Stochastic space-time modeling for agricultural decision support in the Argentine Pampas

Doctoral Dissertation Defense

by
Andrew Verdin

Dept of Civil, Environmental, and Architectural Engineering
University of Colorado at Boulder

Advised by Prof. Balaji Rajagopalan

April 8, 2016

Table of Contents

Motivation

"It is arguable that the artificiality of agricultural production systems make them less flexible, and therefore more vulnerable to climatic change than the naturally occurring species of the ecosystem within which they fit, and that the more unstable the climate the greater this vulnerability is likely to be."

- Oram (1985)

THE PAMPAS

Alternating wet and dry epochs:

- Early 1900s floods.
- 1930-50 drought.
- Increasing precipitation 1960-2000.

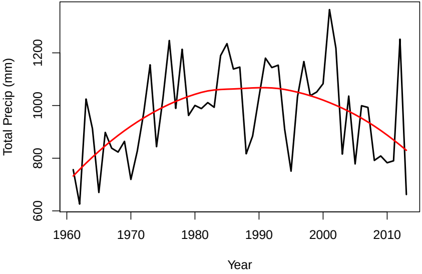
  $\rightarrow$ Extremes & frequency.

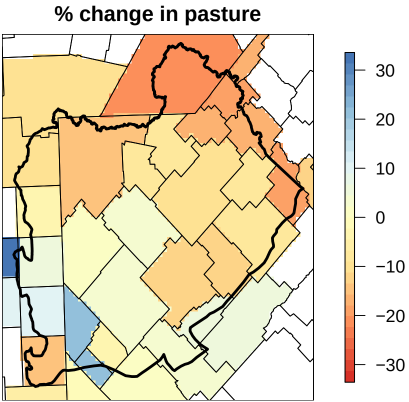
- Agricultural expansion.

  $\rightarrow$ Reduced annual ET.

  $\rightarrow$ Rising water table.

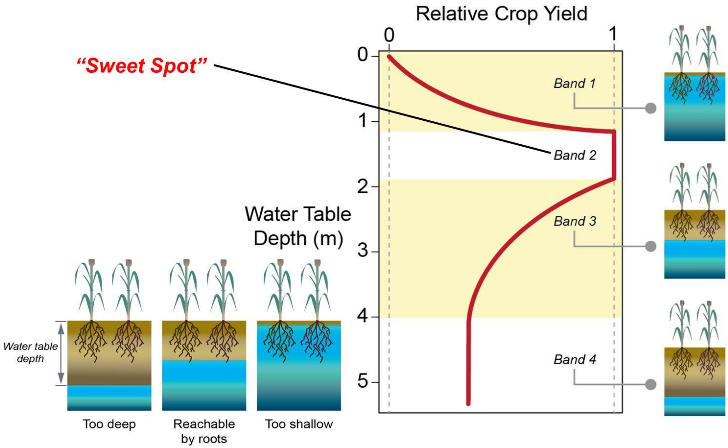
- 1997-2003 flooding.
- 2008 drought.





$$\Delta = (1981\text{-}2013) - (1961\text{-}1980)$$

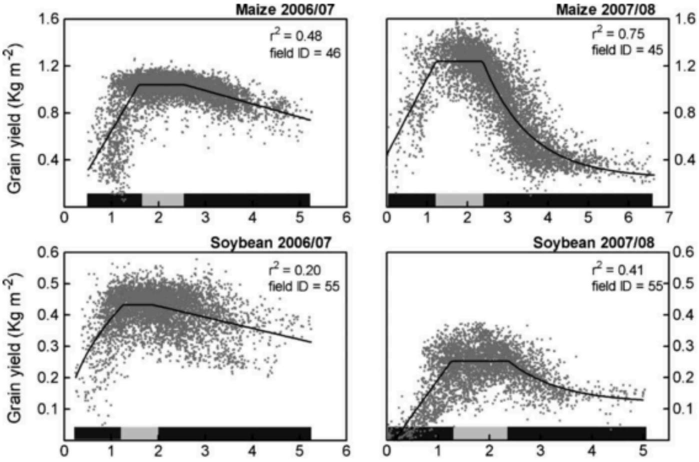
CROP YIELD MAPPING



courtesy: Guillermo Podestá

WTD has been shown to affect crop yields.

Nosetto et al. [2009]



Agricultural planning in the Pampas

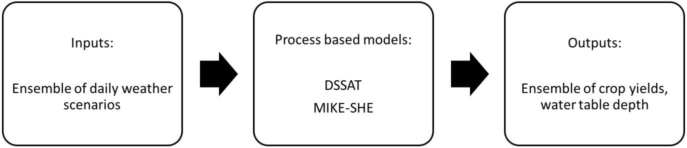
- Crop simulation models (DSSAT).
- Analyze different management strategies.
- Correct crop yields with WTD (MIKE-SHE).

  → MIKE-SHE has great computational costs. (5km x 5km, daily)

…the need for data.

- Observed record limited.
- Variety of scenarios.
- Specific trends; traits of seasonal forecast.
- At unobserved locations.

→ **Stochastic weather generation!**



Single site weather generators

Parametric:

[Richardson, 1981; Richardson & Wright, 1984; Semenov & Barrow, 1997]

- Precipitation occurrence modeled as Markov chain. *[Katz, 1977]*
- Precipitation amount modeled as Gamma random variable.
- Temperatures modeled using autoregressive time series.
- *Fit separate models for each month (or finer).*
- *Captures climatology and linear relationships; fails to capture extremes.*

Nonparametric:

[Lall & Sharma, 1996; Rajagopalan & Lall, 1999; Harrold et al., 2003]

- Kernel density estimators of precipitation.
- K-nearest neighbor time series bootstrap.
- *Non-normal features captured.*
- *Resamples historic record, not simulating unobserved values.*

Multiple site weather generators

Parametric:

[Wilks, 1998, 1999; Qian et al., 2002; Baigorria & Jones, 2010; Khalili et al., 2009]

- Climate variables modeled consistent with single site.
- Estimate parameters for each variable at each location.
- *Unwieldy to preserve temporal statistics and spatial dependence.*
- *Difficult to simulate at <u>unobserved</u> locations, or large network of stations.*

Nonparametric:

[Yates et al., 2003; Apipattanavis et al., 2007; Sharif & Burn, 2007]

- Resample observed vector of variables.
- Preserve spatial dependence and climatological statistics.
- *Due to resampling, not simulating unobserved values.*
- *Difficult to simulate at <u>unobserved</u> locations, or large network of stations.*

Generalized linear models:

[Stern & Coe, 1984; Chandler, 2005; Furrer & Katz, 2007; Kleiber et al., 2012, 2013]

Model and simulate skewed or discrete variables.

$Y \sim f(\mu, \sigma)$

$f(.) \rightarrow$ binomial, Gamma, poisson, gaussian.

$\eta(\mu) = \mathbf{X}'\boldsymbol{\beta}$

$\eta(\mu) \rightarrow$ logit, probit, inverse, log, identity.

- Time series models at each location.
- Covariates capture seasonality, temporal dependence.
  $\rightarrow$ e.g., sine, cosine, autoregressive.

- GLM coefficients (i.e., β_i) are fixed.
  $\rightarrow$ Uncertainty is neglected.
  $\rightarrow$ Variability from residuals process.

Research objectives

1. Develop a weather generator for simulating at unobserved locations.
   - Space-time GLM weather generator.

2. Generate space-time weather scenarios consistent with seasonal forecasts and multidecadal trends, for resources planning and management.
   - Modification of GLM weather generator.

3. A generalized weather generator that can model and propagate uncertainty.
   - Bayesian space-time weather generator.

4. An efficient water table depth model for seasonal agricultural planning.
   - Hydrologic model emulator (i.e., "metamodel").

Table of Contents

1. **Develop a weather generator for simulating at unobserved locations.**

Fit GLMs at each location, obtain $\boldsymbol{\beta}(\boldsymbol{s})$.

$\rightarrow$ Captures temporal variability.

Spatial models to estimate $\hat{\boldsymbol{\beta}}(\boldsymbol{s})$ coefficients at unobserved locations.

$\rightarrow$ Captures spatial structure.

Exponential covariance functions defined by residuals.

Occurrence $O(\boldsymbol{s}, t) = \mathbb{1}_{[W_O(\boldsymbol{s},t) \geq 0]}$ modeled using probit regression.

Latent Gaussian process $W_O(\boldsymbol{s}, t) = \mathbf{X}_O(\boldsymbol{s}, t)'\boldsymbol{\beta}_O(\boldsymbol{s})$ is linear regression.
By definition, probit residual variance $\sigma_O^2 = 1$

If $W_O(\boldsymbol{s}, t) \geq 0$, $O(\boldsymbol{s}, t) = 1$
Else, $O(\boldsymbol{s}, t) = 0$
$\mathbf{X}_O(\boldsymbol{s}, t) = (1, O(\boldsymbol{s}, t-1), \cos(2\pi t/365), \sin(2\pi t/365))$

Amounts $A(\boldsymbol{s}, t)$ modeled as Gamma$(\alpha_A(\boldsymbol{s}), \alpha_A(\boldsymbol{s})/\mu_A(\boldsymbol{s}, t))$.

Gamma shape $\alpha_A(\boldsymbol{s})$ varies in space.
Gamma scale $\alpha_A(\boldsymbol{s})/\mu_A(\boldsymbol{s}, t)$ varies in space and time.

$\mu_A(\boldsymbol{s}, t) = \exp(\mathbf{X}_A(\boldsymbol{s}, t)'\boldsymbol{\beta}_A(\boldsymbol{s}))$
$\mathbf{X}_A(\boldsymbol{s}, t) = (1, \cos(2\pi t/365), \sin(2\pi t/365))$

Model structure

Temperatures modeled using linear regression:

$$Z_N(\boldsymbol{s},t) = \mathbf{X}_N(\boldsymbol{s},t)'\boldsymbol{\beta}_N(\boldsymbol{s}) + W_N(\boldsymbol{s},t)$$

$$Z_X(\boldsymbol{s},t) = \mathbf{X}_X(\boldsymbol{s},t)'\boldsymbol{\beta}_X(\boldsymbol{s}) + W_X(\boldsymbol{s},t)$$

$$\begin{aligned}\mathbf{X}_i(\boldsymbol{s},t) = (1, Z_N(\boldsymbol{s},t-1), Z_X(\boldsymbol{s},t-1), \cos(2\pi t/365),\\ \sin(2\pi t/365), r(t), O(\boldsymbol{s},t)) \qquad \text{for } i = N, X\end{aligned}$$

The mean $\mathbf{X}_i(\boldsymbol{s},t)'\boldsymbol{\beta}_i(\boldsymbol{s})$ is climate.

Residuals $W_i(\boldsymbol{s},t)$ is like daily weather.

Parameter estimation & simulation

Model fitting:

1. Four GLMs are fitted at each location.
2. Maximum likelihood estimates (MLE) of $\hat{\boldsymbol{\beta}}_i(\boldsymbol{s})$ at each location.
3. Spatial models on $\hat{\boldsymbol{\beta}}_i(\boldsymbol{s})$ to simulate at any location.
4. Exponential covariance functions defined from monthly residuals.

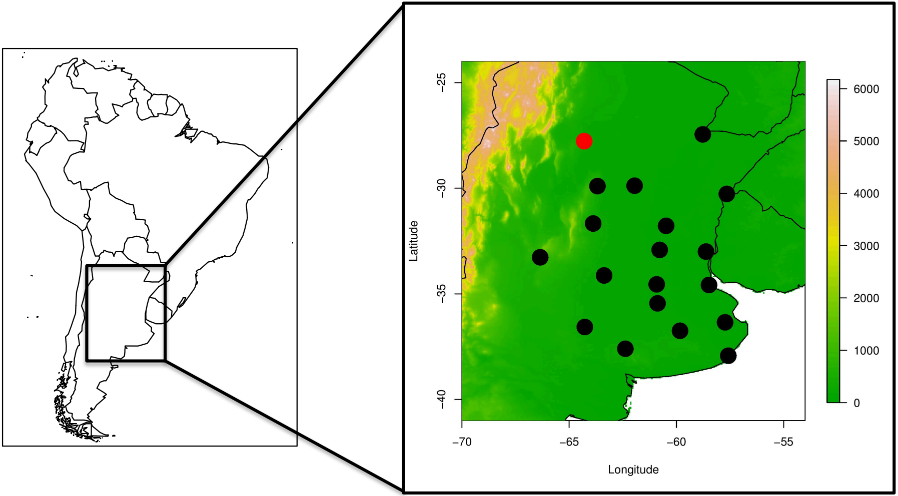
Parameter estimation & simulation

Model fitting:

1. Four GLMs are fitted at each location.
2. Maximum likelihood estimates (MLE) of $\hat{\boldsymbol{\beta}}_i(\boldsymbol{s})$ at each location.
3. Spatial models on $\hat{\boldsymbol{\beta}}_i(\boldsymbol{s})$ to simulate at any location.
4. Exponential covariance functions defined from monthly residuals.

Simulation:

1. Simulate occurrence.
2. Where occurrence is positive, simulate amounts.
3. Update temperature covariates using simulated occurrence.
4. Simulate temperatures.
5. Repeat for all days to simulate.

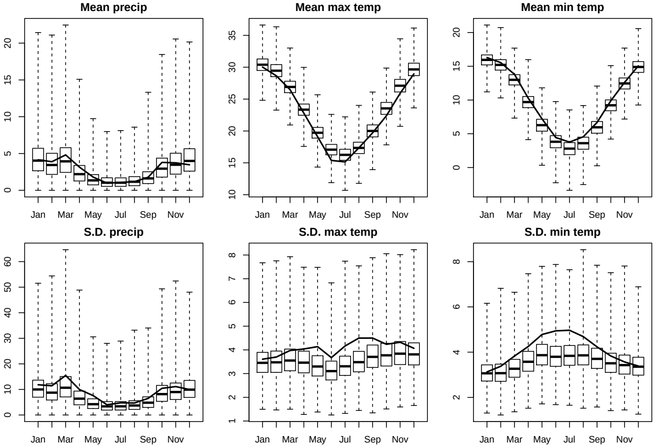


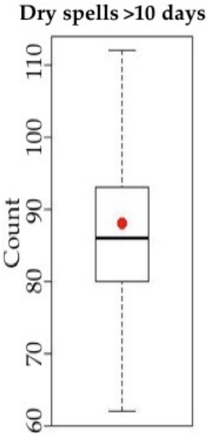
1908–2010 daily precipitation, maximum and minimum temperature.

Simulation & validation

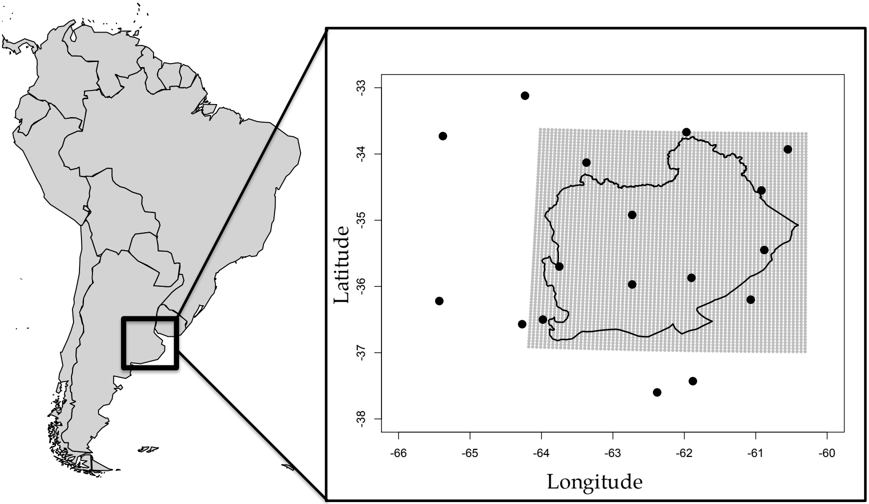
100 simulations: 1 Jan 1908 - 31 Dec 2010.

Cross-validation





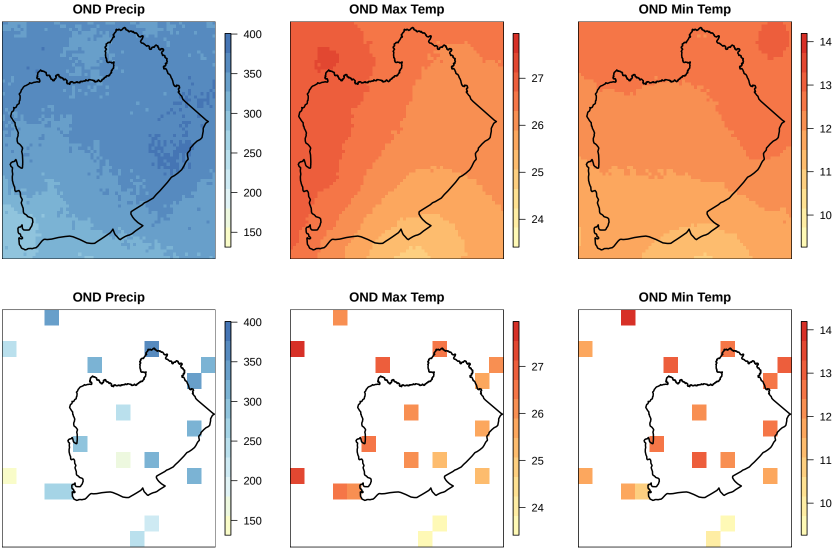
Gridded simulation



1961–2013 daily precipitation, maximum and minimum temperature.

Gridded simulation

Ensemble mean OND seasonal total precip, mean max temp, mean min temp.



CONTRIBUTIONS

✓ Developed GLM based stochastic weather generator.
   $\rightarrow$ Spatial models on GLM coefficients.
   $\rightarrow$ Residuals as monthly Gaussian processes.

Contributions

✓ Developed GLM based stochastic weather generator.
$\rightarrow$ Spatial models on GLM coefficients.
$\rightarrow$ Residuals as monthly Gaussian processes.

✓ Address problem of simulating daily weather at <u>unobserved</u> locations.

Contributions

✓ Developed GLM based stochastic weather generator.
$\rightarrow$ Spatial models on GLM coefficients.
$\rightarrow$ Residuals as monthly Gaussian processes.

✓ Address problem of simulating daily weather at unobserved locations.

✓ Ease modeling of skewed & discrete variables.

Verdin et al., (2015). Coupled stochastic weather generation using spatial and generalized linear models. *Stochastic Environmental Research and Risk Assessment*, 29(2), 347–356. doi:10.1007/s00477-014-0911-6

Table of Contents

2. **Generate space-time weather scenarios consistent with seasonal forecasts and multidecadal trends, for resources planning and management.**

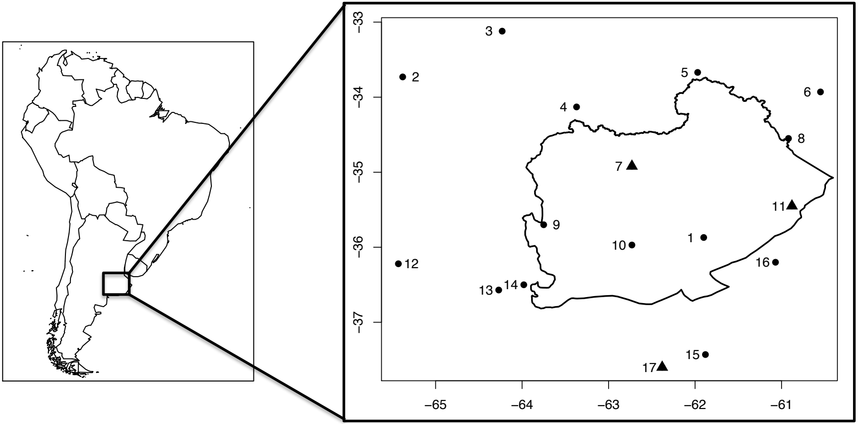
Propose the use of GLM weather generator for planning purposes.

Additional covariates, i.e., seasonal climate information.

Enables translation of seasonal forecasts or climate model projections.

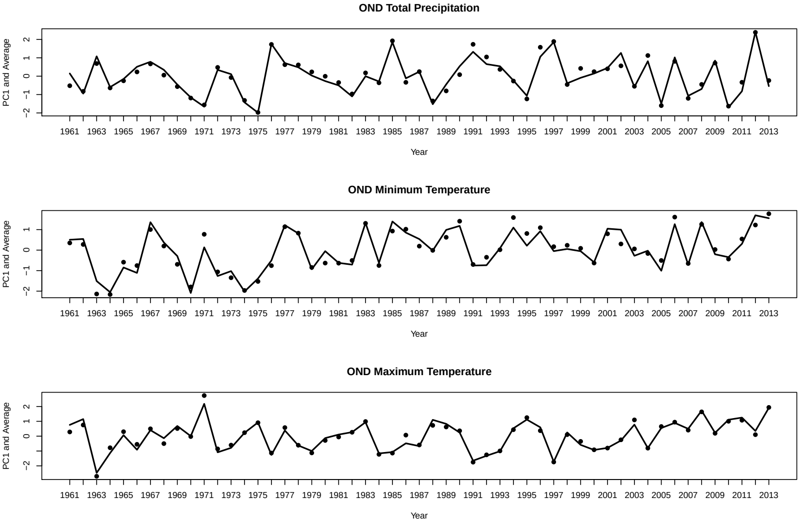
$\rightarrow$ Downscaling coarse scale information.

$\rightarrow$ Address the disconnect between data & decision.



1961–2013 daily precipitation, maximum and minimum temperature.

Domain average OND (lines) and PC1 of all stations.



Additional covariates

Precipitation models:
Domain average seasonal total precipitation.

$$\mathbf{X}_O(\boldsymbol{s}, t) = (1, \ldots, ST1(t), ST2(t), ST3(t), ST4(t))$$

$$\mathbf{X}_A(\boldsymbol{s}, t) = (1, \ldots, ST1(t), ST2(t), ST3(t), ST4(t))$$

Temperature models:
Domain average seasonal mean maximum and minimum temperature.

$$\begin{aligned}\mathbf{X}_i(\boldsymbol{s}, t) = (1, \ldots, &SMN1(t), SMN2(t), SMN3(t), SMN4(t),\\ &SMX1(t), SMX2(t), SMX3(t), SMX4(t))\\ &\text{for } i = N, X\end{aligned}$$

Validate by comparison to unconditional GLM weather generator.

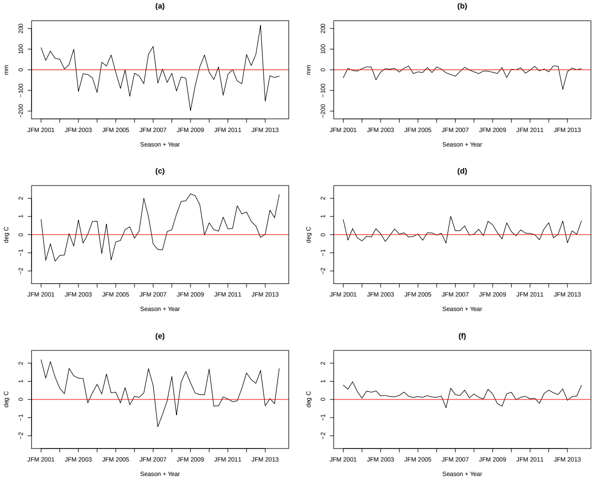
13 years: JFM 2001 – OND 2013

Unconditional: Conditional:

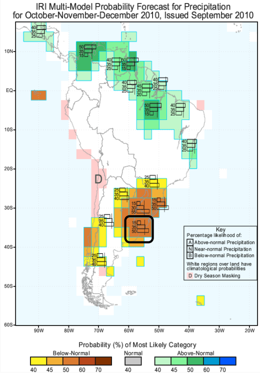
Precip Precip

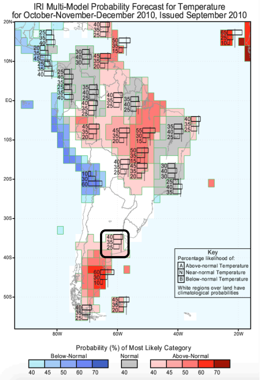
Max Temp Max Temp

Min Temp Min Temp



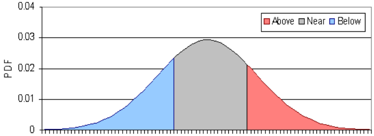
SEASONAL FORECAST





OND 2010 Precipitation: (15:35:50) (A:N:B)

OND 2010 Temperature: (40:35:25) (A:N:B)



1. Classify observed OND precipitation and temperatures as A:N:B.

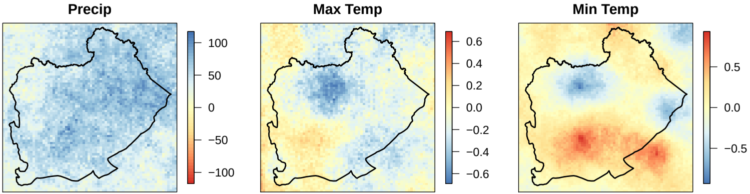
2. Bootstrap ensemble of OND covariates, A:N:B as weights.

3. One simulation per ensemble member.

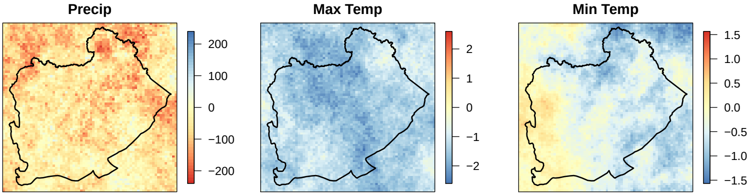
4. Ensemble of weather reflects seasonal forecast.

Seasonal forecast

Differences in ensemble mean (unconditional minus conditional):



Differences in 95% ensemble spread (unconditional minus conditional):



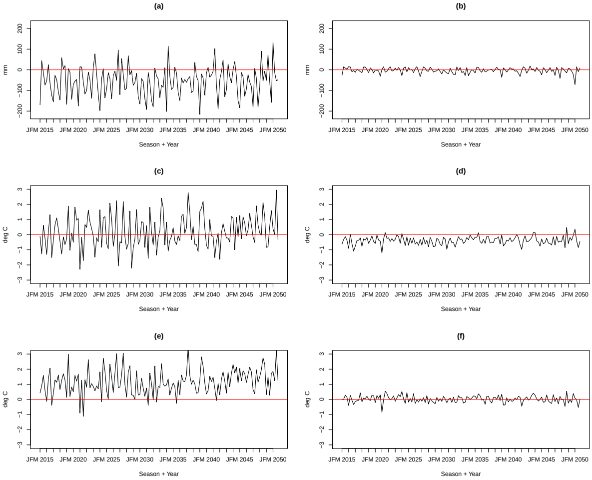
CORDEX-CMIP5 RCM (2015–2050)

Unconditional: | Conditional:

Precip | Precip

Max Temp | Max Temp

Min Temp | Min Temp



✓ Condition GLM weather generator on seasonal forecasts, trends.

Contributions

✓ Condition GLM weather generator on seasonal forecasts, trends.

✓ Translate coarse seasonal forecasts.

Contributions

- ✓ Condition GLM weather generator on seasonal forecasts, trends.
- ✓ Translate coarse seasonal forecasts.
- ✓ Downscale regional climate model output.

CONTRIBUTIONS

✓ Condition GLM weather generator on seasonal forecasts, trends.

✓ Translate coarse seasonal forecasts.

✓ <u>Downscale</u> regional climate model output.

✓ Addressed the disconnect between data & decision.

Verdin et al., (2016). A conditional stochastic weather generator for seasonal to multi-decadal simulations. *Journal of Hydrology*.

doi:10.1016/j.jhydrol.2015.12.036

Table of Contents

3. **A generalized weather generator that can model and propagate uncertainty.**

Address non-treatment of uncertainty in weather generation.

"Why do I care about uncertainty?"
$\rightarrow$ Variability between ensemble members.
$\rightarrow$ Better estimates of crop production risk.

Bayesian methods widely used to quantify uncertainty.
$\rightarrow$ Bayesian model averaging *[Leamer, 1978]*.
$\rightarrow$ Coupling with weather generators *[Pezzulli et al., 2006; Hashmi et al., 2009]*.

Bayesian weather generator propagate uncertainty to simulations.

Consider the model from Chapter 2:

$$O(\boldsymbol{s}, t) = \mathbb{1}_{[W_O(\boldsymbol{s},t) \geq 0]} \tag{1}$$

$$W_O(\boldsymbol{s}, t) = \mathbf{X}_O(\boldsymbol{s}, t)'\boldsymbol{\beta}_O(\boldsymbol{s}) \tag{2}$$

$$A(\boldsymbol{s}, t) \sim Gamma(\alpha_A(\boldsymbol{s}), \alpha_A(\boldsymbol{s})/\mu_A(\boldsymbol{s}, t) \tag{3}$$

$$\mu_A(\boldsymbol{s}, t) = \exp(\mathbf{X}_A(\boldsymbol{s}, t)'\boldsymbol{\beta}_A) \tag{4}$$

$$Z_N(\boldsymbol{s}, t) = \mathbf{X}_N(\boldsymbol{s}, t)'\boldsymbol{\beta}_N(\boldsymbol{s}) + W_N(\boldsymbol{s}, t) \tag{5}$$

$$Z_X(\boldsymbol{s}, t) = \mathbf{X}_X(\boldsymbol{s}, t)'\boldsymbol{\beta}_X(\boldsymbol{s}) + W_X(\boldsymbol{s}, t) \tag{6}$$

Model structure

Now model parameters are treated as random variables.
Monte Carlo sampling produces distribution of model parameters.
→*Posterior distribution.*

Data layer:

$$O(\boldsymbol{s},t) = \mathbb{1}_{[W_O(\boldsymbol{s},t)\geq 0]} \tag{7}$$

$$A(\boldsymbol{s},t) \sim Gamma(\alpha_A(\boldsymbol{s}), \alpha_A(\boldsymbol{s})/\mu_A(\boldsymbol{s},t)) \tag{8}$$

$$Z_N(\boldsymbol{s},t) \sim GP(\mu_N(\boldsymbol{s},t), C_N(t)) \tag{9}$$

$$Z_X(\boldsymbol{s},t) \sim GP(\mu_X(\boldsymbol{s},t), C_X(t)) \tag{10}$$

*Priors centered on MLE.

Process layer:

$$W_O(\boldsymbol{s},t) \sim N(\mu_O(\boldsymbol{s},t), \sigma_O^2) \tag{11}$$

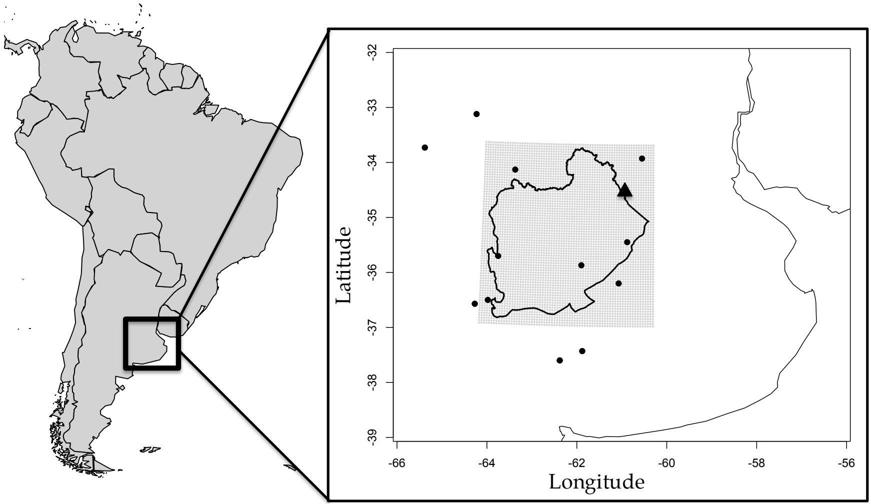
$$\mu_A(\boldsymbol{s},t) = \exp(\mathbf{X}_A(\boldsymbol{s},t)'\boldsymbol{\beta}_A(\boldsymbol{s})) \tag{12}$$

$$\mu_i(\boldsymbol{s},t) = \mathbf{X}_i(\boldsymbol{s},t)'\boldsymbol{\beta}_i(\boldsymbol{s}) \quad \text{for } i = O, N, X \tag{13}$$

$$\boldsymbol{\beta}_j(\boldsymbol{s}) \sim GP(\hat{\beta}_j(\boldsymbol{s}), C_{\beta_j}) \quad \text{for } j = O, A, N, X \tag{14}$$

$$\boldsymbol{\alpha}_A(\boldsymbol{s}) \sim GP(\hat{\boldsymbol{\alpha}}_A(\boldsymbol{s}), C_{\boldsymbol{\alpha}_A}) \tag{15}$$

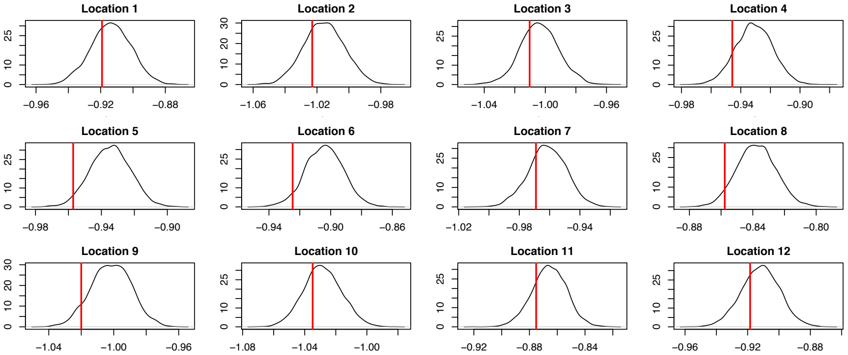
Study region & data



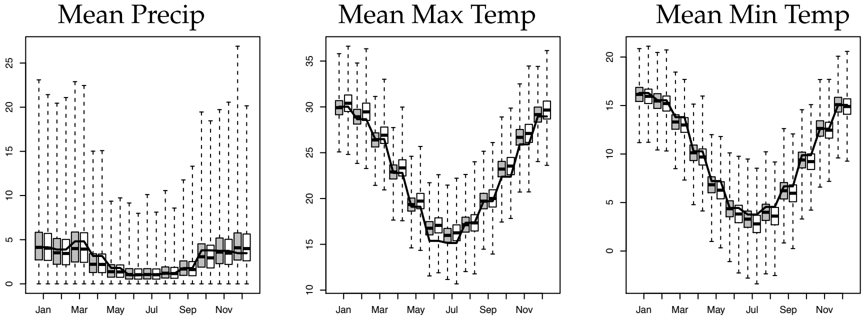
1961–2013 daily precipitation, maximum and minimum temperature.

Validate BayGEN by comparison to GLMGEN.

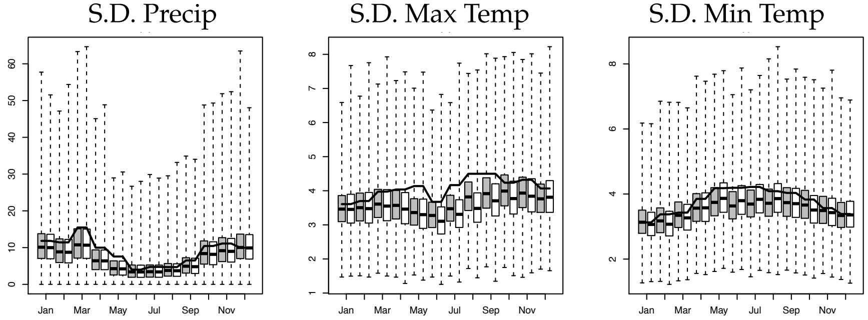
Posterior distribution



Posterior distribution and MLE of precipitation occurrence intercept: (β_0)



BayGEN as grey boxplots; GLMGEN as white boxplots; observed as lines.



BayGEN as grey boxplots; GLMGEN as white boxplots; observed as lines.

DSSAT

The Decision Support System for Agrotechnology Transfer

Required inputs:

- Land use and soil type
- Minimum, maximum temperatures (*daily*)
- Precipitation (*daily*)
- Solar radiation (*daily*)

Optional inputs:

- Dew point (*daily*)
- Wind speed (*daily*)

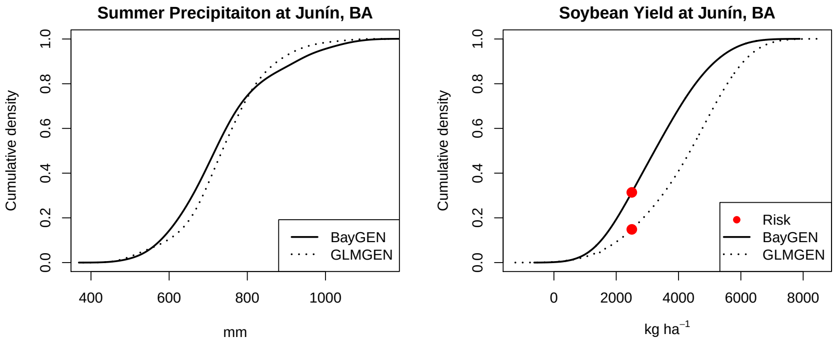
Outputs:

- Crop yield (*seasonal*)
- Flowering and harvest date
- Biomass (*daily*)
- Soil water content (*daily*)
- Root depth (*daily*)

[Jones et al., 1998]

Coupling with crop simulation model

100 simulations from BayGEN, GLMGEN:



"Break even" production risk: 30% (BayGEN), 16% (GLMGEN).

Nonlinear relationship between weather and crop yield.

CONTRIBUTIONS

✓ Development of Bayesian weather generator.
$\rightarrow$ Address non-treatment of uncertainty.

Contributions

✓ Development of Bayesian weather generator.
$\rightarrow$ Address non-treatment of uncertainty.

✓ "Cascade of uncertainty."

Contributions

- ✓ Development of Bayesian weather generator.
  $\rightarrow$ Address non-treatment of uncertainty.
- ✓ "Cascade of uncertainty."
- ✓ Uncertainty propagates to production risk.

Verdin et al., (2016, In Preparation). BayGEN: A Bayesian space-time stochastic weather generator. *Water Resources Research*.

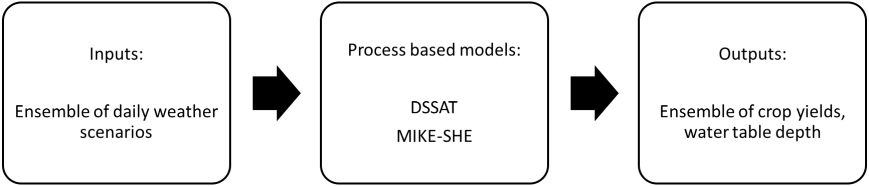
Table of Contents

Process based models

Agricultural planning in the Pampas:

- Crop simulation models (DSSAT).
- Analyze different management strategies.
- Correct crop yields with WTD (MIKE-SHE).

  → MIKE-SHE has great computational costs. (5km x 5km, daily)

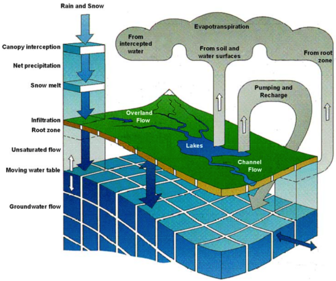


MIKE SHE

Rainfall-runoff model

- Deterministic
- Physically based
- Distributed
- Solves PDE for mass flow & momentum transfer.

MIKE by DHI (2012)



courtesy: DHI

Figure 1.1 Hydrologic processes simulated by MIKE SHE

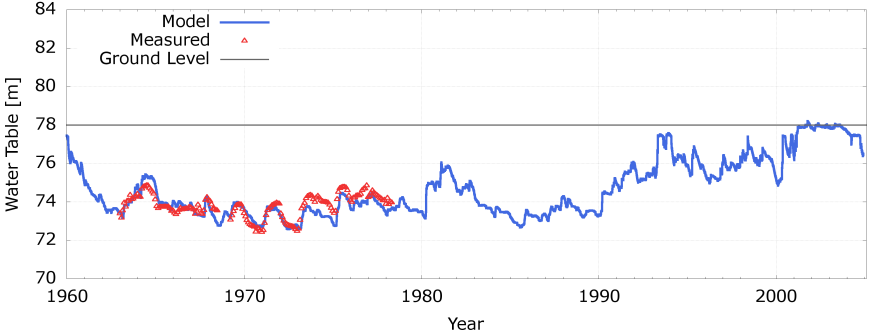
Inputs:

- Spatially-distributed land use (e.g., leaf-area index, root depth), soil data, precipitation, reference evapotranspiration, river networks

Outputs:

- Water table depth (WTD), real evapotranspiration (ET), overland flow, unsaturated zone flow, etc.

MIKE-SHE CALIBRATION



Motivating issues.

- Computationally expensive (5km x 5km, daily).
- Only 4 simulations at a time.
- Proprietary.
- For agricultural planning, interested in WTD.

4. **An efficient water table depth model for seasonal agricultural planning.**

Model WTD using weather and land use data?
$\rightarrow$ Reduce computational costs of distributed hydrologic models.

Propose stochastic hierarchical metamodel for *monthly* ET and WTD.

Level 1: ET modeled with land use and weather, PET.

Level 2: WTD modeled with ET and weather.

Assumptions:

Land use constant for calendar year (from database).

$\rightarrow$ Pasture, soybean, maize, wheat, soybean/wheat, sunflower.

Homogeneous soil type (i.e., not parameterized.)

Equally spaced "knot" locations $\boldsymbol{s}$. (20km x 20km)
$\rightarrow$MIKE-SHE grid cells. (5km x 5km)
$\rightarrow$Analog to stations.

$$Y_{ET}(\boldsymbol{s},t) = \mathbf{X}_{ET}(\boldsymbol{s},t)'\boldsymbol{\beta}_{ET}(\boldsymbol{s}) \tag{16}$$

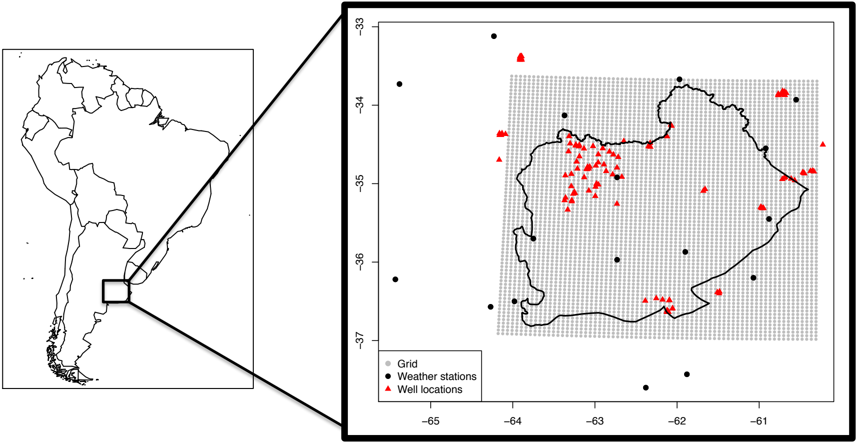
$$Y_{WT}(\boldsymbol{s},t) = \mathbf{X}_{WT}(\boldsymbol{s},t)'\boldsymbol{\beta}_{WT}(\boldsymbol{s}) + \epsilon \tag{17}$$

$$\epsilon \sim GP(0, C(t)) \tag{18}$$

$$\begin{aligned}\mathbf{X}_{ET}(\boldsymbol{s},t) = (&X_P(\boldsymbol{s},t), X_X(\boldsymbol{s},t), X_N(\boldsymbol{s},t), X_E(\boldsymbol{s},t), X_{L1}(\boldsymbol{s},t),\\ &X_{L2}(\boldsymbol{s},t), X_{L3}(\boldsymbol{s},t), X_{L4}(\boldsymbol{s},t), X_{L5}(\boldsymbol{s},t), X_{L6}(\boldsymbol{s},t))\end{aligned} \tag{19}$$

$$\mathbf{X}_{WT}(\boldsymbol{s},t) = (Y_{WT}(\boldsymbol{s},t-1), X_P(\boldsymbol{s},t), X_X(\boldsymbol{s},t), X_N(\boldsymbol{s},t), Y_{ET}(\boldsymbol{s},t)) \tag{20}$$

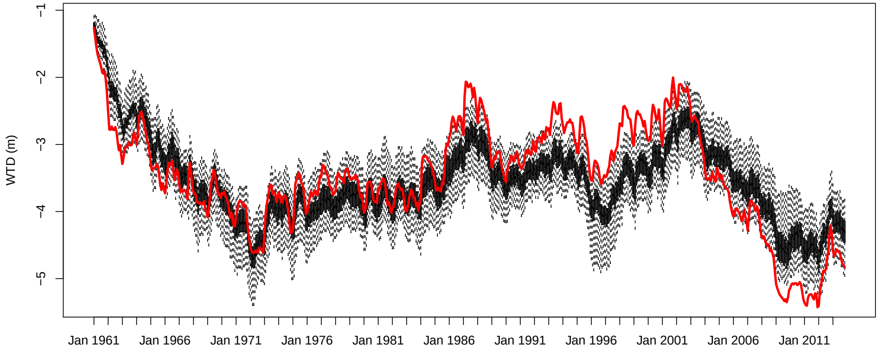
STUDY REGION & DATA



1961–2013 daily precipitation, maximum and minimum temperature, interpolated to MIKE-SHE grid; MIKE-SHE estimates of ET and WTD.

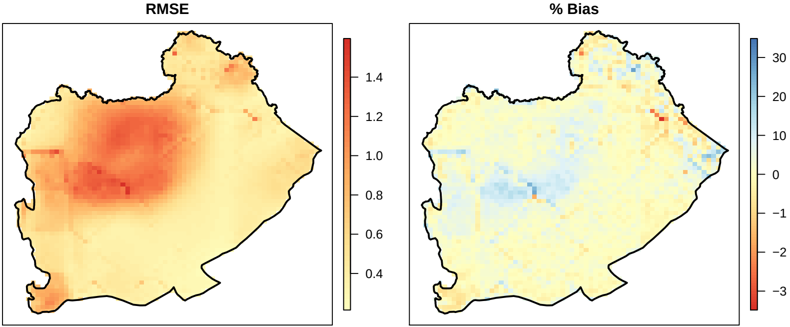
Basin average behavior

Jan 1961 – Dec 2013



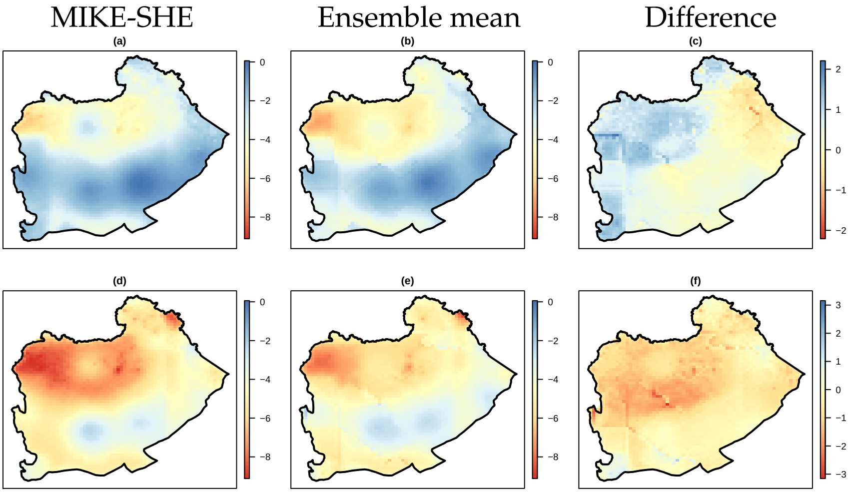
Ensemble of 100 simulations vs. MIKE-SHE for historic period.

Jan 1961 – Dec 2013

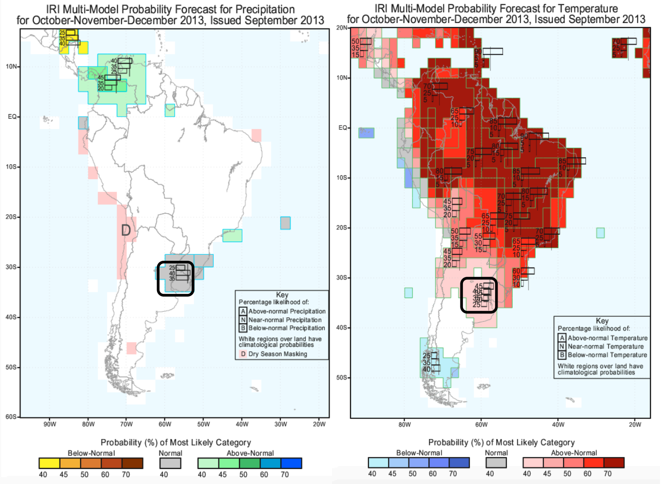


Ensemble mean vs. MIKE-SHE.

Wet & dry periods



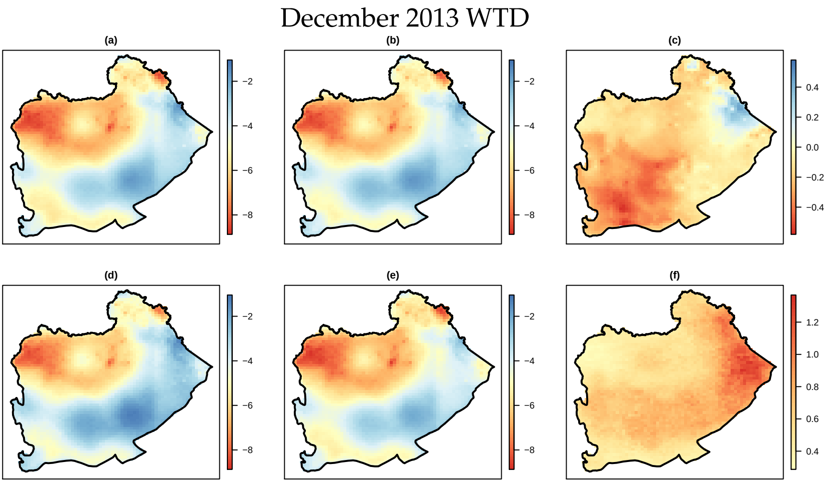
Top row: Dec 2000 (SE floods); Bottom row: Feb 2012 (NW depletion).



Probabilistic forecasts:

OND 2013 Precipitation: (25:40:35) (A:N:B)

OND 2013 Temperature: (40:35:25) (A:N:B)



(a) MIKE-SHE, (b) ensemble mean, (c) difference, (d) ensemble 97.5^{th} percentile, (e) 2.5^{th} percentile, (f) 95% ensemble spread.

Contributions

✓ Development of statistical metamodel for monthly WTD.
$\rightarrow$ Use in agricultural planning.

Contributions

- ✓ Development of statistical metamodel for monthly WTD.
  $\rightarrow$ Use in agricultural planning.
- ✓ Address computational costs of distributed hydrologic models.

CONTRIBUTIONS

✓ Development of statistical metamodel for monthly WTD.
$\rightarrow$ Use in agricultural planning.

✓ Address computational costs of distributed hydrologic models.

✓ Capture dominant mechanisms of WTD variability.

Contributions

- ✓ Development of statistical metamodel for monthly WTD.
  $\rightarrow$ Use in agricultural planning.
- ✓ Address computational costs of distributed hydrologic models.
- ✓ Capture dominant mechanisms of WTD variability.
- ✓ Downscaling seasonal forecasts (i.e., Chapter 3).

Verdin et al. (2016, In Preparation). A statistical metamodel for monthly groundwater fluctuations. *International Journal of Agricultural Management*.

Table of Contents

SUMMARY

✓ Developed GLM weather generator.

→ Spatial models on GLM coefficients.

→ Address problem of simulating at unobserved locations.

SUMMARY

✓ Developed GLM weather generator.

$\rightarrow$ Spatial models on GLM coefficients.

$\rightarrow$ Address problem of simulating at unobserved locations.

✓ Illustrated GLM weather generator for planning purposes.

$\rightarrow$ Translate seasonal forecasts; downscale projections.

$\rightarrow$ Address disconnect between data & decision.

Summary

✓ Developed GLM weather generator.
  → Spatial models on GLM coefficients.
  → Address problem of simulating at unobserved locations.

✓ Illustrated GLM weather generator for planning purposes.
  → Translate seasonal forecasts; downscale projections.
  → Address disconnect between data & decision.

✓ Developed Bayesian weather generator.
  → Quantify uncertainty neglected with traditional methods.
  → "Cascade of uncertainty" to decision space.

Summary

- ✓ Developed GLM weather generator.
  - $\rightarrow$ Spatial models on GLM coefficients.
  - $\rightarrow$ Address problem of simulating at unobserved locations.
- ✓ Illustrated GLM weather generator for planning purposes.
  - $\rightarrow$ Translate seasonal forecasts; downscale projections.
  - $\rightarrow$ Address disconnect between data & decision.
- ✓ Developed Bayesian weather generator.
  - $\rightarrow$ Quantify uncertainty neglected with traditional methods.
  - $\rightarrow$ "Cascade of uncertainty" to decision space.
- ✓ Developed statistical metamodel for hydrologic model.
  - $\rightarrow$ Use in agricultural planning, downscaling seasonal forecasts.
  - $\rightarrow$ Address computational costs of distributed hydrologic models.

CONTRIBUTIONS

- GLM space-time weather generator (GLMGEN).[1]

  → Simulate at unobserved locations.
- Conditional GLMGEN.[2]

  → Downscaling tool; useful for climate change studies.
- Bayesian space-time weather generator (BayGEN).[3]

  → New, novel, and a significant contribution to the literature.
- Statistical emulator for water table depth.[4]

  → Conditioned on climate, land use; downscale seasonal forecasts, make cropping decisions.

[1] Verdin et al., (2015). Coupled stochastic weather generation using spatial and generalized linear models. *Stochastic Environmental Research and Risk Assessment*, 29(2), 347–356.

[2] Verdin et al., (2016). A conditional stochastic weather generator for seasonal to multi-decadal simulations. *Journal of Hydrology*.

[3] Verdin et al., (2016, In Preparation). BayGEN: A Bayesian space-time stochastic weather generator. *Water Resources Research*.

[4] Verdin et al. (2016, In Preparation). A statistical metamodel for monthly groundwater fluctuations. *International Journal of Agricultural Management.*

FUTURE WORK

- Development of CRAN package for GLMGEN.
  $\rightarrow$ Conditional & unconditional simulation.
- Refinement of metamodel.
  $\rightarrow$ Consider LAI & root depth.
  $\rightarrow$ Finer scales if needed.
- Graduate??
- Get married!!

Thank you!

Committee:

- Balaji Rajagopalan
- Guillermo Podestá
- William Kleiber
- Ben Livneh
- Joseph Kasprzyk

NCAR folks:

- Rick Katz
- Linda Mearns
- Seth McGinnis

Buenos Aires team:

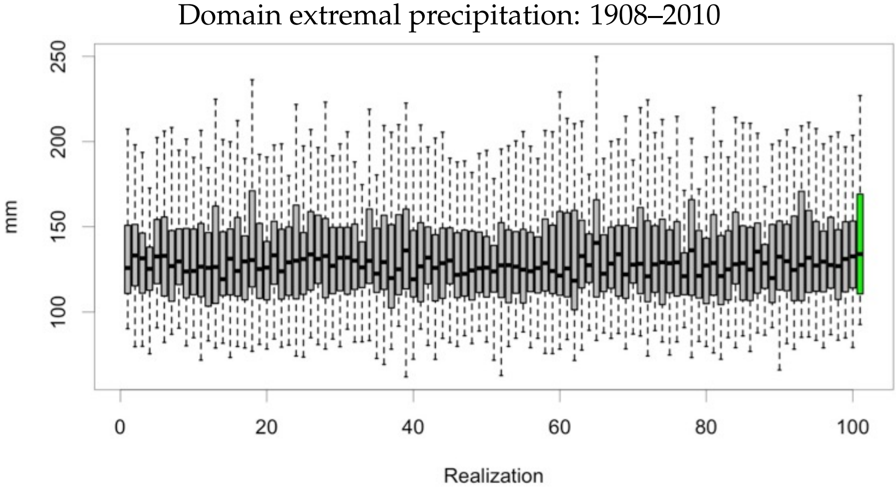
- Federico Bert
- Pablo García
- Angel Menéndez
- Santiago Rovere

Funding:

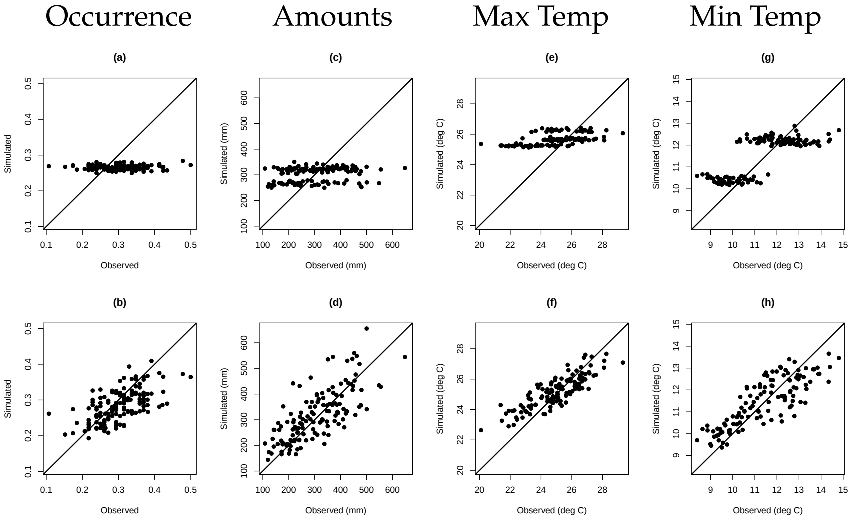
- NSF: The Dynamics of Coupled Natural and Human Systems (CNH) Program

Carsen, my family, my friends, …
and all I forgot to mention.

Additional slides


Domain extremal precipitation: 1908–2010
mm
250
200
150
100
0
20
40
60
80
100
Realization

Spatial validation: OND (1961–2013)



Define $\boldsymbol{\theta}$ as a vector of all model parameters:

$$\boldsymbol{\theta} = (\boldsymbol{\beta}_O, \boldsymbol{\beta}_A, \boldsymbol{\beta}_N, \boldsymbol{\beta}_X, C_X(t), C_N(t), C_{\boldsymbol{\beta}_O}, C_{\boldsymbol{\beta}_A}, C_{\boldsymbol{\beta}_X}, C_{\boldsymbol{\beta}_N}, \boldsymbol{\alpha}_A, C_{\boldsymbol{\alpha}_A}), \tag{21}$$

and $\mathbf{Y}$ as a vector of the data:

$$\mathbf{Y} = (O(\boldsymbol{s}, t), A(\boldsymbol{s}, t), Z_N(\boldsymbol{s}, t), Z_X(\boldsymbol{s}, t)). \tag{22}$$

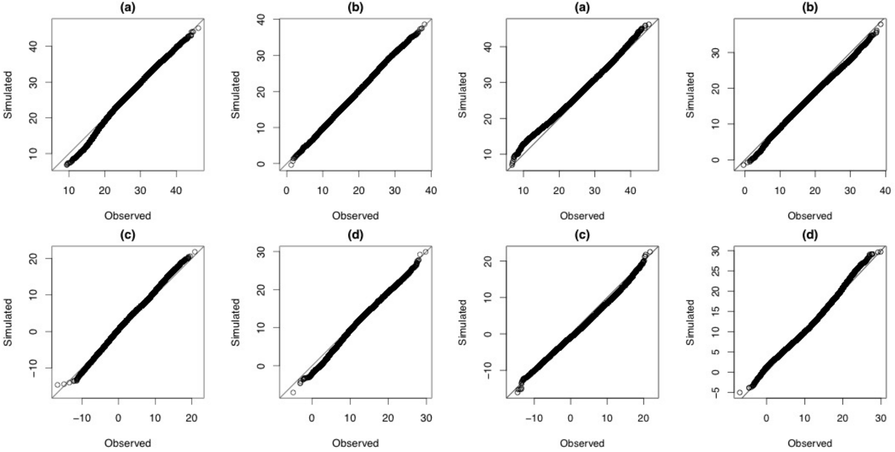
Seek the *posterior distribution*:

$$p(\boldsymbol{\theta}|\mathbf{Y}) = \frac{p(\mathbf{Y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{Y})} \propto p(\mathbf{Y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{23}$$

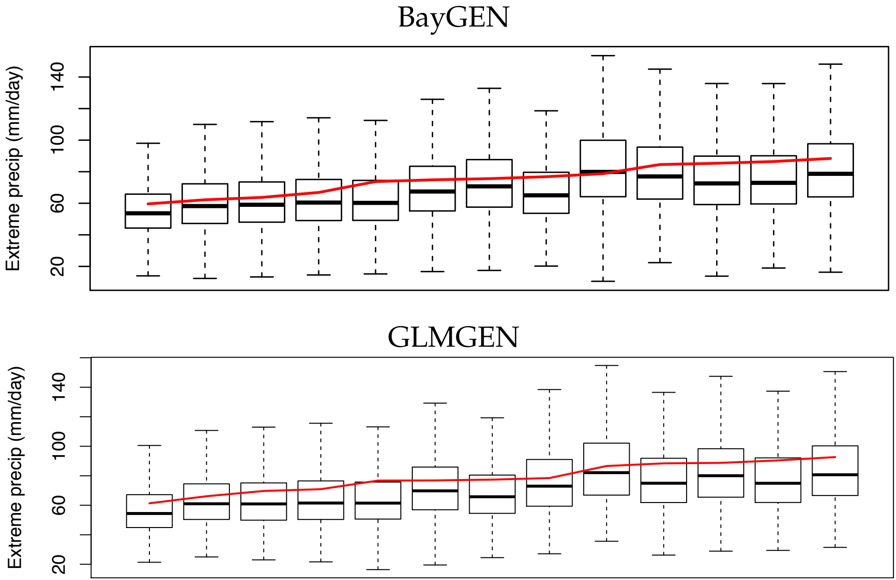
$$\tag{24}$$

$p(\mathbf{Y}|\boldsymbol{\theta})$ = likelihood function.
$p(\boldsymbol{\theta})$ = prior joint distribution.

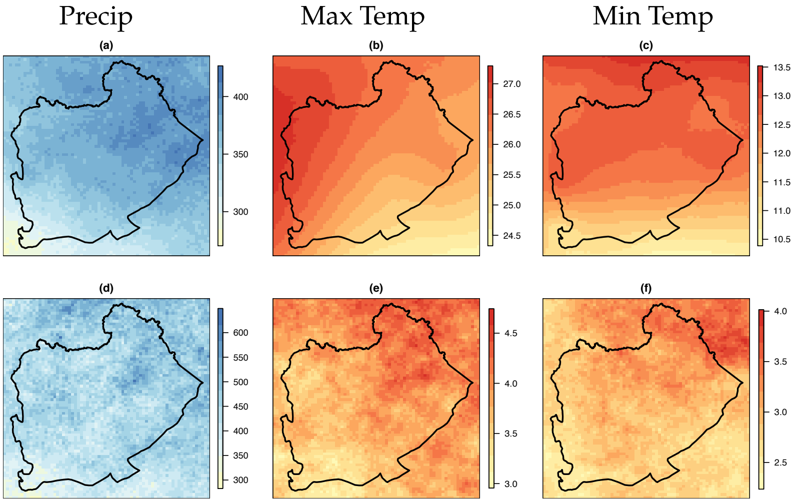
BayGEN GLMGEN



Domain daily extreme temperature.

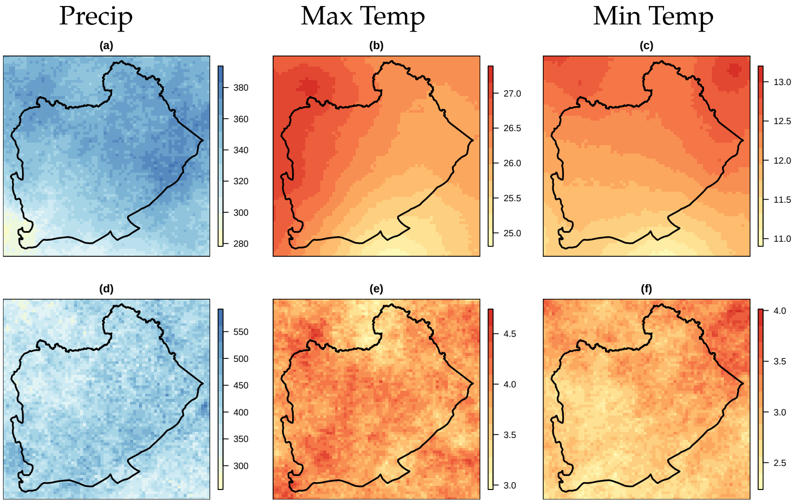

BayGEN
Extreme precip (mm/day)
20
60
100
140
GLMGEN
Extreme precip (mm/day)
20
60
100
140

BayGEN:

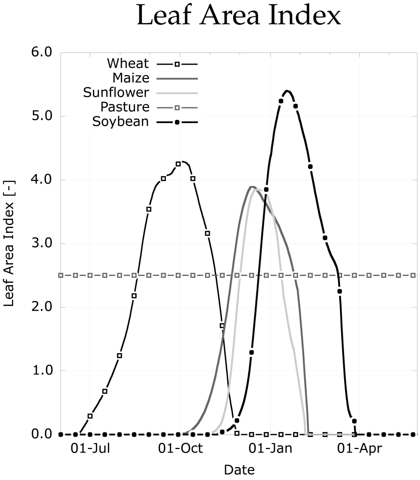


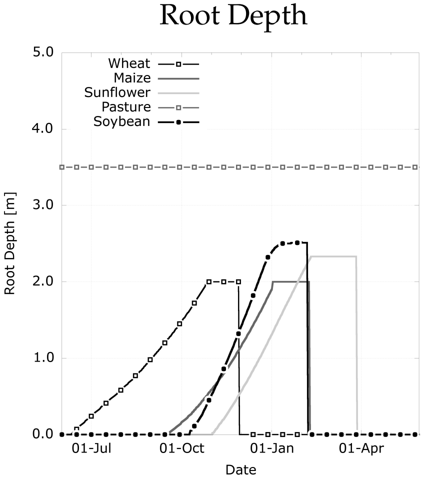
OND: Ensemble mean (top) and 95% ensemble spread (bottom).

GLMGEN:



OND: Ensemble mean (top) and 95% ensemble spread (bottom).


Leaf Area Index
Wheat
Maize
Sunflower
Pasture
Soybean
Leaf Area Index [-]
6.0
5.0
4.0
3.0
2.0
1.0
0.0
01-Jul
01-Oct
01-Jan
01-Apr
Date


Root Depth
Wheat
Maize
Sunflower
Pasture
Soybean
Root Depth [m]
5.0
4.0
3.0
2.0
1.0
0.0
01-Jul
01-Oct
01-Jan
01-Apr
Date

MIKE-SHE calibration