A Framework for Daily Spatio-Temporal Stochastic Weather Simulation

Will Kleiber, Rick Katz, Balaji Rajagopalan

Geophysical Statistics Project
Institute for Mathematics Applied to Geosciences
National Center for Atmospheric Research
Boulder, CO

Roscoff, France, May 28, 2012
Agricultural, ecological, hydrological models often require daily weather

- On grid
- In the past or future

Stochastic Weather Generators (SWGs) are statistical models whose simulated values “look like” observed weather

SWGs can be used to produce infinitely long series of synthetic weather, for observation network infilling, or climate model downscaling
General Framework

For a spatially consistent stochastic weather generator, think of generating

Local Climate + Weather

- Local Climate: Generalized linear model (include covariates, model parameters spatially varying, portable between domains)

- Weather: Spatial Gaussian process (flexible spatial model)
We say $W(s)$, indexed by location $s \in \mathbb{R}^d$, is a isotropic Gaussian process if

i) For any $s_1, \ldots, s_n \in \mathbb{R}^d$, $(W(s_1), \ldots, W(s_n))'$ is multivariate normal

ii) $E W(s) = \mu$ for all $s \in \mathbb{R}^d$

iii) $\text{Cov}(W(s + h), W(s)) = C(||h||)$ for all $s, h \in \mathbb{R}^d$

Model is complete with $\mu$ and $C(||h||)$. 

Will Kleiber
Stochastic Weather Generators
Kriging

Suppose we observe $W(s)$ at $s = s_1, \ldots, s_n$, and want $W(s_0)$.

The kriging estimator or interpolator for $W(s_0)$ is

$$\hat{W}(s_0) = c'\Sigma^{-1}(W - \mu)$$

with interpolation variance

$$C(s_0, s_0) - c'\Sigma^{-1}c$$

(these coincide with conditional expectation and conditional variance of a multivariate normal).
Daily Precipitation
Precipitation at NCAR
Data: 22 stations from the United States Historical Climatology Network over the years 1893-2009.
Location $s$, day $t$, occurrence $O$:

$$O(s, t) = \begin{cases} 0 & \text{if } W(s, t) < 0 \\ 1 & \text{if } W(s, t) \geq 0 \end{cases}$$

where $W$ is a Gaussian process with mean

$$\mu(s, t) = \beta'_\mu X_\mu(s, t)$$

( = local climate)

and covariance

$$C(h, t) = \exp(-\|h\|/A(t)).$$
Occurrence

Location $s$, day $t$, occurrence $O$:

$$O(s, t) = 0 \quad \text{if} \quad W(s, t) < 0$$
$$O(s, t) = 1 \quad \text{if} \quad W(s, t) \geq 0$$

where $W$ is a Gaussian process with mean

$$\mu(s, t) = \beta_\mu ' X_\mu(s, t)$$

($\mu$ = local climate)

and covariance

$$C(h, t) = \exp(-\|h\|/A(t))$$
Location $s$, day $t$, occurrence $O$:

$$O(s, t) = 0 \quad \text{if} \quad W(s, t) < 0$$

$$O(s, t) = 1 \quad \text{if} \quad W(s, t) \geq 0$$

where $W$ is a Gaussian process with mean

$$\mu(s, t) = \beta_{\mu}(s)'X_{\mu}(s, t)$$

( = local climate)

and covariance

$$C(h, t) = \exp(-\|h\|/A(t)).$$
We suppose

\[ \beta \sim \text{Gaussian process} \]

with mean \( \mu \) and covariance function

\[
\text{Cov}(\beta(s + h), \beta(s)) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (a\|h\|)^\nu \mathcal{K}_\nu(a\|h\|) + \tau^2 \mathbb{I}_{\|h\|=0}
\]

aka a Matérn covariance with nugget effect.
We suppose

\[ \beta \sim \text{Gaussian process} \]

with mean \( \mu \) and covariance function

\[
\text{Cov}(\beta(s + h), \beta(s)) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (a\|h\|)^\nu K_\nu(a\|h\|) + \tau^2 \mathbb{I}_{\|h\| = 0}
\]

aka a Matérn covariance with nugget effect.

Locally, estimate \( \hat{\beta}(s) \) by maximum likelihood, and conditional on these, estimate \( \mu, \sigma^2, \nu, a, \tau^2 \)

by maximum likelihood.
Kriging

Estimates

Kriged

Will Kleiber
Stochastic Weather Generators
Precipitation

Temperature

Conclusion

Daily Precipitation

Occurrence

Amounts

Occurrence

Mean

Residual

Mean + Residual

Occurrence

Will Kleiber

Stochastic Weather Generators
Model amount $Y(s, t)$ with gamma cdf $G_{s,t}$ with scale $\alpha$ and shape $\gamma$:

$$\log \alpha(s, t) = \beta_\alpha(s)'X\alpha(s, t)$$
$$\log \gamma(s, t) = \beta_\gamma(s)'X\gamma(s, t)$$
Amounts

Model amount $Y(s, t)$ with gamma cdf $G_{s,t}$ with scale $\alpha$ and shape $\gamma$:

$$\log \alpha(s, t) = \beta_\alpha(s)'X_\alpha(s, t)$$
$$\log \gamma(s, t) = \beta_\gamma(s)'X_\gamma(s, t)$$

For spatially correlated fields of precipitation, use Gaussian process $W(s, t)$

$$Y(s, t) = G_{s,t}^{-1} (\Phi(W(s, t)))$$

where $\Phi$ is the standard normal cdf.
**Amounts**

### Shape Parameter

- Shape Parameter values range from 0.6 to 1.4.
- Different colors represent varying parameter values.

### Gaussian Realization

- Gaussian realization values range from -1 to 3.
- Colors indicate different realizations.

### Probability Integral Transform

- Probability integral transform values range from 0.2 to 0.8.
- Colors show variations in the transform.

### Rain

- Rain values range from 10 to 50.
- Colors illustrate different rainfall amounts.
Daily Simulation

Day 1

Day 2

Day 3

Day 4

Will Kleiber

Stochastic Weather Generators
Colorado

Data: 145 stations from the Global Historical Climatology Network. Daily minimum and maximum temperature between 1893-2011.
Model min and max temperature $Z_N$ and $Z_X$ as

$$Z_N(s,t) = \beta_N(s)'X_N(s,t) + W_N(s,t)$$
$$Z_X(s,t) = \beta_X(s)'X_X(s,t) + W_X(s,t)$$

($ = \text{Local Climate} + \text{Weather}$).

As with precipitation,

$$\beta_{N,i}(s), \beta_{X,i}(s) \sim \text{Gaussian process}.$$
Min/Max Temperature

Complicated relationship between min and max temperature:

Evidence of **nonstationary cross-correlation**.
The matrix-valued covariance function

\[ C(x, y) = \begin{pmatrix} C_{NN}(x, y) & C_{NX}(x, y) \\ C_{XN}(x, y) & C_{XX}(x, y) \end{pmatrix}, \]

with diagonals direct-covariance functions and off diagonal cross-covariance functions, e.g.

\[ C_{NN}(x, y) = \text{Cov}(Z_N(x), Z_N(y)) \]
\[ C_{NX}(x, y) = \text{Cov}(Z_N(x), Z_X(y)). \]
Covariance for Temperature

For $i, j = N, X$, at sites $x, y$ and day $t$,

$$\text{Cov}(W_i(x, t), W_j(y, t + 1)) = 0$$
Covariance for Temperature

For $i, j = N, X$, at sites $x, y$ and day $t$,

$$\text{Cov}(W_i(x, t), W_j(y, t + 1)) = 0$$

$$\text{Cov}(W_N(x, t), W_N(y, t)) = C_{NN}(x, y, t) + \tau_N(x, y)^2 \mathbb{I}_{[x=y]}$$
Covariance for Temperature

For \( i, j = N, X \), at sites \( x, y \) and day \( t \),

\[
\text{Cov}(W_i(x, t), W_j(y, t + 1)) = 0
\]

\[
\text{Cov}(W_N(x, t), W_N(y, t)) = C_{NN}(x, y, t) + \tau_N(x, y)^2 \mathbb{I}[x=y]
\]

\[
\text{Cov}(W_N(x, t), W_X(y, t)) = C_{NX}(x, y, t)
\]
For $i, j = N, X$, estimator of $C_{ij}(x, y, t_0)$ is

$$\frac{\sum_{t=1}^{n_i} \sum_{k=1}^{n} \sum_{\ell=1}^{n} K_{\lambda_i}(t - t_0) K_{\lambda}(\|x - s_k\|) K_{\lambda}(\|y - s_{\ell}\|) W_i(s_k, t) W_j(s_{\ell}, t)}{\sum_{t=1}^{n_i} \sum_{k=1}^{n} \sum_{\ell=1}^{n} K_{\lambda_i}(t - t_0) K_{\lambda}(\|x - s_k\|) K_{\lambda}(\|y - s_{\ell}\|)}$$

for kernel functions $K$ and bandwidths $\lambda$ and $\lambda_t$. 
Spatial Correlation

Min Temperature Correlation

Max Temperature Correlation

Cross-Correlation

Longitude

Latitude

Min Temperature Correlation

Max Temperature Correlation

Cross-Correlation

Longitude

Latitude

Min Temperature Correlation

Max Temperature Correlation

Cross-Correlation

Longitude

Latitude

Will Kleiber
Stochastic Weather Generators
Conclusions

- Stochastic Weather Generator framework: Local Climate + Weather

- Generalized linear models make SWG portable

- Gaussian processes are flexible spatial models for climate and weather
