. In this context, it is worth mentioning, that in later parts of the depth-time plot (see Fig. 4 >30 days after rain onset), no depth variation nor a clear backfront of seismicity as described by Shapiro and Dinske (2007) can be seen. Because of this obvious discrepancy of the 1-D diffusion model, we apply different, purely statistical models in the main part of this paper in order to evaluate the influence of various parameters (rain, air temperature, tidal parameters) on the forcing of earthquake swarms. The common way of parameterizing earthquake swarm data into a daily bases results in a pronounced lag time and significant coefficients in case of correlation with rain data. Surprisingly and in contrast to the finding of Kraft et al. (2006), the most important time lag is constant over all depth intervals in the daily data set, while the statistical significance and the influence of the computed coefficients is smaller in deeper source regions. This might point to either a very unsharp time dependence of the earthquake depth relationship or possibly reflect uncertainties of the hypocenter location and the chosen depth intervals. Nonetheless, the increase in the number of earthquakes depending on rain influx can be estimated to be approx. 10 % per 10 mm rain at its peak (i.e. after 13 days). The estimated delay is a bit larger than it was estimated by Kraft et al. (2006a), which used linear cross-correlation and the data set of 2002. While the tidal parameters (strain and tilt) are currently not measured with the existing monitoring network, it is possible to take them into account by calculating approximate values using a theoretical model (SPOTL: Agnew, 1996). This synthetic model does not account for any local geology or even topographic effects and might therefore not reflect the true situation at Mt. Hochstaufen. We furthermore didn't apply a depth dependent model as this would just lead to a more complex statistical model without any physical constraint. It becomes apparent, however, that the influence of a tidal component is, when present at all, very weak. None of the modeled parameters show a statistically significant influence on the daily earthquake rate.

The results of the hourly analysis shows a similar picture. The model parameters chosen again by the AIC forward and backward selection now show only a significantly negative influence of the east-west tilt component. How this can be explained in the geological or tectonic context at Mt. Hochstaufen needs to be answered in future studies where true tilt and strain may be measured in situ.
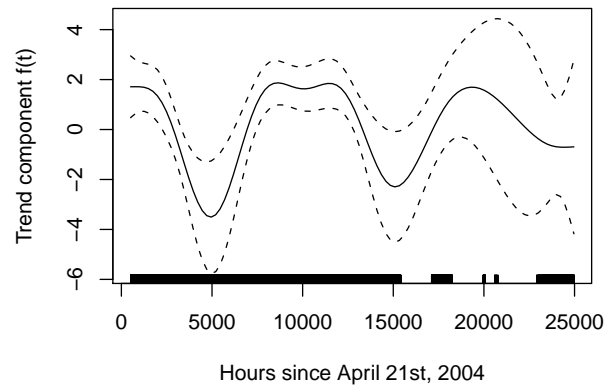
Fig. 11. Estimated trend component ($f(t)$) of the hourly data set.

With the use of hourly data, the time lag of rain influence changes. Here, the shallow depth intervals show a reduced time lag of reaction of about 11 days, which is close to the value of 10 days estimated by Kraft et al. (2006a) using a purely linear regression model. In this model, also a more pronounced time-depth relation can be found.

Overall, this study independently emphasizes the importance of rainfall on the triggering of earthquake swarms at Mt. Hochstaufen. What remains is a significant influence of temperature with no time delay on the occurrence of earthquake swarms in the daily data set. This raises questions regarding whether the temperature should be modeled with a possible time delay. Following the model of Ben-Zion and Leary (1986) the topography of Mt. Hochstaufen may influence the temperature field in a way that a local build-up of inhomogeneous thermo-elastic strain is possible. This hypothesis should be tested in the future.

In order to answer questions about the precise influence of tides on the swarm quake occurrence, the build-up of local stress transforming the crust from stable to critical condition and the mechanism of water transport to the fault systems, the installation of tilt and strain components to the existing monitoring network as well as a local study of persistent scattering In-SAR and a detailed modeling of the hydrology is needed as a next step at Mt. Hochstaufen. The significant influence of a trend function in both the daily and hourly data set with a similar period (approx. 400–500 days) makes a non seasonal loading process likely. This, however, needs independent strain magnitude measurement and possible accumulation of water and hydrological discharge modeling.

However, applying the Poisson model with distributed rain lags, we are able to quantify the influence and importance of various external parameters on triggering earthquake swarms at Mt. Hochstaufen without using an explicit physical model. In doing so, we were also able to confirm results of Kraft et al. (2006a) and Hainzl et al. (2006) about a strong influence of rain on the earthquake triggering process.

Edited by: G. Zöller
Reviewed by: S. Husen and another anonymous referee

References

Agnew, D. C.: SPOTL: Some programs for ocean-tide loading, SIO Ref. Ser. 96-8, 35 pp., Scripps Institution of Oceanography, La Jolla, CA, 1996.

Akaike, H.: Information theory and an extension of the maximum likelihood principle, in: 2nd International Symposium on Information Theory, edited by: Ptrov, B. N. and Csaki, F., Budapest, 276–81, 1973.

Ben-Zion Y. and Leary, P.: Thermoelastic strain in a half-space covered by unconsolidated material, BSSA, 76, 1447–1460, 1986.

Dietrich, J. H.: A constitutive law for rate of earthquake production and its application to earthquake clustering, J. Geophys. Res., 9, 2601–2618, 1994.

Fahrmeir, L. and Tutz, G.: Multivariate Statistical Modeling based on Generalized Linear Models, Springer, Berlin, 2001.

Guillas, S., Day, S., and McGuire, B.: Statistical analysis of the El Niño–Southern Oscillation and sea-floor seismicity in the eastern tropical Pacific, Phil. Trans. R. Soc., 368, 2481–2500, 2010.

Hainzl, S., Kraft, T., Wassermann, J. and Igel, H.: Evidence for rain-triggered earthquake activity, Geophys. Res. Lett, 33, L19303, doi:10.1029/2006GL027642, 2006.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning, Springer, Berlin, 2009.

Husen, S., Bachmann, C., and Giardini, D.: Locally triggered seismicity in the central Swiss Alps following the large rainfall event of August 2005, Geophys. J. Int., 171, 1126–1134, 2007.

Kraft, T., Wassermann, J., Schmedes, E., and Igel, H.: Meteorological triggering of earthquake swarms at Mt. Hochstaufen, SE-Germany, Tectonophysics, 424, 245–258, 2006a.

Kraft, T., Wassermann, J., and Igel, H.: High-precision relocation and focal mechanism of the 2002 rain-triggered earthquake swarms at Mt. Hochstaufen, SE-Germany, Geophys. J. Int., 167, 1513–1528, 2006b.

Miller, S. A.: Note on rain-triggered earthquakes and their dependence on karst geology, Geophys. J. Int., 173, 334–338, 2008.

Roth, P., Pavoni, N., and Deichmann, N.: Seismotectonics of the Eastern Swiss Alps and evidence for precipitation-induced variations of seismic activity, Tectonophysics, 207, 183–197, 1992.

Saar, M. O. and Manga, M.: Seismicity induced by seasonal groundwater recharge at Mt. Hood, Oregon, Earth Planet. Sci. Lett., 214, 605–618, 2003.

Schmedes, E.: Die seismische Aktivität im Raum Bad Reichenhall, Geol. Jahrb. C, 22, 91–102, 1979.

Schwartz, J.: The Distributed Lag between Air Pollution and Daily Deaths, Epidemiology, 11, 320–326, 2000.

Schwarzmann, A.: Untersuchungen der seismischen Aktivitaet im Raum Bad Reichenhall, Diploma thesis, University of Munich, 1996.

Shapiro, S. A. and Dinske, C.: Violation of the Kaiser effect by hydraulic-fracturing-related microseismicity, J. Geophys. Eng., 4, 378–383, doi:10.1088/1742-2132/4/4/003, 2007.

Shapiro, S. A., Huenges, E., and Borm, G.: Estimating the crust permeability from fluid-injection-induced seismic emission at the KTB site, Geophys. J. Int., 131, 15–18, 1997.

Woessner, J. and Wiemer, S.: Assessing the Quality of Earthquake Catalogues: Estimating the Magnitude of Completeness and Its Uncertainty, B.S.S.A, 95, 684–698, doi:10.1785/0120040007, 2005.