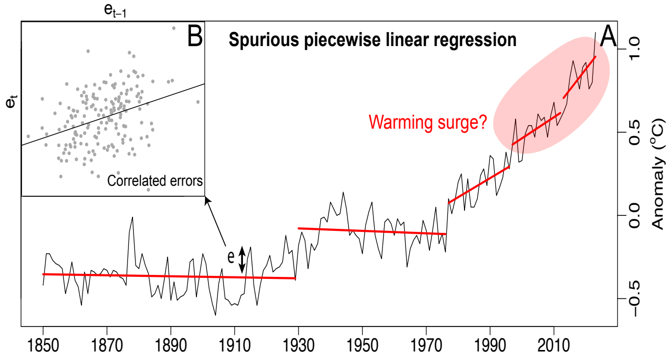
A Statistical View of Some Climate Change Issues.

Robert Lund
Statistics
Cal, Santa Cruz
Rolund@ucsc.edu

October 16, 2024
San Jose State University

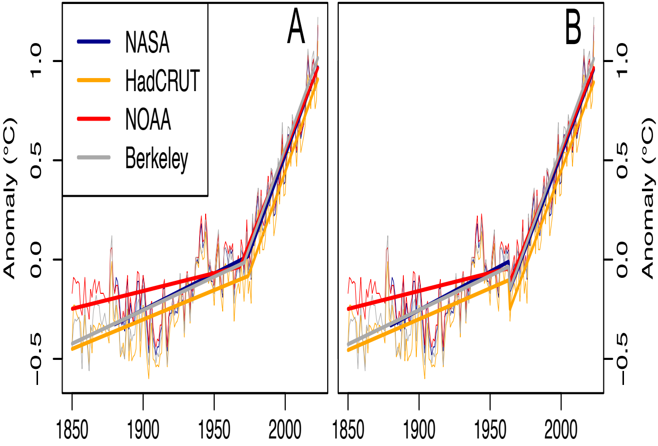
Global Warming

Is the Earth Experiencing a Recent Acceleration in Warming?



Four Global Temperature Series

2023 was the hottest year in recent times for the Earth, checking in way above the 1950s baseline mean. Some say we're entering in a new era of increased warming.



Joinpin Changepoint Models

A simple linear regression model for series $\{X_t\}$ with m changepoints is

$$X_t = E[X_t] + \epsilon_t.$$

The errors $\{\epsilon_t\}$ are a time series with $E[\epsilon_t] \equiv 0$. The mean is parametrized linearly in a piecewise manner:

$$E[X_t] = \begin{cases} \alpha_1 + \beta_1 t, & \tau_0 < t \leq \tau_1, \\ \alpha_2 + \beta_2 t, & \tau_1 < t \leq \tau_2, \\ \vdots & \vdots \\ \alpha_{m+1} + \beta_{m+1} t, & \tau_m < t \leq N. \end{cases}$$

This is called a changepoint model. We do not know m or $\tau_1, \ldots, \tau_m$.

Joinpin Models (continued)

Continuity of $E[X_t]$ imposes

$$\alpha_i + \beta_i \tau_i = \alpha_{i+1} + \beta_{i+1} \tau_i, \quad i = 1, 2, \ldots, m.$$

Fitting this model requires estimating the number of changepoints m, which is a subject of much current statistical research.

In the modern record (1850-Current), we find only one changepoint ($m = 1$) circa $\tau_1 = 1970$ in all series! No recent changepoint is found.

It is too early to statistically confirm any acceleration of warming! But 2024 is looking quite hot too!

A comment we extracted from our article

Just trust us. Just because nothing is actually changing doesn't mean it's not changing. And if you'll send our campaigns money, we'll fix it. Oh and also, one million people died from "COVID", the borders are secure and masks and vaccines prevent "COVID".
- The Democrats

Hurricane and Tropical Cyclone Changes

As the Earth warms, what happens to the heat imbalance between the Equator and the North Pole?

Will this produce more hurricanes, stronger hurricanes, or both?

Circa 2017.....Climatologist Kelvin Droegemeier opined that hurricanes will become stronger, but there will not be more of them in the future. This was the prevailing opinion of climatologists up to about 2020.

Rationale for this statement: Physics, GCM models, Trump's science adviser.

Popular Book: *Storm World* (2007) by Chris Mooney narrates the political mudslinging.

Poisson Likelihoods

Tropical cyclone counts are well-modeled by Poisson marginal distributions.

No autocorrelation is seemingly needed for this data (people vehemently argue about this).

See Livsey, Lund, Kechagais, and Pipiras (2018) *Annals of Applied Statistics* for negatively correlated count series in the Atlantic/Pacific Basins.

See Robbins, Lund, Gallagher, and Lu (2010), *Journal of the American Statistical Association* for more on this pre 2010 Atlantic data and categorical changepoints.

Poisson Dynamics

The Poisson count distribution:

$$P[X = k] = \frac{e^{-\lambda}\lambda^k}{k!}, k = 0, 1, \ldots$$

Our model for the data: independent Poisson counts, no trends, with a mean shift at each changepoint time.

$$X_t \overset{\text{Indep}}{\sim} \text{Poisson}(\lambda_t);$$

$$\lambda_t = \begin{cases} \mu_1, & 1 \leq t < \tau_1, \\ \mu_2, & \tau_1 \leq t < \tau_2, \\ \vdots & \vdots \\ \mu_{m+1}, & \tau_m \leq t \leq N \end{cases}.$$

Poisson Model

Let $\boldsymbol{\Theta} = (m; \mu_1, \tau_1; \mu_2, \tau_2; \ldots; \mu_m, \tau_m; \mu_{m+1})'$ denote all model parameters.

The model likelihood of getting the data $X_1 = x_1, \ldots, X_N = x_N$ is

$$L(\boldsymbol{\Theta}) = \prod_{t=1}^{N} P[X_t = x_t] = \prod_{t=1}^{N} \frac{e^{-\lambda_t} \lambda_t^{x_t}}{x_t!}.$$

To prevent overfitting, we penalize the likelihood for the model parameters with the MDL penalty

$$\text{MDL}(\boldsymbol{\Theta}) = \sum_{i=1}^{m+1} \ln(\tau_i - \tau_{i-1}) + 2\ln(m) + 2\sum_{i=1}^{m} \ln(\tau_i).$$

A penalized likelihood scheme minimizes

$$-2\ln(L(\boldsymbol{\Theta})) + \text{MDL}(\boldsymbol{\Theta})$$

over all parameters in $\boldsymbol{\Theta}$.

Penalized Likelihood Minimization

For a fixed changepoint configuration $(m; \tau_1, \ldots, \tau_m)$, the best estimates of the means are the sample averages

$$\hat{\mu}_i = \frac{1}{\tau_i - \tau_{i-1}} \sum_{t=\tau_{i-1}}^{\tau_{i-1}} X_t,$$

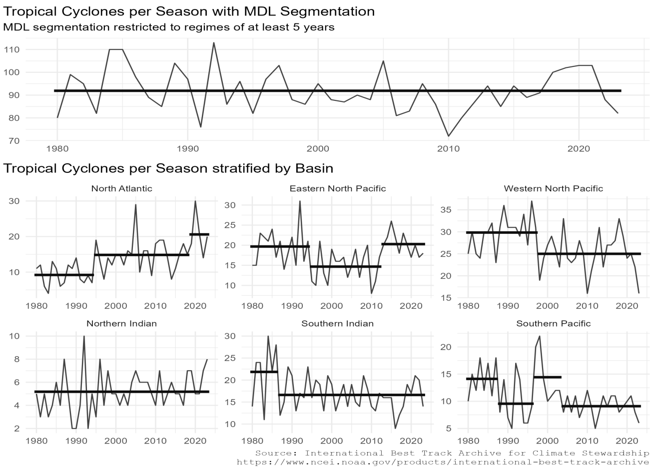
This reduces our problem to minimizing

$$2 \sum_{i=1}^{m+1} (\tau_i - \tau_{i-1}) \hat{\mu}_i [1 - \ln(\hat{\mu}_i)] + \mathrm{MDL}(m; \mu_1, \tau_1, \ldots, \tau_m)$$

over all changepoint configuration parameters $m; \tau_1, \ldots, \tau_m$.

This minimization is done with a genetic algorithm search.

Recent Global Tropical Cyclone Count Changes



Snow Presence/Absence Trends

- We have about 1700 grid cells over land that see some snow in the Northern Hemisphere.
- For each cell, we have a weekly time series $\{X_t\}$ indicating whether snow is present ($X_t = 1$) or absent ($X_t = 0$) during week t.
- The series is extracted from weekly satellite flyover pictures spanning the years 1967-2021 inclusive.

Goal: Model, quality check, and estimate trends in this data and quantify their uncertainties. Communicate the results effectively to the layman.

Comments

- The study is for trends in snow presences, not snow depths.
- The data are highly seasonal with a period of $T = 52$ weeks.
- Existing statistical methods to estimate snow presence trends in the hydrology community are primitive.

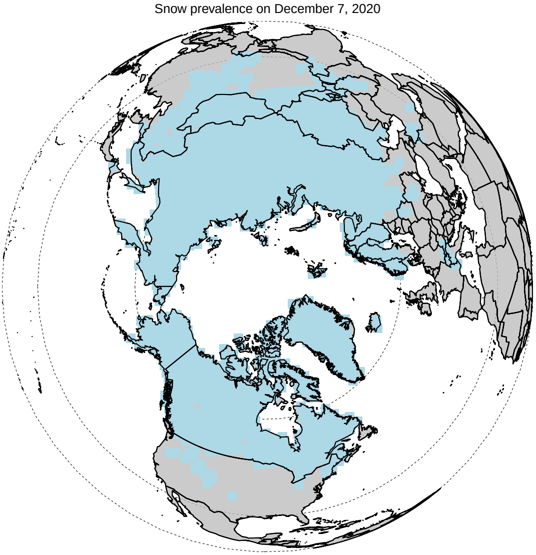


Figure: NH snow presence reported for the first week of December, 2020.

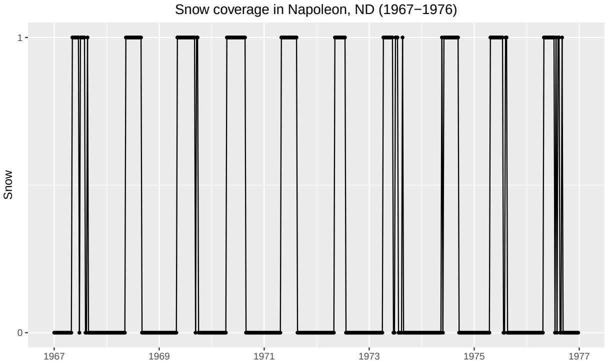


Figure: Napoleon, ND snow presences

The Model

The binary series $\{X_t\}$ will be modeled as a two-state Markov chain with the time varying transition matrix

$$\mathbf{P}(t) = \begin{bmatrix} p_{0,0}(t) & p_{0,1}(t) \\ p_{1,0}(t) & p_{1,1}(t) \end{bmatrix}.$$

We work in winter center years that begin on August 1 and start each cell with bare ground:

Trends and Seasonality

Transition probabilities are modeled via the logit link

$$p_{0,1}(t) = \frac{1}{1+\exp(-m_t)}, \quad p_{1,0}(t) = \frac{1}{1+\exp(-m_t^*)},$$

where m_t and m_t^* contain seasonal effects and trends.

These quantities are posited to have the additive form

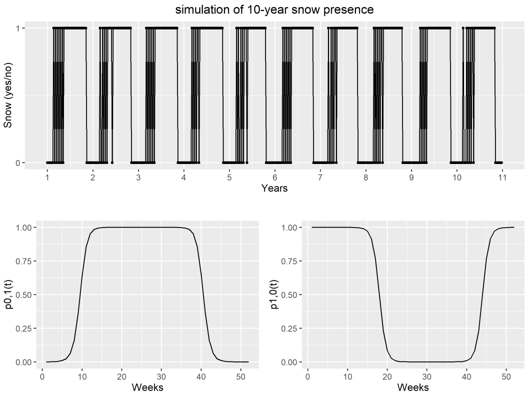
$$m_t = \mu_t + \alpha t, \quad m_t^* = \mu_t^* + \alpha^* t.$$

The parameters μ_t and μ_t^* contain seasonal effects with period $T = 52$:

$$\mu_t = A_0 + A_1 \left[\cos\left(\frac{2\pi(t-\kappa)}{T}\right)\right], \quad \mu_t^* = A_0^* + A_1^* \left[\cos\left(\frac{2\pi(t-\kappa^*)}{T}\right)\right].$$

A Simulated Series

- For $p_{0,1}(t)$ set $A_0 = 3$, $A_1 = 10$, $\tau = 25$, and $\alpha = 0$.
- For $p_{1,0}(t)$ set $A_0^* = 0$, $A_1^* = 10$, $\tau^* = 5$, and $\alpha^* = 0$.



Let $\boldsymbol{\Theta}$ denote all model parameters in $\{m_t\}$ and $\{m_t^*\}$ for a fixed cell. These include A_0, A_1, κ, α and their starred counterparts.

The likelihood of $\boldsymbol{\Theta}$, denoted by $L(\boldsymbol{\Theta})$, is derived from the Markov property:

$$\ln(L(\boldsymbol{\Theta})) = \sum_{t=2}^{N} \ln\left(p_{X_{t-1},X_t}(t)\right).$$

The $p_{i,j}(t)$ depend on $\boldsymbol{\Theta}$. Numerically maximizing this likelihood provides estimates of the components in $\boldsymbol{\Theta}$, which will be used in assessing variability (uncertainty) margins of the trends.

Estimating $\boldsymbol{\Theta}$ via gradient step and search is relatively easy.

Let S_n be the number of weeks of snow observed during year n:

$$S_n = \sum_{\nu=1}^{T} 1_{[X_{(n-1)T+\nu}=1]}.$$

The linear rate of snow cover change is estimated via

$$\hat{\beta} = \frac{\sum_{k=1}^{d} S_k(k-\bar{k})}{\sum_{k=1}^{d}(k-\bar{k})^2}.$$

The units of β are weeks of snow cover gained/lost per year, but can be rescaled. Here, $\bar{k} = (d+1)/2$ is the average year index and $d = 53$ is the number of years in the study.

The variance of $\hat{\beta}$ can be obtained from $\text{Cov}(S_k, S_\ell)$ for every $k, \ell \in \{1, 2, \ldots, N\}$. $\text{Cov}(S_k, S_\ell)$ depends on parameters in $\boldsymbol{\Theta}$

The variance $\hat{\beta}$ is estimated from parameters estimated for $\boldsymbol{\Theta}$ in the likelihood fit.

Asymptotic normality allows one to test that $\beta = 0$ against alternatives.

$$Z = \frac{\hat{\beta}}{\text{Var}(\hat{\beta})^{1/2}}.$$

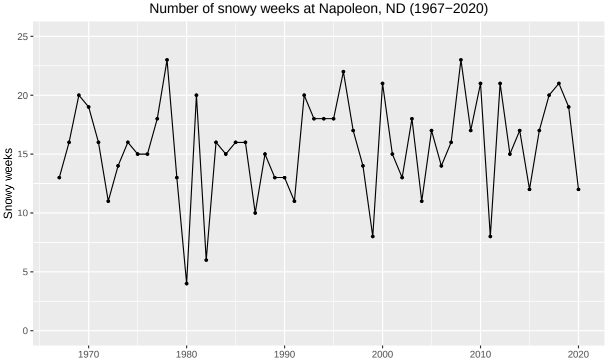


Figure: Number of snow covered weeks during the 1967-2020 period for the Napoleon, ND cell.

Table: Model parameter estimates and their standard errors for the Napoleon, ND grid.

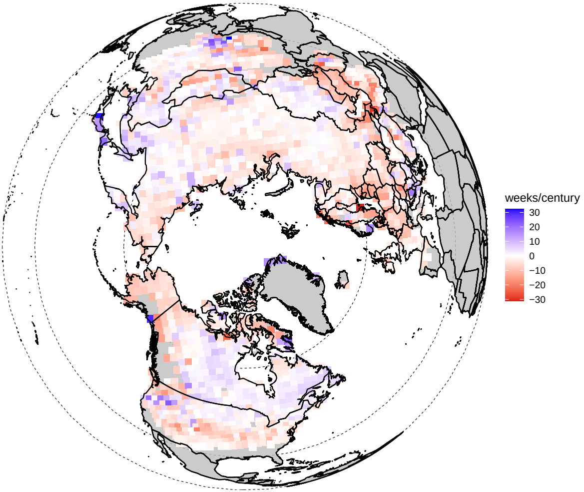
Parameter	A_0	A_1	κ	α
Estimate	-3.2016	4.1499	24.3492	0.0000382
Standard Error	0.2538	0.2936	0.26460	0.0001315
Parameter	A_0^*	A_1^*	κ^*	α^*
Estimate	1.7258	3.7889	49.8375	-0.0004935
Standard Error	0.3774	0.4139	0.3800	0.0001273

- For this grid, $\hat{\beta} = 0.038613$, $\text{SE}(\hat{\beta}) = 0.0247$.
- This translates to 3.86 additional snowy weeks per century.
- The test statistics is $Z = 1.5633$, with a two sided p-value of 0.1180.

Results for the Entire Northern Hemisphere

- 1,613 cells were deemed to have useable data.
- 578 cells (35.83%) report a positive $\hat{\beta}$.
- 1,035 cells (64.16%) report a negative $\hat{\beta}$.
- The average trend over all 1,613 analyzed cells entails 1.901 weeks of snow cover lost per century.
- These results are comparable to those in the 4^{th} IPCC report.

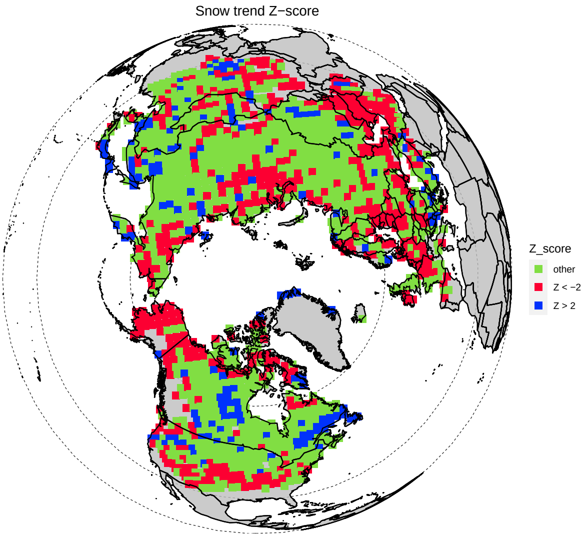
Snow cover trends 1967−2020



Hypothesis Test Results

Test H_0: $\beta = 0$ versus H_A: $\beta \neq 0$.

- Z-scores are deemed significant if they exceed 2.0 in absolute value.
- Red colored Z-scores indicate declining snow presence.
- Blue colored Z-scores indicate increasing snow presence.


Snow trend Z−score
Z_score
other
Z < −2
Z > 2

Conclusions

- Statisticans are the arbiters of climate change.
- There are many important climate issues to resolve.
- UCSC awaits your PhD application.
- Statisticians can "play in anyone's back yard". Data is everywhere in the modern age and there is much to explore beyond climate! You will need every bit of math you've ever learned and much more!

The California, Santa Cruz Statistics Department

- We have ten faculty and two lecturers.
- We conduct research in Bayesian methods, data science, time series, spatial statistics, census problems, human health, and environmental and climatological applications.
- Our department has significant research funding.
- The PhD degree requires one year of classes and typically 3-4 years of research.

Some Websites

General Information:

- http://grad.soe.ucsc.edu
- http://grad.soe.ucsc.edu/stats
- https://catalog.ucsc.edu/en/current/General-Catalog/Academic-Programs
- https://engineering.ucsc.edu/departments/statistics

Apply by mid December at:

https://applygrad.ucsc.edu/apply

Tell us that Robert Lund invited you to apply!

Some References

- Davis, Lee, and Rodriguez-Yam (2006), Structural Break Estimation for Non-stationary Time Series, *Journal of the American Statistical Association*, **101**, 223-239.
- Mooney (2007), *Storm World*, Harcourt Books.
- Lu, Lund, and Lee (2010), An MDL Approach to the Climate Segmentation Problem, *Annals of Applied Statistics*, **4**, 299-319.
- Robbins, Lund, Gallagher, and Lu (2011), Changepoints in the North Atlantic Tropical Cyclone Record, *Journal of the American Statistical Association*, **106**, 89-99.
- Li and Lund (2012), Multiple Changepoint Detection via Genetic Algorithms, *Journal of Climate*, **25**, 674-686.

More References

- Livsey, Lund, Kechagais, and Pipiras (2018), Multivariate Poisson Count Time Series and their Application to Major Hurricane Counts, *Annals of Applied Statistics*, **12**, 408-431.
- Fisher, Lund, and Robbins (2020), A Statistical Analysis of North American Tropical Cyclone Changes, In: *Evaluating Climate Change Impacts*, Edited by Lyubchich, Gel, Kilbourne, Miller, Newlands, and Smith, CRC Press, 25-41.
- Jia, Lund, Kong, Dyer, Woody, and Marron (2023), Trends in Northern Hemispheric Snow Cover, *Journal of Hydrometeorology*, **24**, 1137-1154.
- Beaulieu, Gallagher, Killick, Lund, and Shi (2024). Is a Recent Surge in Global Warming Detectable?, To appear, *Nature Communications, Earth & Environment*.
- Diawara, Fisher, Lund, and Wehner (2024+). Recent Changes in the Earth's Tropical Cyclone Environment, to be submitted to *Nature*.