A conditional stochastic weather generator for seasonal to multi-decadal simulations

Andrew Verdin
University of Colorado, Boulder

February 22, 2016

Table of Contents

Overview
Motivation

Data

Methodology

Results

Summary

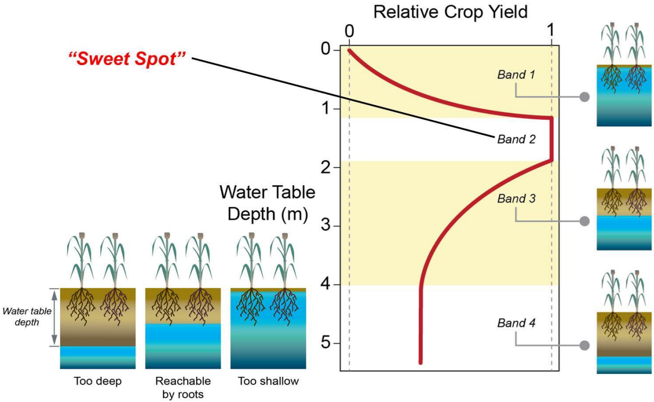
Motivation

"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

FLOODPLAIN HYDROLOGY



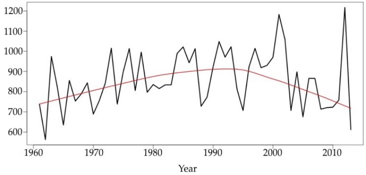
courtesy: Guillermo Podestá

PAMPAS CLIMATE

Alternating droughts and floods since colonial times:

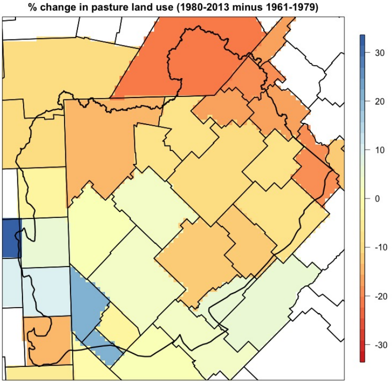
- floods during late 19th and early 20th centuries
- extensive droughts during 1930s-1950s
  - dry and windy conditions led to "dust bowl"

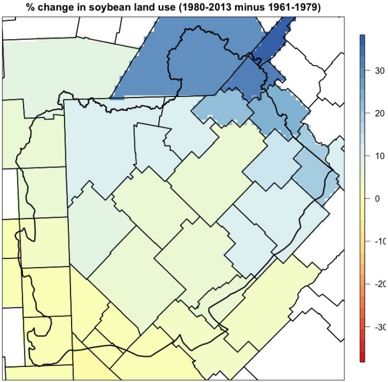
Total precipitation
(mm year^{-1})



- increasing precipitation from 1970-2000
- flooding in western Pampas from 1997-2003 left 27% of the landscape underwater, halved grain production, damaged infrastructure and soil quality
- rising water table attributed to land use change (perennial pasture for cattle vs. seasonal crops/fallow land for agriculture

LAND USE CHANGE

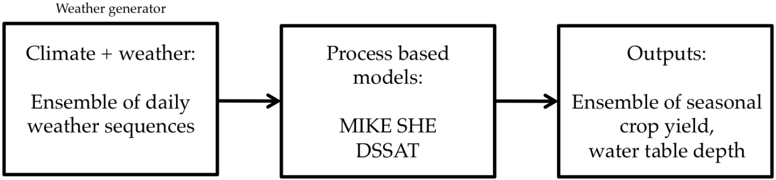




Risk management

Process based models useful to assess likely impacts on climate-sensitive sectors and evaluate outcomes of alternative adaptive actions.

- Risk based approaches using process based models require ensembles of *complete* input sequences.
- Historical daily weather often limited in space and time, and provide a solution based on only one realization of the weather process
- Climate information (seasonal forecasts, climate model projections) available, but generally coarse scale
- Stochastic weather generation!



Climate information

Scientific and technological advances, together with awareness of the importance of climate on human endeavors, are creating increased worldwide demand for climate information.

Potential outcomes of adaptation actions more relevant to stakeholders than raw climate information (e.g., 25, 35, and 40% chances of experiencing below-, near-, or above-normal conditions).

Global/regional climate models are generally coarse in space & time

To support public and private adaptation and mitigation responses, climate information must be credible and salient (e.g., relevant to the needs of decision makers.) - *Cash et al. (2003)*

Thus, an enhanced capacity to "translate" climate information into distributions of outcomes is needed for risk assessment and management. - *Hansen et al. (2006)*

Table of Contents

Overview

Data

Methodology

Results

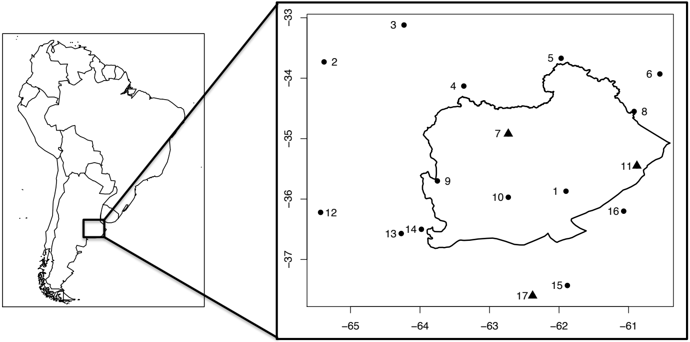
Summary

DATA

Network of 17 weather stations in and around Salado A basin

- Precipitation, minimum and maximum temperatures

Near complete record (1 January 1961 - 31 December 2013)



Data provided by the Met Service of Argentina.

Proposed grid

5km x 5km grid, Gauss-Krueger coordinates, (74 x 70) = 5180 pixels

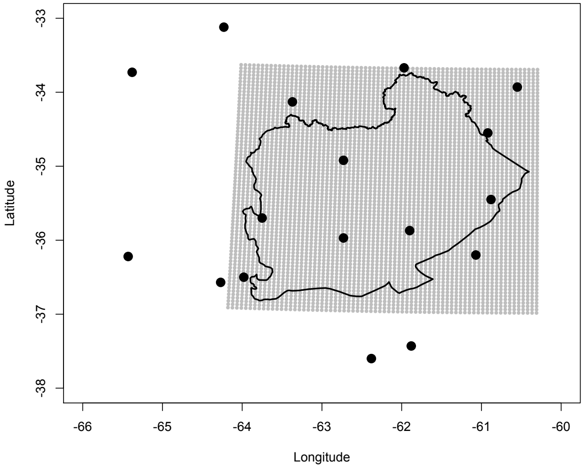


Table of Contents

Overview

Data

Methodology

Results

Summary

COUPLED STOCHASTIC WEATHER GENERATOR

Simultaneous simulation of precipitation and temperature at arbitrary locations (i.e., on a grid)

Generalized linear models (GLMs) ease the modeling of non-normal variables, can include any number of relevant covariates.

Spatial process models on GLM parameters alleviate efforts in simulating gridded weather sequences.

Fitting GLM at each location captures local (temporal) features – spatial process models on GLM parameters capture spatial features.

See *Verdin et al. (2015)* for more details.

MODEL STRUCTURE

Precipitation is decomposed into occurrence and amounts. Occurrence is modeled using probit regression:

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

The latent process for occurrence is a local regression where

$$\mathbf{X}_O = (1, \cos(2\pi t/365), \sin(2\pi t/365), O(\boldsymbol{s}, t-1), \ldots)' \tag{2}$$

Amounts are modeled as a gamma random variable:

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

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

The gamma parameters are estimated using a transformed gamma local regression where

$$\mathbf{X}_A = (1, \cos(2\pi t/365), \sin(2\pi t/365), \ldots)' \tag{5}$$

Model structure

Temperatures are decomposed as follows:

$$Z_N(\boldsymbol{s}, t) = \beta_N(\boldsymbol{s})'\mathbf{X}_N(\boldsymbol{s}, t) + W_N(\boldsymbol{s}, t) \tag{6}$$
$$Z_X(\boldsymbol{s}, t) = \beta_X(\boldsymbol{s})'\mathbf{X}_X(\boldsymbol{s}, t) + W_X(\boldsymbol{s}, t) \tag{7}$$

The first component is a local regression on some covariate vector $\mathbf{X}_i(\boldsymbol{s}, t)$, which represents the expected climate at a point.

Covariate vectors for Z_N and Z_X are identical and defined as

$$\begin{aligned}\mathbf{X}_i(\boldsymbol{s}, t) = (&1, \cos(2\pi t/365), \sin(2\pi t/365), r(t),\\ &Z_N(\boldsymbol{s}, t-1), Z_X(\boldsymbol{s}, t-1), O(\boldsymbol{s}, t), \dots)'\end{aligned} \tag{8}$$

The weather component (denoted by $W_i(\boldsymbol{s}, t)$ for weather) generates variability and spatial correlation via mean-zero multivariate normal Gaussian processes.

Parameter estimation & simulation

Four GLMs are fitted at each location.

Obtain maximum likelihood estimates.

Covariance parameters for $W_i(\boldsymbol{s}, t)$ estimated from residuals.

Simulated daily weather obtained from covariate vectors, MLE parameters, and $W_i(\boldsymbol{s}, t)$.

First occurrence, then amounts, then temperatures (conditional on occurrence and previous day's temperatures).

CONDITIONAL WEATHER GENERATOR

GLM framework allows modification to model structure.

Areal total precipitation included in occurrence and amounts models.

$$X_O = (1, \cos(2\pi t/365), \sin(2\pi t/365), O(\boldsymbol{s}, t-1), ST1(t), ST2(t), ST3(t), ST4(t))' \quad (9)$$

$$X_A = (1, \cos(2\pi t/365), \sin(2\pi t/365), ST1(t), ST2(t), ST3(t), ST4(t))' \quad (10)$$

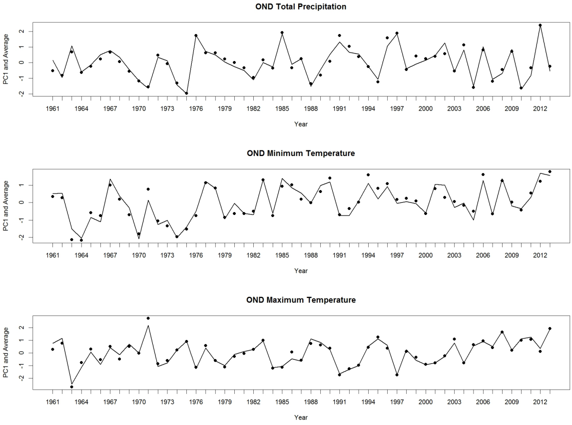
Areal average minimum & maximum temperature included in both temperature models.

$$\mathbf{X}_i(\boldsymbol{s}, t) = (1, \cos(2\pi t/365), \sin(2\pi t/365), r(t), Z_N(\boldsymbol{s}, t-1), Z_X(\boldsymbol{s}, t-1), O(\boldsymbol{s}, t),$$

$$SMN1(t), SMN2(t), SMN3(t), SMN4(t), SMX1(t), SMX2(t), SMX3(t), SMX4(t))' \quad (11)$$

PRINCIPAL COMPONENT ANALYSIS

Principal component analysis on seasonal weather justifies areal covariates.



All areal covariates are significant at 99% level.

Table of Contents

Overview

Data

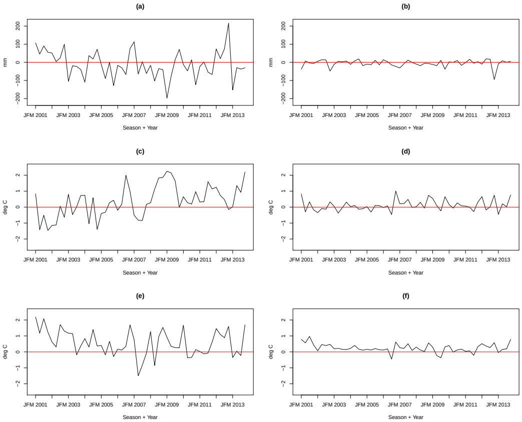
Methodology

Results

Summary

TEMPORAL VALIDATION

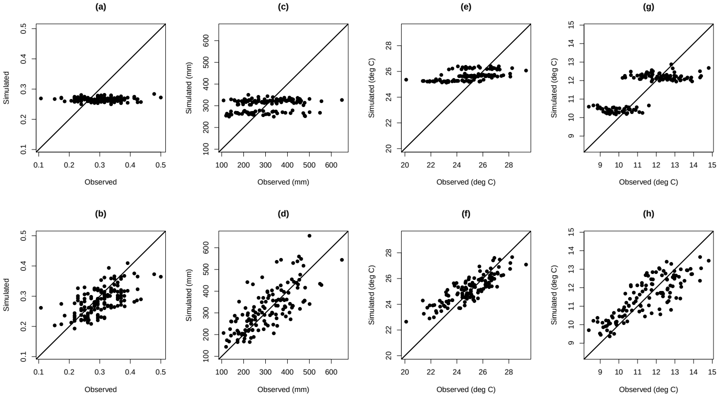
JFM 2001 – OND 2013 (simulated minus observed)



Ensemble mean precip RMSE reduced from 77 to 21 mm (a-b); max (c-d) and min (e-f) temp RMSEs reduced from 1.05° to 0.37° and 0.99° to 0.37°C.

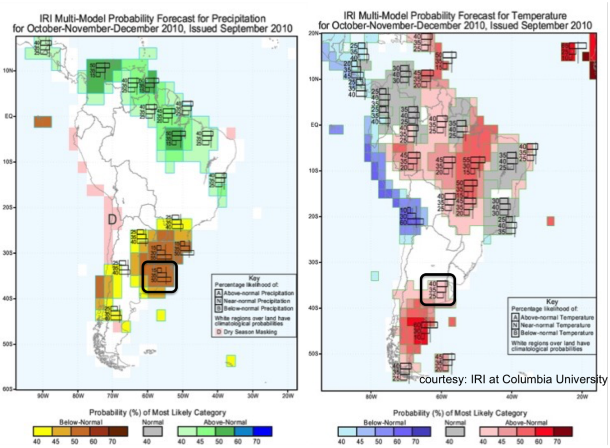
Spatial validation

OND 1961 – OND 2013 (three withheld stations)



Spatial validation: (a-b) OND 1961-2013 observed versus ensemble mean simulated probability of precipitation occurrence, (c-d) total precipitation, (e-f) mean maximum temperature, and (g-h) mean minimum temperature.

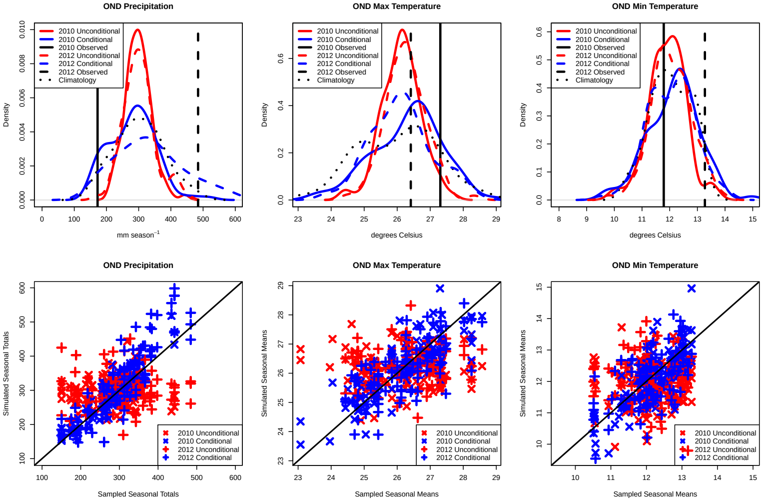
IRI SEASONAL FORECASTS



Probabilistic forecasts:

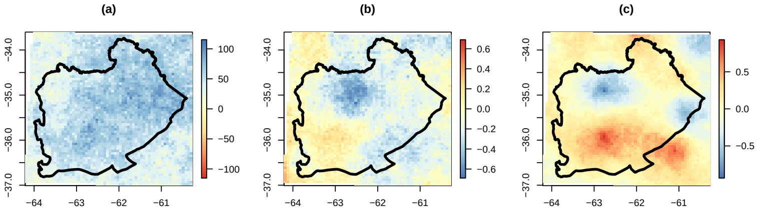
OND 2010 Precipitation	15:35:50 (A:N:B)
OND 2010 Temperature	40:35:25 (A:N:B)

IRI SEASONAL FORECASTS

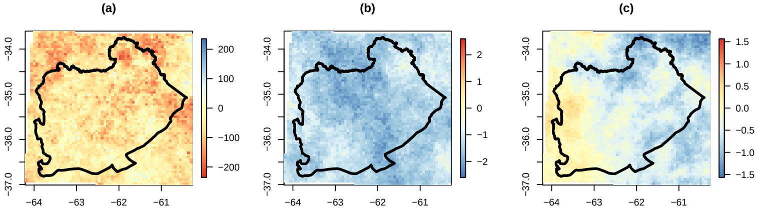


IRI seasonal forecasts

Difference in ensemble mean:

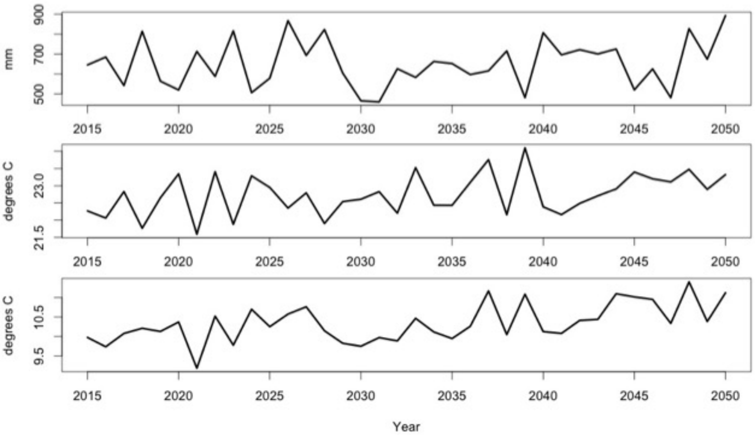


Difference in 95% ensemble spread:

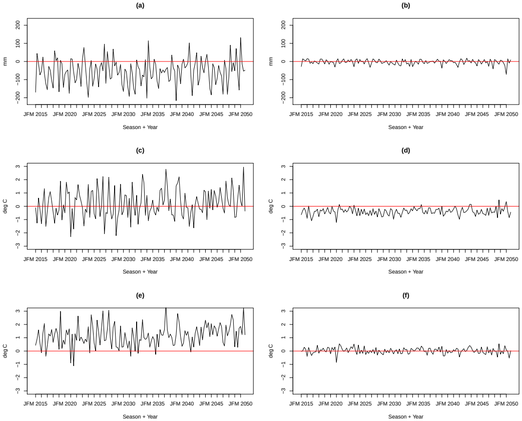


MULTI-DECADAL PROJECTIONS

CORDEX-CMIP5 RCM (2015-2050)



CORDEX-CMIP5 RCM (2015-2050)



Ensemble mean precip RMSE reduced from 90 to 14 mm (a-b); max (c-d) and min (e-f) temp RMSEs reduced from 1.09° to 0.48° and 1.43° to 0.23°C.

TABLE OF CONTENTS

OVERVIEW

DATA

METHODOLOGY

RESULTS

SUMMARY

Summary

- GLM framework for stochastic weather generator allows the inclusion of conditional covariates
- Sub-basin scale validates use of areal average covariates, propagates to local scale
  - Non-homogeneous regions should consider clustering or site-specific covariates
- Seasonal forecasts and multi-decadal projections may be translated to daily sequences, subsequently used to drive crop simulation, hydrologic models
  - Translated to more salient information for decision making

Thank you!

Ongoing and future work

- Identify teleconnections (i.e., ENSO, PDO, AMO) for nonstationary weather generation
  - Current model framework enables inclusion of relevant covariates
  - Strong ENSO signal for crop yield, may be relevant for weather scenarios
- Development of Bayesian weather generator
  - Account for spatial correlation of model parameters
  - Propagate uncertainty in parameter estimates to simulation space
  - Implemented in Stan modeling language
- Model response of water table to climatic inputs
  - Water table depth has strong influence on seasonal crop yield

Funding

National Science Foundation (NSF)

The Dynamics of Coupled Natural and Human Systems (CNH) Program promotes interdisciplinary analyses of relevant human and natural system processes and complex interactions among human and natural systems at diverse scales.