Time Series HW 5

Allison Theobold, Andrea Mack, and Kenny Flagg

October 6, 2016

Groups of up to 3 allowed.
Read Vincent and Mekis's 2006 paper and answer the following questions. I am also posting Zhang et al. (2000) for the reference on the method of trend estimation used.
When you get to Table 3, it is helpful to add a vertical linear between the first "positive" column and the second "negative" column to understand what is presented.
For the following questions, ignore the "national trends" part of it except for possibly the last question. There is plenty to discuss without that piece of the paper.

1. *What is/are the research questions in Vincent and Mekis? You can be a little bit general here – just a couple of sentences on what they are trying to explore.*

2. *How many response variables are they analyzing? Explain one temperature and one precipitation based response that they are using in detail.*

3. *How are they assessing trends? You will need to read pages 402-405 in Zhang et al. (2000) to get all the details. I just want the general discussion of the method – do this in a few sentences. Then note their reason(s) for using this method?*

4. *I think I successfully extracted their counts from Table 3 into the provided Table3.csv file but you should check my results. Use that information to provide the total number of tests they considered for the shorter and then longer series analyses and the total number of each of those that were "significant". What is the minimum and maximum number of tests performed across the different response variables for each length of series?*

   The total number of tests done on temperature indicies was 1849 and the total number of tests done on the precipitation indicies was 2880 . The total number of tests done on the shorter time period, 1950-2003 was 3441 and the total number of tests done on the longer time period, 1900-2003 was 1288 . The total number of significant tests found in the shorter time period, 1950-2003 was 919 and the total number of significant tests found in the longer time period, 1900-2003 was 580 .

	Longer	Shorter
Total Tests	1288	3441
Total Significant	580	919

What does he mean max and min number of tests?

5. *Again focused on Table 3, if all tests are performed at the 5% level, overall how many results should they have expected to be "significant" within the shorter and then longer trends being considered? How does that compare to their total number of "significant" results on each? Do the same comparison for each response variable in the table (total expected for that variable and number observed for shorter and then longer series). Are there any response variables where the difference in observed "significant" results and expected is such that you don't think there really is much overall evidence on that variable?*

   At the 95% confidence level, we cannot say that exactly 95% of the tests done in a study will detect actual differences for a fixed number of tests. But in the long run, we expect 95% of tests to come to correct conclustions. By chance, we can estimate that 5% of the tests will falsely find significant results. For the tests done in the shorter time frame, that is about 172.05 and in the longer time frame, that is about 64.4 that we expect to wrongly yield significant results.

	Estimated FDR	Total Significant
Longer	1224	580
Shorter	3269	919

Total Tests	Total Signif	Expected Signif	Percent Signif	Index	Type
159	49	8	31	Frost	Temp
167	41	8	25	Cold	Temp
167	71	8	43	Cold	Temp
148	30	7	20	Summer	Temp
167	38	8	23	Warm	Temp
167	58	8	35	Warm	Temp
155	36	8	23	Diurnal	Temp
162	19	8	12	Standard	Temp
221	44	11	20	Annual	Precip
209	48	10	23	Snow	Precip
208	107	10	51	Days	Precip
214	146	11	68	Days	Precip
208	66	10	32	Simple	Precip
214	71	11	33	Simple	Precip
208	36	10	17	Max	Precip
239	16	12	7	Highest	Precip
214	16	11	7	Very	Precip
214	27	11	13	Heavy	Precip

longer_var.tot	longer_var.sig	longer_var.expected	longer_var.prop	index	type
65	40	3	25	Frost	Temp
75	32	4	19	Cold	Temp
74	62	4	37	Cold	Temp
62	13	3	9	Summer	Temp
75	18	4	11	Warm	Temp
74	40	4	24	Warm	Temp
59	34	3	22	Diurnal	Temp
73	16	4	10	Standard	Temp
76	22	4	10	Annual	Precip
72	24	4	11	Snow	Precip
72	52	4	25	Days	Precip
73	54	4	25	Days	Precip
72	42	4	20	Simple	Precip
73	41	4	19	Simple	Precip
72	47	4	23	Max	Precip
75	10	4	4	Highest	Precip
73	12	4	6	Very	Precip
73	21	4	10	Heavy	Precip

The lower the number of significant tests, the more I am concerned. Across all indicies for both temperature and precipitation measurements, by chance we should expect 7-12 significant results. Particularly in the *Very Wet Days Index* for precipitation, we see that 7% of the tests were significant, which is getting very close to the 5% that we expect to be significant by chance when setting the 95% confidence level. Corrections for multiple testing should have been used.

6. *Consider one panel of one of the figures that displays the trend and test results on a map of Canada. Note which figure you are discussing.*

   (a) *What is being displayed and what does that suggest about changes on this metric? Is there any issue with their display choices?*

   (b) *Does it appear that there is some sort of spatial structure to the trend results? What does this suggest about the tests on that variable?*

7. *Provide at least two things you might have done differently if you were involved in this paper. This can be in terms of methods used, the way things were interpreted, or graphical presentation of results. Assume that the data set is more or less what you have to work with and you want to address the same research questions.*

References

Vincent, L. and Mekis, E. (2006) Changes in Daily and Extreme Temperature and Precipitation Indices for Canada over the Twentieth Century. *Atmosphere-Ocean*, 44(2): 177-193.

Zhang, X., Vincent, L., Hogg, W., and Niitsoo, A. (2000) Temperature and precipitation trends in Canada during the 20th century. *Atmosphere-Ocean*, 38(3): 395-429.

R Code

4.

5.