UNSW SYDNEY

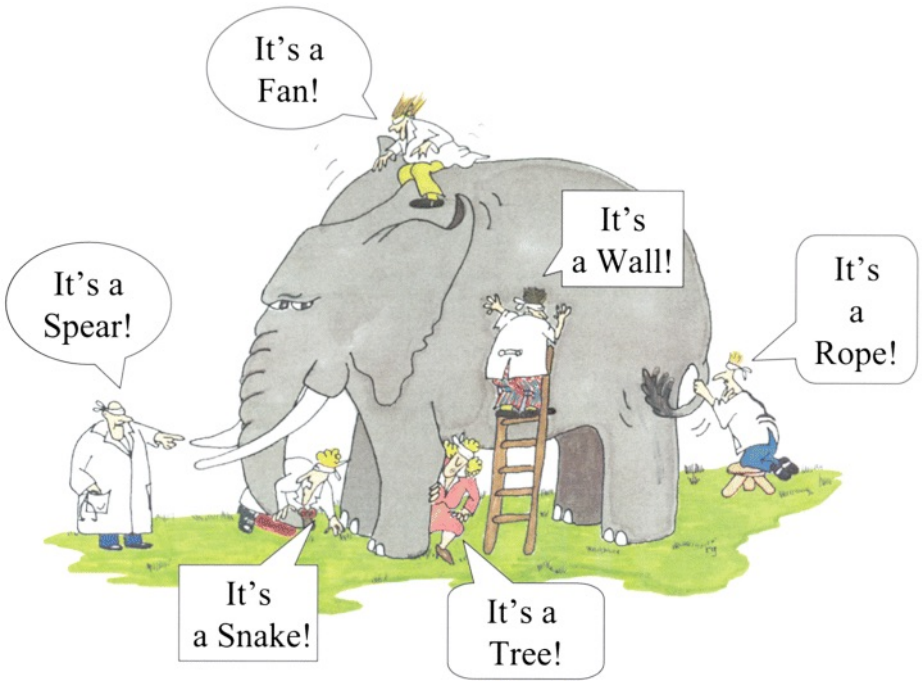
Composite likelihood and logistic regression models for aggregated data

Tom Whitaker, Boris Beranger, Scott A. Sisson,

UNSW & ACEMS

Mathematics Colloquium, UTS, August 14th

Talk Outline



1. **What is Symbolic Data Analysis?**
   - Symbolic likelihood
   - . . . and its limitations
2. Symbolic composite likelihoods
   - Methods
   - Applications to spatial extremes
3. SCL for logistic regression
   - Methods
   - Application to satellite crop prediction
4. Discussion

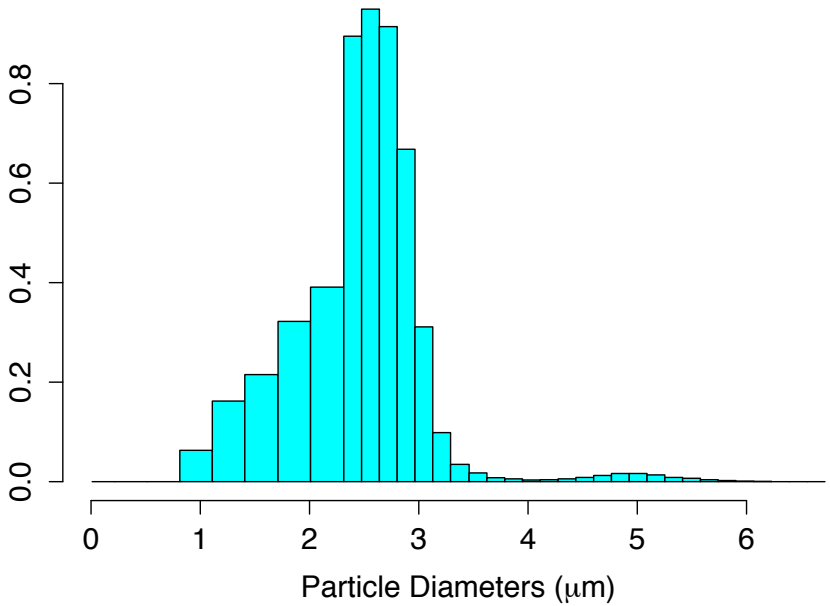
Main idea

Standard statistical methods analyse classical datasets

E.g. $x_1, \ldots, x_n$ where $x_i \in \mathcal{X} = \mathbb{R}^d$

However there is a rise of *non-standard* data forms:

- As a result of measurement process;
- Blood pressure recorded as (low, high) interval;
- Particulate matter recorded as counts within particle diameter ranges i.e. histogram;



Main idea

'Big Data' context:

- Symbolic data points to summarise a complex & very large dataset in a compact manner.

$$S = \pi(X_{1:N}) : [\mathcal{X}]^N \to \mathcal{S} \text{ such that } x_{1:N} \mapsto \pi(x_{1:N})$$

- Retaining maximal relevant information in original dataset.
- Collapse over data not needed in detail for analysis.
- Summarised data have own internal structure, which must be taken into account in any analysis.

⊕ Big data $\to$ small (symb) data
⊕ Possible use in data privacy? Individual can't be identified

Statistical question:
How to do statistical analysis for this form of data?

Likelihood-based SDA (Beranger, Lin & Sisson, 2018)

The general approach:

$$L(S|\theta,\phi) \propto \int_x g(S|x,\phi)L(x|\theta)dx$$

where

- $L(x|\theta)$ – standard, classical data likelihood
- $g(S|x,\phi)$ – probability of obtaining S given classical data x
- $L(S|\theta)$ – new symbolic likelihood for parameters of classical model

Gist: Fitting the standard classical model $L(x|\theta)$, when the data are viewed only through symbols S as summaries.

As $S_i \to x_i$, then $g(S_i|x,\phi) = g(x_i|x) = \delta_{x_i}(x)$ and so

$$L(S_i|\theta,\phi) \propto \int_x \delta_{x_i}(x)L(x|\theta)dx = L(x_i|\theta) \qquad \text{(classical likelihood)}$$

Different symbols give different forms for $g(S|x,\phi)$ (and $\therefore$ $L(S|\theta,\phi)$).

Specific case: Random histograms

Underlying data $X_1, \ldots, X_N \in \mathbb{R}^d \sim g(x|\theta)$ collected into random counts histogram, with fixed bins $\mathcal{B}_1, \ldots, \mathcal{B}_B$.

Aggregation:

$S = \pi(X_{1:N}) : \mathbb{R}^{d \times N} \to \mathcal{S} = \{0, \ldots, N\}^{B^1 \times \cdots \times B^d}$ such that
$x_{1:N} \mapsto \left(\sum_{i=1}^n \mathbb{I}\{x_i \in \mathcal{B}_1\}, \ldots, \sum_{i=1}^n \mathbb{I}\{x_i \in \mathcal{B}_B\}\right)$.

$$g(S|x, \phi) = \begin{cases} 1 & \text{if } s_b \text{ observations in bin } b\text{; for each } b = 1, \ldots, B \\ 0 & \text{else} \end{cases}$$

The symbolic likelihood is then (multinomial):

$$L(S|\theta) \propto \int_x g(S|x) \prod_{k=1}^n g(x_k|\theta) dx \propto \prod_{b=1}^B \left(\int_{B_b} g(z|\theta) dz \right)^{s_b}$$

$\Rightarrow$ generalises univariate result of McLachlan & Jones (1988). ✓

Specific case: Random histograms

- Can recover classical likelihood as $B \to \infty$

$$\begin{aligned} \lim_{B\to\infty} L(S|\theta) &\propto \lim_{B\to\infty} \prod_{b=1}^{B} \left[\int_{B_b} g(z|\theta) dz \right]^{s_b} \\ &= L(X_1, \ldots, X_n|\theta) \end{aligned}$$

  So recover classical analysis as we approach classical data. ✓

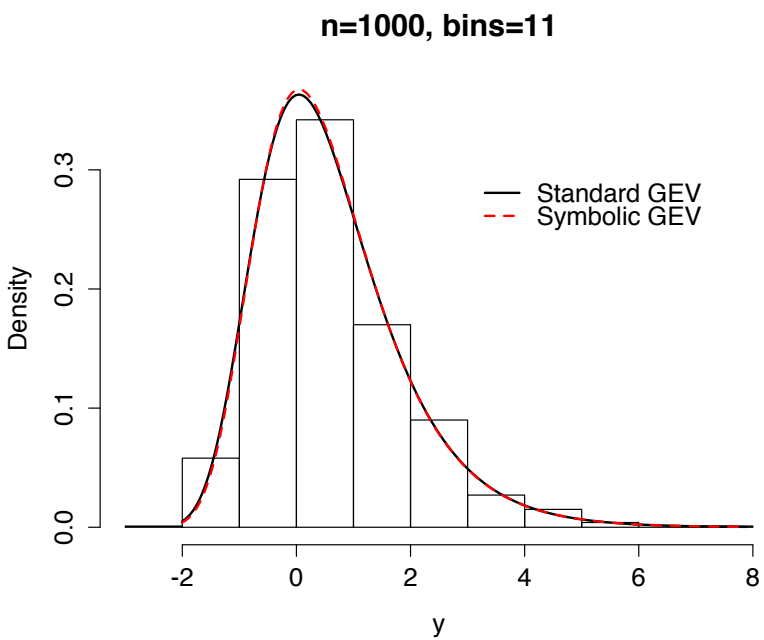
- Consistency: Can show that with a sufficient number of histogram bins can perform analysis arbitrarily close to analysis with full dataset.
- Some approximation of $L(S|\theta)$ to $L(x|\theta)$ depending on level of discretisation. Work needed to quantify this.
- More complicated if data are not *iid* but exchangeable

  (Zhang, Beranger & Sisson (2020), SJS)

Fitting a GEV

Suppose $X_1, \ldots, X_n \sim \text{GEV}(\mu, \sigma, \xi)$ for large n.
Create histogram of counts $s = (s_1, \ldots, s_B)$.
Symbolic log-likelihood function is then

$$\ell(s|\mu, \sigma, \xi) \propto \sum_{b=1}^{B} s_b \log\left[G(a_{b+1}|\mu, \sigma, \xi) - G(a_b|\mu, \sigma, \xi)\right]$$



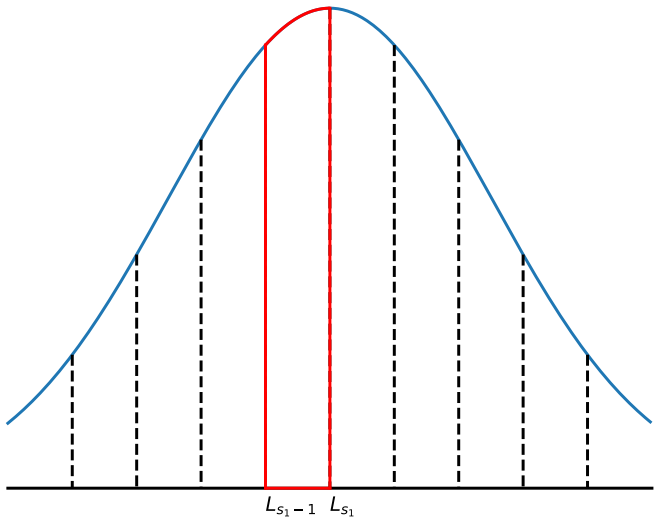
Note that $\ell(s|\mu, \sigma, \xi)$ tends to standard likelihood as # bins gets large (so 1 or 0 observations per bin)

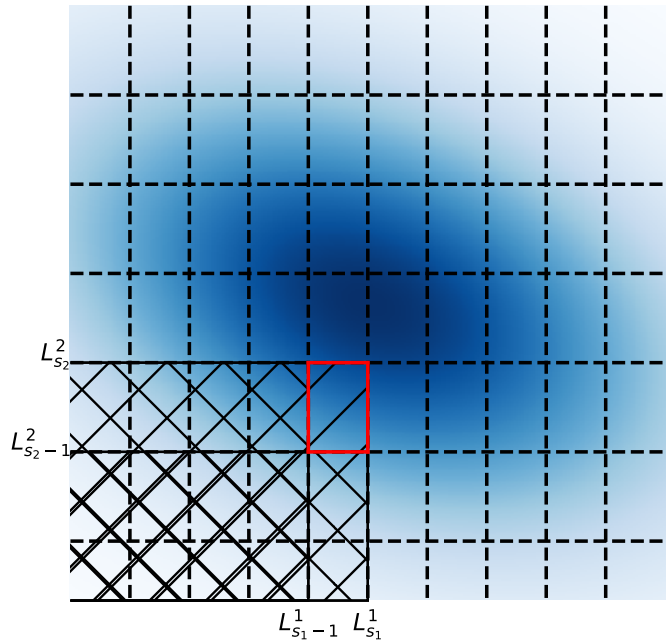
Computation:

- Optimisation of ℓ (v. quick)
- Creation of histogram s (slower)

$\leftarrow$ good fits with moderate bin numbers

Limitations: Calculating $\int_{B_b} g(z|\theta)dz$





1D: Probability of X_i falling in bin $(L_{s_1}, L_{s_1-1}]$ is $F(L_{s_1}|\theta) - F(L_{s_1-1}|\theta)$.

2D: Probability of X_i falling in bin $(L^{(1)}_{s_1}, L^{(1)}_{s_1-1}] \times (L^{(2)}_{s_1}, L^{(2)}_{s_1-1}]$ is

$$F\left(L^{(1)}_{s_1-1}, L^{(2)}_{s_2-1}|\theta\right) - F\left(L^{(1)}_{s_1-1}, L^{(2)}_{s_2}|\theta\right) - F\left(L^{(1)}_{s_1}, L^{(2)}_{s_2-1}|\theta\right) + F\left(L^{(1)}_{s_1}, L^{(2)}_{s_2}|\theta\right).$$

d-D: Has 2^d components – viable for low d.

Limitations for histograms

For the symbolic log likelihood (using multivariate histograms)

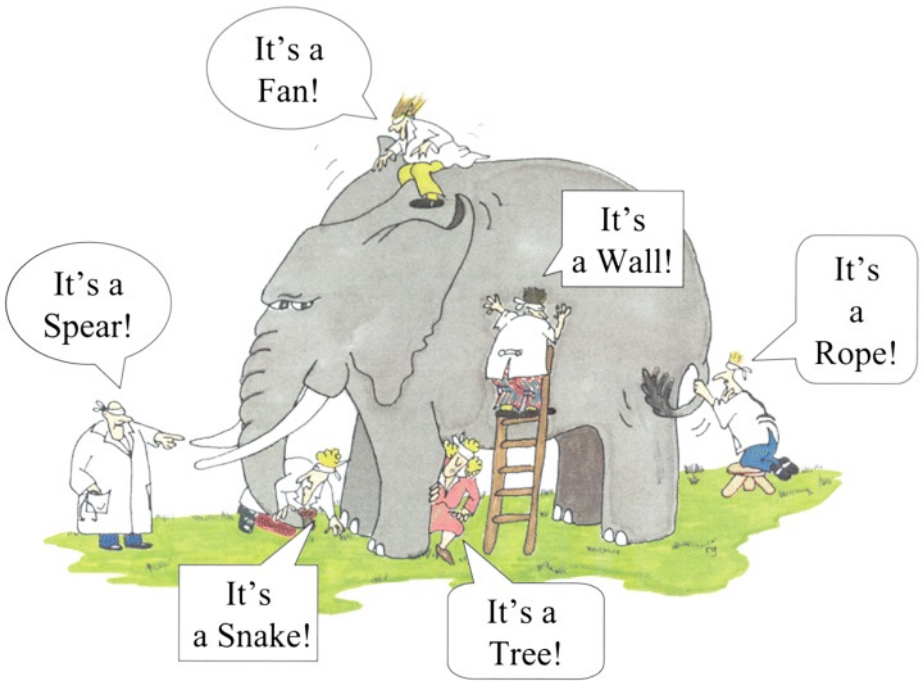
$$\log L(s|\theta) \propto \sum_{b=1}^{B} s_b \log \left[\int_{B_b} g(z|\theta) dz \right]$$

there are some limitations:

- Multivariate histograms become inefficient as d gets large
  – number of bins to cover d dimensions accurately gets large fast.
- $\int_{B_b} g(z|\theta) dz$ has 2^d components *for each bin*
  – only viable for low d.
- This means that we are limited to low-dimensional symbols.
  - SDA traditionally uses 1- or 2-dimensional symbols, so maybe this is ok;
  - But in principle there will be an analysis that requires higher-D information
    So should resolve this problem if possible.

One option: Composite likelihoods.

Talk Outline

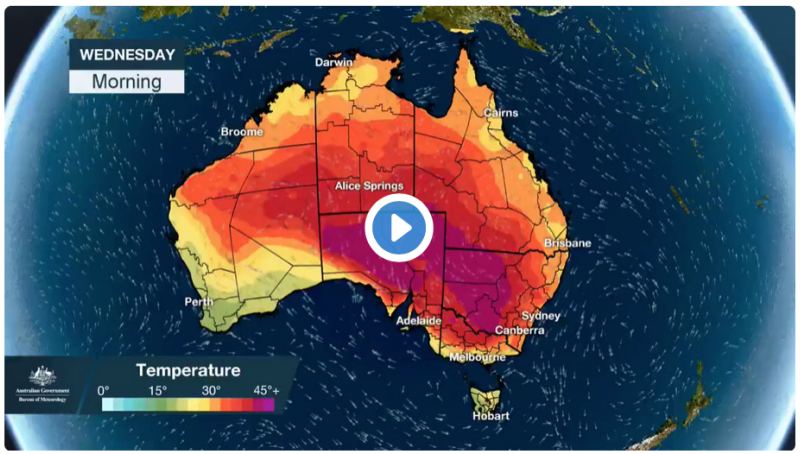


1. What is Symbolic Data Analysis?
   - Symbolic likelihood
   - . . . and its limitations
2. **Symbolic composite likelihoods**
   - Methods
   - Applications to spatial extremes
3. SCL for logistic regression
   - Methods
   - Application to satellite crop prediction
4. Discussion

Spatial Extremes

Bureau of Meteorology, New South Wales
@BOM_NSW

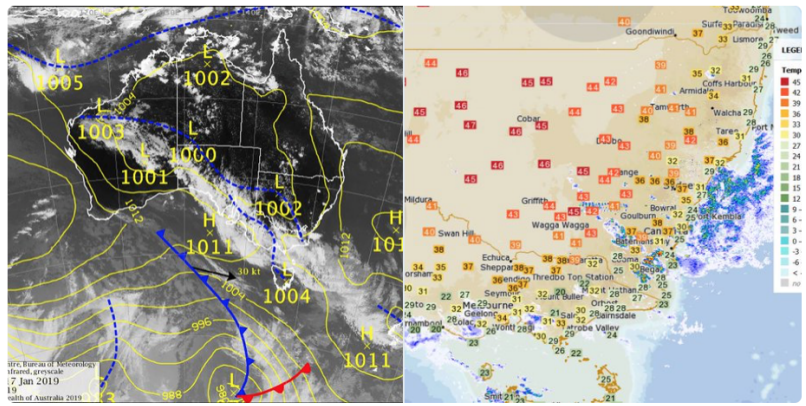
Fri marks peak day for some of #NSW most heavily populated areas.Temps in western #Sydney well into the 40's, regional western towns similar after many broke records this week, CBD likely to have 5th consecutive day above 30 for 1st time in 8 yrs ow.ly/E9QY50ke6I7 #heatwave



♡ 59 4:23 PM - Jan 17, 2019

Bureau of Meteorology, Australia
@BOM_au

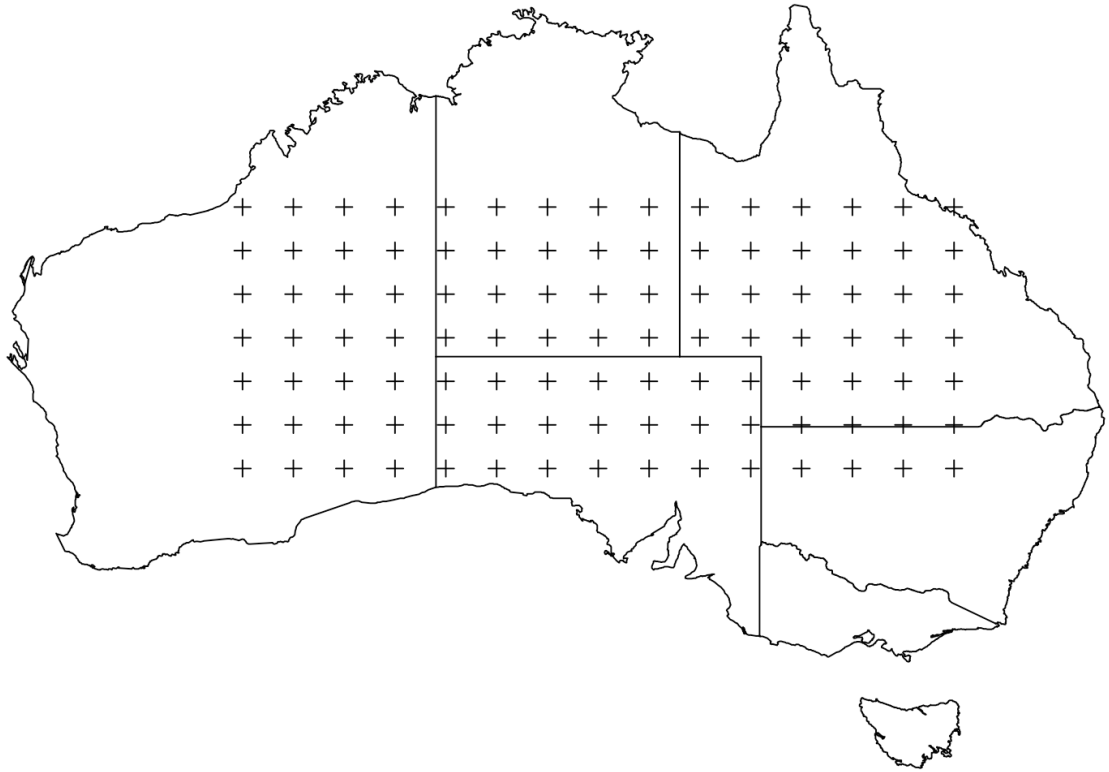
"Severe to extreme heatwave conditions across the southeast interior". Temperatures exceeding 45oC for many locations through western NSW and central Australia this afternoon. Latest at ow.ly/3W6s30nlrdY



♡ 18 6:11 PM - Jan 17, 2019

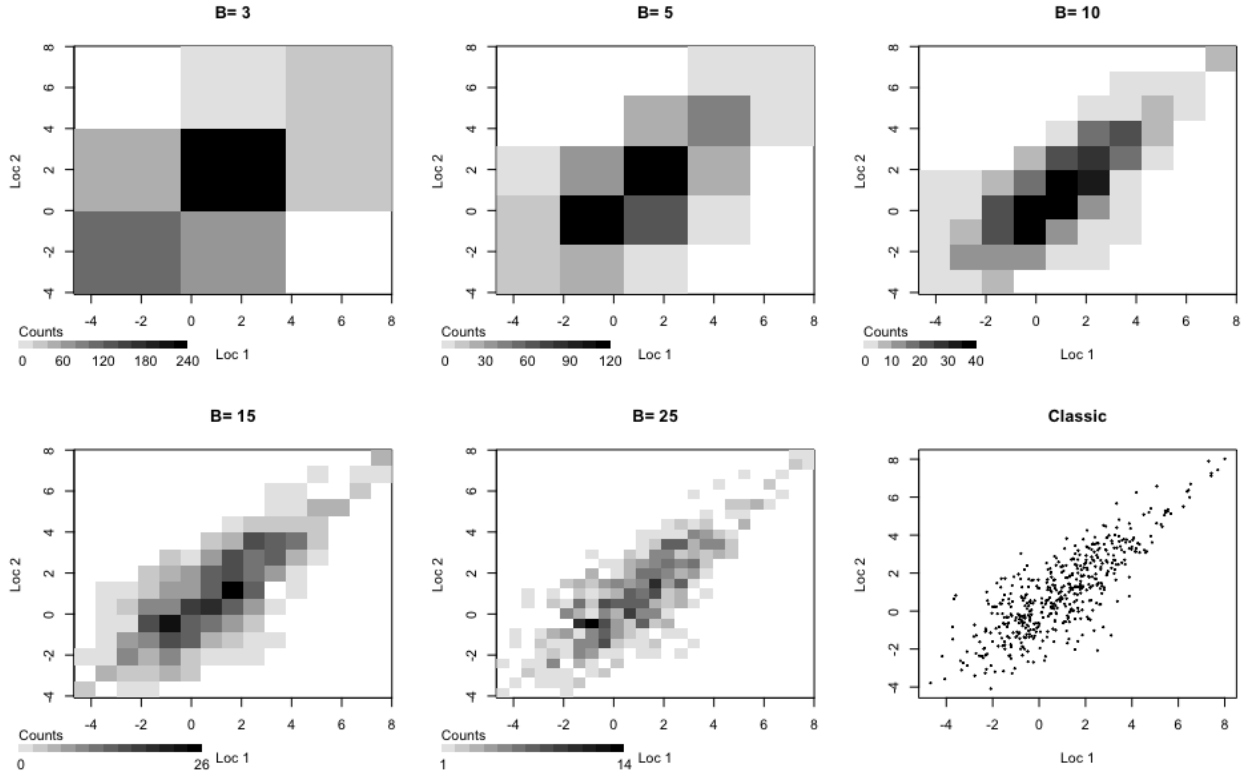
- What is the maximum value that a process (Temperature) is expected to reach over some region of interest (NSW/Australia) within the next 20, 50 years?
- Whitaker, Beranger & Sisson (2020, Stat. Comput.)

Modelling Australian temperature spatial extremes



- 105 spatial locations with temperature observation, over time
- Want to fit spatial model to temperature extremes.
  Spatial (multivariate) information is important!
- Lots of data – can form 105-dimensional histogram(!)
- Can't fit this using $L(S|\theta)$. **What can we do?**

Modelling Australian temperature spatial extremes



- 105-dimensional histograms are completely infeasible.
- But lower-dimensional histograms could still be very informative.
- E.g. 2-dimensional.
- But how to do inference with $105 \times 104/2$ bivariate histograms?
- One answer: Composite likelihoods

Composite likelihoods

Standard likelihood $L(\boldsymbol{x}|\theta)$ where $\boldsymbol{x} = (x_1, \ldots, x_d)$
where e.g. x_i = data at i-th spatial location.

Suppose $L(\boldsymbol{x}|\theta)$ is computationally intractable except for e.g. $d = 2$
(as for the spatial extremes model we are using).

Then can construct (say) **pairwise composite likelihood**

$$L_{CL}^{(2)}(\boldsymbol{x}|\theta) \propto \prod_i \prod_{j>i} L(x_i, x_j|\theta)$$

from all bivariate marginal events.

Works if $\ell(x_i, x_j|\theta)$ is an unbiased estimating equation for θ (as then log likelihood is a sum of these and so is also an unbiased estimating equation for θ). So works well for spatial models.

Similarly j-wise composite likelihoods.

Composite likelihoods

Behaviour of composite MLE

$\hat{\theta}^{(j)}_{CL}$ is asymptotically ($N \to \infty$) consistent and distributed as

$$\sqrt{N}\left(\hat{\theta}^{(j)}_{CL} - \theta\right) \to N\left(0,\, G^{(j)}(\theta)^{-1}\right)$$

where

- $G^{(j)}(\theta)^{-1} = H^{(j)}(\theta) J^{(j)}(\theta)^{-1} H^{(j)}(\theta)$ is Godambe information matrix
- $H^{(j)}(\theta) = -\mathbb{E}(\nabla^2 \ell^{(j)}_{CL}(\theta; \boldsymbol{x}))$ is the sensitivity matrix
- $J^{(j)}(\theta) = \mathbb{V}(\nabla \ell^{(j)}_{CL}(\theta; \boldsymbol{x}))$ is the variability matrix.

- For standard likelihoods $j = d$ and $H(\theta) = J(\theta)$ and so $G(\theta) = H(\theta) = I(\theta)$ is the Fisher information matrix.
- How can this help us with $L(\boldsymbol{S}|\theta)$ when $\boldsymbol{S}$ is a 105-dimensional histogram?

Composite symbolic likelihoods

As with $L^{(2)}_{CL}(\boldsymbol{x}|\theta) \propto \prod_i \prod_{j>i} L(x_i, x_j|\theta)$ we may similarly have

$$L^{(2)}_{CL}(\boldsymbol{S}|\theta) \propto \prod_i \prod_{j>i} L(S_{ij}|\theta)$$

where S_{ij} is the bivariate *marginal* histogram for dimensions (i, j).

In this setting

$$L(S_{ij}|\theta) \propto \prod_b \left(\int_{B_b} g(z_1, z_2|\theta) dz_1 dz_2 \right)^{s_b}$$

as before.

Composite symbolic likelihoods

From $L(S|\theta)$ we have (for a single histogram):

$\hat{\theta}$ is asymptotically consistent and distributed as

$$\sqrt{N}\left(\hat{\theta} - \theta\right) \to N\left(0,\, I(\theta)^{-1}\right)$$

when

- $N \to \infty$
- Number of bins $\to \infty$ and volume of each bin $\to 0$ (because then $L(S|\theta) \to L(x|\theta)$)

But when the bins (number and volume) are fixed then

$$\sqrt{N}\left(\hat{\theta} - \theta\right) \to N\left(??(\theta, bins),\, ??(\theta, Bins)^{-1}\right).$$

- Currently working on non-asymptotic (in bins) distribution of MLE

Composite symbolic likelihoods

From $L^{(j)}_{SCL}(\boldsymbol{S}|\theta)$ we have (for a single histogram):

$\hat{\theta}^{(j)}_{SCL}$ is asymptotically consistent and distributed as

$$\sqrt{N}\left(\hat{\theta}^{(j)}_{SCL} - \theta\right) \to N\left(0,\, G(\theta)^{-1}\right)$$

when

- $N \to \infty$
- Number of bins $\to \infty$ and volume of each bin $\to 0$ (because then $L^{(j)}_{SCL}(\boldsymbol{S}|\theta) \to L^{(j)}_{CL}(\boldsymbol{x}|\theta)$)

But when the bins (number and volume) are fixed then, as before

$$\sqrt{N}\left(\hat{\theta}^{(j)}_{SCL} - \theta\right) \to N\left(??(\theta, bins),\, ??(\theta, Bins)^{-1}\right).$$

- Similarly work in progress.

Simulated spatial extremes

(Mean) Pairwise symbolic composite likelihood estimates ($\hat{\theta}^{(2)}_{SCL}$):

- Consider $N = 1\,000$ observations at $K = 15$ spatial locations and $T = 1$ random histogram
- Spatial dependence of Gaussian max-stable model is $\sigma_{11} = 300$, $\sigma_{12} = 150$ and $\sigma_{22} = 200$

B	σ_{11}	σ_{12}	σ_{22}
2	321.6 (360.0)	162.3 (210.6)	210.8 (131.2))
3	296.1 (30.6)	147.4 (20.1)	197.9 (19.9)
5	298.8 (23.3)	149.4 (15.3)	199.6 (15.4)
10	299.0 (19.3)	149.6 (12.3)	199.7 (12.9)
15	299.5 (18.7)	149.8 (11.6)	199.8 (12.1)
25	299.7 (17.8)	150.0 (11.2)	200.0 (11.8)
Classic	**300.7 (16.4)**	**150.6 (10.2)**	**200.6 (10.9)**

Table: Mean (and standard errors) of the symbolic composite MLE $\hat{\theta}^{(2)}_{SCL}$ and composite MLE $\hat{\theta}^{(2)}_{CL}$ (Classic) from 1000 replications of the Gaussian max-stable process model, for $B \times B$ histograms for varying values of B.

- As "bins $\to \infty$" performance approaches classical composite likelihood (also estimated the marginal parameters).
- "Acceptable" results for $B = 10$

Simulated spatial extremes

(Mean) Time comparisons for increasing N

- Consider $B = 25$ bins, $K = 10, 100$ spatial locations and $T = 1$ random histogram. Repetitions $= 10$

N	$K = 10$				$K = 100$			
	t_c	t_s	t_{histDR}	t_{histR}	t_c	t_s	t_{histDR}	t_{histR}
1 000	71.9	22.5	0.8	0.1	–	2 238.0	78.8	12.0
5 000	291.8	19.0	0.8	0.3	–	2 650.2	81.7	30.9
10 000	591.7	23.8	0.9	0.5	–	2 356.6	85.8	54.1
50 000	2 626.8	24.2	1.7	2.1	–	2 300.6	131.6	237.0
100 000	5 610.7	25.4	2.4	4.2	–	2 766.9	188.2	461.8
500 000	31 083.1	23.2	7.5	20.6	–	3 111.5	627.1	2 243.5

Table: Mean computation times (seconds) for different components involved in computing $\hat{\theta}^{(2)}_{CL}$ and $\hat{\theta}^{(2)}_{SCL}$.

- Classical composite likelihood rapidly not feasible as spatial dimensions increases ($K = 20$)
- Symbolic approach much more efficient

Simulated spatial extremes

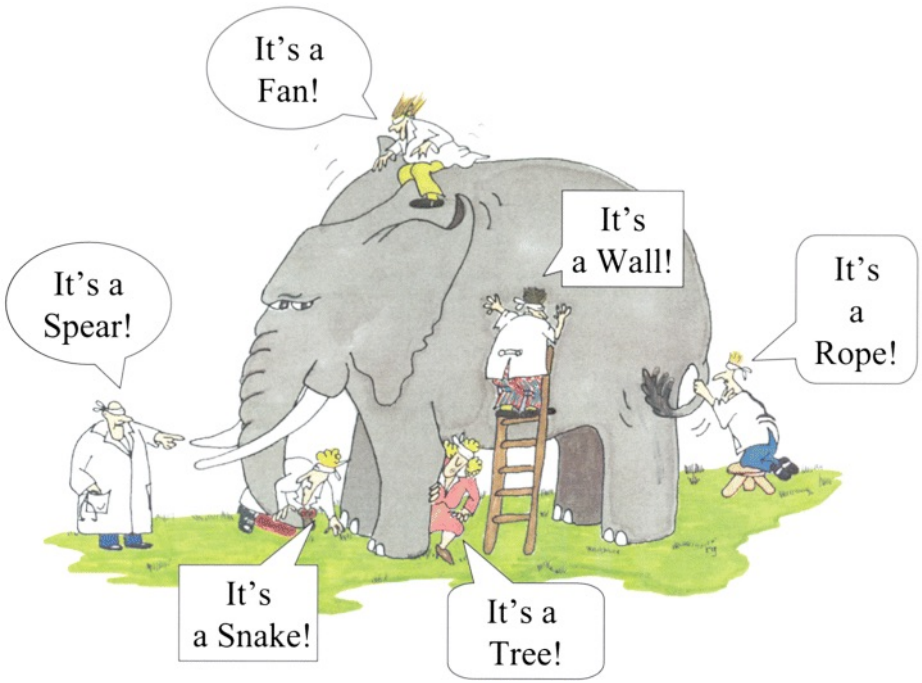
Plenty more simulations and interesting results in the paper[1]

- Including effect of number of histograms
- . . . and allocation of micro-data between them;
- Comparing bivariate SCL and trivariate SCL.

Now consider a different problem with "high" dimensional histograms:
$\longrightarrow$ logistic regression $\longleftarrow$

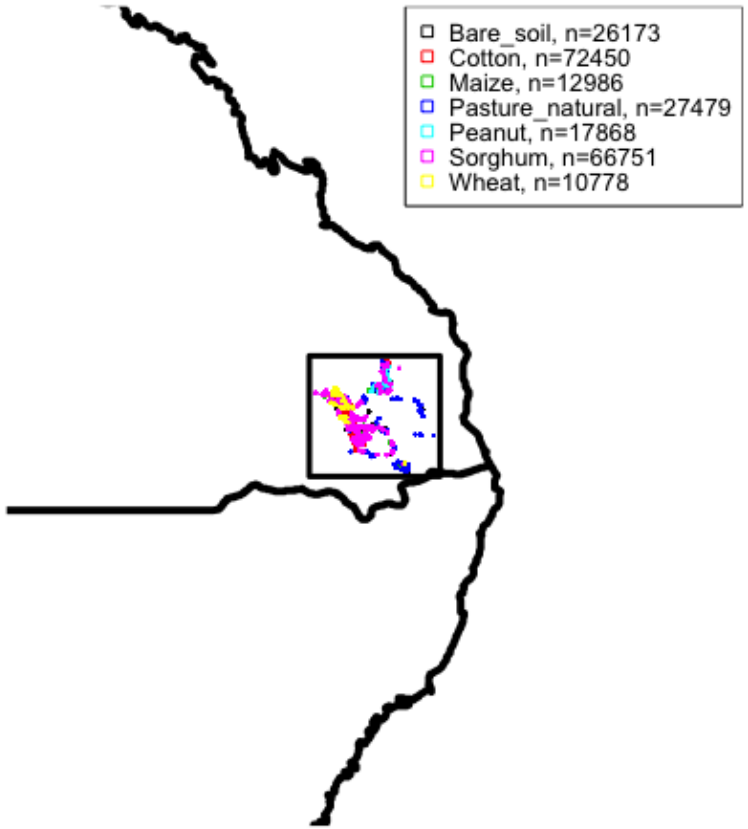
[1] Whitaker T., B. Beranger and S. A. Sisson (2020). Composite likelihood methods for histogram-valued random variables. Stat. Comput., In Press.

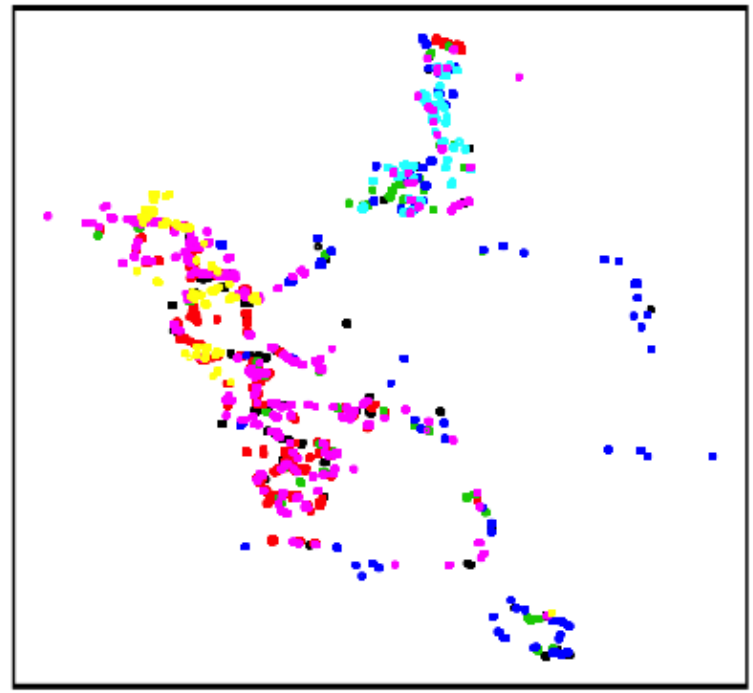
Talk Outline



1. What is Symbolic Data Analysis?
   - Symbolic likelihood
   - . . . and its limitations
2. Symbolic composite likelihoods
   - Methods
   - Applications to spatial extremes
3. **SCL for logistic regression**
   - Methods
   - Application to satellite crop prediction
4. Discussion

Prediction of crop types from satellite images





- ~250K pixels with 7-dimensional predictor variable x
- 7 response categories with known ground truth
- Multinomial logistic regression, but computational to fit (8+ hours)
- Can we use SDA to speed things up while maintaining prediction quality?

Logistic regression models

(Y_i, X_i) pairs, $Y_i \in \Omega = \{1, \ldots, K\}$ and $X \in \mathbb{R}^d$, $i = 1, \ldots, N$

Want to predict category Y_i given vector X_i.

One-versus-Rest logistic regression

Odds ratio linear model

$$\log\left(\frac{P(Y=k|X)}{P(Y\neq k|X)}\right) = \beta_{0k} + \beta_k^\top X$$

so that

$$P(Y = k|X) = \frac{\exp\{\beta_{0k}+\beta_k^\top X\}}{1+\exp\{\beta_{0k}+\beta_k^\top X\}}$$

and so the standard classical likelihood is

$$L(\boldsymbol{Y}, \boldsymbol{X}|\beta) \propto \prod_{i=1}^N \left(P(Y = Y_i|X = X_i) \prod_{j\in\Omega\setminus\{Y_i\}} P(Y \neq j|X = X_i)\right).$$

Logistic regression models

(Y_i, X_i) pairs, $Y_i \in \Omega = \{1, \ldots, K\}$ and $X \in \mathbb{R}^d$, $i = 1, \ldots, N$

Want to predict category Y_i given vector X_i.
Construct 7-dimensional predictor histogram $\boldsymbol{S}^j$ from $\{X_i : Y_i = j\}$, for each crop type $j = 1, \ldots, 7$.

The histogram-based likelihood for $\boldsymbol{S} = (\boldsymbol{S}^1, \ldots, \boldsymbol{S}^7)$ is then

$$L_S(\boldsymbol{S}|\beta) \propto \prod_{k\in\Omega} \prod_{b_k} \left(\int_{B_{b_k}} P(Y = k|X = x)dx \prod_{j\in\Omega\setminus\{k\}} \int_{B_{b_k}} P(Y \neq k|X = x)dx \right)^{s_{b_k}}.$$

- Standard application of symbolic likelihood
- Can only do this integral if each category has only one predictor X. (Will return to this shortly...)

Logistic regression models

Interesting result on existence of MLEs

Standard logistic regression:

1) $\hat{\beta} = \arg\max L(\boldsymbol{Y}, \boldsymbol{X}|\beta)$ exists and is unique if their is neither complete nor quasi-complete separation of the data (Albert and Anderson, 1984).

Histogram-based logistic regression:

2) $\hat{\beta}_S = \arg\max L_S(\boldsymbol{S}|\beta)$ exists and is unique if the set of histograms $(\boldsymbol{S}^1, \ldots, \boldsymbol{S}^K)$ does not exhibit complete nor quasi-complete separation[2] of the data (Whitaker, Beranger & Sisson, 2019; arxiv).

- 2) is stronger than 1), so 2) $\rightarrow$ 1)
- So if $\nexists\hat{\beta}_S \Rightarrow \nexists\hat{\beta}$ (i.e. if $\exists\hat{\beta} \Rightarrow \exists\hat{\beta}_S$)
- However $\hat{\beta}_S$ can exist where $\hat{\beta}$ does not
- So can do something in SDA that you can't with classical data[3]
- (Gets a bit crazy when $\boldsymbol{S} \rightarrow \boldsymbol{X}$... as then $\exists\hat{\beta}_S \rightarrow \nexists\hat{\beta}$!!)

[2] For modified definitions of separation compared to Albert and Anderson (1984)

[3] Not sure this is useful though!

Logistic regression models

Recall $L_S(\boldsymbol{S}|\beta)$ is proportional to

$$\prod_{k\in\Omega}\prod_{b_k}\left(\int_{B_{b_k}} P(Y=k|X=x)dx \prod_{j\in\Omega\setminus\{k\}}\int_{B_{b_k}} P(Y\neq k|X=x)dx\right)^{s_{b_k}}$$

- This works as standard application of symbolic likelihood
- However, can only do this integral if each category has only one predictor X.
- So either need to do d-dimensional computational integration or
- . . . abuse ideas from composite symbolic likelihoods

Logistic regression models

. . . Abuse ideas from composite symbolic likelihoods

- As we can integrate $L_S(\boldsymbol{S}|\beta)$ for univariate predictor
- Construct composite likelihood over all univariate predictor likelihoods
- Or over all 2-dimensional predictor likelihoods
- Or . . . etc.

The Good:

- Gets around high-dimensional integration (1-d is particularly good)

The Bad:

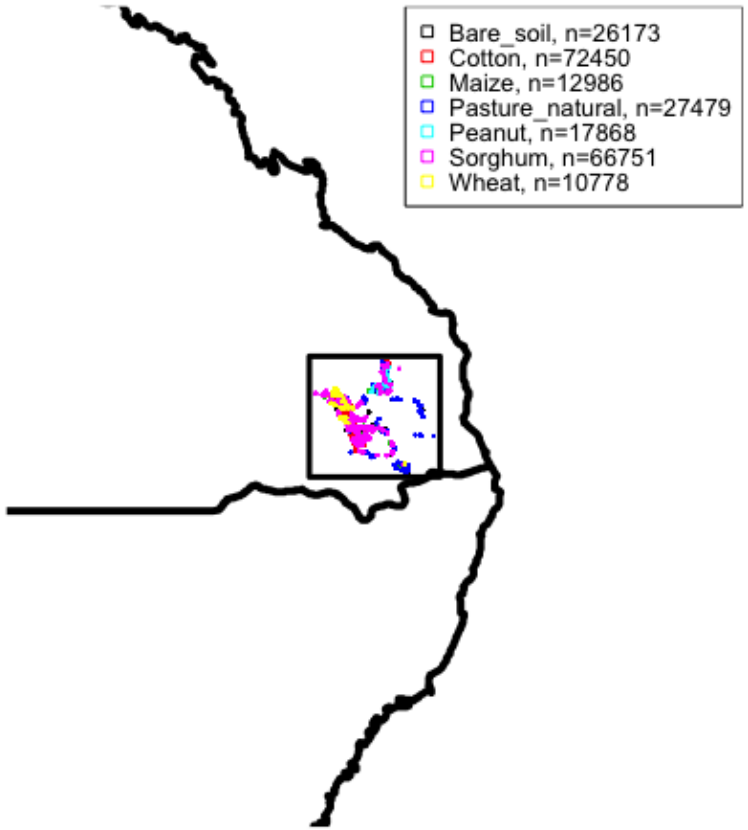
- Each marginal event is not an unbiased estimating equation
- So this is not a "true" composite likelihood
- The estimates of β will be biased
- All parameters depressed $\beta \downarrow 0$ (known result)

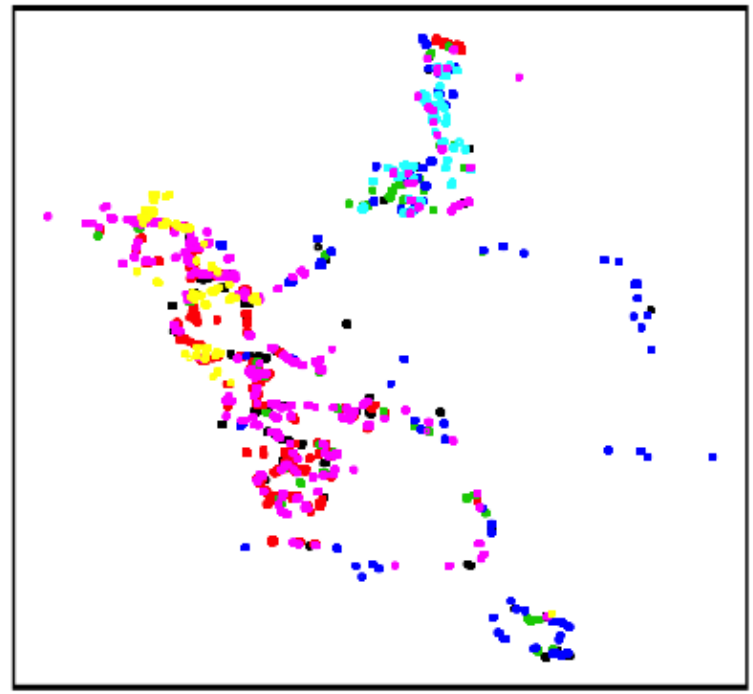
Logistic regression models

The (partial) Fix:

- Can reduce the bias using some modifications to this composite likelihood following ideas in a related context by Cramer (2007)
- Does **not** eliminate it
- However prediction can still be good if reduction in $\beta \downarrow 0$ is similar for all parameters
- This is what we found to happen in practice

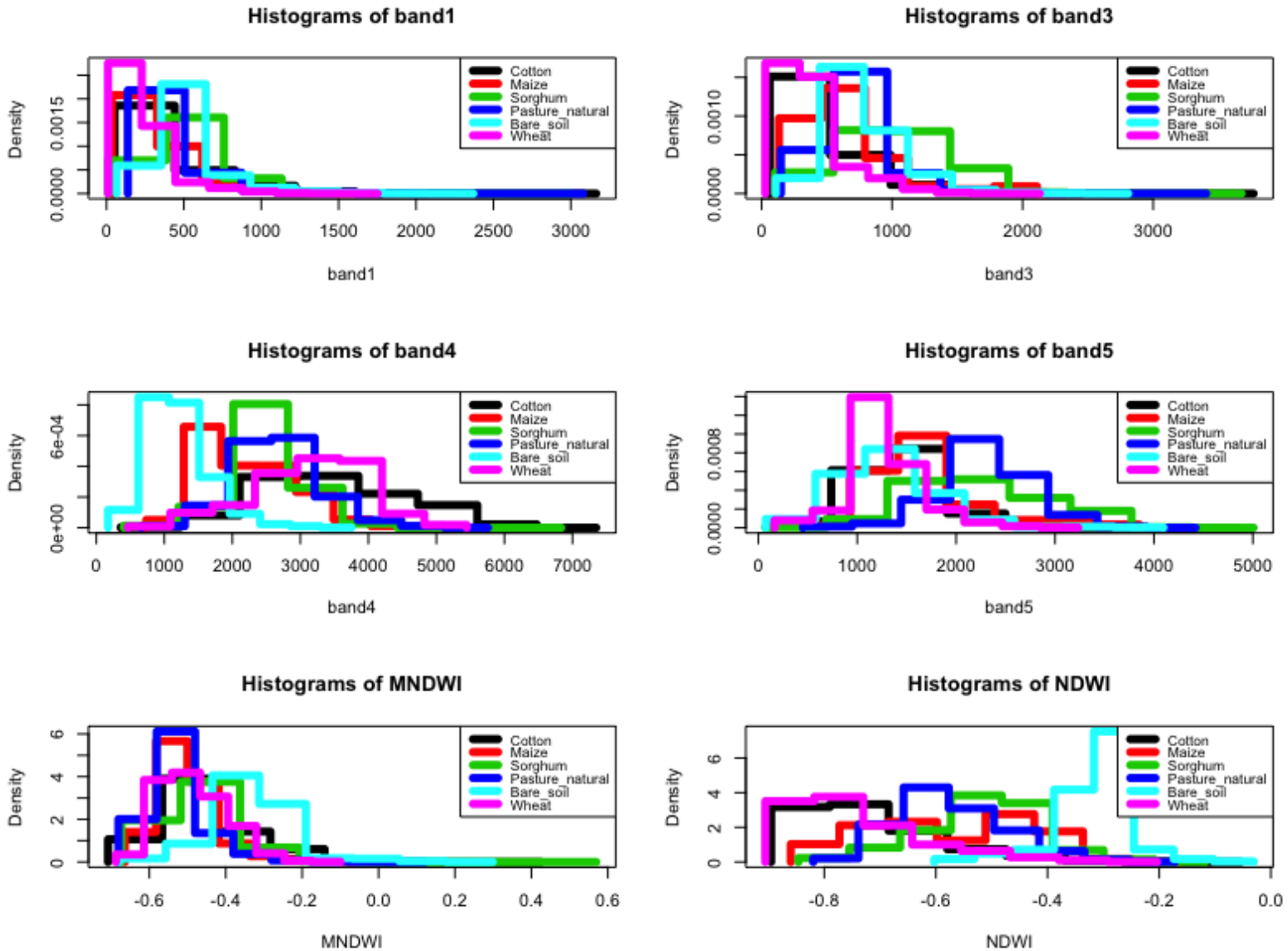
Prediction of crop types from satellite images





- $\sim$250K pixels with 7-dimensional predictor variable x
- 7 response categories with known ground truth
- Use (modified) symbolic composite likelihood over all 1-D predictors
- (Lasso regularisation included)

Prediction of crop types from satellite images



- Some predictors clearly identify crops even in 1-d (e.g. Bare soil)

Prediction of crop types from satellite images

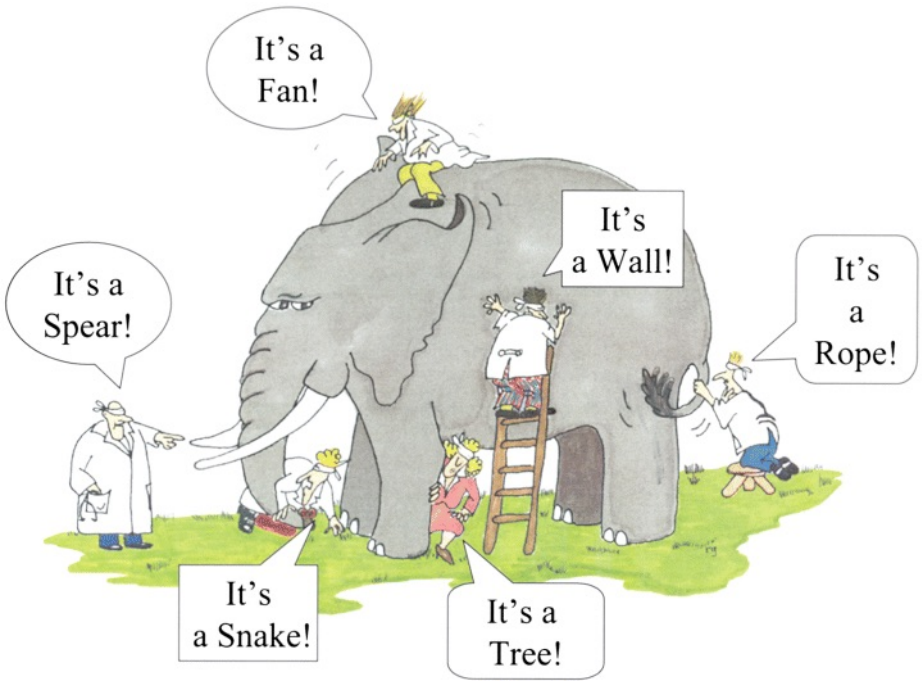
		Bins						
Crop type	N_k	6	8	10	12	15	20	$L_M(\boldsymbol{X}, \boldsymbol{Y}\|\boldsymbol{\beta})$
Cotton	72 450	90.5	90.6	92.8	93.6	94.0	94.1	92.2
Sorghum	66 751	74.6	74.8	75.7	76.4	76.2	76.3	80.3
Pasture Natural	27 479	75.7	75.4	76.0	76.8	77.0	77.1	77.6
Bare Soil	26 173	88.0	89.6	89.2	90.0	89.5	90.1	91.0
Peanut	17 868	81.2	81.3	81.5	81.5	81.9	81.6	82.9
Maize	12 986	9.7	9.9	10.2	10.4	10.3	10.4	14.2
Wheat	10 778	3.4	4.0	4.8	5.0	5.2	5.7	10.3
Overall	234 485	74.6	75.5	76.4	77.1	77.2	77.2	78.1
Time (secs)		(164)	(162)	(221)	(229)	(276)	(508)	(6071)

Table: Crop specific and overall prediction accuracies (%) using univariate marginal histograms with B bins. The likelihood optimisation times (in seconds) are reported in the last row. The full model is the standard multinomial likelihood $L_M(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{\beta})$ with LASSO regularisation.

- Pretty good results with 10 bins
  - Overall accuracy 76.4% (histogram) versus 78.1% (classical)
  - Poorer performance for less numerous crops (wheat, maize)
- . . . and 27× faster

More simulations and details in the paper.

Talk Outline



1. What is Symbolic Data Analysis?
   - Symbolic likelihood
   - . . . and its limitations
2. Symbolic composite likelihoods
   - Methods
   - Applications to spatial extremes
3. SCL for logistic regression
   - Methods
   - Application to satellite crop prediction
4. **Discussion**

Summary

Summary

- Symbolic composite likelihoods a natural extension of symbolic likelihoods
- Can avoid likelihood integration issues for some models
- Was useful for prediction even in models for which composite likelihoods are not suited (logistic regression)

Questions

- Other ways to avoid problematic integration in high dimensional histograms?
- Other ways to do high-dimensional regression (general) with symbolic likelihood?
- Etc.

UNSW SYDNEY

THANK YOU

Relevent Manuscripts:

- *Beranger, Lin & Sisson (2018). New models for symbolic data analysis.*
  `https://arxiv.org/abs/1809.03659`
- *Whitaker T., B. Beranger and S. A. Sisson (2020). Composite likelihood functions for histogram-valued random variables.* *Stat. Comput., In press.*
- *Whitaker T., B. Beranger and S. A. Sisson (2019). Logistic regression models for aggregated data.*
  `https://arxiv.org/abs/1912.03805`

Contact:

B.Beranger@unsw.edu.au
www.borisberanger.com