Climate and statistics
Outline

Anomalies
Comparing climate models to data
Downscaling and bias correction
The fit of models

Climate is the distribution of weather
To reasonably estimate a distribution (from data or from models) need a relatively long stretch of data—WMO suggests at least 30 years
How well does the CMIP5 experiment used in the recent IPCC assessment work for describing annual global mean temperature?
The greenhouse effect

Joseph Fourier (1768-1830) realized that Earth ought to be a lot cooler than it is.
John Tyndall (1820-1893) found that water vapor and CO₂ are greenhouse gases
Svante Arrhenius (1859-1927) calculated how changes in CO₂ can heat the planet
A simple climate model

What comes in

\[ S\pi r^2 (1 - a) \]

- Solar constant
  - 1365 W/m\(^2\)
- Earth's albedo
  - 0.3

must go out

\[ 4\pi r^2 \varepsilon \sigma T^4 \]

- Effective emissivity
  - (greenhouse, clouds)
  - 0.61
- Stefan's constant
  - \(5.67 \times 10^{-8} \text{ W/(K}^4 \cdot \text{m}^2)\)
Solution

\[ T^4 = \frac{1365 \times 0.7}{4 \times 0.61 \times 5.67} \times 10^8 \]

Average earth temperature is \( T = 288 \text{K} \) (15°C)

One degree Celsius change in average earth temperature is obtained by changing

- solar constant by 1.4%
- Earth’s albedo by 4.5%
- effective emissivity by 1.4%
But in reality…

The solar constant is not constant
The albedo changes with land use changes, ice melting and cloudiness
The emissivity changes with greenhouse gas changes and cloudiness
Need to model the three-dimensional (at least) atmosphere
But the atmosphere interacts with land surfaces…
…and with oceans!
The climate engine I

If Earth did not rotate:
tropics get higher solar radiation
hot air rises, reducing surface pressure
and increasing pressure higher up
forces air towards poles
lower surface pressure at poles
makes air sink
moves back towards tropics
The climate engine II

Since earth does rotate, air packets do not follow longitude lines (Coriolis effect)
Speed of rotation highest at equator
Winds travelling polewards get a bigger and bigger westerly speed (jet streams)
Air becomes unstable
Waves develop in the westerly flow (low pressure systems over Northern Europe)
Mixes warm tropical air with cold polar air
Net transport of heat polewards
Modeling the atmosphere

Lagrangian approach

Conservation of mass:
\[
\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial p} = 0
\]

Conservation of momentum:
\[
\frac{dc}{dt} = -2\Omega \times c - \frac{\text{grad}(p)}{\rho} + \text{grad}(\phi) + \text{friction}
\]

\( c = (u, v, w); \rho \text{ is geopotential}; \phi \text{ is density} \)
Thermodynamics:
\[
\frac{dT}{dt} = \frac{Q}{c_p} + \kappa \frac{T_w}{p}
\]

Conservation of water vapor:
\[
\frac{dq}{dt} = s(q) + D
\]

Hydrostatic equilibrium:
\[
\frac{\partial \phi}{\partial p} = -\frac{RT}{p}
\]
The issue of gridding

Hurricanes
Clouds
Glaciers
Comparing two distributions

Location shift: $X + \delta \sim Y$

Location-scale: $\alpha X + \delta \sim Y$

More general: $X + \Delta(X) \sim Y$

Having $m$ observations from $Y$ and $n$ observations from $X$ we estimate

$$\hat{\Delta}(x) = G_m^{-1}(F_n(x)) - x$$
Comparing global climate models to data
Anomalies

Comparison to “normal”
Normal = 30 yr average
Different baselines
Helps for regional trends
Really residuals
So fit a model (trend + seasonal + covariates + variability)
Global mean temperature

Reference period 1951-1980
30-year distributions

1930-1959

CMIP5 models

GISS data

1970-1999

CMIP5 models

GISS data
Comparing location and spread
Effect of anomalies

Scaled KS Statistic

End year

NCEI
Dynamical downscaling

Global models are very coarse
Regional models are driven by boundary conditions given by global model runs
Swedish temperature minima

SMHI synoptic stations in south central Sweden, 1961-2008. SMHI regional model (open air & snow)
Seasonal minima (d=1 DJF, d=2 MAM, d=3 JJA, d=4 SON).

Due to an unusual series of collection dates, data at Borlänge Flygplats was kept aside for crossvalidation purposes.
Spatial models

\[-m_t(s) \sim \text{GEV}(\mu_t(s), \sigma(s), \xi)\]

where

\[\mu_t(s) = \beta_0(s) + \beta_1(s)(t - 1961) / 50 + \sum_{d=2}^{4} \beta_d(s)1(d_t = d) \]

\[\beta_i(s) \sim \text{GP}(\mu_i, \sigma_i(1 - \exp(-\theta_i d(s))))\]

\[\log \sigma(s) \sim \text{GP}(\mu, \sigma(1 - \exp(-\theta d(s))))\]

Both for data and model output.
RCM temperature minima

DJF

MAM

JJA

SON

Site 1

Site 2

Site 3
Fit of GEV-distribution

Fit is better when $\sigma$ depends on $s$.
Seasonal fits

Obs. Stations

<table>
<thead>
<tr>
<th>Post mean SON</th>
<th>Post mean DJF</th>
<th>Post mean MAM</th>
<th>Post mean JJA</th>
</tr>
</thead>
</table>

RCM

| Post mean SON | Post mean DJF | Post mean MAM | Post mean JJA |
(Dis)agreement between RCM and data

Seasonal effects quite similar
Similar spatial scale
Similar shape parameter
Temporal trend substantially lower in model output
Data trend about 0.4-1°C /decade
(lower than the annual model)
Norwegian winter precip
Bias correction

Need downscaled precipitation projections for adaptation plans

Bias correction for downscaled reanalysis

Apply to downscaling historical GCM

If works, apply to downscaling GCM projections

Correction more important for large quantiles than for entire distribution
Full quantile correction

Applying the Doksum shift we get

\[
z_{it'}^{\text{cal}} = z_{it'}^{H} + \hat{\Delta}_i(z_{it'}^{H})
\]

\[
= z_{it'}^{H} + \hat{G}_i^{-1}(\hat{F}_i(z_{it'}^{H})) - z_{it'}^{H}
\]

\[
= \hat{G}_i^{-1}(\hat{F}_i(z_{it'}^{H}))
\]

Rejections (fit to 80%, tested on 20%):

Raw 77%
Corrected 18%
Corrected GCM 79%
Single quantile correction

\[ \log(Y^q) = \alpha + \text{diag}(\log(X^q))\beta + \epsilon \]

\[ \beta \sim \text{GP}(0, \Sigma(\nu, \kappa, \phi)) \]

Replace with smoothed version
Bias-correction of 95\textsuperscript{th} precipitation percentile