ECCOMAS

Proceedia

A BAYESIAN HIERARCHICAL MODEL FOR CLIMATIC LOADS UNDER CLIMATE CHANGE

P. Croce[1], P.Formichi[1] and F. Landi[1]

[1] Department of Civil and Industrial Engineering – Structural Division, University of Pisa
Largo Lucio Lazzarino 1, Pisa, Italy
e-mail: name@e-mail.address

Abstract

In the mid-term future, climate change could determine significant alterations of the frequency and magnitude of climate extremes, so affecting the design of new structures and infrastructures, and the reliability of existing ones designed according to the provisions of present or past Codes.

In this work, a Bayesian hierarchical model for the characterization of climate extremes under non stationary climate conditions is presented starting from the analysis of an ensemble of future climate projections. The Bayesian Hierarchical Model is formulated through the classical three-level formulation, in which the standard extreme value representation at each site is combined with a spatial latent process, and collects the main sources of uncertainties regarding climate projections.

A Metropolis Hastings algorithm within a Gibbs sampler is implemented to update model parameters, and from the posterior probability density functions of the extreme value distribution parameters, return levels that serve as basis for structural design are estimated. The implementation of the model in different time windows combined with the Bayesian framework allows the probabilistic assessment of time evolution of extreme value parameters and return levels.

The results obtained for a relevant case study demonstrate the possibilities of the proposed methodology to describe climate extremes under climate change and to provide guidance for potential amendments in the current definition of climatic actions on structures.

Keywords: Climate Change, Climatic Actions, Structural Design, Bayesian Hierarchical Model, MCMC algorithm.

1   INTRODUCTION

In the mid-term future, climate change could determine significant alterations of the frequency and the magnitude of climate extremes. Since structural design is often governed by climatic actions such as thermal, wind, snow and ice loads, alteration of them caused by climate change could significantly affect the design of new structures and infrastructures as well as the reliability of the existing ones designed in accordance to the provisions of current or past Codes [1]. Indeed, the current definition of climatic actions on structures is based on the extreme value analysis of the underlying natural phenomena (daily temperatures, ground snow load, wind velocities) under the assumption of stationary climate conditions [2].

As consequence of global warming this assumption is becoming more and more arguable and a better evaluation of climate extremes and their evolution over time is needed to evaluate the potential consequences for infrastructures and buildings.

Dealing with climate extremes, generally recorded at a spatial scale, a key strategy in extreme value analysis to overcome difficulties caused by the scatter of data is the spatial modelling [3]. The main advantage in spatial modelling is the pooling of information but it can be also useful for interpolation to sites where little or no data may have been collected. Then, the implementation in a Bayesian framework, enables inferences and predictions to incorporate uncertainties in process variation and parameter estimates.

In order to characterize the spatial behavior of the extreme value process, a Bayesian hierarchical model for climate extremes derived from the analysis of Regional Climate Model (RCM) output is proposed. The model is able to incorporate physical and spatial information through covariates and random effects and is implemented on different time windows of forty years long to assess the time evolution of extreme value parameters. From the posterior PDFs of extreme value parameters, the characteristic values of climatic loads, used for structural design, are evaluated assessing their changes with time and considering the uncertainty in the predictions.

The proposed methodology will be presented showing the results obtained for extreme ground snow loads in the Italian Mediterranean region [4], considering an ensemble of six different RCMs for the period 1951-2100 and two different emission scenarios.

2   METHODOLOGY

There has been considerable recent interest in spatial hierarchical models to characterize the spatial behavior of climate data. Aim of these models is to describe how the marginal distribution of a quantity of interest varies with its location. The key idea is that rather than applying a spatial model directly to the data, it is assumed that there is a latent spatial process characterized by a spatial model for the parameters of the marginal distributions at each location. An extensive review of such models for spatial data can be found in [5].

Hierarchical spatial modelling for extremes has begun to be studied recently, one of the first work in this field is found in [3], while successive developments and applications are available in [6] for extreme precipitations and in [7] for extreme precipitations obtained by regional climate models. They are increasingly used for the capability to borrow strength from neighboring locations when estimating parameters in extreme value analysis, usually characterized by small amount of data. The Bayesian Hierarchical Model is formulated through what has now become the standard three-level hierarchical formulation [8]:

- Data Layer, which is the base layer where data, e.g. the yearly maxima of the investigated climate variables, are modelled at each location according to the Extreme Value theory;

- Process Layer, where the latent process that drives the extremes for the study region is formulated;
- Prior Layer, where information about the parameters controlling the latent process are given in terms of prior distributions.

The model is flexible and able to incorporate covariate information, variability due to spatial effects and micro-scale variability due to climate model uncertainty. Each layer of the model will be fully described in the next paragraphs.

2.1 Data Level

At data level, series of yearly maxima derived from the analysis of climate projections provided by each RCM r, are available for each cell i in the study region. In order to evaluate the evolution in time of the extreme value process, data are divided in subsequent time windows of 40 years shifted by ten years, thus obtaining eleven time window t (1951-1990, 1961-2000,..., 2041-2080 and 2051-2090). The time window length is set to 40 years to be consistent with the actual definition of climatic loads on structural codes, which is based on the analysis of observed data series of climate extremes of about forty years [9], while the shift of ten years is defined to properly evaluate the evolution in time of climatic loads.

For each time window t, $N=40$ yearly maxima are thus given at each cell i in the study region and assuming an Extreme Value Distribution Type I as marginal distribution, the random variable Y_{itr} is described by the cumulative distribution function $F(y)$

$$F(Y_{i,t,r} < y) = \exp\left\{-\exp\left[-\frac{y - \mu_{i,t,r}}{\sigma_{i,t,r}}\right]\right\} \tag{1}$$

and the probability density function $f(y)$ is

$$f(Y_{i,t,r} < y) = \frac{1}{\sigma_{i,t,r}}\exp\left\{-\left[\frac{y - \mu_{i,t,r}}{\sigma_{i,t,r}} + \exp\left(-\frac{y - \mu_{i,t,r}}{\sigma_{i,t,r}}\right)\right]\right\} \tag{2}$$

where $\mu_{i,t,r}$ and $\sigma_{i,t,r}$ are the location and scale parameter for cell i, time window t and RCM r. The first level of the hierarchical model structure, for each climate model r, will be described by

$$Y_t(s) \mid \theta_t \sim \text{EVI}(\mu_t(s,\omega), \exp(\log(\sigma_t)(s,\omega))) \tag{3}$$

with

$Y_t(s)$ are the yearly maxima of climate data at the location s in the study region for the time window t;

θ_t are the random parameter of the model in the time window t;

$\mu_t(s,\omega)$ is a random field describing the spatial variation of location parameter of EV Type I distribution in the time window t, where $\omega \in \Omega$ express the random event;

$\log(\sigma_t)(s,\omega)$ is a random field describing the spatial variation of the log-scale parameter of EV Type I distribution in the time window t, where $\omega \in \Omega$ express the random event.

If $Y_{i,t}$ is a vector of the yearly maxima in the investigated time window t for the cell i in the study region and $Y_t = (Y^T_{1,t}, ..., Y^T_{D,t})$ contains all the maxima for the D cells in the region, then assuming the conditional independence of Y_i for all location, common assumption in hierarchical modelling [5], starting from eq. 2 the likelihood function becomes

$$p(Y_t \mid \theta_t) = \prod_{i=1}^{D} \prod_{k=1}^{40} \frac{1}{\sigma_i} \exp\left\{-\left[\frac{y_{ik} - \mu_i}{\sigma_i} + \exp\left(-\frac{y_{ik} - \mu_i}{\sigma_i}\right)\right]\right\} \tag{4}$$

2.2 Process Level

In the hierarchical model, at the process level, the latent spatial process is formulated by constructing a structure that relates the parameter of the data level to the characteristics of the region. In particular, a Gaussian random field is proposed to model spatial variation of location and log-scale parameters according the following formulas

$$\mu_t(s,\omega) \sim N(X\beta_{\mu,t} + W_{\mu,t}(s,\omega), \tau_{\mu,t}^2) \tag{5}$$

$$\log(\sigma_t)(s,\omega) \sim N(X\beta_{\sigma,t} + W_{\sigma,t}(s,\omega), \tau_{\sigma,t}^2) \tag{6}$$

with

$W_{\mu,t}(s,\omega)$ is a spatial random effect described by a zero mean Gaussian random field $N(0, \sum_\mu(l_{\mu,t}, s_{\mu,t}))$ with covariance matrix $\sum_\mu$.

$W_{\sigma,t}(s,\omega)$ is a spatial random effect described by a zero mean Gaussian random field $N(0, \sum_\sigma(l_{\sigma,t}, s_{\sigma,t}))$ with covariance matrix $\sum_\sigma$.

X is a matrix of covariate information;

$\beta_{\mu,t}$ and $\beta_{\sigma,t}$ are vectors of regression coefficients for μ_t and σ_t given X;

τ_μ^2 and τ_σ^2 are precision terms for the location and the log-scale fields

Different models may be set for the covariance structure, considering stationarity or non stationarity in the covariance function as described in [10]. In this work an exponential model with parameter correlation length $l_{\mu,t}$ and sill $s_{\mu,t}$ has been considered. ;

Covariate information are spatially-varying, physical features or observable quantities that can either be collected at all prediction locations of interest or in some way interpolated from nearby observations [11] (for example, elevation, or geographical feature such latitude or longitude but also wind speed or direction).

The precision terms, $\tau_{\mu,t}^2$ and $\tau_{\sigma,t}^2$ in eq. 5 and 6, can be viewed as a noise associated with replication of measurements at location s, and in this case represents the variability of the data related to internal climate model uncertainty. However, the availability of few realizations of climate model run, often only one, due to the enormous computational demand doesn't allow a direct assessment of this source of uncertainty.

A possibility to assess the uncertainty related to the RCM internal variability is the methodology described by the authors in [12], where an ad hoc weather generator is proposed able to generate new consistent climate projections directly from RCM output. Analyzing the generated series, an evaluation of the noise associated to the EV parameters becomes possible and the constant precision terms $\tau_{\mu,t}^2$ and $\tau_{\sigma,t}^2$, associated at each investigated climate model r, depending on the cell i and the time window t, are defined.

2.3 Prior Level

Prior distribution are finally assigned to the hyperparameters of the model at each time window t, $\theta_t(\beta_{\mu,t}, \beta_{\sigma,t}, l_{\mu,t}, s_{\mu,t}, l_{\sigma,t}, s_{\sigma,t})$. Where possible, uninformative priors are assigned to these parameters and conjugate priors are used to facilitate the use of Gibbs sampling in the model implementation.

Normal distribution with mean defined as the mean of the point estimates of parameters in the region and large variance are set for the intercept terms of the regression coefficients ($\beta_{0,\mu}$

and $\beta_{0,\sigma}$), while normal distribution with zero mean and large variance are set for the other regression coefficients β.

However, informative priors are generally needed for the sill ($s_{\mu,t}$, $s_{\sigma,t}$) and correlation length ($l_{\mu,t}$, $l_{\sigma,t}$) parameters to avoid improper posteriors [5]. Since these parameters are not observable quantities, a preliminary analysis should be carried out to characterize the behavior of the experimental semi-variogram for μ and σ. Following the procedure proposed in [13] maximum likelihood estimates of μ and σ are computed at each location in the study region, and prior distributions for the parameters are chosen to define a wide envelope around the experimental semi-variogram given by the ML estimates.

2.4   Implementation of the model

In order to update each parameter θ_t of the described model a Metropolis–Hastings algorithm within a Gibbs sampler has been implemented. This hybrid MCMC algorithm [14] consists of a Gibbs sampler where a Metropolis step is used in order to sample from conditional distributions which are not known. Parameters of the model, which will be implemented for each time window t, are collected at each step i of the algorithm in the vector $\theta_t^{(i)}(\beta_{\mu,t}^{(i)}, \beta_{\sigma,t}^{(i)}, l_{\mu,t}^{(i)}, s_{\mu,t}^{(i)}, l_{\sigma,t}^{(i)}, s_{\sigma,t}^{(i)})$. Then, applying the Gibbs sampler, we partition the sampling for location μ and log-scale $log(\sigma)$ parameters and the next point in the chain $i + 1$, is generated in the following steps:

- Updating of correlation length parameter;
- Updating of sill parameter;
- Updating of regression parameters;
- Updating of EV parameter at each site;
- Repetition of the previous four steps for log-scale parameter.

A complete description of each step of the algorithm can be found in [15]. The algorithm is iterated checking the convergence for each parameters and finally, posterior densities of parameters θ_t are obtained. Implementing the model in the subsequent time windows t, the variation over time of posterior densities can be easily assessed, especially for EV parameters and consequently for return levels. In particular, for the definition of climatic actions on structures, we are interested in the evaluation of climate change impact on characteristic values c_k, i.e. value having a probability of 2% to be exceeded in one year (mean return period of 50 years) [16]. Therefore, posterior samples are easily computed for c_k according to

$$c^i_{k,t} = \mu^i_t + \sigma^i_t \left\{ -\log\left[-\log(1-0.02)\right]\right\} \tag{7}$$

and updated return level maps for characteristic loads can be easily drawn evaluating changes in the different time windows.

3   APPLICATION FOR GROUND SNOW LOADS

3.1   Study area and dataset

This section shows an application of the methodology presented in the previous section, on extreme ground snow loads considering the Zone 3-4 of the Italian Mediterranean climatic region defined by the Annex C to EN1991-1-3 [4]. The study region is shown in Figure 1 and comprises D=272 cells at which climate projections are provided by the highest resolution Regional Climate Models developed by the EUROCORDEX initiative [17].
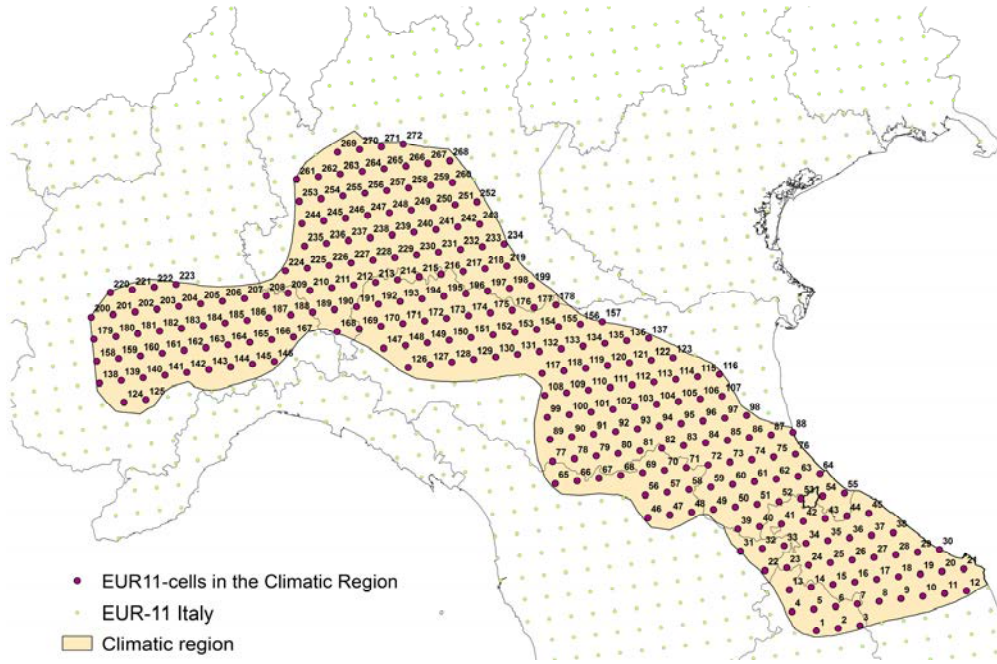
Figure 1: Investigated region.

Climate projections provided by an ensemble of *r*=6 RCMs for the period 1951-2100 have been analyzed, considering a medium emission scenario RCP4.5 and the highest emission scenario RCP8.5 [18]. The main characteristics of the investigated climate projections are reported in Table 1.

Institute	RCM	GCM	Period	Experiment
DMI	HIRHAM5	EC-EARTH	1951-2100	Historical,RCP4.5,RCP8.5
CLMcom	CCLM4-8-17	CNRM-CM5-LR	1951-2100	Historical,RCP4.5,RCP8.5
CLMcom	CCLM4-8-171	EC-EARTH	1951-2100	Historical,RCP4.5,RCP8.5
KNMI	RACMO22E	EC-EARTH	1951-2100	Historical,RCP4.5,RCP8.5
MPI-CSC	REMO2009	MPI-ESM-LR	1951-2100	Historical,RCP4.5,RCP8.5
IPSL-INERIS	WRF331F	CM5A-MR	1951-2100	Historical,RCP4.5,RCP8.5

Table 1: Overview on the analyzed climate projections and their main characteristics.

3.2 Implementation and results

In order to derive ground snow loads from regional climate models output such as daily temperatures and precipitation, the procedure described in [1] and [15] has been implemented deriving series of N=140 yearly maxima snow load for each cell in the study region.

Among possible covariate information, altitude shows most significant influence on extreme snow loads, it has been then considered as the only covariate and a quadratic model has been chosen as defined in the Eurocode EN1991-1-3 [4] for characteristic ground snow load in Mediterranean region. Then, covariate matrix X and the vectors of regression coefficients $\beta_{\mu,t}$ and $\beta_{\sigma,t}$ in eq. 5 and 6 become

$$X = \begin{bmatrix} 1 & alt_1 & alt_1^2 \\ \vdots & \vdots & \vdots \\ 1 & alt_D & alt_D^2 \end{bmatrix}; \beta_{\mu,t} = \begin{bmatrix} \beta_{\mu,0,t} \\ \beta_{\mu,1,t} \\ \beta_{\mu,2,t} \end{bmatrix}; \beta_{\sigma,t} = \begin{bmatrix} \beta_{\sigma,0,t} \\ \beta_{\sigma,1,t} \\ \beta_{\sigma,2,t} \end{bmatrix} \tag{7}$$

The model have been implemented for each time window t, and the MCMC algorithm has been iterated 40 000 times, obtaining posterior densities of random parameters $\theta_t(\beta_{\mu,t}, \beta_{\sigma,t}, l_{\mu,t}, s_{\mu,t}, l_{\sigma,t}, s_{\sigma,t})$. As an example in Figure 2 the results in terms of posterior densities of location μ and scale σ EV parameters, but also q_k, are presented for one cell, i=160, in the study region in different time windows (t = 1, 4, 8, 10) according to one of the investigated climate model (first RCM in Table 1).
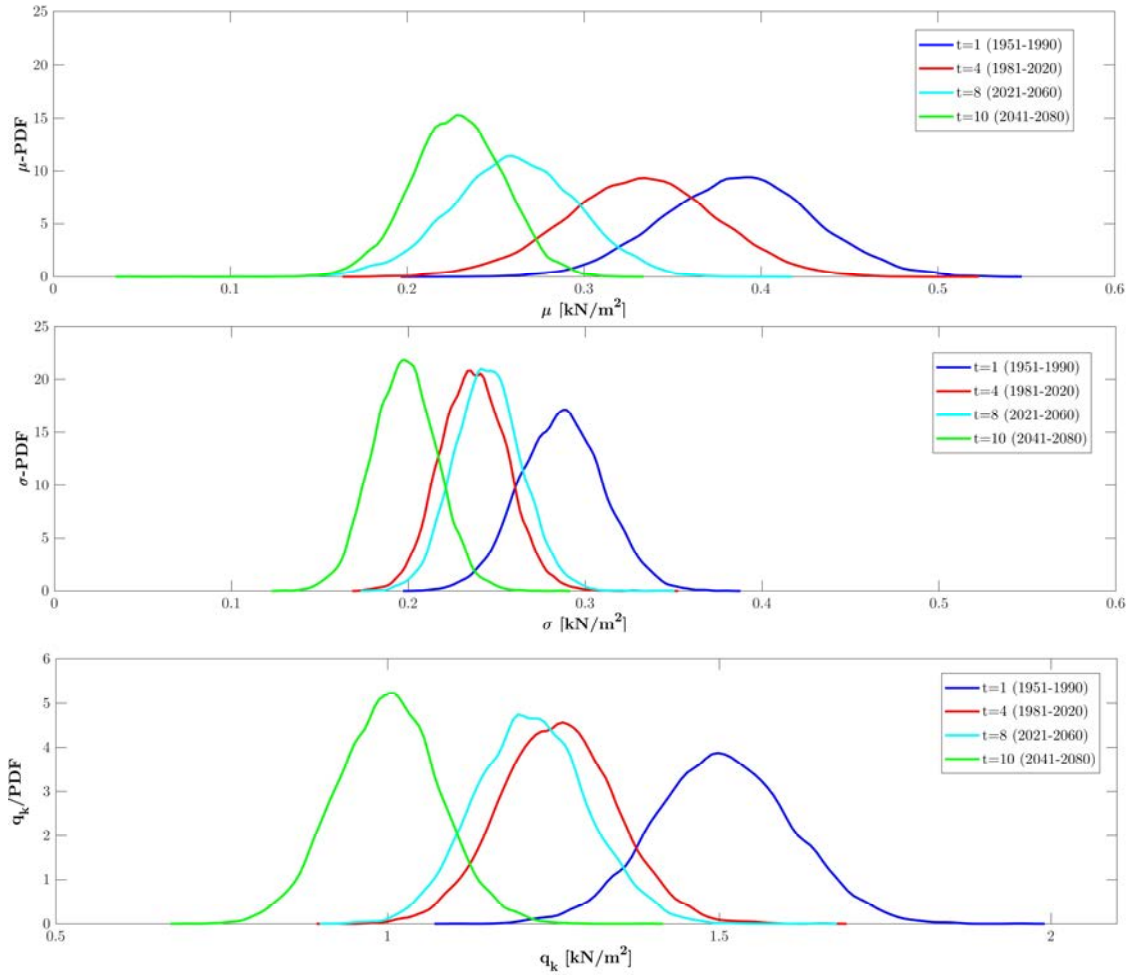

Figure 2: Changes in posterior PDFs of μ, σ and q_k with time t.

The hierarchical model combined with the Bayesian approach enables a direct assessment of the uncertainties affecting the extreme value process using the posterior distribution of parameters and return values, as shown in Figure 2. Moreover, the implementation of the model in subsequent time window allows a direct estimation of the effect of climate change on extreme ground snow loads by means of the analysis of changes in posterior densities of EV parameters and return values.

The spatial pooling of the data provides an added value in comparison with classical approach based on maximum likelihood estimates at point level leading to more precise and less variable estimates [3]. The reduced uncertainty in the estimation is shown in Figure 2 where

q_k estimates obtained by the presented spatial model are compared with the classical site by site analysis according the maximum likelihood method, for some cells at increasing distance in the study region. The results in terms of 95% confidence interval clearly show the reduced uncertainty for the illustrated spatial model confirming the advantages of spatial pooling for tail estimation.
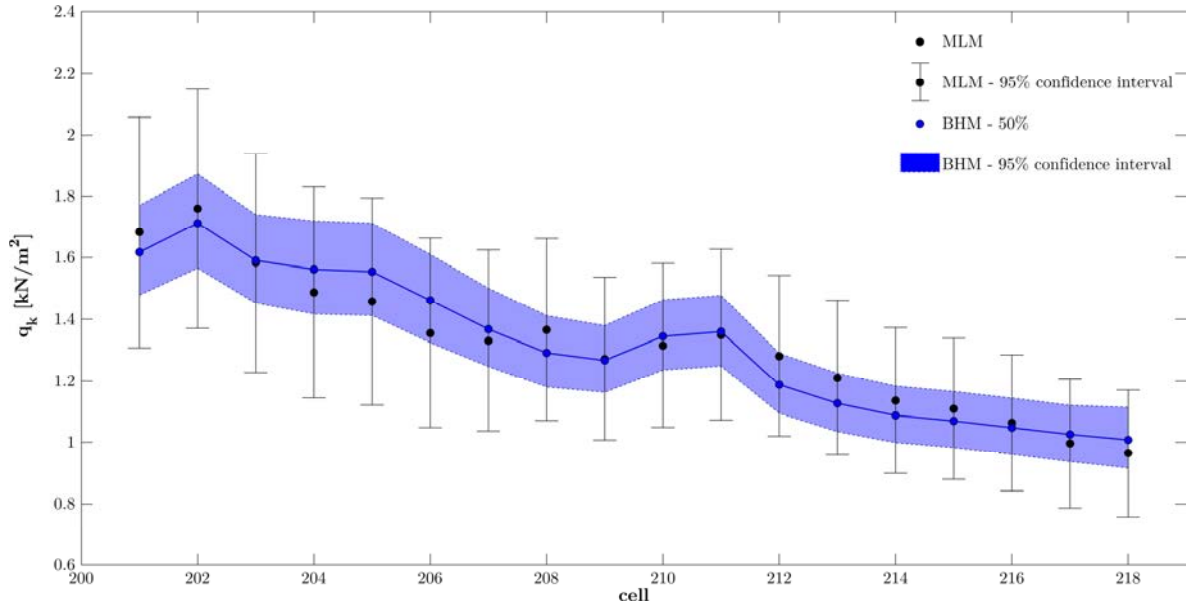

Figure 3: q_k estimates by maximum likelihood method (MLM) and Bayesian hierarchical model (BHM) with 95% confidence intervals.

3.3 Return level and Factor of Change Maps

Return level maps can also be drawn from the posterior samples of q_k obtained according the investigated climate models and scenarios. However, more information about climate change impact can be derived by the definition of factors of change (FC) as the difference or the ratio of predictions from RCM in the future period and the historical period.

The factor of change approach has a long history in climate change impact studies, it is based on the assumption that changes in the observed climate variables form present to future are the same than changes predicted by the climate models not requiring to apply bias correction methods.

Factor of change maps represent a good solution for the assessment and the visualization of future trends in climatic actions since the estimated changes can be easily applied to the current version of climatic load maps in structural Codes.

Therefore, from the posterior samples of q_k, mean and standard deviation for *FC* are computed

$$\overline{FC}(q_{k,t}) = \frac{\overline{q_{k,t}}}{q_{k,1}}; \sigma_{FC(q_{k,t})} = \frac{\sigma_{q_{k,t}}}{q_{k,1}} \tag{8}$$

As an example maps for mean FC and standard deviation are reported in Figure 4 and 5 respectively, considering t=8 (2021-2060) and the six RCMs in Table 1 run according the RCP4.5 scenario.
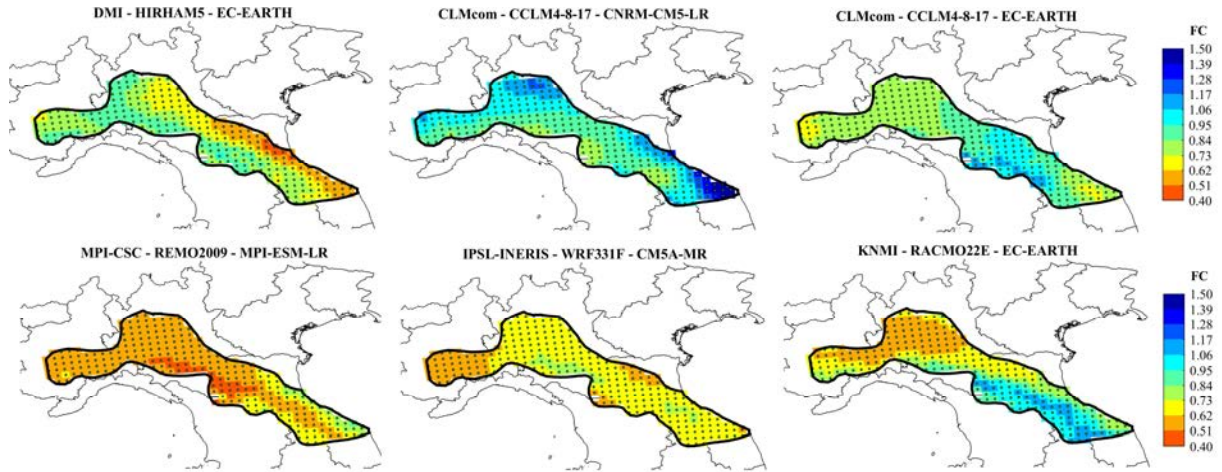
Figure 4: Posterior mean of q_k Factor of Change for 2021-2060 w.r.t. 1951.1990 according to the climate models in Table 1, Scenario RCP4.5.
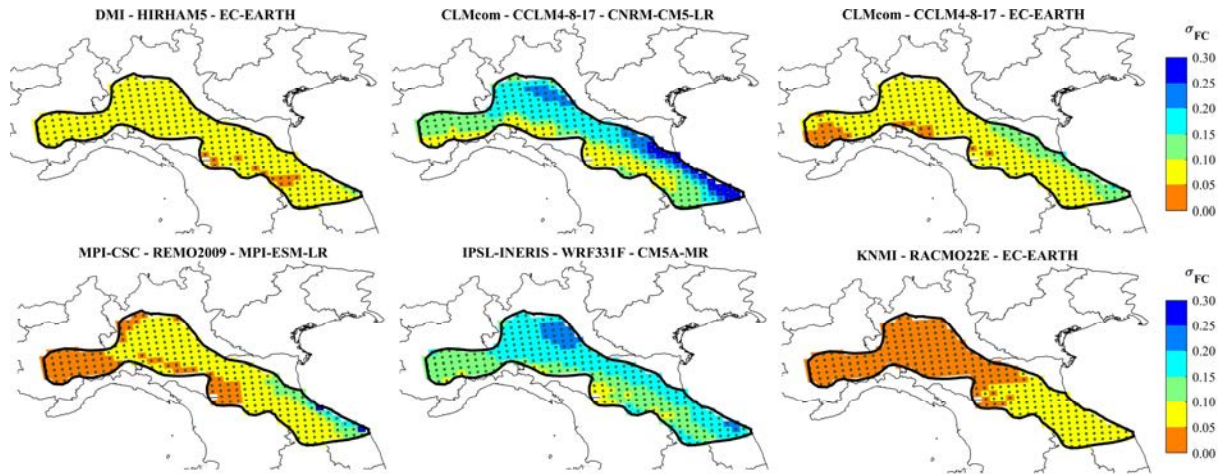

Figure 5: Posterior standard deviation of q_k Factor of Change for 2021-2060 w.r.t. 1951.1990 according to the climate models in Table 1, Scenario RCP4.5.

The results obtained for the different climate models can be finally combined considering each climate model of the ensemble as an equally likely representation of future climate. In this way, a complete probabilistic description of future changes in characteristic loads is obtained providing guidance for potential amendments of the current version of climatic load maps in structural Codes.

In Figure 6, the results in terms of factor of change maps for characteristic ground snow load q_k are presented in a bivariate map, which consider the 25-75% prediction interval for FC, for three time windows (1991-2030, 2011-2050, and 2031-2080) according the RCP4.5 and RCP8.5 scenario (second and third row respectively). In the same Figure, on the top row, the current snow load map for the study region, obtained implementing the load altitude relationship given in the Annex C of EN1991-1-3 [4] and based on the results of the European Snow Load Research Project[19], which analyzed observed data series of ground snow loads in the period 1951-1990, is also reported.
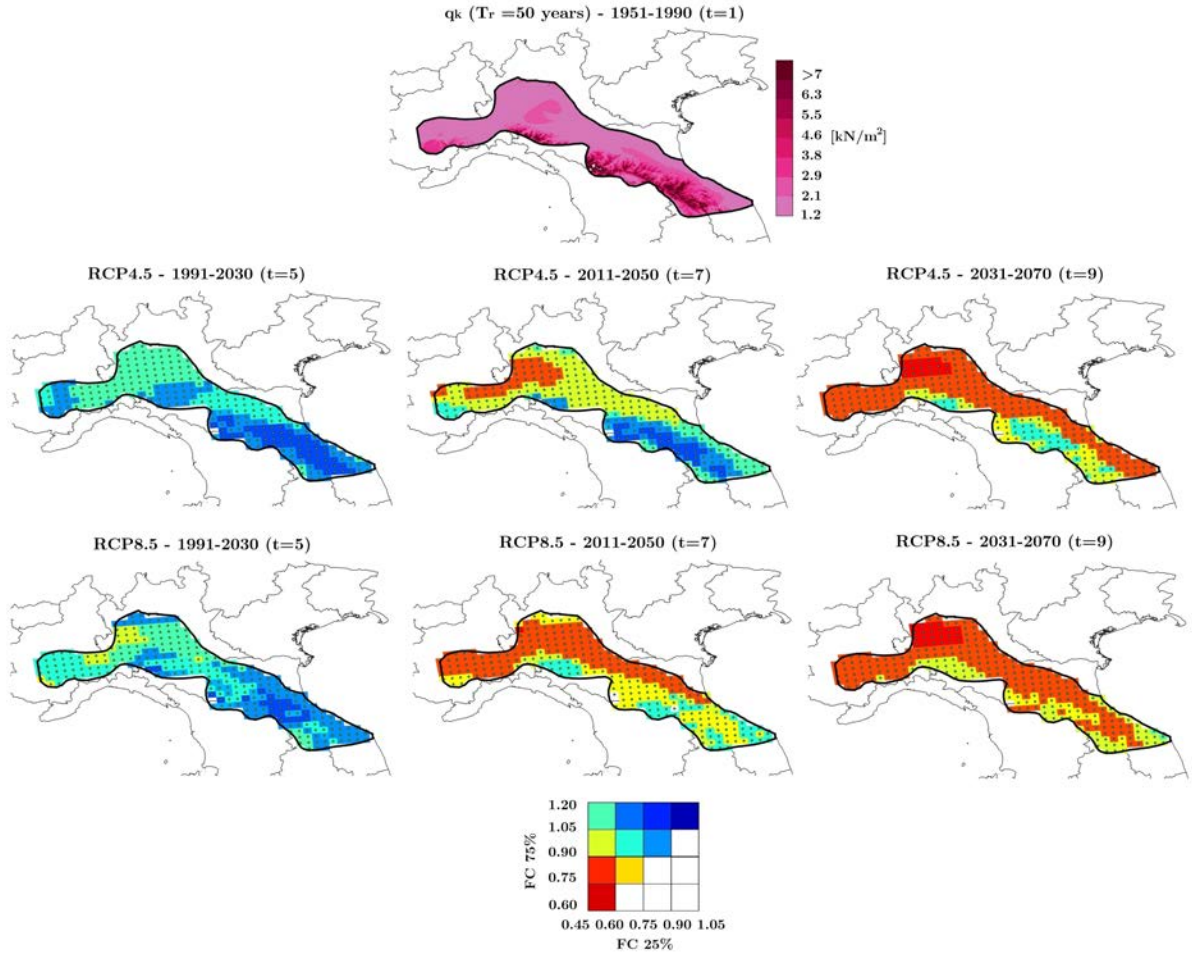
Figure 6: Factors of Change for q_k - Confidence interval [25-75%] Map (Scenario RCP4.5).

4 CONCLUSIONS

In order to estimate future changes in climatic actions on structures, a methodology based on the construction of a Bayesian hierarchical model for the characterization of climate extremes derived from the analysis of high-resolution climate model output has been presented.

The model is formulated through the classical three-level formulation, in which the standard extreme value representation at each site is combined with a spatial latent process, and it is implemented in different time windows to assess climate change effects on the extreme value process.

An application on ground snow loads has been carried out to illustrate the capabilities of the proposed methodology. The results shows that the Bayesian framework enables a direct assessment of the uncertainties affecting the prediction of the extreme value parameters and return levels. Moreover, the spatial pooling of the data leads to more precise and less variable estimates with respect to classical approaches based on maximum likelihood estimates at point level.

Finally, combining the results obtained for each climate model, suitable factors of change uncertainty maps are drawn providing guidance for potential amendments of the climatic load maps in structural Codes.

REFERENCES

[1] P. Croce, P. Formichi, F. Landi, F. Marsili, Climate change: Impact on snow loads on structures. *Cold Regions Science and Technology*, 150, 35–50, 2018.

[2] P. Croce, P. Formichi, F. Landi, F. Marsili, Evaluating the effect of climate change on thermal actions on structures. In *Life-Cycle Analysis and Assessment in Civil Engineering: Towards an Integrated Vision*, 1751–1758, 2019.

[3] E. Casson, S. Coles, Spatial Regression Models for Extremes. *Extremes*, 449–468, 1999.

[4] CEN. EN 1991-1-3:2003 - Eurocode 1: Actions on structures - Part 1-3: General actions - Snow loads. 2003.

[5] S. Banerjee, B. P. Carlin, A. E. Gelfand, *Hierarchical modeling and analysis for spatial data*. CRC Press LLC, 2004.

[6] D. Cooley, D. Nychka, P. Naveau, Bayesian Spatial Modeling of Extreme Precipitation Return Levels. *Journal of the American Statistical Association*, 479, 824–840, 2007.

[7] D. Cooley, S. R. Sain, Spatial hierarchical modeling of precipitation extremes from a regional climate model. *Journal of Agricultural, Biological, and Environmental Statistics*, 15, 381–402, 2010.

[8] S. R. Sain, R .Furrer, N. Cressie, A spatial analysis of multivariate output from regional climate models. *Annals of Applied Statistics*, 1, 150–175, 2011.

[9] P. Formichi, et al.; *Eurocodes: background and applications. Elaboration of maps for climatic and seismic actions for structural design with the Eurocodes*, JRC Science for Policy Report.

[10] H. Xu, P. Gardoni, Improved latent space approach for modelling non-stationary spatial–temporal random fields, *Spatial Statistics*, 23, 160–181, 2018.

[11] M. D. Risser, C. A. Calder, Regression-based covariance functions for nonstationary spatial modeling, *Environmetrics*; 26, 284–297, 2015.

[12] P. Croce, F. Landi, P. Formichi, R. Castelluccio, Use of weather generators to assess impact of climate change : thermal actions on structures. In *Proc. of the Fifth Intl. Conf. Advances in Civil, Structural and Mechanical Engineering - CSM 2017*, 32–36, 2017.

[13] D. Cooley, D. Nychka, P. Naveau. Bayesian Spatial Modeling of Extreme Precipi-tation Return Levels. J*ournal of the American Statistical Association*, 102, 824–40, 2007.

[14] C. P. Robert, , G. Casella. *Monte Carlo Statistical Methods*, Springer, New York., 2004.

[15] F.Landi, A General Methodology for the Assessment of the Impact of Climate Change on Snow Loads on Structures, Phd Thesis, University of Pisa – TU Braunschweig.

[16] CEN. EN1990 Eurocode - Basis of structural design. 2002.

[17] D. Jacob, et al. EURO-CORDEX: New High-Resolution Climate Change Projections for European Impact Research. *Regional Environmental Change*, 14, 563–78, 2014.

[18] Van Vuuren, et al., The representative concentration pathways: an overview, *J. Climatic Change*, 109, 5-31, 2011.

[19] L. Sanpaolesi, et al. *Phase 1 Final Report to the European Commission, Scientific Support Activity in the Field of Structural Stability of Civil Engineering Works: Snow Loads*. Tech. rep., Department of Structural Engineering, University of Pisa, 1998.