Optimisation

2024-10-21

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 $t_0 = 1$.

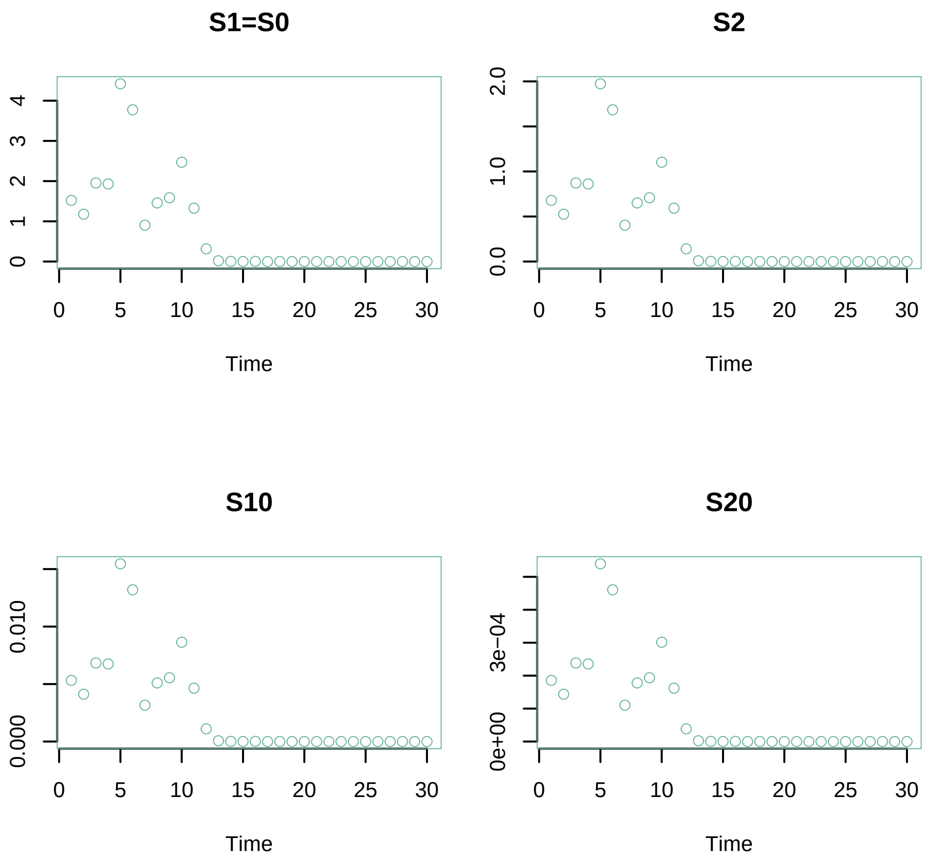

Figure 1: Time series for 4 sites of the first realization

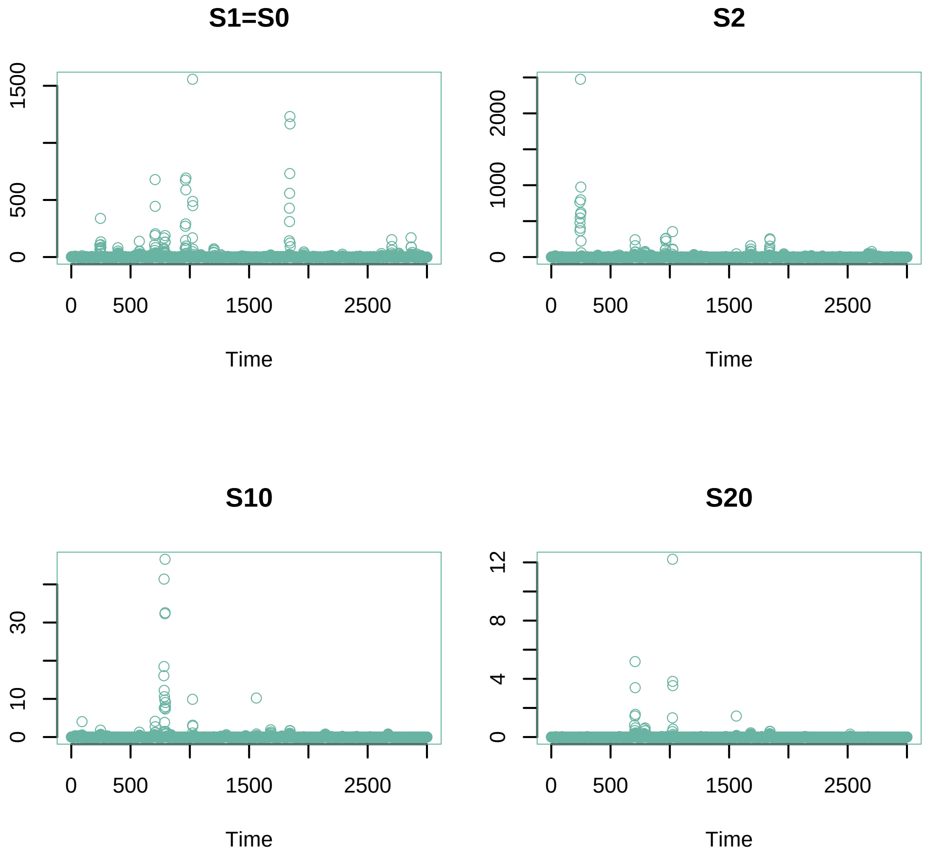

Figure 2: Time series for 4 sites of the all replicates together

Estimation of the variogram parameters

Table 1: 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

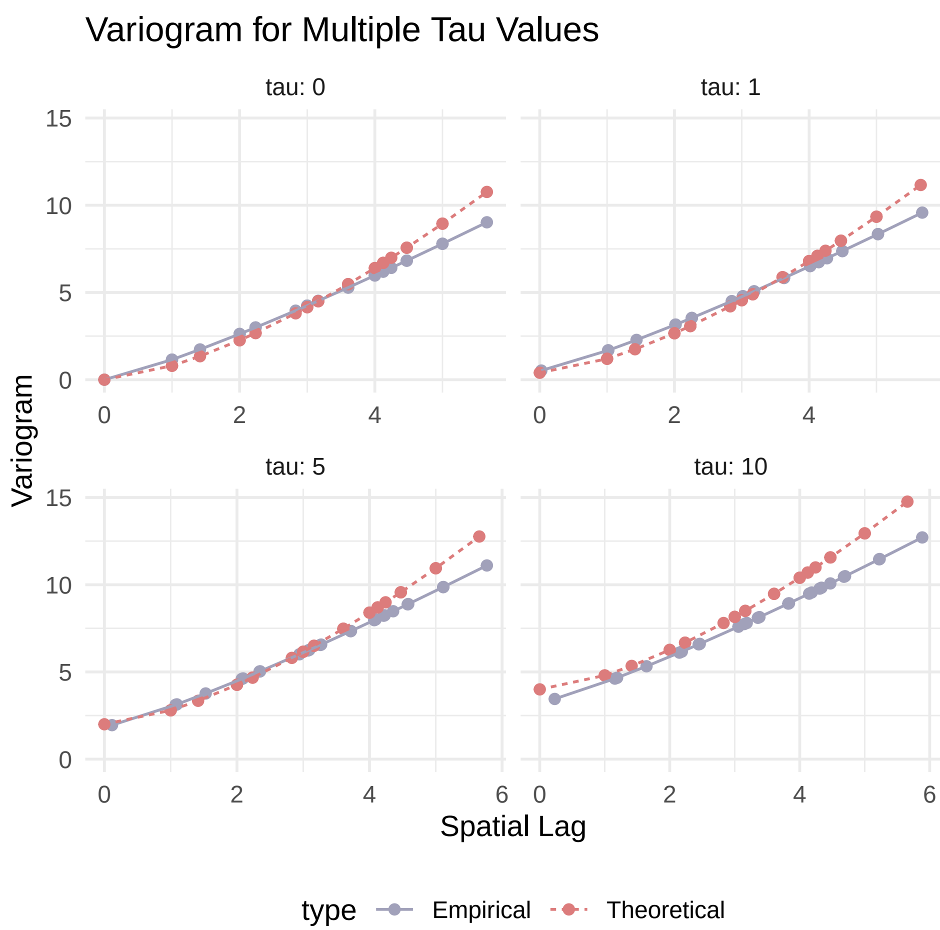

Figure 3: Variogram of the estimated parameters

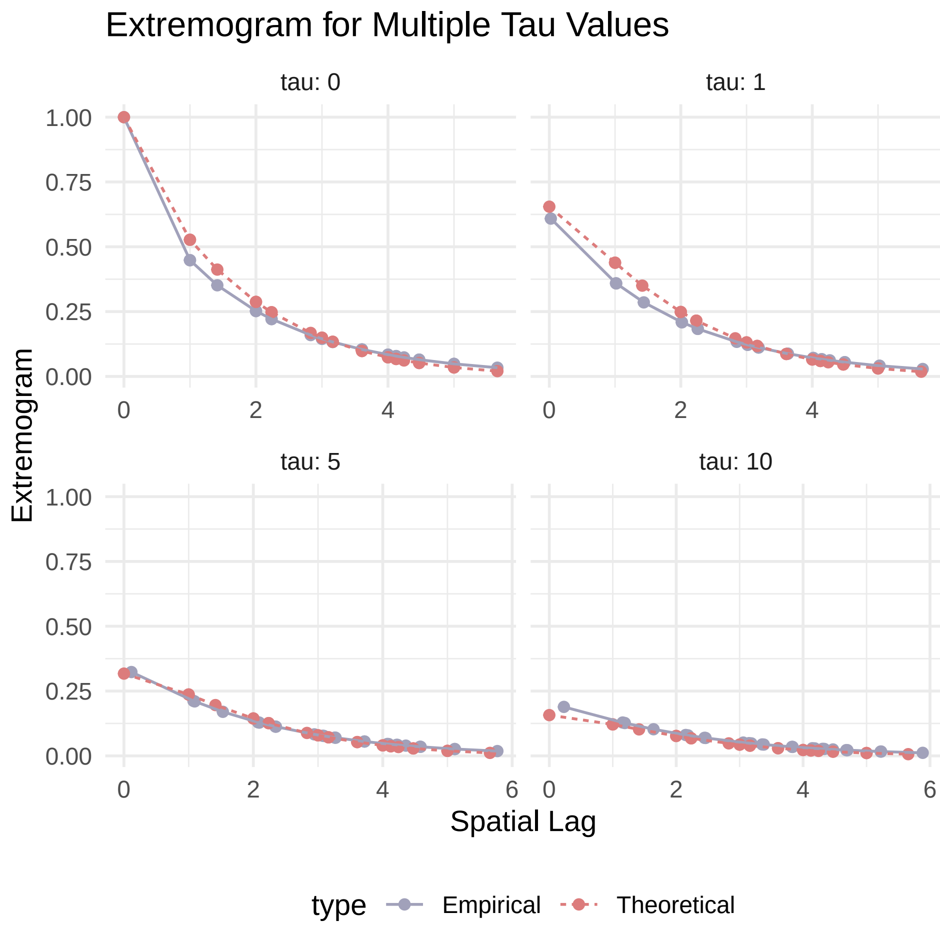

Figure 4: Boxplot of the estimated parameters

For multiple sets of replicates

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

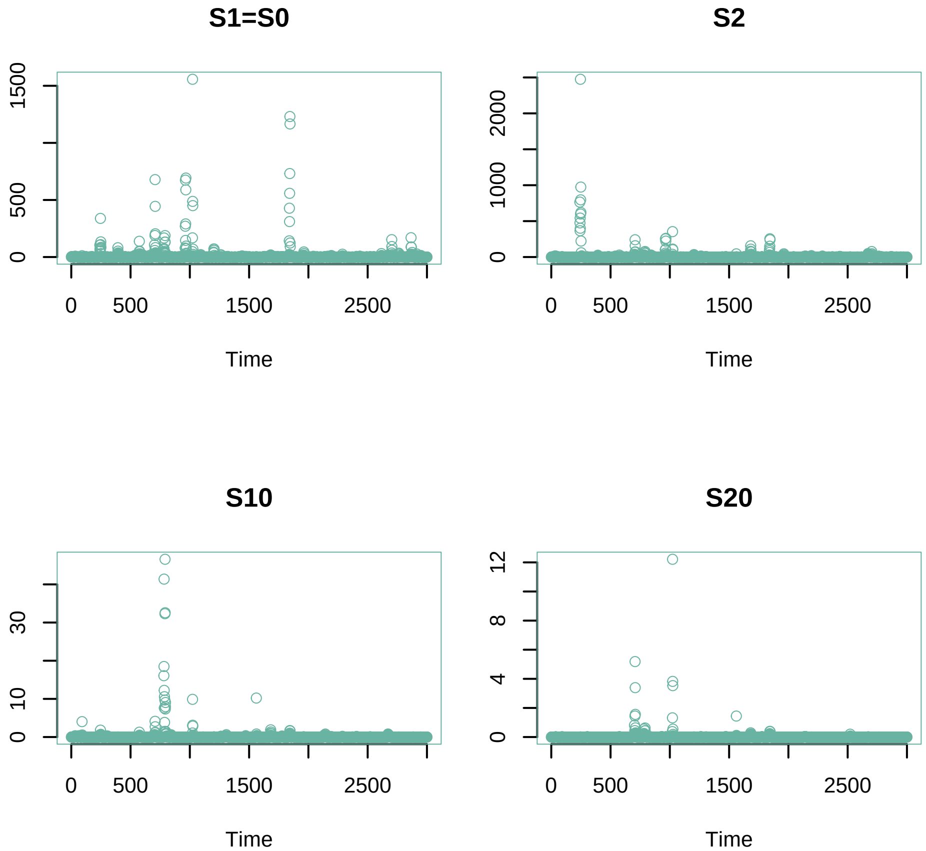

Figure 5: Time series for 4 sites of the first realization

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

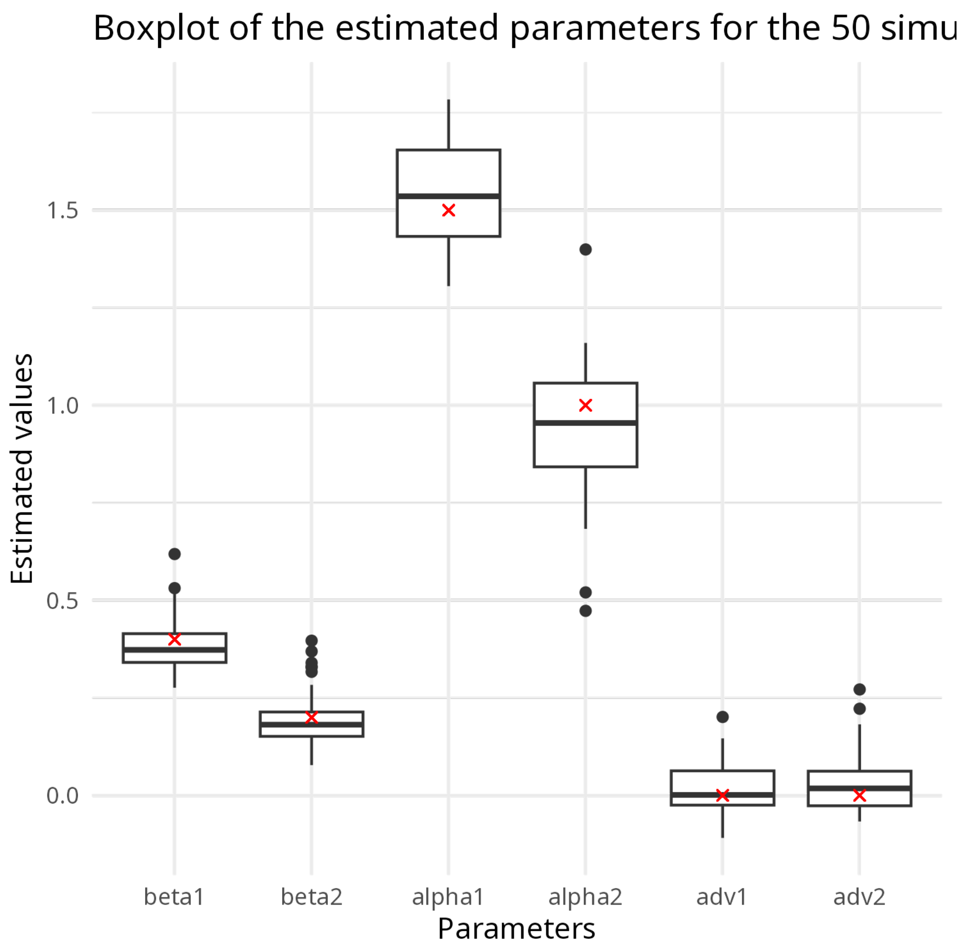

Figure 6: Boxplot of the estimated parameters for the 100 simulations of 100 replicates each

Table 2: 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

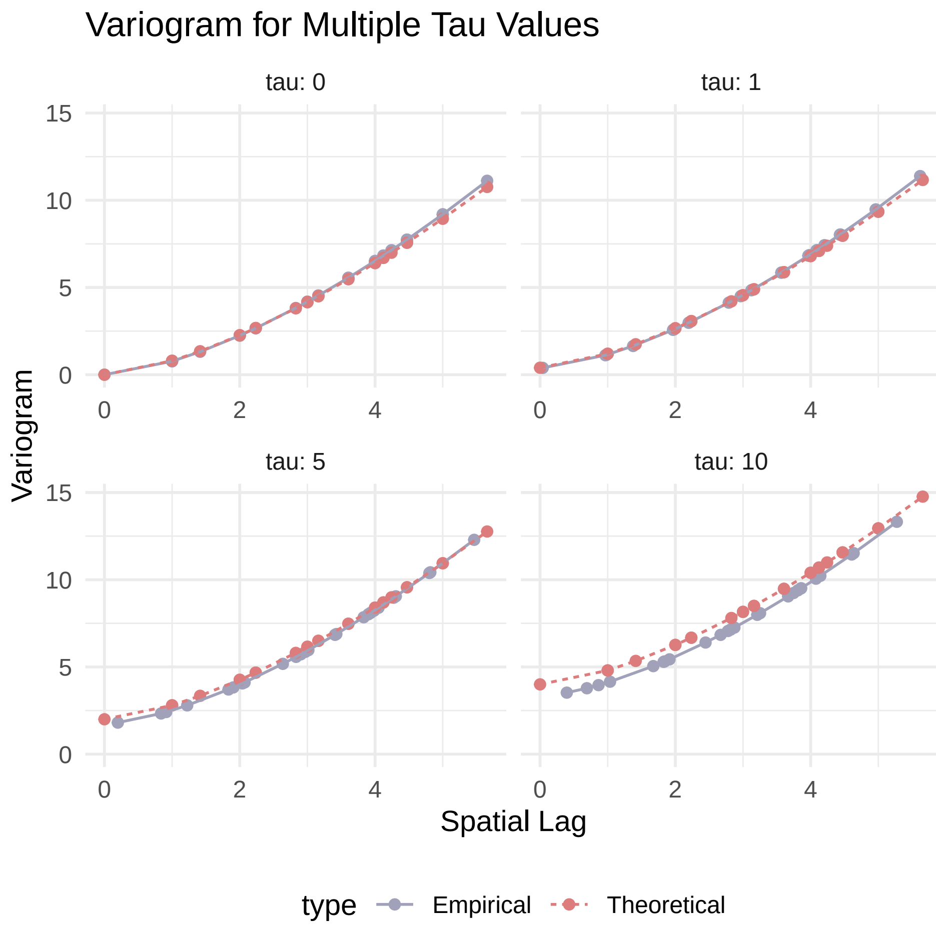

Figure 7: Variogram for the mean estimated parameters for the 50 simulations of 100 replicates each

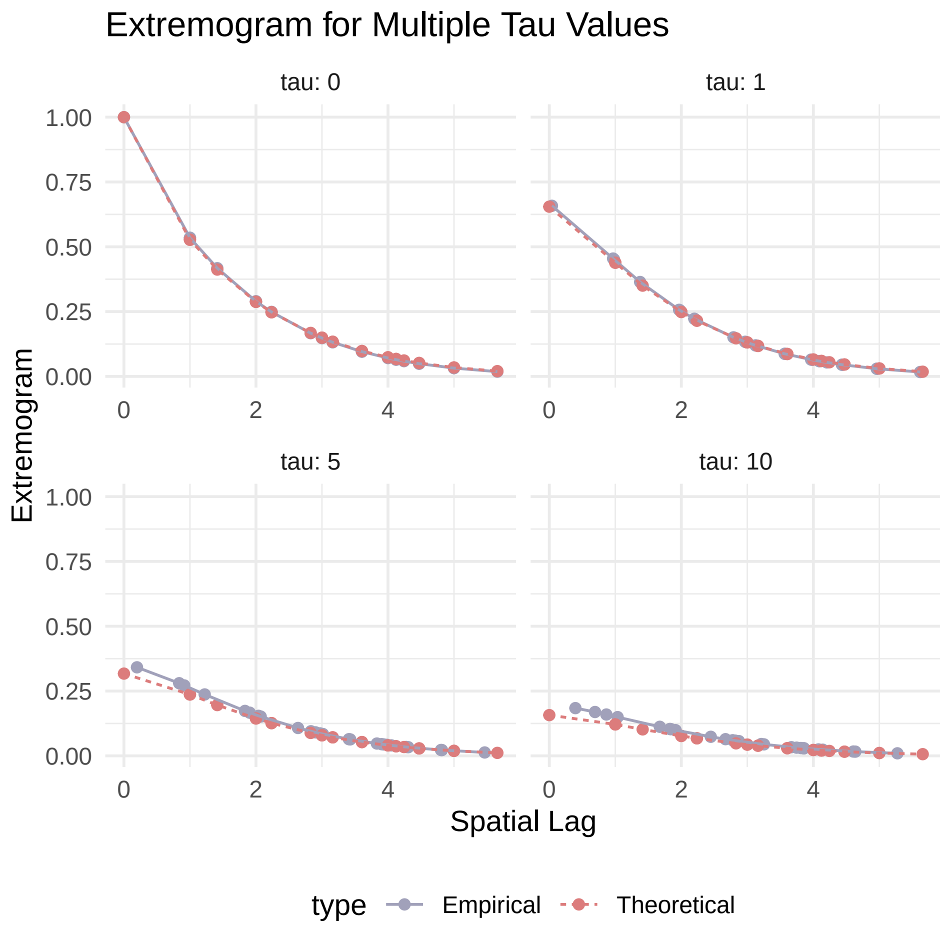

Figure 8: Extremogram for the mean estimated parameters for the 50 simulations of 100 replicates each

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 $t_0 = 1$.

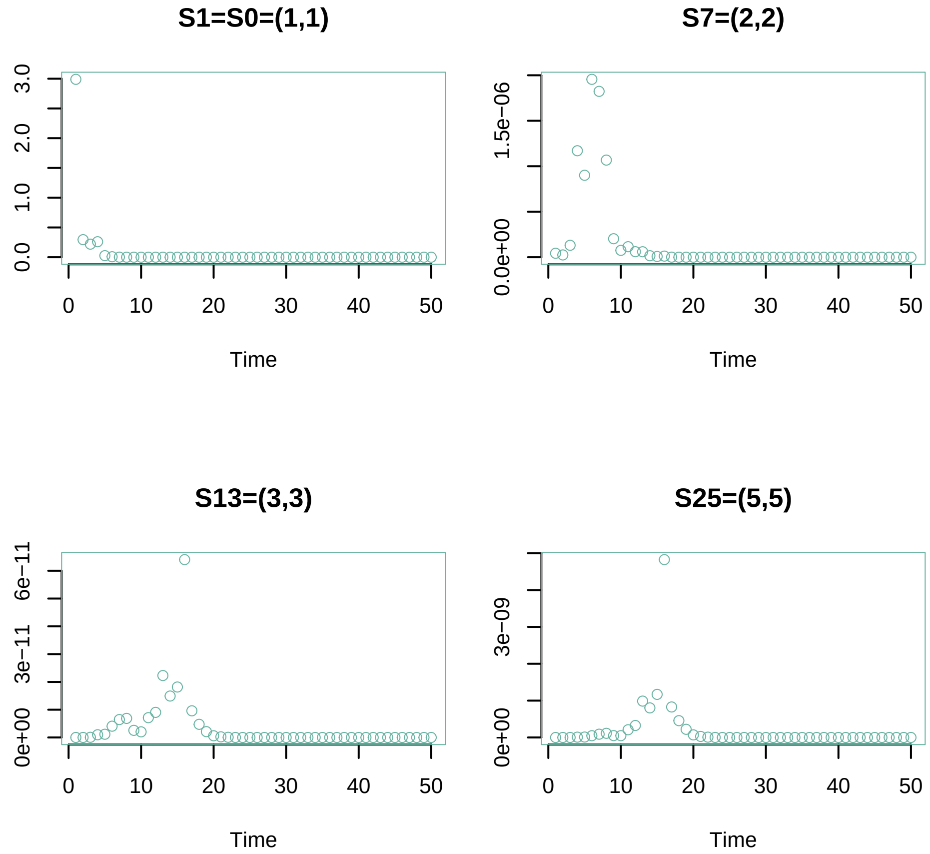

Figure 9: Time series for 4 sites of the first realization in the advection direction

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).

Table 3: Result and RMSE for all replicates together with advection=(0.5,0.3)

X.1	X	beta1	beta2	alpha1	alpha2	adv1	adv2
1	estim	0.4701942	0.0743631	1.2271052	0.8879485	0.5888297	0.3743178
2	rmse	0.0701942	0.1256369	0.2728948	0.1120515	0.0888297	0.0743178

Plot the estimated variogram vs the true theorical variogram.

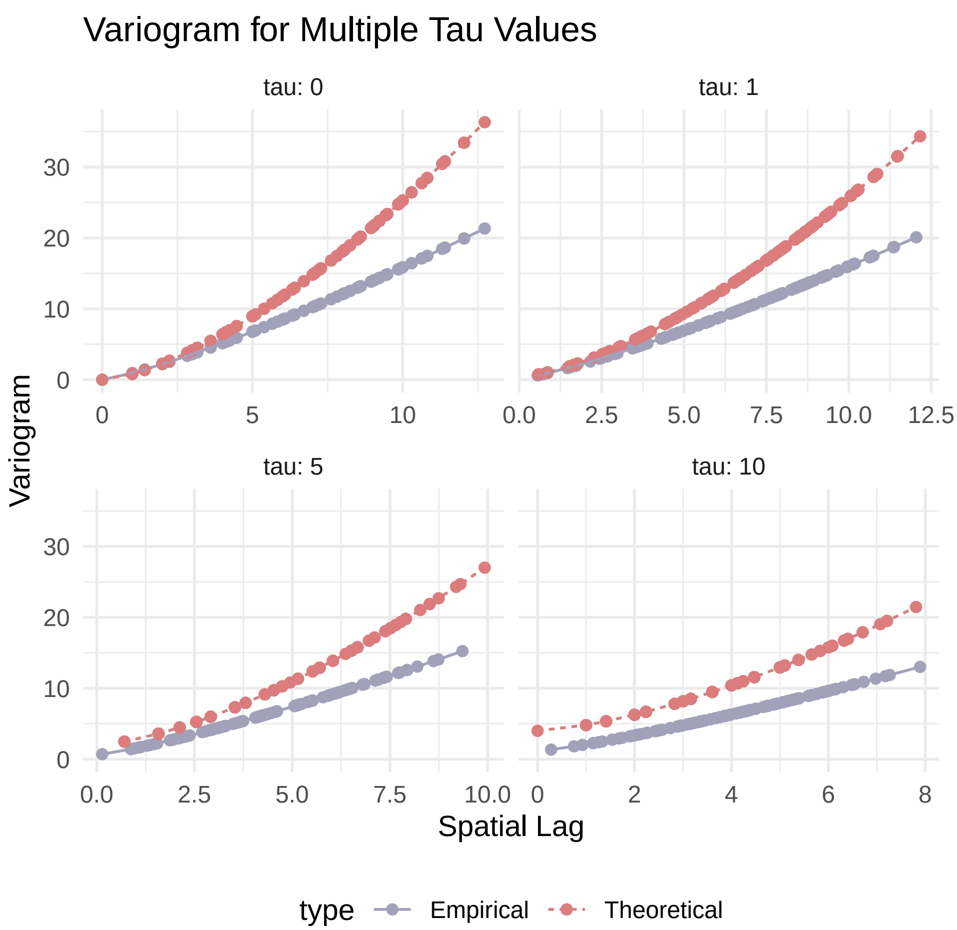

Figure 10: Variogram for the estimated parameters

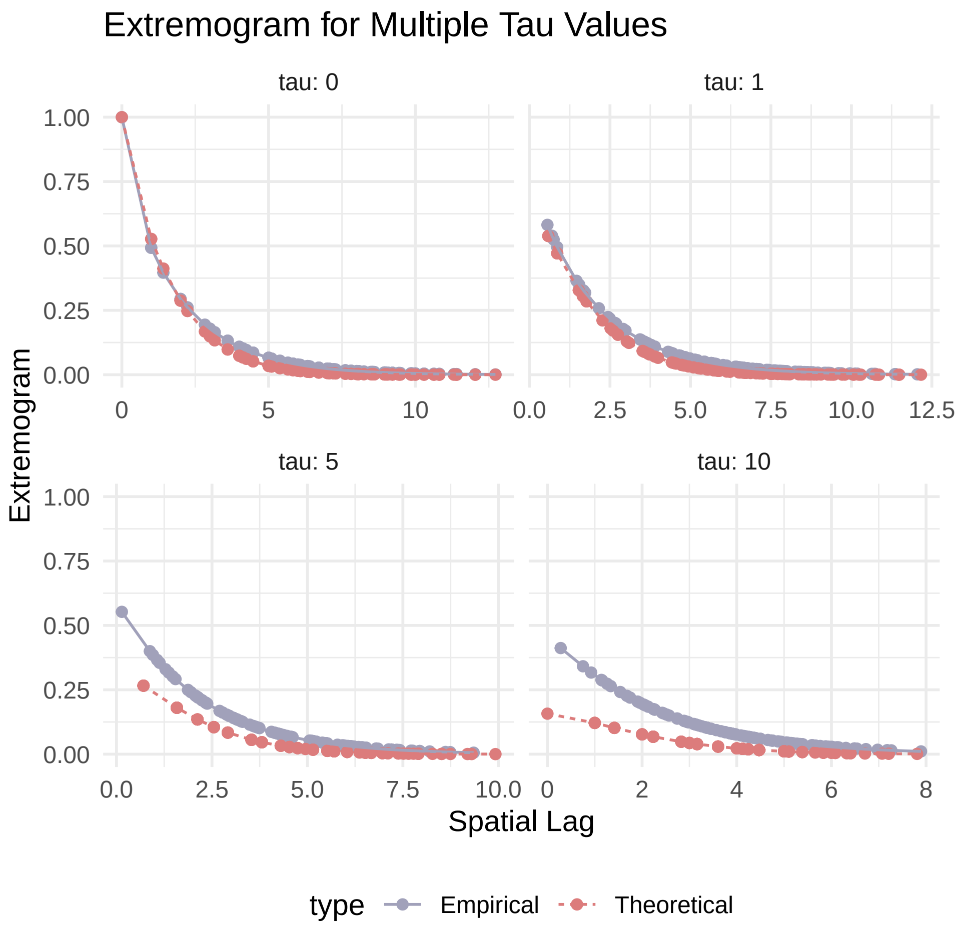

Figure 11: Extremogram for the estimated parameters

For $M = 50$ simulations with $m = 100$ replicates each, we have the following results:

Table 4: Criterion for all simulations together with advection=(0.5,0.3)

X	beta1	beta2	alpha1	alpha2	adv1	adv2
1	0.4701942	0.0743631	1.227105	0.8879485	0.5888297	0.3743178
2	0.1852269	0.2926831	1.848575	0.4968551	0.5892619	0.3948056
3	0.2912299	0.3052413	1.414356	0.3386320	0.6199299	0.4360714
4	0.2818271	0.5292684	1.692426	0.2617114	0.5887619	0.3634654
5	0.2357891	0.6459569	1.707784	0.1507806	0.6247320	0.3768638
6	0.2279789	0.3232258	1.659084	0.4672459	0.6140588	0.3956194
7	0.2709530	0.3558674	1.553302	0.2966033	0.6007885	0.3943267
8	0.3192591	0.4841837	1.516040	0.2168593	0.6594828	0.3565864
9	0.2666533	0.3547496	1.553606	0.1688214	0.6101359	0.3968675
10	0.3302486	0.3286228	1.410353	0.3204853	0.5974691	0.4094110
11	0.1717251	0.3322628	1.999000	0.3084378	0.5999374	0.3614012
12	0.3252657	0.4198885	1.479647	0.3349486	0.6763840	0.3566516
13	0.1923014	0.4395445	1.994682	0.3100067	0.6090555	0.3545787
14	0.3844048	0.2782952	1.469004	0.5186611	0.6443674	0.3092158
15	0.2433071	0.4665873	1.623385	0.3582028	0.6656058	0.3800642
16	0.2326921	0.2491094	1.631257	0.5252802	0.6189961	0.3719391
17	0.2713734	0.4994697	1.656910	0.2024826	0.5982931	0.4061981
18	0.4500202	0.1065181	1.258560	0.8558439	0.5880627	0.3573642
19	0.1919163	0.4176922	1.814177	0.0703362	0.6100217	0.3744885
20	0.3753814	0.2230166	1.606734	0.6419615	0.5614813	0.3531595
21	0.4089541	0.3054141	1.370406	0.4177205	0.5653139	0.3744095
22	0.3845732	0.3258613	1.468368	0.4796903	0.6013654	0.3783558
23	0.2390112	0.3310346	1.702254	0.3569962	0.6246198	0.3434380
24	0.2129293	0.4499330	1.801005	0.3200933	0.6032899	0.4241674
25	0.5222824	0.1046243	1.172084	1.0180339	0.5608246	0.3922067
26	0.2659028	0.1118464	1.582399	0.9096498	0.6119623	0.4001188
27	0.3701741	0.4108020	1.387099	0.1893471	0.6466748	0.3759947
28	0.4295174	0.2673939	1.275485	0.4150817	0.6289350	0.3745415
29	0.3443805	0.3254002	1.537961	0.5192858	0.5914948	0.3654938
30	0.2170074	0.4525599	1.718901	0.2132053	0.6451599	0.3879724
31	0.4668649	0.2902245	1.308953	0.5056728	0.6341103	0.3830401
32	0.3593518	0.4468944	1.516342	0.3226480	0.5858087	0.4112183
33	0.2636468	0.4348231	1.601368	0.1231507	0.6325751	0.3959167
34	0.1948892	0.3286029	1.972314	0.3494113	0.5859722	0.3901956
35	0.3785312	0.1508004	1.310452	0.8289942	0.5825460	0.4571330
36	0.3315365	0.1941845	1.533726	0.5456114	0.5846252	0.3629330
37	0.5542052	0.2843644	1.150229	0.3920096	0.6229395	0.3959007
38	0.3844880	0.4240237	1.294169	0.0599317	0.6249619	0.4014876
39	0.2026465	0.5409250	1.771636	0.2632778	0.6355527	0.3706261
40	0.2182205	0.3530851	1.698866	0.3963078	0.5859084	0.3953993
41	0.6929593	0.3715051	0.974486	0.1607151	0.6129945	0.3653339
42	0.2081209	0.3815551	1.769909	0.3899288	0.6369371	0.3699439
43	0.2533809	0.5448967	1.772084	0.3215167	0.6039127	0.3974596
44	0.4951736	0.4784777	1.340373	0.3231508	0.6669492	0.3552347
45	0.2168851	0.3653940	1.652908	0.2656720	0.6376136	0.3944501
46	0.2594478	0.3845065	1.648828	0.3688261	0.6202433	0.3787198
47	0.4198156	0.5704889	1.406643	0.1706529	0.5953578	0.3783525
48	0.1986329	0.2995115	1.817112	0.4730907	0.5913795	0.3846523
49	0.2144575	0.4671443	1.861907	0.1475304	0.5756778	0.4142869
50	0.2885053	0.3182926	1.654623	0.4540367	0.6094200	0.3779523

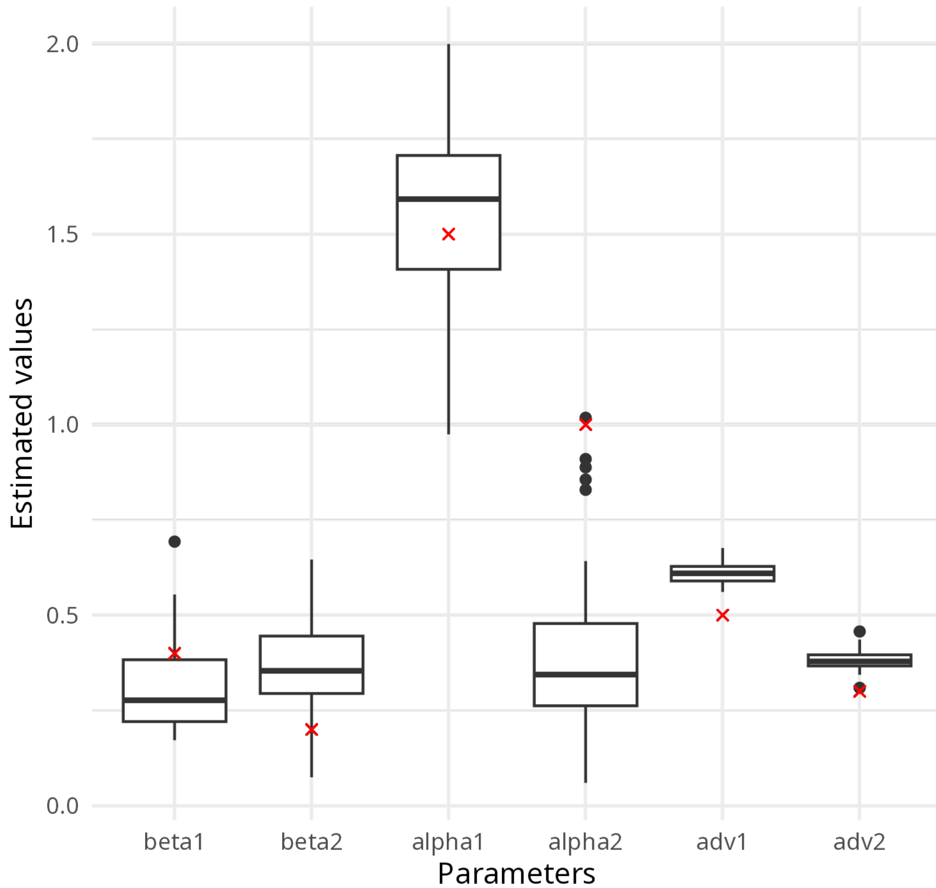

Figure 12: Boxplot of the estimated parameters for the 50 simulations of 100 replicates each

For one simulation:

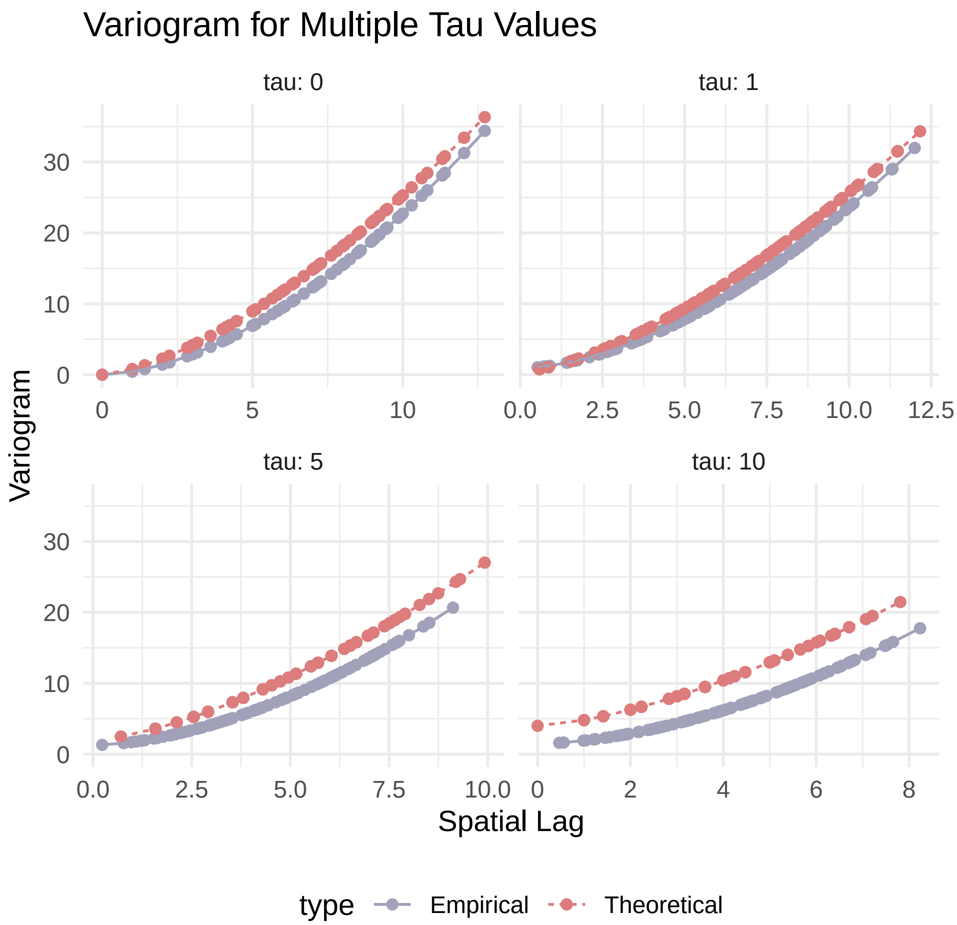

Figure 13: Variogram for the estimated parameters for the 30th simulation of 100 replicates each

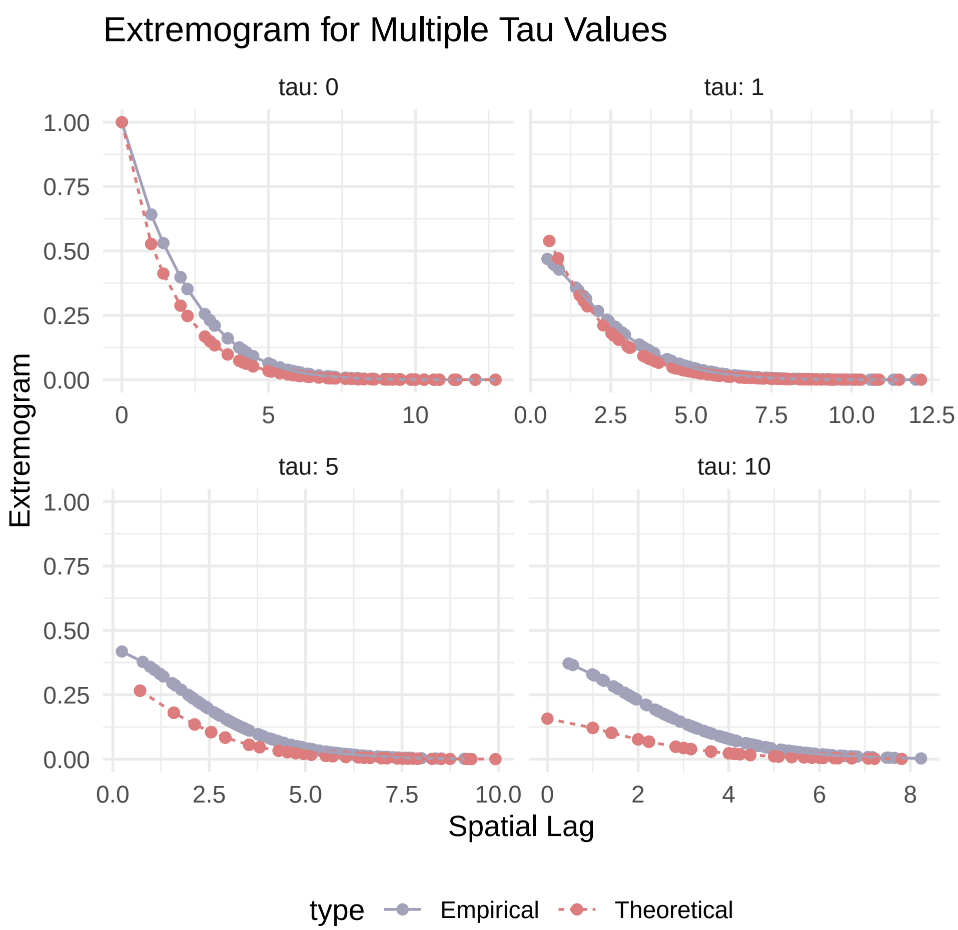

Figure 14: Extremogram for the estimated parameters for the 30th simulation of 100 replicates each

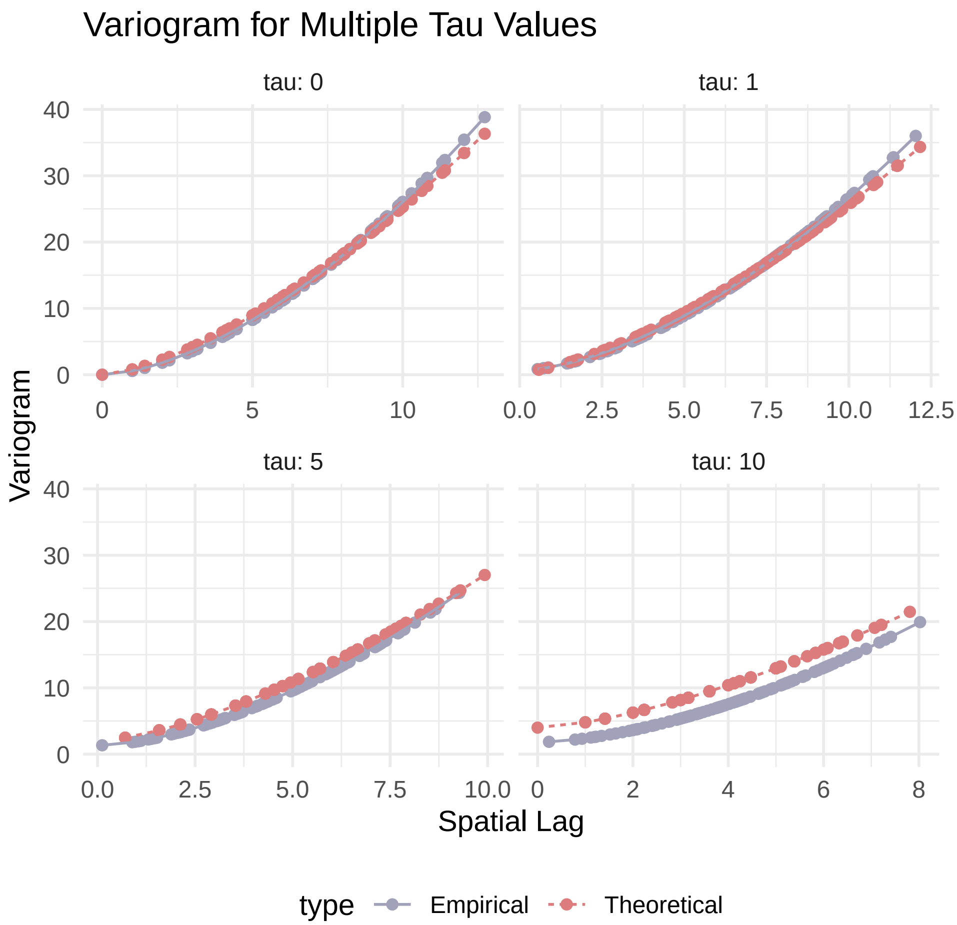

Figure 15: Variogram for the estimated parameters for the 50th simulation of 100 replicates each

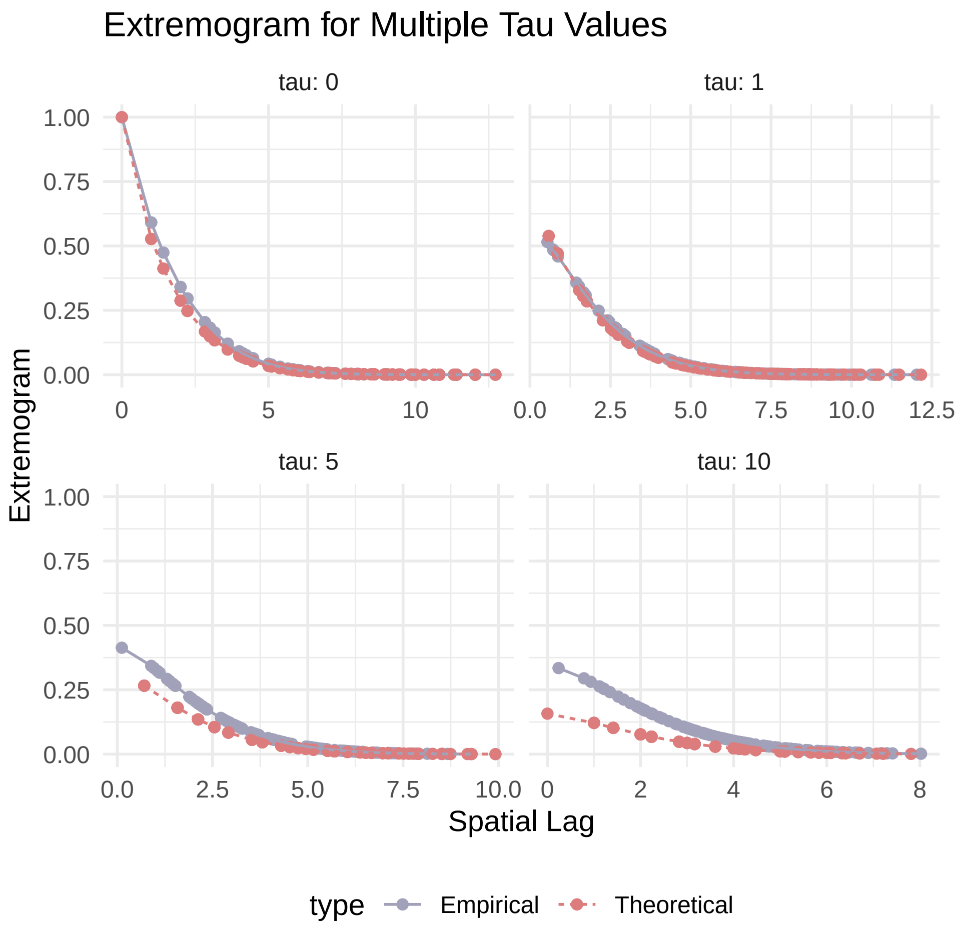

Figure 16: Extremogram for the estimated parameters for the 50th simulation of 100 replicates each

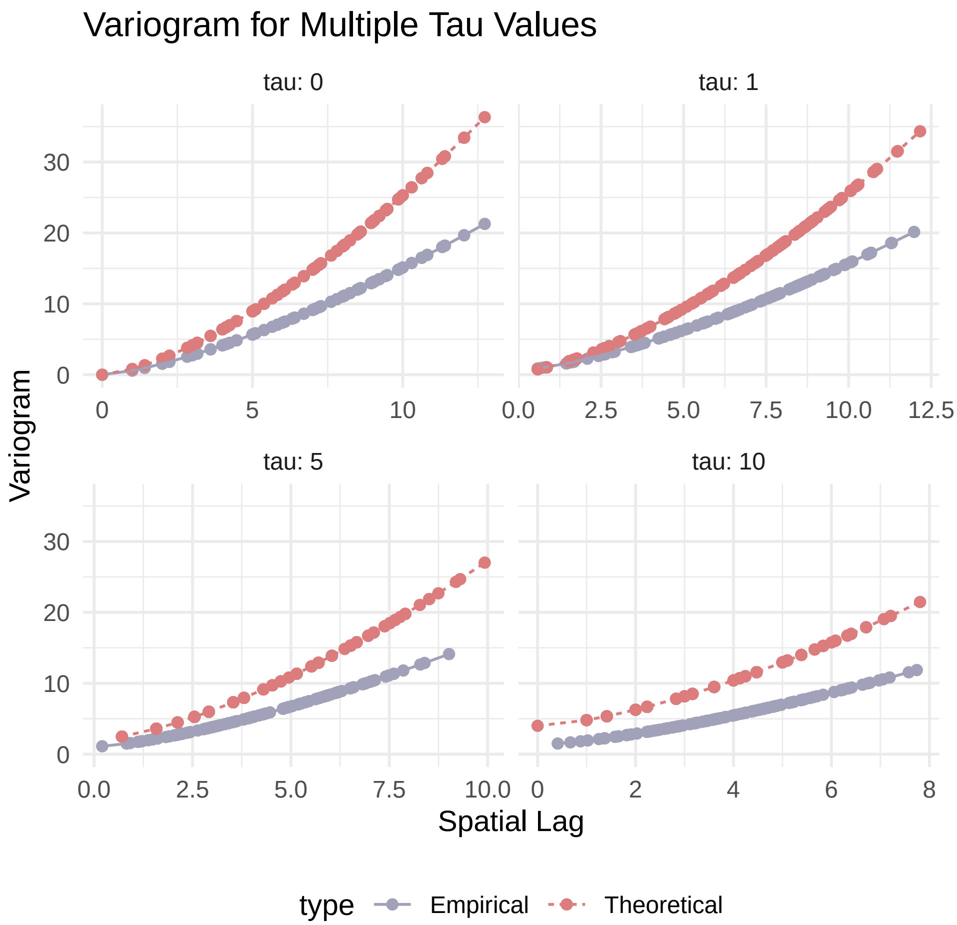

Figure 17: Variogram for the estimated parameters for the third simulation of 100 replicates each

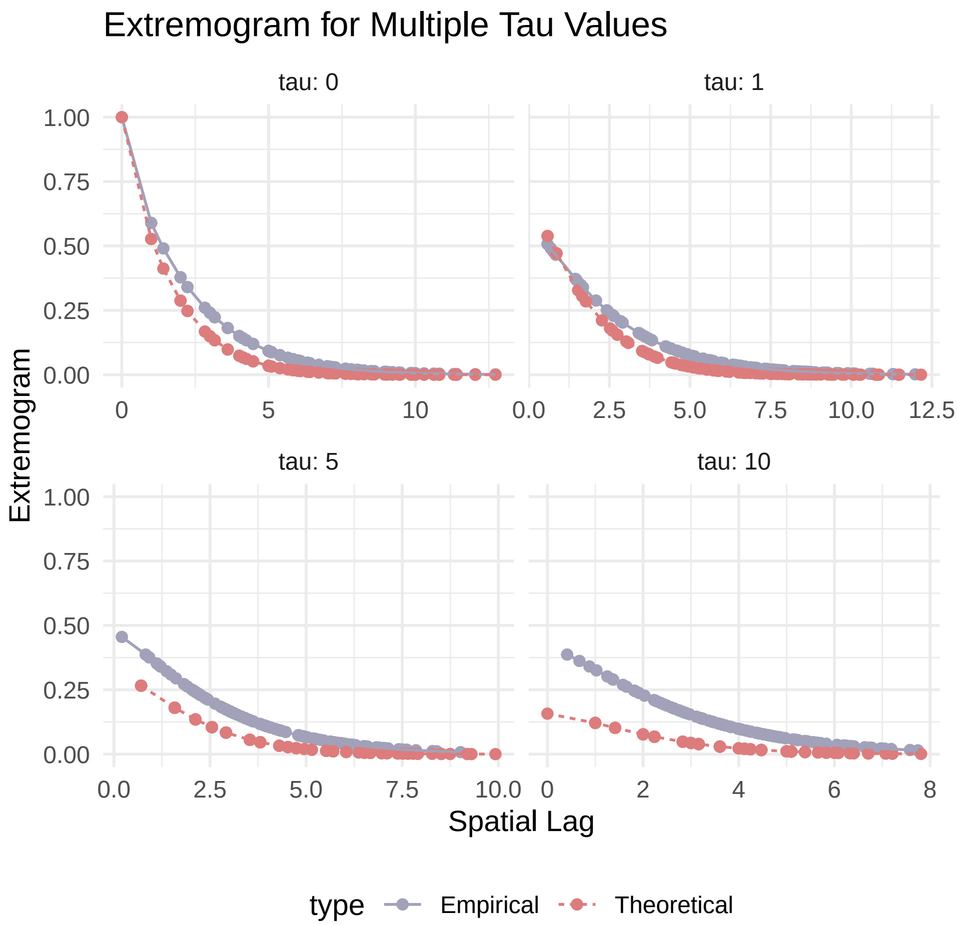

Figure 18: Extremogram for the estimated parameters for the third simulation of 100 replicates each

With mean estimation for the $M = 50$ simulations with $m = 100$ replicates each.

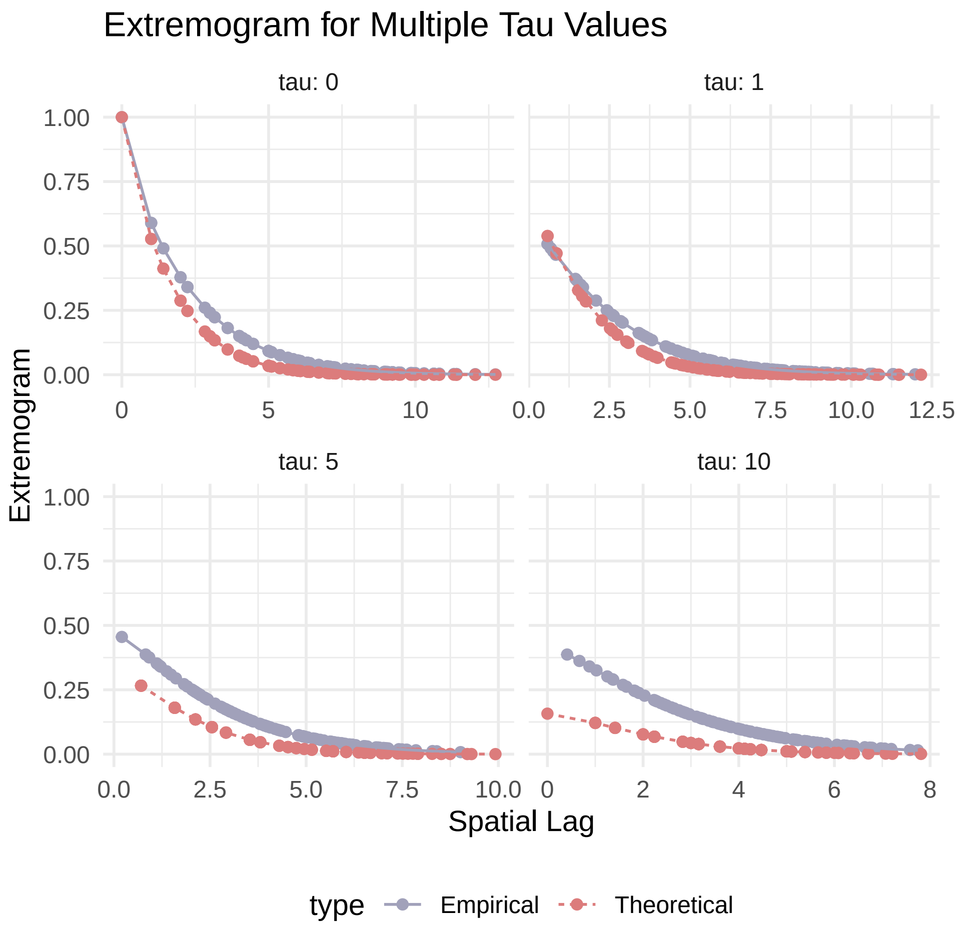

Figure 19: Extremogram for the mean estimated parameters for the 50 simulations of 100 replicates each

Plot by fixing advection parameters

With another advection with $V_x = 0$

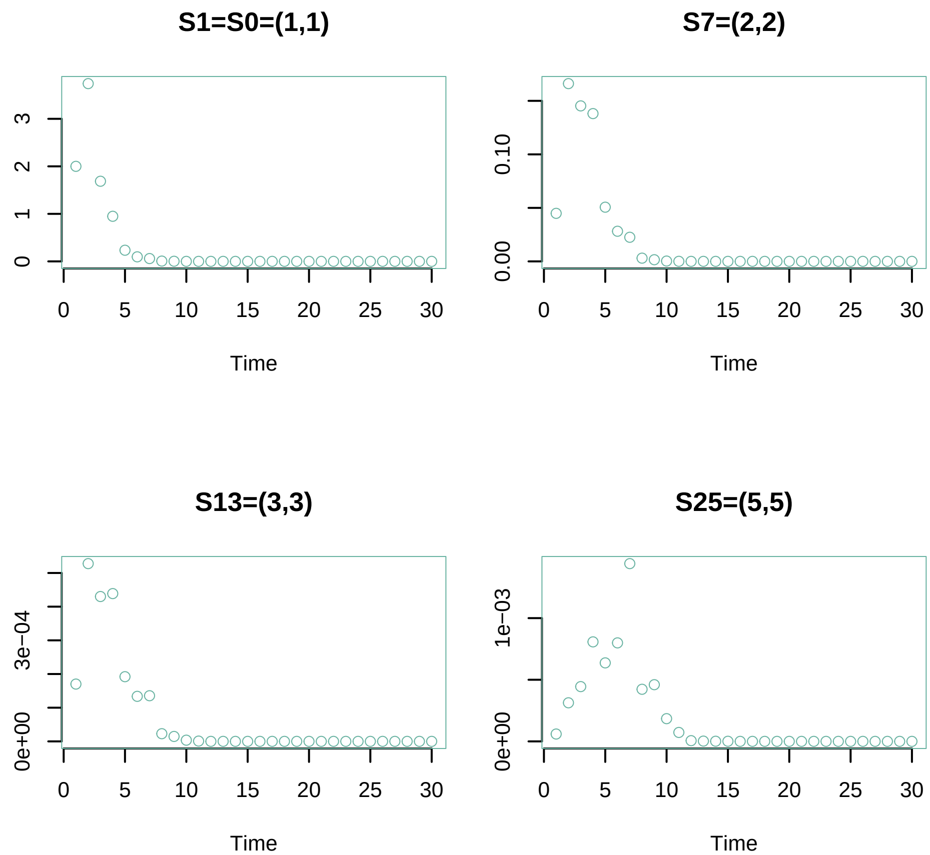

Figure 20: Time series for 4 sites of the first realization in the advection direction

Table 5: Criterion for all simulations together with advection=(0,0.5)

	mean	rmse	mae
beta1	0.2966012	0.1297489	0.1130374
beta2	0.3551559	0.1714348	0.1551559
alpha1	1.7291264	0.2821017	0.2437736
alpha2	0.4379397	0.5796602	0.5620603
adv1	0.0647144	0.0689657	0.0647144
adv2	0.6287663	0.1307650	0.1287663

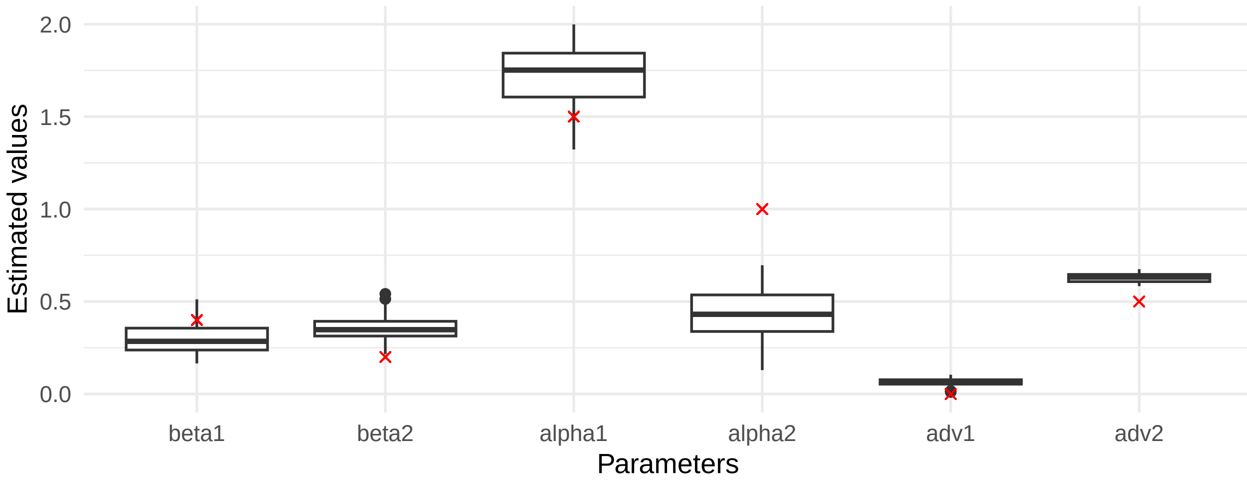

Figure 21: Boxplot of the estimated parameters for the 50 simulations of 100 replicates each

For one simulation:

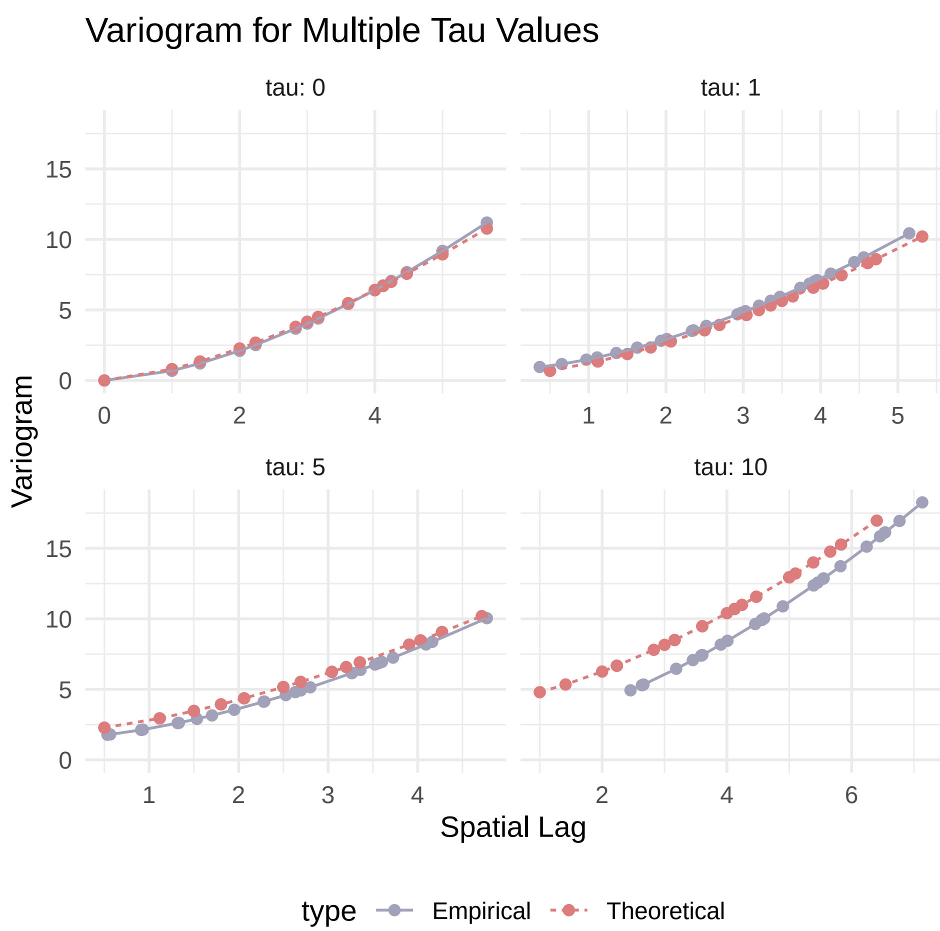

Figure 22: Variogram for the estimated parameters for the 30th simulation of 100 replicates each

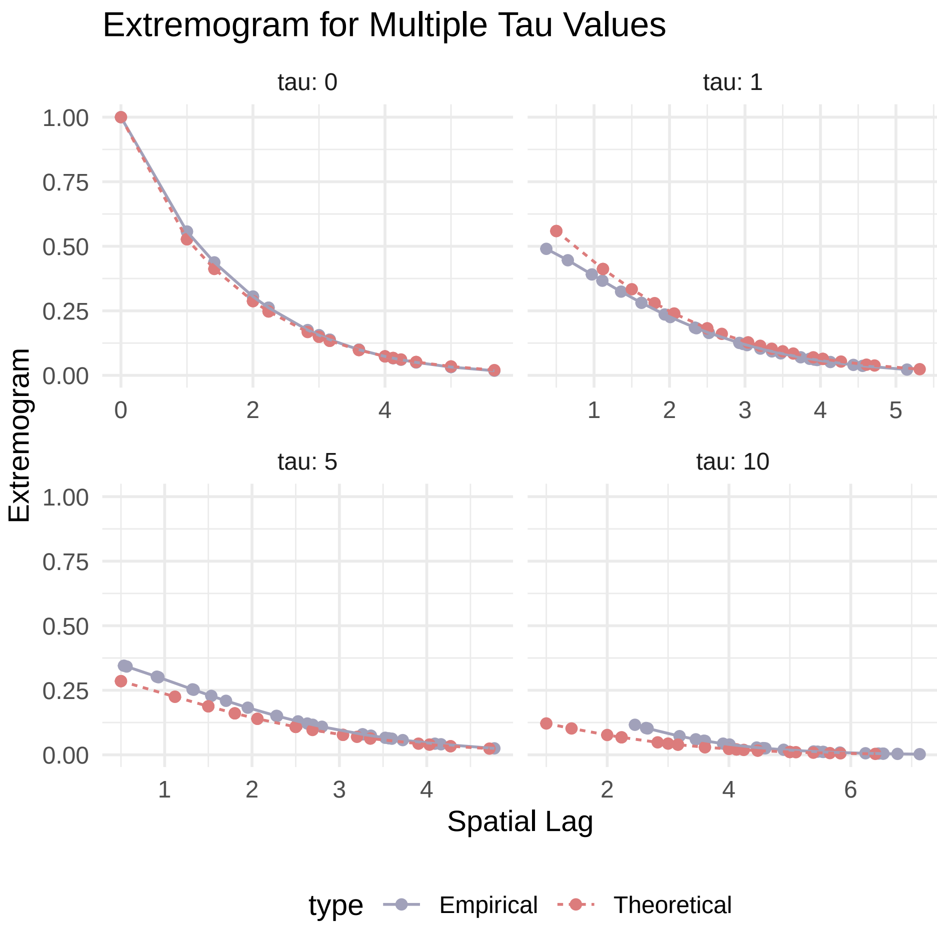

Figure 23: Extremogram for the estimated parameters for the 30th simulation of 100 replicates each

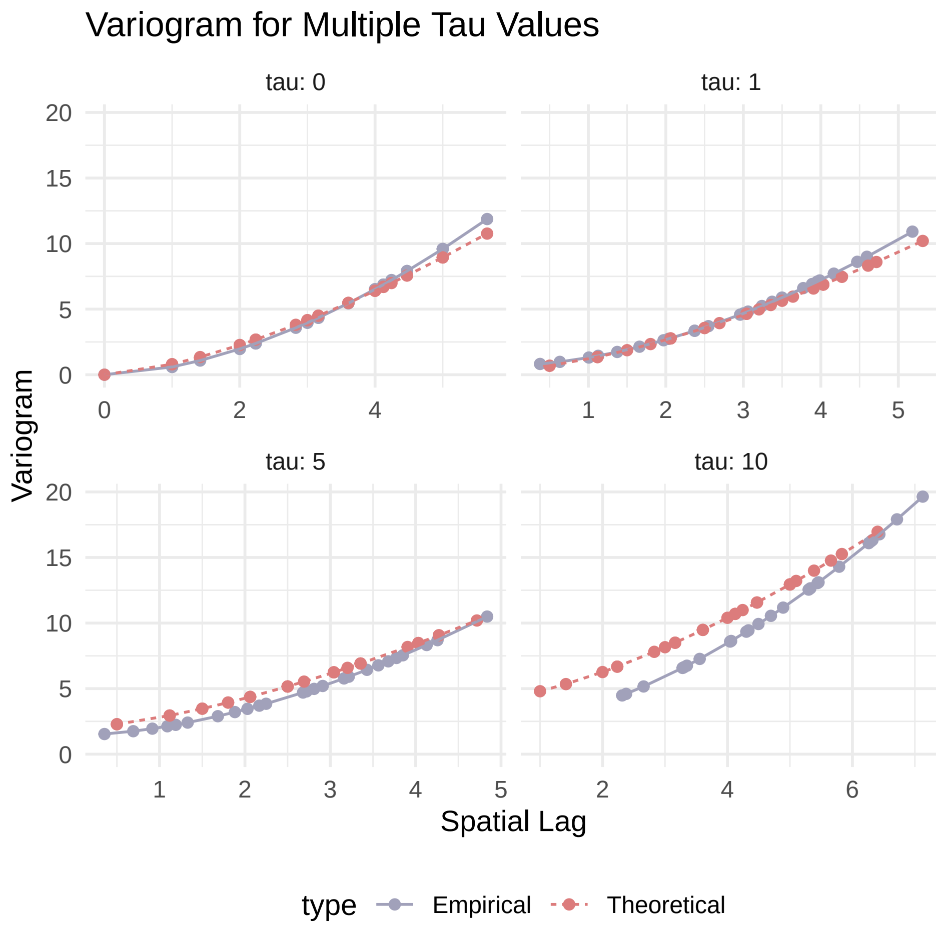

Figure 24: Variogram for the mean estimated parameters for the 50 simulations of 100 replicates each

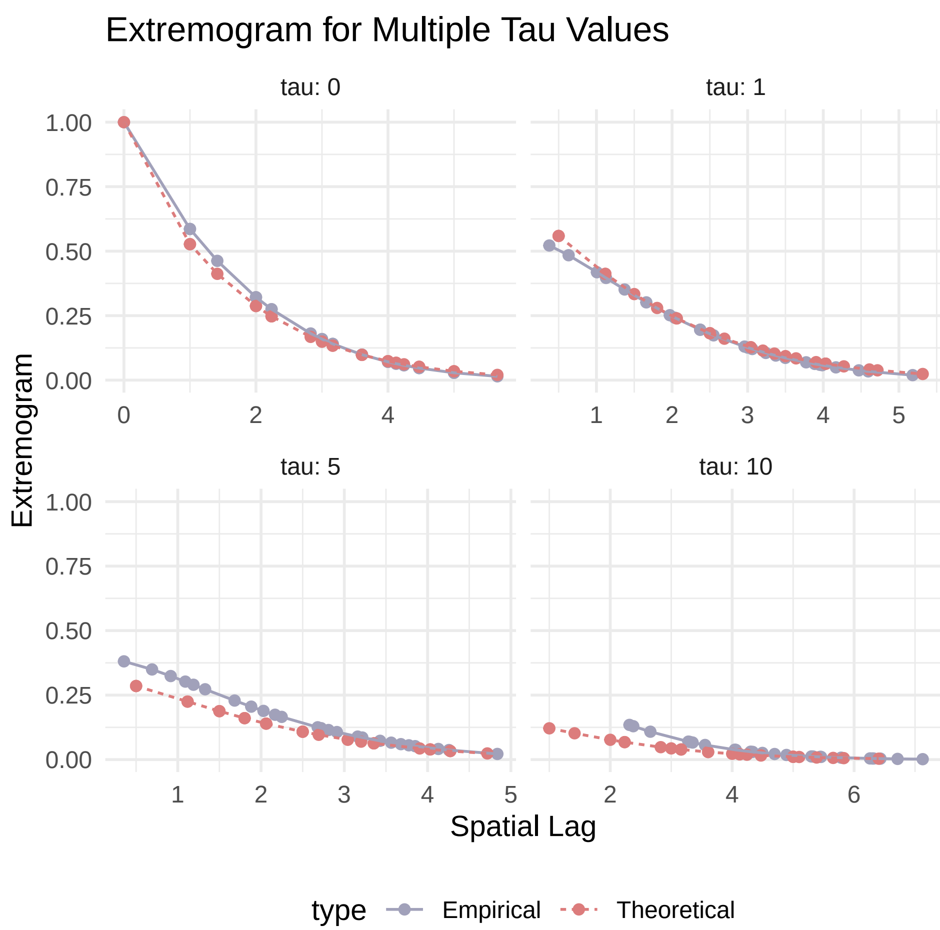

Figure 25: Extremogram for the mean estimated parameters for the 50 simulations of 100 replicates each