Prediction of regional wildfire activity in the probabilistic Bayesian framework of Firelihood

François Pimont, Hélène Fargeon, Thomas Opitz, Julien Ruffault, Renaud Barbero, Nicolas Martin-StPaul, Eric Rigolot, Miguel Riviere, Jean-Luc Dupuy

1   Running Head: Probabilistic fire activity modeling

2

3   **Title**

4   Prediction of regional wildfire activity in the probabilistic Bayesian framework of *Firelihood*

5

6   **Author names and affiliations**

7   François Pimont[a,e], Hélène Fargeon[a], Thomas Opitz[b], Julien Ruffault[a], Renaud Barbero[c],

8   Nicolas Martin StPaul[a], Eric Rigolot[a], Miguel Rivière[d], Jean-Luc Dupuy[a]

9   [a] Ecologie des Forêts Méditerranéennes (URFM), INRAe, 84914 Avignon, France

10   [b] Biostatistics and Spatial Processes, INRAe, 84914 Avignon, France

11   [c] Ecosystèmes Méditerranéens et Risques, INRAe, 13182 Aix-en-Provence, France

12   [d] Université de Lorraine, Université de Strasbourg, AgroParisTech, CNRS, INRAe, BETA,

13   54000 Nancy, France

14   [e] Corresponding author. E-mail: francois.pimont@inrae.fr

15

16

17

18 **Abstract**

19 • Modelling wildfire activity is crucial for informing science-based risk management and

20 understanding the spatio-temporal dynamics of fire-prone ecosystems worldwide. Models help

21 disentangle the relative influences of different factors, understand wildfire predictability and

22 provide insights into specific events.

23 • Here, we develop *Firelihood*, a two-component Bayesian hierarchically-structured

24 probabilistic model of daily fire activity, which is modelled as the outcome of a marked point

25 process: individual fires are the points (occurrence component), and fire sizes are the marks

26 (size component). The space-time Poisson model for occurrence is adjusted to gridded fire

27 counts using the integrated nested Laplace approximation (INLA) combined with the Stochastic

28 Partial Differential Equation (SPDE) approach. The size model is based on piecewise-estimated

29 Pareto and Generalized-Pareto distributions, adjusted with INLA. The Fire Weather Index

30 (FWI) and Forest Area are the main explanatory variables. Temporal and spatial residuals are

31 included to improve the consistency of the relationship between weather and fire occurrence.

32 • The posterior distribution of the Bayesian model provided 1000 replications of fire activity

33 that were compared with observations at various temporal and spatial scales in Mediterranean

34 France. The number of fires larger than 1ha across the region was coarsely reproduced at the

35 daily scale, and was more accurately predicted on a weekly basis or longer. The regional weekly

36 total number of larger fires (10 to 100 ha) was predicted as well, but the accuracy degraded with

37 size, as the model uncertainty increased with event rareness. Local predictions of fire numbers

38 or burnt areas also required a longer aggregation period to maintain model accuracy.

39 • The estimation of fires larger than 1ha was also consistent with observations during the

40 extreme fire season of the 2003 unprecedented heat wave, but the model systematically

41 underrepresented large fires and burnt areas, which suggests that the FWI does not consistently

42 rate the actual danger of large fire occurrence during heat waves.

43   • *Firelihood* enabled a novel analysis of the stochasticity underlying fire hazard, and offers a

44   variety of applications, including fire hazard predictions for management and projections in the

45   context of climate change.

46

47

48

49   **Keywords:** Bayesian; fire; *Firelihood*; INLA; Mediterranean; spatiotemporal; Fire Weather

50

51

1. INTRODUCTION

Wildfires contribute to shape ecosystems across large parts of the world and threaten human lives and properties. Mapping features of fire regimes such as frequency, size, intensity, severity or pattern of fires across time and space is useful for planning fire and natural resource management, assessing risk and evaluating ecological conditions (Morgan et al. 2001). Indeed, fire regimes vary substantially over time and space at multiple scales, in response to weather, climate, vegetation, orography, as well as local and regional human influences (e.g. Bradstock 2010, Bowman et al. 2011, Parks et al. 2012). Understanding fire regimes and their economic, social and ecological consequences is a major challenge for scientists, especially in the context of climate change, which is expected to increase fire activity in many regions of the world (e.g. Flannigan et al. 2009, Barbero et al. 2015a, Turco et al. 2018, Dupuy et al. 2020).

Fire regimes are strongly influenced by contemporary fire management, which often aims at reducing fire activity. In some locations (US), burnt areas increased substantially despite large suppression expenditures that led to increased fire hazard through fuel accumulation, which suggests the need to reexamine policies (Stephens and Ruth 2005, Calkin et al. 2015). By contrast, fire suppression policies have likely been effective for reducing burnt areas in many regions of the Mediterranean basin (Turco et al. 2016), but the long-term adequacy of such policies in the context of climate warming and fuel build-up is currently debated (Moreira et al. 2020). In this context, the design and application of new policies require reinforced land management and planning, while fire suppression must continue to play a key role in the protection of human lives and assets. For planning purposes, managers and policy makers need to anticipate future scenario-based fire regimes, while for preparedness and response actions, fire managers need to be informed on daily, weekly and seasonal bases of the expected number, size, duration and spread rate of fires (Taylor et al. 2013; Xi et al. 2019).

76   While wildfire regimes depend on multiscale interactions between climate, vegetation and

77   humans (Moritz et al. 2005), weather has long been recognized as the main factor driving

78   regional fire activity from daily to seasonal scale (Abatzoglou and Kolden 2013, Barbero et al.

79   2015b, Turco et al. 2017). Much effort has been dedicated to developing and evaluating

80   weather-based fire danger rating systems, including the widely used Canadian Fire Weather

81   Index (FWI, Van Wagner 1987), the Australian McArthur index (FFDI, Noble et al. 1980) or

82   the American National Fire Danger Ratings System (NFDRS, Deeming et al. 1978). These

83   indices operate at the daily time scale and can be computed in real time from local weather

84   variables to inform managers, or they can be projected under future climatic scenarios to

85   anticipate the effect of climate change (Dupuy et al. 2020). However, the link between fire

86   danger rating systems and observed fire activity is not straightforward. Indeed, fire events are

87   fairly rare at local and daily scales, and hence, highly random in nature. To handle this

88   stochasticity, observations are often aggregated over time and space prior to examining

89   empirical relationships between fire activity and average indices, typically using weekly to

90   monthly bases (e.g. Krawchuck et al. 2009, Barbero et al. 2014, Turco et al. 2018).

91   Unfortunately, these correlative approaches cannot appropriately account for a number of

92   operational and research applications that require daily predictions on fine scales. Indeed,

93   climate, land cover and human variables can vary substantially over short distances in some

94   regions (Fréjaville and Curt 2015). Likewise, weather processes, such as wind or hot

95   temperature events, can influence fire activity on daily or even sub-daily timescales. This is

96   typically the case in the Mediterranean region, where most fires spread during less than a day

97   and the final fire size is less than 1000 ha, contrary to other regions where fires can spread over

98   several weeks, for which daily variations would be less relevant.

99   The rareness and the stochastic nature of individual fire events can be addressed in a formalized

100   probabilistic framework (Brillinger et al. 2003, Preisler et al. 2004, Preisler and Westerling

101   2007, Turner 2009, Vilar et al. 2010, Woolford et al. 2014, Serra et al. 2014a&b). In this

102   approach, observed patterns of fire occurrence are viewed as realizations of a spatio-temporal

103   point process, where points correspond to locations and times of ignition of a fire, and the burnt

104   area is used as a mark for the points. The latent (*i.e.*, unobserved) spatio-temporal intensity

105   function that has generated the observed point pattern is then estimated. In practice, this point

106   process is often approximated by a Bernoulli probability of fire presence in discrete and fairly

107   small space-time cells (called voxels, typically some km^2 X days) in which at most one fire has

108   generally been observed. The notion of intensity (*i.e.*, expected counts) is crucial since it

109   provides more information than only susceptibility (*i.e.*, presence-absence); in particular,

110   intensities can be additively aggregated within any spatio-temporal unit. Such fire occurrence

111   modelling can be combined with fire size distribution models, typically expressed as the

112   probability for a fire to exceed a given size, to simulate fire hazard (Preisler et al. 2004, 2011).

113   These probabilistic models have most commonly been adjusted within the framework of

114   generalized linear modeling (GLM), or of related extensions such as generalized additive

115   modeling (GAMs, Wood et al. 2006), where the latter have been shown to perform better

116   (Woolford et al. 2011), since they allow replacing linear effects of explanatory variables (such

117   as fire danger and/or human activity metrics) by more flexible shapes in nonlinear effects.

118   Besides accounting for non-linear effects of explanatory variables – as many other techniques-,

119   GAM can include model components to account for spatial residuals (Preisler et al. 2004)),

120   *i.e.*, spatial coordinates are used as explanatory variables with smooth nonlinear effects.

121   More recently, Bayesian methods have also been used as an alternative to these frequentist

122   methods (Serra et al. 2014a&b, Joseph et al. 2019). They allow including and accurately

123   estimating random components in the predictor to capture variation in components of fire

124   activity that cannot be explained by a deterministic influence of other available explanatory

125   variables. Moreover, expert knowledge can guide the choice of prior distributions of predictor

126   components and smoothing parameters (*e.g.*, variances and dependence ranges). In particular,

127   the complexity of sophisticated model components can be controlled by shrinking them towards

128   simpler baselines when no strong signal in the data exists. Therefore, Bayesian analysis

129   provides a convenient and flexible setting for inference in hierarchically structured models. In

130   particular, spatially correlated data can be handled, for example with the Stochastic Partial

131   Differential Equations approach that allows for highly resolved spatial random effects (SPDE,

132   Lindgren et al. 2011). Moreover, posterior distributions of parameters allow for interpretation

133   of uncertainties and provide decision support thanks to credible intervals. Finally, predictive

134   distributions for unavailable observations (*e.g.*, future observations) -not only point predictions-

135   can be naturally generated from the posterior distributions and new explanatory variables.

136   Probabilistic models in general -but mostly in a frequentist setting- have been used for a variety

137   of applications, including forecasts of large fires (Preisler et al. 2008), the projection of future

138   fire activity (Ager et al. 2018), the estimation of suppression costs (Preisler et al. 2011), or the

139   estimation of extreme fire size (Joseph et al. 2019).

140

141   Despite their potential for wildfire predictions, probabilistic approaches still present some

142   challenges and limitations, and some of which have not been fully addressed. First, the

143   evaluation of the underlying model performance in a probabilistic framework is not

144   straightforward. Indeed, it requires checking the goodness-of-fit and model parsimony in the

145   model-building framework through various approaches including information criteria,

146   comparisons of predictions and uncertainty bounds with observations aggregated on various

147   temporal and spatial scales, and external validation using hold-out data (Xi et al. 2019). Second,

148   even if early probabilistic approaches combined models of occurrence and exceedance of fire

149   size above high fixed thresholds, they did not simulate the size of fire events. Notable

150   exceptions are Westerling et al. 2011 and Ager et al. (2014, 2018), who fitted generalized Pareto

151   distributions with parameters depending on explanatory variables to simulate the size of

152   individual fires. Third, probabilistic approaches have seldom been used to evaluate the potential

153   predictability of fire activity (i.e. the degree to which observations can be deterministically

154   predicted) across a range of spatial and temporal scales. Indeed, it is expected that fire activity

155   is less predictable at short temporal and/or fine spatial scales and for rare events (large fires),

156   than more frequent events (small fires) over longer and/or broader scales. Probabilistic

157   approaches provide a suitable framework to quantify this predictability, which should help

158   managers to understand observed activity patterns. Moreover, probabilistic models help

159   understand the extent to which catastrophic events are (un)expected, and can therefore provide

160   useful information regarding their likelihood of occurrence, such as return periods and levels.

161

162   The objective of the present study is to assess the predictability of fire activity at various

163   temporal and spatial scales in the French Mediterranean region, through a Bayesian

164   probabilistic approach. To this aim, we present and use a full framework of fire activity

165   modelling, called *Firelihood*, which simulates potential scenarios of daily fires occurring in

166   small pixels (8 x 8 km). We then assess the overall model performance, the relative importance

167   of selected explanatory variables, and the predictability at scales ranging from the pixel to the

168   region, and from days to periods of multiple years. The assessment of model performance

169   includes a specific focus on the catastrophic 2003 year characterized by a severe synoptic-scale

170   heat wave in summer following a prolonged drought (Trigo et al. 2005). We finally discuss the

171   strength and weaknesses of the current model and its potential applications for wildfire-related

172   research avenues and the improvement of operational fire suppression and management.

173

174   **2. METHODS**

175   ***2.1. Data and site description***

176   ***Study site and fire activity.*** The study area consists of 15 NUTS3-level French administrative

177   units located in southeastern France (Fig. 1a, 75,560 km^2), which concentrate the vast majority

178   of burnt area during the summer season in France. The climate of this area is mostly

179   Mediterranean, characterized by cool and moist winters and hot and dry summers, but exhibits

180   strong variations with orography, from the high mountains in the Alps to coastal plains. The

181   population is mostly concentrated near the Mediterranean coast and the Rhône river valley.

182   These climatic and socio-economic contrasts strongly shape variations in fire activity over time

183   and space. Fire activity is the highest near the coast and in the Corsican island, where human

184   activities, drought and wind bursts come together (Fig. 1c). Burnt area shows a bimodal

185   seasonal pattern, with a first peak in spring associated with agricultural, pastoral and forestry

186   practices, during which fires are generally not a major threat, and a more important second peak

187   during the summer dry season, during which most large fires occur (Fig. 1b). At the interannual

188   scale, fire activity is highly variable (Fig. 1d) and mostly dictated by annual drought conditions

189   (Ruffault et al. 2016, Barbero et al. 2019). The outstanding burnt area of the 2003 summer was

190   due to several extreme fires that occurred during an unprecedented heatwave (Trigo et al. 2005,

191   Ruffault et al. 2018a). Following these 2003 extreme fires, fire prevention and fighting was

192   enhanced with a modernization of the fire management law in 2004. This might explain the

193   decrease in the number of fires larger than 1 ha and in burnt areas after 2003 (Fig. 1d, Curt et

194   al. 2018). In France, fires larger than 1 ha are of special interest, as limiting fire size to 1 ha is

195   a goal of fire suppression services during the dry season. Fires larger than 1ha will therefore be

196   referred to as "escaped fires" in the remainder of the article.

197   Fire records were extracted from the Prométhée fire database (http://www.promethee.com/) for

198   the period from 1995 to 2018. This period was selected so that the dataset was large enough to

199 allow the fitting of robust models. We discarded, however, the pre-1995 period, because of the

200 lack of consistency of the weather data prior to 1995 (monitoring station number had evolved

201 until 1995, Vidal et al. 2010) and of reliability and completeness issues in earlier fire records.

202 Similarly, to limit the uncertainties associated with small fires in fire databases (Turco et al.

203 2013, Ruffault and Mouillot 2015), only fires larger (or equal) than 1 ha (or escaped fires) were

204 retained. One should note, however, that the increasing precision of size records over time has

205 led to a temporal decline of the proportion of fires exactly equal to 1ha among small fires.

206 We focused our analysis on the summer season (weeks 22-44, 25th may to 31th October, Fig.

207 1b), as most burnt areas occur during summer, and because the causes and the factors behind

208 spring fires are quite different and would have blurred the fire-climate relationships we sought

209 to explore.

210 ***Explanatory variables.*** The main explanatory variable was the daily Fire Weather Index (FWI),

211 which represents temporal and spatial variations in meteorological fire danger. FWI was

212 computed onto an 8 km-resolution grid from 12:00 LST meteorological variables (24h-

213 cumulated precipitation, mean wind speed, mean temperature and minimum relative humidity,

214 calculated using specific humidity and maximum temperature) following Bedia et al. (2014),

215 using the 'cffdrs' R package (Wang et al. 2017). These variables were extracted from the

216 SAFRAN reanalysis (Vidal et al. 2010).

217 The second explanatory variable was the forest area in each 8-km pixel, which is expected to

218 affect both the number and size of fires. It shows significant spatial variability (Appendix S1:

219 Fig. S1). Forest area was obtained from the CORINE land-cover database (CLC,

220 https://land.copernicus.eu/pan-european/corine-land-cover), by merging the patch areas

221 covered by sublevels "Forests" and "Scrub and/or herbaceous vegetation association" in each

222 pixel. This forest area (FA, in ha or in % cover of the pixel) was estimated on a yearly basis by

223 linear interpolation of CLC inventories available in 1990, 2000, 2006, 2012 and 2018.

224

2.2. Probabilistic model of fire activity

Model overview. *Firelihood* consisted of two hierarchically structured components: one describing the occurrence of escaped fires, and another describing the size of each fire event conditional to its occurrence (Fig. 2). For the occurrence component, the response variable was the daily number of escaped fires (i.e. fires larger than 1 ha), for each pixel of the FWI grid. For the size component, the response was a continuous positive quantity (size of each escaped fire event) modelled with a piecewise distribution, for flexibility.

Both the occurrence and size components included FWI and forest area as explanatory variables. The occurrence model also included two temporal factors and a spatial model. Models were fitted in a Bayesian framework, using the integrated nested Laplace approximation (INLA) implemented in R software (www.r-inla.org) and described in (Rue et al. 2009, Lindgren and Rue 2015). INLA can be applied to large datasets using sophisticated hierarchical structure and provides accurate and (relatively) fast inference by means of analytical approximations of the posterior model, in contrast to standard, simulation-based Bayesian approaches (Markov-Chain Monte-Carlo). It allows non-linear responses to explanatory variables to be estimated through flexible Gaussian prior distributions for spline functions in combination with spatial models.

Model components were trained with data from 1995-2014 (training sample), the years 2015-2018 being withheld for the evaluation of the predictive performance (validation sample). The dataset for fire occurrence contained fire counts ($\geq$ 1 ha) for approximately 4.44 million pixels-days, whereas the dataset of observed fire sizes contained 7193 fires ($\geq$ 1 ha).

In the next two subsections, we describe the "full" model that includes all explanatory variables (Table 1). To verify the added value of the "full" model and to avoid overfitting (*i.e.* the situation where prediction performance on validation data decreases) (Xi et al. 2019),

249   intermediate models for fire occurrence and size with less explanatory variables were also

250   estimated, and their corresponding information criteria were compared with those of the "full"

251   model.

252   Combining both components of the models enable simulation, according to the estimated

253   posterior model, of an unlimited number of replications (here: 1000) of the potential daily fire

254   activity in each pixel, in the form of escaped fire lists whose number and sizes were simulated

255   with each model component. These simulations can then be aggregated at different spatial and

256   temporal scales for corresponding predictions and evaluations against observations.

257

258   ***Fire occurrence component.*** We built the fire occurrence model for escaped fire ($\geq$ 1 ha)

259   counts in 8 km X 8 km daily voxels, as a Poisson random variable (Fig. 2). Following the

260   approach of Brillinger et al. (2003), we incorporated residual spatial and temporal random

261   effects (at the pixel size and weekly, respectively), to account for unknown sources of variations

262   in escaped fire probability. They can be viewed as spatial and temporal scaling factors between

263   FWI and the observed number of escaped fires.

264   The voxel size was considered as a good approximation for the "true" Poisson distribution

265   resulting from grouping intra pixel variations, since pixel-day probabilities remained small

266   (Brillinger et al. 2003). Contrary to large voxels in which multiple fires can occur more often

267   (e.g. Joseph et al. 2019), our Poisson-based method was applied to an almost binary dataset,

268   and spatial correlations were accounted for with a spatial model, so that over-dispersion was

269   less of a concern (Taylor et al. 2013). For the same reasons, the use of a zero-inflated Poisson

270   model was not required. Moreover, this resolution was fine enough to explicitly link the fire

271   occurrence probability to locally observed fire conditions (weather data and forest area), rather

272   than some average value at a coarser scale (Taylor et al. 2013). To identify the range of variation

273   of spatial biases unexplained by the available predictor variables (FWI, FA, season), the pixel

274   size should be much smaller than the distance at which the correlation in the spatial model drops

275   to near zero. This avoids issues related to within-pixel overdispersion and overestimation of the

276   smoothness of occurrence intensity maps. Model fit showed that this range was approximately

277   30 km, which is indeed substantially larger than pixel size. Finally, the pixel size was consistent

278   with the computational and memory costs of INLA, which strongly increase with the size of the

279   dataset and the resolution of spatial and temporal random effects.

280   The partial effects of the models were assumed to be multiplicative, based on an additive

281   decomposition of the log of expected fire counts, which has been shown to be adequate for time

282   and space in Woolford et al. (2011). The form of the "full" model, including all explanatory

283   variables ("FWI+2003+FA+WEEK+SPATIAL", see Table 2) was:

284    $\log N_i \sim \log(FA) + \beta(YEAR_i > 2003) + f_{FWI}(FWI_i) + f_{FA}(FA_i) + f_{X,Y}(X_i, Y_i) + f_{WEEK}$
285    $(WEEK_i)$ (1)

286   where $log(FA)$ was a deterministic offset, β the fixed effect (with a different value before and

287   after 2003), and the f-terms captured nonlinear influences of the covariates FWI and FA, as

288   well as spatial and temporal effects.

289   Because escaped fires cannot occur in non-forested area (urban areas, crops, etc.), the area of

290   each pixel in which fire "points" could happen was not spatially constant. This variability was

291   incorporated in the Poisson model with an offset equal to forest area (FA). The model allowed

292   non-linear effects of FWI, but also of FA – in addition to the offset – as a land use factor.

293   Indeed, it is expected that the probability to get a fire per area of forest decreased for high forest

294   area, since interface, road and urban densities decreased. Spatial effects were represented using

295   the Stochastic Partial Differential Equation approach (SPDE, Lindgren et al. 2011), which

296   estimates the spatial model for residuals, through continuous spatial random effects at high

297   resolution. Temporal effects were incorporated as a non-linear weekly seasonal factor and a

298   fixed effect "post 2003". This "post 2003" effect should not be interpreted as an actual shift in

299   the relationship that occurred exactly in 2003, but more as a convenient and simple manner to

300   incorporate in the model the temporal evolution of the fire-weather relationship. In this study,

301   we decided to ignore other annual effects in order to develop a model applicable to predictions

302   and projections. Contrary to other studies (e.g. Opitz et al. 2020), we did not seek to model time

303   variation in spatial patterns.

304   The prior distributions of the different predictors were Gaussian processes. The nonlinear f-

305   functions in (1) were modelled with piecewise constant first-order random walks, with 30, 18

306   and 23 segments, for FWI, FA and the seasonal effect, respectively. For each of them, one

307   hyperparameter (called *precision*) governed curve smoothness (i.e., the size of the small steps

308   between consecutive segments), as a Bayesian variant of smoothers used in GAM models (e.g.

309   Preisler et al. 2004). For the spatial component, the SPDE approach consists in implementing a

310   numerically convenient approximation to the Matérn covariance function for the Gaussian

311   random field prior of $f_{X,Y}$ in the 1143 pixels (meshing the study site). Two hyperparameters

312   were estimated for this random field: precision (to control the spatial variability of field values)

313   and range (to control spatial dependence, *i.e.* the smoothness of the spatial surface). For

314   hyperparameters, we specified Penalized Complexity priors (Bakka et al. 2018), which

315   penalized the distance of a model component towards a basic baseline (*i.e.*, absence of effect),

316   and we fixed penalty parameters that ensured fairly smooth estimated posterior effects. In order

317   to limit computational and memory costs, we took advantage of the additivity of the Poisson

318   process to aggregate data in segment classes to reduce dataset size, which initially contained

319   4.44 million voxels. The numerical design described above enabled keeping the number of

320   observed classes below 500,000, which avoids numerical instabilities when running R-INLA,

321   and models are estimated within several minutes to several hours in case of the full model. Such

322   an aggregation is an appealing alternative to approximations often implemented for large

323   datasets (e.g. subsampling of non-fire voxels, in Brillinger et al. 2003), which have shown to

324   decrease model robustness (Woolford et al. 2011).

325

326   ***Fire size component.*** We built a probabilistic model for sizes of escaped fires (conditional on

327   a fire being larger than 1 ha), which corresponded to the marks of the "fire" point-process

328   (Fig. 2). The fire size distribution is usually not well reproduced by any of the commonly used

329   probability distributions over the whole range of observations -in particular for small and large

330   fires- (Cui and Perera 2008). Consequently, we used a piecewise specification of the distribution

331   based on Pareto and Generalized Pareto Distributions (GPD) in the different size segments, as

332   justified by the asymptotic theory of threshold exceedances (Davison and Huser 2015). In each

333   segment, size distributions depended on both FWI and FA of the voxel (8 km X 8 km by daily

334   cell) in which the fire initially spread. In principle, we could have estimated the probability of

335   a given fire to exceed the upper threshold of each segment by using the exceedance probability

336   derived from the fire size distribution within this segment. However, because of the small

337   fraction of fire sizes in the higher parts of each segment (*i.e.*, most fires have size closer to the

338   lower than the upper bound of each segment), we obtained more accurate estimates of

339   exceedance probabilities with specific logistic regressions for each threshold (Bernouilli

340   process, see Fig. 2). In summary, the size model was generative and had hierarchical structure

341   using a piece-wise specification over intervals of burnt area (Fig. 2). First, a logistic-regression-

342   based model determined the segment into which each individual fire should fall (1-10 ha, 10-

343   100 ha, 100-1000 ha or larger than 1000 ha). Then, the exact size was simulated according to

344   the distribution of the corresponding segment.

345   Contrary to other regions of the world where fires can spread over tens of km^2 during several

346   days (e.g. Joseph et al. 2019), most fires in the study area spread for less than a day and were

347   smaller than 1000 ha (which was much smaller than the pixel area of 6400 ha). Therefore, it is

348   appropriate to stick to the voxel scale for fire size modeling, even if a few fires spread over

349   more than one voxel. The rationale for including FA (in addition to the FWI) was that a small

350   forest area is thought to limit fire spread. Contrary to the fire occurrence model, we did not

351   include any other spatio-temporal factors in the size model, as the dataset size was too small to

352   develop robust models. The form of the size component of this "full" model was hence

353   "FWI+FA" (Table 1), except for the GPD (Table 3).

354

355   The piecewise model of fire size distribution was developed using standard modeling

356   techniques suggested by extreme-value theory, based on threshold exceedances. We carried out

357   preliminary analyses of the response of fire size distributions in different FWI classes based on

358   mean excess plots (Hall and Wellner 1981) of the log-transformed escaped fire sizes. The

359   number of exceedances over increasingly high thresholds suggested a slow power-law-like tail

360   decay for most of the thresholds except the highest ones, for which exceedance numbers seem

361   to decrease much faster as in the power-law setting, similar to the findings in Cui and Perera

362   (2008). For our data, the behavior of mean excess curves of log fire sizes, and of related curves

363   (cumulative distributions in log-log scale), tended to change around fire sizes of 10 ha, 100 ha

364   and 1000 ha. Therefore, we assumed that the distribution of fire sizes could be modelled through

365   piecewise Pareto distributions between thresholds $u_1 = 1$, $u_2 = 10$, $u_3 = 100$ and u_4

366    $= 1000$ ha, which depended on both FWI and FA (equivalently, through piecewise exponential

367   distributions for log fire sizes). More precisely, given a threshold u_k, we estimated exponential

368   regression models for $log\left(\frac{S_i}{u_k}\right)$, where S_i corresponds to an observed fire size larger than u_k, for

369   segments k=1, 2, 3; moreover, we censored observations $S_i > u_{k+1}$, so that the model provides

370   a good fit for $u_k \le S_i < u_{k+1}$ by construction. The estimation was conducted with INLA using

371   its survival model framework for handling censoring, and FWI and FA were used as covariates

372   with potentially nonlinear influence:

373    $$log\frac{S_i}{u_k} \sim e^{\eta_i}, \eta_i = \beta_0^{exc,u} + f_{FWI(FWI_i)}^{exc,u} + f_{FA(FA_i)}^{exc,u} \qquad S_i > u \qquad (2)$$

374   where the f-terms captured nonlinear influences of the covariates FWI and FA, with piecewise

375   constant first-order random walks, with 20 and 10 segments, respectively.

376   Finally, for the category with largest fire sizes (exceeding 1000 ha and containing 33 and 7 fires

377   for the periods 1995-2014 and 2015-2018, respectively), we selected a generalized Pareto

378   distribution (GPD), which allows for a finite upper endpoint if its shape parameter is negative,

379   since an upper bound – even a very large one – must necessarily exist for physical

380   considerations. The GPD has shown to perform generally well for large fire sizes (Schoenberg

381   et al. 2003; Westerling et al. 2011). Therefore, we estimated the shape ξ and scale σ parameters

382   of the GPD, by fitting it to $log\left(\frac{S_i}{1000}\right)$ for observations $S_i > 1000$. Since this model was not

383   available within INLA with a negative shape parameter (due to some peculiarities of its density,

384   *e.g.*, de Haan and Ferreira 2007), we estimated the GPD parameters using frequentist maximum

385   likelihood, followed by a careful inspection of the estimated model. Owing to the small sample

386   size, we chose a more parsimonious parametrization of covariate influence using only linear

387   coefficients:

388
$$\begin{cases} log\left(\frac{S_i}{1000} \mid S_i > 1000 \right) \sim GPD\{\xi(FWI_i, FA_i), \sigma(FWI_i, FA_i)\} \\ \quad \xi(FWI_i, FA_i) = \xi_0 + \xi_1 * FWI_i + \xi_2 * FA_i \\ \quad log\sigma(FWI_i, FA_i) = \sigma_0 + \sigma_1 * FWI_i + \sigma_2 * FA_i \end{cases} \quad (3)$$

389   However, because the sample size was small, uncertainty on the FA coefficient was high and

390   confidence intervals and information criteria advised against including FA in this model (see

391   Appendix S2 for details). We hence selected for the "Full" model a GPD with parameters

392   function of FWI only.

393   As mentioned above, we cannot expect a good estimation of exceedance probabilities derived

394   from the three Pareto distribution due to the relatively small sample fraction of fire sizes in the

395   higher segments. For example, a moderate number of fires were greater than 10 ha (1348 fires

396   among 7193), and only few of them (280, approx. 4 % of fires larger than 1 ha) were also larger

397   than 100 ha. Between 1995 and 2018, only 40 fires reached more than 1000 ha (0.6% of fires

398   larger than 1 ha). Therefore, we separately modeled and estimated these exceedance

399   probabilities p_i^u, for a threshold u and a voxel i, with INLA, based on logistic regressions for

400   the indicator variables of threshold exceedances (i.e., 1 if fire size exceeds u_k and 0 otherwise),

401   given FWI and FA:

402    $$log\frac{p_i^u}{1-p_i^u} = \beta_0^{p,u} + f_{FWI}^{p,u} \qquad \text{for } u = 10,100,1000 \text{ (4)}$$

403   where the f-terms captured nonlinear influences of the covariates FWI and FA, with piecewise

404   constant first-order random walks, with 20 and 10 segments, respectively.

405   These three probabilities and the four estimated fire size distributions were hence combined to

406   predict the size of each fire larger than 1 ha, given FWI and FA, through a sequential approach

407   consisting in simulating i) in which segment the size of the fire is, and, conditional to the

408   segment, ii) the exact size (in this segment).

409

410   ***Variable selection and model evaluation.*** The final "full" model was developed by including

411   the different explanatory variables and non-linear functions step by step, checking information

412   criteria of the intermediate probabilistic models. A selection of intermediate models is presented

413   in Tables 1,2,3, ranging from the simple "FWI-linear" to the "full" model. Information criteria

414   aimed to assess goodness-of-fit of models while safeguarding against overly complex models

415   that overfit data (Vehtari et al 2017), and were an appropriate means to check if the structure of

416   each response to an explanatory variable, as implemented in the "full" model, was significant

417   and parsimonious. We used the DIC (Deviance Information Criterion) and the WAIC (Widely-

418   Applicable Information Criterion) for variable selection in submodels, which are

419   generalizations of the well-known Akaike Information Criterion (AIC) for Bayesian models.

420 The WAIC is known to better reflect posterior uncertainty in the models' prediction than DIC,

421 which can sometimes select over-fitted models (Vehtari et al 2017). For the Generalized Pareto

422 Distribution of largest fires, we simply used the AIC of the model fit. The robustness of the

423 occurrence model was checked in a preliminary development, thanks to a 7-fold cross validation

424 procedure, holding out 3 randomly selected years in each fold, which demonstrated little

425 sensitivity to data sample (Fargeon 2019).

426 We evaluated the performance of the model by two different means. First, we evaluated the

427 subcomponents of the model with Area-under-the-Curve measures (AUC, Fawcett 2006).

428 AUCs rate the model ability to diagnose the realization in voxels of different events, here "at

429 least one escaped fire" and a selection of "size exceedances", by verifying that their occurrence

430 probabilities were better predicted when including more explanatory variables and/or non-

431 linear responses. AUC values range between 0 and 1, with 1 indicating perfect prediction of the

432 binary presence/absence information, whereas 0.5 indicates a random prediction. AUCs were

433 computed for both the 1995-2014 (training) and 2015-2018 (validation) periods. Second, we

434 evaluated model performance by comparing simulations with historical observations

435 aggregated on various temporal and spatial scales (Xi et al. 2019). These evaluations were

436 carried out from 1000 replications of fire occurrence per voxel, which were sampled as a

437 Poisson process according to draws from the posterior predictive distributions of the occurrence

438 intensity. Note that INLA (in contrast to Markov Chain MonteCarlo, MCMC) does not provide

439 simulations of the posterior model's component during the estimation process, but sampling

440 from the fitted model is nevertheless straightforward (e.g. Fuglstad and Beguin 2018). A fire

441 size was then randomly assigned to each simulated escaped fire based on the size submodels

442 (Fig. 2) parametrized with posterior mean parameters. This approach allowed considering the

443 inherent variability of the stochastic processes at stake. This variability was used to draw

444 pointwise envelopes showing the spread between 5th and 95th percentiles of fire activity, and to

445   compute central tendencies for the different spatio-temporal aggregations of simulated fire

446   activity.

447   The overall goodness-of-fit between central tendencies and observations were measured with

448   mean absolute error (MAE, in %). These errors were examined with respect to model

449   uncertainty (MU, in %), which quantified the stochasticity of the corresponding trend,

450   expressed as the variability of simulated quantities over the 1000 simulations. MU was

451   computed as the mean absolute deviation of the simulated activities to rate the model spread,

452   expressed in % of the observed value. The last metric used to evaluate the model was the

453   coverage probability (CP) of the 95% confidence interval, which measured how often

454   observations fall within the estimated confidence interval. The CP of a perfect model is exactly

455   equal to 95 %. A coverage significantly different from 95% means that the model is either

456   biased or exhibits an incorrect variability.

457

458   *2.3. Model applications*

459   Once the model has been evaluated, it can be used to analyze the stochasticity in fire activity,

460   given that 1000 replicates of the models can provide more insight than the single realization of

461   observations. Two example applications were developed in the present paper.

462   ***Detailed analysis of year 2003***. For the first application, we provided detailed comparisons of

463   seasonal predictions and observations and fire size distribution in 2003, during which the total

464   burnt area was extremely high for study area.

465   ***Predictability analysis***. In order to better understand the predictability of fire activity, we

466   compared simulated fire activity for years 2015-2018 to observations (validation sample) within

467   a crossover plan of spatio-temporal aggregations, and for a selection of fire sizes. We used six

468   classes for spatial aggregation (ranging from the single 8-km pixel to the whole studied area)

469   that were crossed with seven classes for time aggregation (ranging from a single day to the four

470   2015-2018 years). MAE and MU, as defined in the previous section, were computed based on

471   the 1000 simulations for the corresponding 42 aggregation classes, for fire numbers ranging

472   between 1 and 500 ha and for total burnt area. The predictability was analyzed by comparing

473   the model uncertainty and prediction errors (i.e., expected minus observed value). When both

474   were of the same order of magnitude, the model correctly represented the stochasticity at play.

475   Hence, low uncertainty and error indicated a high predictability, whereas high uncertainty and

476   error revealed low predictability. Uncertainty lower than error meant that a bias was present in

477   the model predictions, making predictability assessment tricky if the bias was not constant, even

478   if a high model uncertainty likely indicated low predictability. This approach allowed to

479   diagnose the fire sizes and aggregation scales for which simulations were in agreement with

480   observations.

481

482   **3. RESULTS**

483   ***3.1. Presentation of the "Full" model***

484   ***Partial effects of the "Full" model.*** The partial effects of both components (occurrence, size)

485   of the "full" fire activity model for the different explanatory variables on escaped fire numbers

486   and on a selection of exceedance probabilities are shown in Fig. 3. The 95th credible intervals

487   were obtained from the posterior predictive distributions. As expected, the FWI had a strong

488   effect on the expected number of escaped fires, which was about 60 times higher for a FWI of

489   60 than one of 5 (Fig. 3a). This effect was however marginal for FWI above 60, with wider

490   credible intervals, due to smaller sample size for the most extreme values. We observed a

491   positive effect of forest area (FA, including both f_{FA} and the offset, see Eq. 1) on the expected

492   number of escaped fires with a maximum around 30%. The slight decrease starting at around

493   40 % reflected a strong decrease in escaped fire density (number of fires per unit of forest area)

494   observed in pixels with the highest FA.

495   The partial effect of the season showed a constant increase between mid-June and the end of

496   August, and then decreased during autumn. Even if the magnitude of this effect was moderate,

497   it indicated that the FWI was not fully consistent to rate escaped fire occurrence over the course

498   of the fire season. For instance, for a same fire danger level, escaped fires were 1.6 times more

499   numerous in late August than in mid-June. The spatial effect was much stronger in magnitude,

500   indicating that very different fire activities were associated with the same FWI level at different

501   locations. The last effect was the "post-2003" effect (Eq. 1), which was equal to 0.46 in

502   posterior mean, meaning that the number of fires was roughly reduced by half after 2003, with

503   a high statistical significance. The transition between the two periods will be further analyzed

504   and discussed below. Among these different effects, the FWI and -in a lesser extent- the spatial

505   model exhibited the strongest magnitudes.

506

507   The effects of FWI and FA on the size of escaped fires showed that the probability to exceed a

508   given size generally increased with both explanatory variables. Moreover, these exceedance

509   probabilities decreased when larger fire size were considered, as expected (Fig. 3b). For

510   example, the probability to exceed 10 ha in a pixel with FA=30 % increased from 0.069 for a

511   FWI of 7.3 to 0.325 for a FWI of 64. Also, the probability to exceed 500 ha could be larger than

512   to exceed 100 ha, depending on the value of FWI, which illustrates that fire size is strongly

513   impacted by this index. It should be noted, however, that the magnitude of these effects was

514   generally much smaller in the size than in the occurrence model.

515   Surprisingly, the exceedance probability decreased for fires larger than 2000 ha at higher FWI,

516   with however little significance because of small sample issues. We further point out (see last

517   subsection of 3.2) that highest FWI values were not equally distributed in space but were most

518   often observed in areas less prone to large fires (e.g. coastal populated areas where suppression

519   is high). The surprising decrease could hence be explained by a confounding spatial effect. The

520   use of a spatial model -as in the occurrence model, see Fig. 9 for details- in the size model could

521   have dampened the impact of missing spatial factors, but the dataset was too small to afford it

522   for the size model. More surprising was the moderate decrease observed for FWI lower than

523   10. In the dataset, a non-negligible number of medium and large fires was recorded for very

524   low FWI (<5), with 60 fires larger than 10 ha occurring with FWI lower than 5. For example,

525   three large fires (936, 2369 and 4378 ha) occurred in 2003 when FWI was lower than 1, 2 and

526   4, respectively. A range of factors might explain this, including the development of burnt area

527   of these fires on the days following ignition (for which the FWI at ignition day is a not relevant

528   fire danger metric for size), by sub-daily scale events (e.g. afternoon thunderstorms following

529   fire events inducing rapid change in FWI during the ignition day), by uncertainties in the

530   weather reanalysis (SAFRAN), or simply, poor rating of actual fire danger conditions by the

531   FWI.

532

533 ***Example simulated scenarios of fire activity with the "Full" fire activity model.*** As an

534 illustration of the model practical utility, fire activity simulations aggregated for the whole zone

535 were compared to historical observations (black dots) at daily (for escaped fires only) or weekly

536 scales (escaped fires, fire number larger than 10, 50 and 100 ha, as well as burnt areas). Results

537 are shown in Fig. 4 for the example year 2001, but similar figures for other years of the study

538 period are available in Appendix S5. As expected, the uncertainty due to stochasticity (MU, in

539 %) was larger for daily than for weekly escaped fires, and tended to increase with the fire size

540 of interest, partly because the numbers to predict were smaller. Although not exactly equal to

541 95%, coverage probabilities (CP) were of the right order of magnitude, even when the width of

542 the confidence intervals was fairly narrow (e.g. weekly escaped fire number, CP=70%). MAE

543 were most often slightly larger than MU, but on the same order, which illustrated model skills,

544 despite high stochasticity in the data.

545 Next, to study in detail the ability of our modeling framework to reproduce observed patterns

546 of fire size distribution, simulated cumulative distributions of fire size were compared to

547 observations for the same example year in Fig. 5. Although observations may deviate from

548 expectations for the largest fires, most exceedance probabilities fell into the simulation-based

549 95th confidence interval. Note that in this example, the simulated trend for 2001 was close to

550 the mean simulation (orange dotted line, for year 1995-2018), but this was not the case in

551 general (e.g. years 1997 or 2002, see Appendix S6 for details).

552

553 *3.2. Model evaluation and importance of each explanatory variable*

554 ***Variable selection and model fits.*** Model information criteria (DIC, WAIC), and AUCs for

555 years 1995-2014 (training sample) and 2015-2018 (validation sample), are systematically

556 reported in Tables of Appendix S2. A selection of these AUCs is presented in Fig. 6. The

557   performance of the "full" occurrence component was high (>0.8) on both training and

558   validation subsets, and better than the simple FWI-linear model. Regarding the size component

559   model, the predictability of medium fire sizes (50 to 500 ha) was highest, with AUC > 0.75.

560   AUCs were in general on the same order for the validation (2015-2018) and the training sample

561   (1995-2014), showing the encouraging performance of the model beyond the training sample.

562   ***Evaluation of central tendencies and spatial patterns.*** In order to more comprehensively

563   evaluate the model, fire activity simulations were compared to the historical observations for

564   spatially aggregated annual and seasonal data as well as temporally aggregated data at the pixel

565   level in Fig. 7 (escaped fire numbers) and Fig. 8 (burnt areas). Orange lines and left maps

566   ("Full" model) compared generally well to observations, contrary to blue lines and right maps,

567   which correspond to different intermediate models, with less explanatory variables. This

568   demonstrates the absence of major bias of the "Full" model (metrics in orange), as well as the

569   limitations of intermediate models (metrics in blue).

570   In particular, annual trends in escaped fire number (Fig. 7a) were poorly predicted without the

571   "Post-2003" effect, with a CP of only 4.2 % and a MAE of 40 %. The systematic overestimation

572   of fire activity after 2003 with this model clearly indicates that the fire-weather relationship

573   changed over time. The "Full" model performed much better than the intermediate one but

574   remained still slightly biased and did not fully accounted for the evolution of the fire-climate

575   relationship over years, or underestimated confidence intervals (CP of 58%, lower than 95%).

576   In particular, we note that the "Full" model underestimated escaped fire numbers in years 2004-

577   2007, suggesting that the transition was probably less abrupt in observations than assumed with

578   a single fixed "Post-2003" effect. Several factors could explain this strong evolution near 2003.

579   This includes the evolutions in fire management after 2003 (modernization of the fire

580   management law, increase in airborne armed-guard funding), as well as the increase in the

581   precision of fire size recordings that has decreased the proportion of fire sizes above 1 ha in

2

582   records, which was marked near 2003. In the end, we considered the "Full" model satisfactory,

583   because tendencies for recent years -and especially those simulated after 2015 (validation

584   sample)- were in good agreement with historical observations.

585   Although moderate, the seasonal correction at the weekly scale enabled to match observations

586   closely, with a CP of 87% (Fig. 7b), which was fairly close to 95 %. The trends observed with

587   the "No seasonal" model showed that FWI explained a large part of the seasonal dynamics, but

588   that the escaped fire number was overestimated until the end of July and underestimated in late

589   August, which was consistent with the partial effect of the Week of Year shown in Fig. 3. When

590   the spatial model component was not included ("No spatial"), the occurrence component did

591   not simulate the spatial patterns of escaped fires well (Fig. 7c, right). In particular, hot spots

592   were missed, whereas the model predicted too many fires in the Alps and in the Camargue

593   region (coastal plain in the Rhône valley).

594   In general, burnt areas exhibited similar results (Fig. 8), albeit with notable differences. First,

595   the observed burnt area in 2003 was strongly underestimated, and was way above the upper

596   bounds of confidence intervals (Fig. 8a). This important point will be further analyzed in section

597   3.4. Second, the confidence interval widths, which expressed the amount of stochasticity, were

598   much larger for burnt areas than for escaped fire numbers, with model uncertainties on the order

599   of 35 % for both annual and seasonal predictions. Although such a high stochasticity was

600   expected due to the flat tail of the fire size distribution, one could argue that the randomness

601   was overestimated by the model. Two main clues indicated that it was not the case. Although

602   no obvious bias was evident in temporal and spatial trends, CP were on the order of 75-80 %,

603   which was not too far from the target (95 %), suggesting that stochasticity was on the right

604   order of magnitude. Moreover, observed mean weekly burnt areas (averaged over 1995-2018)

605   exhibited large fluctuations between consecutive weeks (between weeks 28 and 36, especially

606   between the last week of July and the first week of August: 13.6 km^2 for week 30; 4.0 km^2 for

607   week 31), whereas no obvious mechanisms except randomness could explain such a behavior

608   (Fig. 8b). The magnitude of these fluctuations seems to be consistent with the confident

609   intervals predicted by the model.

610   Expected temporal trends and spatial patterns in simulations looked like a smoothed expression

611   of observations, in which stochasticity would have been removed. However, one should notice

612   a few spurious differences. Simulations seemed to slightly overestimate burnt areas at the

613   beginning and the end of the season, and to exhibit less burnt areas than observed in the Corsica

614   and Var NUTS3 units, where most of the very large fires of the 2003 season occurred.

615   Interestingly, simulated burnt areas were only slightly better predicted when they were

616   simulated from escaped fire observations (using the fire size component only, see Appendix

617   S4: Fig. S1). This revealed that the limitation in burnt area simulations mostly arose from the

618   fire size model and that the full occurrence model performed well.

619   ***Sensitivity of response functions to explanatory variable selection.*** Beyond their limited

620   ability to reproduce observations, intermediate models also revealed that the accuracy and shape

621   of response functions could also be greatly impacted by modeling choices and the non-inclusion

622   of some key effects. The response function of FWI to escaped fire number for the intermediate

623   models were both limited in magnitude and exhibited spurious decreases, when compared to

624   the "Full" model (Fig. 9). In particular, the "Linear-FWI" model (for which the log number of

625   escaped fires has a linear response to FWI) was penalized by both low and high FWI, for which

626   the actual response to FWI was respectively stronger and lower than exponential. For other

627   intermediate models, the decrease observed above the FWI level of 65 could be explained by

628   confounding effects between FWI and space. Indeed, highest FWI values mostly occurred in

629   coastal populated areas where fire density was lower (at constant FWI). For the "Full" model,

630   a small decrease was also observed at 65, but its magnitude was much smaller and was followed

631   by slight increase above 70, thanks to the spatial model that considerably limited the impact of

632   the confounding effect. Hence, the inclusion of the cofactor (FA) and spatio-temporal

633   components in the occurrence model enabled to extract valuable information from FWI relative

634   variations, than what is available in absolute values in this fire danger index.

635

636   *3.3. Insights on an extreme year: the example of 2003*

637   Here, we examine in detail why the model underestimated burnt area for the year 2003. Because

638   most of the burnt area was caused by a small number of large fires, one might hypothesize that

639   i) observed burnt area for 2003 was unlikely considering fire weather, but possible with a low

640   probability ("bad luck"); ii) most fire observations were expected (i.e. consistent with the usual

641   fire-weather relationship), but the occurrence of a few very large fires that disproportionately

642   contributed to the total burnt area, leading to an underprediction of the total burnt area with the

643   model. In this context, arson is sometimes mentioned (and an arsonist was indeed involved in

644   a limited number of large fires in Var NUTS 3 division).

645   According to Fig. 8, the central tendency and the upper bound of the 95th CI (0.975 quantile)

646   for burnt areas predicted by the model were respectively of 213 and 357 km^2, which was well

647   below observations (610 km^2). Similarly, the quantile 0.999, corresponding to a millennial

648   event according to the model was 469 km^2, still below observations. We can then conclude that

649   the model failed to simulate the likelihood of observed burnt area in 2003. We then analyzed

650   time series corresponding to 2003 seasonal fire activities in Fig. 10 (similar to Fig. 4, but with

651   99.9th confidence intervals added to show unlikely events). The predictions of escaped fires

652   were consistent with observations at both the daily and weekly scales, which shows that fires

653   did not escape more frequently than expected all along the season (without any exceptional

654   week or day). However, more than six weeks were largely above the central tendency in

655   numbers of fires larger than 100 ha and in burnt areas. Four weeks were above the 0.975

656   quantile and one was even above the 0.9995 quantile (week 35). This shows that 2003 was

657   atypical during several weeks, with an early start. The analysis of the distribution of fire size

658   (Fig. 11), showed that all fires larger than 10 ha occurred much more often than expected.

659   Hence, even if the presence of a few very large fire constituted most of the burnt area, most

660   fires (larger than 10 ha and during most of the season) were exceptionally large with respect to

661   observed FWI, which invalidates the "bad luck" or arson assumptions. The decrease in fire

662   suppression efficiency with increasing escaped fire number can be invoked to explain the 2003

663   observations as well. However, it is important to acknowledge that it had apparently not affected

664   the number of escaped fires, weakening this assumption, as their number was consistent with

665   expectations.

666

667   **3.4. Predictability of fire activity**

668   We proposed a detailed analysis of the predictability of fire activities at various temporal and

669   spatial aggregation scales to better understand the role of stochasticity in fire activity patterns.

670   In general, stochasticity in observations (fire counts and sizes) typically decreases when

671   aggregating them to larger scales, such that the nature of both observations and model

672   predictions becomes more deterministic. Slight biases of models arising at the voxel scale may

673   then lead to stronger biases at aggregated scales.

674   ***Predictability and confidence intervals at weekly scale.*** As shown in Fig. 4, aggregating fire

675   activity over the whole area at the weekly scale led to reasonable confidence intervals and MAE

676   for year 2001. More generally, the overall predictability of escaped fires at the weekly scale

677   was satisfactory for the whole period (1995-2018), with a MAE of 32% and a CP of 84%

678   (Appendix S3: Fig. S1A). However, the number of escaped fires out of confidence interval

679   (16%) was slightly larger than expected (5%). The majority of these weeks consisted in false

680   "high" fire numbers for small observed numbers, and false "low" fire numbers for high

681   observed numbers. They were explained by stochasticity (fortunate and unfortunate events),

2

682   Bayesian model smoothness, and over-dispersion of observations with respect to aggregated

683   Poisson laws (caused by the decrease in fire suppression efficiency with fire activity for

684   example). More explanations can be found in Appendix S3. Regarding weekly burnt areas, the

685   central tendencies (Appendix S3: Fig. S1B) were positively correlated with observed burnt area

686   for 1995-2018, but the MAE was high (76 %), mostly because of the width of the confidence

687   intervals (MU=91 %). This illustrates the huge role of stochasticity in observed burnt areas at

688   this scale. Most false "high" burnt areas occurred when observed occurrences were lower than

689   expected, whereas false "low" burnt areas mostly occurred in 2003. More details are available

690   in Appendix S3.

691

692   ***Predictability at other scales.*** The MAE and MU are presented in Fig. 12 for 42 spatio-temporal

693   levels of aggregation, ranging from one pixel-day to the whole area during the four years of the

694   validation sample. As expected, the MAE increased for smaller aggregations, and large fire

695   numbers and burnt areas were more uncertain than escaped fire number (higher MU). Beyond

696   this general trend, Fig. 12 allows us to identify which scales led to reasonable predictions

697   (typically, those with MAE -and MU- lower than 30 %) and which were, on the contrary,

698   subject to too much stochasticity for valuable predictions (typically, MAE and MU both above

699   60-70 %). In particular, the spatial aggregation drastically reduced the MAE, while sub-regional

700   predictions remained quite poor, even for escaped fire number, when predictions were made at

701   a shorter scale than the full season. This could be partly explained by the fact that spatial

702   patterns of ignitions have slightly evolved over the 24 years of the study period, resulting in a

703   small spatial bias of the model during recent years. This was suggested by the pattern followed

704   by MU, which was slightly smoother regarding the effect of spatial aggregation and by the

705   lower AUC of the occurrence model in recent years (Appendix S2: Table S1). In particular, less

706   fire activity than expected was observed in North Corsica and more escaped fires in the western

707   part of the basin during the recent years (Not shown). It should also be noted that the fire activity

708   was fairly limited during the validation period, which increased the magnitude of relative errors,

709   as model uncertainty becomes larger when central tendency decreases. It is hence expected that

710   the predictability of a larger number of fire events would be more accurate (even if the model

711   failed to predict 2003).

712

713   **4. DISCUSSION**

714   ***Predictability of fire activity.*** The wildfire phenomenon results from multiple interactions

715   between biophysical and human factors acting at various spatial and temporal scales, which

716   spawn individual events with a high degree of stochasticity. *Firelihood* is a probabilistic model

717   of regional fire activity that simulates replications of individual fires with their size (on a daily

718   basis and at 8km-pixel resolution) with a reasonable accuracy (for example see coverage

719   probabilities and AUC values in Fig. 4 and 5, respectively), thereby offering the opportunity to

720   study the predictability of fire activity, i.e. to determine to what extent observed patterns can

721   be deterministically predicted. This analysis was carried out by comparing the model

722   uncertainty and prediction errors (i.e., expected minus observed value) on data held out during

723   the estimation of the model.

724   Our results showed that the stochasticity of fire activity was quite high, especially for large fire

725   sizes, which exhibited large model uncertainties and equally large errors, unless aggregating

726   fire data at larger temporal and/or spatial scales (Fig. 12). The number of fires larger than small

727   thresholds (1 to 10 ha) were fairly deterministic for the whole study area on a weekly basis and

728   could be well predicted for years following the training period with limited uncertainties and

729   errors. However, the predictability of sub-regional fire activity decreased rapidly at finer spatial

730   scales, with model errors increasing faster than model uncertainties (Fig. 12). Refined analyses

731   are required to better understand the differences between spatial and temporal aggregations in

3

732   terms of predictability, but this suggests that there are prospects for improving the model spatial

733   predictions. Preliminary analysis suggested that time-variations in spatial effects would explain

734   this decrease in model performance.

735   Overall, our results highlighted that stochasticity was a major component of observed fire

736   activity, so that individual events, pixels, years or fire size class, etc. are often far from

737   deterministic. Hence, even long-term (>20 years) fire datasets that are used either for ecological

738   or operational purposes should be used with caution since they do not allow capturing the full

739   variability in fire disturbance. This major point is illustrated for instance by the high degree of

740   stochasticity observed in weekly burnt areas in Fig. 8b. Moreover, observed monthly and annual

741   burnt areas, which are often used as benchmarks in a number of studies, namely for damage

742   assessment exercises or for comparing efficiency of fire suppression strategies, were also shown

743   to be highly random. We recommend to use more robust metrics such as number of escaped

744   fires or fires larger than a given size threshold, instead (e.g. $\geq$ 100 ha in Southeastern France).

745   ***Factors controlling fire activity.*** The FWI and the forest area were included as explanatory

746   variables in the model to represent where and when weather conditions are conducive to fires,

747   and how much area is available for fire spread, respectively. Weather conditions are major

748   drivers of fire activity in Southeastern France and our results showed that FWI is an effective

749   metric to rate fire danger (e.g. Fig. 3a and 3b), as suggested by earlier studies based on FWI

750   subcomponents in the same region (Ruffault et al. 2016, Fréjaville and Curt 2017, Barbero et

751   al. 2019), and in accordance with studies using the FWI in other southern Mediterranean

752   countries (e.g. Padilla and Vega-Garcia 2011, Fernandes 2018, Dupuy et al. 2020). However,

753   our study also highlighted several limitations of the FWI. First and foremost, our results suggest

754   an inconsistent rating of actual fire danger by the FWI for different fire weather types (Ruffault

755   et al. 2020). Specifically, the relative sensitivity of the index to ranges of wind speed on the one

756   hand, and of temperature and drought on the other hand, might not be properly scaled in

757   southeastern France. Indeed, the FWI has an exponential response to wind, which might not be

758   the case of fire activity for high wind values, as suggested by the saturation of the FWI effect

759   in the fire occurrence (FWI > 60) and fire size (FWI > 40) models (Fig. 3), even if the saturation

760   was partly explained by spatial bias (Fig. 9). The FWI could overestimate actual fire danger

761   during these high wind days. By contrast, our results showed a systematic underestimation of

762   the numbers of fires with size larger than 10 ha during the whole fire season during the 2003

763   heatwave (Fig. 10 and 11), suggesting that the explanatory variable for fire danger (FWI) could

764   underestimate the actual fire danger during hot drought conditions. Another weakness of the

765   FWI could be its time lag with respect to seasonal fire observations. Indeed, with the version of

766   the probabilistic model without week correction (model "No seasonal" in Table 1), the fire

767   season predicted by the model starts and ends two to three weeks earlier than the observations

768   (Fig. 5b and 6b, blue line). Such a lag has already been reported for the KBDI in the

769   Mediterranean (Ganatsas et al. 2011), but never for the FWI to our knowledge. This could be

770   explained by the desiccation processes empirically modelled in the Drought Code (a

771   subcomponent of the FWI reflecting monthly variations in fuel moisture content), which only

772   poorly explains Live Fuel Moisture Content (Ruffault et al. 2018b), the latter being increasingly

773   recognized as an important factor of fire behavior (e.g. Pimont et al. 2019). The seasonality -

774   and potentially spatial patterns- of fire danger rating could be improved by the use of more

775   mechanistic models for fuel moisture assessment. Live fuel moisture content dynamics depends

776   on the processes of the water (soil water uptake, plant water storage and transpiration) and the

777   carbon (photosynthesis, respiration, carbon allocation and canopy phenology) cycles (Jolly and

778   Johnson 2018). Key water processes during an extreme drought and heat wave can be modeled

779   according to plant hydraulics, depending on plant traits (e.g. Martin-StPaul et al. 2020). As far

780   as the carbon cycle is concerned, the evolution of fuel moisture linked to the production of new

781   shoots could be taken into account thanks to tree phenological models (e.g. Vitasse et al. 2011).

782

783   Our results also confirmed that forest area is an important factor of spatial variations of fire

784   activity in southeastern France (Ruffault and Mouillot 2017, Ganteaume and Barbero 2019),

785   with less escaped fires in pixels with low forest area (typically below 20 %) than those with

786   moderate to high forest areas (typically higher than 40 %) (Fig. 3). A saturation effect or a slight

787   decrease was however observed above 40 %. This pattern is the result of two opposite

788   mechanisms explicitly accounted for in our model. First, fire numbers in each pixel are limited

789   by the area in which ignition "points" can occur ("offset" effect). Second, fire density (i.e.

790   number of fires per unit of forest area, f_{FA}) decreased with increasing forest area because human

791   activities, which are responsible for most fire ignitions in human-dominated landscapes

792   (Syphard et al. 2007, Hawbaker et al. 2013, Costafreda-Aumedes et al. 2018), are more limited

793   in these gridcells. This confirms that fire activity can be ignition-limited when forest area is

794   high but fuel-area-limited when forest area is low; the resulting effect of these two opposites

795   being likely due to some specificities of the studied area and scale-dependent (Parisien et al.

796   2011).

797   ***Methodological insights to fire activity modelling.*** One of the critical aspects of this work was

798   the determination of the appropriate voxel size. Typically, smaller voxels lead to issues related

799   to both data reliability, as well as computational and memory costs, but larger voxels result in

800   information loss, leading to a challenging trade-off. For example, we aimed at accounting for

801   fine spatio-temporal variations in fire weather, but had to ignore hourly variations in FWI and

802   spatial variations smaller than 8 km. Depending on the objective of a study and the specificity

803   of each region, an appropriate voxel size should be adapted, but the approach of the present

804   study is applicable with a sophisticated Bayesian method to fine scales at a reasonable cost. To

805   bypass prohibitively high computation costs, one critical device in our study was the

806   aggregation of escaped fire counts in classes of same predictor configurations. This aggregation

807   led to a reduction of the dataset's effective size by a factor close to 10, allowing to avoid the

808   resampling of non-fire observations, which can affect the accuracy of partial effect estimates

809   (Woolford et al. 2011).

810   Our study also confirmed the importance of accounting for non-linearities in FWI and co-

811   factors for fire activity modeling, in accordance with Woolford et al. (2011). We found the FWI

812   effect was either underestimated or inconsistent when non-linearities/co-factors were ignored

813   (Fig. 9). In particular, the use of the spatial model was critical to properly address fine scale

814   data in which spatial autocorrelation is present. This spatial model was only involved in the fire

815   occurrence component -not in the fire size, which simply included FWI and FA as explanatory

816   variables-, which was sufficient to obtain a general agreement with burnt area observations.

817   Although 1 ha fires (escaped fires) only represent a minor part of cumulated burnt areas, this

818   finding highlights their importance in explaining patterns of burnt areas. Regarding the fire size

819   component, an innovative aspect of the present work was the piecewise modelling framework

820   for fire size distribution, which allowed to model its dependency to explanatory variables with

821   flexible nonlinear response functions.

822   The ensemble approach allowed us to compute central tendencies, but also confidence intervals

823   and return intervals of specific events. This approach is hence purely probabilistic; in particular,

824   probabilistic statements about the uncertainty of specific components of the model (*e.g.*, the

825   response functions) are possible, and the goodness-of-fit of models can be formally compared

826   through probabilistic information criteria. These ensemble simulations can be evaluated against

827   observations thanks to coverage probabilities of the confidence intervals (Joseph et al. 2019).

828   ***Model limitations.*** A first limitation of the present approach was the limited number of Land

829   Use and Land Cover (LULC) predictors included in both models (Forest Area only). LULC

830   predictors are increasingly used in statistical models of fire activity, especially when intended

831   for long-term predictions (Costafreda-Aumedes et al. 2017). They usually relate to the presence

832   of human settlements through variables such as the presence of roads, urban land-uses and

833   housing or population densities, or to that of human activities with variables encompassing the

834   number of farms or that of touristic infrastructure. Other models also include variables related

835   to landscape composition, such as the amount of shrubland (compared to high forest) or the

836   proportion of coniferous trees, or to landscape structure, such as wildland-urban interfaces,

837   wildland-agriculture interfaces or landscape fragmentation metrics. The influence of such

838   variables on fire activity has been documented in France (Ager et al. 2014; Opitz et al. 2020;

839   Curt et al. 2016; Ruffault and Mouillot 2017; Ganteaume and Jappiot 2013) and other

840   Mediterranean countries (Moreira et al. 2011; Oliveira et al. 2012; Gallardo et al. 2016; Nunes

841   et al. 2016; Vilar et al. 2016; Papakosta and Straub 2017; Costafreda-Aumedes et al. 2017).

842   Other, less frequently encountered variables with similar effects include levels of preparedness

843   (Podschwit and Cullen 2020) or of unemployment and economic welfare metrics (Mancini et

844   al. 2018; González-Olabarria et al. 2015).

845   In the present study, we did not explicitly account for those factors, but relied on the SPDE

846   approach to model spatial variations in escaped fire density. We do not expect large

847   improvement from the inclusion of LULC variables (which are almost static over a decade) in

848   the occurrence model performance, but it would increase the genericity of the model, for a

849   better understanding of effective factors, or for extrapolated predictions to other areas of

850   Southern France where reliable calibration data are not available, for example. Larger benefits

851   could be expected from the inclusion of LULC variables into our fire size model in which no

852   spatial effect was included because of the small amount of data available. For example, large

853   fires are expected to happen more often in shrubland-dominated areas and further away from

854   cities and communication axes (Moreira et al. 2011; Ager et al. 2014; Ganteaume and Jappiot

855   2013). This perspective is all the more important as our results suggested that the main

856   limitations of the "Full" model were rooted in this fire size component. Hence, going this

857   direction appears promising and the inclusion of both LULC and socio-economic predictors is

858   part of our future research agenda. Based on high-resolution rasterized datasets of standard

859   LULC variables, we propose to construct derived predictor variables (e.g., proportions of

860   LULC categories around the location of interest for different sizes of spatial buffers, or

861   proportions of co-occurrence of several categories such as forest cover and buildings in such

862   buffers to represent interfaces), which may have higher predictive power. Since this may lead

863   to a relatively large number of potentially useful but also strongly correlated variables,

864   preliminary analyses (prior to constructing the full stochastic model) may consist of variable

865   selection algorithms, or of dimension reduction steps in the spirit of principal component

866   analysis. The resulting set of variables, identified as possessing relevant predictive potential, is

867   then included into the occurrence and size model components, either in a linear or nonlinear

868   way as for predictor variables in the current model. This approach is expected to allow for

869   attribution of wildfire occurrences to risk factors enclosed in LULC and socio-economic

870   variables. Moreover, it presents a promising lead for improving predictions since LULC

871   variables are available at higher spatial and temporal resolution than the random spatial effect

872   accounting for bias terms in our current model.

873   One challenging aspect of fire activity modelling is related to the changes in the nature and

874   strength of the fire-weather relationship not only over space but also over time, which can be

875   due to various co-factors, including variations in LU-LC factors, in suppression means,

876   detection efficiency (Higuera et al. 2015, Ruffault and Mouillot 2015, Xi et al. 2019). In the

877   present study, we used a sophisticated approach for spatial but static effects, and a crude

878   approach to account for annual evolution of the relationship by simply modelling the abrupt

879   shift in the relationship occurring between 2003 and 2005 (using a single global, not spatially

880   resolved regression coefficient). We noticed that spatial patterns have slightly evolved over

881   time resulting in a slight decrease in AUC after 2015. These elements might call for a reduction

882   of the temporal period used in fire activity studies to increase fire-weather relationship

883   consistency, and hence prediction performance. However, our study also highlights the need

884   for long time series to include exceptional events such as 2003. The best approach would hence

885   be to allow for temporal variation of response functions and spatial bias components in the

886   models. In that regard, the use of a spatio-temporal log-Gaussian Cox process model for fire

887   occurrence is particularly promising to account for time-varying spatial effects (Serra et al.

888   2014b; Opitz et al. 2020); but its implementation on a daily basis still poses considerable

889   methodological challenges due to prohibitively high computational and numerical

890   requirements.

891   Some studies proposed that fire numbers could be better represented with a negative binomial

892   model than with a Poisson model (Marchal et al. 2017; Joseph et al. 2019) as it accounts for the

893   overdispersion of the data, i.e., the situation where variance in observed count values can be

894   higher than the intrinsic variance of the estimated model. This situation can arise due to a too

895   low space-time resolution of the model and explanatory variables (such that important

896   variability within a pixel-time step unit is not properly taken into account), of missing

897   explanatory variables, and of spatio-temporal effects unaccounted for in the model (e.g., time-

898   varying spatial effect). In particular, we identified in section 3.4 that the decrease in suppression

899   efficiency when escaped fire numbers increase could be involved. In this context, preliminary

900   investigations (where we replaced the Poisson distribution with a negative binomial one)

901   confirmed the presence of overdispersion with respect to the Poisson model, but simulations

902   from a model fitted with a negative binomial response showed unrealistically high values for

903   the number of escaped fires per voxels, with simulated counts being up to 10 times larger than

904   the maximum of observed numbers. This implies that the model performance would deteriorate

905   by capturing overdispersion at the pixel-day scale. Therefore, we used the Poisson-based model

906   as the more realistic model for all our analyses, but confidence intervals based on the aggregated

907   simulations can be too narrow in some cases (with CP lower than 95) due to an excess of "false"

908   extremely high and low fire activity. A more robust, but highly challenging approach would be

909   to identify the spatial and temporal scales where overdispersion arises, and to include some

910   corresponding components in the model.

911   ***Model applications and perspectives.*** Beyond the developments and example cases covered in

912   the present study, the *Firelihood* modeling framework offers a variety of promising research

913   applications. Our modeling framework can be used to further investigate the effects of

914   biophysical and human factors on fire dynamics and the variations of these effects over time

915   and space. The model is also able to compute and map the return intervals of fire disturbance,

916   which can cover longer time spans than observations, and be used to increase our understanding

917   of the interactions between wildfire disturbance and vegetation dynamics (Keeley et al. 2012).

918   Another research avenue for *Firelihood* is to run the model for retrospective evaluation (e.g.

919   attribution studies, Barbero et al. 2020) or the anticipation (e.g. Wotton et al. 2003; Fargeon et

920   al. 2019) of the effects of climate change on wildfire activity. Statistical models of fire activity

921   may also help gain insight into the socio-economic impacts of changing fire regimes and drive

922   insurance company strategic prospects. At local scales, they can be used to estimate and forecast

923   suppression costs (Preisler et al. 2011). At larger scales, economic activity and climate

924   mitigation in the forest sector may be affected by disturbances (Lindner et al. 2010; Seidl et al.

925   2014), as well as smoke production (McKenzie et al. 2014). Long-term forecasts often rely on

926   deterministic simulators where the inclusion of risk is a methodological challenge (Chudy et al.

927   2016, Riviere et al. 2020): *Firelihood* may provide a way to consider fire activity in such

928   assessments, provided that they include relevant LULC factors for management.

929   Operational outcomes include short-term forecasts of fire activity across the landscape

930   (Woolford et al. 2011). Because weather forecasts are now quite accurate for periods up to

931   seven days, the present modelling approach could be used to anticipate the number of escaped

3

932   fires and large fires (50-100ha) and to organize suppression means, since the predictability of

933   fire activity was encouraging at weekly regional scales. It should however be acknowledged

934   that uncertainties arising from weather predictions were not included in our study and might

935   degrade the prediction performance, so that additional testing is required. Although this weekly

936   regional scale is relevant for many operational purposes, finer resolutions are also of interest

937   for managers. Our modeling framework offers opportunity for managers to select the temporal

938   and spatial aggregation scales of interest. In the present study, we used the NUTS3 spatial scale,

939   but other can be used, such as the forest massif, which is often used for detailed preparedness.

940   Our predictability study showed that the predictive performance remained high at finer

941   temporal scales, so that daily predictions can be envisioned as a reasonable lead time. Regarding

942   spatial predictions however, we found that the predictive performance of the model tended to

943   decrease faster than expected when the aggregation level was finer than the region (on the

944   evaluation dataset, period 2015-2018, Fig. 12). Since mapping daily expected burnt areas (e.g.

945   at a 8km-resolution) or predicting fire numbers inside forest massifs are of major interest for

946   operational services, we think incorporating temporal trends in spatial effects (i.e. time-varying

947   effects) would be the most promising option to improve operational applications of our model.

948   Another promising application of such probabilistic approaches is to help identify changes in

949   fire-weather relationship over time, related to operational aspects. This includes detection

950   efficiency variation, changes of local regulations and public awareness strategies for risk

951   mitigation, as well as the evolution in strategies, tools and techniques for fire suppression (Xi

952   et al 2019).

953   Finally, the probabilistic approach of *Firelihood* can be applied to many other areas, provided

954   that i) fire data are available; ii) the factors controlling fire activity are identified and

955   measurable; iii) voxel sizes are adjusted following guidelines mentioned above. In particular,

956   this methodology can be used to model occurrence in mixed models, which combine a

957   probabilistic approach for occurrence with a mechanistic approach for fire size (Parisien et al.

958   2013; Finney et al. 2011). The mechanistic approach for fire size is based on the spread of fire

959   contours on a landscape and may better account for fuel load and continuity, topography and

960   land use and cover. However, such mixed approaches could be refined by building on the

961   strengths of the occurrence model presented here (e.g. SPDE approach for spatial variations)

962   and on their mechanistic model for fire spread.

963

964   **Conclusion**

965   This study proposed a comprehensive probabilistic framework, *Firelihood,* for modeling fire

966   activity in Southeastern France. The hierarchically-structured Bayesian model used a space-

967   time Poisson model for fire occurrence and a piecewise-estimated distribution for fire size,

968   which enables the simulation of likely spatial-temporal explicit fire activities. The Bayesian

969   approach allows an accurate estimation of random components in these sophisticated

970   hierarchical models, which can be parametrized in a convenient and interpretable setting thanks

971   to INLA. This ensemble-based methodology is innovative and applicable to spatially-correlated

972   fire observations in other regions or landscapes. In Southeastern France, the model performance

973   was very encouraging, especially for escaped fire numbers, and allowed to better understand

974   the role of stochasticity in fire activity. Further effort is needed to elucidate the fire outbreak

975   that occurred during the 2003 heat wave as well as the limitations of the FWI in fire size

976   estimates during such conditions. We identified and discussed a few methodological challenges,

977   including the time variations in spatial effects or the proper integration of overdispersion in

978   data. We also suggest a variety of ecological, operational and economic applications of

979   *Firelihood*.

980

981   **Author contributions:**

982   HF led an earlier manuscript describing the model development in her PhD dissertation

983   (Fargeon 2019), supervised by FP, JLD and NM. TO, HF and FP built the models and ran the

984   simulations with RINLA. FP, TO and JR wrote the present version of the manuscript with the

985   help of other co-authors.

991

992

References

994   Abatzoglou, J. T., and C. A. Kolden. 2013. Relationships between climate and macroscale area

995   burned in the western United States. International Journal of Wildland Fire 22:1003.

996   Ager, A. A., H. K. Preisler, B. Arca, D. Spano, and M. Salis. 2014. Wildfire risk estimation in

997   the Mediterranean area: MEDITERRANEAN WILDFIRE RISK ESTIMATION.

998   Environmetrics 25:384–396.

999   Ager, A. A., A. M. G. Barros, M. A. Day, H. K. Preisler, T. A. Spies, and J. Bolte. 2018.

1000   Analyzing fine-scale spatiotemporal drivers of wildfire in a forest landscape model.

1001   Ecological Modelling 384:87–102.

1002   Bakka, H., H. Rue, G.-A. Fuglstad, A. Riebler, D. Bolin, J. Illian, E. Krainski, D. Simpson, and

1003   F. Lindgren. 2018. Spatial modeling with R-INLA: A review. WIREs Computational

1004   Statistics 10:e1443.

1005   Barbero, R., J. T. Abatzoglou, E. A. Steel, and N. K Larkin. 2014. Modeling very large-fire

1006   occurrences over the continental United States from weather and climate forcing.

1007   Environmental Research Letters 9:124009.

1008   Barbero, R., J. T. Abatzoglou, N. K. Larkin, C. A. Kolden, and B. Stocks. 2015a. Climate change

1009   presents increased potential for very large fires in the contiguous United States.

1010   International Journal of Wildland Fire 24:892.

1011   Barbero, R., J. T. Abatzoglou, C. A. Kolden, K. C. Hegewisch, N. K. Larkin, and H. Podschwit.

1012   2015b. Multi-scalar influence of weather and climate on very large-fires in the Eastern

1013   United States: WEATHER, CLIMATE AND VERY LARGE-FIRES IN THE EASTERN

1014   UNITED STATES. International Journal of Climatology 35:2180–2186.

1015   Barbero, R., T. Curt, A. Ganteaume, E. Maillé, M. Jappiot, and A. Bellet. 2019. Simulating the

1016   effects of weather and climate on large wildfires in France. Natural Hazards and Earth

1017   System Sciences 19:441–454.

1018   Barbero, R., J. T. Abatzoglou, F. Pimont, J. Ruffault, and T. Curt. 2020. Attributing Increases

1019   in Fire Weather to Anthropogenic Climate Change Over France. Frontiers in Earth

1020   Science 8:104.

1021   Bedia, J., S. Herrera, and J. M. Gutiérrez. 2014. Assessing the predictability of fire occurrence

1022   and area burned across phytoclimatic regions in Spain. Natural Hazards and Earth System

1023   Science 14:53–66.

1024   Bowman, D. M. J. S., J. Balch, P. Artaxo, W. J. Bond, M. A. Cochrane, C. M. D'Antonio, R.

1025   DeFries, F. H. Johnston, J. E. Keeley, M. A. Krawchuk, C. A. Kull, M. Mack, M. A.

1026   Moritz, S. Pyne, C. I. Roos, A. C. Scott, N. S. Sodhi, and T. W. Swetnam. 2011. The

1027   human dimension of fire regimes on Earth: The human dimension of fire regimes on

1028   Earth. Journal of Biogeography 38:2223–2236.

1029   Bradstock, R. A. 2010. A biogeographic model of fire regimes in Australia: current and future

1030   implications: A biogeographic model of fire in Australia. Global Ecology and

1031   Biogeography 19:145–158.

1032   Brillinger, D. R., H. K. Preisler, and J. W. Benoit. 2003. Risk assessment: a forest fire example.

1033   Lecture Notes-Monograph Series:177–196.

1034   Calkin, D. E., M. P. Thompson, and M. A. Finney. 2015. Negative consequences of positive

1035   feedbacks in US wildfire management. Forest Ecosystems 2:9.

1036   Chudy, R. P., H. K. Sjølie, and B. Solberg. 2016. Incorporating risk in forest sector modeling

1037   – state of the art and promising paths for future research. Scandinavian Journal of Forest

1038   Research 31:719–727.

1039   Costafreda-Aumedes, S., C. Comas, and C. Vega-Garcia. 2017. Human-caused fire occurrence

1040   modelling in perspective: a review. International Journal of Wildland Fire 26:983.

1041   Cui, W., and A. H. Perera. 2008. What do we know about forest fire size distribution, and why

1042   is this knowledge useful for forest management? International Journal of Wildland Fire

1043   17:234.

1044   Curt, T., T. Fréjaville, and S. Lahaye. 2016. Modelling the spatial patterns of ignition causes

1045   and fire regime features in southern France: implications for fire prevention policy.

1046   International Journal of Wildland Fire 25:785.

1047   Curt, T., and T. Frejaville. 2018. Wildfire Policy in Mediterranean France: How Far is it

1048   Efficient and Sustainable?: Wildfire Policy in Mediterranean France. Risk Analysis

1049   38:472–488.

1050   Davison, A. C., and R. Huser. 2015. Statistics of Extremes. Annual Review of Statistics and Its

1051   Application 2:203–235.

1052   Deeming, J. E., R. E. Burgan, and J. D. Cohen. 1977. The National Fire-Danger Rating System-

1053   1978. Gen. Tech. Rep. INT-39. Ogden, UT: U.S. Department of Agriculture, Forest

1054   Service, Intermountain Forest and Range Experiment Station. 63 pp

1055   Dupuy, J., H. Fargeon, N. Martin-StPaul, F. Pimont, J. Ruffault, M. Guijarro, C. Hernando, J.

1056   Madrigal, and P. Fernandes. 2020. Climate change impact on future wildfire danger and

1057   activity in southern Europe: a review. Annals of Forest Science 77:35.

1058   Gallardo, M., I. Gómez, L. Vilar, J. Martínez-Vega, and M. P. Martín. 2016. Impacts of future

1059   land use/land cover on wildfire occurrence in the Madrid region (Spain). Regional

1060   Environmental Change 16:1047–1061.

1061   Ganatsas, P., M. Antonis, and T. Marianthi. 2011. Development of an adapted empirical

1062   drought index to the Mediterranean conditions for use in forestry. Agricultural and Forest

1063   Meteorology 151:241–250.

1064   Ganteaume, A., and M. Jappiot. 2013. What causes large fires in Southern France. Forest

1065   Ecology and Management 294:76–85.

1066   González-Olabarria, J. R., B. Mola-Yudego, and L. Coll. 2015. Different Factors for Different

1067   Causes: Analysis of the Spatial Aggregations of Fire Ignitions in Catalonia (Spain). Risk

1068   Analysis 35:1197–1209.

1069   Fargeon, H. 2019. Effet du changement climatique sur l'évolution de l'aléa incendie de forêt

1070   en France métropolitaine au XXI e siècle. Thèse de doctorat, Institut des sciences et

1071   industries du vivant et de l'environnement (AgroParisTech), INRA, UR 629 –Écologie

1072   des forêts méditerranéennes.

1073   Fawcett, T. 2006. An introduction to ROC analysis. Pattern Recognition Letters 27:861–874.

1074   Finney, M. A., C. W. McHugh, I. C. Grenfell, K. L. Riley, and K. C. Short. 2011. A

1075   simulation of probabilistic wildfire risk components for the continental United States.

1076   Stochastic Environmental Research and Risk Assessment 25:973–1000.

1077   Flannigan, M. D., M. A. Krawchuk, W. J. de Groot, B. M. Wotton, and L. M. Gowman. 2009.

1078   Implications of changing climate for global wildland fire. International Journal of

1079   Wildland Fire 18:483.

1080   Fréjaville, T., and T. Curt. 2015. Spatiotemporal patterns of changes in fire regime and climate:

1081   defining the pyroclimates of south-eastern France (Mediterranean Basin). Climatic Change

1082   129:239–251.

1083   Fréjaville, T., and T. Curt. 2017. Seasonal changes in the human alteration of fire regimes

1084   beyond the climate forcing. Environmental Research Letters 12:035006.

1085   Fuglstad, G.-A., and J. Beguin. 2018. Environmental mapping using Bayesian spatial modelling

1086   (INLA/SPDE): A reply to Huang et al. (2017). Science of The Total Environment 624:596–

1087   598.

1088   Ganteaume, A., and R. Barbero. 2019. Contrasting large fire activity in the French

1089   Mediterranean. Natural Hazards and Earth System Sciences 19:1055–1066.

1090   de Haan, L., and A. Ferreira. 2006. Extreme Value Theory. Springer New York, New York, NY.

1091   Hall, W. J. and J. Wellner. 1981. Mean residual life. In Statistics and related topics (Ottawa,

1092   Ont., 1980), pages 169–184. North-Holland, Amsterdam.

1093   Higuera, P. E., J. T. Abatzoglou, J. S. Littell, and P. Morgan. 2015. The Changing Strength

1094   and Nature of Fire-Climate Relationships in the Northern Rocky Mountains, U.S.A.,

1095   1902-2008. PLOS ONE 10:e0127563.

1096   Jolly, W. M., and D. M. Johnson. 2018. Pyro-Ecophysiology: Shifting the Paradigm of Live

1097   Wildland Fuel Research. Fire 1:8.

1098   Joseph, M. B., M. W. Rossi, N. P. Mietkiewicz, A. L. Mahood, M. E. Cattau, L. A. St.Denis, R.

1099   C. Nagy, V. Iglesias, J. T. Abatzoglou, and J. K. Balch. 2019. Spatiotemporal prediction

1100   of wildfire size extremes with Bayesian finite sample maxima. Ecological

1101   Applications:e01898.

1102   Krawchuk, M. A., M. A. Moritz, M.-A. Parisien, J. Van Dorn, and K. Hayhoe. 2009. Global

1103   Pyrogeography: the Current and Future Distribution of Wildfire. PLoS ONE 4:e5102.

1104   Keeley, J. E., W. J. Bond, R. A. Bradstock, J. G. Pausas, and P. W. Rundel. 2012. Fire in

1105   Mediterranean ecosystems—ecology, evolution and management. Cambridge University

1106   Press, Cambridge, UK.

1107   Lindgren, F., and H. Rue. 2015. Bayesian Spatial Modelling with *R* - **INLA**. Journal of

1108   Statistical Software 63.

1109   Lindgren, F., H. Rue, and J. Lindström. 2011. An explicit link between Gaussian fields and

1110   Gaussian Markov random fields: the stochastic partial differential equation approach:

1111   Link between Gaussian Fields and Gaussian Markov Random Fields. Journal of the

1112   Royal Statistical Society: Series B (Statistical Methodology) 73:423–498.

1113   Lindner, M., M. Maroschek, S. Netherer, A. Kremer, A. Barbati, J. Garcia-Gonzalo, R. Seidl,

1114   S. Delzon, P. Corona, M. Kolström, M. J. Lexer, and M. Marchetti. 2010. Climate change

1115   impacts, adaptive capacity, and vulnerability of European forest ecosystems. Forest

1116   Ecology and Management 259:698–709.

1117 Littell, J. S., D. McKenzie, D. L. Peterson, and A. L. Westerling. 2009. Climate and wildfire

1118   area burned in western U.S. ecoprovinces, 1916–2003. Ecological Applications 19:1003–

1119   1021.

1120 McKenzie, D., U. Shankar, R. E. Keane, E. N. Stavros, W. E. Heilman, D. G. Fox, and A. C.

1121   Riebau. 2014. Smoke consequences of new wildfire regimes driven by climate change.

1122   Earth's Future 2:35–59.

1123 Marchal, J., S. G. Cumming, and E. J. B. McIntire. 2017. Exploiting Poisson additivity to predict

1124   fire frequency from maps of fire weather and land cover in boreal forests of Québec,

1125   Canada. Ecography 40:200–209.

1126 Mancini, L. D., P. Corona, and L. Salvati. 2018. Ranking the importance of Wildfires' human

1127   drivers through a multi-model regression approach. Environmental Impact Assessment

1128   Review 72:177–186.

1129 Martin-StPaul, N., J. Ruffault, C. Blackmann, H. Cochard, M. De Cáceres, S. Delzon, J.

1130   Dupuy, H. Fargeon, L. Lamarque, M. Moreno, R. Parsons, F. Pimont, J. Ourcival, J.

1131   Torres-Ruiz, and J. Limousin. 2020. Modelling live fuel moisture content at leaf and

1132   canopy scale under extreme drought using a lumped plant hydraulic model. preprint,

1133   Ecology.

1134 Moreira, F., O. Viedma, M. Arianoutsou, T. Curt, N. Koutsias, E. Rigolot, A. Barbati, P.

1135   Corona, P. Vaz, G. Xanthopoulos, F. Mouillot, and E. Bilgili. 2011. Landscape – wildfire

1136   interactions in southern Europe: Implications for landscape management. Journal of

1137   Environmental Management 92:2389–2402.

1138 Moreira, F., D. Ascoli, H. Safford, M. A. Adams, J. M. Moreno, J. M. C. Pereira, F. X. Catry,

1139   J. Armesto, W. Bond, M. E. González, T. Curt, N. Koutsias, L. McCaw, O. Price, J. G.

1140　　　　Pausas, E. Rigolot, S. Stephens, C. Tavsanoglu, V. R. Vallejo, B. W. Van Wilgen, G.

1141　　　　Xanthopoulos, and P. M. Fernandes. 2020. Wildfire management in Mediterranean-type

1142　　　　regions: paradigm change needed. Environmental Research Letters 15:011001.

1143　Morgan, P., C.C. Hardy, T. Swetnam, M.G. Rollins MG, and D.G. Long DG. 2001. Mapping

1144　　　　fire regimes across time and space: Understanding coarse and fine-scale fire

1145　　　　patterns. International Journal of Wildland Fire 10: 329-342.

1146　Moritz, M. A., M. E. Morais, L. A. Summerell, J. M. Carlson, and J. Doyle. 2005. Wildfires,

1147　　　　complexity, and highly optimized tolerance. Proceedings of the National Academy of

1148　　　　Sciences 102:17912–17917.

1149　Nunes, A. N., L. Lourenço, and A. C. C. Meira. 2016. Exploring spatial patterns and drivers of

1150　　　　forest fires in Portugal (1980–2014). Science of The Total Environment 573:1190–1202.

1151　Noble, I. R., A. M. Gill, and G. A. V. Bary. 1980. McArthur's fire-danger meters expressed as

1152　　　　equations. Austral Ecology 5:201–203.

1153　Oliveira, S., F. Oehler, J. San-Miguel-Ayanz, A. Camia, and J. M. C. Pereira. 2012. Modeling

1154　　　　spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression

1155　　　　and Random Forest. Forest Ecology and Management 275:117–129.

1156　Opitz, T., F. Bonneu, and E. Gabriel. 2020. Point-process based Bayesian modeling of space–

1157　　　　time structures of forest fire occurrences in Mediterranean France. Spatial

1158　　　　Statistics:100429.

1159　Papakosta, P., and D. Straub. 2017. Probabilistic prediction of daily fire occurrence in the

1160　　　　Mediterranean with readily available spatio-temporal data. iForest - Biogeosciences and

1161　　　　Forestry 10:32–40.

1162　Parks, S. A., M.-A. Parisien, and C. Miller. 2012. Spatial bottom-up controls on fire likelihood

1163　　　　vary across western North America. Ecosphere 3:art12.

1164   Parisien, M.-A., S. A. Parks, M. A. Krawchuk, M. D. Flannigan, L. M. Bowman, and M. A.

1165   Moritz. 2011. Scale-dependent controls on the area burned in the boreal forest of Canada,

1166   1980–2005. Ecological Applications 21:789–805.

1167   Parisien, M.-A., G. R. Walker, J. M. Little, B. N. Simpson, X. Wang, and D. D. B. Perrakis.

1168   2013. Considerations for modeling burn probability across landscapes with steep

1169   environmental gradients: an example from the Columbia Mountains, Canada. Natural

1170   Hazards 66:439–462.

1171   Pimont, F., J. Ruffault, N. K. Martin-StPaul, and J.-L. Dupuy. 2019. Why is the effect of live

1172   fuel moisture content on fire rate of spread underestimated in field experiments in

1173   shrublands? International Journal of Wildland Fire 28:127.

1174   Podschwit, H., and A. Cullen. 2020. Patterns and trends in simultaneous wildfire activity in

1175   the United States from 1984 to 2015. International Journal of Wildland Fire.

1176   Preisler, H. K., D. R. Brillinger, R. E. Burgan, and J. W. Benoit. 2004. Probability based

1177   models for estimation of wildfire risk. International Journal of Wildland Fire 13:133.

1178   Preisler, H. K., and A. L. Westerling. 2007. Statistical Model for Forecasting Monthly Large

1179   Wildfire Events in Western United States. Journal of Applied Meteorology and

1180   Climatology 46:1020–1030.

1181   Preisler, H. K., S.-C. Chen, F. Fujioka, J. W. Benoit, and A. L. Westerling. 2008. Wildland fire

1182   probabilities estimated from weather model-deduced monthly mean fire danger indices.

1183   International Journal of Wildland Fire 17:305.

1184   Preisler, H. K., A. L. Westerling, K. M. Gebert, F. Munoz-Arriola, and T. P. Holmes. 2011.

1185   Spatially explicit forecasts of large wildland fire probability and suppression costs for

1186   California. International Journal of Wildland Fire 20:508.

1187   Riviere, M., S. Caurla, and P. Delacote. 2020. Evolving Integrated Models From Narrower

1188   Economic Tools: the Example of Forest Sector Models. Environmental Modeling &

1189   Assessment.

1190   Rue, H., S. Martino, and N. Chopin. 2009. Approximate Bayesian inference for latent

1191   Gaussian models by using integrated nested Laplace approximations. Journal of the

1192   Royal Statistical Society: Series B (Statistical Methodology) 71:319–392.

1193   Ruffault, J., and F. Mouillot. 2015. How a new fire-suppression policy can abruptly reshape the

1194   fire-weather relationship. Ecosphere 6:art199.

1195   Ruffault, J., V. Moron, R. M. Trigo, and T. Curt. 2016. Objective identification of multiple

1196   large fire climatologies: an application to a Mediterranean ecosystem. Environmental

1197   Research Letters 11:075006.

1198   Ruffault, J., and F. Mouillot. 2017. Contribution of human and biophysical factors to the

1199   spatial distribution of forest fire ignitions and large wildfires in a French Mediterranean

1200   region. International Journal of Wildland Fire 26:498.

1201   Ruffault, J., V. Moron, R. M. Trigo, and T. Curt. 2017. Daily synoptic conditions associated

1202   with large fire occurrence in Mediterranean France: evidence for a wind-driven fire

1203   regime: DAILY SYNOPTIC CONDITIONS ASSOCIATED WITH LARGE FIRE

1204   OCCURRENCE. International Journal of Climatology 37:524–533.

1205   Ruffault, J., T. Curt, N. K. Martin-StPaul, V. Moron, and R. M. Trigo. 2018a. Extreme

1206   wildfire events are linked to global-change-type droughts in the northern Mediterranean.

1207   Natural Hazards and Earth System Sciences 18:847–856.

1208   Ruffault, J., N. Martin-StPaul, F. Pimont, and J.-L. Dupuy. 2018b. How well do

1209   meteorological drought indices predict live fuel moisture content (LFMC)? An

1210   assessment for wildfire research and operations in Mediterranean ecosystems.

1211   Agricultural and Forest Meteorology 262:391–401.

1212 Schoenberg, F. P., R. Peng, Z. Huang, and P. Rundel. 2003. Detection of non-linearities in the

1213   dependence of burn area on fuel age and climatic variables. International Journal of

1214   Wildland Fire 12:1.

1215 Seidl, R., M.-J. Schelhaas, W. Rammer, and P. J. Verkerk. 2014. Increasing forest

1216   disturbances in Europe and their impact on carbon storage. Nature Climate Change

1217   4:806–810.

1218 Serra, L., M. Saez, P. Juan, D. Varga, and J. Mateu. 2014a. A spatio-temporal Poisson hurdle

1219   point process to model wildfires. Stochastic Environmental Research and Risk

1220   Assessment 28:1671–1684.

1221 Serra, L., M. Saez, J. Mateu, D. Varga, P. Juan, C. Díaz-Ávalos, and H. Rue. 2014b. Spatio-

1222   temporal log-Gaussian Cox processes for modelling wildfire occurrence: the case of

1223   Catalonia, 1994–2008. Environmental and Ecological Statistics 21:531–563.

1224 Stephens, S. L., and L. W. Ruth. 2005. FEDERAL FOREST-FIRE POLICY IN THE

1225   UNITED STATES. Ecological Applications 15:532–542.

1226 Syphard, A. D., V. C. Radeloff, J. E. Keeley, T. J. Hawbaker, M. K. Clayton, S. I. Stewart,

1227   and R. B. Hammer. 2007. HUMAN INFLUENCE ON CALIFORNIA FIRE REGIMES.

1228   Ecological Applications 17:1388–1402.

1229 Taylor, S. W., D. G. Woolford, C. B. Dean, and D. L. Martell. 2013. Wildfire Prediction to

1230   Inform Fire Management: Statistical Science Challenges. Statistical Science 28:586–615.

1231 Tedim, F., V. Leone, M. Amraoui, C. Bouillon, M. Coughlan, G. Delogu, P. Fernandes, C.

1232   Ferreira, S. McCaffrey, T. McGee, J. Parente, D. Paton, M. Pereira, L. Ribeiro, D. Viegas,

1233   and G. Xanthopoulos. 2018. Defining Extreme Wildfire Events: Difficulties, Challenges,

1234   and Impacts. Fire 1:9.

1235 Trigo, R. M. 2005. How exceptional was the early August 2003 heatwave in France?

1236   Geophysical Research Letters 32:L10701.

1237

1238   Turco, M., M. C. Llasat, A. Tudela, X. Castro, and A. Provenzale. 2013. Brief communication

1239   Decreasing fires in a Mediterranean region (1970&ndash;2010, NE Spain). Natural

1240   Hazards and Earth System Sciences 13:649–652.

1241   Turco, M., J. Bedia, F. Di Liberto, P. Fiorucci, J. von Hardenberg, N. Koutsias, M.-C. Llasat,

1242   F. Xystrakis, and A. Provenzale. 2016. Decreasing Fires in Mediterranean Europe. PLOS

1243   ONE 11:e0150663.

1244   Turco, M., J. von Hardenberg, A. AghaKouchak, M. C. Llasat, A. Provenzale, and R. M.

1245   Trigo. 2017. On the key role of droughts in the dynamics of summer fires in

1246   Mediterranean Europe. Scientific Reports 7:81.

1247   Turco, M., J. J. Rosa-Cánovas, J. Bedia, S. Jerez, J. P. Montávez, M. C. Llasat, and A.

1248   Provenzale. 2018. Exacerbated fires in Mediterranean Europe due to anthropogenic

1249   warming projected with non-stationary climate-fire models. Nature Communications 9.

1250   Turner, R. 2009. Point patterns of forest fire locations. Environmental and Ecological

1251   Statistics 16:197–223.

1252   Vehtari, A., A. Gelman, and J. Gabry. 2017. Practical Bayesian model evaluation using leave-

1253   one-out cross-validation and WAIC. Statistics and Computing 27:1413–1432.

1254   Vidal, J.-P., E. Martin, L. Franchistéguy, M. Baillon, and J.-M. Soubeyroux. 2010. A 50-year

1255   high-resolution atmospheric reanalysis over France with the Safran system. International

1256   Journal of Climatology 30:1627–1644.

1257   Vilar, L., Douglas. G. Woolford, D. L. Martell, and M. P. Martín. 2010. A model for

1258   predicting human-caused wildfire occurrence in the region of Madrid, Spain.

1259   International Journal of Wildland Fire 19:325.

1260   Vilar, L., I. Gómez, J. Martínez-Vega, P. Echavarría, D. Riaño, and M. P. Martín. 2016.

1261   Multitemporal Modelling of Socio-Economic Wildfire Drivers in Central Spain between

1262   the 1980s and the 2000s: Comparing Generalized Linear Models to Machine Learning

1263   Algorithms. PLOS ONE 11:e0161344.

1264   Vitasse, Y., C. François, N. Delpierre, E. Dufrêne, A. Kremer, I. Chuine, and S. Delzon. 2011.

1265   Assessing the effects of climate change on the phenology of European temperate trees.

1266   Agricultural and Forest Meteorology 151:969–980.

1267   Wagner, C. E. V. 1977. Conditions for the start and spread of crown fire. Canadian Journal of

1268   Forest Research 7:23–34.

1269   Wang, X., B. M. Wotton, A. S. Cantin, M.-A. Parisien, K. Anderson, B. Moore, and M. D.

1270   Flannigan. 2017. cffdrs: an R package for the Canadian Forest Fire Danger Rating

1271   System. Ecological Processes 6:5.

1272   Westerling, A. L., B. P. Bryant, H. K. Preisler, T. P. Holmes, H. G. Hidalgo, T. Das, and S. R.

1273   Shrestha. 2011. Climate change and growth scenarios for California wildfire. Climatic

1274   Change 109:445–463.

1275   Wood, S. N. 2006. Low-Rank Scale-Invariant Tensor Product Smooths for Generalized

1276   Additive Mixed Models. Biometrics 62:1025–1036.

1277   Woolford, D. G., D. R. Bellhouse, W. J. Braun, C. B. Dean, D. L. Martell, and J. Sun. (n.d.). A

1278   Spatio-temporal Model for People-Caused Forest Fire Occurrence in the Romeo Malette

1279   Forest. Journal of Environmental Statistics 2:2-16.

1280   Wotton, B. M., D. L. Martell, and K. A. Logan. (n.d.). Climate Change and People-Caused

1281   Forest Fire Occurrence in Ontario:21.

1282   Xi, D. D. Z., S. W. Taylor, D. G. Woolford, and C. B. Dean. 2019. Statistical Models of Key

1283   Components of Wildfire Risk. Annual Review of Statistics and Its Application 6:197–222.

1284

1285　**Tables**

1286　Table 1. List of probabilistic fire activity models (occurrence+size)

Probabilistic model name	Occurrence model	Size model
"Linear FWI"	Linear FWI	Linear FWI
"FWI only"	FWI	FWI
"No Seasonal"	FWI+2003+FA	FWI+FA
"No Spatial"	FWI+2003+FA+WEEK	FWI+FA
"Full"	FWI+2003+FA+WEEK+SPATIAL	FWI+FA

1287

1288

1289   Table 2. Fire occurrence component (number of escaped fires, larger than 1 ha)

Occurrence model component	Effects
Null	No explanatory variable
Linear FWI	The predictor of fire counts in the Poisson model is a linear function of FWI
FWI	The predictor of fire counts in the Poisson model is a non-linear function of FWI
FWI+2003	As above + include also a fixed effect to account for post-2003 difference
FWI+2003+FA	As above + include the offset associated with FA and the non-linear effect of FA
FWI+2003+FA+WEEK	As above + include a weekly-based seasonal correction for occurrence
FWI+2003+FA+WEEK+SPATIAL	Full occurrence model described in section 2.1. As above + spatial model

1290

1291

1292   Table 3. Fire size component (size of escaped fires)

Size model component	Effects
Null	No explanatory variable
Linear FWI	The predictor of fire-size-distribution parameters is a linear function of FWI
FWI	The predictor of fire-size-distribution parameters is a non-linear function of FWI
FWI+FA	As above + include a non-linear effect of FA[1]

1293   [1] Except for the GPD which is a linear function of FWI only

1294

1295 **Figure legends**

1296 FIG. 1. Data site and description: a) Orography (elevation in m) of the study region; Fire activity

1297 (1995-2018): b) Monthly distribution of burnt areas; c) Spatial distribution of fires larger or

1298 equal to 1 ha; d) Yearly burnt areas. Data were extracted from the Prométhée fire database

1299 (http://www.promethee.com/).

1300 FIG. 2. Framework of the "full" probabilistic model of fire activity

1301 FIG. 3. Partial effects of the "Full" fire activity model: a) Occurrence component (number of

1302 escaped fires, i.e. larger than 1 ha): effects of FWI, Forest Area (including offset), Week of

1303 Year, and location; b) Size component (exceedance probabilities of escaped fires for a selection

1304 of thresholds ranging between 1 and 2000 ha): effects of FWI and FA, for a Forest Area of 30 %

1305 (top) and a FWI of 20 (bottom), respectively.

1306 FIG. 4. Simulated fire activity (in orange) and observations (black dots) for year 2001: daily

1307 and weekly escaped fire numbers, as well as weekly number of fires larger than 10, 50, 100 ha

1308 and weekly burnt areas added up over the whole study area. Central tendency (red line) was

1309 surrounded by the 95th confidence interval in orange and was based on averages computed over

1310 1000 simulations of fire activities for all voxels of year 2001.

1311 FIG. 5. Simulated fire size cumulative distribution and observations for year 2001: Central

1312 tendency (orange line) was surrounded by the 95th confidence interval in orange and the 99.9th

1313 confidence interval (light orange), computed from 1000 simulations of fire sizes for year 2001.

1314 The dotted line corresponds to the fire size distribution from 1995-2018.

1315 FIG. 6. "Areas Under the Curve" (AUCs) corresponding to the realization of events diagnosed

1316 according to the two components of the fire activity model. The different series correspond to

1317 the "full" model (on the subset used to fit data, 1995-2014; and for the subset used to test the

1318 model, 2015-2018) and to a more basic "Linear FWI" model (before 2014), which only

1319 implemented the linear effect of the FWI as explanatory variable.

1320   FIG. 7. Evaluation of the occurrence model component (escaped fire number) for a selection of

1321   aggregation scales. In each subplot, the "Full" model is compared to an intermediate model

1322   where one critical effect required for consistent simulations at this aggregation scale was not

1323   included: a) Yearly trends for the whole studied area (the dashed vertical line shows the

1324   separation between training and validation sample); b) Seasonal trends at the weekly scale; c)

1325   Spatial patterns in number of escaped fires.

1326   FIG. 8. Same as Fig. 7 for burnt area.

1327   FIG. 9. Partial effect of the FWI for the different occurrence models

1328   FIG. 10. Same as FIG. 4, for year 2003. Comparison of simulated fire activity (in red) with

1329   observation (black dots): daily and weekly escaped fire numbers, as well as weekly number of

1330   fire larger than 10, 50, 100 ha and weekly burnt areas were summed for the whole study area.

1331   Central tendency (red line) was surrounded by the 95th and 99.9th confidence intervals in orange

1332   and light orange (computed from 1000 simulations of fire activities).

1333   FIG. 11. Same as FIG. 5 for year 2003. Comparison of simulated fire size distribution with

1334   observation: Central tendency (red line) was surrounded by the 95th confidence interval in

1335   orange and the 99.9th confidence interval (light orange), computed from 1000 simulations of

1336   fire sizes for year 2003.

1337   FIG. 12. Mean Absolute Error (top) and Model Uncertainty (bottom) of the model at various

1338   spatial (1, 4, 16, 36 pixels, NUTS 3, all region, with size in km^2 in brackets) and temporal (1

1339   day, 1 week, two weeks, …, one year, four years) aggregations for fire numbers of 1, 10, 100 ha

1340   and burnt area for the period 2015-2018.

1341

1342

1343   FIG. 1. Data site and description: a) Orography (elevation in m) of the study region; Fire activity

1344   (1995-2018): b) Monthly distribution of burnt areas; c) Spatial distribution of fires larger or

1345   equal to 1 ha; d) Yearly burnt areas. Data were extracted from the Prométhée fire database

1346   (http://www.promethee.com/).

1347
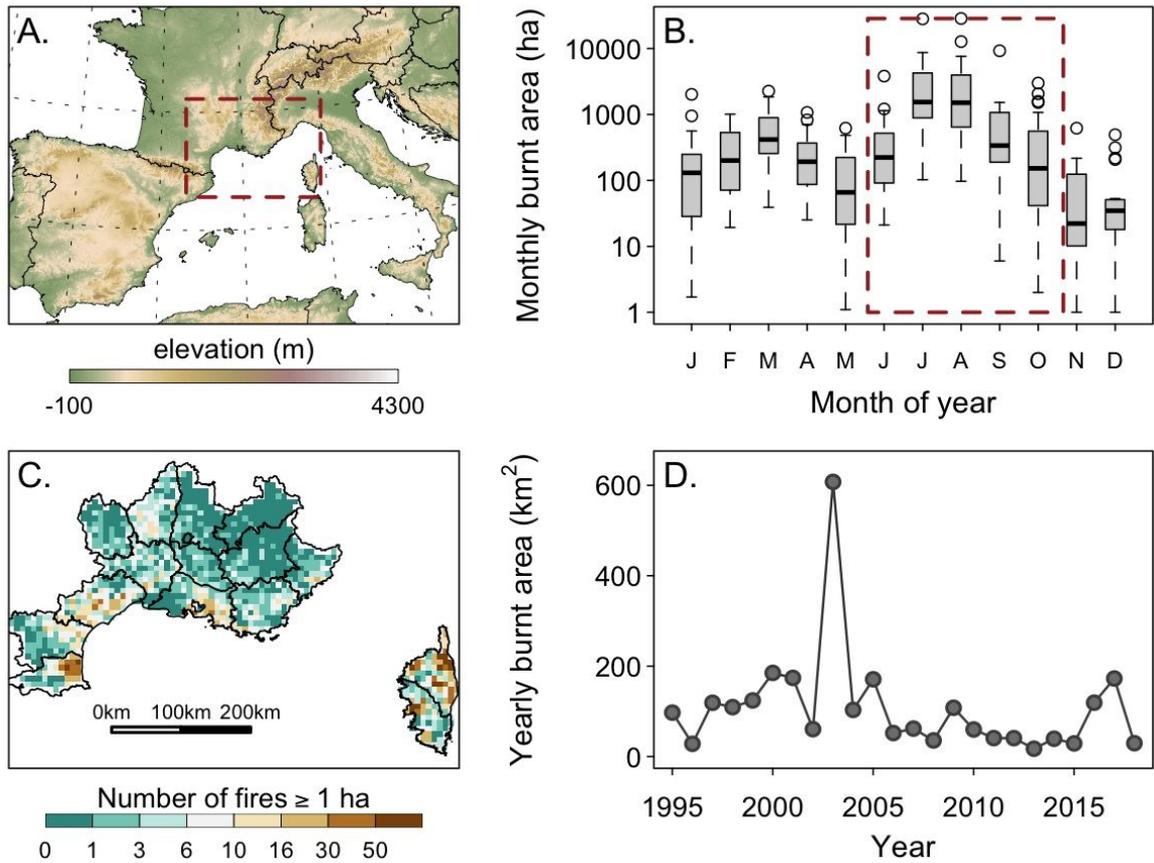
1348

1349

1350

1351   FIG. 2. Framework of the "full" probabilistic model of fire activity. The occurrence and size

1352   components of the model are presented in dashed boxes on the left in yellow and red,

1353   respectively. Random effects are listed in green, whereas statistical processes involved are in

1354   labelled in blue. The occurrence model (in yellow) used a space-time Poisson process to

1355   simulate the daily number of escaped fires (i.e. larger than 1ha) in each pixel. The size of each

1356   escaped fire is individually simulated from the 4-piece distribution estimated from the size

1357   model (in red). The distribution model includes three exceedance thresholds (Bernouilli

1358   process) and the 4 pieces corresponding to either Pareto or Generalized Pareto Distributions

1359   (because fire size cannot be infinite). The blue box on the right illustrates ensemble simulations

1360   that can be done with the model. For each pixel day, the models are used to simulate realizations

1361   (here 1000) of the number of fires (which is most often 0) and their sizes (if any). The likely

1362   fire activities can then be aggregated at larger scale to provide metrics and to compute

1363   confidence intervals.
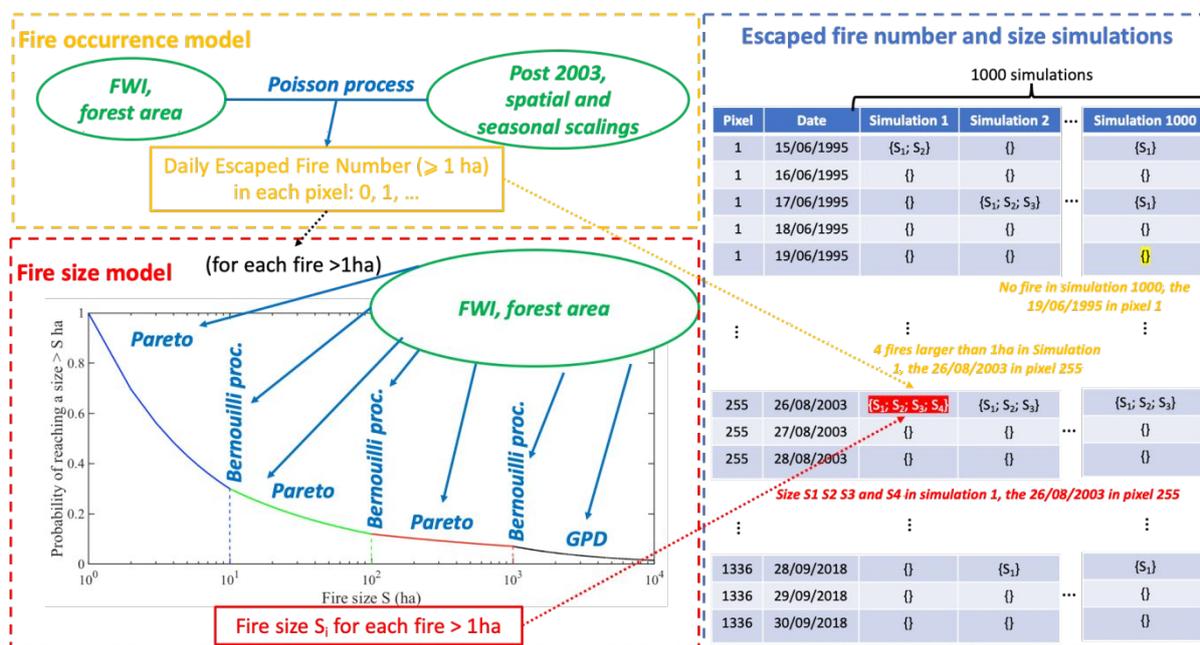
1364

1365

1366   FIG. 3. Partial effects of the "Full" fire activity model: a) Occurrence component (number of

1367   escaped fires, i.e. larger than 1 ha): effects of FWI, Forest Area (including offset), Week of

1368   Year, and location; Because the different effects are multiplicative, the partial effects represent

1369   the multiplicative effect of each factor on the escape fire number, with respect to an arbitrary

1370   reference. We respectively used FWI=5 (low danger value), week=24 (15th of June is the

1371   beginning of the summer fire season), FA=30% (mean value in Mediterranean France), and

1372   mean spatial effect for the spatial model. b) Size component (exceedance probabilities of

1373   escaped fires for a selection of thresholds ranging between 1 and 2000 ha): effects of FWI and

1374   FA, for a Forest Area of 30 % (top) and a FWI of 20 (bottom), respectively.
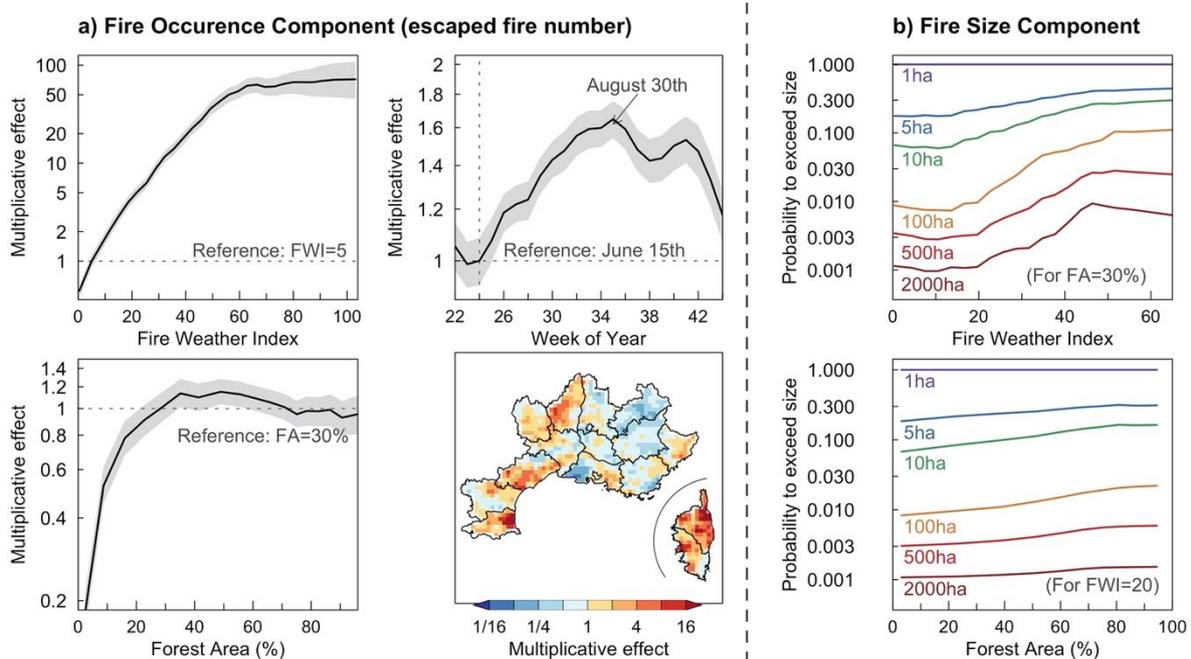
1375

1376

1377

1378

1379

1380

6

1381   FIG. 4. Simulated fire activity (in orange) and observations (black dots) for year 2001: daily

1382   and weekly escaped fire numbers, as well as weekly number of fires larger than 10, 50, 100 ha

1383   and weekly burnt areas added up over the whole study area. Central tendency (red line) was

1384   surrounded by the 95th confidence interval in orange and was based on averages computed over

1385   1000 simulations of fire activities for all voxels of year 2001.
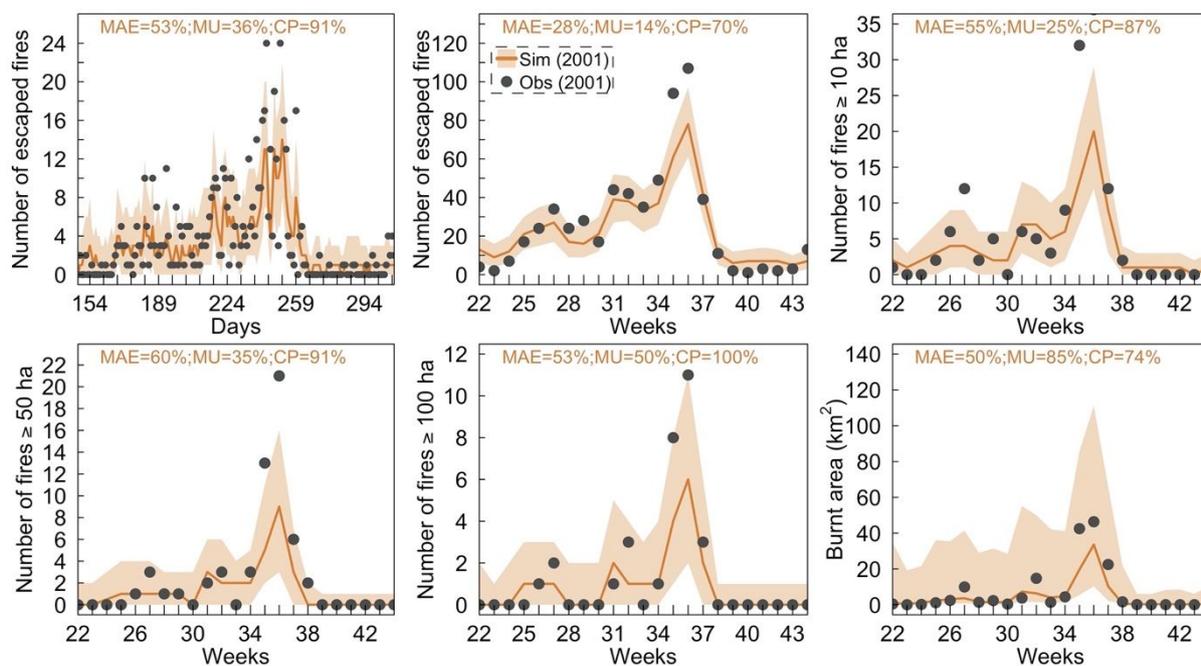
1386

1387

1388

1389

1390   FIG. 5. Simulated fire size cumulative distribution and observations for year 2001: Central

1391   tendency (orange line) was surrounded by the 95th confidence interval in orange and the 99.9[th]

1392   confidence interval (light orange), computed from 1000 simulations of fire sizes for year 2001.

1393   The dotted line corresponds to the fire size distribution from 1995-2018.
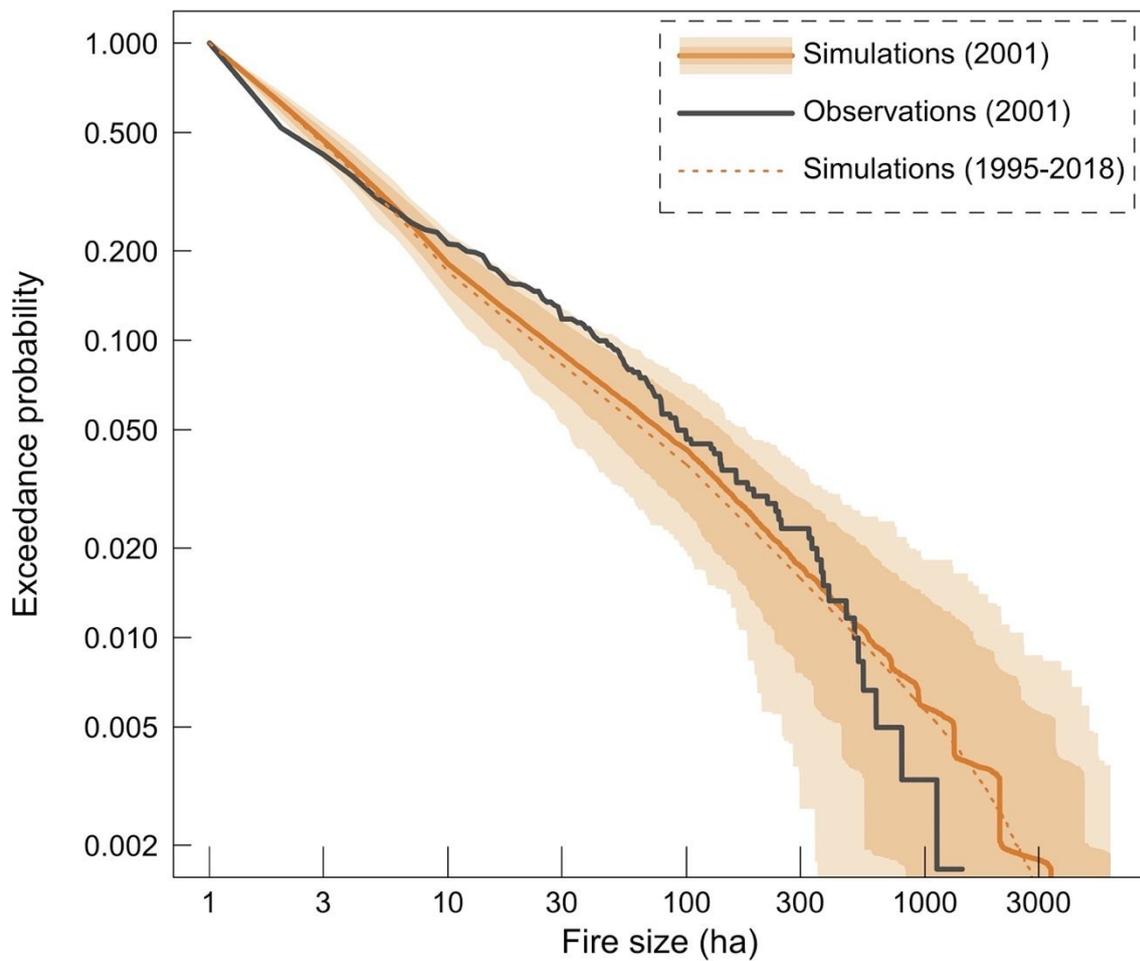
1394

1396   FIG. 6. "Areas Under the Curve" (AUCs) corresponding to the realization of events diagnosed

1397   according to the two components of the fire activity model. The different series correspond to

1398   the "full" model (on the subset used to fit data, 1995-2014; and for the subset used to test the

1399   model, 2015-2018) and to a more basic "Linear FWI" model (before 2014), which only

1400   implemented the linear effect of the FWI as explanatory variable.
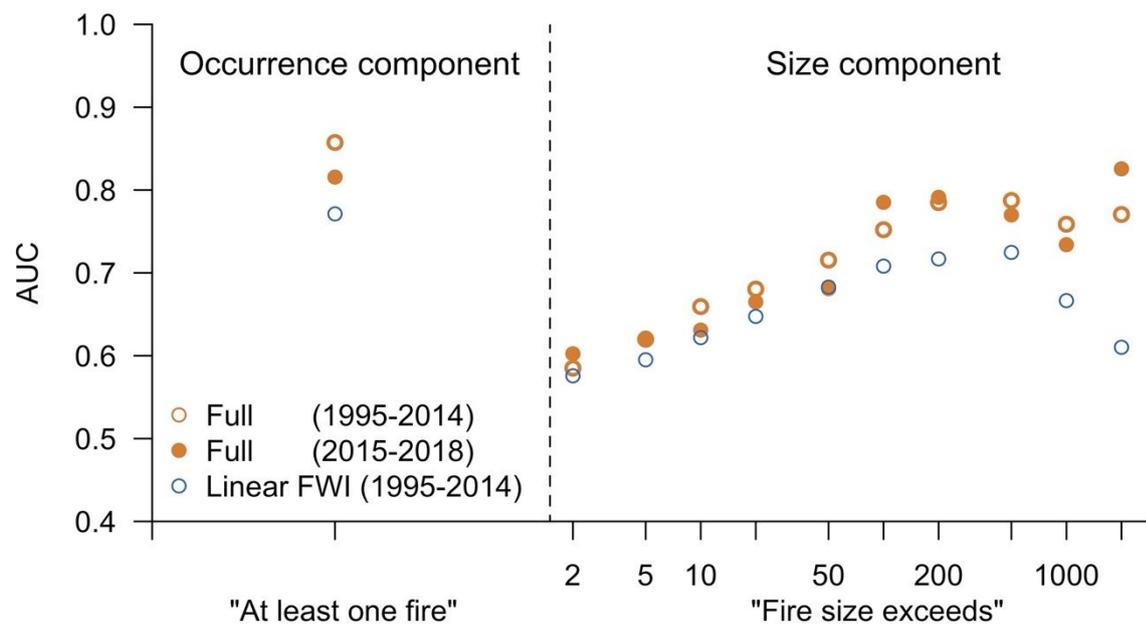
1401

1402

1403   FIG. 7. Evaluation of the occurrence model component (escaped fire number) for a selection of

1404   aggregation scales. In each subplot, the "Full" model is compared to an intermediate model

1405   where one critical effect required for consistent simulations at this aggregation scale was not

1406   included. a) Yearly trends for the whole studied area (the dashed vertical line shows the

1407   separation between training and validation sample); b) Seasonal trends at the weekly scale; c)

1408   Spatial patterns in number of escaped fires.
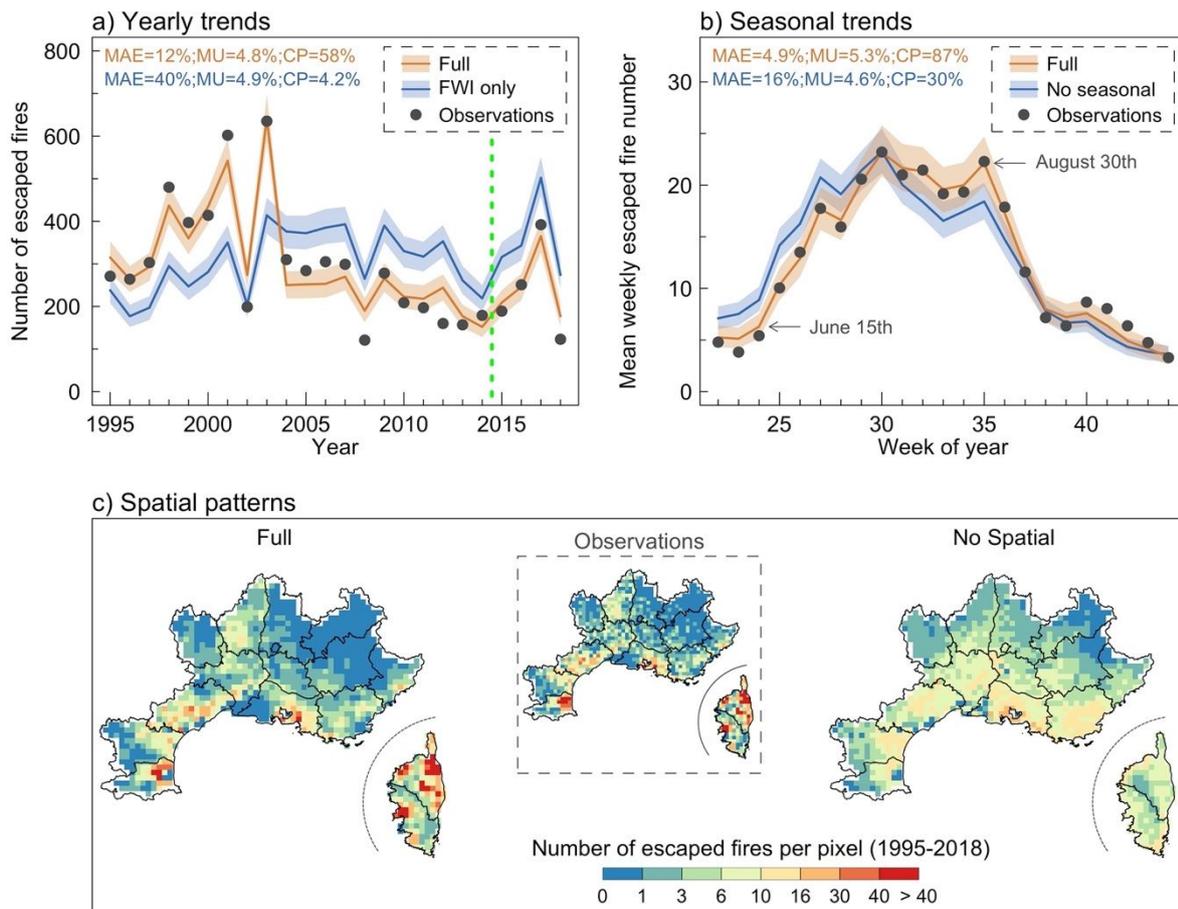
1409

1410   FIG. 8. Same as Fig. 7 for burnt area.

1411

1412

a) Yearly trends

b) Seasonal trends

c) Spatial patterns
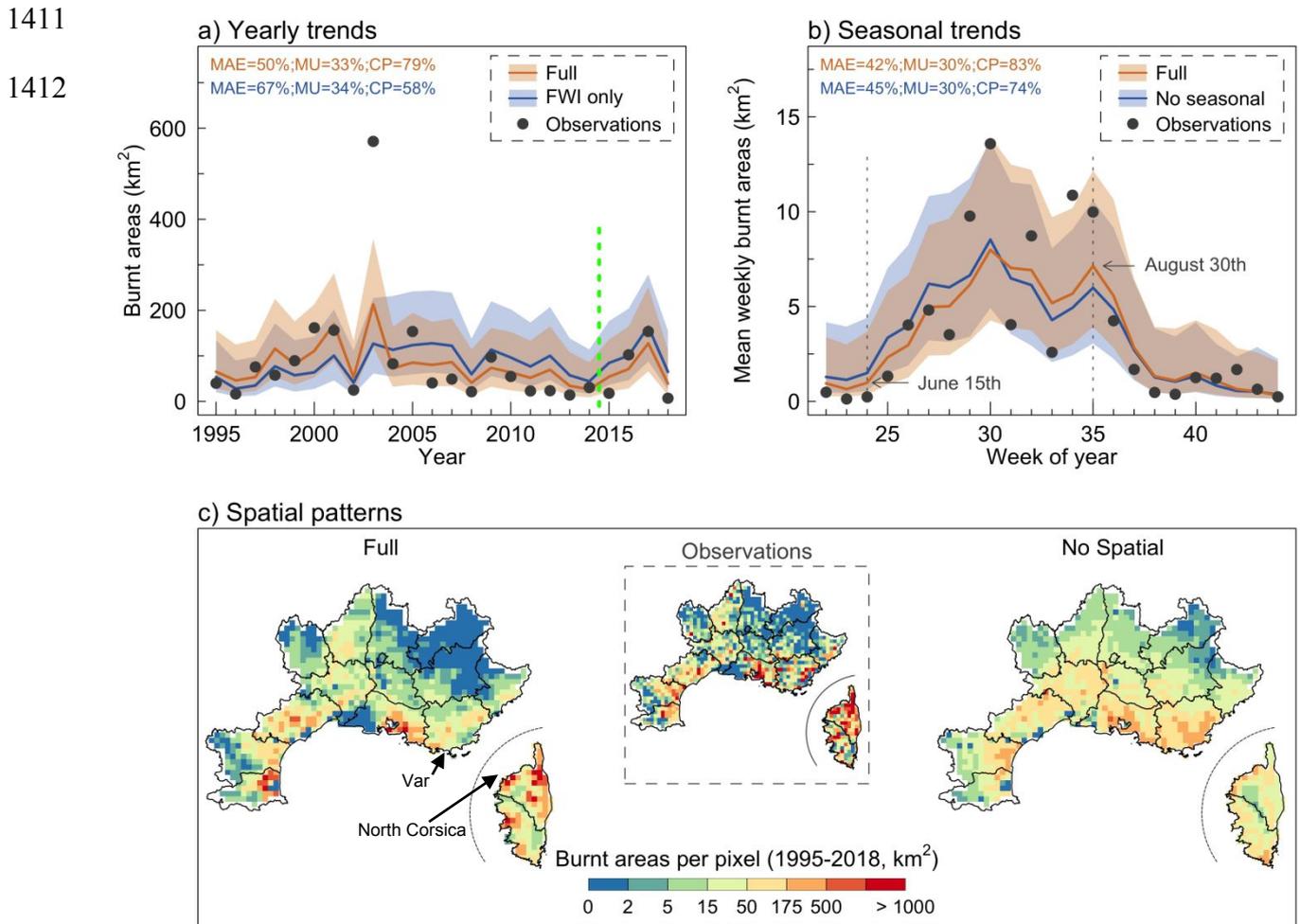
1413   FIG. 9. Partial effect of the FWI for the different occurrence models
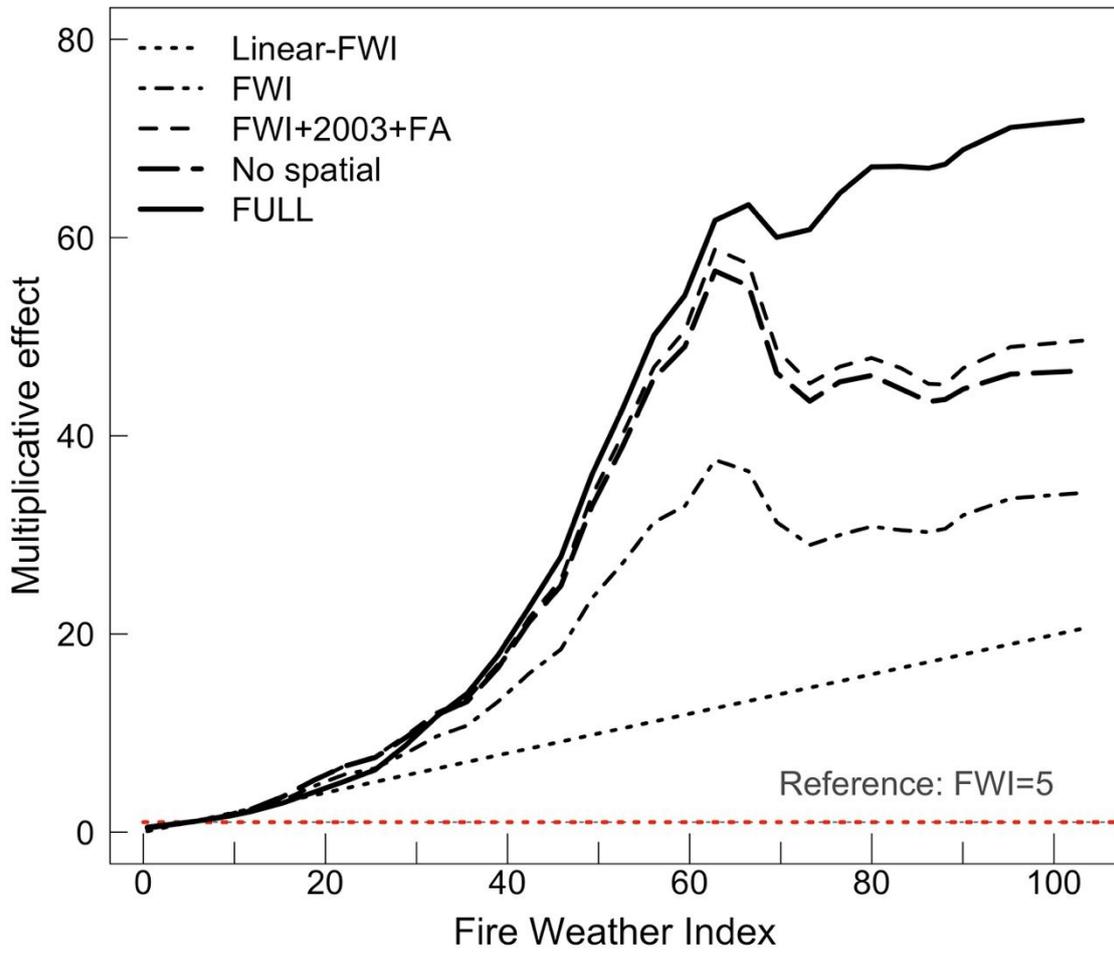
1414

1415
1416

1417   FIG. 10. Same as FIG. 4, for year 2003. Comparison of simulated fire activity (in red) with

1418   observation (black dots): daily and weekly escaped fire numbers, as well as weekly number of

1419   fire larger than 10, 50, 100 ha and weekly burnt areas were summed for the whole study area.

1420   Central tendency (red line) was surrounded by the 95th and 99.9th confidence intervals in orange

1421   and light orange (computed from 1000 simulations of fire activities).
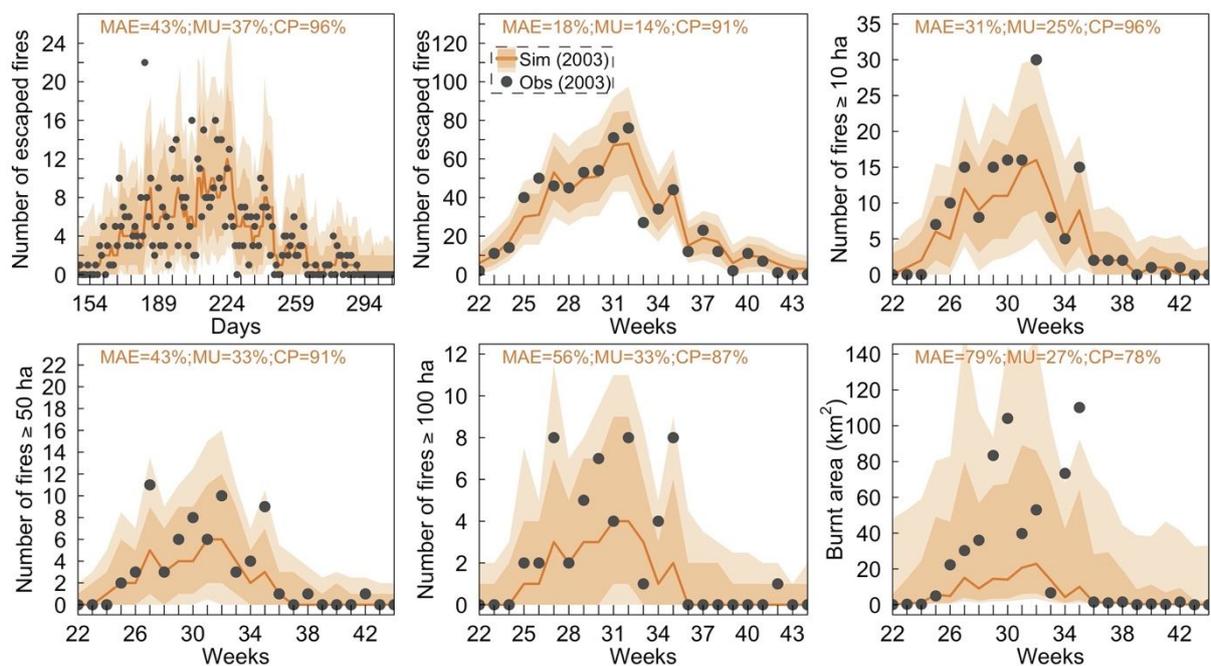
1422

1423

1424

1425   FIG. 11. Same as FIG. 5 for year 2003. Comparison of simulated fire size distribution with

1426   observation: Central tendency (red line) was surrounded by the 95th confidence interval in

1427   orange and the 99.9th confidence interval (light orange), computed from 1000 simulations of

1428   fire sizes for year 2003.
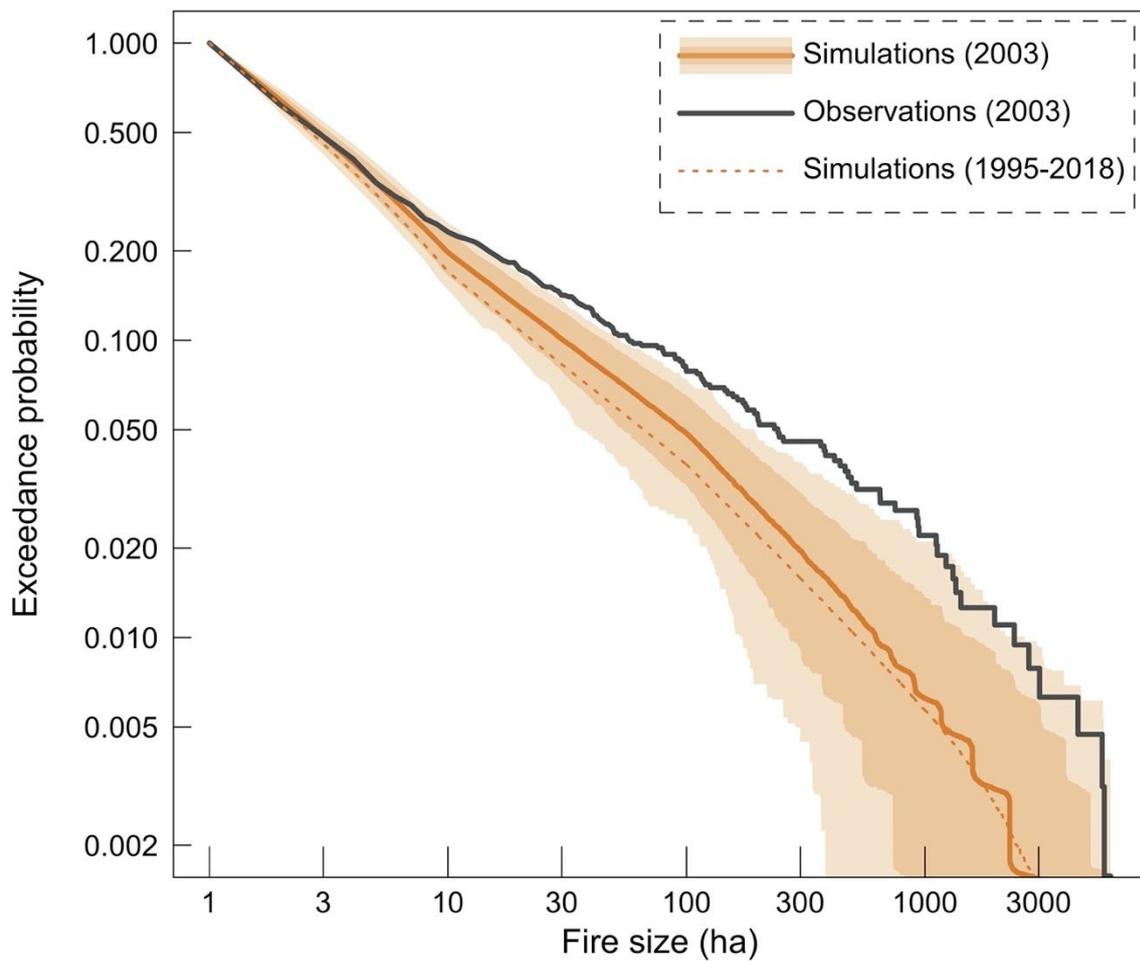
1429

1430

1431

1432   FIG. 12. Mean Absolute Error (top) and Model Uncertainty (bottom) of the model at various

1433   spatial (1, 4, 16, 36 pixels, NUTS 3, all region, with size in km^2 in brackets) and temporal (1

1434   day, 1 week, two weeks, …, one year, four years) aggregations for fire numbers of 1, 10,

1435   100 ha and burnt area for the period 2015-2018.
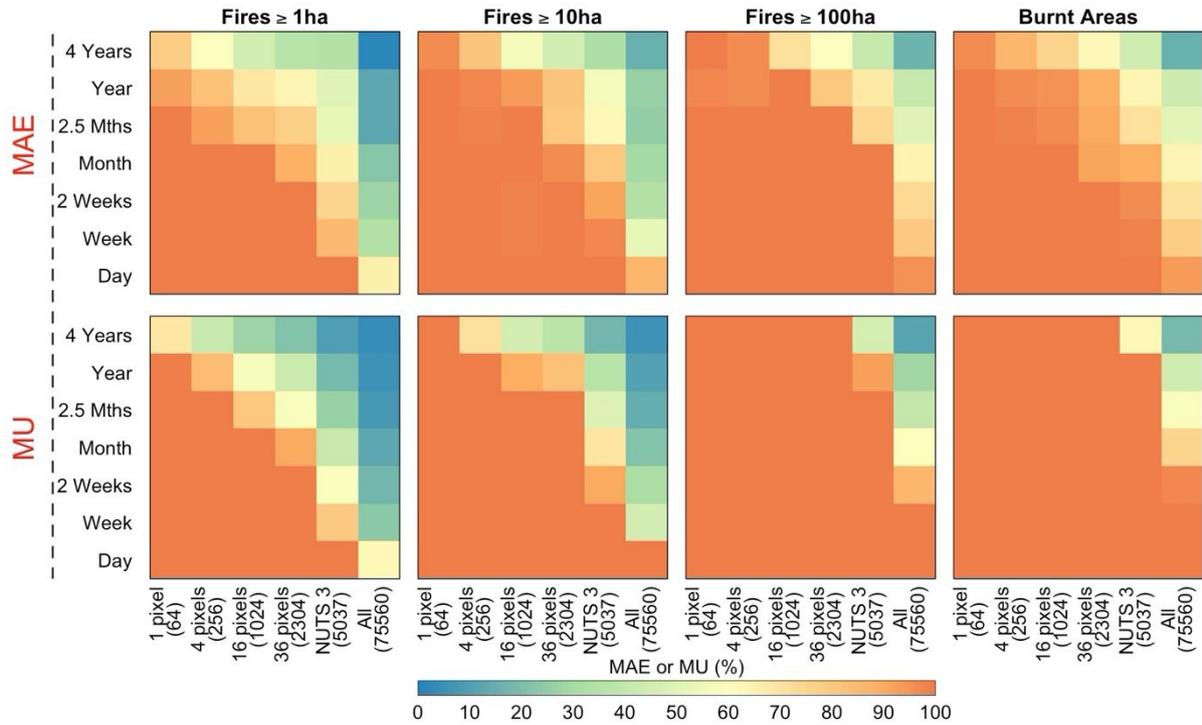
1436
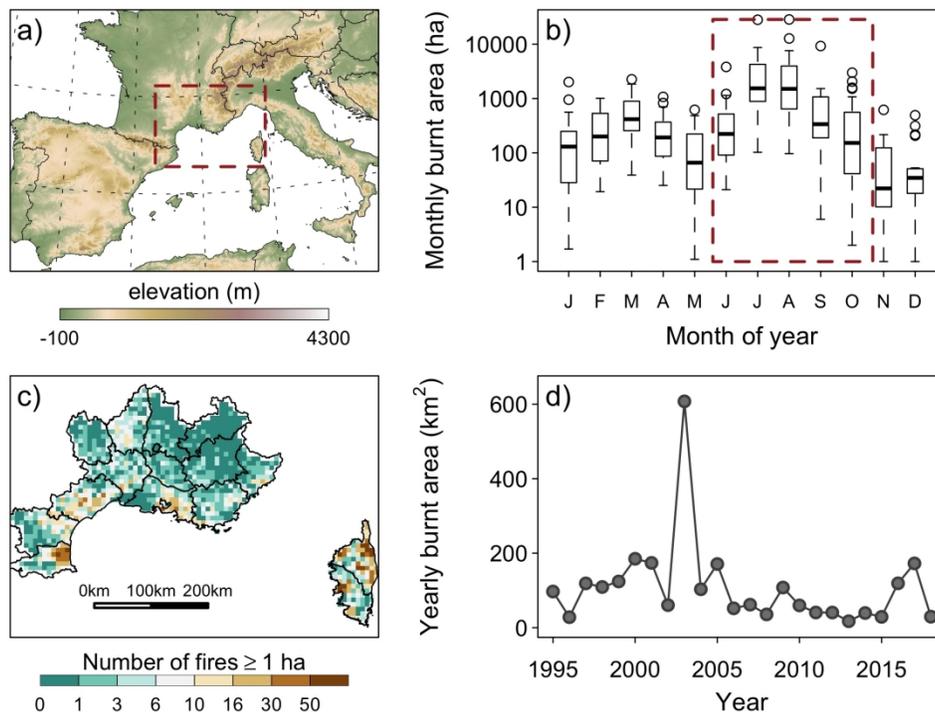
7

FIG. 1. Data site and description: a) Orography (elevation in m) of the study region; Fire activity (1995-2018): b) Monthly distribution of burnt areas; c) Spatial distribution of fires larger or equal to 1 ha; d) Yearly burnt areas. Data were extracted from the Prométhée fire database (http://www.promethee.com/).

846x634mm (72 x 72 DPI)

Fire occurrence model

FWI, forest area

Post 2003, spatial and seasonal scalings

Poisson process

Daily Escaped Fire Number (⩾ 1 ha) in each pixel 0, 1, ...
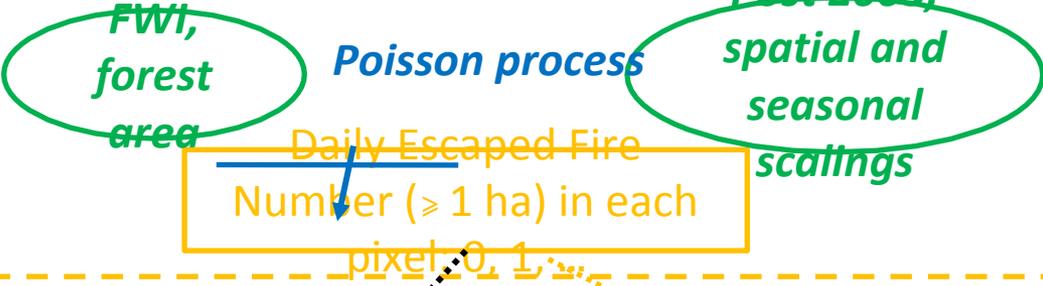
Fire size model

(for each fire >1ha)

FWI, forest area

Pareto

Bernouilli proc.

Bernouilli proc.

Bernouilli proc.

Pareto

Pareto

GPD

Probability of reaching a size > S ha

Fire size S (ha)

Fire size S_i for each fire >1ha
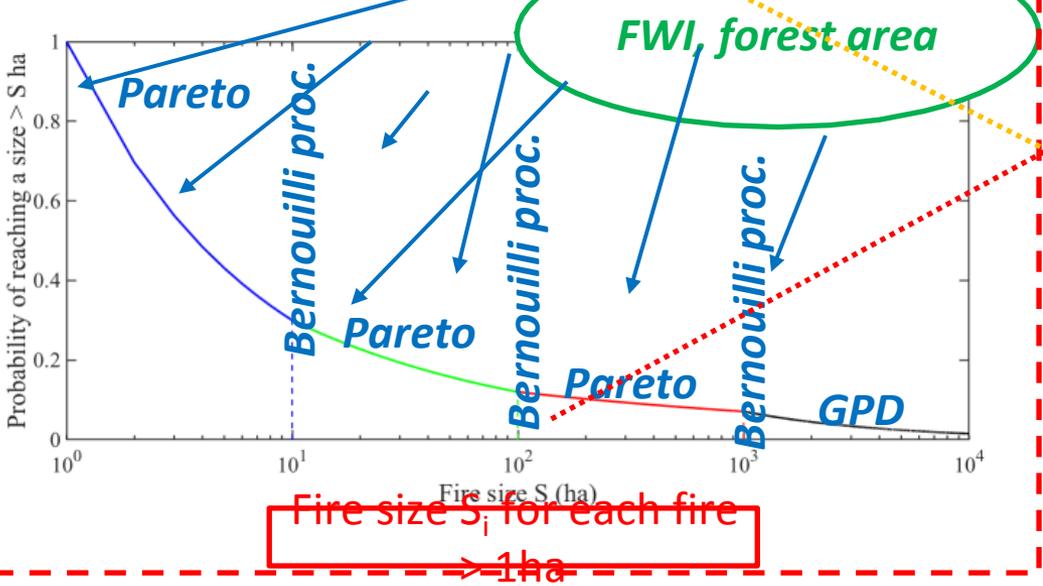
Escaped fire number and size simulations

1000 simulations

Pixel	Date	Simulation 1	Simulation 2	...	Simulation 1000
1	15/06/1995	{S_1; S_2}	{}		{S_1}
1	16/06/1995	{}	{}		{}
1	17/06/1995	{}	{S_1; S_2; S_3}	...	{S_1}
1	18/06/1995	{}	{}		{}
1	19/06/1995	{}	{}		{}

No fire in simulation 1000, the 19/06/1995 in pixel 1

4 fires larger than 1ha in Simulation 1, the 26/08/2003 in pixel 255

Pixel	Date	Simulation 1	Simulation 2	...	Simulation 1000
255	26/08/2003	{S_1; S_2; S_3; S_4}	{S_1; S_2; S_3}		{S_1; S_2; S_3}
255	27/08/2003	{}	{}	...	{}
255	28/08/2003	{}	{}		{}

Size S1 S2 S3 and S4 in simulation 1, the 26/08/2003 in pixel 255

| Pixel	Date	Simulation 1	Simulation 2	...	Simulation 1000
1336	28/09/2018	{}	{S_1}		{S_1}
1336	29/09/2018	{}	{}	...	{}
1336	30/09/2018	{}	{}		{}
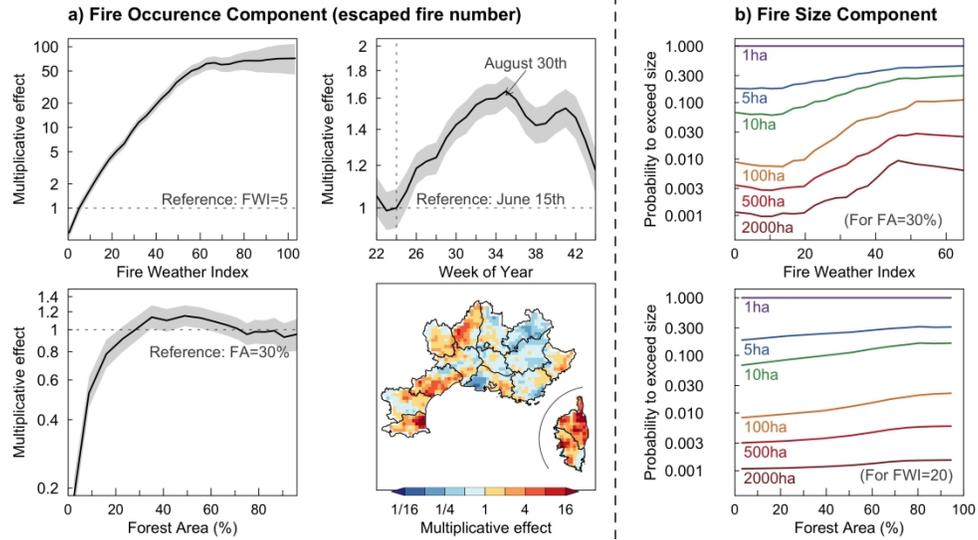
FIG. 3. Partial effects of the "Full" fire activity model: a) Occurrence component (number of escaped fires, i.e. larger than 1 ha): effects of FWI, Forest Area (including offset), Week of Year, and location; Because the different effects are multiplicative, the partial effects represent the multiplicative effect of each factor on the escape fire number, with respect to an arbitrary reference. We respectively used FWI=5 (low danger value), week=24 (15th of June is the beginning of the summer fire season), FA=30% (mean value in Mediterranean France), and mean spatial effect for the spatial model. b) Size component (exceedance probabilities of escaped fires for a selection of thresholds ranging between 1 and 2000 ha): effects of FWI and FA, for a Forest Area of 30 % (top) and a FWI of 20 (bottom), respectively.

1354x740mm (72 x 72 DPI)
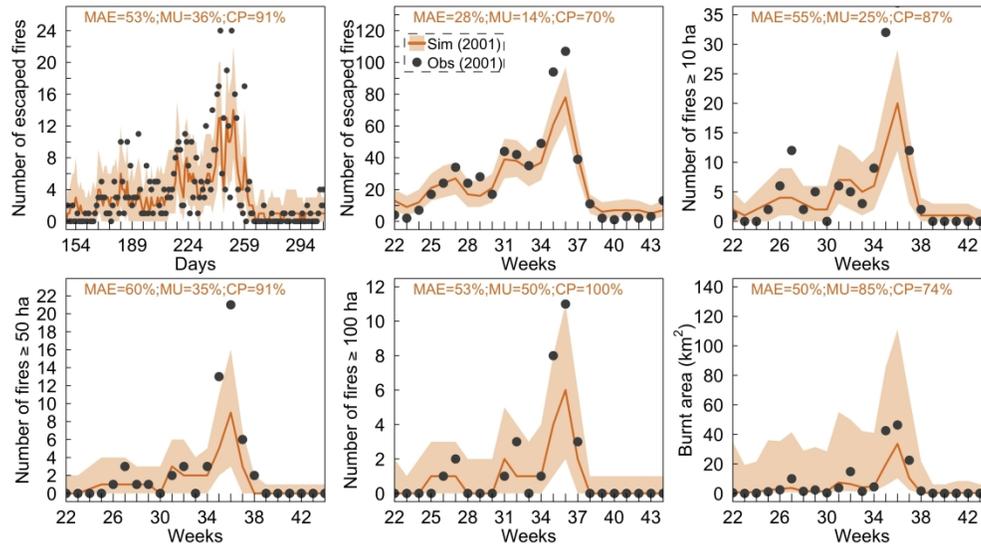
FIG. 4. Simulated fire activity (in orange) and observations (black dots) for year 2001: daily and weekly escaped fire numbers, as well as weekly number of fires larger than 10, 50, 100 ha and weekly burnt areas added up over the whole study area. Central tendency (red line) was surrounded by the 95th confidence interval in orange and was based on averages computed over 1000 simulations of fire activities for all voxels of year 2001.

1354x740mm (72 x 72 DPI)
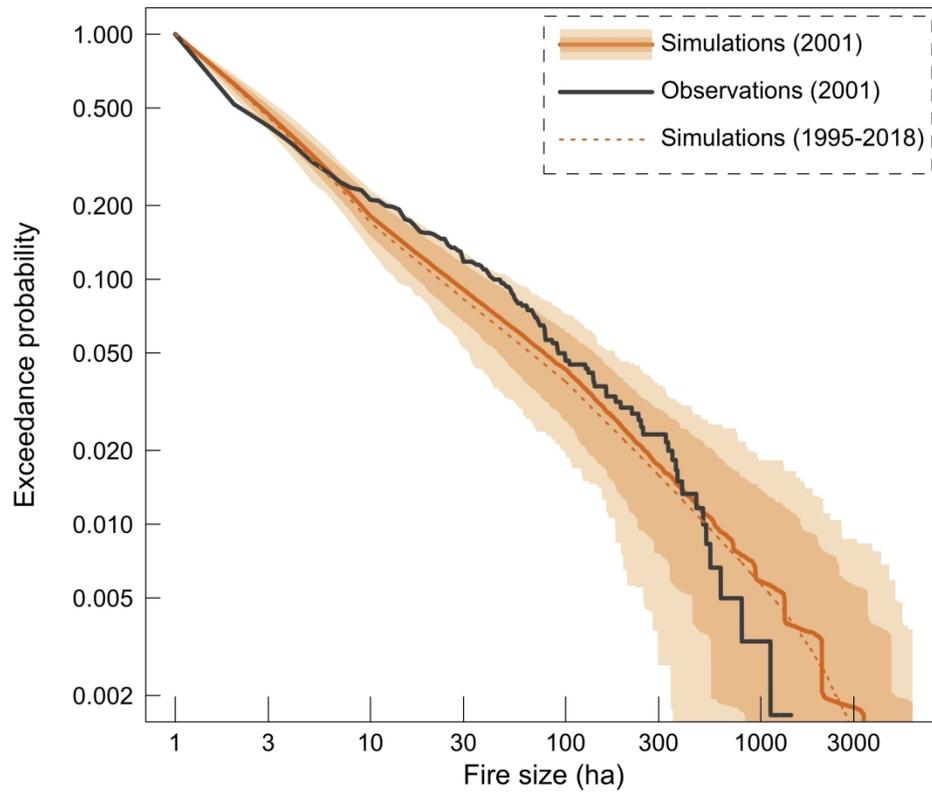
FIG. 5. Simulated fire size cumulative distribution and observations for year 2001: Central tendency (orange line) was surrounded by the 95th confidence interval in orange and the 99.9th confidence interval (light orange), computed from 1000 simulations of fire sizes for year 2001. The dotted line corresponds to the fire size distribution from 1995-2018.

1015x846mm (72 x 72 DPI)
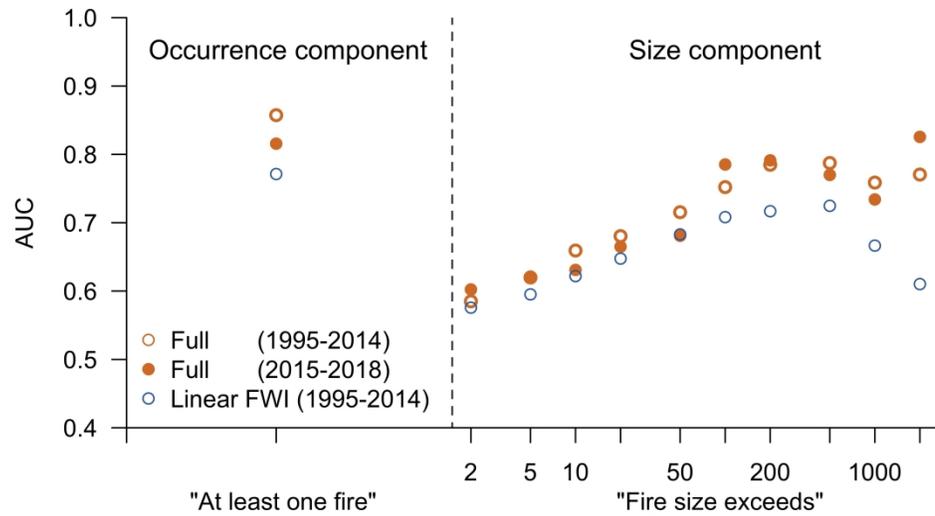
FIG. 6. "Areas Under the Curve" (AUCs) corresponding to the realization of events diagnosed according to the two components of the fire activity model. The different series correspond to the "full" model (on the subset used to fit data, 1995-2014; and for the subset used to test the model, 2015-2018) and to a more basic "Linear FWI" model (before 2014), which only implemented the linear effect of the FWI as explanatory variable.

1438x846mm (72 x 72 DPI)
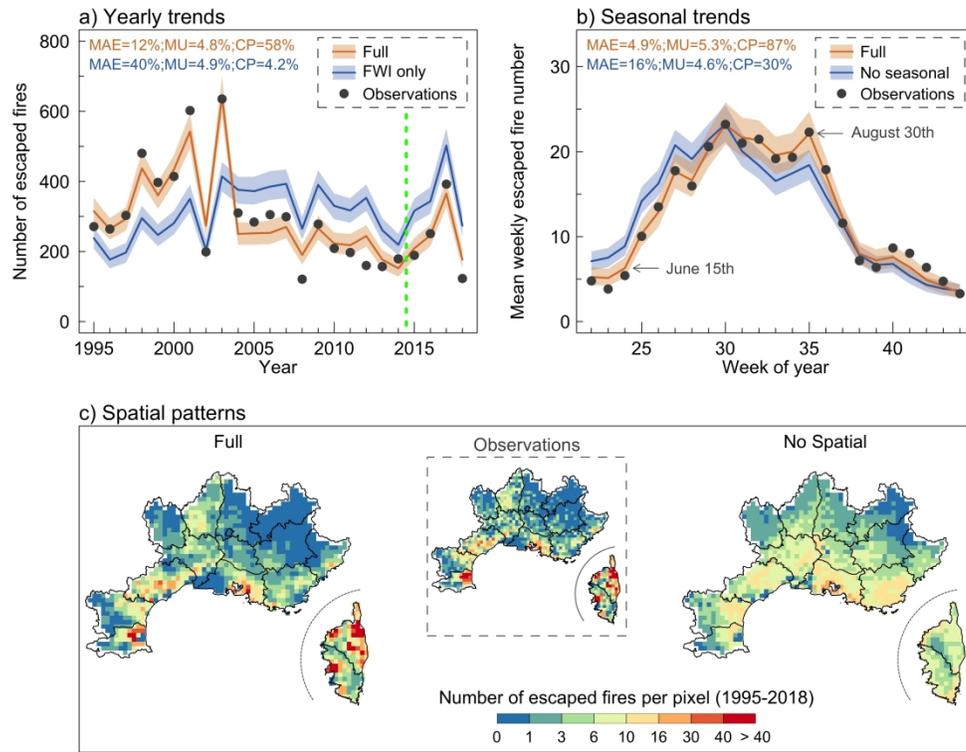
FIG. 7. Evaluation of the occurrence model component (escaped fire number) for a selection of aggregation scales. In each subplot, the "Full" model is compared to an intermediate model where one critical effect required for consistent simulations at this aggregation scale was not included. a) Yearly trends for the whole studied area (the dashed vertical line shows the separation between training and validation sample); b) Seasonal trends at the weekly scale; c) Spatial patterns in number of escaped fires.

1354x1057mm (72 x 72 DPI)

a) Yearly trends

MAE=50%;MU=33%;CP=79%
MAE=67%;MU=34%;CP=58%

— Full
— FWI only
● Observations

b) Seasonal trends

MAE=42%;MU=30%;CP=83%
MAE=45%;MU=30%;CP=74%

— Full
— No seasonal
● Observations

August 30th

June 15th

c) Spatial patterns

Full   Observations   No Spatial

Var
North Corsica

Burnt areas per pixel (1995-2018, km^2)

0   2   5   15   50   175   500   > 1000
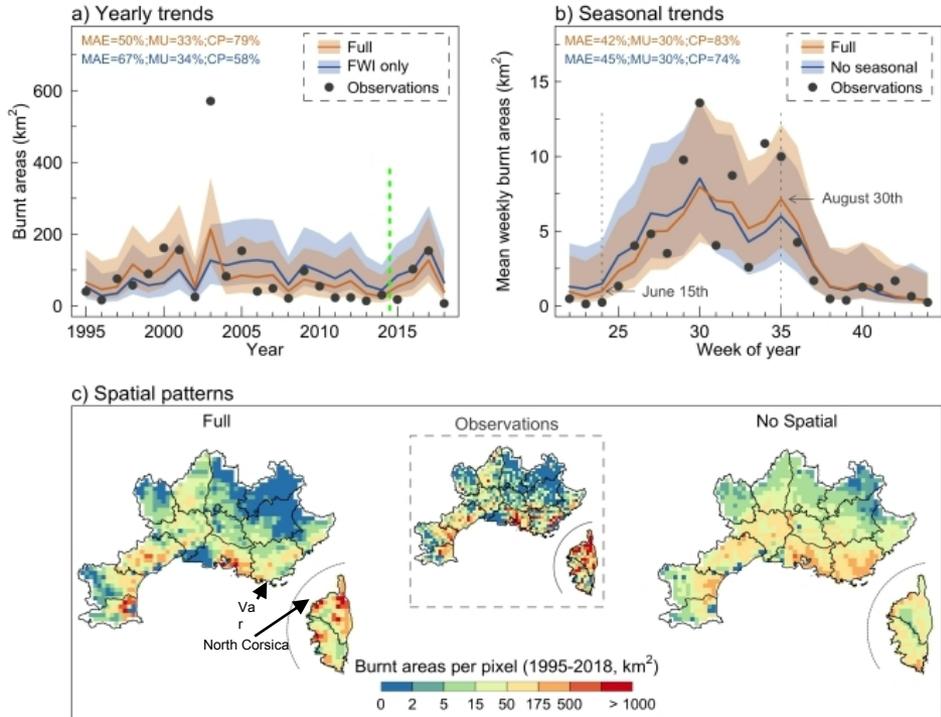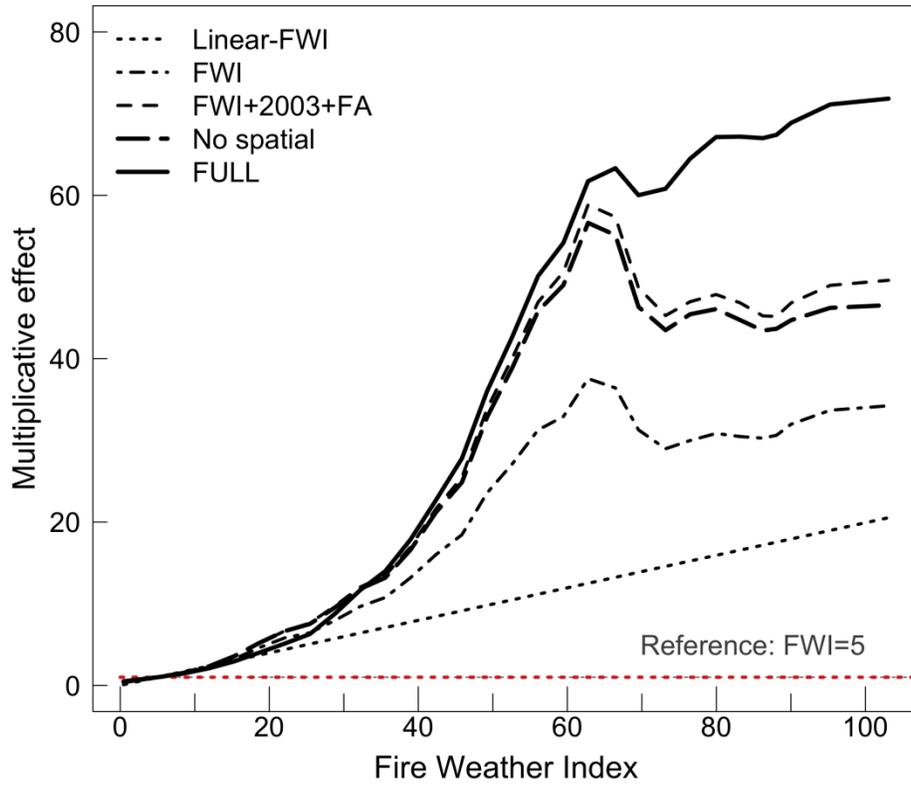
FIG. 9. Partial effect of the FWI for the different occurrence models

1015x846mm (72 x 72 DPI)
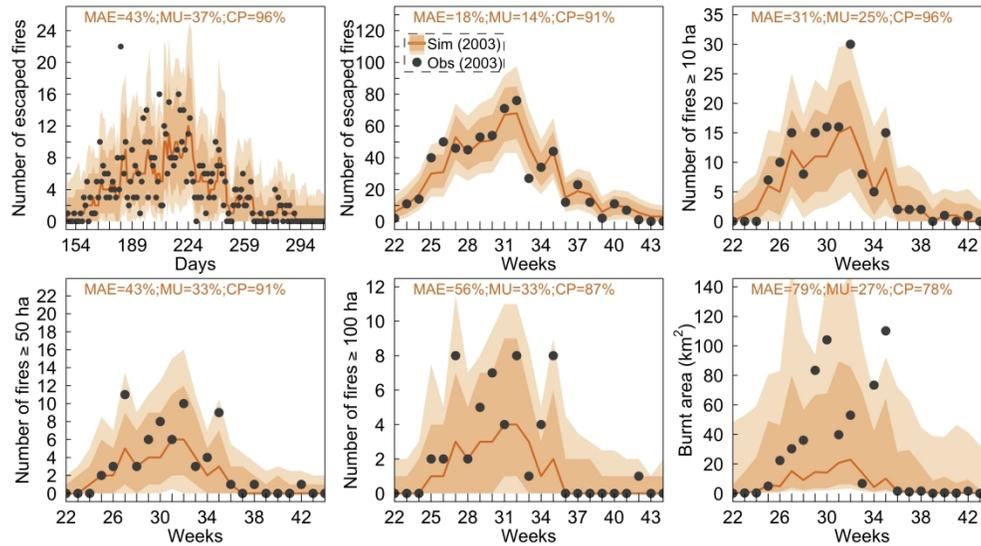
FIG. 10. Same as FIG. 4, for year 2003. Comparison of simulated fire activity (in red) with observation (black dots): daily and weekly escaped fire numbers, as well as weekly number of fire larger than 10, 50, 100 ha and weekly burnt areas were summed for the whole study area. Central tendency (red line) was surrounded by the 95th and 99.9th confidence intervals in orange and light orange (computed from 1000 simulations of fire activities).

1354x740mm (72 x 72 DPI)
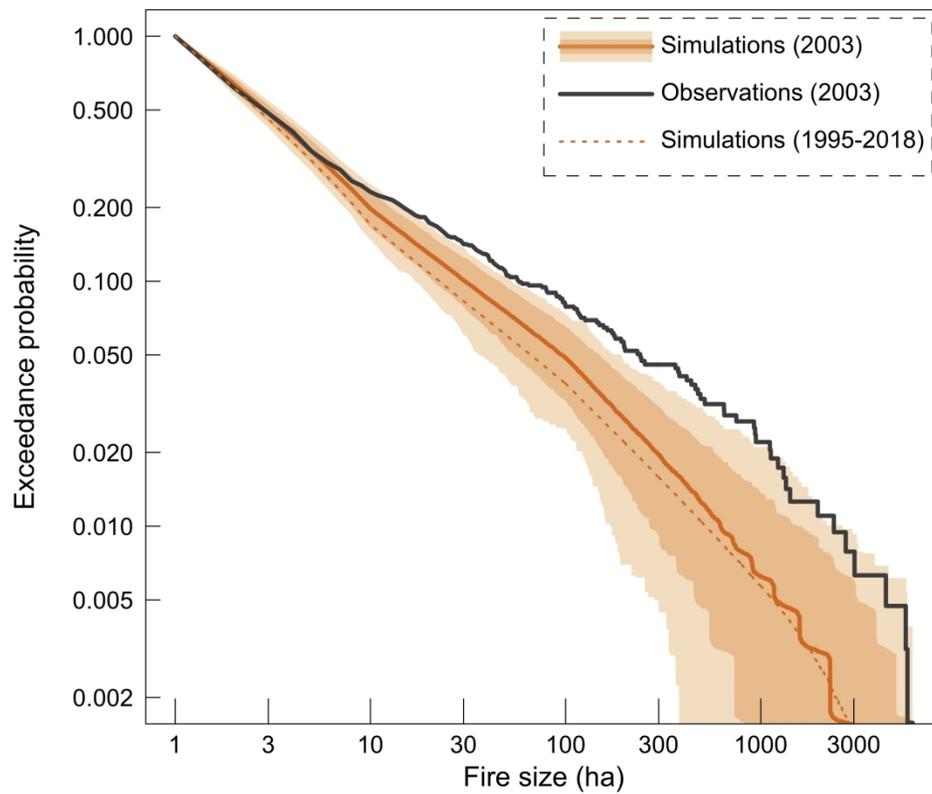
FIG. 11. Same as FIG. 5 for year 2003. Comparison of simulated fire size distribution with observation: Central tendency (red line) was surrounded by the 95th confidence interval in orange and the 99.9th confidence interval (light orange), computed from 1000 simulations of fire sizes for year 2003.

1015x846mm (72 x 72 DPI)
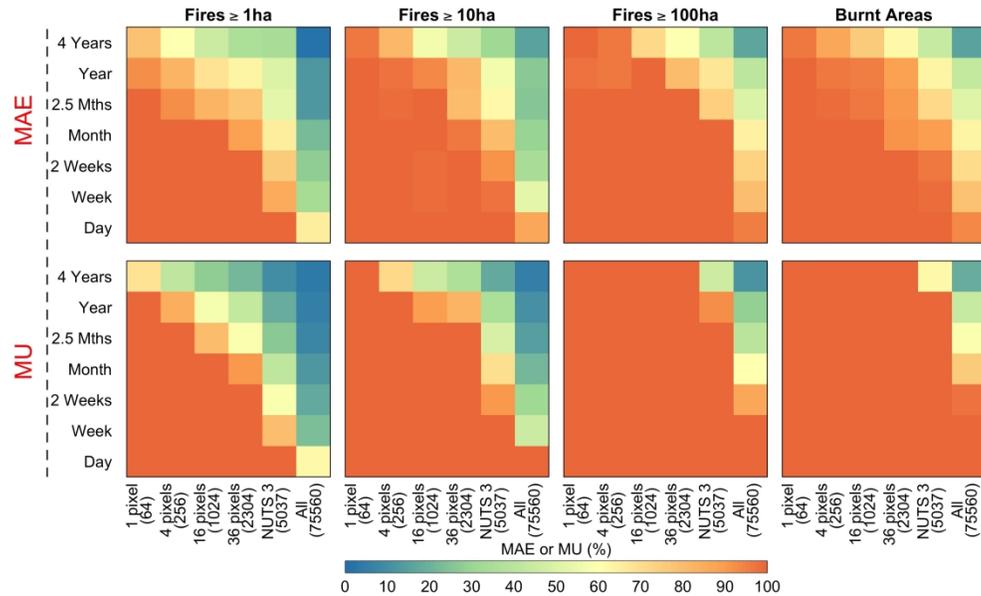
FIG. 12. Mean Absolute Error (top) and Model Uncertainty (bottom) of the model at various spatial (1, 4, 16, 36 pixels, NUTS 3, all region, with size in km2 in brackets) and temporal (1 day, 1 week, two weeks, …, one year, four years) aggregations for fire numbers of 1, 10, 100 ha and burnt area for the period 2015-2018.

1333x803mm (72 x 72 DPI)