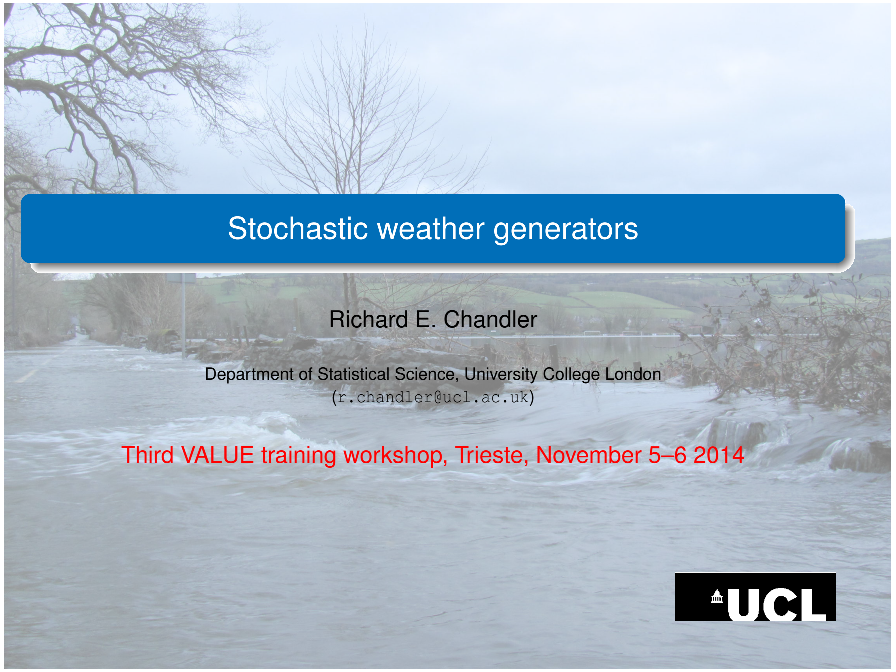


Stochastic weather generators

Richard E. Chandler

Department of Statistical Science, University College London
(r.chandler@ucl.ac.uk)

Third VALUE training workshop, Trieste, November 5–6 2014

Motivating example

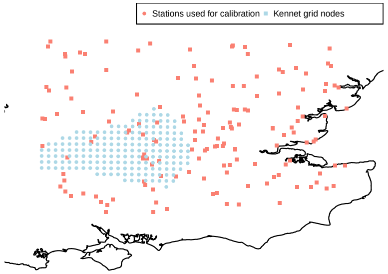
- HydEF project (http://www.bgs.ac.uk/changingwatercycle/hydef.html) looking at hydro(geo)logical impacts of climate change in UK
- Detailed hydro(geo)logical models require high-resolution weather inputs, consistent with changing large-scale synoptic conditions as obtained e.g. from reanalysis products or GCMs

E.g. variables needed by JULES:

Rainfall rate	*Air pressure*	*Snowfall rate*	*Air temperature*
Wind speed	*Specific humidity*	*Downward short-wave radiation*	*Downward long-wave radiation*

Case study: the Thames

- Largest catchment in UK ($\sim$ 10000km^2)
- Modellers wanted hourly sequences, 8 variables, 1km^2 resolution throughout catchment



- Negotiated settlement: daily sequences, 5×5km^2 resolution, Kennet subcatchment (186 grid nodes)
- Data on (most) variables nominally available from 157 stations, 1970 onwards

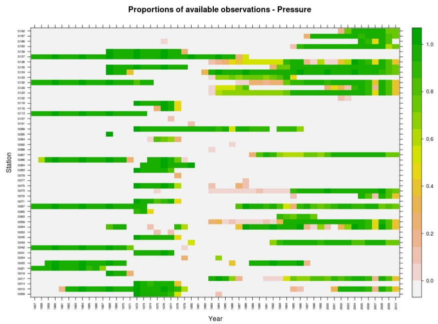
---

Data availability (I)

- Hourly data obtained from British Atmospheric Data Centre (BADC), MIDAS Met Office dataset
- Available variables: rainfall, snow, air pressure, air temperature, wind speed, downward SW radiation
- Missing variables: specific humidity and downward LW radiation
  - Can be derived from other variables using standard procedures from literature
- **BUT** ...

Data availability (II)

Numbers of stations with data (out of 157)				
Rainfall	Pressure	Temperature	Wind speed	SWR
71	52	140	135	22



- Many stations have short / incomplete/ patchy records

Motivating example: summary of requirements

- Need to generate daily time series for ...
- Several variables simultaneously, with different distributions and preserving inter-variable relationships ...
- at many locations simultaneously, preserving inter-site relationships ...
- ... including locations for which no observations are available ...
- ... and substantial amounts of missing data at locations where observations *are* available ...
- including a realistic climate change signal.

Structure of session

Part 1

- Weather generators: what and why?
- Weather generators vs RCMs
- ‘Classical’ generators
- Other types of weather generator
- Incorporating climate change information

Part 2

- Issues in multisite generation
- Classes of multisite generator
- Data requirements
- Software packages available, including `Rglimclim`
- The Thames revisited

Part 1: Introduction to weather generators

What is a weather generator?

IPCC summary
(www.ipcc-data.org/guidelines/pages/weather_generators.html)

A stochastic weather generator (WG) produces synthetic time series of weather data of unlimited length for a location based on the statistical characteristics of observed weather at that location.

(remainder of IPCC summary strictly correct but potentially misleading — and no mention of multi-site generation)

- Additional requirement here: ability to capture climate change signal using information from GCMs
  - **NB** tacit assumption that GCMs do not provide useful information at resolution required by users (classic example: Abourgila 1992)
  - 'Perfect Prognosis' approach to downscaling: GCM outputs taken as correct (possibly after processing)

Why time series?

- Interest in assessing response of complex systems to climate change
- System response depends on how weather effects are aggregated:
  - UK flooding, Boscastle, August 2004: localised intense rainfall in one day (Met Office, 2005)
  - UK flooding, winter 2000–2001: two-month rainfall totals exceeding 200-year return period (Finch et al., 2004)
  - European heatwave, 2003: excess deaths associated with extended periods of extreme heat without night-time cooling (http://en.wikipedia.org/wiki/2003_European_heat_wave)
  - Crop growth sensitive to quantity and timing of precipitation (Kniveton et al., 2009)
  - etc. etc.

Boscastle, August 16th 2003

Generic requirements

- Aim to reproduce some subset of time series features ("aspects" in VALUE vocabulary)
- Subset depends on context

Examples:

Marginal aspects : mean, variance, frequency of threshold exceedances, return levels, . . .

Temporal aspects : trends, seasonality, autocorrelation, spell lengths, . . .

Spatial aspects : systematic regional variation, residual inter-site dependence, simultaneous threshold exceedances, . . .

Inter-variable relationships : correlations, frequency of joint events, . . .

Weather generators versus RCMs

WGs	RCMs
Empirically based	Physically based
Stochastic in nature	Deterministic in nature
Cheap to simulate	Expensive to simulate
Require calibration (fitting) on case-by-case basis	No user calibration required
Can choose method and tune to meet application requirements	Limited options for application-specific tuning
Rely on empirical relationships persisting into future	Rely on laws of physics persisting into future

The 'classical' weather generator

- First 'weather generator': WGEN based on Richardson (1981) for daily weather sequences
- Built on earlier models for daily precipitation going back to Gabriel and Neumann (1962)
- Model precipitation first, then other variables conditional on precipitation — because precipitation has challenging statistical properties
- Markov chain for precipitation occurrence, gamma distribution for intensity, separate parameters for each month of the year
- (Some) other variables conditioned on precipitation status e.g. separate distributions for wet and dry days, cosine functions fitted to parameters for seasonality

Markov models for precipitation

The basic Markov precipitation model

- Let $Y_t = 1$ if day t is 'wet', 0 otherwise
- **Markov assumption**:
  $P(Y_t = y | Y_{t-1}, Y_{t-2}, \ldots) = P(Y_t = y | Y_{t-1})$ for $y = 0, 1$
- Leads to 2-state Markov chain for precipitation occurrence
- Characterised by transition probabilities:
  $\pi_{11} = P(Y_t = 1 | Y_{t-1} = 1)$, $\pi_{01} = P(Y_t = 1 | Y_{t-1} = 0)$
- Wet-day intensities assumed independent and to follow some distribution (exponential, gamma, ...)

Properties of Markov chains

- Temporal dependence characterised via transition probabilities:
  - If $\pi_{11} \simeq 1$ then one wet day will very likely follow another
  - If $\pi_{01} \simeq 0$ then one dry day will very likely follow another
  - etc.
- 2-state Markov chain has equilibrium distribution: long-run proportion of wet days is

$$P(Y_t = 1) = \frac{\pi_{01}}{1 + \pi_{01} - \pi_{11}}.$$

- So transition probabilities also characterise marginal aspects of precipitation occurrence
- Higher-order chains give more flexibility e.g. specifying $P(Y_t = 1 | Y_{t-1} = y_1, Y_{t-2} = y_2)$.
- See **Exercise 1**.

Deficiencies of basic WG

From IPCC guidelines:

One criticism of the Richardson-type WG is its failure to describe adequately the length of dry and wet series (i.e. persistent events such as drought and prolonged rainfall). These can be very important in some applications (e.g. agricultural impacts).

Other common problems:

- Tendency to underestimate variability of seasonal means / totals ("overdispersion" — see, e.g., Katz and Parlange 1998).
- Underestimation of high return levels e.g. 100-year daily maxima (independent exponential / gamma intensity distributions do not yield 'heavy tailed' extreme distributions observed in daily rainfall data e.g. Katz et al. 2002)

Approaches to remedying deficiencies in basic WG structure

Many suggestions in literature:

- Higher-order Markov chains to improve wet and dry spell performance
- Heavy-tailed intensity distributions to improve extremal behaviour
- Introduce latent classes with separate parameter sets, to increase variability in seasonal means
- Nonparametric modelling to avoid specific distributional assumptions
- Etc. etc.

An elephant in the room?

What about correlation between successive days' precipitation intensities?

Other approaches to weather generation

- Approaches based on spell lengths
- Resampling methods
- Generalised linear models
- Subdaily weather generators

See also:

VALUE inventory and review of statistical downscaling methods — summary at `http://convection.zmaw.de/fileadmin/user_upload/convection/Convection/WG_Presentations/2014.01.29-30/SDS_COST_Inventory_AFischer.pdf`

Generators based on spell lengths

- **Idea**: resolve problems with spell-length distributions by placing these at heart of generator:
  - Start by generating wet and dry spell lengths
  - Then proceed similarly to 'classic' generator
- Approach common in agricultural applications where spell lengths are important
- LARS-WG is best-known example:
  - Uses 'semi-empirical' spell length distributions fitted separately for each month
- See Semenov et al. (1998) for summary and comparison with WGEN

Resampling methods

- **Idea**: for each day of simulation, choose values at random from observations on days 'similar to' current day
- 'Similarity' could be, e.g.:
  - All values on same day of year (seasonality)
  - Values from days with similar previous days' weather (autocorrelation)
  - Values from days with similar large-scale synoptic conditions
- Nonparametric approach makes minimal assumptions
- Inter-variable dependencies automatically preserved
- Cannot generate values outside range of those previously observed
- Cannot consider too many factors in determining similarity ('curse of dimensionality')
- More details: Buishand and Brandsma (2001)

Generalised linear models (GLMs)

- **Idea**: embed 'classical' generator within wider class of models
- Grunwald and Jones (2000) showed that Markov-based models are special cases of Generalised Linear Models (GLMs)
  - GLMs first applied to daily rainfall by Coe and Stern (1982).
  - Cornerstone of modern statistical practice in all application areas

GLM for precipitation occurrence

- Common to use logistic regression model:

$$\ln\left(\frac{p_t}{1-p_t}\right) = \mathbf{x}_t'\beta = \eta_t^{\text{occ}}, \text{ say} \quad \Rightarrow \quad p_t = [1+\exp(-\eta_t)]^{-1}$$

where:
  - p_t is probability of precipitation on day t
  - $\mathbf{x}_t$ is vector of covariates (predictors)
  - β is coefficient vector
- E.g. set $\mathbf{x}_t = (1 \; Y_{t-1})'$, $\beta = (\beta_0 \; \beta_1)'$, then $p_t = \left[1+e^{-(\beta_0+\beta_1 Y_{t-1})}\right]^{-1}$
  - When $Y_{t-1} = 0$, $p_t = \left[1+e^{-\beta_0}\right]^{-1}$ — this is π_{01} in Markov formulation
  - When $Y_{t-1} = 1$, $p_t = \left[1+e^{-(\beta_0+\beta_1)}\right]^{-1}$ this is π_{11}

GLMs for other variables

- Generic formulation of arbitrary (now generic) variable Y_t
  - $\{Y_t\}$ considered drawn from common family of distributions (normal, gamma, Poisson, Bernoulli, ...)
  - Conditional on covariate vector $\mathbf{x}_t$, expected value of Y_t is $\mu_t = \mathbb{E}(Y_t|\mathbf{x}_t)$
  - μ_t related to linear predictor $\eta_t = \mathbf{x}_t\beta$ via relationship $g(\mu_t) = \eta_t$ for link function $g(\cdot)$.
- Extends linear regression model (normal distributions, $g(\mu_t) = \mu_t$).
- E.g. use gamma GLM with log link for precipitation intensity
  - Can model temporal dependence in intensity by including Y_{t-1} in $\mathbf{x}_t$ — maybe resolve overdispersion problem?
- Unified approach for all variables — differences only in choice of distribution
- Coefficients estimated using maximum likelihood
- Assumptions can be checked

Interactions

- With two covariates x_{1t}, x_2, suppose $\eta_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t}$.
- Suppose also that x_2 modulates effect of x_1: $\beta_1 = \gamma_0 + \gamma_1 x_{2t}$. Then

$$\begin{aligned} \eta_t &= \beta_0 + (\gamma_0 + \gamma_1 x_{2t})x_{1t} + \beta_2 x_{2t} \\ &= \beta_0 + \gamma_0 x_{1t} + \beta_2 x_{2t} + \gamma_1 x_{1t} x_{2t}. \end{aligned}$$

- Easily handled in usual framework: just define extra covariate $x_{1t}x_{2t}$.
- Higher-order interactions can be handled similarly.

Consequence for weather generators

- Seasonal variation in parameters is just an interaction between seasonal and other covariates
- Eliminates need for separate fitting to different months / seasons

Subdaily generators

- Subdaily precipitation structure too complex for many of previous model types
- Subdaily models attempt to represent underlying mechanisms in more or less simplified form
- Two broad classes of subdaily precipitation generator:
  - **Poisson cluster models** — represent precipitation as superposition of 'cells' clustered within 'storms'
  - **Multiscaling models** — exploit systematic variation of precipitation summary statistics with temporal resolution

  Up-to-date review in Chandler et al. (2014).
- Limited work on subdaily generation for other variables
- Subdaily generation not considered further here.

Incorporating climate change information

- Weather generation in climate change context requires ability to connect WG parameters / outputs with large-scale atmospheric structure
- Various heuristic schemes e.g. additive / multiplicative change factors based directly on GCM changes in variables of interest
  - Inappropriate for (e.g.) precipitation because change factors do not affect wet / dry properties
  - Some more considered applications apply change factors to relevant model parameters e.g. (Kilsby et al., 2007) — approach used in UKCP09 national climate projections for UK (http://ukclimateprojections.metoffice.gov.uk/).
- More formally: integrate indices of large-scale structure formally into model specification

Predictor selection

Requirements for indices of large-scale structure

- Indices must have genuine relationship with local variable(s) of interest
  - Relationship must be robust to changes in climate
  - Relationship must capture climate change signal
- Indices must be well simulated by GCMs

- See also IPCC guidelines at www.ipcc-data.org/guidelines/dgm_no2_v1_09_2004.pdf (but **NB** review of weather generators now out-of-date)
- Requirements unverifiable(!) Pragmatic response:
  - Focus on variables and scales at which GCMs might reflect reality; and acknowledge difficulty (Smith, 2002).
  - Try to incorporate known mechanisms into WG structure

Synoptic indices

- One possibility: construct indices of large-scale structure and incorporate directly into weather generator model
- Examples of indices:
  - Teleconnection indices: ENSO, NAO, . . .
  - Means of relevant fields e.g. MSLP, temperature, . . . over relevant area
  - Principal modes of relevant fields (e.g. EOFs) — but **NB** can be hard to align modes from GCMs with those from observations
- Typically need measures of moisture availability where precipitation is concerned (Charles et al., 1999b)
- Relevant indices may vary with region and season

Incorporating synoptic indices into WG models

- **'Classical' WG models**: difficult, mostly done by parameter perturbation or weather classification (next slide)
- **Resampling methods**: incorporate indices in metric used to select candidate days for resampling
- **GLMs**: incorporate directly as additional covariates
  - Interactions account for regional / seasonal variation in effect size

Weather classification

- Alternative to direct use of indices: classify days into 'weather types' based on circulation patterns
  - **Examples**: Jenkinson-Collinson, Großwetterlagen, etc.
  - Classification may also depend on predictand(s) (see practical session)

Incorporating weather types into WGs

- **Most WG models**: fit separate parameters for each type
- **Resampling methods**: resample from days with same type
- **GLMs**: define 'dummy' 0 / 1 covariates to select type for each day
  - With G types (groups), need $G-1$ dummy covariates
  - Coefficients are deviations from remaining 'reference type'
- **NB** can be parameter-intensive if many types are used

- **Useful resource** for European applications: COST733 intercomparison project (http://cost733.met.no/)

Summary of Part 1

- Weather generators are stochastic models to produce (usually daily) time series of one or more variables
- Precipitation is fundamental due to modelling challenges
- 'Classical' structure based on Markov chain for precipitation occurrence; performs poorly with respect to spell lengths, interannual variability and extremes
- Other suggestions designed to address deficiencies directly or to make minimal assumptions about distributions etc.
- GLMs encompass 'classical' structures within flexible framework that permits many extensions to basic model structure (including ease of incorporating large-scale information)
- In climate change work, predictor selection requires care

Part 2: Multisite generators

Multisite generation

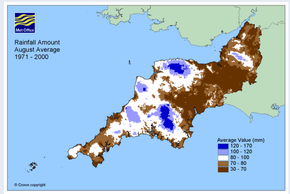
- Methods in Part 1 primarily developed for series at single site
- Some applications need simultaneous time series at multiple sites
  - E.g. hydrological studies of large catchments
  - Development of national energy infrastructure to respond to local variation in energy demand / risk of damage to generators etc.
  - Strategies for health provision or wildfire management in heatwaves
- In all examples above, spatial organisation of weather is important:
  - Do all sites experience similar weather simultaneously?
  - Or are only one or two sites affected at any one time?
- May also need to generate at ungauged sites (cf Thames example)

Additional benefit of multisite analysis:

Pooling data across sites can increase modelling precision (“space-for-time” / “borrowing strength”)

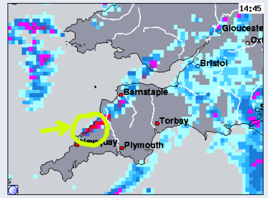
‘Spatial dependence’: a key distinction

Systematic regional variation:



From UK Met Office

Residual inter-site dependence:



From www.weatheronline.co.uk

1. Systematic regional (spatial) variation ≡ ‘climatology’
2. Residual inter-site dependence ≡ ‘spatial organisation of anomalies’

Implications of distinction

- A truly multisite weather generator must address both aspects of spatial structure
- Relatively few truly multisite WGs widely available . . .
- . . . and *very* few multisite, multivariate WGs
- **Aim here**: review most promising options that are truly multisite
  - Deliberately exclude those that do not address residual inter-site dependence
- Focus inevitably on precipitation since few multisite WGs available for other variables

Multisite extensions of classical generators

- Most multisite extensions of classical generator follow Wilks (1998)
- Fit standard generator at each location separately
  - Systematic variation captured by different parameters at each site (so cannot use directly at ungauged locations)
- Residual inter-site dependence captured by using correlated random numbers in simulations
  - Exploit ease of generating correlated Gaussian random numbers
  - **Occurrence**: use correlations for latent Gaussian variables (next slide)
  - **Intensity**: work with intensities transformed to Gaussianity, then back-transform
- Correlations estimated by matching to observed correlations
  - Occurrence: ‘trial and error’ simulation-based scheme — unsuitable for large numbers of sites

Latent Gaussian variables

Convenient way to generate correlated vector $\mathbf{Y} = (Y_1, \ldots, Y_S)'$ of binary (0/1) variables:

- Generate vector $\mathbf{Z} = (Z_1, \ldots, Z_S)'$ of correlated Gaussian variables, with $Z_s \sim N(0,1)$ for $s = 1, \ldots, S$.
- For each s, set $Y_s = 1$ if $Z_s > \lambda_s$, $Y_s = 0$ otherwise
- Choose thresholds $\lambda_1, \ldots, \lambda_S$ to obtain desired probabilities of occurrence at each site
- Choose correlations among (Z_s) to obtain desired dependence in (Y_s)
  - 'Standard' approach in WG literature: match to observed correlations
  - Easier approach: match to joint occurrence probabilities (enables direct numerical calibration, see Ambrosino et al. 2014)
- **Difficulty**: estimated correlations may not be mutually compatible
  - **Solution**: use spatial correlation model fitted to estimates

Other extensions of single-site models

- **Resampling methods**: conceptually identical to single-site case
  - Automatically reproduces distributions, dependence between sites & variables etc.
  - Cannot resample at ungauged locations
- **GLMs**: add extra covariates to represent systematic regional variation, then use e.g. correlation models for residual dependence (Chandler and Wheater, 2002; Yang et al., 2005b).
  - Extra covariates: altitude, functions of geographical coordinates etc.
  - Interactions allow regional variation of other model parameters
  - Regional covariates and correlation functions allow simulation at ungauged locations
  - Models fitted under 'working' assumption of independence, with subsequent adjustments to uncertainty assessments (see practical session)

Additional multisite class: transformed Gaussian variables

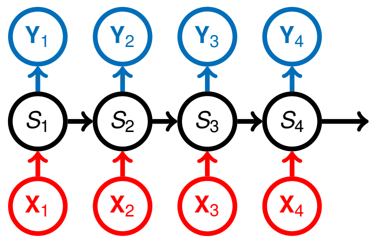
- **Idea**: let $\mathbf{X}_t$ be vector of correlated Gaussian variables on day t, and generate vector $\mathbf{Y}_t$ of precipitation values as

$$Y_{st} = \begin{cases} X_{st}^{\beta} & \text{if } X_{st} > 0 \\ 0 & \text{otherwise.} \end{cases}$$

- Similar to latent Gaussian approach for occurrence, but generates occurrence and intensity simultaneously
- Parameter β controls shape of intensity distribution
- Mean vector and covariance matrix of $\mathbf{X}_t$ simultaneously control occurrence probabilities, mean intensity and inter-site dependence.
- Key reference: Stehlík and Bárdossy (2002).
- **Caveat**: in reality, different processes control occurrence and intensity

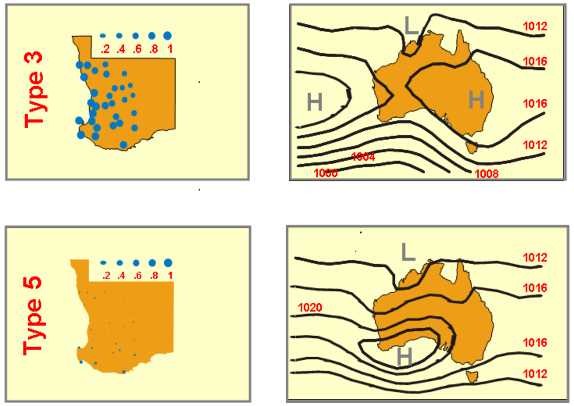
Hidden Markov Models

- **Idea** (Charles et al., 1999a): extension of weather typing
- Sequence of weather states $S_1, S_2, \ldots$ associated both with typical patterns of precipitation occurrence $\mathbf{Y}_1, \mathbf{Y}_2, \ldots$ and large-scale circulation patterns $\mathbf{X}_1, \mathbf{X}_2, \ldots$



- State sequence is Markov chain with transition probabilities determined by large-scale circulation
- Precipitation usually assumed conditionally independent given state
  - Assumption probably reasonable for large study areas with few sites
  - Assumption relaxed by Ailliot et al. (2009).

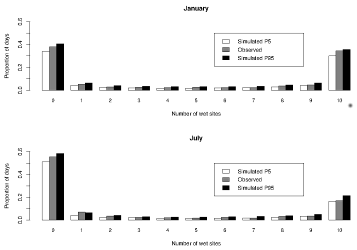
Example of HMM states and precip patterns



(Joint work with Bryson Bates and Steve Charles)

Small study areas

- Small study areas often have very high inter-site dependence
- Occurrence models based on latent Gaussian correlations can struggle to capture this . . .
- . . . but correlation is not the only measure of dependence



From Yang et al. (2005b)

- Alternative (Yang et al., 2005b): model distribution of # of wet sites
  - Beta-binomial is flexible and interpretable family of distributions for this purpose
  - Allows tendency for most sites to be either wet or dry

Data requirements for multisite weather generators

- Weather generators require calibration to observed data . . .
- . . . but some or all observations are often missing:
  - Individual observations / blocks missing from otherwise complete record
  - Different record lengths (short records have missing ends)
  - Absence of recording stations at required locations (e.g. subcatchment centres, nodes of regular grid)
- Possible solutions:
  - Work just with data available if WG calibration scheme allows it
  - Interpolation: estimate missing values (e.g. kriging, inverse distance weighting, splines etc.)

Strong recommendation:

NEVER, on any account, work with interpolated precipitation data!!!

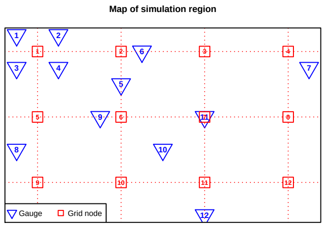
Interpolation: what's the problem?

"Interpolation" here means using 'best' estimates of missing values

- Interpolated values are smoothed ⇒ variability reduced (affects, e.g., extremes)
- Interpolation introduces artificial inhomogeneities e.g. due to different distances from nearest neighbouring gauges . . .
- and it gives false impression of reduced uncertainty

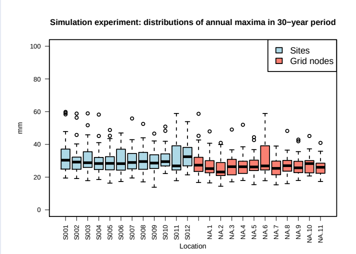
Example: simulation experiment

- Simulate 30-year sequences at 12 locations (blue triangles):
  - Multi-site GLM used: identical structure at all sites
  - Sequences 'typical' of SE England
  - Spatial scale: $\sim 75\%$ of days have sites all wet or all dry, wet-day inter-site correlations ~ 0.6–0.8.
- Use kriging to create gridded daily dataset from simulations
- Regular grid: 12 nodes (red squares)
- Compare annual maxima / return levels for original & gridded data



Results of simulation experiment

Distributions of annual maxima, and pooled return level estimates



Return period	Estimate (mm) Original	Gridded
10 yr	44.0	38.0
50 yr	57.8	49.4
100 yr	63.9	54.4

Actual return periods for gridded estimates: 5, 19 and 34 years

- Maxima for gridded data are smaller and less variable
- Gridding reduces return level estimates by $\sim 15\%$

Handling missing data

- When fitting WG models to sites with missing data, ideally choose approach that does not require complete records
- Multisite model classes for which this is straightforward:
  - Multisite extensions of 'classical' models (calibration done site-by-site)
  - GLMs
  - Models based on transformed Gaussian fields
- For simulation at ungauged locations: better to interpolate WG parameters than data values
  - GLM does this automatically via interactions with 'spatial' covariates

Software packages for weather generation

Name & URL	Notes
`LARS-WG` (`www.rothamsted.ac.uk/mas-models/larswg.php`)	Single-site, multivariate. Based on wet and dry spell length distributions.
`SDSM` (`co-public.lboro.ac.uk/cocwd/SDSM/`	Single-site, multivariate. Based on 'classical' WG formulation.
`WeaGETS` (`www.mathworks.co.uk/matlabcentral/fileexchange/29136-stochastic-weather-generator--weagets-`)	Single-site, multivariate, based on 'classical' WG formulation.
`MulGETS` (`www.mathworks.co.uk/matlabcentral/fileexchange/47537-multi-site-stochstic-weather-generator--mulgets`)	Multi-site, multivariate. Extension of `WeaGETS`, based on Wilks (1998) approach.
`UKCP09` (`ukclimateprojections.metoffice.gov.uk/22540`)	Single-site, multivariate, 'classical' WG formulation but with Poisson cluster model for precipitation component.
`Rglimclim` (`www.homepages.ucl.ac.uk/~ucakarc/work/glimclim.html`)	Multi-site, multivariate, based on GLMs.
`NHMM` (`iamrandom.com/nhmm-package`)	Multi-site, univariate, based on hidden Markov models.

Rglimclim

- Software package for developing multivariate, multisite daily weather generators using GLMs
- Runs under R (http://www.R-project.org) on all platforms
- Based on earlier Glimclim package — Fortran 77(!), multisite but univariate weather generator
- Adds graphical facilities and diagnostics as well as multivariate modelling / simulation capability
- Flexible model structures allow development based on physical understanding rather than statistical convenience
- Allows imputation of missing values (see later)

Modelling capability (I)

- Distributions currently available:
  - Normal (not very useful)
  - Heteroscedastic normal (suitable for, e.g., temperature)
  - Gamma (suitable for, e.g., wind speed, precipitation intensity)
  - Bernoulli (suitable for, e.g., precipitation occurrence)
- Covariate classes:
  - 'Site effects': flexible representation of systematic regional variation ('climatology')
  - Seasonality: various options available
  - Autocorrelation: functions of lagged values
  - Inter-variable dependence: functions of simultaneous and lagged values of other variables
  - 'External' influences e.g. indices of large-scale climate
  - Interactions: allow effects of one variable to be modulated by others

Modelling capability (II)

- Several structures available for representing residual inter-site dependence to ensure spatial coherence
- Most based on correlation structures for standardised / Anscombe residuals (defined so as to have "almost Gaussian" distribution)
- Additional options available for Bernoulli distributions — needed for realistic generation of spatial rainfall occurrence:
  - Thresholding of latent Gaussian field with spatial correlation structure — suitable for large regions
  - Beta-binomial representation for distribution of 'wet area' — suitable for small catchments where inter-site dependence is uniformly high
  - Model based on simple binary weather state process (original Glimclim model — other options preferable)

Model fitting and comparison

- Models fitted using maximum likelihood under (incorrect) assumption of independence between sites
  - Standard IWLS fitting algorithm, augmented to allow estimation of parameters in nonlinear covariate transformations
  - Computationally fast $\Rightarrow$ feasible to fit & compare many different models on large datasets
  - Lose some estimation efficiency compared with fully-specified spatial model — unimportant for large datasets
  - Usual standard errors adjusted for inter-site dependence ('sandwich covariance estimation')
- Model comparison using likelihood ratio tests adjusted for inter-site dependence (methodology of Chandler & Bate, *Biometrika*, 2007)
- Extensive summary and diagnostic information to identify lack-of-fit and guide model-building process

Simulation and imputation

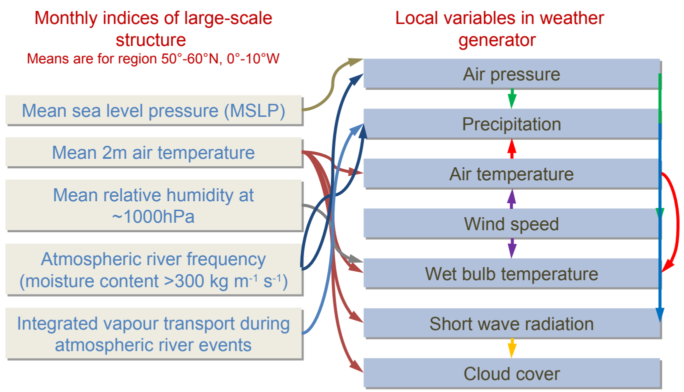
- Simulated sequences can be either unconstrained (conventional WG) or conditioned on all available observations:
  - Allows for multiple imputation of missing observations ⇒ quantifies uncertainty in historical properties
  - Can also be used to ‘interpolate’ to regular grid — alternative to gridded datasets
- Summary and plot methods check ability to reproduce wide variety of properties
- Examples in practical sessions

Example: the Thames again

Variables modelled and distributions used

Variable	Distribution
Air pressure	Normal distribution with changing mean and variance
Rainfall	Logistic regression for occurrence (wet / dry), gamma distribution with changing mean & constant coefficient of variation (CV) for wet-day amounts
Air temperature	Normal distribution with changing mean and variance
Wind speed	Gamma distribution with changing mean & constant CV
Wet bulb temperature	Normal distribution with changing mean and variance
Short wave radiation	Gamma distribution with changing mean & constant CV
Cloud cover	Gamma distribution with changing mean & constant CV

Thames: structure of multivariate model



Summary of Part 2

- Key issue is distinction between systematic regional variation and residual inter-site dependence
- Multi-site methods in literature tend to be designed with specific types of problem in mind, e.g.:
  - Hidden Markov Model (in usual form) suitable for widely separated locations in large regions
  - In small areas, distribution of # of wet sites may better characterise dependence in precipitation occurrence
- Data availability may constrain types of multi-site WG that are appropriate
  - Beware interpolation / gridded datasets!
- Limited software available for multi-site, multivariate weather generation

Part 3: Assessing weather generator performance

Structure of session

- Motivation
- Assessing stochastic models
- Extremes
- Multisite performance

Assessing weather generator performance

Questions:

A user wants to drive an impacts model with a weather generator.

1. How to choose from wide range of generators available?
2. How to determine whether a given generator is fit for purpose?

Issues to consider:

- Ease of use & level of technical sophistication required
- Applicability of key assumptions in user's context
- Ability to calibrate using available data
- Credibility of mechanism for incorporating climate change effects (in user's context)
- Ability to reproduce key features of interest in past observations

What are 'key features of interest'?

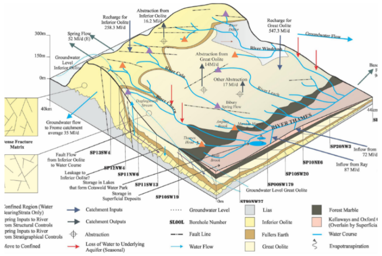
- Relevant features / properties are context-dependent
- From user perspective, ultimate test is realism of impacts model output
  - But this requires user to build WG & run impacts model — may be time-consuming
  - Also, deficiencies may be due to impacts model rather than WG

Aim therefore:

Provide information that enables user to judge whether WG has potential to provide suitable inputs to, e.g., impacts model

Example: distributed hydrological modelling

- Complex hydro(geo)logical models convert spatial rainfall into runoff / groundwater levels etc.
- Precise details depend on land use, soil type, geology, current soil state, river levels etc.



Thanks to colleagues at British Geological Survey

- But to zero-order approximation: need realistic areal average rainfall and realistic rainfall at each individual location — hence focus on these quantities to assess WG performance in this application

The VALUE framework

Decision tree for validating downscaling methods

1. Identify phenomena of interest (precipitation, heatwaves, weather during growing season etc.
2. Identify relevant aspects of weather distribution that are relevant (marginal, temporal, spatial, inter-variable)
3. Identify relevant indices to quantify performance with respect to each aspect
4. Identify performance measures to assess ability of downscaling method to reproduce indices

Application of framework to hydrological modelling example

Phenomena : precipitation and evapotranspiration over catchment

Aspects : marginal (distributions), temporal (spell lengths, seasonality), spatial and intervariable

Indices : e.g. mean, variance, proportion of dry days, autocorrelations, phase and amplitude of seasonal cycle, spatial maps of other properties, variability of areal mean, inter-site correlations, inter-variable correlations

Measures : e.g. bias or relative error

Issues in the assessment of stochastic models

- Means, variances, threshold exceedances, correlations etc. often cannot be deduced from weather generator structure — must use simulations to estimate WG properties
- Stochastic weather generators produce random realisations ⇒ do not expect exact match between WG properties and observations
- Question is not 'does WG output match observations?', but 'do observations look like a realisation from the WG?'

Example: simple temperature generator

Hypothetical example

- Phenomenon: temperature
- Aspects: marginal distribution
- Weather generator is

$$Y_t = \beta_0 + \beta_1 \cos\left[\frac{2\pi \times \text{day of year}}{365}\right] + \beta_2 \sin\left[\frac{2\pi \times \text{day of year}}{365}\right] + \beta_3 Y_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim N(0, \sigma^2)$$

- Daily observations available 1980–2010
- Index: mean
- Performance measure: ???

Temperature example: 'obvious' approach?

- Fit model to observations:
  - Suppose you get $\hat\beta_0 = 3$, $\hat\beta_1 = 3$, $\hat\beta_2 = 0.5$, $\hat\beta_3 = 0.75$, $\sigma^2 = 1$, so model is

$$Y_t = 3 + 3\cos\left[\frac{2\pi \times \text{day of year}}{365}\right] + \frac{1}{2}\sin\left[\frac{2\pi \times \text{day of year}}{365}\right] + \frac{3}{4}Y_{t-1} + \varepsilon_t$$

- Figure out mean temperature for fitted model ($\beta_0/(1-\beta_3) = 12°$ — obvious?).
  - **NB** if interested: mean seasonal cycle for this model given in equation (19) of Yang et al. (2005a) — not at all obvious! See **Exercise 2**
- Compare observed and modelled means — perhaps use t-test?

Problems with 'obvious' approach

- Usually infeasible to derive properties of interest directly from model specification $\Rightarrow$ must use simulations
  - For nonstationary weather generators, use many simulations corresponding to same time period as observations
- Same data used to fit and check model — means guaranteed to be similar!
  - Need independent dataset for testing
  - E.g. fit to data from 1980–2000, test on data from 2001–2010
  - More sophisticated approach: block cross-validation as in VALUE framework

Second attempt

- Fit model to observations 1980-2000
- Carry out many simulations of 2001–2010 period to find mean temperature for this period under model
- Compare with observed mean temperature

How to make comparison?

- Test hypothesis $H_0 : \mu_{\text{sim}} = \bar{Y}_{\text{obs}}$? (**WRONG!**)
- Test null hypothesis $H_0 : \mathbb{E}\left(\bar{Y}_{\text{obs}}\right) = \mu_{\text{sim}}$? (✓?)
  - Care required with interpretation: relevant question is not 'is $\mu_{\text{obs}} = \mu_{\text{sim}}$?', but 'is $|\mu_{\text{obs}} - \mu_{\text{sim}}|$' small enough for WG to be useful?'
  - Also, standard test assumptions unlikely to hold (independence etc.)
- Some role for informal approach

Informal approaches

Key question:

Does observed series ‘look like’ weather generator realisation?

- **Idea**: look at distribution of selected indices across many simulations
  - E.g. 100 simulations give 100 different mean temperatures to form simulated distribution
- If observations were produced by weather generator, observed index should be sampled from this distribution
  - Implication: pool observed index with n simulated indices, rank of observation equally likely to be $1, 2, \ldots$ or $n+1$
  - Basis for Probability Integral Transform (PIT):

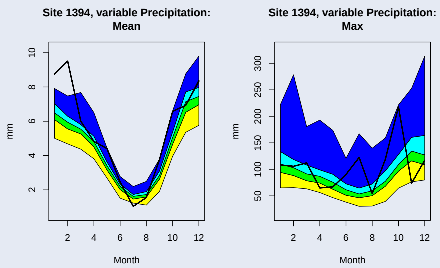
$$\text{PIT} = \frac{\text{rank of observed index}}{n+1}$$

PIT and related techniques

- If many ‘replicate’ indices are computed, can produce PIT histogram — should be flat within sampling error
  - E.g. annual means over 50-year period
- Alternative: for ‘similar but unreplicated’ indices, plot simulated distributions overlain with observations (‘caterpillar plots’):
  - E.g. summary statistics for each month of year

Example: northern Iberia precipitation

Monthly indices for period 1960–1990:

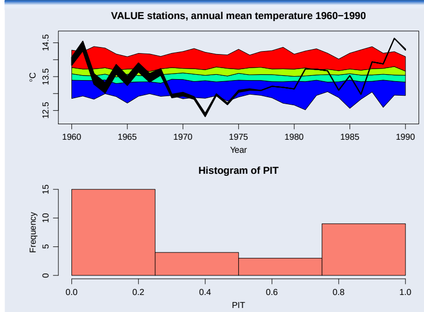


- *Distributions from 100 simulations of 1960–1990 period, with observed statistics superimposed*
- *Coloured bands show range, median and quartiles of simulated distributions*

- Shows underestimation of mean precipitation in January & February
- Note skewed simulation distribution of monthly maxima — typical for precipitation (and realistic according to observations)

Another example: northern Iberia temperature

Annual means for period 1960–1990:



- ***NB** uncertainty in observations due to missing data — uncertainty envelope produced using multiple imputation in* `Rglimclim`
- *39 imputations used for 95% uncertainty interval on observations*
- *Only 31 annual values ⇒ coarse resolution chosen for PIT histogram*

- WG here fails to capture trend (no atmospheric predictors) — does this matter? (is this aspect important?)

Distribution comparisons: quantile-quantile plots

- Further option to assess overall distribution:
  - Compute selected quantiles of observations
  - Compute corresponding quantiles of pooled distribution from all simulations
  - Plot against each other — should be roughly equal
- Quantile estimates are biased near 0 and 1, especially with small samples in observations ⇒ avoid extreme quantiles here
- Can use to assess agreement in, e.g., overall distribution of annual maxima throughout simulation period
  - Example in practical session

Assessing extremes — motivation

- Many applications support decisions with implications over long periods e.g.

  Flood defences : design lifetime 30–50 years
  Investment in energy infrastructure : returns over 10–20 year periods
  Agricultural development : adaptation strategies with 5–20 year horizons
  Safety of nuclear waste repositories : silly time scales
- Risk-based approach: plan for specified chance of coping with worst scenario in decision horizon
  - E.g. flood defences: 10% chance of failure in 50 years (say)
- Leads to consideration of very rare events:
  - E.g. $\sim$ '1 in 500 year' event in flood defence example
- Compare with 'extremes' often studied in downscaling e.g. 95th percentile of daily distribution ('1 in 20 day')

Extreme value theory

Question:
How to assess credibility of rare events in weather generator simulations?

- **Possible approach**: compare simulated and observed distributions of (e.g.) annual maxima
  - **Problem**: want (e.g.) 99th percentile of distribution of annual maximum, have (say) 30-year record ⇒ 30 observations
- Need principled basis for heroic extrapolation!
- Extreme value theory provides such a basis — analogous to Central Limit Theorem for means

Extreme Value Theory in one slide

Key result (paraphrase)
In almost all situations of practical interest, the maximum of a large collection of independent, identically distributed random variables has approximately a Generalised Extreme Value (GEV) distribution

- Parameters of distribution: shape ξ, scale σ, location μ
- Result also holds for dependent sequences
- Can also argue that it should hold for, e.g., annual maxima even though variables are not identically distributed (Chandler and Scott, 2011, §6.4)
- Hence common to fit GEV distributions to annual maxima (mximum likelihood preferred) and use fitted distributions for extrapolation
- GEV result underpins all mathematically justified alternative methods e.g. peaks-over-threshold, point process likelihood — see Coles (2001) for more details

Implication of EV theory

Recall the question . . .

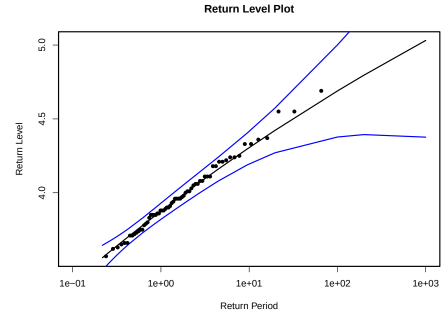
How to assess credibility of rare events in weather generator simulations?

. . . and the previously suggested answer:

Compare simulated and observed distributions of (e.g.) annual maxima

- Extreme Value Theory provides defensible alternative: replace observed distribution with GEV distribution fitted to observed maxima
  - Need to account for uncertainties in GEV-based extrapolation — maximum likelihood estimation enables this
  - Uncertainties usually shown on return level plot: shows estimate of values exceeded with frequencies from once per year to once every *N* years
  - Observations added to plot as check on GEV fit
- Possibility for weather generator assessment: add *simulated* maxima to 'observed' return level plot (example in practical session)

Example of return level plot



Return level plot for annual maximum sea levels at Port Pirie, South Australia, 1923–1987 (data from `ismev` *library in* `R`*, originally in Coles (2001))*

The GEV shape parameter

- Shape parameter ξ plays crucial role in behaviour of extremes:
  - $\xi < 0$: finite upper limit
  - $\xi = 0$: infinite upper limit but light tail
  - $\xi > 0$: infinite upper limit and heavy tail (potential for 'black swans')
- If using weather generator for extremes, minimal requirement is that associated value of ξ is roughly correct
- **Fact**: for independent sequences, underlying distribution determines value of ξ e.g.

  Normal distributions : lead to $\xi = 0$
  Gamma distributions : lead to $\xi = 0$
  Pareto distributions : lead to $\xi > 0$
- **But**: tail behaviour can be different in dependent sequences specified via conditional distributions (see **Exercise 3**)

Assessment of multi-site performance

- If spatial aspects are important then these must be assessed
- Systematic variation: use single-site measures at selected sites
  - May want to map single-site measures or plot against (e.g.) site altitude — but would need to reduce previous graphs to single measure e.g. mean bias over all simulations
  - **NB** also mapping involves interpolation — beware artefacts!
- 'Residual inter-site dependence' now better characterised via indices of joint distributions at sets of sites e.g.
  - Correlations / variograms of (standardised?) anomalies — similar comments apply
  - Probabilities of simultaneous threshold exceedances e.g. Yan et al. (2006)
- Alternative approach: work with spatially aggregated daily series
  - Easier to apply & tests for realistic spatial coherence in WG output
  - More user-relevant in some applications e.g. hydrological modelling

Summary of Part 3

- Many judgements can be made without assessing WG performance (what was it designed for, what data are required, ...)
- Different WGs appropriate depending on key features of interest in application
- Aim of performance assessment: determine whether WG has potential to provide suitable inputs to (e.g.) impacts model
- VALUE decision tree (Phenomena $\rightarrow$ Aspects $\rightarrow$ Indices $\rightarrow$ Measures) helps to structure assessment exercise
- Question for stochastic WGs framed as 'Do observations look like realisation from WG?'
- Need independent test data / block cross-validation for credible assessments
- Clear role for informal / graphical assessments of performance: not 'is it right?' but 'is it good enough?'

References

Abourgila, A. E. (1992). Large-scale hydrological modelling of the Nile basin for the assessment of the impact of climate and land use changes. MSc thesis, Department of Civil Engineering, Imperial College London.

Ailliot, P., Thompson, C., and Thomson, P. (2009). Spacetime modelling of precipitation using a hidden Markov model and censored Gaussian distributions. *Appl. Statist.*, 58:405–426.

Ambrosino, C., Chandler, R. E., and Todd, M. C. (2014). Rainfall-derived growing season characteristics for agricultural impact assessments in South Africa. *Theor. Appl. Climatol.*, 115:411–426. DOI: 10.1007/s00704-013-0896-y.

Buishand, T. and Brandsma, T. (2001). Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling. *Water Resources Research*, 37:2761–2776.

Chandler, R., Isham, V., Northrop, P., Wheater, H., Onof, C., and Leith, N. (2014). Uncertainty in rainfall inputs. In ad J. Hall, K. B., editor, *Applied Uncertainty Analysis for Flood Risk Management*, pages 101–152. Imperial College Press, London.

Chandler, R. E. and Scott, E. M. (2011). *Statistical methods for trend detection and analysis in the environmental sciences*. John Wiley & Sons, Chichester.

Chandler, R. E. and Wheater, H. S. (2002). Analysis of rainfall variability using Generalized Linear Models — a case study from the West of Ireland. *Water Resources Research*, 38, No.10:doi:10.1029/2001WR000906.

Charles, S., Bates, B., and Hughes, J. (1999a). A spatiotemporal model for downscaling precipitation occurrence and amounts. *J. Geophys. Res — Atmospheres*, 104 (D24):31657–31669.

Charles, S., Bates, B., Whetton, P., and Hughes, J. (1999b). Validation of downscaling models for changed climate conditions: case study of southwestern Australia. *Clim. Res.*, 12:1–14.

Coe, R. and Stern, R. D. (1982). Fitting models to daily rainfall. *J. Appl. Meteorol.*, 21:1024–1031.

Coles, S. (2001). *An introduction to the statistical modelling of extreme values*. Springer series in statistics. Springer-Verlag, London.

Finch, J. W., Bradford, R. B., and Hudson, J. A. (2004). The spatial distribution of groundwater flooding in a chalk catchment in southern England. *Hydrol. Process*, 18:959–971.

Gabriel, K. R. and Neumann, J. (1962). A Markov chain model for daily rainfall in Tel Aviv. *Q. J. Roy. Meteor. Soc.*, 88:90–95.

Grunwald, G. K. and Jones, R. H. (2000). Markov models for time series with mixed distribution. *Environmetrics*, 11:327–339.

Katz, R. and Parlange, M. (1998). Overdispersion phenomenon of stochastic modeling of precipitation. *J. Climate*, 11:591–601.

Katz, R., Parlange, M., and Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, 25:1287–1304.

Kilsby, C., Jones, P., Burton, A., Ford, A., Fowler, H., Harpham, C., James, P., Smith, A., and Wilby, R. L. (2007). A daily weather generator for use in climate change studies. *Environ. Modell. Software*, 22(12):1705–1719.

Kniveton, D. R., Layberry, R., Williams, C. J. R., and Peck, M. (2009). Trends in the start of the wet season over Africa. *Int. J. Climatol.*, 29:1216–1225.

Met Office (2005). Boscastle and North Cornwall post flood event study — meteorological analysis of the conditions leading to flooding on 16 August 2004. In Goulding, B., editor, *Forecasting Research Technical Report N. 459*. UK Meteorological Office, Exeter.

Richardson, C. W. (1981). Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research*, 17:182–190.

Semenov, M., Brooks, R., Barrow, E., and Richardson, C. (1998). Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Clim. Res.*, 10:95–107.

Smith, L. (2002). What might we learn from climate forecasts? *PNAS*, 99:2487–2492.

Stehlík, J. and Bárdossy, A. (2002). Multivariate stochastic downscaling model for generating daily precipitation series based on atmospheric circulation. *J. Hydrol*, 256:120–141.

Wilks, D. (1998). Multisite generalization of a daily stochastic precipitation generation model. *J. Hydrol.*, 210:178–191.

Yan, Z., Bate, S., Chandler, R. E., Isham, V. S., and Wheater, H. S. (2006). Changes in extreme wind speeds in NW Europe simulated by Generalized Linear Models. *Theoretical and Applied Climatology*, 83:121–137.

Yang, C., Chandler, R. E., Isham, V. S., Annoni, C., and Wheater, H. S. (2005a). Simulation and downscaling models for potential evaporation. *J. Hydrol.*, 302:239–254.

Yang, C., Chandler, R. E., Isham, V. S., and Wheater, H. S. (2005b). Spatial-temporal rainfall simulation using generalized linear models. *Water Resources Research*, 41:doi:10.1029/2004WR003739.