Optimisation

2024-10-31

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 $t_0 = 1$.

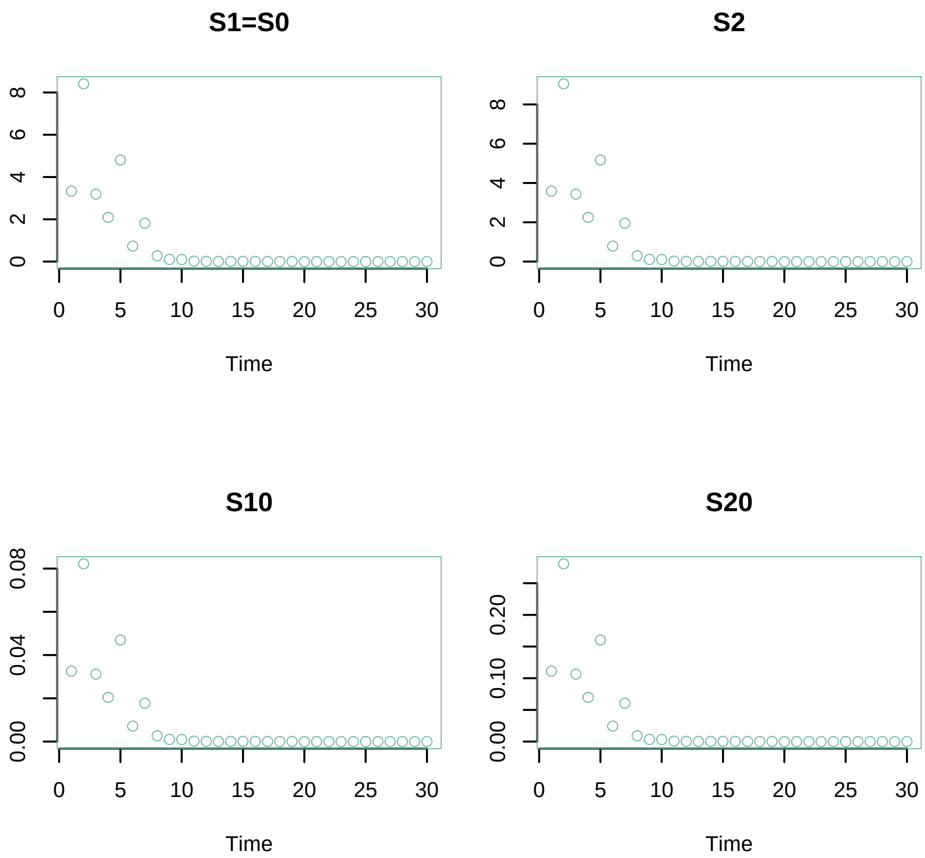

Figure 1: Time series for 4 sites of the first realization

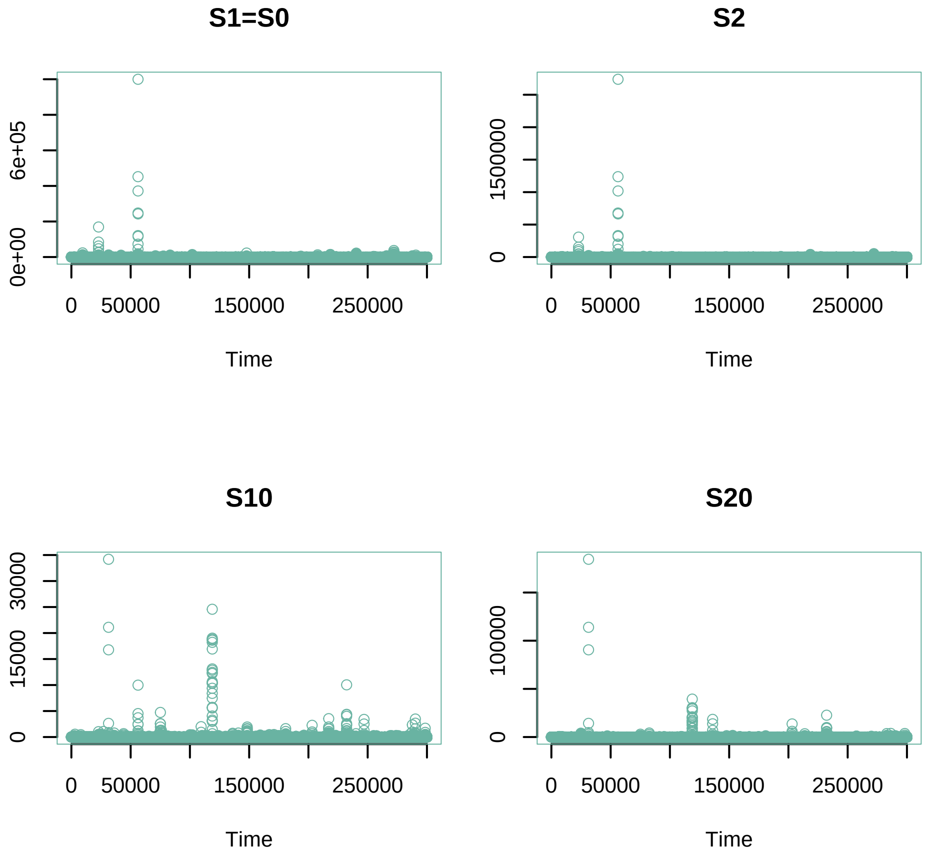

Figure 2: Time series for 4 sites of the all replicates together

Estimation of the variogram parameters

Table 1: Estimation and RMSE for 1 simulation of 10000replicates with advection = (0,0)

	beta1	beta2	alpha1	alpha2	adv1	adv2
estim	0.4225438	0.1902072	1.4108126	1.0270925	0.0110135	0.0049317
rmse	0.0225438	0.0097928	0.0891874	0.0270925	0.0110135	0.0049317

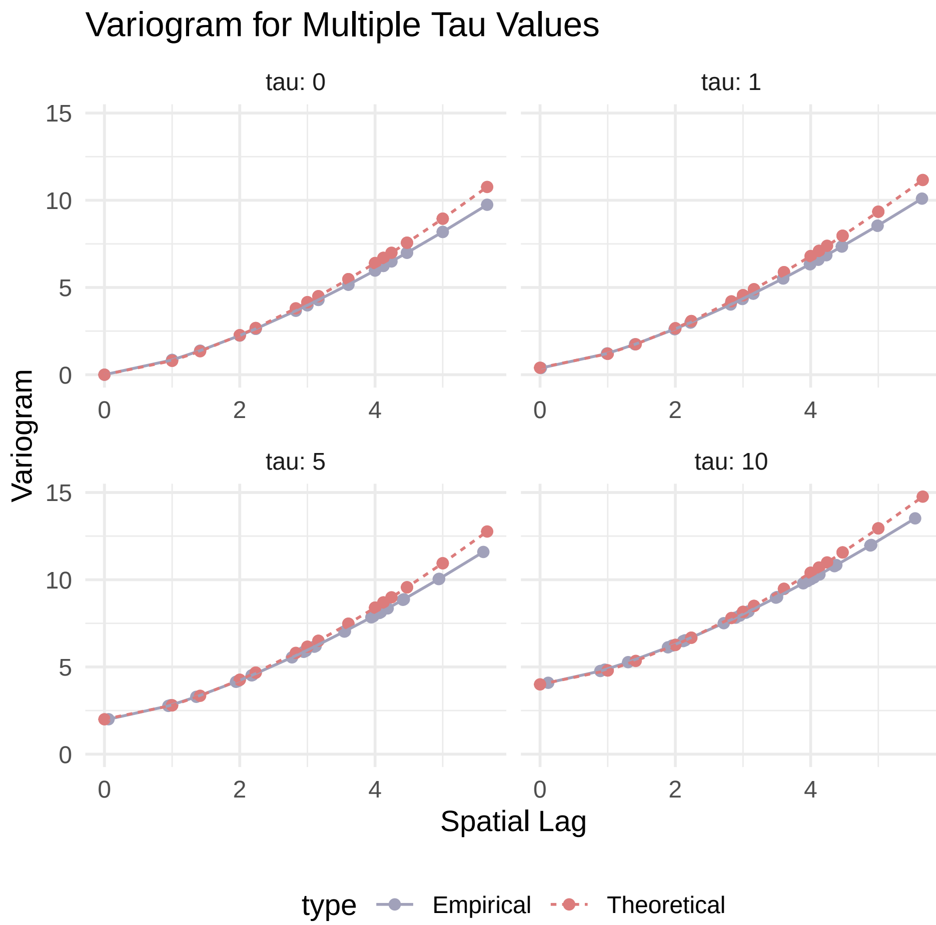

Figure 3: Variogram of the estimated parameters for 1 simulation of 10000 replicates

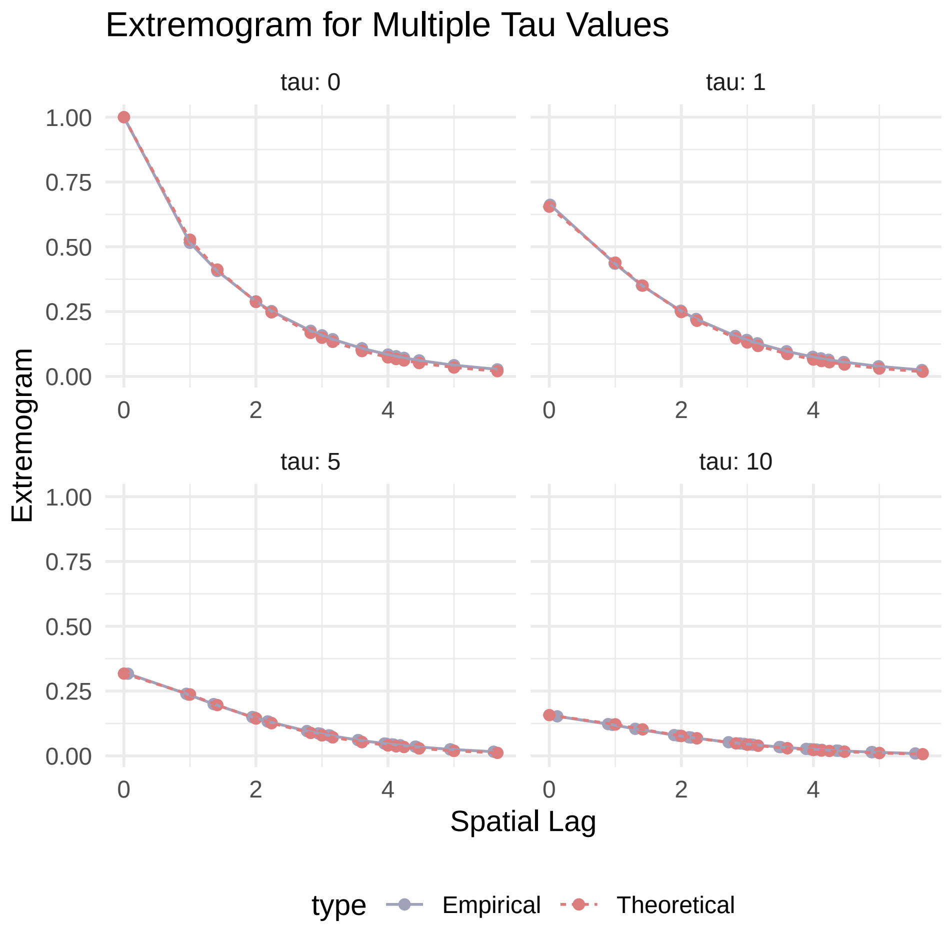

Figure 4: Boxplot of the estimated parameters for 1 simulation of 10000 replicates

For multiple sets of replicates

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

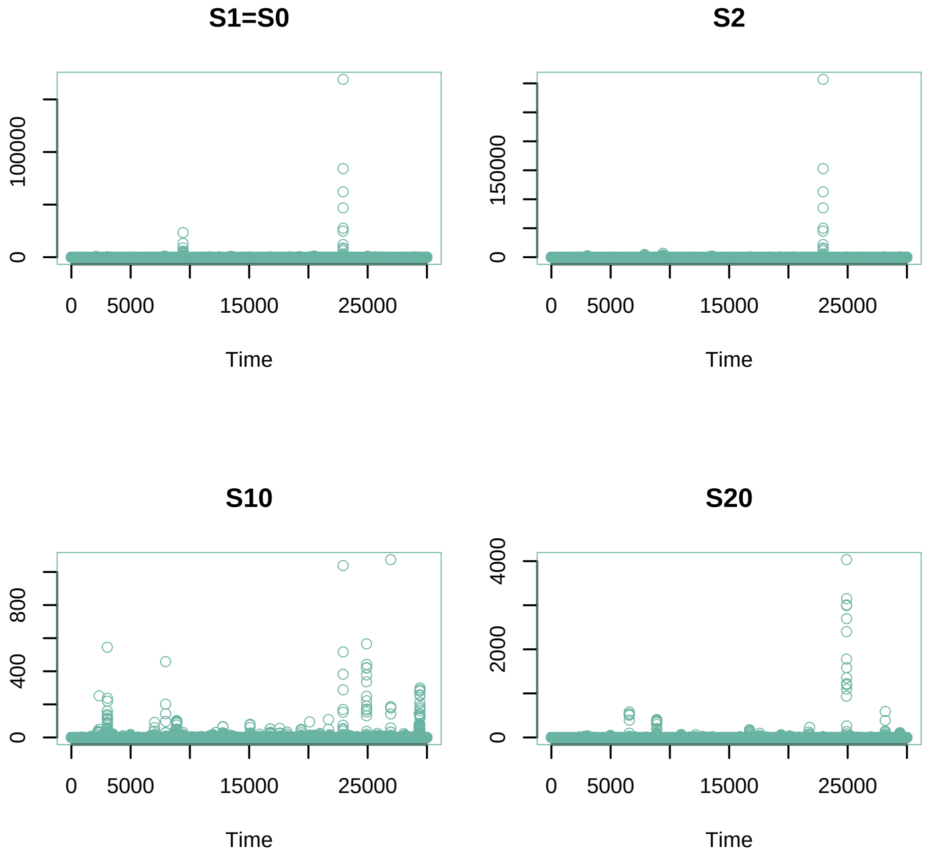

Figure 5: Time series for 4 sites of the first realization

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

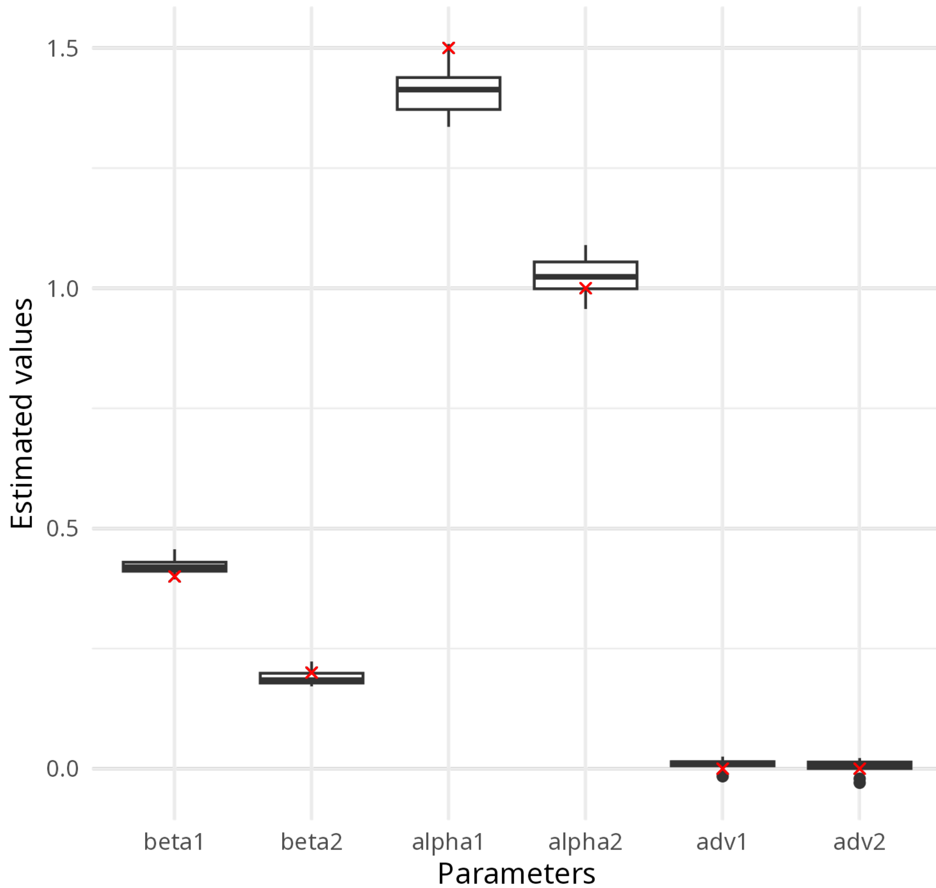

Figure 6: Boxplot of the estimated parameters for the 100 simulations of 100 replicates each

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

	Mean	RMSE	MAE
beta1	0.4212807	0.0276765	0.0224957
beta2	0.1901275	0.0187398	0.0167625
alpha1	1.4142911	0.0996325	0.0873605
alpha2	1.0249641	0.0470608	0.0374813
adv1	0.0097327	0.0142902	0.0129081
adv2	0.0029947	0.0157180	0.0132076

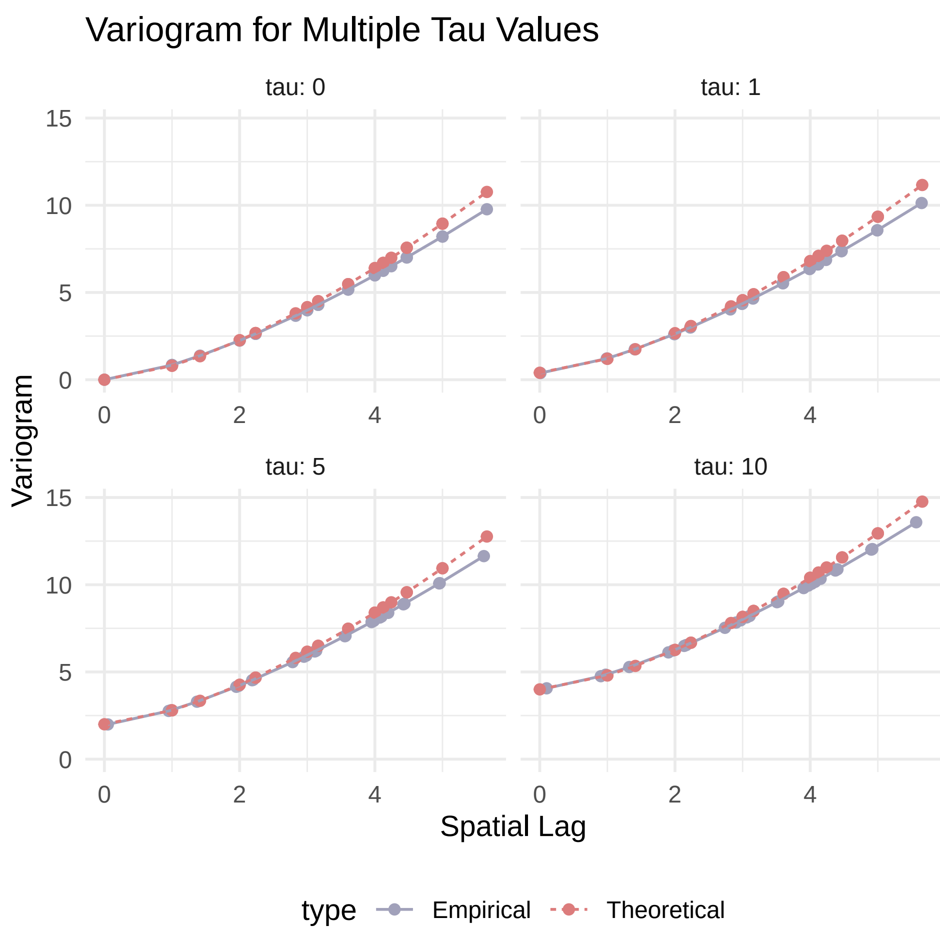

Figure 7: Variogram for the mean estimated parameters for the 50 simulations of 100 replicates each

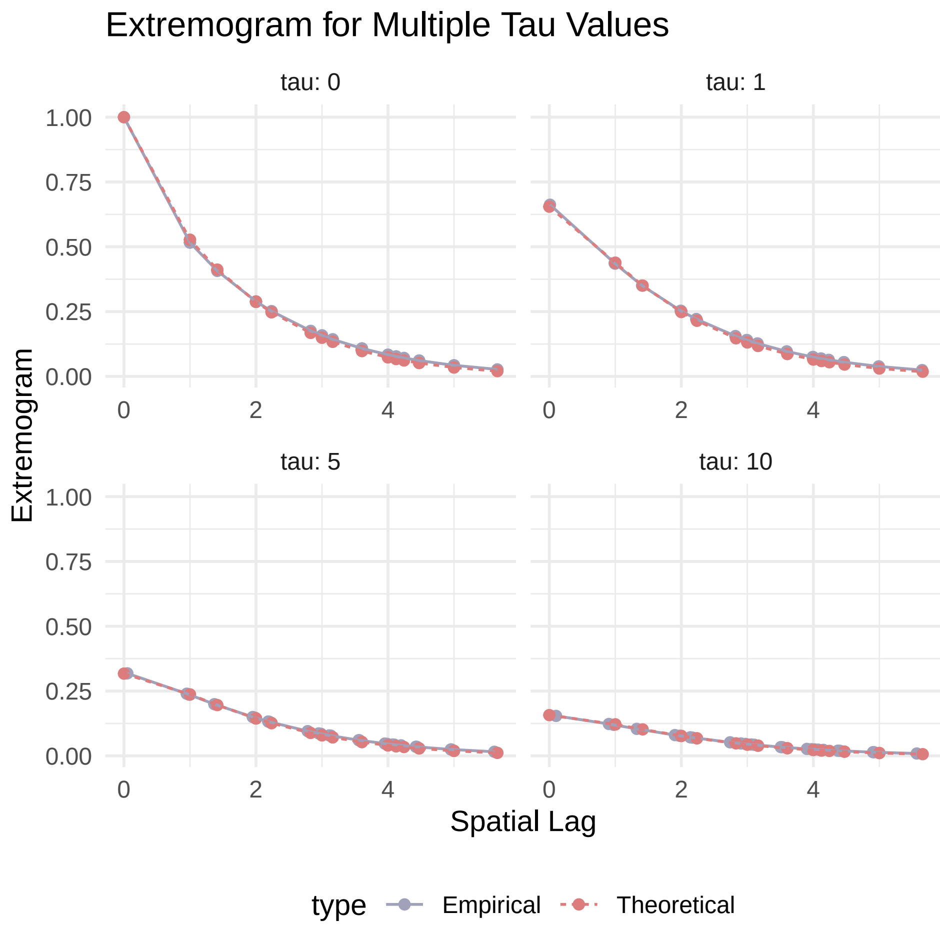

Figure 8: Extremogram for the mean estimated parameters for the 50 simulations of 100 replicates each

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 $t_0 = 1$.

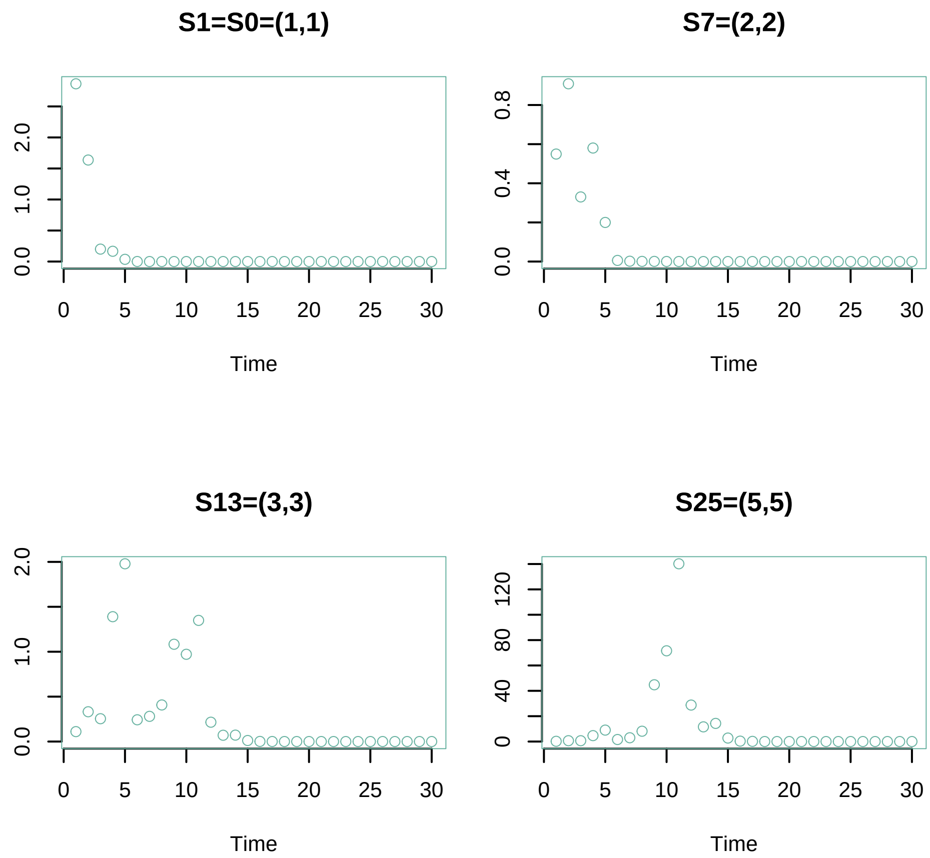

Figure 9: Time series for 4 sites of the first realization in the advection direction

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	beta1	beta2	alpha1	alpha2	adv1	adv2
estim	0.285501	0.3784384	1.726048	0.3751419	0.5891621	0.3659891
rmse	0.114499	0.1784384	0.226048	0.6248581	0.0891621	0.0659891

Plot the estimated variogram vs the true theorical variogram.

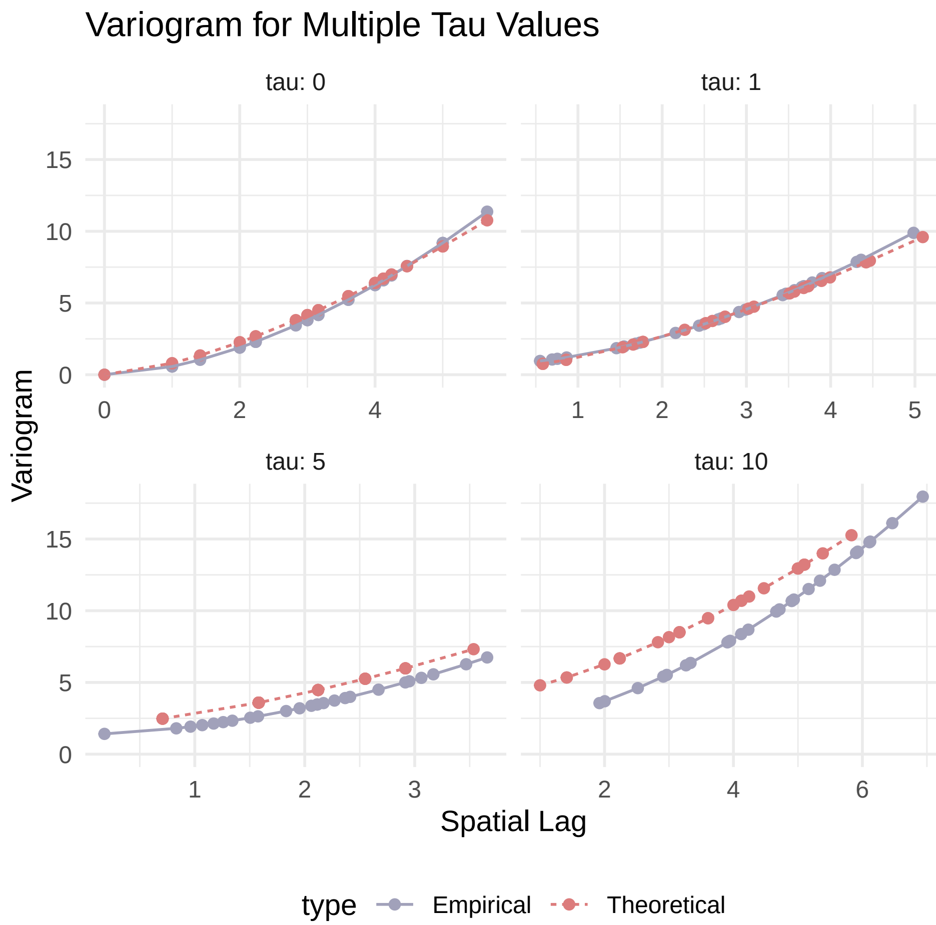

Figure 10: Variogram for the estimated parameters

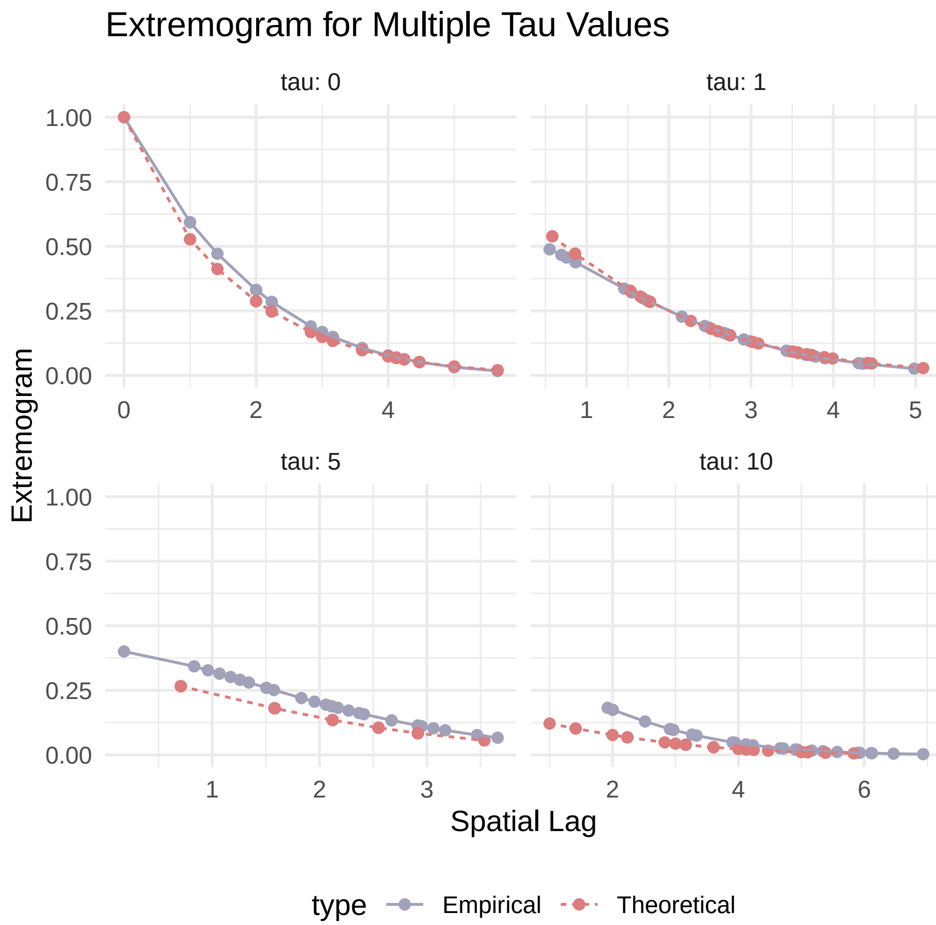

Figure 11: Extremogram for the estimated parameters

For $M = 10$ simulations with $m = 1000$ replicates each, we have the following results:

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

	mean	rmse	mae
beta1	0.2679667	0.1335807	0.1320333
beta2	0.3663510	0.1683415	0.1663510
alpha1	1.7778590	0.2813716	0.2778590
alpha2	0.3740939	0.6270630	0.6259061
adv1	0.5903870	0.0905009	0.0903870
adv2	0.3685796	0.0689144	0.0685796

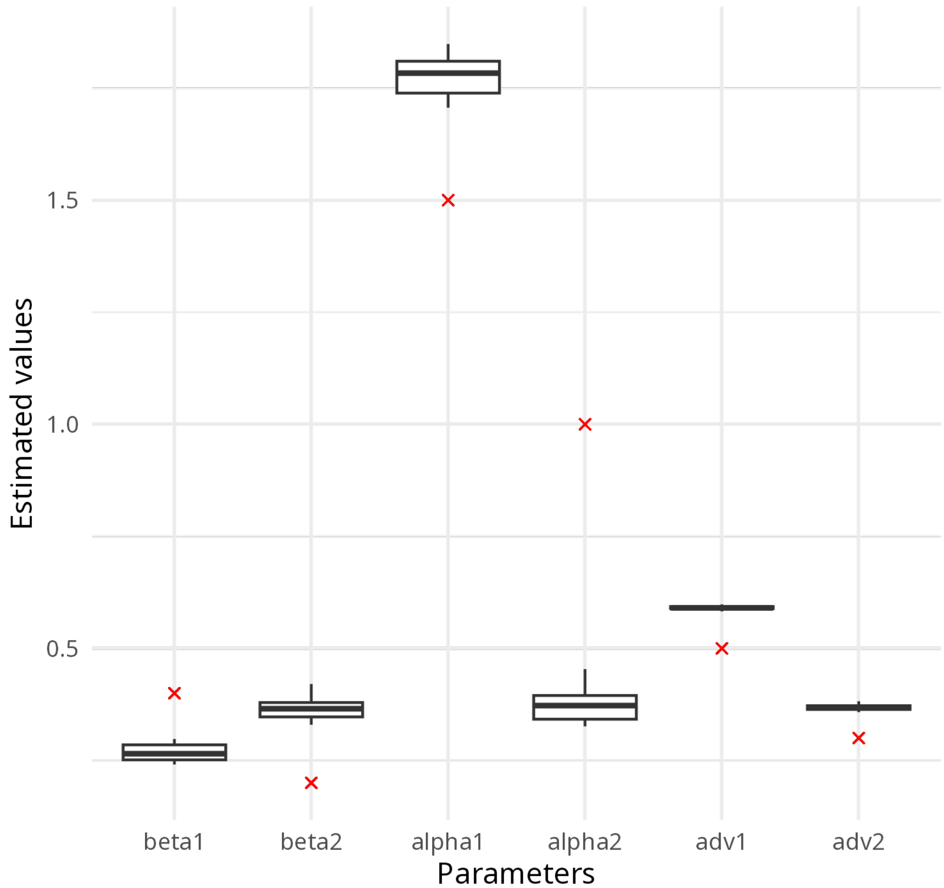

Figure 12: Boxplot of the estimated parameters for the 50 simulations of 100 replicates each

For one simulation:

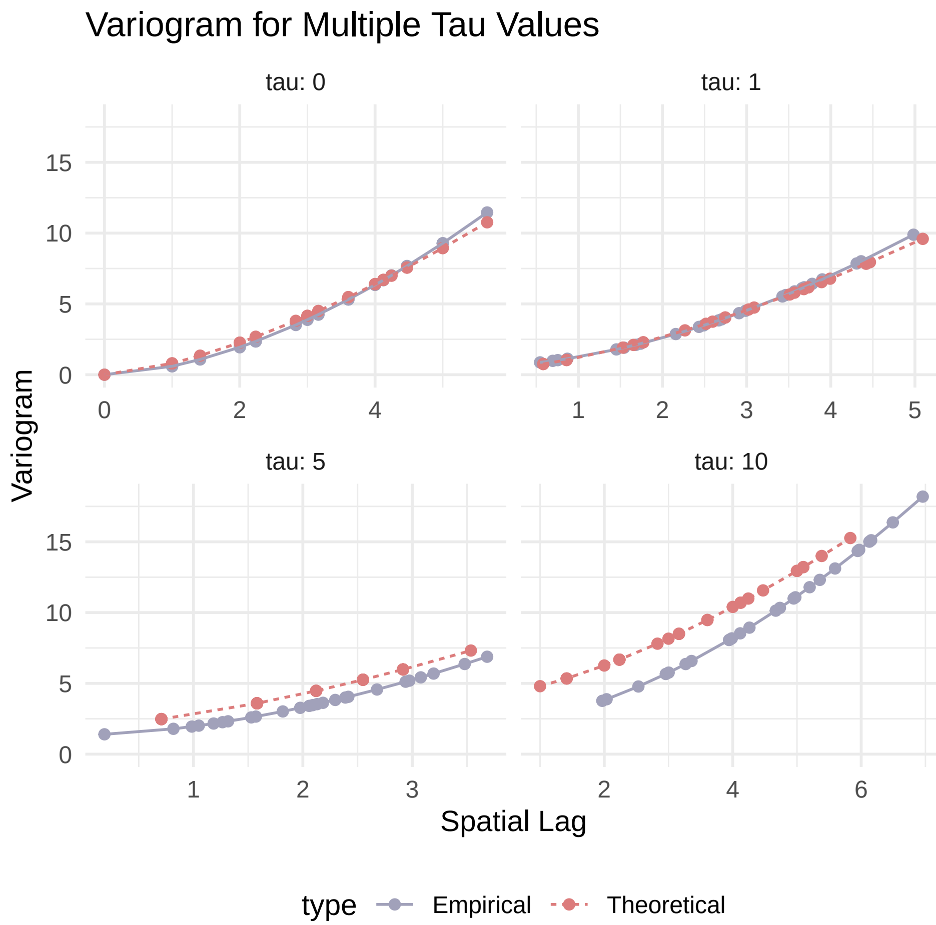

Figure 13: Variogram for the estimated parameters for the 5th simulation of 100 replicates each

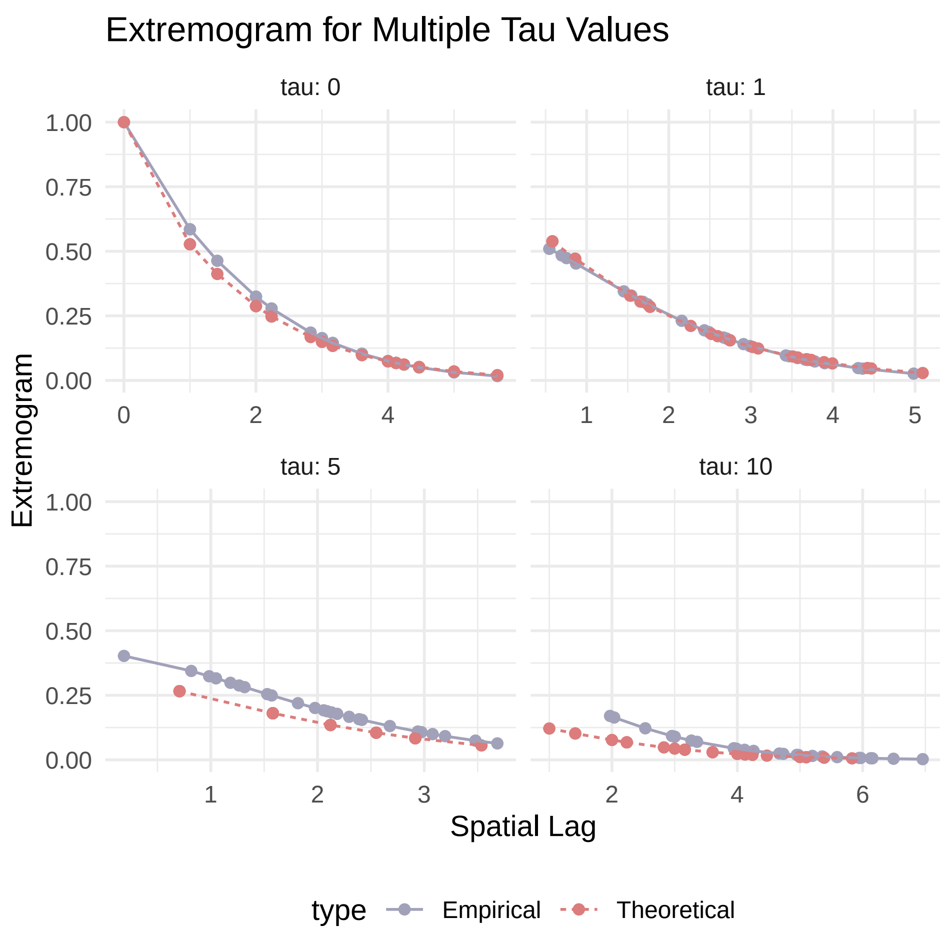

Figure 14: Extremogram for the estimated parameters for the 30th simulation of 100 replicates each

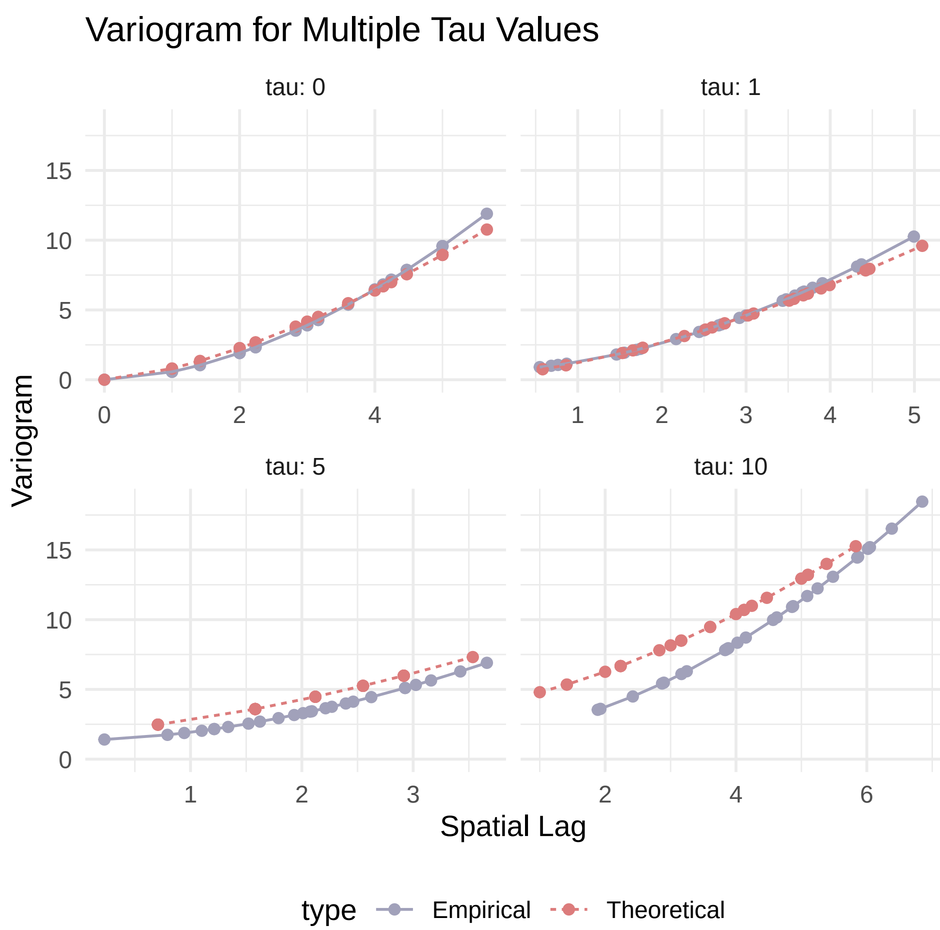

Figure 15: Variogram for the estimated parameters for the 10th simulation of 100 replicates each

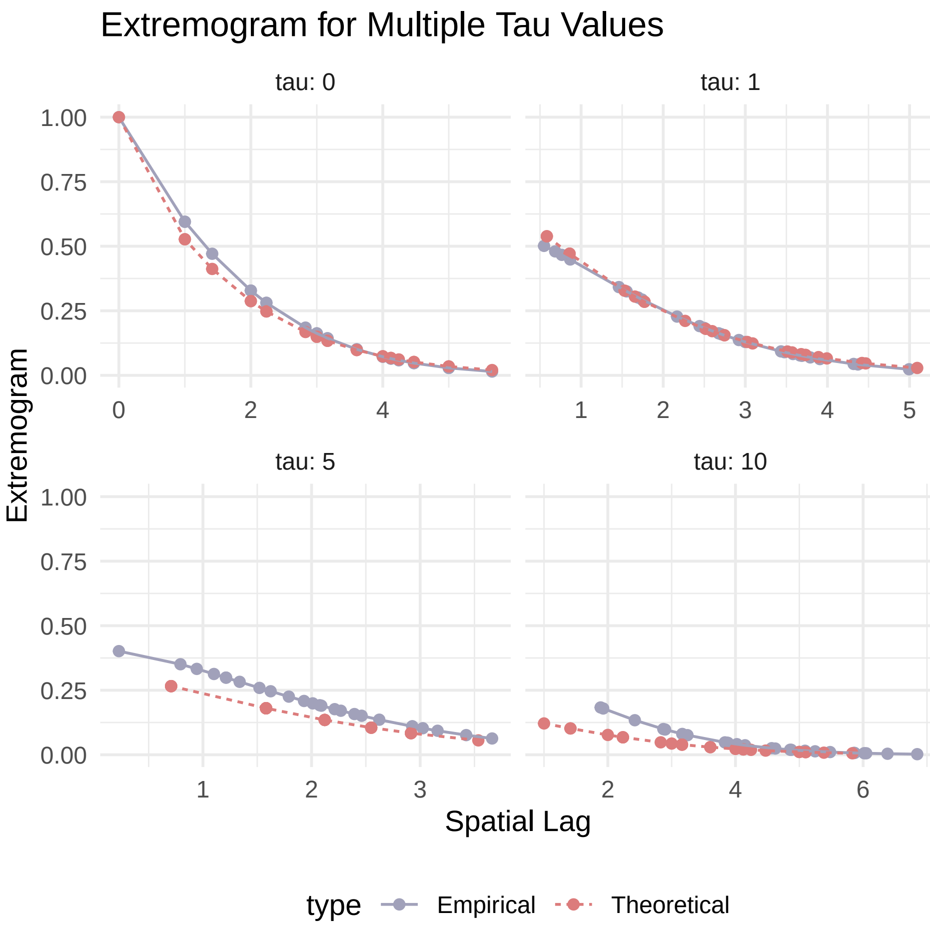

Figure 16: Extremogram for the estimated parameters for the 10th simulation of 100 replicates each

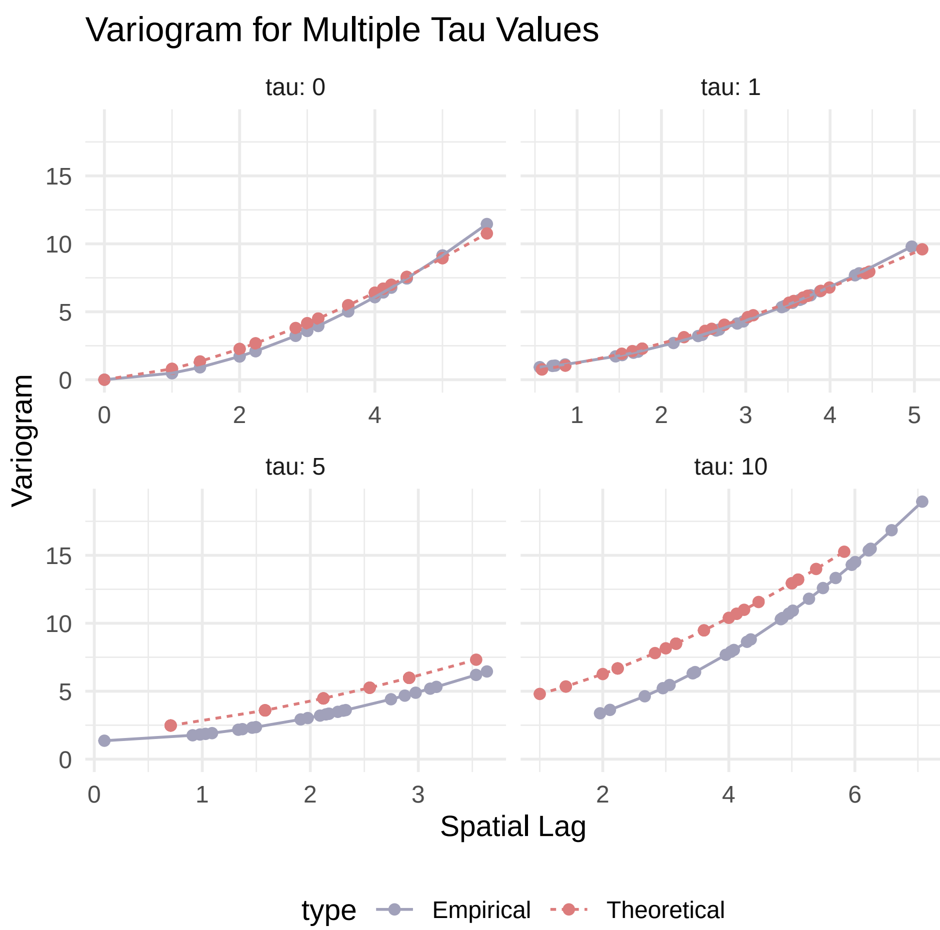

Figure 17: Variogram for the estimated parameters for the third simulation of 100 replicates each

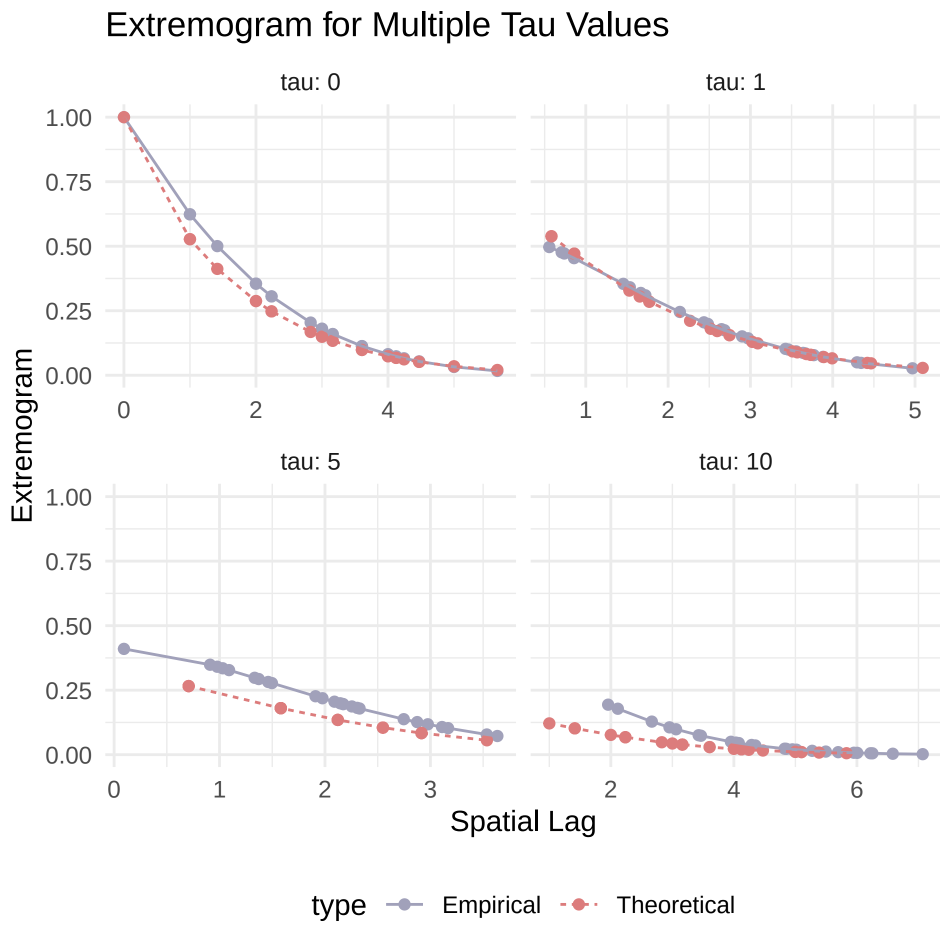

Figure 18: Extremogram for the estimated parameters for the third simulation of 100 replicates each

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

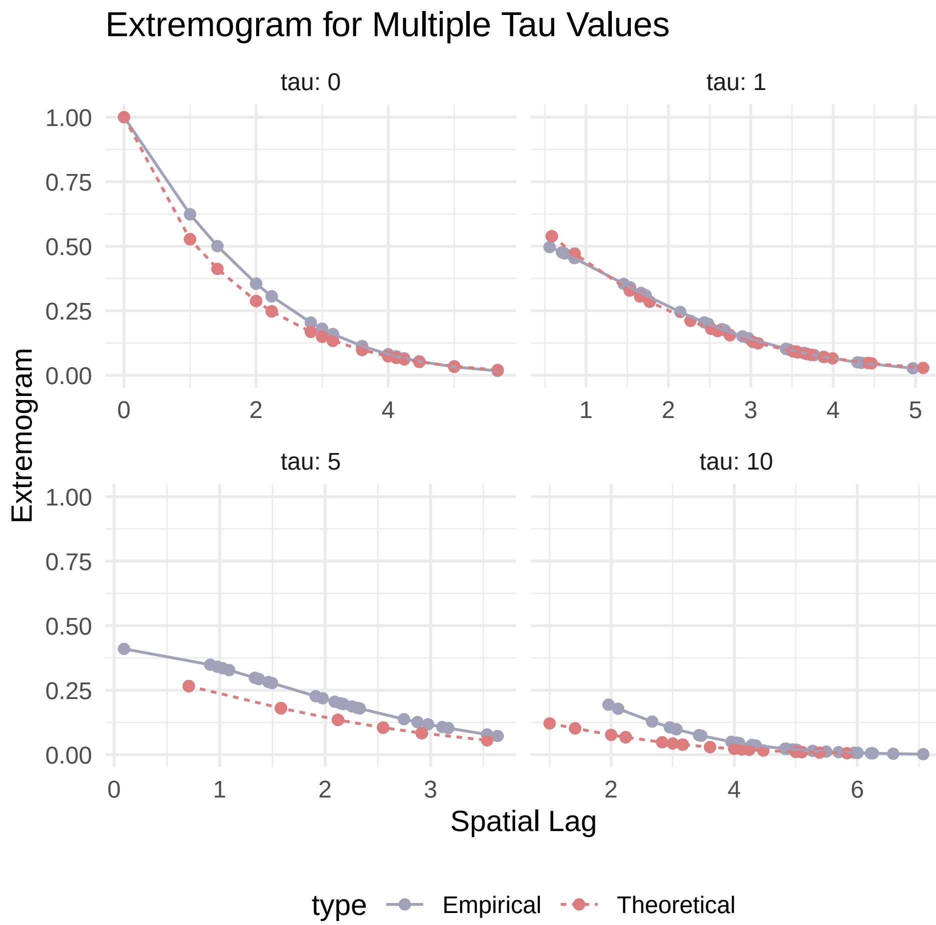

Figure 19: Extremogram for the mean estimated parameters for the 50 simulations of 100 replicates each

Plot by fixing advection parameters

With another advection with $V_x = 0$

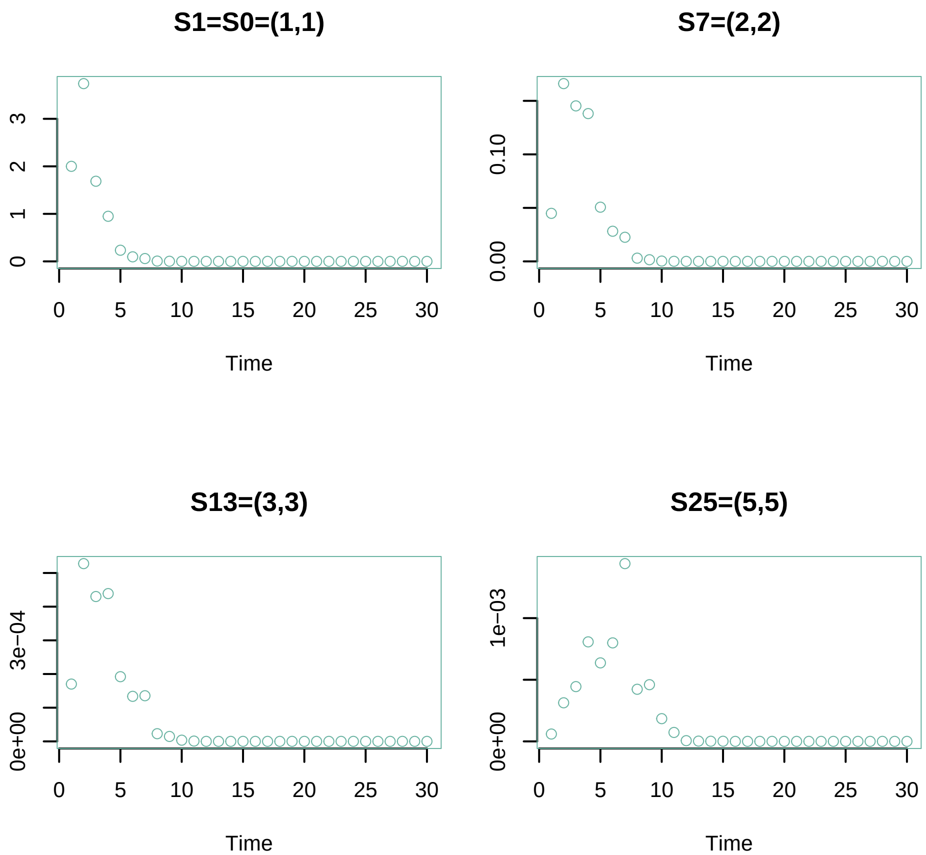

Figure 20: Time series for 4 sites of the first realization in the advection direction

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

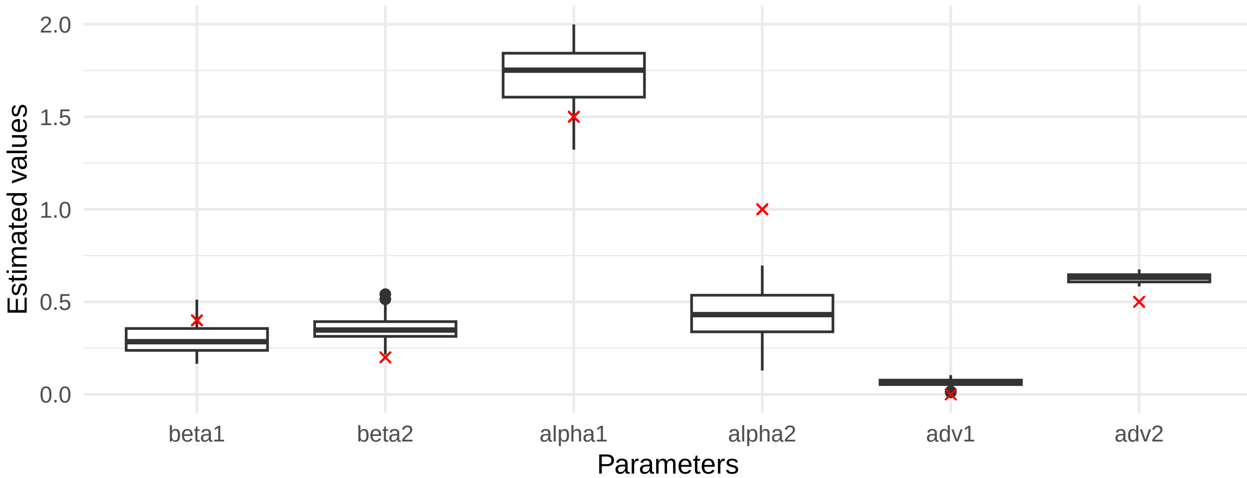

Figure 21: Boxplot of the estimated parameters for the 50 simulations of 100 replicates each

For one simulation:

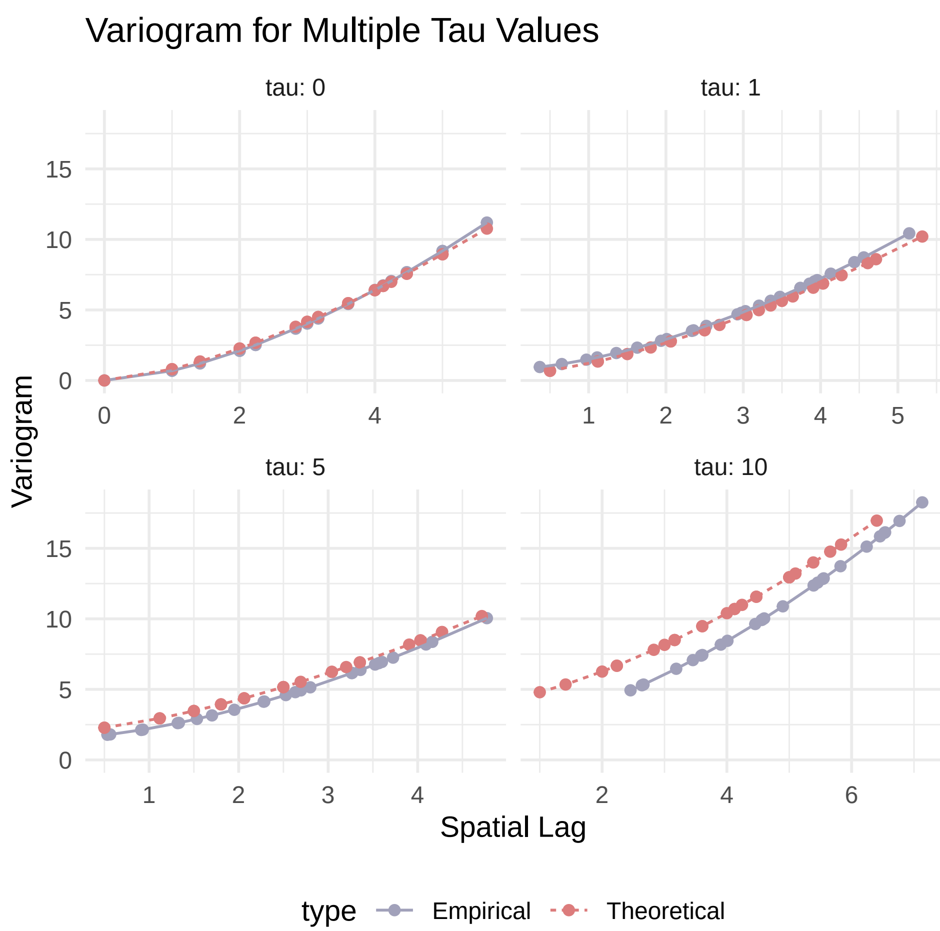

Figure 22: Variogram for the estimated parameters for the 30th simulation of 100 replicates each

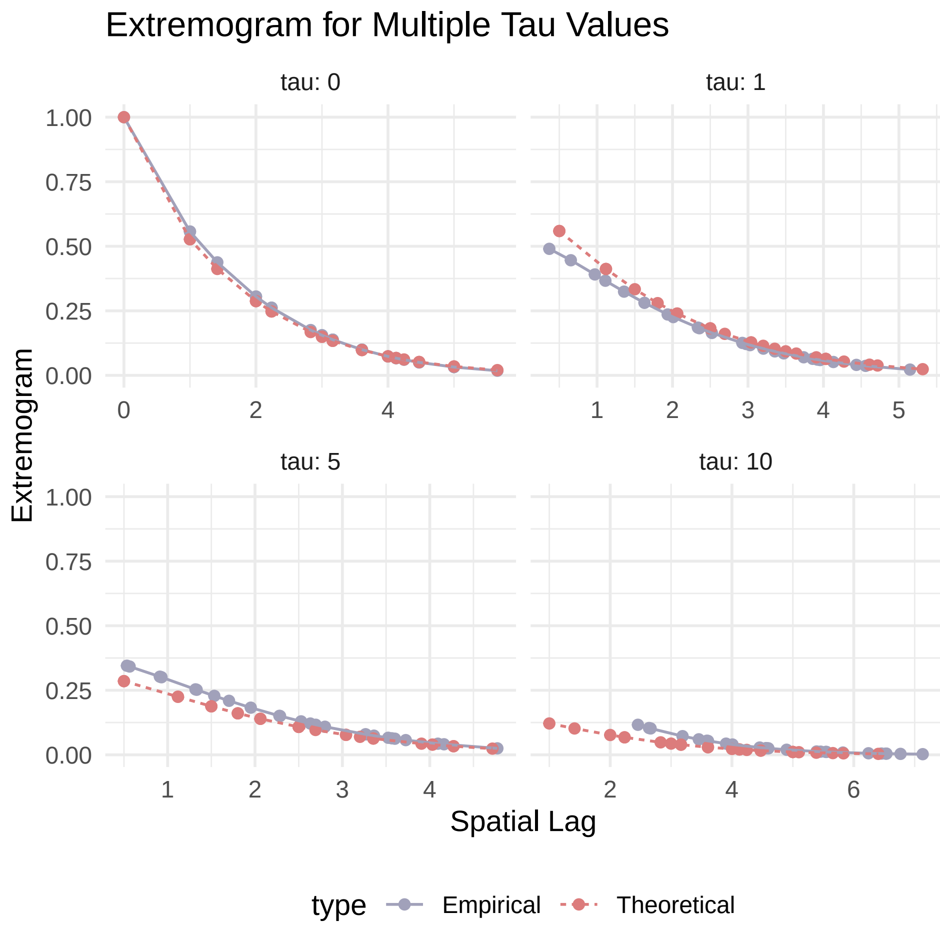

Figure 23: Extremogram for the estimated parameters for the 30th simulation of 100 replicates each

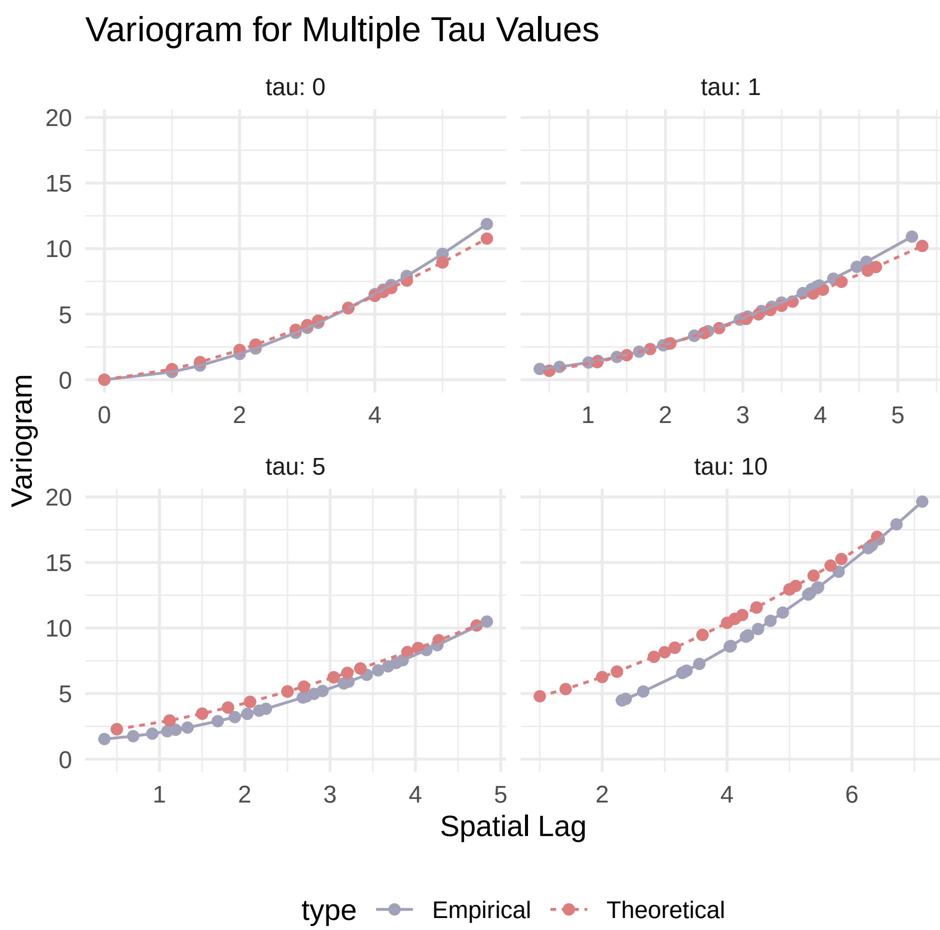

Figure 24: Variogram for the mean estimated parameters for the 50 simulations of 100 replicates each

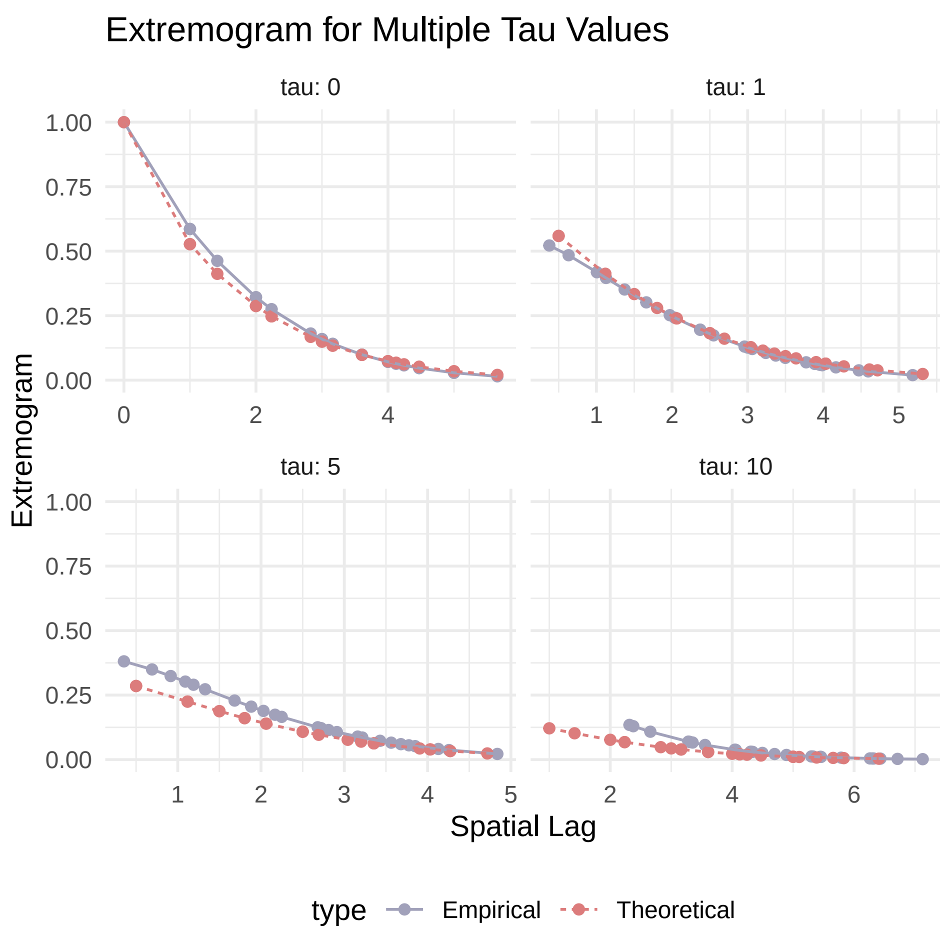

Figure 25: Extremogram for the mean estimated parameters for the 50 simulations of 100 replicates each