where,

$$t(x) = \begin{cases} (1+\xi(\frac{x-\mu}{\sigma}))^{-1/\xi} & \xi \neq 0 \\ \exp\left(-\frac{x-\mu}{\sigma}\right) & \xi = 0. \end{cases}$$

Then the log-likelihood can be given by,

$$\begin{aligned} l(\theta|Y_1, Y_2, Z_1, X) = \sum_{i=1}^{n} & - \log(\sigma_{y_{2i}}) + (\xi_{y_{2i}} + 1)\log(t(y_{2i})) - t(y_{2i}) \\ & - \log(\sigma_{y_{1i}}) + (\xi_{y_{1i}} + 1)\log(t(y_{1i})) - t(y_{1i}) \\ & - \log(\sigma_{z_{1i}}) + (\xi_{z_{1i}} + 1)\log(t(z_{1i})) - t(z_{1i}). \end{aligned}$$

Our choice of priors are essentially non-informative with variances chosen to ensure proper coverage of the sample space and reasonably good acceptance rates in the Metropolis-Hastings algorithm. These are:

$$\begin{aligned} \alpha_i &\overset{iid}{\sim} \mathrm{N}(0, 10^2), i = 1, \ldots, p \\ \beta_i &\overset{iid}{\sim} \mathrm{N}(0, 10^3), i = 1, \ldots, p+1 \\ \gamma_i &\overset{iid}{\sim} \mathrm{N}(0, 10^2), i = 1, \ldots, p+2 \\ \sigma_{z_1}, \sigma_{Y_1}, \sigma_{Y_2} &\overset{iid}{\sim} \mathrm{IG}(\alpha = 1, \beta = 3) \\ \xi_{z_1} &\sim \mathrm{Unif}(-1, 1) \\ \xi_{Y_1}, \xi_{Y_2} &\overset{iid}{\sim} \mathrm{Unif}(-0.55, 0.5). \end{aligned}$$

The notations N, IG and Unif respectively stand for the Normal/Gaussian distribution, the inverse Gamma distribution and the Uniform distribution.

For the Bayesian computations, we use the Metropolis Hastings sampling scheme to obtain the Markov Chain Monte Carlo (MCMC) chain for $N = 10^6$ steps and the step-sizes are chosen to achieve about 20% acceptance rate. For starting values in the MCMC algorithm, we use the frequentist estimates of each of individual GEV models, (9), (10) and (11).

Variable selection: Since all 11 covariates may not be relevant for each layer of the hierarchy, we select relevant variables based on the posterior results obtained from the full hierarchical model. In order to establish relevance of the covariates in the hierarchical model layers, we use the concept of 1-D depth. We look at the marginal posterior distribution for each parameter, and calculate the one-dimensional depth. Let F_β be the (one-dimensional) marginal posterior distribution of a parameter, say, β, then we compute an empirical estimate of $4F_\beta(0)(1 - F_\beta(0))$, where we plug-in the following empirical estimator in place of F_β:

$$\hat{F}_\beta(0) = \frac{\sum_{j=1}^{M} I\{\beta^{(j)} \leq 0\}}{M},$$

where M is the number of MCMC samples obtained for the parameter β, i.e., $\{\beta^{(j)}, j = 1, \ldots, M\}$. The closer this value is to 0, the farther away zero (thinking about it as a hypothesis testing problem to test for $\beta \neq 0$) is in the tails of the distribution and the more relevant that variable is. The closer this value gets to 1, that is indicative of zero being the median of the distribution. We fit this hierarchical model for the entire dataset prior to the test year to select the important variables. In Table 1, the blue and purple colored variables are relevant using this metric for the entire dataset (1960-2022). We shade in blue the variables that seem to have a depth value 0 or close to 0 (suggesting them being far off in the tails) while in purple are the variables with values between 0.2 and 0.6, which suggest that these estimates are still far in the tail ($\sim$ between the 80 and 95 percentile). Next we refit the heirarchical Bayesian model on these selected variables (blue and purple colored ones). Then, we calculate the posterior means and standard deviations for this selective model, which are then tabulated in Table 2.

Variables	Min CP	Max WS	Damages
Intercept	0.0000	0.0000	0.0000
Min CP (scaled)	NA	0.0000	0.0177
Max WS (scaled)	NA	NA	0.2479
Avg. Latitude	0.0000	0.8318	0.2290
Avg. Longitude	0.0009	0.9698	0.9965
StartMonth	0.3456	0.9949	0.3351
Year	0.1122	0.9620	0.5877
NAO	0.9956	0.9973	0.9991
SOI	0.9912	1.0000	0.4244
AMO	0.4748	0.9967	0.5665
ANOM.3.4	0.9513	0.9974	0.0486
Atl_SST	0.0021	0.9996	0.0091
Sunspots	0.9461	0.9830	0.9998
ξ	0.0000	0.0000	0.0018
σ	0.0000	0.0000	0.0000

Table 1: 1D data depth to measure how relevant a variable is in each of the layers of the model.

We highlight that the shape (ξ) and scale (σ) parameters for all the three models are highly significant, thus validating the use of the GEV models for modeling the extreme behavior of these natural events. The negative estimate for the shape parameters signifying reverse Weibull distributions for the marginals of each of the three variables, log(minCP), log(maxWS) and log(damages), respectively. In addition, average latitude, average longitude, AMO, and Atlantic SST are statistically significant in modeling the location parameter for minCP. Similarly, the effect of minCP (β_1) in modeling location parameter for maximum wind speed seems to be significant. The effect of maximum wind speed, minimum central pressure, average latitude, ANOM 3.4, and Atlantic SST are significant in modeling the location parameter in the GEV model for log(damages). We also note that the significant variables are mostly the same across the three different modeling schemes that we have employed, namely, the hierarchical generalized extreme value Bayesian model, the trivariate generalized extreme value model (Section A.1), and the hierarchical generalized extreme