PRICING AREA YIELD CROP INSURANCE CONTRACTS
SPATIO-TEMPORAL APPROACHES

Conference 2

1º Workshop do projeto PROCAD:
Seguro Agrícola: Modelagem Estatística e Precificação

UFMG - Belo Horizonte, Novembro 25-27 de 2009

This is joint work with:

- **Camilo Teixeira de Andrade** (EMBRAPA - Milho)
- **Elias Teixeira Krainski** (UFPR)
- **Juan Carlos Vivar Rojas** (Duke Univ.)
- **Juliana Garcia Céspedes** (UNIFEI)
- **Rogério Teixeira de Faria** (IAPAR)
- **Vitor Augusto Ozaki** (ESALQ/USP)

Outline of the talk

- Motivation
- Crop Insurance
    - main problems
    - recent alternatives
- Spatio-temporal modelling
    - main challenges
    - strategies
- Some results
- Future work

Outline of the talk

- Motivation
- Crop Insurance
    - main problems
    - recent alternatives
- Spatio-temporal modelling
    - main challenges
    - strategies
- Some results
- Future work

Outline of the talk

- Motivation
- Crop Insurance
  - main problems
  - recent alternatives
- Spatio-temporal modelling
  - main challenges
  - strategies
- Some results
- Future work

Outline of the talk

- Motivation
- Crop Insurance
  - main problems
  - recent alternatives
- Spatio-temporal modelling
  - main challenges
  - strategies
- Some results
- Future work

Outline of the talk

- Motivation
- Crop Insurance
  - main problems
  - recent alternatives
- Spatio-temporal modelling
  - main challenges
  - strategies
- Some results
- Future work

Traditional crop insurance

Main Problems:

- **lack of methods to properly quantify agricultural risk,**
- **inadequate pricing techniques,**
- **insufficient sources of data,**
- **systemic nature of the risk,**
- **information asymmetries (moral hazard, adverse selection),**
- **greater ruin probability.**

Traditional crop insurance

Main Problems:

- **lack of methods to properly quantify agricultural risk,**
- **inadequate pricing techniques,**
- **insufficient sources of data,**
- **systemic nature of the risk,**
- **information asymmetries (moral hazard, adverse selection),**
- **greater ruin probability.**

Traditional crop insurance

Main Problems:

- **lack of methods to properly quantify agricultural risk,**
- **inadequate pricing techniques,**
- **insufficient sources of data,**
- **systemic nature of the risk,**
- **information asymmetries (moral hazard, adverse selection),**
- **greater ruin probability.**

Traditional crop insurance

Main Problems:

- **lack of methods to properly quantify agricultural risk,**
- **inadequate pricing techniques,**
- **insufficient sources of data,**
- **systemic nature of the risk,**
- **information asymmetries (moral hazard, adverse selection),**
- **greater ruin probability.**

Traditional crop insurance

Main Problems:

- lack of methods to properly quantify agricultural risk,
- inadequate pricing techniques,
- insufficient sources of data,
- systemic nature of the risk,
- information asymmetries (moral hazard, adverse selection),
- greater ruin probability.

Traditional crop insurance

Main Problems:

- **lack of methods to properly quantify agricultural risk,**
- **inadequate pricing techniques,**
- **insufficient sources of data,**
- **systemic nature of the risk,**
- **information asymmetries (moral hazard, adverse selection),**
- **greater ruin probability.**

Information assimetries

Moral hazard:

- Farmers change their behavior in a way that increases the risks beyond what the insurer believed they would be when the insurance was developed.

Adverse selection:

- It occurs when an individual's demand for insurance is positively correlated with the individual's risk of loss, and the insurer is unable to allow for this correlation in the price of insurance.

Index based crop insurance contracts

Two main types:

- **area yield insurance:**

  farmers collect an indemnity whenever the county average yield falls beneath a yield guarantee, regardless of the farmers' actual yields.

- **weather based insurance:**

  It is based on the events of a weather variable measured at a given location. The payoff is based on the difference between an underlying weather index over a specified period and an agreed strike value.

Index based crop insurance contracts

Two main types:

- **area yield insurance:**

  farmers collect an indemnity whenever the county average yield falls beneath a yield guarantee, regardless of the farmers' actual yields.

- **weather based insurance:**

  It is based on the events of a weather variable measured at a given location. The payoff is based on the difference between an underlying weather index over a specified period and an agreed strike value.

Index based crop insurance contracts

Main features:

- ✓ their payoff depends on values of an index, related to the risk being hedged against
- ✓ indexes can be measured objectively (do not depend on individual actions)
- ✓ loss adjustment costs at individual level are eliminated
- ✓ Their risk-reducing potential depends strongly on the extent to which individual farmers are affected by *basis risk*

Index based crop insurance contracts

Main features:

- ✓ their payoff depends on values of an index, related to the risk being hedged against
- ✓ indexes can be measured objectively (do not depend on individual actions)
- ✓ loss adjustment costs at individual level are eliminated
- ✓ Their risk-reducing potential depends strongly on the extent to which individual farmers are affected by *basis risk*

Index based crop insurance contracts

Main features:

- ✓ their payoff depends on values of an index, related to the risk being hedged against
- ✓ indexes can be measured objectively (do not depend on individual actions)
- ✓ loss adjustment costs at individual level are eliminated
- ✓ Their risk-reducing potential depends strongly on the extent to which individual farmers are affected by *basis risk*

Index based crop insurance contracts

Main features:

- ✓ their payoff depends on values of an index, related to the risk being hedged against
- ✓ indexes can be measured objectively (do not depend on individual actions)
- ✓ loss adjustment costs at individual level are eliminated
- ✓ Their risk-reducing potential depends strongly on the extent to which individual farmers are affected by *basis risk*

Spatio-temporal modelling of crop yield

Main challenges:

- **crop yield estimates in some regions have a considerable error**
- **short length of time series**
- **missing values are very frequent**
- **there are change of support problems**
- **crop yield statistics are released with a lag of two years**
- **the number of areas is not constant through time**
- **planting date interval may be large ⇒ weak correlation covariates : crop yield**
- **seasonal forecasting (some months ahead) for climatic covariates is needed**
- **computational methods are time consuming.**

Spatio-temporal modelling of crop yield

Main challenges:

- crop yield estimates in some regions have a considerable error
- short length of time series
- missing values are very frequent
- there are change of support problems
- crop yield statistics are released with a lag of two years
- the number of areas is not constant through time
- planting date interval may be large ⇒ weak correlation covariates : crop yield
- seasonal forecasting (some months ahead) for climatic covariates is needed
- computational methods are time consuming.

Spatio-temporal modelling of crop yield

Main challenges:

- **crop yield estimates in some regions have a considerable error**
- **short length of time series**
- **missing values are very frequent**
- **there are change of support problems**
- **crop yield statistics are released with a lag of two years**
- **the number of areas is not constant through time**
- **planting date interval may be large ⇒ weak correlation covariates : crop yield**
- **seasonal forecasting (some months ahead) for climatic covariates is needed**
- **computational methods are time consuming.**

Spatio-temporal modelling of crop yield

Main challenges:

- **crop yield estimates in some regions have a considerable error**
- **short length of time series**
- **missing values are very frequent**
- **there are change of support problems**
- **crop yield statistics are released with a lag of two years**
- **the number of areas is not constant through time**
- **planting date interval may be large ⇒ weak correlation covariates : crop yield**
- **seasonal forecasting (some months ahead) for climatic covariates is needed**
- **computational methods are time consuming.**

Spatio-temporal modelling of crop yield

Main challenges:

- **crop yield estimates in some regions have a considerable error**
- **short length of time series**
- **missing values are very frequent**
- **there are change of support problems**
- **crop yield statistics are released with a lag of two years**
- **the number of areas is not constant through time**
- **planting date interval may be large ⇒ weak correlation covariates : crop yield**
- **seasonal forecasting (some months ahead) for climatic covariates is needed**
- **computational methods are time consuming.**

Spatio-temporal modelling of crop yield

Main challenges:

- **crop yield estimates in some regions have a considerable error**
- **short length of time series**
- **missing values are very frequent**
- **there are change of support problems**
- **crop yield statistics are released with a lag of two years**
- **the number of areas is not constant through time**
- **planting date interval may be large ⇒ weak correlation covariates : crop yield**
- **seasonal forecasting (some months ahead) for climatic covariates is needed**
- **computational methods are time consuming.**

Spatio-temporal modelling of crop yield

Main challenges:

- crop yield estimates in some regions have a considerable error
- short length of time series
- missing values are very frequent
- there are change of support problems
- crop yield statistics are released with a lag of two years
- the number of areas is not constant through time
- planting date interval may be large ⇒ weak correlation covariates : crop yield
- seasonal forecasting (some months ahead) for climatic covariates is needed
- computational methods are time consuming.

Spatio-temporal modelling of crop yield

Main challenges:

- crop yield estimates in some regions have a considerable error
- short length of time series
- missing values are very frequent
- there are change of support problems
- crop yield statistics are released with a lag of two years
- the number of areas is not constant through time
- planting date interval may be large ⇒ weak correlation covariates : crop yield
- seasonal forecasting (some months ahead) for climatic covariates is needed
- computational methods are time consuming.

Spatio-temporal modelling of crop yield

Main challenges:

- crop yield estimates in some regions have a considerable error
- short length of time series
- missing values are very frequent
- there are change of support problems
- crop yield statistics are released with a lag of two years
- the number of areas is not constant through time
- planting date interval may be large ⇒ weak correlation covariates : crop yield
- seasonal forecasting (some months ahead) for climatic covariates is needed
- computational methods are time consuming.

Spatio-temporal modelling of crop yield

Main challenges:

- **crop yield estimates in some regions have a considerable error**
- **short length of time series**
- **missing values are very frequent**
- **there are change of support problems**
- **crop yield statistics are released with a lag of two years**
- **the number of areas is not constant through time**
- **planting date interval may be large $\Rightarrow$ weak correlation covariates : crop yield**
- **seasonal forecasting (some months ahead) for climatic covariates is needed**
- **computational methods are time consuming.**

Spatio-temporal modelling of crop yield

Modelling strategies:

- a Bayesian hierarchical approach
- a Bayesian dynamic approach

Covariates:

- drought indexes
- change of technology indexes

Spatio-temporal modelling of crop yield

Modelling strategies:

- a Bayesian hierarchical approach
- a Bayesian dynamic approach

Covariates:

- drought indexes
- change of technology indexes

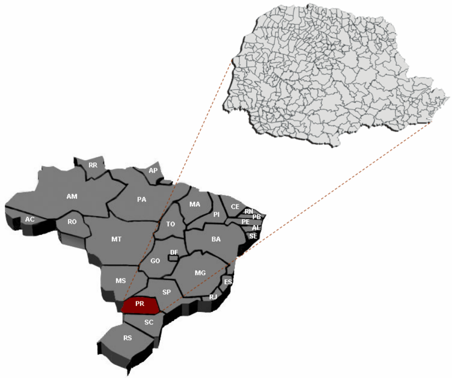
Study region and available data sets

- Crop yield data:
  average annual county yield (1980 – 2007).
  source: IBGE / SEAB

- Meteorological data:

  daily precipitation series for 503 stations (01/01/76 – 31/12/08).
  source: ANA / SUDHERSA / IAPAR / SIMEPAR / INMET

  daily temperature series for 87 stations (01/01/76 – 31/12/08).
  source: INMET / IAPAR / SIMEPAR



State: Paraná
N^{o} counties: 399
planted area (grains): 8.45 mill Ha

Solving the change of support problem

- Let $\{\boldsymbol{D}(g) : g \in G\}$, with $G \in \Re^2$, be a spatial random field, with g describing the latitude and longitude of a point within the region of interest.

- for a block (county) $B \subset G$,
$$\boldsymbol{D}(B) = \mid B \mid^{-1} \int_B \boldsymbol{D}(g) dg \tag{1}$$

  where $|B|$ represents the area of B.

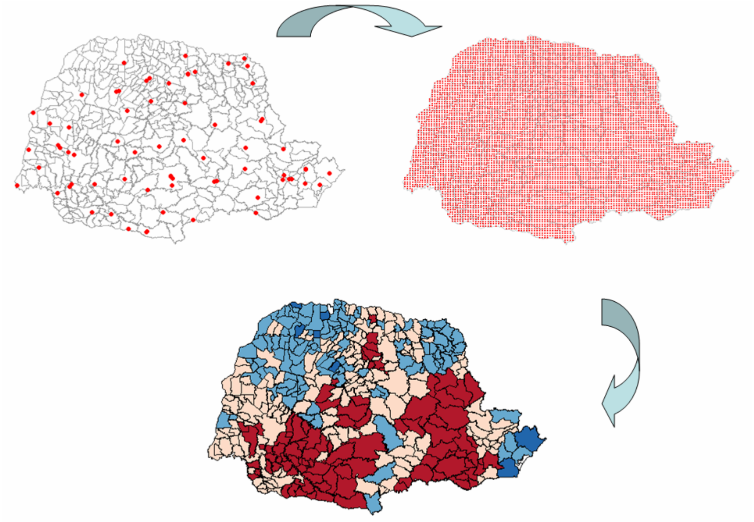
- predictions are obtained by approximating the integral in (1) by a sum, using standard interpolation techniques.

- areal predicted values are estimated as the mean of the predicted values at points of a dense grid that fall within each county s.

- the procedure is independently repeated at each time point, $t = 1, \cdots, T$.

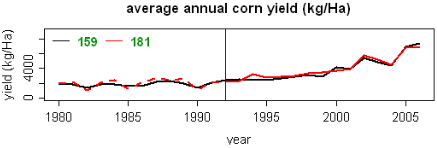
Solving the change of support problem

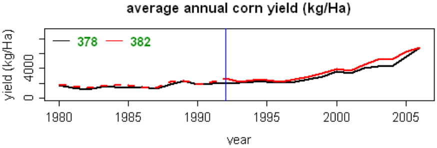
Interpolation (GAM, Geostatistics)

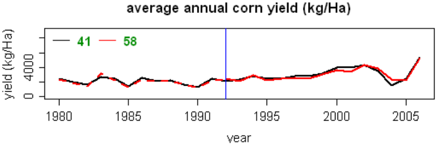


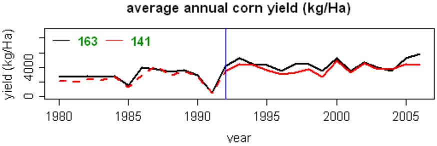
Recovering the crop yield time series

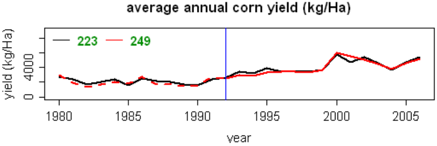
- 109 counties were created between 1983 and 1997 from existing ones.

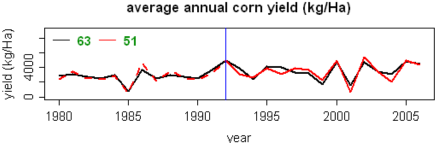

average annual corn yield (kg/Ha)


average annual corn yield (kg/Ha)


average annual corn yield (kg/Ha)

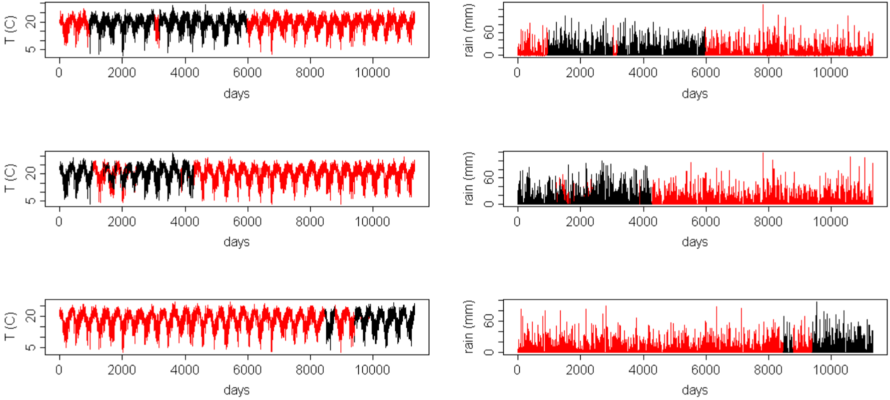

average annual corn yield (kg/Ha)


average annual corn yield (kg/Ha)

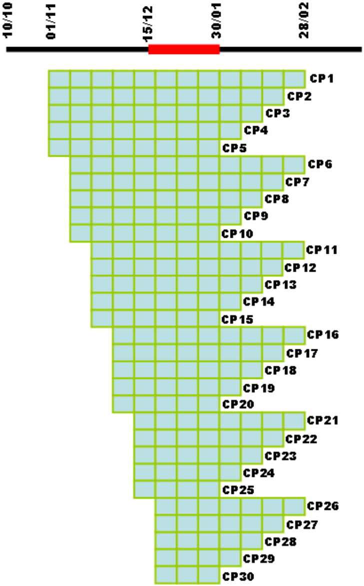

average annual corn yield (kg/Ha)

Imputing the missing values

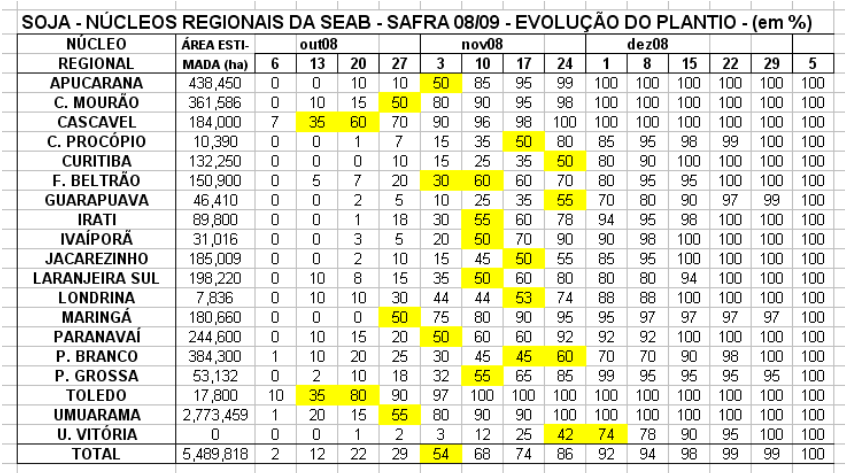
- We use a modified version of the EM algorithm for imputation of missing values in multivariate time series (Junger et al, 2003)
- The temporal pattern was modeled with a cubic spline.



Critical periods



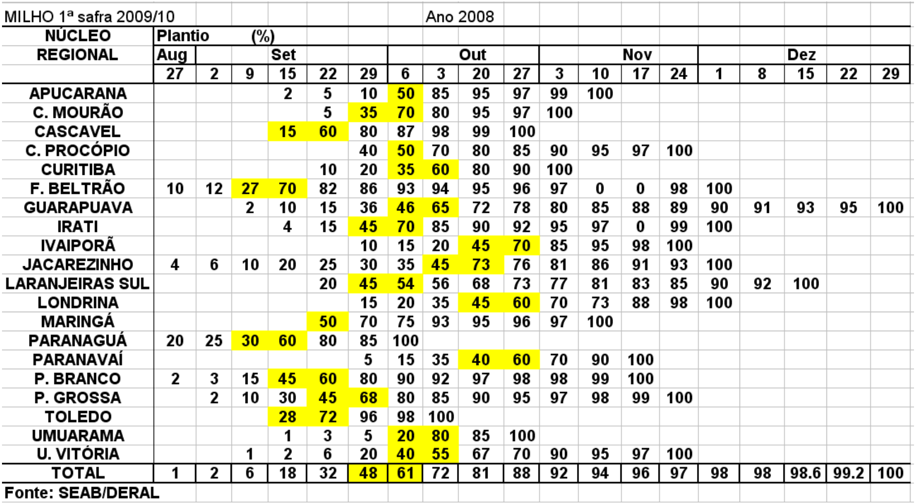
evolution of planting (soybean)



SOJA - NÚCLEOS REGIONAIS DA SEAB - SAFRA 08/09 - EVOLUÇÃO DO PLANTIO - (em %)

NÚCLEO	ÁREA ESTI-	out08				nov08				dez08					
REGIONAL	MADA (ha)	6	13	20	27	3	10	17	24	1	8	15	22	29	5
APUCARANA	438,450	0	0	10	10	50	85	95	99	100	100	100	100	100	100
C. MOURÃO	361,586	0	10	15	50	80	90	95	98	100	100	100	100	100	100
CASCAVEL	184,000	7	35	60	70	90	96	98	100	100	100	100	100	100	100
C. PROCÓPIO	10,390	0	0	1	7	15	35	50	80	85	95	98	99	100	100
CURITIBA	132,250	0	0	0	10	15	25	35	50	80	90	100	100	100	100
F. BELTRÃO	150,900	0	5	7	20	30	60	60	70	80	95	95	100	100	100
GUARAPUAVA	46,410	0	0	2	5	10	25	35	55	70	80	90	97	99	100
IRATI	89,800	0	0	1	18	30	55	60	78	94	95	98	100	100	100
IVAÍPORÃ	31,016	0	0	3	5	20	50	70	90	90	98	100	100	100	100
JACAREZINHO	185,009	0	0	2	10	15	45	50	55	85	95	100	100	100	100
LARANJEIRA SUL	198,220	0	10	8	15	35	50	60	80	80	80	94	100	100	100
LONDRINA	7,836	0	10	10	30	44	44	53	74	88	88	100	100	100	100
MARINGÁ	180,660	0	0	0	50	75	80	90	95	95	97	97	97	97	100
PARANAVAÍ	244,600	0	10	15	20	50	60	60	92	92	92	100	100	100	100
P. BRANCO	384,300	1	10	20	25	30	45	45	60	70	70	90	98	100	100
P. GROSSA	53,132	0	2	10	18	32	55	65	85	99	95	95	95	95	100
TOLEDO	17,800	10	35	80	90	97	100	100	100	100	100	100	100	100	100
UMUARAMA	2,773,459	1	20	15	55	80	90	90	100	100	100	100	100	100	100
U. VITÓRIA	0	0	0	1	2	3	12	25	42	74	78	90	95	100	100
TOTAL	5,489,818	2	12	22	29	54	68	74	86	92	94	98	99	99	100

evolution of planting (maize)

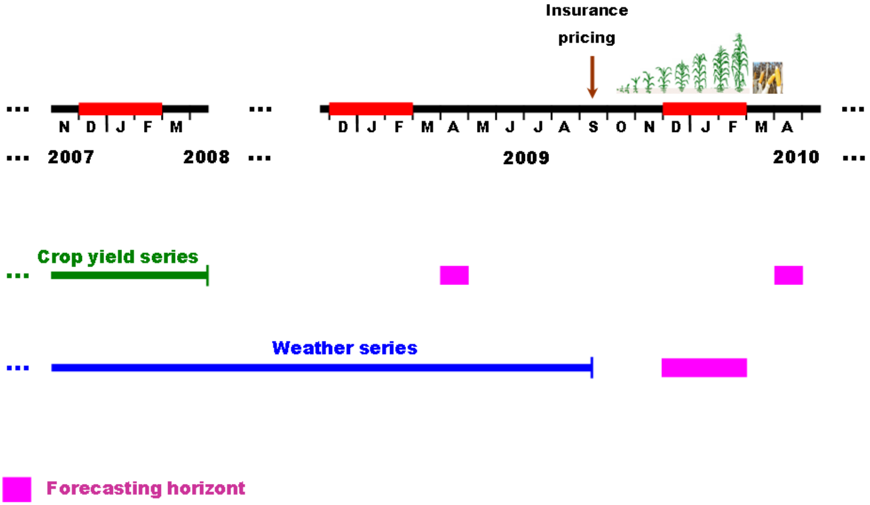


MILHO 1ª safra 2009/10 Ano 2008

NÚCLEO REGIONAL	Plantio (%)																		
	Aug	Set					Out				Nov				Dez				
	27	2	9	15	22	29	6	3	20	27	3	10	17	24	1	8	15	22	29
APUCARANA				2	5	10	50	85	95	97	99	100							
C. MOURÃO					5	35	70	80	95	97	100								
CASCAVEL				15	60	80	87	98	99	100									
C. PROCÓPIO						40	50	70	80	85	90	95	97	100					
CURITIBA					10	20	35	60	80	90	100								
F. BELTRÃO	10	12	27	70	82	86	93	94	95	96	97	0	0	98	100				
GUARAPUAVA			2	10	15	36	46	65	72	78	80	85	88	89	90	91	93	95	100
IRATI				4	15	45	70	85	90	92	95	97	0	99	100				
IVAIPORÃ						10	15	20	45	70	85	95	98	100					
JACAREZINHO	4	6	10	20	25	30	35	45	73	76	81	86	91	93	100				
LARANJEIRAS SUL					20	45	54	56	68	73	77	81	83	85	90	92	100		
LONDRINA						15	20	35	45	60	70	73	88	98	100				
MARINGÁ					50	70	75	93	95	96	97	100							
PARANAGUÁ	20	25	30	60	80	85	100												
PARANAVAÍ						5	15	35	40	60	70	90	100						
P. BRANCO	2	3	15	45	60	80	90	92	97	98	98	99	100						
P. GROSSA		2	10	30	45	68	80	85	90	95	97	98	99	100					
TOLEDO				28	72	96	98	100											
UMUARAMA				1	3	5	20	80	85	100									
U. VITÓRIA			1	2	6	20	40	55	67	70	90	95	97	100					
TOTAL	1	2	6	18	32	48	61	72	81	88	92	94	96	97	98	98	98.6	99.2	100

Fonte: SEAB/DERAL

Chronological scheme



The Bayesian hierarchical approach

$$y_{it} \sim N(\mu_{it}, \tau_i)$$
$$\mu_{it} = \rho_i y_{i,t-1} + \beta_{0_i} + \beta_{1_i} t + \beta_{2_i} t^2 + \beta_{3_i} cov_{it}$$
$$\tau_i = \exp(\mu_\tau + h_i)$$

where:

$\rho_i \sim N(\alpha_\rho, \tau_\rho)$, $\alpha_\rho \sim N(0, \tau_\alpha)$

$\beta_{j_i} \mid \beta_{j_{-i}} \sim N\left(\bar{\beta}_{j_{(i)}}, \frac{\tau_{\beta_j}}{r_i}\right)$ with $\bar{\beta}_{j_{(i)}} = \sum_{k \in \partial_i} \beta_{j_k} / r_i$ for $j \in \{0, 1, 2, 3\}$

$h_i \mid h_{-i} \sim N\left(\bar{h}_{(i)}, \frac{\tau_h}{r_i}\right)$ with $\bar{h}_{(i)} = \sum_{k \in \partial_i} h_k / r_i$; $\mu_\tau \sim N(0, b_\tau)$

$\beta_{j_{-i}}$ and h_{-i} are the vectors of all β_j's and h's excluding β_{j_i} and h_i, respectively

∂_i: set of neighbors of area i ; r_i: number of neighbors of area i

τ_h, τ_ρ and τ_{β_j} are inverse gamma distributed

The Bayesian dynamic approach

We use a Gaussian dynamic spatio-temporal model (Vivar and Ferreira, 2009):

$$\boldsymbol{y}_t = \boldsymbol{F}'_t \boldsymbol{x}_t + \varepsilon_t \quad , \qquad \varepsilon_t \sim PGMRF\left(\mathbf{0s}, \boldsymbol{V}_t^{-1}\right)$$

$$\boldsymbol{x}_t = \boldsymbol{G}_t \boldsymbol{x}_{t-1} + \omega_t \quad , \qquad \omega_{it} \sim PGMRF\left(\mathbf{0s}, \boldsymbol{W}_i^{-1}\right)$$

where,

$$\boldsymbol{x}_t = \begin{pmatrix} \boldsymbol{x}_{1t} \\ \boldsymbol{x}_{2t} \\ \beta_{1t} \\ \beta_{2t} \end{pmatrix}, \boldsymbol{F}_t = \begin{pmatrix} \boldsymbol{Is} \\ \mathbf{0s} \\ \mathrm{cov1} \\ \mathrm{cov2} \end{pmatrix}, \boldsymbol{G}_t = \begin{pmatrix} \rho_1 \boldsymbol{Is} & \rho_1 \boldsymbol{Is} & \mathbf{0s} & \mathbf{0s} \\ \mathbf{0s} & \rho_2 \boldsymbol{Is} & \mathbf{0s} & \mathbf{0s} \\ \mathbf{0s} & \mathbf{0s} & \rho_3 \boldsymbol{Is} & \mathbf{0s} \\ \mathbf{0s} & \mathbf{0s} & \mathbf{0s} & \rho_4 \boldsymbol{Is} \end{pmatrix}, \boldsymbol{W}^{-1} = \begin{pmatrix} \boldsymbol{W}_1^{-1} & \mathbf{0s} & \mathbf{0s} & \mathbf{0s} \\ \mathbf{0s} & \boldsymbol{W}_2^{-1} & \mathbf{0s} & \mathbf{0s} \\ \mathbf{0s} & \mathbf{0s} & \boldsymbol{W}_3^{-1} & \mathbf{0s} \\ \mathbf{0s} & \mathbf{0s} & \mathbf{0s} & \boldsymbol{W}_4^{-1} \end{pmatrix}$$

$$\boldsymbol{V}_t^{-1} = \tau_0(\boldsymbol{Is} + \phi \boldsymbol{M}) \qquad \boldsymbol{W}_i^{-1} = \tau_i(\boldsymbol{Is} + \phi_i \boldsymbol{M})$$

$$M_{k,j} = \begin{cases} m_k & \text{if} \quad k = j, \\ -h_{k,j} & \text{if} \quad k \in N_j, \\ 0 & \text{otherwise.} \end{cases}$$

$h_{k,j} > 0$ is a measure of similarity between regions

$\phi \geq 0$ controls the degree of spatial correlation

N_j is the set of neighbours of region j

τ is a scale parameter

$m_k = \sum_{j \in N_k} h_{k,j}$

- y_{ts} represents the annual average crop yield for each year t and county s,
- $\boldsymbol{x}_{1t}$ represents the level and $\boldsymbol{x}_{2t}$ represents the velocity of the process at time t,
- $\boldsymbol{F}_t$ connects the latent process to the observations,
- $\boldsymbol{G}_t$ describes the spatio-temporal evolution of the process,
- The prior for $\boldsymbol{x}_0$ is a multivariate normal with zero mean vector.
- The prior for τ_i and ϕ_i, $i \in \{1, 2, 3, 4\}$, is the joint reference prior for PGMRFs.
- Posterior inference is performed in a MCMC framework, with an embedded forward filter backward sampler (FFBS) algorithm.

The Kalman filter

Let $D_t = \{y_1, ..., y_t\}$

and suppose that at $t = 0$, $X_0 \mid D_0 \sim N(m_0, C_0)$, with known m_0 and C_0

- Posterior at time $t-1$: $X_{t-1} \mid D_{t-1} \sim N(m_{t-1}, C_{t-1})$
- Prior at time t: $X_t \mid D_{t-1} \sim N(a_t, R_t)$

  with $a_t = G_t m_{t-1}$, $R_t = G_t C_{t-1} G_t' + W_t$
- Predictive at time t: $y_t \mid D_{t-1} \sim N(f_t, Q_t)$

  with $f_t = F_t' a_t$, $Q_t = F_t' R_t F_t + V_t$
- Posterior at time t: $X_t \mid D_t \sim N(m_t, C_t)$

  with $m_t = a_t + A_t e_t$, $C_t = R_t - A_t Q_t A_t'$, $A_t = R_t F_t Q_t^{-1}$, $e_t = y_t - f_t$

The FFBS algorithm

- to sample X_T from $N(m_T, C_T)$ (Forward filtering)
- to sample X_t from $(X_t \mid X_{t+1}, V_t, W_t, D_t)$, for $t = T-1, T-2, ..., 2, 1$ (Backward smoothing),

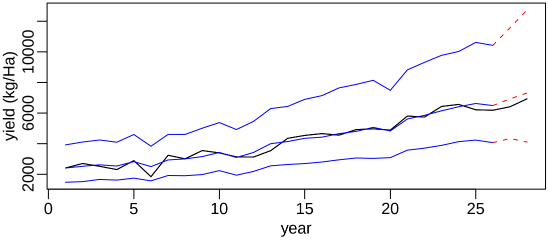
  with $(X_t \mid X_{t+1}, V_t, W_t, D_t) \sim N(mean, var)$,

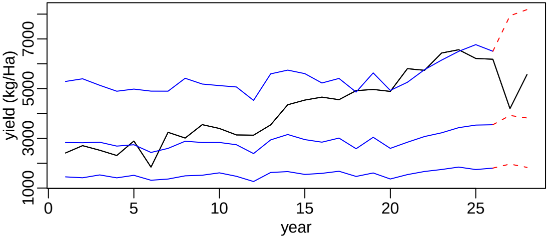
  where

  $$mean = (G_t' W_t^{-1} G_t + C_t^{-1})^{-1} (G_t' W_t^{-1} X_{t+1} + C_t^{-1} m_t)$$

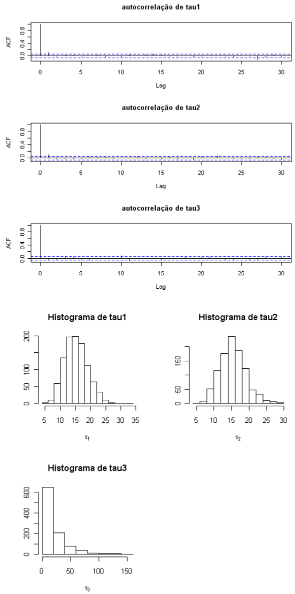
  $$var = (G_t' W_t^{-1} G_t + C_t^{-1})^{-1}$$

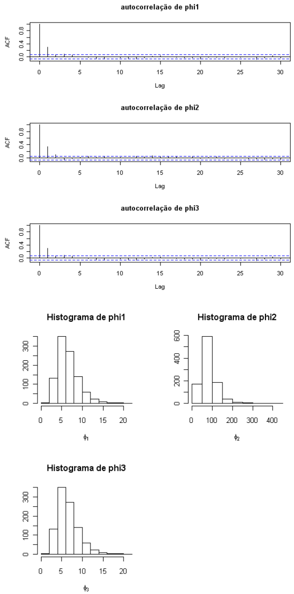
Preliminary results





Preliminary results



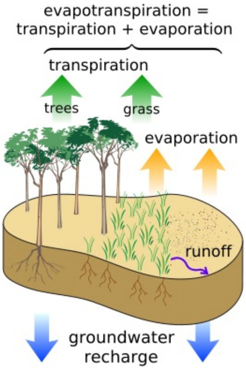


The covariates

The relationship between weather and yield variability is taken into account through agricultural drought indexes:

- *The drought index* (Mota, 1981): $DI = 1 - [ETa/ETm]$
  where ETa and ETm are the daily actual and maximum evapotranspiration accumulated during the critical period of the crop in terms of water deficit.
- *The standardized actual evapotranspiration index* (Blain et al., 2006):
  it quantifies agricultural drought in a 10-days scale, based on the fit of the ETa series to the beta distribution.
- *P - ET0:*
  the accumulated difference between precipitation and reference evapotranspiration through the critical period.

Evapotranspiration:



- *reference evapotranspiration (ET0):*

  is the evapotranspiration rate from a hypothetical reference surface under optimal soil water conditions.

  It is a climatic parameter that expresses the evaporation power of the atmosphere independently of vegetation characteristics and soil factors.

- *potential evapotranspiration (ETP):*

  refers to the evapotranspiration of a specific crop from well-watered fields that achieve full production under the given climatic conditions.

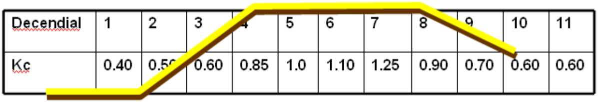
  $ETP = kc * ET0$

- *actual evapotranspiration (ETA):*

  is the evapotranspiration from a crop grown under management and environmental conditions that differ from the standard conditions.

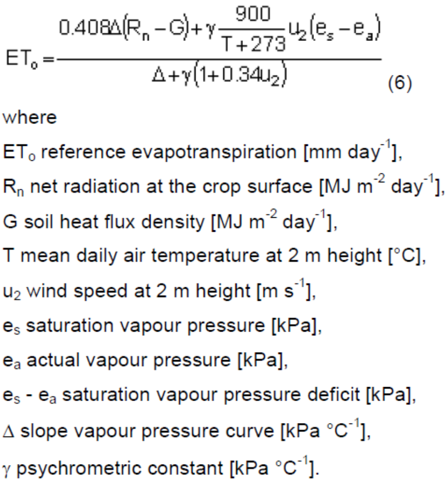
  $ETA = ks * ETP$

15 days	45 days	40 days	20 days
Sowing and establishment	Vegetative growth	Flowering and Grain formation	Maturity



Decendial	1	2	3	4	5	6	7	8	9	10	11
Kc	0.40	0.50	0.60	0.85	1.0	1.10	1.25	0.90	0.70	0.60	0.60

Reference evapotranspiration (Penman)



$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34 u_2)} \qquad (6)$$

where

ET_o reference evapotranspiration [mm day^{-1}],

R_n net radiation at the crop surface [MJ m^{-2} day^{-1}],

G soil heat flux density [MJ m^{-2} day^{-1}],

T mean daily air temperature at 2 m height [°C],

u_2 wind speed at 2 m height [m s^{-1}],

e_s saturation vapour pressure [kPa],

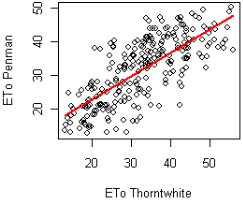
e_a actual vapour pressure [kPa],

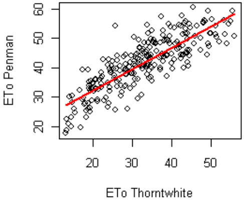
e_s - e_a saturation vapour pressure deficit [kPa],

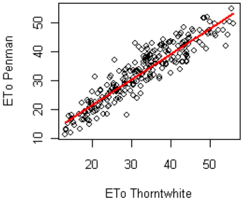
Δ slope vapour pressure curve [kPa °C^{-1}],

γ psychrometric constant [kPa °C^{-1}].

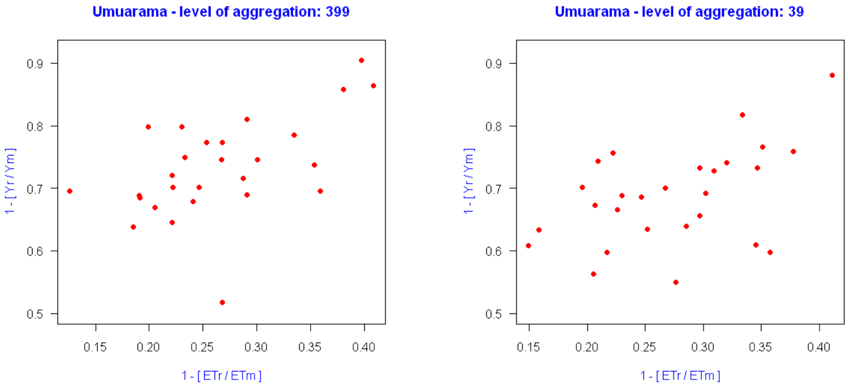
Reference evapotranspiration







A weather-yield relationship



ETa/ETm = 0 ⇒ [1 - ETa/ETm] = 1 ⇒ maximum water deficit ⇒ [1 - Y/Ym] HIGH

ETa/ETm = 1 ⇒ [1 - ETa/ETm] = 0 ⇒ no water deficit ⇒ [1 - Y/Ym] ≈ 0

Model selection criteria

- We used a slightly different version of the Gelfand and Ghosh (1998) criterion, which is based on the posterior predictive distributions.
- It considers the mean square predictive error *(MSPE)* relative to the number of regions used in the analysis.
- Additionally, the mean amplitude of the 95% credible interval for the predictions of the models at $T+2$, relative to the number of regions used in the analysis were used.

Rating the crop insurance contract

When a proportion λ $(0 \leq \lambda \leq 1)$ of the expected crop yield y^e is used to form the basis of insurance, the premium rate (PR) is given by:

$$PR = \frac{F(\lambda y^e) E_Y[\lambda y^e - y \mid y < \lambda y^e]}{\lambda y^e},$$

E is the expectation operator and F is the cumulative distribution function of yields.

If we reparameterize y, such that $y^* = y / \lambda y^e$, and considering $w = 1 - y^*$, then:

$$PR = P(w > 0)[1 - E_w(1 - w \mid w > 0)] = P(w > 0) E_w[w \mid w > 0],$$

which after some simplification reduces to:

$$PR = E[wI(0 < w < 1)].$$

The standard errors of the premium rate estimates obtained under the Bayesian approach provide a natural metric to guide the specification of loading factors.

Empirical pricing of crop insurance contracts

The empirical premium rate (EPR) as currently calculated by the insurers is given by:

$$EPR = \frac{E[y - \lambda y^{(h)}]}{\lambda y^{(h)}},$$

where:

$y^{(h)}$ is the average historic crop yield

$E[y - \lambda y^{(h)}] = \sum_{i=1}^{n} (y_i - y^{(h)})/n$

y_i is the observed crop yield for year i, and

λ is the coverage level.

Preliminary results

Tabela: Results of the model selection criteria for the Bayesian hierarchical approach

aggregation	covariate	MCIA* (95%)		MCIA* (90%)		MSPE** $x\ 10^3$
		mean	sd	mean	sd	
399	without	5588.4	2127.3	4609.2	1745.8	3454
	with	**4174.5**	**1445.1**	**3429.7**	**1186.3**	**2319**

* MCIA: mean credible interval amplitude

** MSPE: mean squared predictive error

Preliminary results

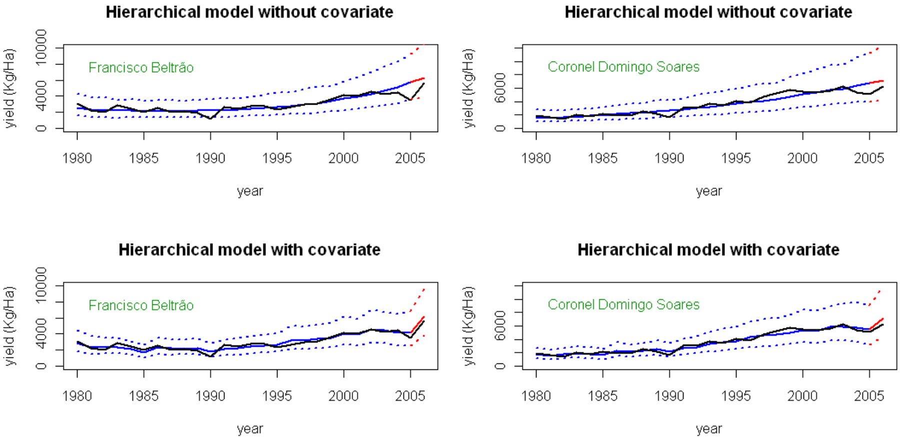
Tabela: Premium rates obtained from the Bayesian hierarchical approach

area	aggregation level	coverage (70%) mean	coverage (70%) CI range*	coverage (90%) mean	coverage (90%) CI range*
Francisco Beltrao	399	0.0735	0.3010	0.2019	0.4563
Domingo Soares	399	0.0571	0.2905	0.1777	0.4482
Manoel Ribas	399	0.1589	0.4259	0.318	0.5535
Cerro Azul	399	0.0440	0.2523	0.1465	0.4184

* CI range: 90% credible interval amplitude

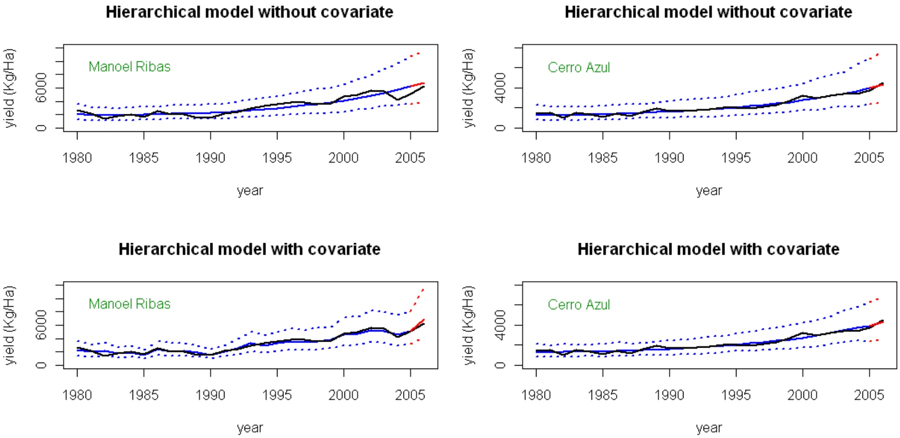
Preliminary results

Some predictions obtained with the Bayesian hierarchical model:



Preliminary results

Some predictions obtained with the Bayesian hierarchical model:



Hierarchical model without covariate

Hierarchical model without covariate

Hierarchical model with covariate

Hierarchical model with covariate

Future work

- spatio-temporal non-parametric modelling of crop yield through Dirichlet processes;
- zone-based crop insurance;
- crop simulation models coupled into spatio-temporal models to pricing crop insurance;
- estimation of crop yields through satellite images;
- risk reduction measures (stochastic dominance, VaR, etc.) to evaluate the performance of the modelling approaches;
- ENSO as an additional covariate in crop yield models;
- sequential Monte Carlo and Laplace approximations to overcome computational problems with MCMC in high dimensions;
- “IC” or “CI”? that’s the question.

Future work

- spatio-temporal non-parametric modelling of crop yield through Dirichlet processes;
- zone-based crop insurance;
- crop simulation models coupled into spatio-temporal models to pricing crop insurance;
- estimation of crop yields through satellite images;
- risk reduction measures (stochastic dominance, VaR, etc.) to evaluate the performance of the modelling approaches;
- ENSO as an additional covariate in crop yield models;
- sequential Monte Carlo and Laplace approximations to overcome computational problems with MCMC in high dimensions;
- “IC” or “CI”? that’s the question.

Future work

- spatio-temporal non-parametric modelling of crop yield through Dirichlet processes;
- zone-based crop insurance;
- crop simulation models coupled into spatio-temporal models to pricing crop insurance;
- estimation of crop yields through satellite images;
- risk reduction measures (stochastic dominance, VaR, etc.) to evaluate the performance of the modelling approaches;
- ENSO as an additional covariate in crop yield models;
- sequential Monte Carlo and Laplace approximations to overcome computational problems with MCMC in high dimensions;
- “IC” or “CI”? that’s the question.

Future work

- spatio-temporal non-parametric modelling of crop yield through Dirichlet processes;
- zone-based crop insurance;
- crop simulation models coupled into spatio-temporal models to pricing crop insurance;
- estimation of crop yields through satellite images;
- risk reduction measures (stochastic dominance, VaR, etc.) to evaluate the performance of the modelling approaches;
- ENSO as an additional covariate in crop yield models;
- sequential Monte Carlo and Laplace approximations to overcome computational problems with MCMC in high dimensions;
- “IC” or “CI”? that’s the question.

Future work

- spatio-temporal non-parametric modelling of crop yield through Dirichlet processes;
- zone-based crop insurance;
- crop simulation models coupled into spatio-temporal models to pricing crop insurance;
- estimation of crop yields through satellite images;
- risk reduction measures (stochastic dominance, VaR, etc.) to evaluate the performance of the modelling approaches;
- ENSO as an additional covariate in crop yield models;
- sequential Monte Carlo and Laplace approximations to overcome computational problems with MCMC in high dimensions;
- "IC" or "CI"? that's the question.

Future work

- spatio-temporal non-parametric modelling of crop yield through Dirichlet processes;
- zone-based crop insurance;
- crop simulation models coupled into spatio-temporal models to pricing crop insurance;
- estimation of crop yields through satellite images;
- risk reduction measures (stochastic dominance, VaR, etc.) to evaluate the performance of the modelling approaches;
- ENSO as an additional covariate in crop yield models;
- sequential Monte Carlo and Laplace approximations to overcome computational problems with MCMC in high dimensions;
- “IC” or “CI”? that’s the question.

Future work

- spatio-temporal non-parametric modelling of crop yield through Dirichlet processes;
- zone-based crop insurance;
- crop simulation models coupled into spatio-temporal models to pricing crop insurance;
- estimation of crop yields through satellite images;
- risk reduction measures (stochastic dominance, VaR, etc.) to evaluate the performance of the modelling approaches;
- ENSO as an additional covariate in crop yield models;
- sequential Monte Carlo and Laplace approximations to overcome computational problems with MCMC in high dimensions;
- “IC” or “CI”? that’s the question.

Future work

- spatio-temporal non-parametric modelling of crop yield through Dirichlet processes;
- zone-based crop insurance;
- crop simulation models coupled into spatio-temporal models to pricing crop insurance;
- estimation of crop yields through satellite images;
- risk reduction measures (stochastic dominance, VaR, etc.) to evaluate the performance of the modelling approaches;
- ENSO as an additional covariate in crop yield models;
- sequential Monte Carlo and Laplace approximations to overcome computational problems with MCMC in high dimensions;
- “IC” or “CI”? that’s the question.

Some References

- Ferreira, M. A. R. and De Oliveira, V. (2007) Bayesian reference analysis for Gaussian Markov Random Fields. *Journal of Multivariate Analysis*, **98**, 789–812.

- Gelfand, A. E. and Ghosh, S.K. (1998) Model Choice: A minimum posterior predictive loss approach. *Biometrika*, **85**, 1–11.

- Gelfand, A. E., Zhu, L. and Carlin, B. (2001) On the change of support problem for spatiotemporal data. *Biostatistics*, **2**, 31–45.

- Junger, W.L., Ponce de Leon, A. and Santos, N. (2003) Missing data imputation in multivariate time series via EM algorithm. *Cadernos do IME*, **15**, 8–21.

- Ozaki, V.A., Ghosh, S.K., Goodwin, B.K. and Shirota, R. (2008) Spatio-Temporal Modeling of Agricultural Yield Data with an Application to Pricing Crop Insurance Contracts. *American Journal of Agricultural Economics*, **90**, 951–961.

- Vivar, J. C. and Ferreira, M. A. R. (2009) Spatiotemporal models for gaussian areal data. *Journal of Computational and Graphical Statistics*, **18**, 658–674.

Some References

Rue, H. and Martino, S. (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society series B*, **71**, 319–392.

Carvalho, C.M., Johannes, M., Lopes, H.F. and Polson, N. (2009) Particle Learning and Smoothing. Technical report, University of Chicago, Booth School of Business. 44p.

Yan, J. (2007) Spatial Stochastic Volatility for Lattice Data. *Journal of Agricultural, Biological, and Environmental Statistics*, **12**, 25–40.