Optimisation

2024-10-16

The r-Pareto process

Our rainfall process is $X = \{X_{s,t}, (s,t) \in S \times T\}$, where S is the spatial domain and T the temporal domain.

We have $X_{s,t} \mid r(X_{s,t}) > u \xrightarrow{d} Z_{s,t}$ with a risk function $r(X_{s,t}) = X_{s_0,t_0}$, a threshold $u > 1$ and $Z = \{Z_{s,t}, (s,t) \in S \times T\}$ a r-Pareto process.

With $W = \{W_{s,t}, (s,t) \in S \times T\}$ the Gaussian process of the Brown-Resnick process and its variogram γ, we can define the r-Pareto process as

$$Z_{s,t} = R_{s,t} e^{W_{s,t} - W_{s_0,t_0} - \gamma(s-s_0, t-t_0)}$$

where $R_{s,t}$ is a random variable following a simple Pareto distribution.

The variogram γ is defined as

$$\gamma(ds, dt) = \beta_1 |ds|^{\alpha_1} + \beta_2 dt^{\alpha_2}$$

with $ds = s - s'$, $dt = t - t'$, $\beta_1, \beta_2 > 0$ and $\alpha_1, \alpha_2 \in (0,1)$.

If we add an advection vector $V = (v_x, v_y)$, the variogram becomes

$$\gamma(ds, dt) = \beta_1 |ds - V dt|^{\alpha_1} + \beta_2 dt^{\alpha_2}.$$

The r-Pareto process without advection

Simulation

We simulate the r-Pareto process with the parameters $\beta_1 = 0.4$, $\beta_2 = 0.2$, $\alpha_1 = 1.5$, $\alpha_2 = 1$ and without advection. We simulate the process on a 5×5 grid with 30 time steps and $m = 100$ realizations. We use a conditonal point $s_0 = (1,1)$ at time $t_0 = 1$.

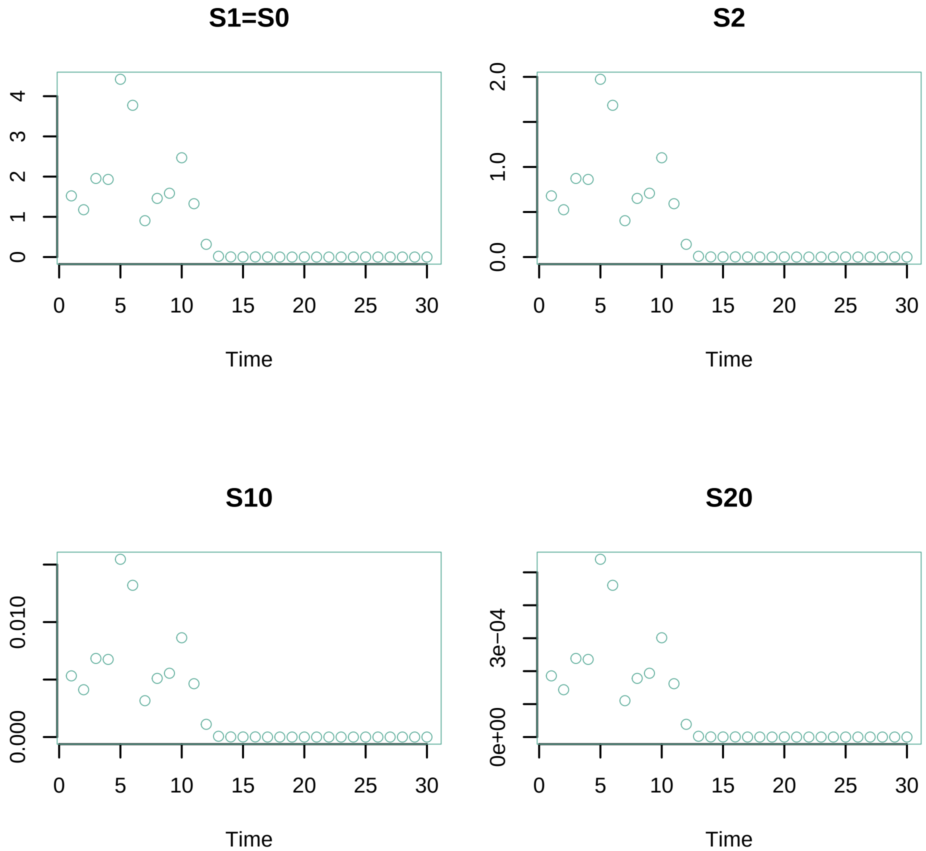


Figure 1: Time series for 4 sites of the first realization

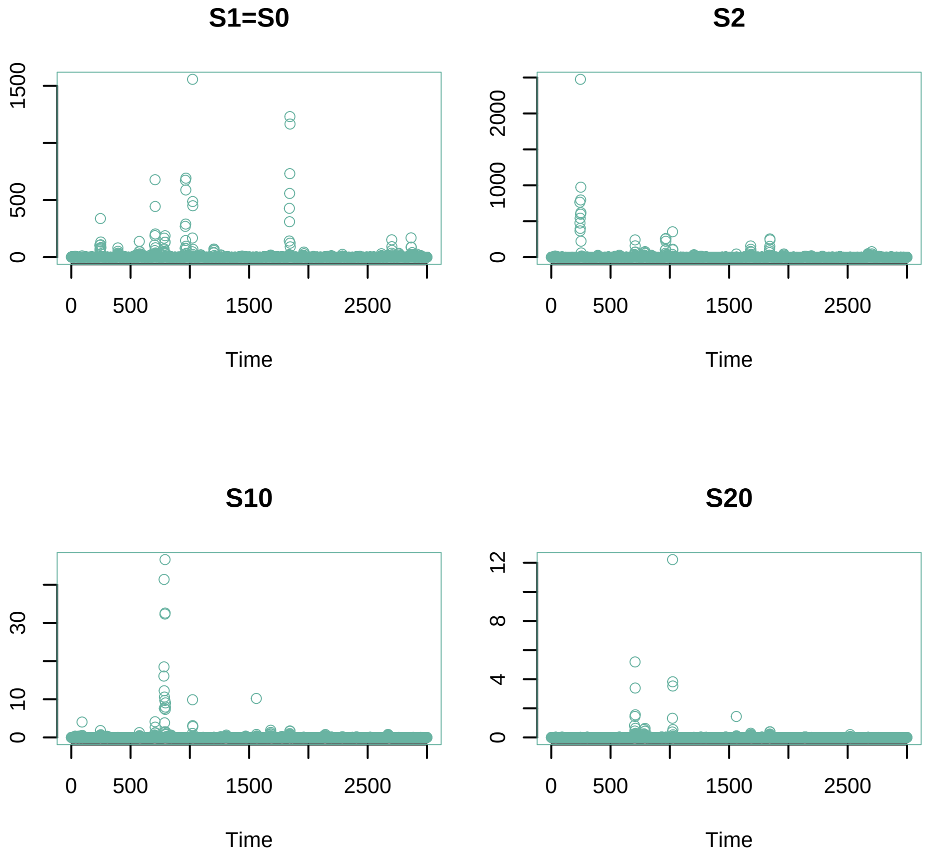


Figure 2: Time series for 4 sites of the all replicates together

Probability of exceedances for the r-Pareto process without advection

For (s_0, t_0)

Verification that $P(X_{s_0,t_0} > u) = 1$ because of the r-Pareto process construction.

```
## [1] "X_s0_t0 = 1.521 > u = 1"
```

```
## [1] "P(X_s0,t > u) = 0.04"
```

For all replicates we can see that the value of the process at the conditional point:

```
## [1] "Minimum value of the process at the conditional point: 1.007 > u = 1"
```

```
## [1] "Maximum value of the process at the conditional point: 104.665"
```

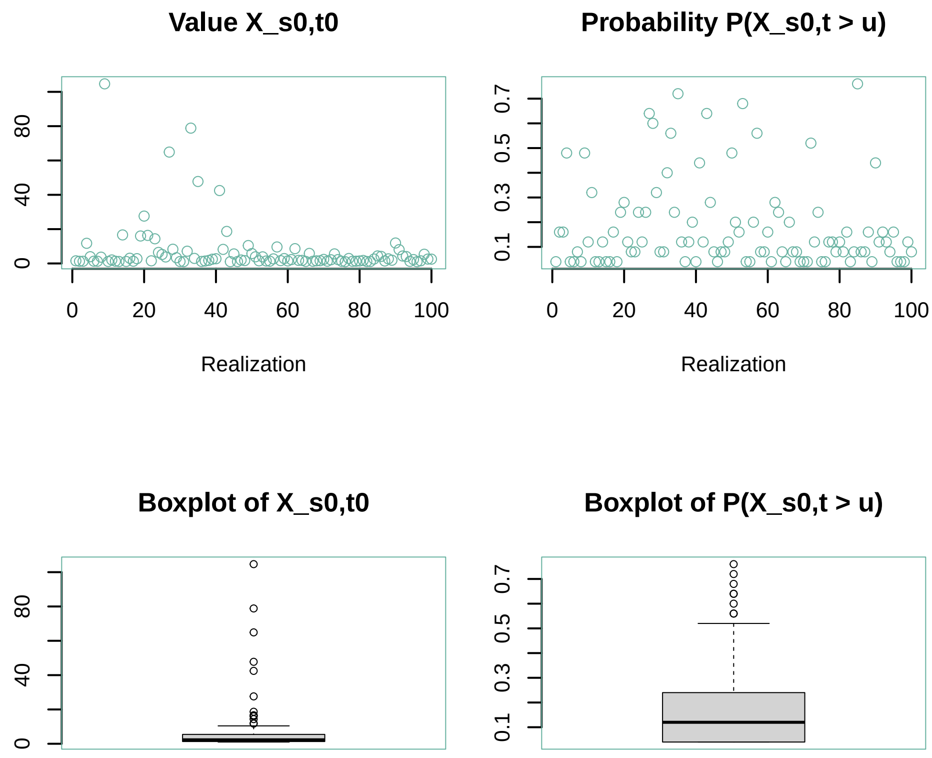


Figure 3: Value and probability at conditional point for all replicates

Marginal exceedance probability

We have different probability of exceedances for each site as we can see in the following plot.

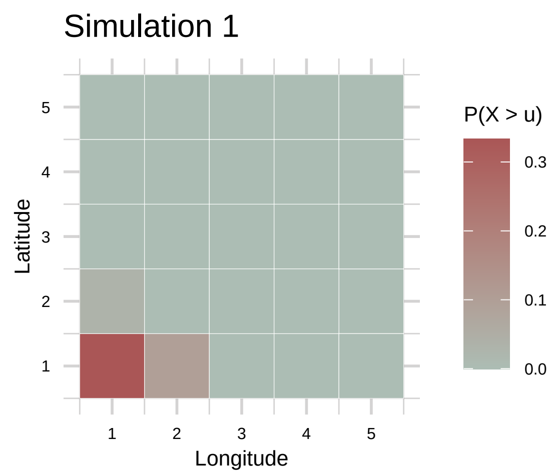


Figure 4: Marginal probability of exceedances for each site for one simulation

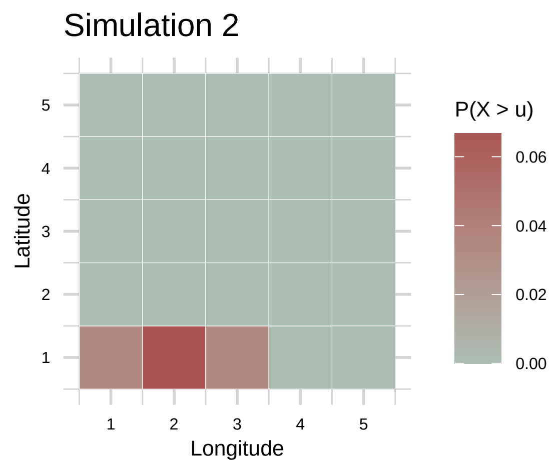


Figure 5: Marginal probability of exceedances for each site for one simulation

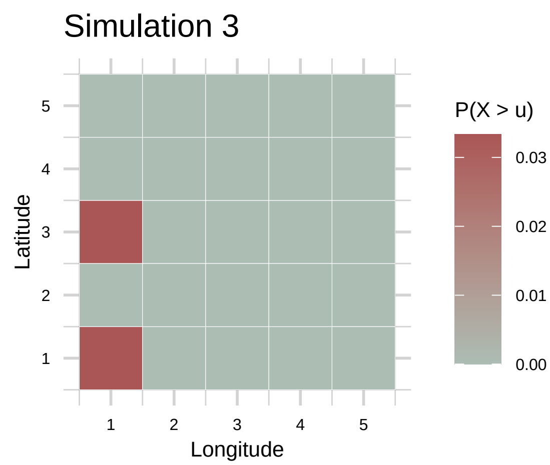


Figure 6: Marginal probability of exceedances for each site for one simulation

Multiple replicates

We have \$m=\$100 replicates of the r-Pareto process. We concate the m replicates to have a single simulation of the r-Pareto process. Here we can see that the higher probability of exceedances are around the conditional point $(1, 1)$.

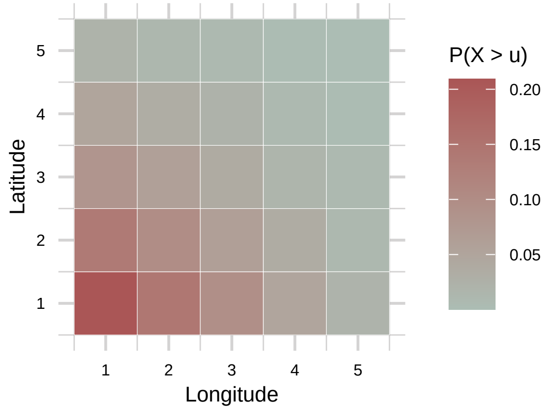


Figure 7: Marginal probability of exceedances for m replicates in a single dataframe

Estimation of the variogram parameters

Fixing advection at $(0, 0)$

Table 1: Result and RMSE for all replicates together with advection=(0.5,0.3)

	beta1	beta2	alpha1	alpha2
estim	0.5440019	0.2749222	1.2150634	0.764817
rmse	0.1440019	0.0749222	0.2849366	0.235183

Without fixing advection

Table 2: Result and RMSE for all replicates together with advection=(0,0)

	beta1	beta2	alpha1	alpha2	adv1	adv2
estim	0.5748966	0.2555284	1.1889737	0.8036342	-0.0181456	-0.0141513
rmse	0.1748966	0.0555284	0.3110263	0.1963658	0.0181456	0.0141513

For multiple sets of replicates

We can look at $M = 100$ simulations with $m = 100$ replicates each.

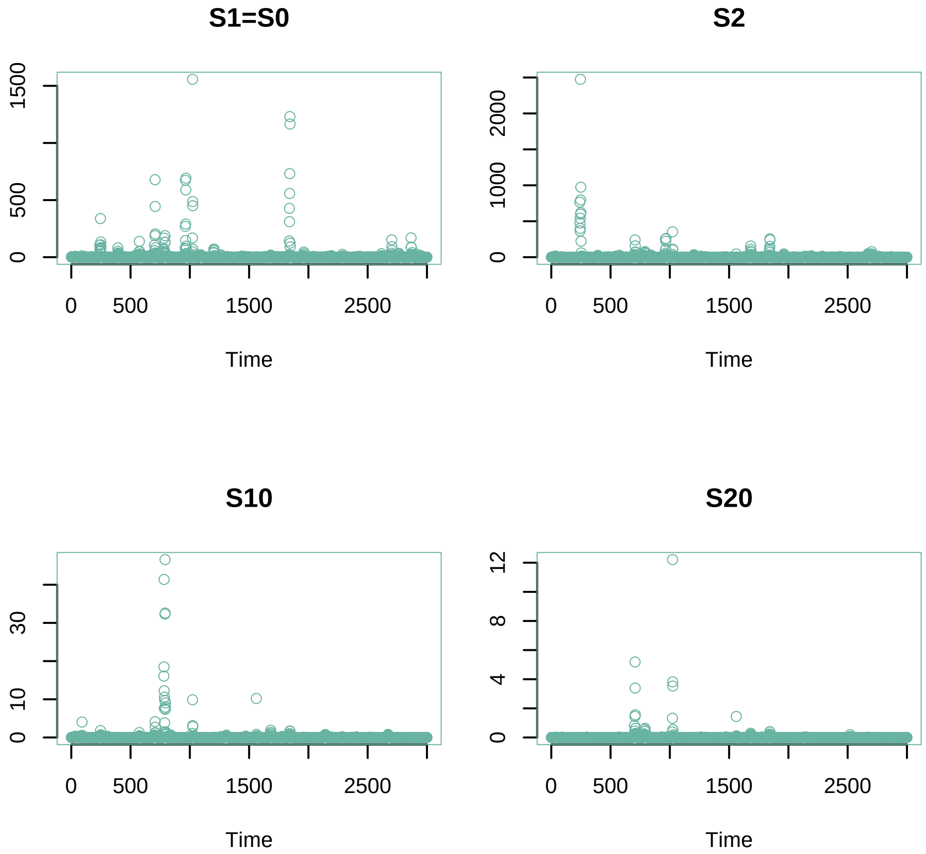


Figure 8: Time series for 4 sites of the first realization

Optimisation of the variogram parameters for the $M = 100$ simulations with $m = 100$ replicates each.

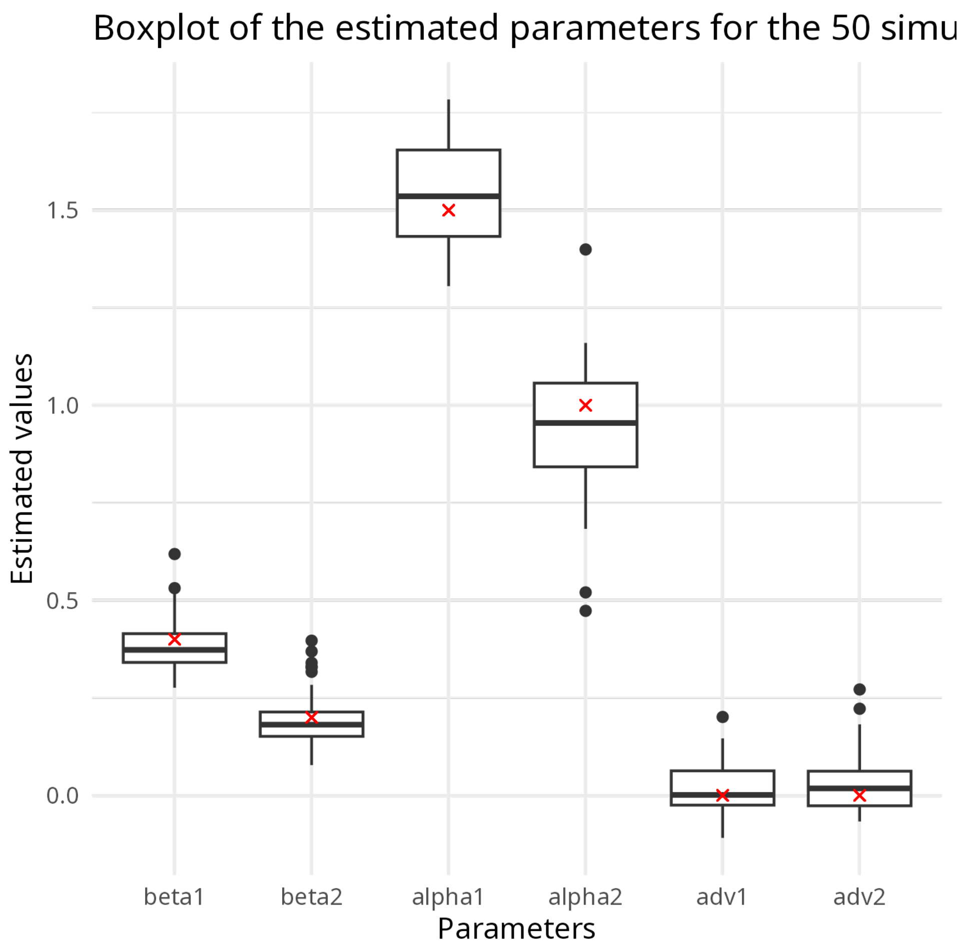


Figure 9: Boxplot of the estimated parameters for the 100 simulations of 100 replicates each

Table 3: Result and RMSE for all replicates together with advection=(0,0)

	Mean	RMSE	MAE
beta1	0.3856815	0.0702273	0.0554053
beta2	0.1930979	0.0707383	0.0523818
alpha1	1.5396520	0.1284241	0.1083304
alpha2	0.9371214	0.1726222	0.1264208
adv1	0.0204490	0.0665498	0.0506481
adv2	0.0338401	0.0841809	0.0583286

The r-Pareto process with advection

Simulation

We simulate the r-Pareto process with the parameters $\beta_1 = 0.4$, $\beta_2 = 0.2$, $\alpha_1 = 1.5$, $\alpha_2 = 1$ and the advection vector $V = (0.5, 0.3)$. We simulate the process on a 5×5 grid with 30 time steps and 100 realizations. We use a conditonal point $s_0 = (1, 1)$ at time $t_0 = 1$.

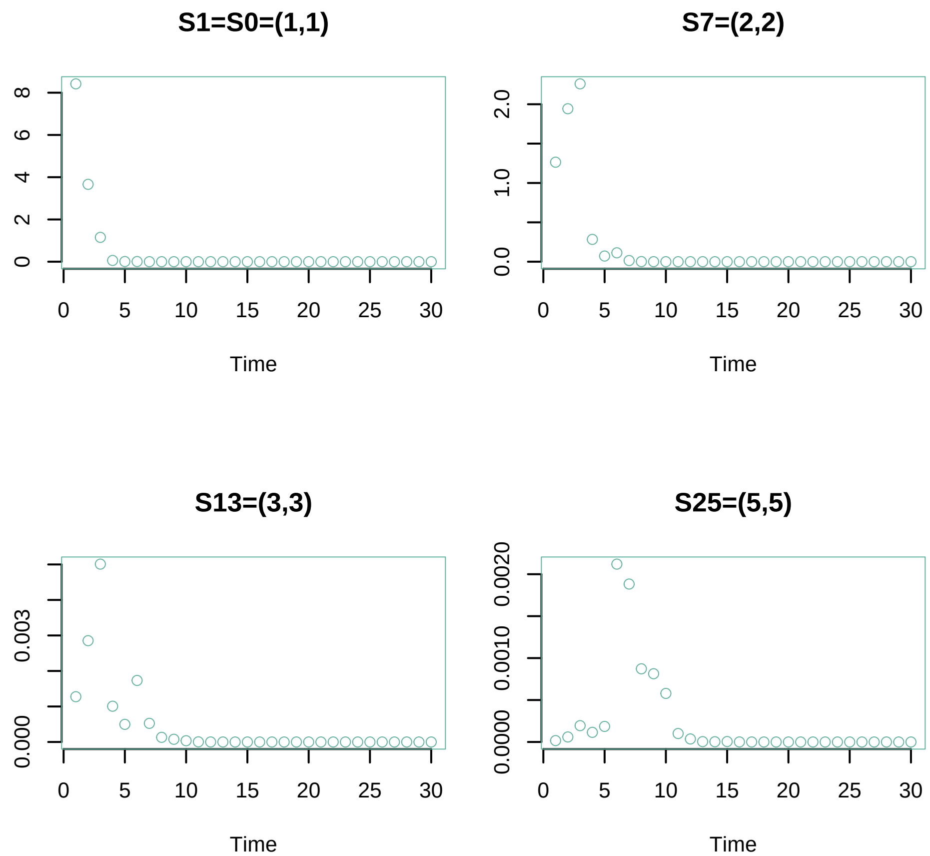


Figure 10: Time series for 4 sites of the first realization in the advection direction

Probability of exceedances for the r-Pareto process

For (s_0, t_0)

Verification that $P(X_{s_0,t_0} > u) = 1$ because of the r-Pareto process construction.

```
## [1] "X_s0_t0 = 12.631 > u = 1"
```

```
## [1] "P(X_s0,t > u) = 0.04"
```

For all replicates we can see that the value of the process at the conditional point:

```
## [1] "Minimum value of the process at the conditional point: 1.038 > u = 1"
```

```
## [1] "Maximum value of the process at the conditional point: 40.598"
```

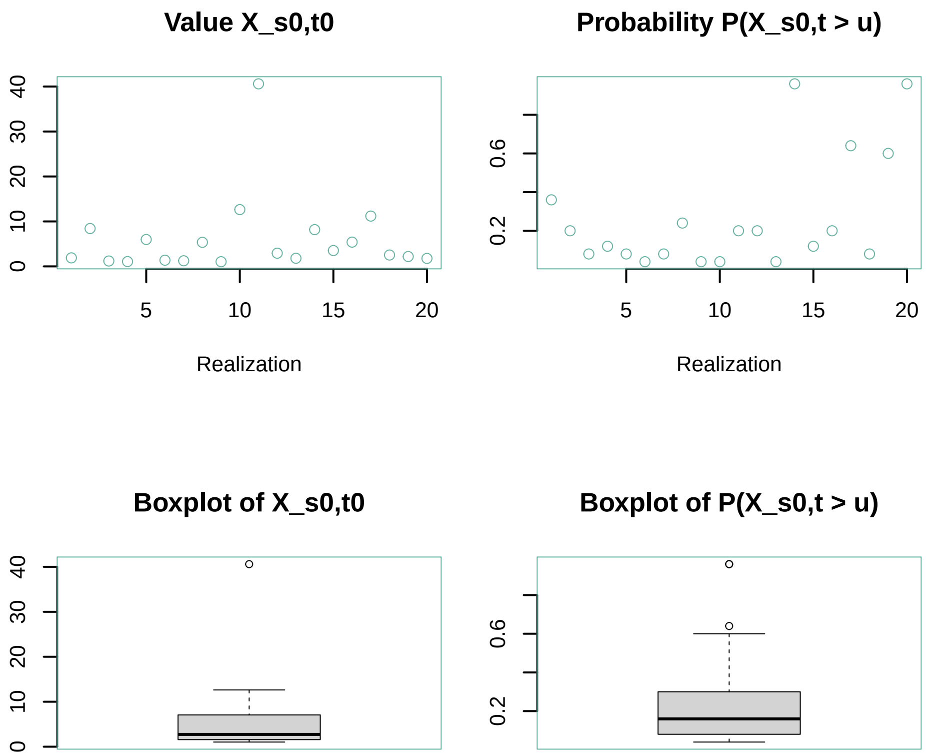


Figure 11: Value and probability at conditional point for all replicates

Marginal exceedance probability

We have different probability of exceedances for each site as we can see in the following plot.

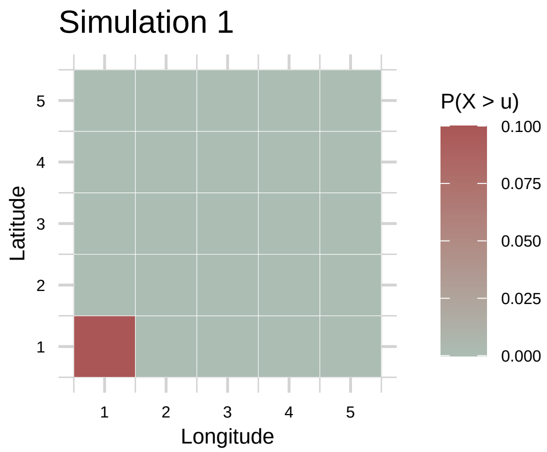


Figure 12: Marginal probability of exceedances for each site for one simulation

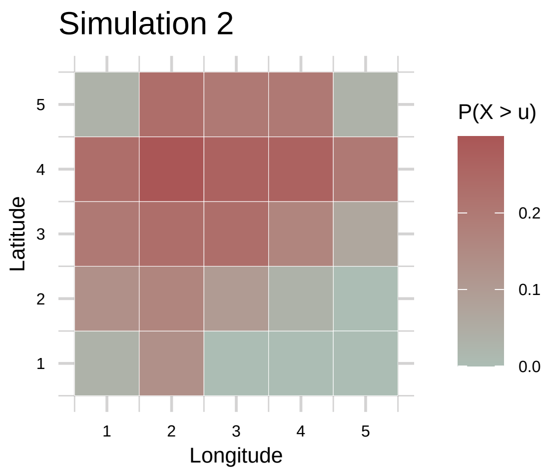


Figure 13: Marginal probability of exceedances for each site for one simulation

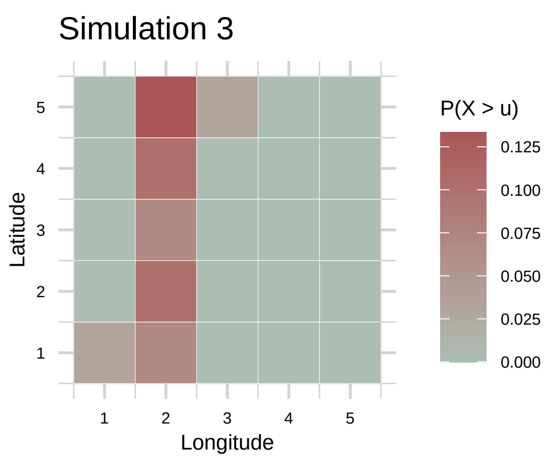


Figure 14: Marginal probability of exceedances for each site for one simulation

TODO : verifier la forme de ma grille dans mes gifs, j'ai un doute

For multiple replicates of the r-Pareto process

We have \$m=\$20 replicates. We concate the m replicates to have a single simulation of the r-Pareto process.

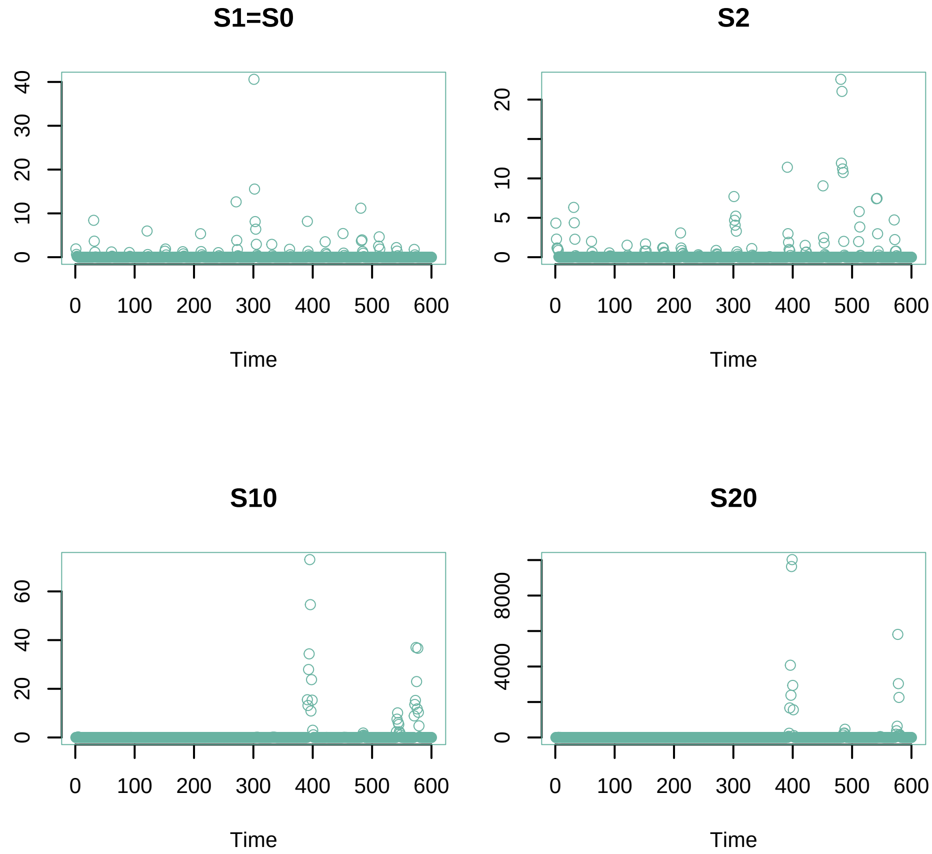


Figure 15: Time series for 4 sites of the replicates together

The probability of exceedances for each site is computed for all replicates.

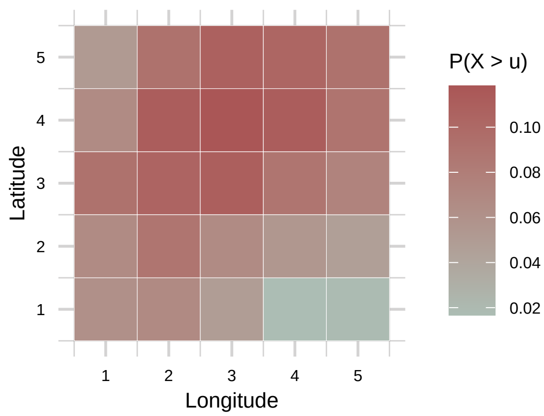


Figure 16: Marginal probability of exceedances for m replicates in a single dataframe

Optimisation

We want to estimate the parameters β_1, β_2, α_1 and α_2 of the r-Pareto process. We use the maximum likelihood estimation with the composite likelihood method.

We compute the number of joint excesses for each replicate i, $k_{s,t}^{(i)} = \sum_{t=1}^{T} \mathbb{1}_{\{X_{s,t}>u, X_{s_0,t_0}>u\}}$ and $k_{s-s_0,t-t_0}^{(i)} \sim Bin(T-t-t_0, \chi(s-s_0, t-t_0))$ with T the number of observations within a replicate (same for all replicates).

Plot by fixing params

Brown-Resnick process

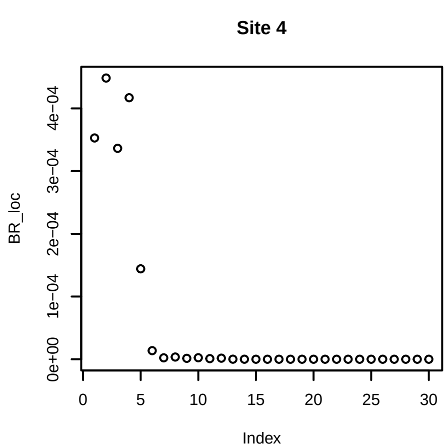
Simulation without advection

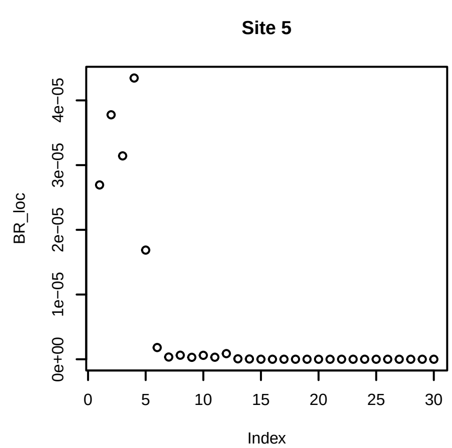
We simulate the Brown-Resnick process with the parameters $\beta_1 = 0.4$, $\beta_2 = 0.2$, $\alpha_1 = 1.5$, $\alpha_2 = 1$. We simulate the process on a 5×5 grid with 300 time steps and 100 realizations. First without advection.

Gumbel margins

```
## $conv
## [1] 0
##
## $nllh
## [1] 133.4754
##
## $mle
## [1] -37.35053  18.98189
##
## $se
## [1] 3.686181 2.541837
```

```
## $conv
## [1] 0
##
## $nllh
## [1] 131.1445
##
## $mle
## [1] -36.42685  17.66557
##
## $se
## [1] 3.432226 2.347046
```





Marginal exceedance probability

Do we have same exceedance probability for each site?

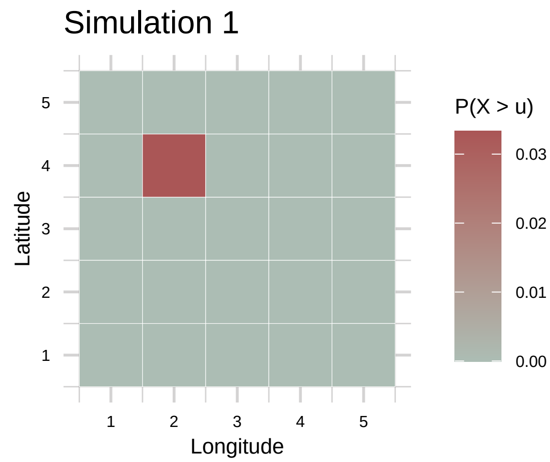


Figure 17: Marginal probability of exceedances for each site for one simulation

Bulh model WLSE

Validation du modèle de Buhl avec WLSE, cela ne marche pas bien pour toutes les simus, pq ?

```
## [1] 0.02865715 2.04001704
```

```
## [1] 1.1447848 0.3390315
```

	estim	rmse	mae
beta1	0.0286571	0.3713429	0.3713429
beta2	1.1447848	0.9447848	0.9447848
alpha1	2.0400170	0.5400170	0.5400170
alpha2	0.3390315	0.6609685	0.6609685

Optimisation

Table 4: Optim estimations for each quantile for one simulation

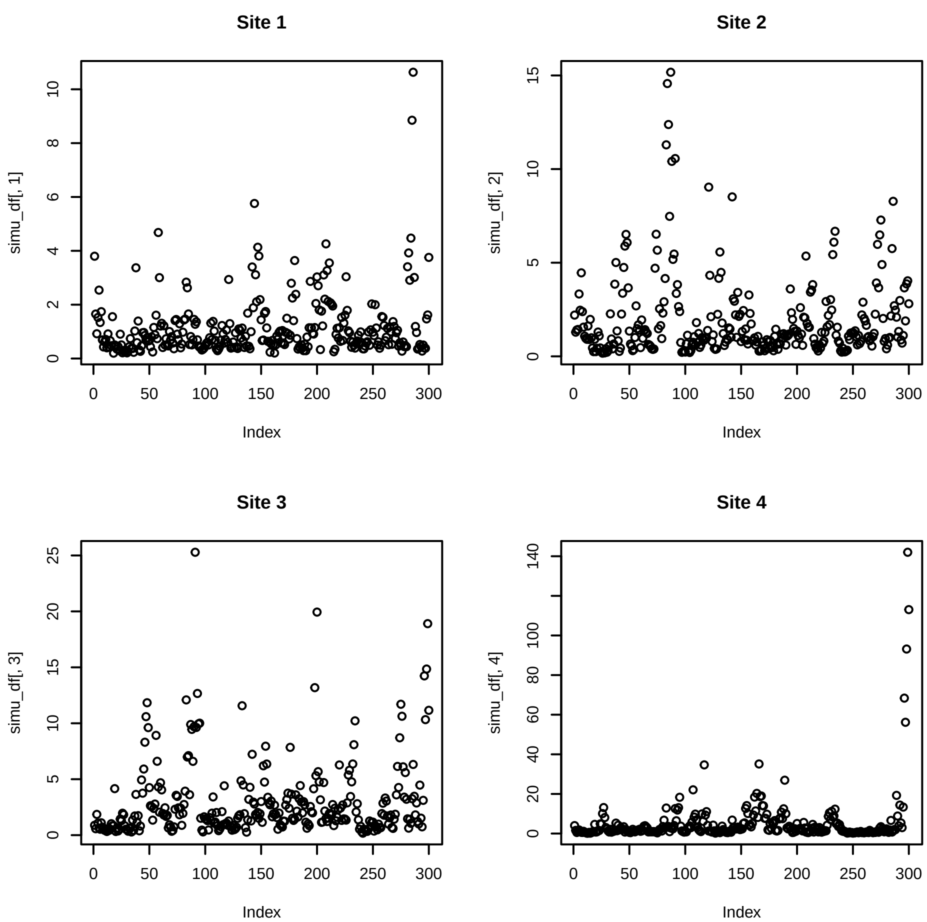
q	beta1	beta2	alpha1	alpha2	adv1	adv2
0.80	0.04491	0.02374	1.99819	1.08996	4.20809	4.45610
0.85	0.13023	0.00057	1.22463	1.07291	0.38356	3.93086
0.90	0.15133	0.17541	1.68175	0.98920	7.73834	7.53216
0.95	0.47691	0.98003	1.30139	1.32691	-7.92567	1.26492

Table 5: RMSE for each parameter and different quantiles for one simulation

q	rmse_beta1	rmse_beta2	rmse_alpha1	rmse_alpha2	rmse_adv1	rmse_adv2
0.80	0.35509	0.17626	0.49819	0.08996	3.70809	4.15610
0.85	0.26977	0.19943	0.27537	0.07291	0.11644	3.63086
0.90	0.24867	0.02459	0.18175	0.01080	7.23834	7.23216
0.95	0.07691	0.78003	0.19861	0.32691	8.42567	0.96492

Simulation with advection

We simulate the Brown-Resnick process with the parameters $\beta_1 = 0.4$, $\beta_2 = 0.2$, $\alpha_1 = 1.5$, $\alpha_2 = 1$. We simulate the process on a 5×5 grid with 300 time steps and 100 realizations. We take into account the advection with $\text{adv} = (0.1, 0.5)$.

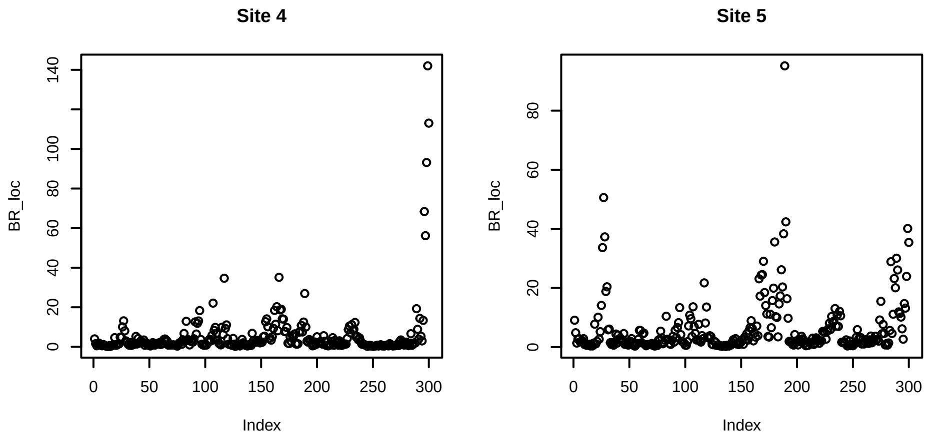


Gumbel margins

```
## $conv
## [1] 0
##
## $nllh
## [1] 468.0574
```

```
##
## $mle
## [1] 0.2536208 1.0078038
##
## $se
## [1] 0.06145882 0.04422571

## $conv
## [1] 0
##
## $nllh
## [1] 481.7131
##
## $mle
## [1] 0.4844236 1.0637453
##
## $se
## [1] 0.06496573 0.04643145
```



Marginal exceedance probability

Do we have same exceedance probability for each site?

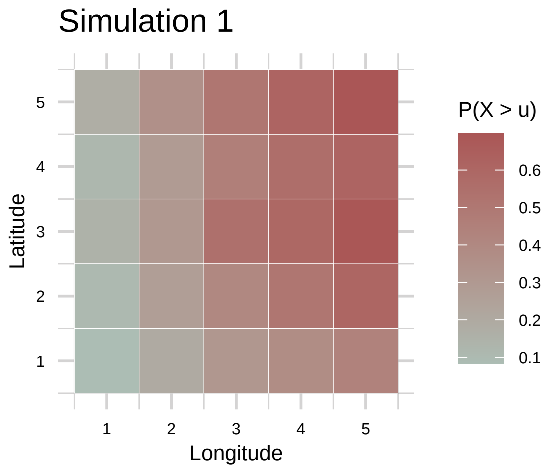


Figure 18: Marginal probability of exceedances for each site for one simulation

Bulh model WLSE

Validation du modèle de Buhl avec WLSE, cela ne marche pas bien pour toutes les simus, pq ?

```
## [1] 0.4993708 1.4859984
```

```
## [1] 0.3173238 0.9565931
```

	estim	rmse	mae
beta1	0.4993708	0.0993708	0.0993708
beta2	0.3173238	0.1173238	0.1173238
alpha1	1.4859984	0.0140016	0.0140016
alpha2	0.9565931	0.0434069	0.0434069

Optimisation