STATISTICAL DOWNSCALING WITH GLIMCLIM: 'Generalised LInear Modelling of daily CLIMate sequences'

Chiara Ambrosino
c.ambrosino@ucl.ac.uk

Many Thanks to Dr. Richard E. Chandler

CONTENTS OF THE LECTURE and PRACTICAL

- GLIMCLIM RATIONALE

CONTENTS OF THE LECTURE and PRACTICAL

- GLIMCLIM RATIONALE
- PROBABILISTIC DAILY RAINFALL MODELLING USING GLMs

CONTENTS OF THE LECTURE and PRACTICAL

- GLIMCLIM RATIONALE
- PROBABILISTIC DAILY RAINFALL MODELLING USING GLMs
- MODELLING APPROACH

CONTENTS OF THE LECTURE and PRACTICAL

- GLIMCLIM RATIONALE
- PROBABILISTIC DAILY RAINFALL MODELLING USING GLMs
- MODELLING APPROACH
- OVERVIEW OF THE GLIMCLIM SOFTWARE

CONTENTS OF THE LECTURE and PRACTICAL

- GLIMCLIM RATIONALE
- PROBABILISTIC DAILY RAINFALL MODELLING USING GLMs
- MODELLING APPROACH
- OVERVIEW OF THE GLIMCLIM SOFTWARE
- EXAMPLE AND PRACTICE WITH GLIMCLIM (this afternoon)

GLIMCLIM RATIONALE

- Long time series not available
- Inputs for impact studies and impact models (e.g. hydrological models, crop models): continuous in time, at fine spatial and temporal scale.
- Poor performance of general circulation models (GCMs) in simulating secondary variables (e.g. precipitation)
- Poor performance of GCMs at local and regional scales (below 200km)

GLIMCLIM

- Open source software for research purposes by Dr. Richard E. Chandler (UCL, London)
- Downloadable at the page: `http://www.homepages.ucl.ac.uk/~ucakarc/work/glimclim.html`
- It incorporates the theory of generalised linear models (GLMs)
- Originally written to model and simulate daily rainfall series
- Fits logistic and gamma regression models to time series
- Simulates sequences using the fitted models

STOCHASTIC MODELLING OF RAINFALL

- Investigate the relationships among rainfall and other components of the climate system
- Generate synthetic sequences conditional on those factors at point location or area averages for impact studies
- Simulate multiple precipitation sequences consistent with any given set of atmospheric drivers
- Generate future projections for impact assessments downscaling all available large-scale GCMs outputs

MODELLING RELATIONSHIPS AMONG COMPONENTS OF THE CLIMATE SYSTEM

One response variable Y of n values $y_1, \ldots, y_n$ and associated vectors $x_1, \ldots, x_n$ each containing p explanatory variables (covariates)

MODELLING RELATIONSHIPS AMONG COMPONENTS OF THE CLIMATE SYSTEM

One response variable Y of n values $y_1, ..., y_n$ and associated vectors $x_1, ..., x_n$ each containing p explanatory variables (covariates)

$$y_i = \beta_0 + \beta_1 x_{i1} + ... + \beta_p x_{ip} + \epsilon_i$$

with $\epsilon i \; N \; (0; \, \sigma^2)$ independently for each i

MODELLING RELATIONSHIPS AMONG COMPONENTS OF THE CLIMATE SYSTEM

One response variable Y of n values $y_1, ..., y_n$ and associated vectors $x_1, ..., x_n$ each containing p explanatory variables (covariates)

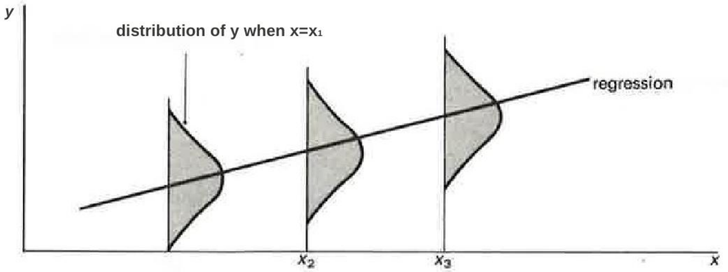
$$y_i = \beta_0 + \beta_1 x_{i1} + ... + \beta_p x_{ip} + \epsilon_i$$

with $\epsilon i \; N \; (0; \sigma^2)$ independently for each i

The regression line describes how the mean response variable (μ_y) changes with the explanatory variables.

The observed values of the dependent variable changes around their μ.

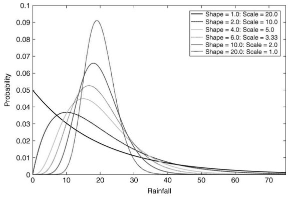
MODELLING RELATIONSHIPS AMONG COMPONENTS OF THE CLIMATE SYSTEM



(Chatfield, 1983)

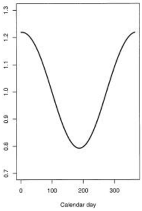
- Each response comes from its own (normal) probability distribution;
- The relationship among the dependent and independent variable(s) is linear.

MODELLING RELATIONSHIPS AMONG RAINFALL AND OTHER COMPONENTS OF THE CLIMATE SYSTEM



(Husak et al., 2006)

- Daily rainfall is better approximated with a γ distribution function.



(Yan et al., 2002)

- The relationship among dependent and independent variables may not be linear.

GENERALIZED LINEAR MODELS

Y_i is assumed to be generated from the same family of distribution with mean μ_i,

$$g(\mu_i) = x_i\beta = \eta_i$$

The GLM is composed of the three elements:

- A choice of distribution with mean μ_i.
- A linear predictor $\eta = X\beta$, a linear combination of unknown parameters β.
- A link function $g(.)$ such that $E(Y) = \mu = g^{-1}(\eta)$.

GENERALIZED LINEAR MODELS FOR RAINFALL

Rainfall occurrence and amounts modelled separately.

Logistic regression for Rainfall occurrence (pattern of dry/wet days):

$$\ln \frac{p_i}{1 - p_i} = \mathbf{x_i}^{'} \beta$$

where p_i is the probability of rain for the i^{th} case in the data set conditional on a covariate vector $\mathbf{x}_i$ with coefficient vector β.

Gamma distribution for Rainfall amounts during wet days.

The rainfall amount for i^{th} wet day has, conditional on a covariate vector ξ_i and coefficient vector γ, a gamma distribution with mean μ_i, where:

$$\ln(\mu_i) = \xi_i^{'} \gamma$$

CHOICE OF PREDICTORS

- account for a significant proportion of the local climate variance
- physically meaningful in explaining the variability
- their relationship with the predictand should be time-invariant
- should not be strongly correlated to each others
- represented by GCMs with relative good skills
- long and reliable records
- able to capture the climate change signal

COVARIATES CATEGORIES FOR CLIMATOLOGICAL APPLICATIONS

- Seasonal effects – 1-year, half-year seasonal cycle
- Geographical effects – Latitude, Longitude, Altitude
- Autocorrelation effects – include ‘autoregressive’ terms (i.e. functions of previous values)
- Temporal effects – simple trend functions, climate indices e.g. NAO, ENSO

INTERACTIONS

- Some predictors may modulate effect of others – incorporated via interactions. e.g. El-Niño or NAO effect may be regional and seasonal dependent (alternative to, e.g., tting separate models in each month of year)
- Simple model: suppose linear predictor is

$$\eta_i = \beta_0 + \beta_1 x_{1i} + \beta_{2i} x_{2i},$$

$$\text{with} \quad \beta_{2i} = \gamma_0 + \gamma_1 x_{1i}$$

therefore, replacing in the first equation:

$$\eta_i = \beta_0 + \beta_1 x_{1i} + (\gamma_0 + \gamma_1 x_{1i}) x_{2i} = \beta_0 + \beta_1 x_{1i} + \gamma_0 x_{2i} + \gamma_1 x_{1i} x_{2i}$$

- Final term represents interaction – easily incorporated into GLM framework

MODELLING STRATEGY

- Progressive addition of terms.

MODELLING STRATEGY

- Progressive addition of terms.
- Estimation of model parameters β with maximum likelihood method.

MODELLING STRATEGY

- Progressive addition of terms.
- Estimation of model parameters β with maximum likelihood method.
- Selection of models with likelihood ratio test, adjusted for inter-site dependence.

MODELLING STRATEGY

BASELINE MODEL $\rightarrow$ Climatology
- regional effects
- seasonal effects
- autocorrelation effects

$\Downarrow$ plus time varying climate covariates (external covariates)

FINAL MODEL $\rightarrow$ Time-varying variability
- external climate covariates

INTER-SITE and INTRA-SITE DEPENDENCE

- Likelihood theory assumes responses are independent given covariates.
- Autoregressive terms ensure no temporal dependence; however, no easy solution for spatial data.
- Possible approaches for spatial dependence:
  - Model spatial dependence explicitly
  - Fit models as though sites are independent and make retrospective adjustments

  Latter approach adopted for computational reasons.

MODEL DIAGNOSTICS

- Need to check:
  - Representation of systematic structure
  - Distributional assumptions (normal, Poisson, gamma etc.)
- Checks based on residuals:
  - Pearson residuals (for unexplained systematic structure):

    $$r_i = \frac{(Y_i - \mu_i)}{\sigma_i}$$

    where Y_i is the observed response for the i^{th} case, μ_i the modelled mean and σ_i the modelled standard deviation. If model is correct, expected value is 0 and variance is 1 – check by computing mean and variance over subgroups of observations.
  - Anscombe residuals (for probability assumption): defined so as to be approximately normally distributed (for a gamma distribution) if model is correct (check quantile-quantile plots)

MODELLING RECOMMENDATION

- The model represents associations in the climate systems: those should be as much as possible physically-based
- Data issues: avoid over-detailed interpretation if there are quality issues
- Prefer parsimonious models: careful addition of terms to avoid overfitting
- Inspect Pearson residuals after every stage

SIMULATION

- Occurrence and amounts models for monthly rainfall can be used jointly to simulate sequences of rainfall.

- Since the model is stochastic, multiple realisations will provide an envelope of simulation to represent uncertainty.

It allows to compare summary statistics from observed data with distribution obtained by simulation (best performance check)

Select and independent period of time for validation purposes...

GLIMCLIM for DOWNSCALING

- For downscaling, need to relate fine-scale rainfall to coarse-scale GCM outputs
- Use coarse-scale atmospheric variables as covariates in GLMs
- Choose atmospheric variables that (a) have demonstrable relationship with fine-scale rainfall (b) are better represented in GCMs
- Fit models using historical data (rainfall and atmospheric variables) to describe relationships
- For future scenarios, simulate from fitted models driven by GCM-generated atmospheric variables

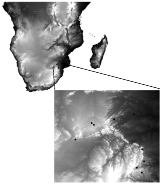
in conclusion the use of GLMs/GLIMCLIM...

- offers a clear approach;
- is flexible environment and do not require the predictand variable to be normally distributed;
- allows multiple climate factors to be considered simultaneously and to represent the climate system using one model;
- discriminates among competing predictors and isolate key drivers;
- is relatively computational inexpensive (can run multiple simulations quickly).

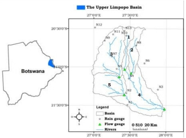
EXAMPLE of APPLICATIONS

Daily rainfall modelling and simulation in:

- Sekhukhune District in South Africa for agricultural applications



- Northern Uganda and
- Limpopo Basin in Botswana for hydrological applications



- ...and many more (U.K., Peru,...)

MORE TO COME – RGlimclim

- multivariate extension allowing sequential modelling of multiple variables (e.g. daily local temperature dependent upon daily pressure)
- R-package working on both Windows and Linux operating systems
- built-in command for model comparison allowing selection among nested models
- built-in command for residuals plotting (e.g. averaged over months, years or sites)

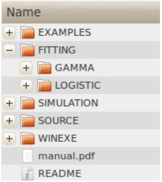
REFERENCES

- Chandler, R. E., 2002: GLIMCLIM: Generalized linear modelling for daily climate time series - User Guide, Department of Statistical Science, University College London. http://www.ucl.ac.uk/~ucakarc/work/rain_glm.html
- Ambrosino, C., Chandler, R. E. and Todd, M. C., 2011: Southern African Monthly Rainfall Variability: An Analysis Based on Generalized Linear Models. J. Climate, 24, 4600–4617.
- Chandler, R. E. and Wheater, H. S., 2002: Analysis of rainfall variability using generalised linear models: A case study from the west of Ireland. Water Res. Res. 38, doi:10.1029/2001/WR000906.
- Kenabatho, P. K., McIntyre, N. R., Chandler, R. E. and Wheater, H. S., 2012: Stochastic simulation of rainfall in the semi-arid Limpopo basin, Botswana. Int. J. Climat., 32, 1113–1127

GLIMCLIM

GLIMCLIM – Software

- Zip archive of the latest version on the Internet (`rain_glm.zip`)
  - EXAMPLES: Contains sample definition files and an artificial data file
  - FITTING: Contains 2 sub-folders: LOGISTIC/ containing blank definition files necessary to fit a logistic regression model for daily rainfall occurrence; GAMMA/ containing files necessary to fit gamma distributions to rainfall amounts on wet days
  - SIMULATION: Contains blank definition files required for simulation of fitted models
  - SOURCE: Contains source code (FORTRAN 77) which can be used to generate executables for fitting and simulation; some simple Unix scripts to compile the code on most Unix machines, and move the resulting executable to the appropriate place.
  - WINEXE: Pre-compiled executables for use with a Windows system (to be run from an MS-DOS prompt)



USE OF THE PACKAGE

`fit_logi`, `fit_gamm` and `simrain` are the 3 executables used to: fit logistic and gamma regression models and simulate a joint occurrence/amounts model

Run them from within a command window changing to the directory containing the executables

The possible covariates are divided into categories, for which a number of parametrisation choices are offered by the package

Each programme needs a number of input files or it will terminate with an error message: suffixed `.def` are definition files, `.dat` data files

INPUT FILES

PROGRAMME	PURPOSE	REQUIRED INPUT FILES
fit_logi	Fit logistic regression model for zero/non-zero values	siteinfo.def gaugvals.dat logistic.def
fit_gamm	Fit gamma distributions to positive amounts	siteinfo.def gaugvals.dat gammamdl.def
simrain	Simulate daily values using combined occurrence and amounts models	regions.def siteinfo.def gaugvals.dat logistic.def gammamdl.def

EXTERNAL COVARIATES: dy_preds.dat, mn_preds.dat and yr_preds.dat

INTER-SITE CORRELATIONS: cor_logi.dat and cor_gamm.dat

SETTING UP GLIMCLIM and DEFINITION FILES

- Install GLIMCLIM:
    - Linux: open `fit_logi.f` and `fit_gamm.f` and change BYTELN value to 8; type chmod 755 $*$; run `fit_logi_compile`, `fit_gamm_compile` and `simrain_compile`
    - Windows: move the executables found in WINEXE folder to the appropriate folders. The programme should be run from a command line (type `cmd` from start menu)
- Inspect example files: `gaugvals.dat`, `siteinfo.def`, `logistic.def`
- `dy_preds.dat`, `mn_preds.dat` and `yr_preds.dat` for ‘external’ covariates

Component	Code 1	Code 2	Code 3
0 (constant)	No codes used		
1 (site effects)	Number of attribute, according to order of definition in file siteinfo.def	If present, label of nonlinear transformation (see Table 4)	Not used
2 (year effects)	**Up to 50:** Label of trend function (see Table 4) **51 upwards:** $(x-50)$th variable defined in file yr_preds.dat (x being the code entered).	Optional selection of lagged values of external covariate (if Code 1 > 50).	Not used.
3 (month effects)	**1:** $\cos(2\pi \times \text{month}/12)$ **2:** $\sin(2\pi \times \text{month}/12)$ **3:** $\cos(2\pi \times \text{month}/6)$ **4:** $\sin(2\pi \times \text{month}/6)$ **11–22:** Individual month indicators (11 = Jan, 12 = Feb etc.) **51 upwards:** $(x-50)$th variable defined in file mn_preds.dat.	E.g. to use covariate value 2 years/months ago, set this field to 2. To use next year's/month's value set to -1.	Not used
4 (day effects)	**See Table 3, page 17**		
5 (2-way interactions)	Indices of interacting main effects (first site effect is 1)		Not used
6 (3-way interactions)	Indices of interacting main effects (first site effect is 1)		
7 (parameters in nonlinear transformations)	Index of covariate for which transformation is being defined. See note 3, page 17.	Parameter being defined (1, 2 or 3 — see Tables 4 and 6)	**0:** treat parameter as known **1:** find ML estimate of parameter
8 (global quantities)	**1:** Threshold for defining 'small' positive values. See note 4, page 17.	Method for dealing with such values (See Table 7)	Not used
9 (dispersion parameter)	No codes required. NB logistic models have no dispersion parameter. This field is ignored by model fitting programs.		
10 (spatial structure)	Label of spatial dependence structure used (see Table 8)	Number of parameter (see Table 8)	Not used

Table 2: Codes for specification of models in files logistic.def and gammamdl.def. To be used in conjunction with Tables 3–8.

Code 1	Code 2	Code 3
1–10: value x days ago **21**: $\cos(2\pi \times \text{day of year}/366)$ (see note 1, page 15) **22**: $\sin(2\pi \times \text{day of year}/366)$ **23**: $\cos(2\pi \times \text{day of year}/183)$ **24**: $\sin(2\pi \times \text{day of year}/183)$ **31–42**: Smooth month adjustments (31 = Jan, 32 = Feb etc.). See note 2 on page 17. **51 upwards**: $(x-50)$th variable defined in file dy_preds.dat	Optional. If present and Code 1 $\leq$ 10, selects a transformation of previous days' values (see Tables 5 and 6). If Code 1 > 50, selects lagged covariate values as in rows 2 and 3 of Table 2.	If present and equal to k, and Code 1 $\leq$ 10, cases with missing values at the same site for any of the previous k days are discarded by the fitting programs. If two records contain different values here, the highest is taken. The maximum allowable value is 10.

Table 3: Codes for specifying 'daily' effects in files logistic.def and gammamdl.def. This is row 4 of Table 2. See Tables 5 and 6 for further details on transformations and averaging.

Component	Label	Function	Parameter 1	Parameter 2
1 (site effects)	1	Box-Cox power transform: $f(x) = \begin{cases} \ln x & \lambda = 0 \\ \frac{x^\lambda - 1}{\lambda} & \text{otherwise} \end{cases}$	λ	Not used
	2	Exponential transform: $f(x) = e^{ax}$	a	Not used
	3	Arctan transform: $f(x) = \arctan\left(\frac{x-a}{b}\right)$	a	b
	11–30	Fourier series representation of effect over the range (a, b). 11 and 12 are sine and cosine terms at the first Fourier frequency, 13 and 14 at second etc. *Odd* numbers correspond to *sine* terms (i.e. *odd* part of function). a and b can be specified *once only* for each site attribute. All sites must lie within the range $[a, b]$. Both a and b must always be treated as known.	a	b
	31–40	Legendre polynomial representation of effect over the range (a, b). 31 is linear, 32 is quadratic etc. a and b can be specified *once only* for each site attribute. All sites must lie within the range $[a, b]$. Both a and b must always be treated as known.	a	b
2 (year effects)	1	Linear: $f(t) = (t - 1950)/10$ (t is year)	No parameters required	
	2	Piecewise linear: $f(t) = \begin{cases} (t-a)/10 & \text{if } t > a \\ 0 & \text{otherwise} \end{cases}$	a	Not used
	3	Cyclical: $f(t) = -\cos\left(2\pi\frac{t-b}{a}\right)$	a	b

Table 4: Labels for nonlinear transformations of covariates (excluding previous days' values) in files logistic.def and gammamdl.def. This table should be used in conjunction with Table 2.

Label	Transformation
1	$\ln\left(Y_{t-k}^{(s)}\right)$
2	$\ln\left(1+Y_{t-k}^{(s)}\right)$
3	$I\left(Y_{t-k}^{(s)}>0\right)$ (i.e. indicator taking the value 1 if $Y_{t-k}^{(s)}$ was non-zero, 0 otherwise).
4	$I\left(0<Y_{t-k}^{(s)}<\tau\right)$, where τ is a 'trace' threshold (defined in row 8 of Table 2).
5	'Persistence' indicator: 1 if $Y_{t-1}^{(s)},\ldots,Y_{t-k}^{(s)}$ were all >0, 0 otherwise.
10–15	Transformations as above, but averaged over all sites with available data. Covariate is $S^{-1}\sum_r f\left(Y_{t-k}^{(r)}\right)$, where S is the number of contributing sites and $f(.)$ is the transformation. Code 10 is an average of untransformed values: $S^{-1}\sum_r Y_{t-k}^{(r)}$.
20–25	Transformations as above, but averaged over all sites with available data using weights that decrease exponentially with distance from the current site s. Covariate is $\sum_r w_{r,s} f\left(Y_{t-k}^{(r)}\right)$, where the weights $\{w_{r,s}\}$ sum to 1 and are proportional to $\exp\left[-ad_{r,s}\right]$. The value of a must be specified in the 'nonlinear parameters' section of the definition file — see Table 6.
30–35	Weighted averages of transformed values; weights proportional to $$\exp\left\{-a\left[(u_r-u_s-ku_0)^2+(v_r-v_s-kv_0)^2\right]^{1/2}\right\}.$$ See note 5, page 20 for an interpretation of this scheme.The values of a, u_0 and v_0 must be specified in the 'nonlinear parameters' section of the definition file — see Table 6.
110–115, 120–125 and 130–135	As 11–15, 21–25 and 31–35 but with the order of transformation and averaging reversed. Covariates are $f\left(\sum_r w_{r,s} Y_{t-k}^{(r)}\right)$ i.e. transformations of averages rather than averages of transformations.

Table 5: Labels for specifying nonlinear transformations of previous days' values, in files logistic.def and gammamdl.def. $Y_t^{(s)}$ denotes the value at site s on day t; u_s and v_s are the geographical co-ordinates of site s (in terms of the **first two** site attributes defined in file siteinfo.def); and d_{s_1,s_2} is the distance between sites s_1 and s_2, again calculated from these two site attributes. The expressions in the table relate to prediction of the value at site s on day t; k is the lag (in days), defined as described in Table 3.

Weighting scheme	Parameters		
	1	2	3
1: Equal weights at all sites	No parameters required		
2: Distance-based exponential decay: $w_{r,s} \propto \exp\left[-ad_{r,s}\right]$	a	Not required	
3: Distance-based exponential decay with shift in origin: $w_{r,s} \propto \exp\left\{-a\left[(u_r - u_s - ku_0)^2 + (v_r - v_s - kv_0)^2\right]^{1/2}\right\}$	u_0	v_0	a

Table 6: Parameters in schemes for computing weighted averages of previous days' values. This table should be used in conjunction with Tables 3 and 5. $w_{r,s}$ is the weight associated with site r when predicting for site s. All other notation is the same as for Table 5.

Value of Code 2 (Table 2, row 8)	Treatment of 'small' values
1	Treat values as 'trace amounts'. This option is designed with rainfall in mind. Any 'small' value will be regarded as non-zero (and hence will count as a 'wet' day in a logistic regression model for rainfall, for example), but will be treated as a left-censored observation in any models for non-zero amounts.
2	'Soft' thresholding. If the original variable of interest is Y and the threshold is τ then models are fitted to Y^*, where $Y^* = 0$ if $Y < \tau$, $Y - \tau$ otherwise.
3	'Hard' thresholding. If the original variable of interest is Y and the threshold is τ then models are fitted to Y^*, where $Y^* = 0$ if $Y < \tau$, Y otherwise.

Table 7: Methods for dealing with 'small' positive values. The threshold below which a value is regarded as 'small' is defined as a 'global' quantity in the main model definition file (see row 8 of Table 2).

Table 8: Labels for specifying spatial structures in files logistic.def and gammamdl.def. For the occurrence model, the correlations in structures 1 through 20 are between latent Gaussian variables. For the amounts model, correlations are between Anscombe residuals at each site. For the correlation-based structures, d_{ij} denotes the Euclidean distance between sites i and j in terms of the first two site attributes defined in file siteinfo.def. This table should be used in conjunction with Table 2.

Label	Model description	Parameter 1	Parameter 2	Parameter 3	Parameter 4
0 (default)	Independence	No parameters required			
1	Empirical correlations between each pair of sites	No parameters required (correlations read from file cor_logi.dat or cor_gamm.dat)			
2	Constant correlation between each pair of sites: $\rho(i,j) = \rho$	ρ	Not used		
3	Exponential correlation function: $\rho(i,j) = \exp\left[-\phi d_{ij}\right]$	ϕ	Not used		
4	Correlation function $\rho(i,j) = \alpha + (1-\alpha)\exp\left[-\phi d_{ij}\right]$	ϕ	α	Not used	
5	Powered exponential correlation function: $\rho(i,j) = \exp\left[-\phi d_{ij}^{\kappa}\right]$	ϕ	κ	Not used	
6	Correlation function $\rho(i,j) = \alpha + (1-\alpha)\exp\left[-\phi d_{ij}^{\kappa}\right]$	ϕ	κ	α	—
21	Conditional independence given weather state X, which is 0 for a 'dry' day and 1 for a 'wet' day. $P(X=1) = 1 - P(X=0) = \alpha$, the mean of the site probabilities predicted by the occurrence model. When $X = x$, the log odds for a non-zero value at site i is $\ln\left(\frac{p_i}{1-p_i}\right) + x\ln a - \ln b(\alpha, p_i)$, where p_i is the probability of rain at site i according to a logistic regression model, and $b(\alpha, p_i)$ is chosen to make the unconditional probability at the site equal to p_i. This structure is valid for the logistic model only.	$\ln a$	Not used		
22	Dependence induced by specifying a Beta-Binomial distribution for the number of wet sites on any day. The mean of the distribution, θ, is fixed at the mean of the individual site probabilities, and the shape parameter ϕ is estimated from data. This structure is valid for the occurrence model only.	ϕ	Parameters 2 and 3 are optional, and control program behaviour for days when the specified Beta-Binomial distribution is incompatible with the probabilities of rain at the sites. When this occurs, the probabilities are shrunk towards θ. p_i becomes $p_i - \lambda(p_i - \theta)$, for $\lambda \in (0,1)$. The default value of λ is 0.01: to change this, enter it as parameter 2 for this model. By default, an error message will be printed to screen whenever such shrinking occurs. To suppress this, set parameter 3 equal to zero.		