Precipitation and Crop Yield: A Statistical Model

Lynne Seymour, Department of Statistics, University of Georgia, USA. seymour@uga.edu

Also: Rosalind Bueckert, Department of Plant Sciences, University of Saskatchewan

Ross Iaci, Department of Mathematics, College of William & Mary

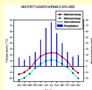
QiQi Lu, Department of Mathematics & Statistics, Mississippi State University

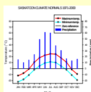
The University of Georgia

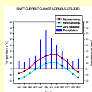
The Question

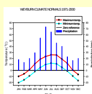
How has climate change in Saskatchewan affected crop yield?

Saskatchewan is a Canadian prairie province situated at the top of the Northern Great Plains between Manitoba and Alberta, and above North Dakota and Montana. The climate is semi-arid, with a five month cold winter (extreme temperature of -47 °C) and relatively warm dry continental summer (extreme temperature of 42 °C), low annual precipitation (range from 370 to 420 mm) with the majority of rainfall (265 to 320 mm) received in April to August in the cropping regions. The last 20 years' climate normals are shown at right for Melfort (cool/wet), Saskatoon, Swift Current (hot/dry) and Weyburn (hot/dry). Saskatchewan agriculture is predominantly rainfed and the province is known for variable weather and famous droughts including the "Dirty Thirties", the 1950s, the 1980s, the 1990s, 2001 and 2003.



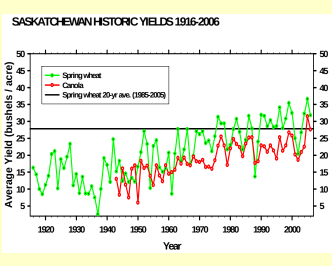


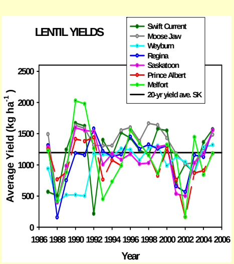




The Crop Yield Data

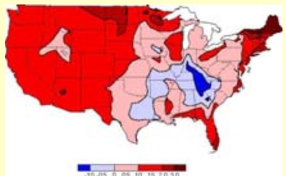
Since agricultural settlement, Saskatchewan crops have been dominated by spring wheat. The traditional rotation of wheat and fallow is still practiced in the brown soil zone. Canola (oil-seed rape) was introduced in the 1940s and is common in the dark brown and black soil zones. In the mid 1970s to mid 1980s zero tillage has been adapted by the majority of farms, with decreasing frequency of fallow. Pulses (pea, lentil, chickpea, dry bean) are also an increasing part of the cropping rotation; pea and lentil are the more dominant pulse crops. Field pea has been produced since settlement; lentil was introduced in the 1970s, while chickpea in the late 1980s.



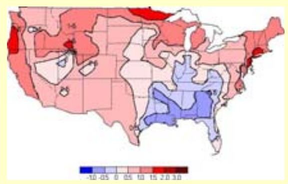


The Climate Data

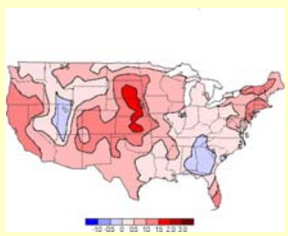
Lu, Lund, and Seymour (2005) demonstrated that average monthly temperatures are changing at different rates depending on the month in question. As some climatologists have predicted, they found that winter months are warming faster, in general, than summer months in the 48 contiguous United States.



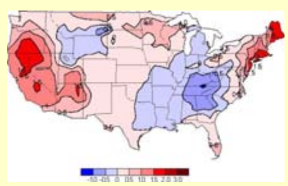
Winter



Spring



Summer



Fall

These maps show *trends* in °C for each of the four seasons. Clearly, the seasonal trends are not the same!

Ideally, we would like to build trend maps such as the above for all of North America. However, for the immediate question, we only need such a map for Saskatchewan.

In fact, we do have annual climate data, since annual averages or totals are relatively easy to find. However, to build these maps and account for the seasonal differences, we need monthly climate data, which tends to be harder to find.

A Challenge: We expect monthly precipitation changes to behave differently, and this needs to be determined in a statistically rigorous fashion. Research presented at this 4th Workshop of the IPWG will likely provide a suitable answer to this question.

Again ideally, we would like to have trend maps for monthly precipitation such as these temperature maps.

Other climate-based variables that should be included in an initial statistical model include light intensity, ground temperature, etc. Dr. Bueckert is spending her sabbatical year gathering the pertinent data for this modelling adventure.

The Statistical Challenge

There are several statistical challenges to modelling crop yield. The statistical model for crop yield that we will begin with will be multivariate in nature, but also possibly nonlinear:

```
AnnualStandardizedYield  =
f(MonthlyTemperatures, MonthlyPrecipitation,
etc.) + error
```

Note that this model is multivariate, and that the inputs are temporally and spatially correlated!

In fact, the variables do not need to be restricted to vector quantities: we will also entertain the idea that they are in fact maps! We give the two cases below.

Case 1: The variables are vector time series. Then, the resulting analysis will be a multivariate time series problem with possibly nonlinear relationships between the seasonal climate components and the yield.

Case 2: The variables are geographical maps indexed by time. In this case, we will take a page from the statistical analysis of fMRI (brain imaging) data, and reshape the maps into vectors; that is, we will take an $n \times n$ matrix of values that constitute the map, and stack the columns so that we have an $n \cdot n$ vector of values. We may then proceed as in Case 1.

`AnnualStandardizedYield`

These yields need to be standardized to control for several factors that contribute to the increase in yield over time:

- <> Changes in the number of hectares planted over the years;
- <> Changes in farming practices; and
- <> The introduction of "high-yield" cultivars.

This sort of standardization may not be easy to accomplish, and may require its own modelling adventure.

`MonthlyTemperatures` and `MonthlyPrecipitation`

These climate variables need to feature monthly averages, since climate change in fact varies greatly across the months.

The Methodology

The main tool to be used in building these models will be a new *Generalized Correlation Analysis* (GCA; Iaci, 2007). This dimension-reduction technique is an extension of Canonical Correlation Analysis, which is common and is used to uncover linear relationships among multivariate observations.

GCA is able to uncover even non-linear relationships! This data will allow us to extend GCA to multivariate time series data, which is the next step in its statistical evolution.

Yes... And then???

The data required for the modelling adventure is being collected, as we noted in the panel to the left. Thus, to our great disappointment, it was not available in time for this presentation.

For the 5th Workshop if the IPWG, we do expect to have some interesting results to present.

Bibliography

R. Iaci (2007). Doctoral Dissertation, Department of Statistics, University of Georgia.

Q. Lu, R. Lund and L. Seymour (2005). An Update of United States Temperature Trends. *Journal of Climate* **18** 4905-4913.