Approximately Linear INGARCH Models for Spatio-Temporal Counts

Malte Jahn, Christian H. Weiß

Department of Mathematics & Statistics,

Helmut Schmidt University, Hamburg

Hee-Young Kim

Department of Applied Statistics,

Korea University, Sejong

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Approximately Linear INGARCH Models for Count Time Series

Introduction

Count processes (X_t): range contained in $\mathbb{N}_0$,

i. e., set of *non-negative* integers (see Weiß, 2018).

Distinguish **two cases:**

- **unbounded** counts (X_t), i. e., full $\mathbb{N}_0$ as range;
- **bounded** counts (X_t), i. e., upper bound $n \in \mathbb{N}$,

  so range $\{0, \ldots, n\}$ also bounded from above!

(Exactly linear) **INGARCH model** (Ferland et al., 2006)

for **unbounded** counts (X_t): Let $M_t = E[X_t|\mathcal{F}_{t-1}]$, then

$$M_t = a_0 + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=1}^{q} b_j M_{t-j}.$$

Counts $X_t|\mathcal{F}_{t-1}$ emitted by, for example, Poi(M_t).

(Exactly linear) **BINGARCH model** (Ristić et al., 2016)

for **bounded** counts (X_t): Let $P_t = \frac{1}{n} M_t$, then

$$P_t \;=\; a_0 \;+\; \sum_{i=1}^{p} a_i\, X_{t-i}/n \;+\; \sum_{j=1}^{q} b_j\, P_{t-j}.$$

Counts $X_t|\mathcal{F}_{t-1}$ emitted by, for example, $\mathrm{Bin}(n, P_t)$.

Problem: *In both cases, rather limiting parameter constraints that prevent negative parameter and ACF values.*

Possible solution: approximately linear (B)INGARCH models by Weiß et al. (2022), Weiß & Jahn (2021), with negative ACF values while (nearly) linear model structure.

Solution by Weiß et al. (2022) for unbounded counts:

softplus INGARCH model with response function $f = \mathrm{sp}_c$.

Softplus function

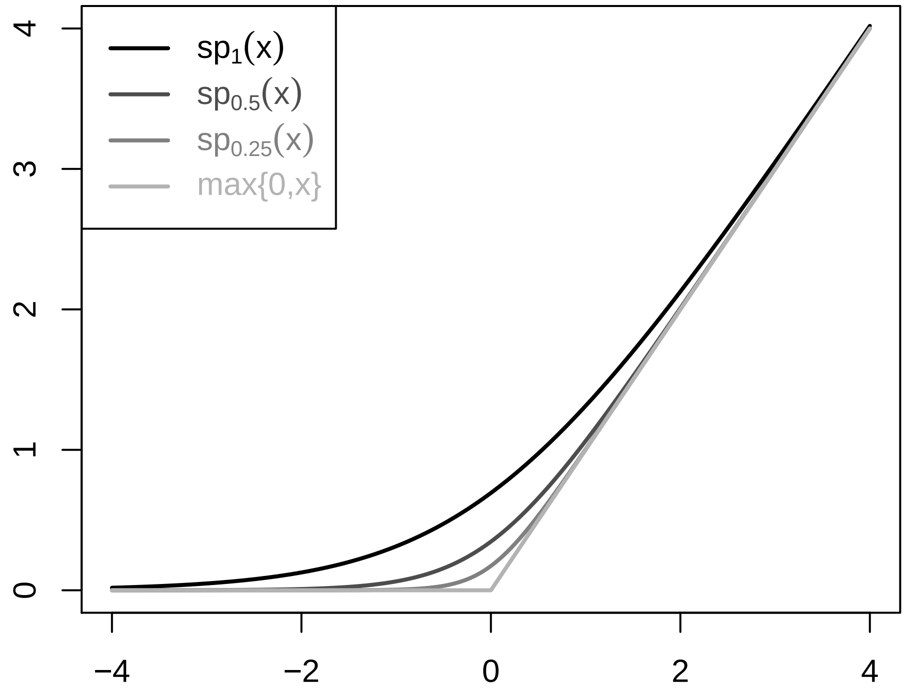
$\mathrm{sp}_c(x) = c\,\ln\big(1 + \exp(x/c)\big)$

with tuning parameter $c > 0$,

truly positive

and differentiable;

spINGARCH model:



$$M_t = \mathrm{sp}_c\Big(\alpha_0 + \sum_{i=1}^{p} \alpha_i\, X_{t-i} + \sum_{j=1}^{q} \beta_j\, M_{t-j}\Big).$$

Solution by Weiß & Jahn (2021) for bounded counts:

soft-clipping BINGARCH with response fct. $f = \mathrm{sc}_c$.

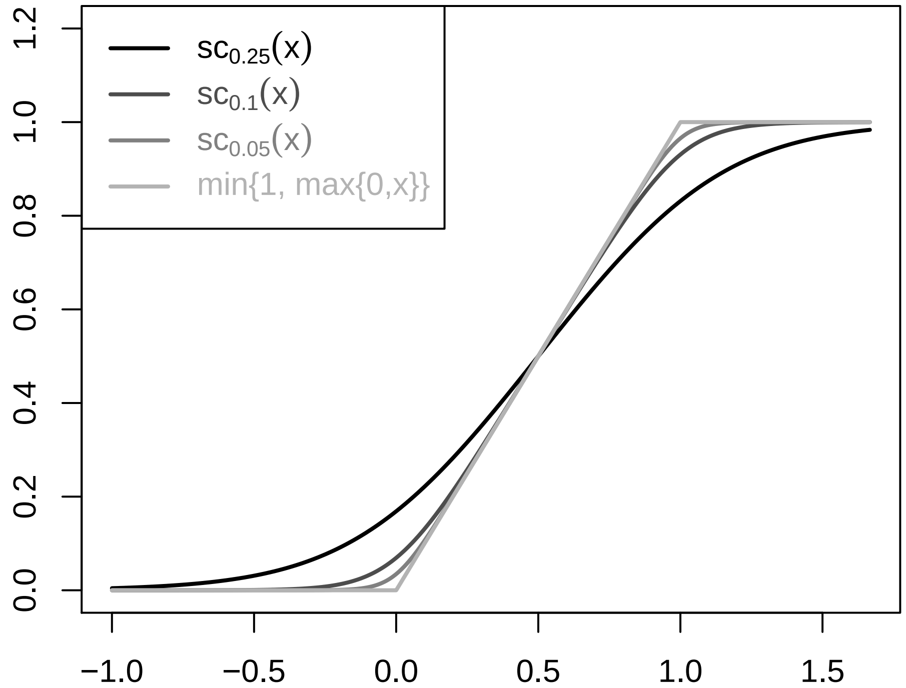
Soft-clipping function

$$\mathrm{sc}_c(x) = c \ln\left(\frac{1+\exp(\frac{x}{c})}{1+\exp(\frac{x-1}{c})}\right)$$

with tuning parameter $c > 0$,

truly in $(0;1)$

and differentiable;



scBINGARCH model:

$$P_t = \mathrm{sc}_c\Big(\alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}/n + \sum_{j=1}^{q} \beta_j P_{t-j}\Big).$$

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

On Linear Spatio-Temporal INGARCH Models

Existing Literature

Up to now, spatio-temporal (ST) models for *un*bounded counts, mainly applications in epidemiology (disease surveillance).

Let $\boldsymbol{X}_t = (X_{t,1},\ldots,X_{t,m})^\top \in \mathbb{N}_0^m$, $m \in \mathbb{N}$, express count $X_{t,i}$ at time t in unit i. For INGARCH-type models, denote conditional mean $M_{t,i}$, used for generating $X_{t,i}|\boldsymbol{\mathcal{F}}_{t-1}$.

ST-INARCH(1) model (Held et al., 2005; Paul et al., 2008):

$$M_{t,i} \;=\; \lambda_i\, X_{t-1,i} \;+\; \phi_i \textstyle\sum_{j\neq i} w_{ji}\, X_{t-1,j} \;+\; \gamma_i,$$

where, e. g., spatial weights $w_{ji} = 0$ if unit j not neighbour of i, otherwise $w_{ji} = 1/n_j$ (row-normalized, n_j=neighbours of j).

ST-INARCH(1) model in matrix notation: $\boldsymbol{M}_t = \boldsymbol{\Gamma}\, \boldsymbol{X}_{t-1} + \boldsymbol{\gamma}$,

where $\boldsymbol{\Gamma}$ diagonal entries λ_i and off-diagonal entries $\phi_i\, w_{ji}$,

see Held & Paul (2012).

Armillotta & Fokianos (2021): unique λ, ϕ, γ across locations i.

Higher-order extension by Bracher & Held (2020):

$$\boldsymbol{M}_t = \sum_{k=1}^{p} u_k\, \boldsymbol{\Gamma}\, \boldsymbol{X}_{t-k} + \boldsymbol{\gamma},$$

where AR-weights $u_1, \ldots, u_p$ with $\sum_{k=1}^{p} u_k = 1$, e. g.,

$u_k \sim \kappa^{k-1}$ such that INGARCH$(1,1)$-type model for $p \to \infty$.

ST-INGARCH$(1,1)$ by Clark & Dixon (2021):

$\boldsymbol{M}_t = \gamma_t + a_1\, \boldsymbol{X}_{t-1} + b_1\, \boldsymbol{M}_{t-1}$, i. e., with scalar coefficients.

Problems:

- All model parameters non-negative (to ensure $M_{t,i} > 0$), thus only positive ACF values can be captured.
- No proposal for bounded counts.

Our contributions:

- ST softplus INGARCH$(p, r; q, s)$ (ST-spINGARCH) for unbounded counts − simulations and illustration;
- ST soft-clipping BINGARCH$(p, r; q, s)$ (ST-scBINGARCH) for bounded counts, applied to meteorological time series, solving problems we are confronted with in practice.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

The Spatio-Temporal Softplus INGARCH Model for Unbounded Counts

Definition & Discussion

ST-spINGARCH$(p, r; q, s)$ model with range $\mathbb{N}_0^m$:

$$M_{t,i} = \mathrm{sp}_c\Big(\alpha_0 + \sum_{k=1}^{p} \alpha_k\, X_{t-k,i} + \sum_{g=1}^{r} \lambda_g \sum_{j\neq i} w_{ij}\, X_{t-g,j} + \sum_{l=1}^{q} \beta_l\, M_{t-l,i} + \sum_{h=1}^{s} \phi_h \sum_{j\neq i} w_{ij}\, M_{t-h,j}\Big), \quad i = 1, \ldots, m,$$

where $\boldsymbol{X}_t|\boldsymbol{\mathcal{F}}_{t-1}$ emitted by

multivariate unbounded-counts distribution with mean $\boldsymbol{M}_t$.

Default choice: (conditionally) independent Poisson.

Possible extensions: copula for stronger cross-correlation; different distribution family, e. g., negative binomial; time- or unit-specific parameters.

Model fitting: CML estimation.

If feedback terms included in model,

then initialization of $(\boldsymbol{M}_t)$ by mean $\overline{\boldsymbol{X}}$.

CML estimates computed by numerical optimization routine,

approximate standard errors via inverse Hessian.

Model selection via Akaike's information criterion (AIC).

Simulation study: 1,000 repl., "k-nearest neighbours" (knn).

$m = 20$ Cartesian coordinates $\boldsymbol{u}_i = (x_i, y_i)$ randomly per simulation run, then knn spatial weights with $k = 5$:

$w_{ji} = 1/5$ if $\boldsymbol{u}_j$ one of five nearest points of $\boldsymbol{u}_i$.

Simulation scenarios $(T = 100, 250, 500, 1000)$:

Model		α_0	α_1	α_2	λ_1	λ_2	β_1	ϕ_1
ST-spINGARCH$(1,1;0,0)$	model I	3	0.4		0.2			
	model II	3	-0.4		0.2			
ST-spINGARCH$(2,2;0,0)$	model III	1.5	0.1	0.15	0.2	0.3		
	model IV	1.5	-0.1	0.15	0.2	0.3		
ST-spINGARCH$(1,1;1,1)$	model V	1.5	0.1		0.15		0.2	0.3
	model VI	1.5	0.3		0.15		0.2	-0.1

Summary of findings:

Bias and RMSE decrease as T increases (consistency).

CML estimates for AR-type models $(1,1;0,0)$ and $(2,2;0,0)$ converge faster than for feedback model $(1,1;1,1)$.

Illustrative application: Chicago crime data.

Monthly number of burglaries ($T = 72$) in $m = 552$ census block groups, see Clark & Dixon (2021). Two blocks are neighbours if they share border. Spatial weights $w_{ji} = 1/n_j$ if i is adjacent to j (n_j = number of neighbours of j).

Armillotta & Fokianos (2021) fit first- and second-order "Poisson network autoregression" (PNAR) model (linear or log-linear). Information criteria prefer linear models, best performance by PNAR(2) model.

Illustrative application: Chicago crime data.

Using Poi-ST-spINGARCH$(1,1;0,0)$–$(2,2;0,0)$ with $c = 0.01$, nearly identical estimates as for exactly linear PNAR models:

	α_0	α_1	λ_1	α_0	α_1	α_2	λ_1	λ_2
est. coeff.	0.4550	0.2835	0.3215	0.3207	0.2286	0.1626	0.2077	0.1191
approx. SE	0.0082	0.0048	0.0070	0.0088	0.0049	0.0048	0.0079	0.0078

Thus, Poi-ST-spINGARCH model behaves linear in close approximation.

(. . .)

Illustrative application: Chicago crime data.

By contrast to PNAR model, we can include feedback term:

	α_0	α_1	α_2	λ_1	λ_2	β_1	ϕ_1
est. coeff.	0.0064	0.1004	-0.0892	0.1620	-0.1160	0.9824	-0.0515
approx. SE	0.0012	0.0037	0.0030	0.0054	0.0057	0.0024	0.0031

Softplus link useful as negative values for $\alpha_2, \lambda_2, \phi_1$.

Large drop in magnitude of (non-)spatial AR coefficients,

nearly all “dependency mass” into non-spatial feedback term β_1,

so slowly decaying ACF.

Seems to indicate non-stationarity: negative trend in data.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

The Spatio-Temporal Soft-clipping BINGARCH Model for Bounded Counts

Definition & Discussion

ST-scBINGARCH$(p, r; q, s)$ model with range $\{0, \ldots, n\}^m$:

$$P_{t,i} = \mathrm{sc}_c\Big(\alpha_0 + \sum_{k=1}^{p} \alpha_k\, X_{t-k,i}/n + \sum_{g=1}^{r} \lambda_g \textstyle\sum_{j\neq i} w_{ij}\, X_{t-g,j}/n$$
$$+ \sum_{l=1}^{q} \beta_l\, P_{t-l,i} + \sum_{h=1}^{s} \phi_h \textstyle\sum_{j\neq i} w_{ij}\, P_{t-h,j}\Big),\quad i = 1, \ldots, m,$$

where $\boldsymbol{X}_t|\boldsymbol{\mathcal{F}}_{t-1}$ emitted by

multivariate bounded-counts distribution with mean $n\cdot\boldsymbol{P}_t$.

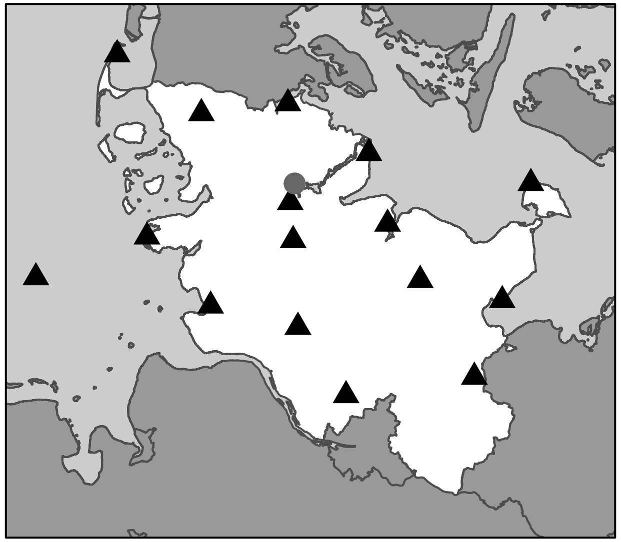
Default choice: (conditionally) independent binomials.

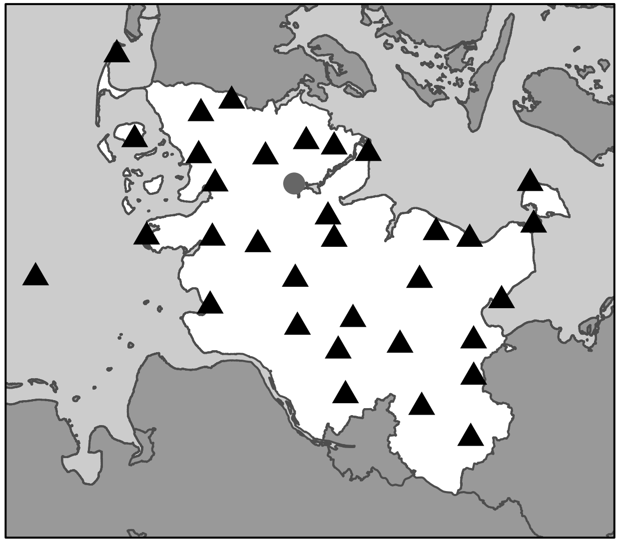
Possible extensions: copula for stronger cross-correlation; different distribution family, e. g., zero-inflated binomial (ZIB).

Parameter estimation by conditional ML approach.

Applied to meteorological t. s. (DWD Climate Data Center), for weather stations in Schleswig-Holstein:

cloud coverage data ($m = 17$): precipitation data ($m = 34$):





grey circle = weather station "Schleswig"

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Cloud Coverage: Dealing with Missing Values

Data Application

Hourly cloud-coverage data (2009–2019), i. e.,

share of visible sky covered by clouds.

Counts (“okta”) from 0 (no clouds) to 8 (sky fully overcast),

i. e., counts $X_{t,i}$ with bound $n = 8$.

Irregular observations:

flag “9” (no data), “-1” (sky not observable),

treated as “NA” (not available, i. e., missing data).

5.5 % of NA observations in $\{x_{t,i}\}$. But

57.2 % of vectors $\boldsymbol{x}_t$ with at least one NA component $x_{t,i}$.

$\Rightarrow$ We have to **impute integer values**!

Our solution: If $x_{t,i}$ missing, then impute with value $x_{t,j}$ from most correlated non-NA station at time t (thus preserving integer nature of data).

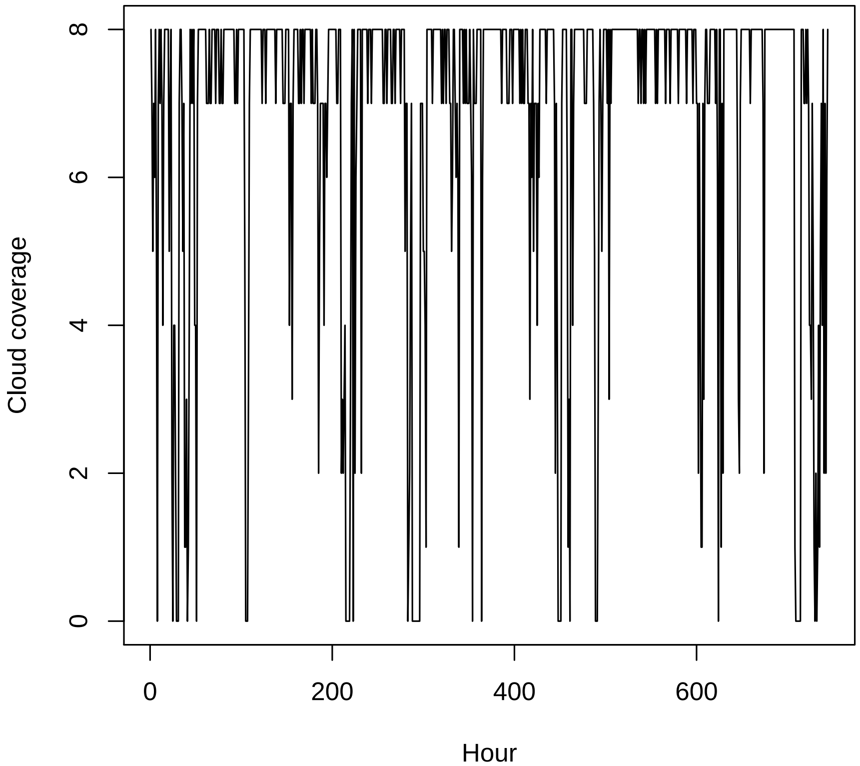
Selected $x_{t,j} \approx$ "nearest neighbour" in terms of cross-correlation.

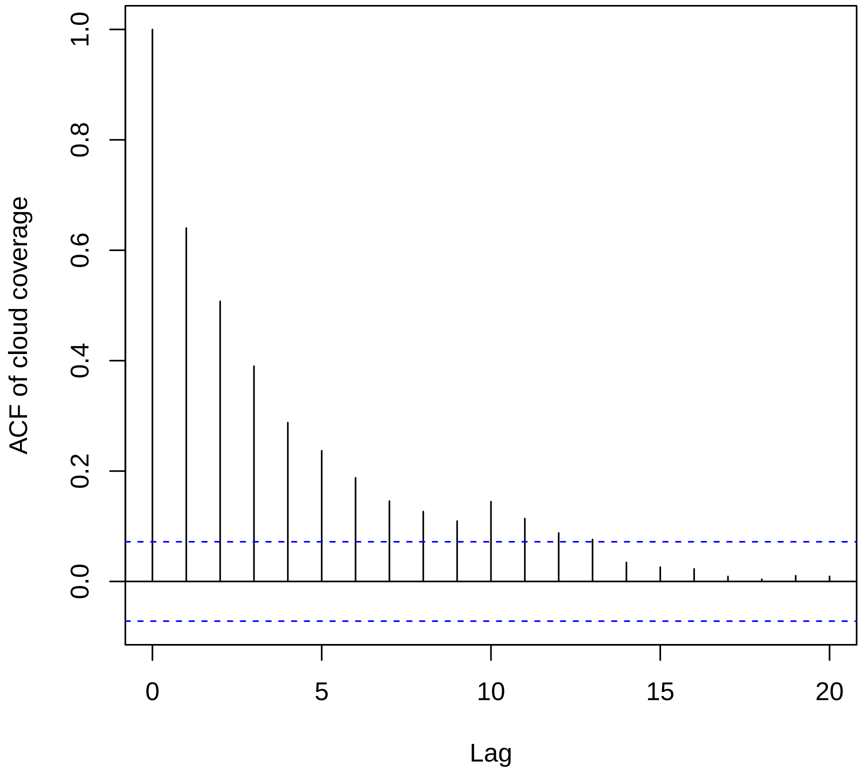
Average maximum corr. between stations over time is 0.709.

Our proposal for **spatial weight matrix W** $= (w_{ij})_{i,j=1,\ldots,m}$: rely on great circle distance d_{ij}, with two extreme cases:

- sparse $\mathbf{W}^{(1)}$: only nearest neighbour, ${w_{ij}}^{(1)} = 1$ if j geographically closest to i;
- dense $\mathbf{W}^{(2)}$: all neighbours, ${w_{ij}}^{(2)} = 1/d_{ij}$ plus row-normalization.

Example: Hourly cloud coverage at "Schleswig", Dec. 2019:





$\Rightarrow$ slowly decaying memory,

use feedback terms in ST-scBINGARCH model.

Cloud Coverage: Dealing with Missing Values

For both $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$, model selection via AIC.

$\mathbf{W}^{(2)}$ generally produces lower AIC values.

Cloud-coverage counts (Dec. 2019):

ML-fitted ST-scBINGARCH$(1,3;1,1)$ model with $\mathbf{W} = \mathbf{W}^{(2)}$.

	α_0	α_1	λ_1	λ_2	λ_3	β_1	ϕ_1
est. coeff.	0.0725	0.3496	0.6544	0.0909	-0.0637	0.2284	-0.3474
approx. SE	0.0060	0.0070	0.0124	0.0511	0.0089	0.0152	0.0515

Note: negative coefficients, soft-clipping link relevant.

Feedback terms β_1 and ϕ_1 have opposing signs:

"compensation effect", for ACF of "foreign" stations,

we have faster decay than under pure AR structure.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Rainy Hours per Day: Accounting for Zero-Inflation

Data Application

Rainy Hours: Accounting for Zero-Inflation

Rainy hours per day: $x_{t,i}$ with upper limit $n = 24$.

Again missing data (0.36 % of total hourly observations),

but in particular, strong **zero inflation** (38 % of zeros).

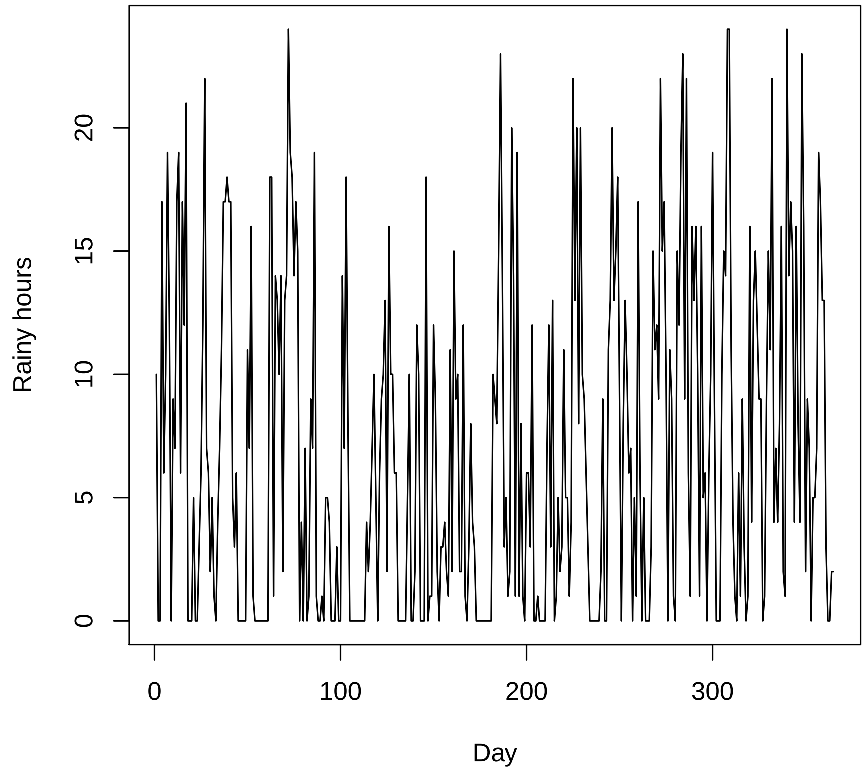
Thus, **conditional ZIB distribution** with parametrization

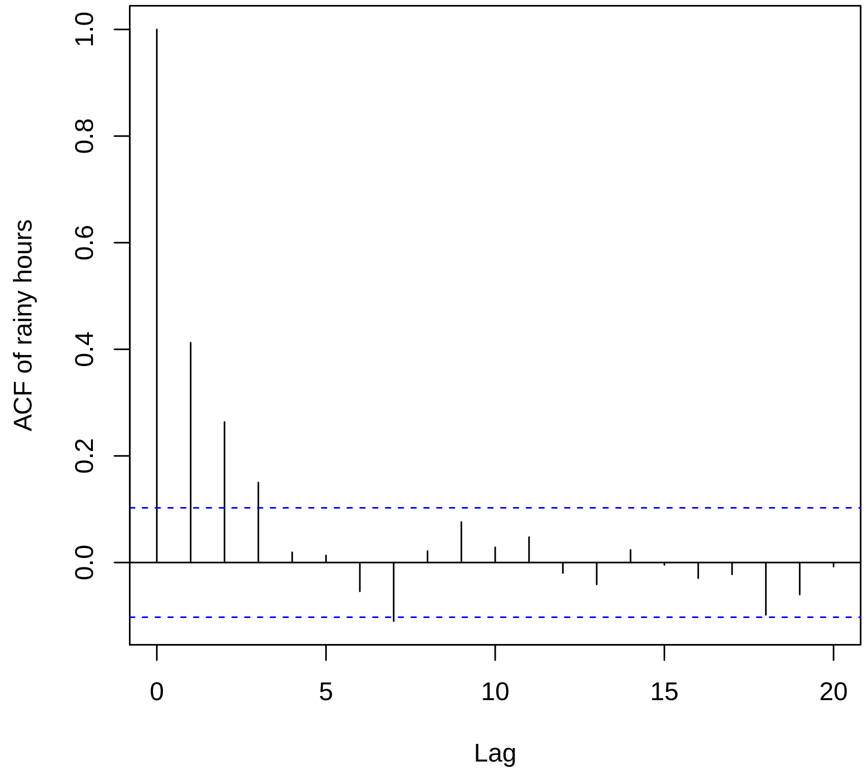
$$\begin{aligned} P\big(X_{t,i} = x \mid \mathcal{F}_{t-1}\big) \;=\; & \omega\,(1 - P_{t,i})\,\mathbb{1}_{\{x=0\}} \\ & + \big(1 - \omega\,(1 - P_{t,i})\big)\, p_{\text{Bin}}\Big(x \mid n, \tfrac{P_{t,i}}{1-\omega\,(1-P_{t,i})}\Big). \end{aligned}$$

(. . .) see plots on next slide!

$\mathbf{W} = \mathbf{W}^{(2)}$	ω	α_0	α_1	α_2	α_3	λ_1	λ_2	λ_3
est. coeff.	0.4359	0.1370	0.1863	0.0986	0.1033	0.0336	-0.0396	-0.1057
approx. SE	0.0047	0.0016	0.0074	0.0082	0.0075	0.0086	0.0096	0.0087

Example: Rainy hours per day at "Schleswig" in 2019:





⇒ ACF drops quickly towards zero,

feedback-terms not needed this time.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

Cloud Coverage: Accounting for Cross-Correlation

Data Application

MATH STAT

Default model (conditionally independent binomials)

only moderate cross-correlation.

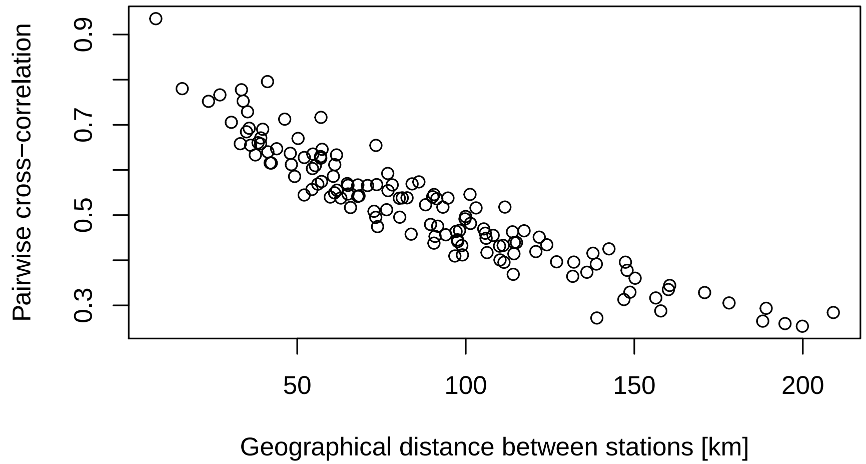
Cloud-coverages (Dec. 2019) with **strong cross-correlation**:

average cross-corr. ≈ 0.52,

negative relationship
to geographical distance.

Default model:

cross-correlation ≈ 0.25.



Idea: like in Armillotta & Fokianos (2021),

employ **copula** distributions with count marginals.

Gaussian copula

$$C(u_1,\ldots,u_m) \;=\; \Phi_{m;\mathbf{R}}\big(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_m)\big)$$

with corresponding copula density

$$c(u_1,\ldots,u_m) \;=\; \det(\mathbf{R})^{-1/2}\,\exp\Big(\tfrac{1}{2}\,\Phi^{-1}(\boldsymbol{u})^{\top}\,\big(\mathbf{I}-\mathbf{R}^{-1}\big)\,\Phi^{-1}(\boldsymbol{u})\Big).$$

With univariate count CDFs $F_1,\ldots,F_m$, joint multivariate CDF

$$F(x_1,\ldots,x_m) \;=\; C\big(F_1(x_1),\ldots,F_m(x_m)\big).$$

Problem: joint PMF $p(x_1,\ldots,x_m)$ requires
2^m discrete differences of joint CDF, not feasible!

Gaussian copula (. . .)

We use approximate solution of Kazianka & Pilz (2010):

$$p(x_1, \ldots, x_m) \;\approx\; c(u_1, \ldots, u_m) \cdot \prod_{i=1}^{m} P(X_i = x_i),$$

where $u_i = \frac{1}{2}\big(F_i(x_i) + F_i(x_i - 1)\big)$ ("mid-point approximation").

How to choose $\mathbf{R}$ for $\Phi_{m;\mathbf{R}}(\cdot)$?

Standard solution:

"exchangeable correlation structure", $\mathbf{R} = (1-\rho)\,\mathbf{I} + \rho\,\mathbf{E}$.

Cloud Coverage: Accounting for Cross-Correlation

How to choose $\mathbf{R}$ for $\Phi_{m;\mathbf{R}}(\cdot)$? $\quad (\ldots)$

To explicitly account for spatial pattern in cross-correlation, we resort to concept of spatially correlated errors.

Spatial error model (SEM): $\quad \boldsymbol{y} = f(\boldsymbol{x}, \boldsymbol{\beta}) + \boldsymbol{u}$, where $\boldsymbol{u} = \rho\,\mathbf{W}\boldsymbol{u} + \boldsymbol{\epsilon}$ with standard normal $\boldsymbol{\epsilon}$ (LeSage & Pace, 2009).

Error $\boldsymbol{u}$ as $\quad \boldsymbol{u} = (\mathbf{I} - \rho\,\mathbf{W})^{-1}\,\boldsymbol{\epsilon}$.
With $\mathbf{B} = (\mathbf{I} - \rho\,\mathbf{W})^{-1}$, resulting covariance matrix $\mathbf{B}\mathbf{B}^\top$.

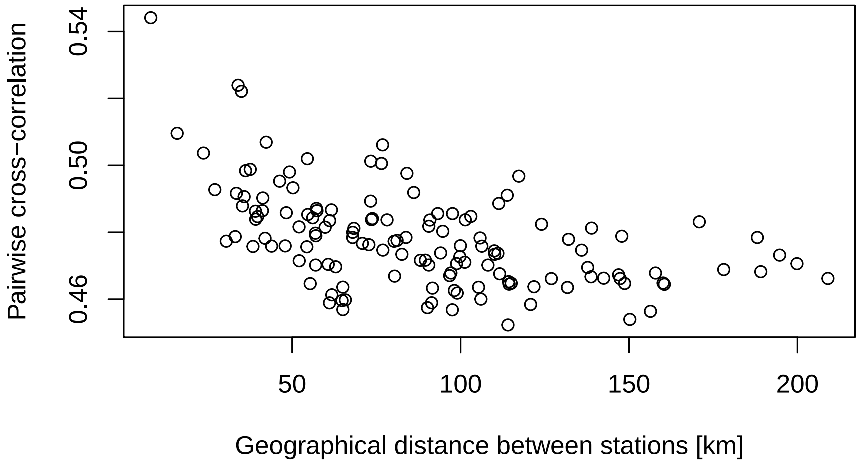
Idea: "SEM copula" with covariance matrix $\mathbf{B}\mathbf{B}^\top$, again single dependence parameter ρ.

SEM-copula ST-scBINGARCH$(1,3;1,1)$ with $\mathbf{W}=\mathbf{W}^{(2)}$:

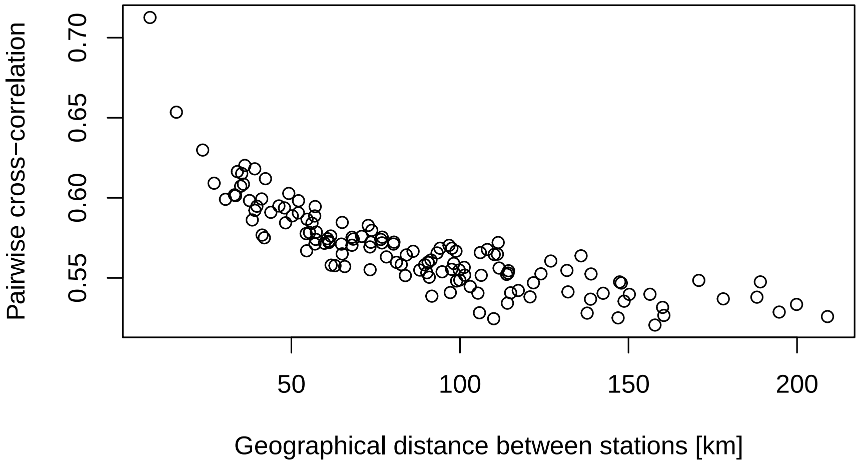
	ρ	α_0	α_1	λ_1	λ_2	λ_3	β_1	ϕ_1
est. coeff.	0.4149	0.0566	0.3452	0.6567	0.0644	-0.0496	0.2335	-0.3181
approx. SE	0.0010	0.0105	0.0079	0.0195	0.0242	0.0211	0.0094	0.0276

Relation between correlation and distance for

standard copula:



SEM copula:



Thank You

for Your Interest!

Helmut Schmidt Universität
Universität der Bundeswehr Hamburg

MATH STAT

Christian H. Weiß

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

`weissc@hsu-hh.de`

Literature

Armillotta & Fokianos (2021) Poisson network . . . *arXiv:2104.06296v2*.

Bracher & Held (2020) Endemic-epidemic models . . . *IJF*, forthc.

Clark & Dixon (2021) A class of spatially . . . *Spat Stat* **43**, 100493.

Ferland et al. (2006) Integer-valued GARCH processes. *JTSA* **27**, 923–942.

Held et al. (2005) A statistical framework . . . *Stat Mod* **5**, 187–199.

Held & Paul (2012) Modeling seasonality . . . *Biom J* **54**, 824–843.

Kazianka & Pilz (2010) Copula-based geostat. . . *SERRA* **24**, 661–673.

LeSage & Pace (2009) *Introduction to spatial econometrics*. T & F.

Paul et al. (2008) Multivariate modelling . . . *Stat Med* **27**, 6250–6267.

Ristić et al. (2016) A binomial int. ARCH model. *Int. J. Biostatistics* **12**.

Weiß (2018) *An Introduction to Discrete-Valued Time Series*. Wiley.

Weiß et al. (2022) Softplus INGARCH mod. *Statistica Sinica* **32**, 1099–1120.

Weiß & Jahn (2021) Soft-clipping INGARCH models . . . *Working Paper*.