Supplementary material for "A model for space-time threshold exceedances with an application to extreme rainfall"

1 Proof of results on spatial dependence

Let Z_t be the space-time process defined by equation (2.1) with innovations given by equation (2.3). Let also e be a stationary random field on S with $F_e(x) = \Pr(e(s) \leq x)$, satisfying the asymptotic dependence condition

$$\lim_{p \to 1^-} \Pr(F_e(e(s_2)) > p, \ldots, F_e(e(s_d)) > p | F_e(e(s_1)) > p) > 0 \tag{S.1}$$

We need to prove that with these specifications Z_t is also asymptotically dependent in space, i.e.

$$\lim_{p \to 1^-} \Pr(\Phi\left(Z_t(s_2)\right) > p, \ldots, \Phi\left(Z_t(s_d)\right) > p | \Phi\left(Z_t(s_1)\right) > p) > 0 \tag{S.2}$$

for all t, all $s_1, \ldots, s_d$, and all $d = 2, 3, \ldots$, where Φ is the CDF of a standard Gaussian random variable,

Note that (S.2) is equivalent to

$$\lim_{u \to \infty} \Pr(Z_t(s_2) > u, \ldots, Z_t(s_d) > p | Z_t(s_1) > u) > 0$$

We have

$$\begin{aligned} &\Pr(Z_t(s_2) > u, \ldots, Z_t(s_d) > u | Z_t(s_1) > u) = \\ &\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \Pr(Z_t(s_2) > u, \ldots, Z_t(s_d) > u | Z_t(s_1) > u, Z_{t-1}(s_1) = z_1, \ldots, Z_{t-1}(s_d) = z_d) \cdot \\ &\cdot f_{s_1, \ldots, s_d}(z_1, \ldots, z_d) dz_1 \ldots dz_d \end{aligned} \tag{S.3}$$

where $f_{s_1,\dots,s_d}(z_1,\dots,z_d)$ is the joint density of $(Z_t(s_1),\dots,Z_t(s_d))$ having $N(0,1)$ univariate margins. The integral in (S.3) is equivalent to

$$\int_{\mathbb{R}}\cdots\int_{\mathbb{R}} \Pr(\varepsilon_t(s_2) > u-\alpha z_2,\dots,\varepsilon_t(s_d) > u-\alpha z_d|\varepsilon_t(s_1) > u-\alpha z_1)$$

$$f_{s_1,\dots,s_d}(z_1,\dots,z_d)dz_1\dots dz_d =$$

$$\int_{\mathbb{R}}\cdots\int_{\mathbb{R}} \Pr\left(\varepsilon_t^*(s_2) > \frac{u-\alpha z_2}{\sqrt{1-\alpha^2}},\dots,\varepsilon_t^*(s_d) > \frac{u-\alpha z_d}{\sqrt{1-\alpha^2}} \mid \varepsilon_t^*(s_1) > \frac{u-\alpha z_1}{\sqrt{1-\alpha^2}}\right)\cdot$$

$$\cdot f_{s_1,\dots,s_d}(z_1,\dots,z_d)dz_1\dots dz_d =$$

$$\int_{\mathbb{R}}\cdots\int_{\mathbb{R}} \Pr\left(\varepsilon_t^*(s_2) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1-z_2),\dots,\varepsilon_t^*(s_d) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1-z_d)|\varepsilon_t^*(s_1) > u^*\right)\cdot$$

$$\cdot f_{s_1,\dots,s_d}(z_1,\dots,z_d)dz_1\dots dz_d \tag{S.4}$$

where $\varepsilon_t^*(s) = \varepsilon_t(s)/\sqrt{1-\alpha^2} \sim N(0,1)$ and $u^* = (u-\alpha z_1)/\sqrt{1-\alpha^2}$. The limit as $u \to \infty$ of (S.3) is equivalent to the limit as $u^* \to \infty$ of (S.4). The integrand in (S.4) is bounded above by $f_{s_1,\dots,s_d}(z_1,\dots,z_d)$ which is integrable. Hence, by the dominated convergence theorem we can take the limit as $u^* \to \infty$ inside the integral, i.e.

$$\lim_{u\to\infty} \Pr(Z_t(s_2) > u,\dots,Z_t(s_d) > u|Z_t(s_1) > u) =$$

$$\int_{\mathbb{R}}\cdots\int_{\mathbb{R}} \lim_{u^*\to\infty} \Pr\left(\varepsilon_t^*(s_2) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1-z_2),\dots,\varepsilon_t^*(s_d) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1-z_d)|\varepsilon_t^*(s_d) > u^*\right)\cdot$$

$$\cdot f_{s_1,\dots,s_d}(z_1,\dots,z_d)dz_1\dots dz_d \tag{S.5}$$

For $k_2,\dots,k_d > 0$, we have

$$\lim_{u^*\to\infty} \Pr(\varepsilon_t^*(s_2) > u^* - k_2,\dots,\varepsilon_t^*(s_d) > u^* - k_d|\varepsilon_t^*(s_1) > u^*) \geq c > 0$$

which follows from

$$\lim_{u^*\to\infty} \Pr(\varepsilon_t^*(s_2) > u^*,\dots,\varepsilon_t^*(s_d) > u^*|\varepsilon_t^*(s_1) > u^*) =$$

$$= \lim_{u^*\to\infty} \Pr(F_e(e(s_2)) > \Phi(u^*),\dots,F_e(e(s_d)) > \Phi(u^*)|F_e(e(s_1)) > \Phi(u^*)) =$$

$$= \lim_{p\to 1^-} \Pr(F_e(e(s_2)) > p,\dots,F_e(e(s_d)) > p|F_e(e(s_1)) > p) = c > 0$$

where the final inequality holds by property (S.1).

The integral (S.5) can be decomposed as follows

$$= \int_{\mathbb{R}} \int_{x_1}^{\infty} \cdots \int_{x_d}^{\infty} \lim_{u^* \to \infty} \Pr\left(\varepsilon_t^*(s_2) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1 - z_2), \ldots, \right.$$

$$\left. \varepsilon_t^*(s_d) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1 - z_d) | \varepsilon_t^*(s_1) > u^* \right) f_{s_1,\ldots,s_d}(z_1, \ldots, z_d) dz_2 \ldots dz_d dz_1 + \tag{S.6}$$

$$\int_{\mathbb{R}} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} \lim_{u^* \to \infty} \Pr\left(\varepsilon_t^*(s_2) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1 - z_2), \ldots, \right.$$

$$\left. \varepsilon_t^*(s_d) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1 - z_d) | \varepsilon_t^*(s_1) > u^* \right) f_{s_1,\ldots,s_d}(z_1, \ldots, z_d) dz_2 \ldots dz_d dz_1 \geq$$

$$\int_{\mathbb{R}} \int_{x_1}^{\infty} \cdots \int_{x_d}^{\infty} \lim_{u^* \to \infty} \Pr\left(\varepsilon_t^*(s_2) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1 - z_2), \ldots, \right.$$

$$\left. \varepsilon_t^*(s_d) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1 - z_d) | \varepsilon_t^*(s_1) > u^* \right) f_{s_1,\ldots,s_d}(z_1, \ldots, z_d) dz_2 \ldots dz_d dz_1$$

Consider the case $\alpha > 0$. Then, by the arguments above,

$$\int_{\mathbb{R}} \int_{x_1}^{\infty} \cdots \int_{x_d}^{\infty} \lim_{u^* \to \infty} \Pr\left(\varepsilon_t^*(s_2) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1 - z_2), \ldots, \right.$$

$$\left. \varepsilon_t^*(s_d) > u^* + \frac{\alpha}{\sqrt{1-\alpha^2}}(z_1 - z_d) | \varepsilon_t^*(s_1) > u^* \right) f_{s_1,\ldots,s_d}(z_1, \ldots, z_d) dz_2 \ldots dz_d dz_1 \geq$$

$$\int_{\mathbb{R}} \int_{x_1}^{\infty} \cdots \int_{x_d}^{\infty} c \cdot f_{s_1,\ldots,s_d}(z_1, \ldots, z_d) dz_2 \ldots dz_d dz_1 = c \cdot \Pr\left(Z_t(s_1) \leq \min(Z_t(s_2), \ldots, Z_t(s_d)) \right)$$

By the stationarity of $Z_t(s)$, we have $\Pr\left(Z_t(s_1) \leq \min(Z_t(s_2), \ldots, Z_t(s_d)) \right) > 0$, from which the result follows. The case $\alpha < 0$ can be dealt with in a similar manner by focussing on the region $(-\infty, x_1)$ in the inner integral of (S.6).

We now need to prove that the space-time process Z_t in (2.1) with innovations given by (2.3) has asymptotically independent non-simultaneous exceedances, that is

$$\lim_{p \to 1^-} \Pr(\Phi\left(Z_{t_2}(s_2)\right) > p, \ldots, \Phi\left(Z_{t_d}(s_d)\right) > p | \Phi\left(Z_{t_1}(s_1)\right) > p) = 0$$

for all $d > 1$, all $s_1, \dots, s_d$ and all $t_1, \dots, t_d$, with at least one time different from the others. This is equivalent to proving that

$$\lim_{u\to\infty} \Pr(Z_{t_2}(s_2) > u, \dots, Z_{t_d}(s_d) > u | Z_{t_1}(s_1) > u) = 0$$

The statement is true if $s_1 = \dots = s_d$ by the asymptotic independence properties of a first-order Gaussian autoregressive process. Therefore, we will assume that at least one of the sites is different from the others and without loss of generality take $s_2 \neq s_1$. Then,

$$\Pr(Z_{t_2}(s_2) > u, \dots, Z_{t_d}(s_d) > u | Z_{t_1}(s_1) > u) \leq \Pr(Z_{t_2}(s_2) > u | Z_t(s_1) > u)$$

Hence, if we show that $\Pr(Z_{t_2}(s_2) > u | Z_{t_1}(s_1) > u) \to 0$ as $u \to \infty$, the result follows. As $Z_t(s) \sim N(0,1)$ for any $s \in S$ and any $t \in \mathbb{Z}$, we have

$$\Pr(Z_{t_2}(s_2) > u | Z_{t_1}(s_1) > u) = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\Pr(Z_{t_2}(s_2) > u, Z_{t_1}(s_1) > u | Z_{t_2-1}(s_2) = x, Z_{t_1-1}(s_1) = y)}{1 - \Phi(u)} \cdot$$

$$\cdot f_{Z_{t_2-1}(s_2), Z_{t_1-1}(s_1)}(x,y) dx dy =$$

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\Pr(\varepsilon_{t_2}(s_2) > u - \alpha x, \varepsilon_{t_1}(s_1) > u - \alpha y)}{1 - \Phi(u)} f_{Z_{t_2-1}, Z_{t_1-1}}(x,y) dx dy =$$

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\left(1 - \Phi\left(\frac{u-\alpha x}{\sqrt{1-\alpha^2}}\right)\right)\left(1 - \Phi\left(\frac{u-\alpha y}{\sqrt{1-\alpha^2}}\right)\right)}{1 - \Phi(u)} f_{Z_{t_2-1}(s_2), Z_{t_1-1}(s_1)}(x,y) dx dy \qquad \text{(S.7)}$$

where $f_{Z_{t_2-1}(s_2), Z_{t_1-1}(s_1)}(x,y)$ denotes the joint density of $(Z_{t_2-1}(s_2), Z_{t_1-1}(s_1))$. From the Mills ratio approximation, $1 - \Phi(z) \sim \phi(z)/z$ for large z. Therefore, for large u,

$$\frac{\left(1 - \Phi\left(\frac{u-\alpha x}{\sqrt{1-\alpha^2}}\right)\right)\left(1 - \Phi\left(\frac{u-\alpha y}{\sqrt{1-\alpha^2}}\right)\right)}{1 - \Phi(u)} \sim$$

$$\sim \exp\left\{-\frac{1}{2(1-\alpha^2)}\left(u^2(1+\alpha^2) - 2\alpha u(x+y) + \alpha^2(x^2+y^2)\right)\right\} \cdot O(u^{-1}) \to 0 \qquad \text{as } u \to \infty$$

As the integrand in (S.7) is bounded above by the integrable function $f_{Z_{t_2-1}(s_2), Z_{t_1-1}(s_1)}$, we can take the limit as $u \to \infty$ inside the integral and the result follows.

2 Simulation study on the Indirect Inference Algorithm

To assess the performance of the Indirect Inference Algorithm (IIA), a simulation study was carried out, assuming equation (2.1) as the data-generating process and using Model 1 of Section 4 as the auxiliary model. Two different scenarios were considered, which, in the notation of Section 4, correspond to Model 1 with $\alpha = 0.35$, $\psi_1 = 8000$ and $\psi_2 = 0.40$ and Model 4 with $\alpha = 0.35$, $\psi_1 = 350$ and $\psi_2 = 0.40$. Model 1 allows a comparison between the direct estimator of $\theta = (\alpha, \psi_1, \psi_2)$ obtained by maximizing PL^G and the indirect estimator derived from IIA. It also provides the most favourable setting for IIA as the auxiliary model coincides with the target one. By contrast, Model 4 is structurally the furtherest away from the auxiliary model among the four formulations of Section 4 and no direct estimator of θ is available. For each scenario, reproducing the locations of the North Brabant rainfall stations, a space-time dataset over 30 sites and with 4000 time points was simulated. The simulation was repeated 100 times in order to reconstruct the estimators' sampling distribution. For IIA, we set $M = 10$. Table S.1 summarizes the simulation results, reporting the estimators' bias and root mean squared error divided by the true value of the parameters to ease comparisons across models. For Model 1, when moving from the direct to the indirect estimator, we observe an increase in bias, but a limited loss in the overall efficiency measured by the relative root mean square error. For Model 4, the comparison with the direct estimator is unavailable, nonetheless, the summary quantities suggest a good performance of the indirect estimator for the parameter α and ψ_1 and a slight worsening of the accuracy and precision of the estimator of ψ_2.

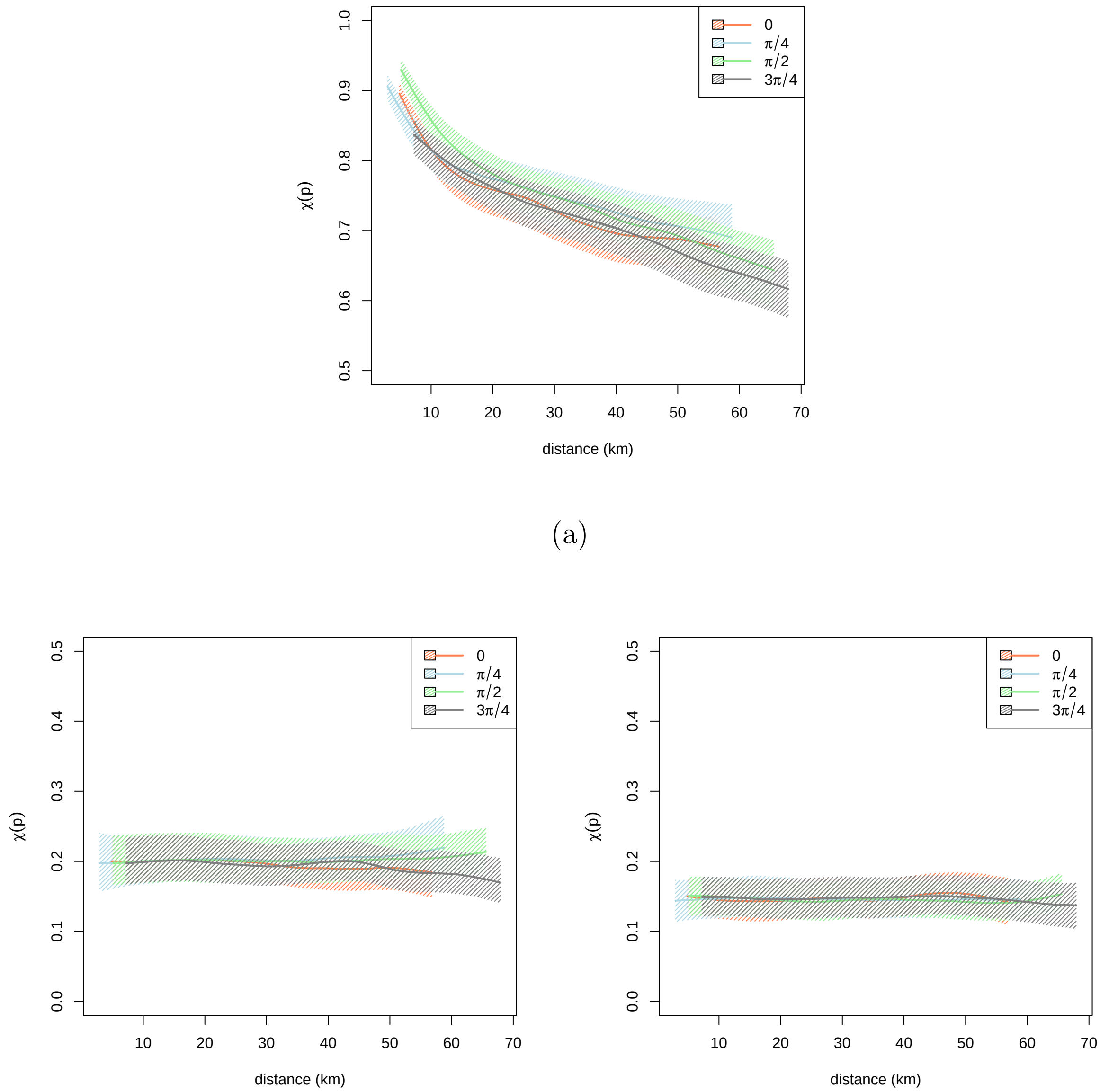
Target model - estimator	$\frac{\text{Bias}(\hat{\alpha})}{\alpha}$	$\frac{\text{Bias}(\hat{\psi}_1)}{\psi_1}$	$\frac{\text{Bias}(\hat{\psi}_2)}{\psi_2}$	$\frac{\text{RMSE}(\hat{\alpha})}{\alpha}$	$\frac{\text{RMSE}(\hat{\psi}_1)}{\psi_1}$	$\frac{\text{RMSE}(\hat{\psi}_2)}{\psi_2}$
Model 1 - direct estimator	0.0012	-0.011	0.0042	0.090	0.15	0.024
Model 1 - indirect estimator	0.020	0.050	-0.0082	0.098	0.21	0.031
Model 4 - indirect estimator	-0.0049	0.068	-0.029	0.092	0.14	0.049

Table S.1: For Model 1 and Model 4, relative bias and root mean square error of the estimators of α, ψ_1 and ψ_2. For Model 1 summaries for both the direct and indirect estimators are reported.

3 Analysis of the North Brabant data

To explore anisotropy at extreme levels for the North Brabant data of Section 4, we computed empirical estimates of $\chi(p)$ in (1.1) for $p = 0.90$ for pairs of sites at increasing distances along four directions corresponding to the angles $0, \pi/4, \pi/2$ and $3\pi/4$. For each of the time lags $0, 1$ and 2, smoothed estimates and 95% pointwise confidence bands are shown in Figure S.1. No significant differences are detected across directions for any of the time lags considered.

Figure S.2 shows smoothed empirical estimates of $\chi(p; l, \|h\|)$ as defined in (4.4) for $l = 2$, as a function of $\|h\|$ and $p \in \{0.90, 0.95, 0.99, 0.995\}$. Also shown are the corresponding model-based estimates from Models 1, 2, 4 and SW. All formulations perform well.



(a)

(b)

(c)

Figure S.1: Smoothed empirical estimates of $\chi(p)$ for $p = 0.90$, and associated approximate 95% pointwise confidence bands computed from pairs of sites at increasing distances (displayed on the x-axis) along four directions: angles 0, $\pi/4$, $\pi/2$ and $3\pi/4$. Pairs of sites are studied, on the same day in (a), with a one-day lag in (b), and with a two-day lag in (c), respectively.

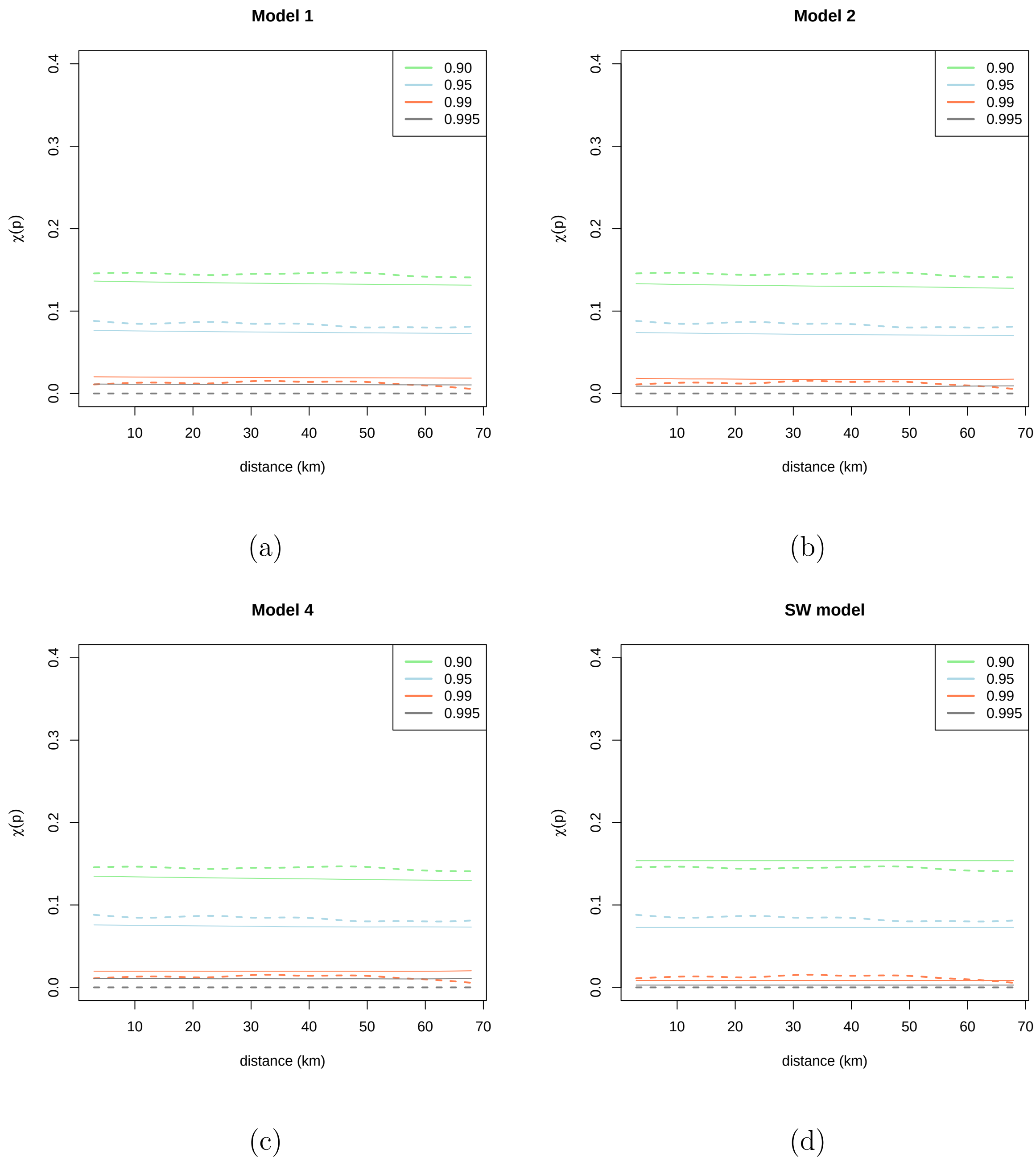


Figure S.2: Estimates of $\chi(p; 2, \|h\|)$ as a function of $\|h\|$, for $p \in \{0.90, 0.95, 0.99, 0.995\}$. In each plot, the dashed line corresponds to smoothed empirical estimates. In (a) the continuous line corresponds to Model 1 estimates, in (b) to Model 2 estimates, in (c) to Model 4 estimates and in (d) to SW model estimates, respectively.