Pan Arctic Outlook

Christopher Randles

Public outlook, statistical method

1. **Extent Projection = 4.45 m km^2 +/- 95% confidence interval of 0.32 m**

2. **Method** – New statistical method.

   The new method predicts changes from a weighted average of extent and area at 31 August to September average extent using a relative measure of incoming solar radiation that could be captured by ocean that in other years might have been ice and reflected the radiation.

3. **Rationale**

   My attempts seem to indicate that it is important to use the curved shape of the decline in extent. My old adjusted gompertz fit method very explicitly used curve fitting but it doesn't have to be so explicit and my new method uses the curve much less explicitly. The change from 31 Aug to September average does not show a particularly curved shape whereas at 31 Aug and September average both have a curved shape of decline. By predicting the difference, this removes much of the need to predict the curved shape leaving the curve in the decline at 31 Aug built in.

   <u>New predict changes from 31 August Method</u>

   (31 Aug Extent minus September average extent) is a noisy number to predict. I found that using a weighted average of extent and area at 31 Aug made the falls easier to predict. For the middle of August, it was better to give more weight to area than extent. Unsurprisingly the best weights to apply moved more to extent than area as we approach the minimum. By 22 Aug more weight should be given to extent and for 31 Aug data, the extent was given 6.6 weight to 1 for area.

   The falls from this weighted average could be predicted using year or extent or area or the weighted average or various other predictors. The area seems to perform well. If carefully tuned, the weighted average may do better but this could easily be over-fitting which could be a serious problem with the alternatives available and considered below.

   Why does area perform so well as a predictor when we are trying to predict extent? A major component of the changing system is the albedo feedback. So a major part of the answer could well be that the extra energy captured by the ocean that was formerly ice is related to area of ocean not extent.

Area at 31 Aug might therefore be doing well as it is acting as a proxy for the extra area of ocean over the last few months. It is also known that bottom melt dominates extent losses in September which supports this possibility. This would suggest that cumulative area over the last few months might perform better as a predictor. There are many ways to do the averaging and overfitting becomes a serious problem. Therefore I decided to only incorporate an improvement if my first attemt without any tuning was beneficial. This also helps with limited time available.

My first guess at the period for the averaging was 92 days (June, July Aug) with a weight factor using a sine wave peaking at 21 June and reaching 0 by 21 Sept to reflect the sun angle. My first guess at the relevant area was to use max(0,14-CT area). This produced an improvement so this style of predictor was accepted. On reflection, the area nearest the ice is more important and (14-CT) area was too high. I settled on using max(0,10-CTarea) as this was better than using 9 or 11 and I didn't want to do too much overtuning. The 92 day period was about right and left at that.

More recent day's energy captured may be more relevant to future extent loss as older energy captured may already have been used in melting ice. I tried using a linearly declining to 0 factor multiplied by the sine wave factor but rejected this despite hints that linearly declining to halve the weight may have been better.

I also tried and rejected:

1. Multiplying the weighed average area by (1+AO for period weighted towards August AO values) as a measure of cloud-less-ness.

2. Dividing by thickness calculated as PIOMAS volume divided by CT area.

3. Multiplying by (Extent - area) as a measure of the low concentration and isolated ice that is more likely to be susceptible to bottom melt that dominates losses in September.

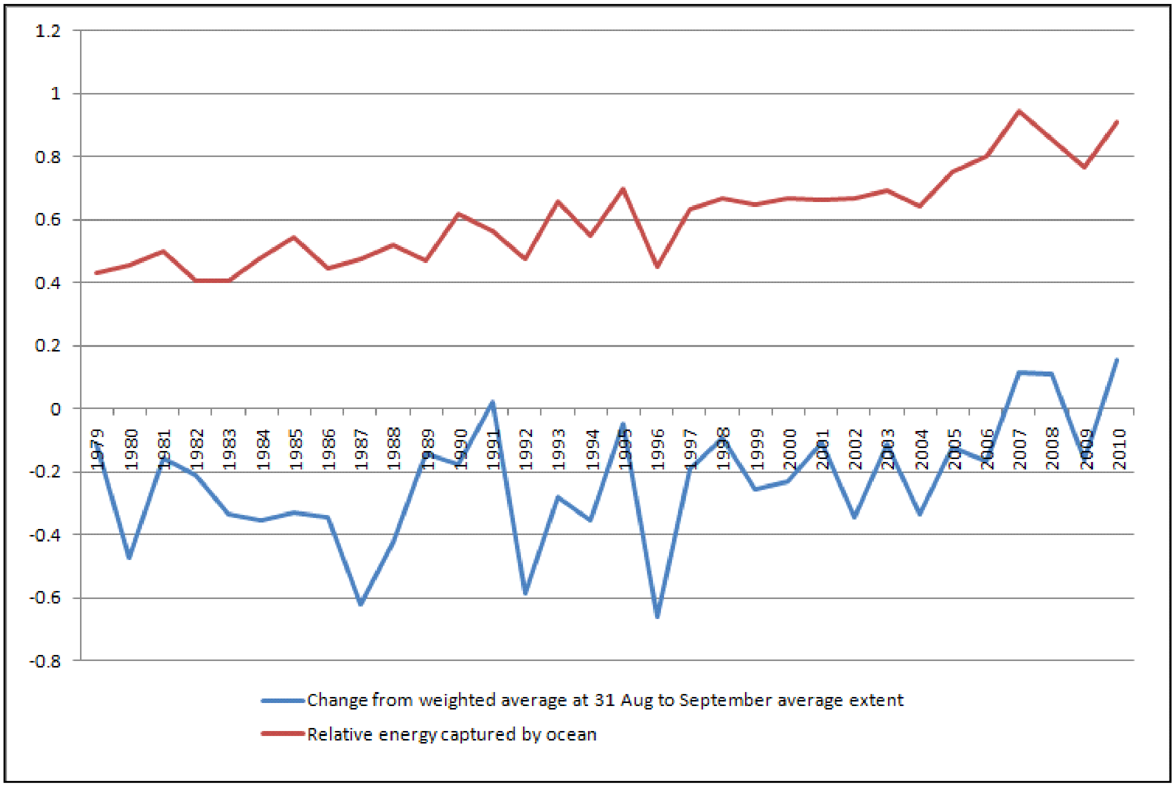
<u>Combining two methods predictions</u>

Averaging my two predictions weighted by skill was considered but the new method was significantly better and averaging the two methods predictions did not improve the prediction of the better new method.

<u>Adjusted gompertz fit method</u>

This is not used but for completeness the prediction using this old method was for 4.35 m km^2 +/- 2 standard deviation after removing hindsight information range of 0.4 m.

Predicted september change method graph



Steele et al 2009 in "Mechanisms of summertime upper Arctic Ocean warming and the effect on sea ice melt" looking at Pacific sector say

"*Our analysis shows that top melt dominates total melt early in the summer, while bottom melt (and in particular, bottom melt due to ocean heat transport) dominates later in the summer as atmospheric heating declines. Bottom melt rates in summer 2007 were 34% higher relative to the previous 7 year average. The modeled partition of top versus bottom melt closely matches observed melt rates obtained by a drifting buoy. Bottom melting contributes about 2/3 of total volume melt but is geographically confined to the Marginal Ice Zone, while top melting contributes a lesser 1/3 of volume melt but occurs over a much broader area of the ice pack.*"

The graph above seems to indicate that some of the September variability from year to year can be predicted from albedo feedback rather than being dominated by ocean heat transport. This isn't necessarily much of a contradiction, if any, of the above quote but might merit further investigation.

Neither of my methods makes any attempt at evaluating impact of likely weather over next six weeks which appears to be the dominant cause of variation in the change in extent between 31 August and the minimum.

4. Executive Summary

The data appears to have a curved shape which it appears advantageous to use whether that is explicitly or implicitly. My new method predicting change from 31 August weighted average of extent and area using culmulative energy that could be captured by ocean that was formerly ice.

The prediction of average September extent is 4.45 m km^2 with a 95% confidence interval after removing hindsight information of +/- 0.32 m km^2 .

5. Estimate of Forecast Skill

A 95% confidence interval of +/- 0.32 m is calculated after removing hindsight information. This estimate is higher than the inappropriately tuned RMSE figure of as low as 0.13m.

<u>Change from 31 Aug method</u>

The standard deviation in the change from 31 Aug GSFC-JAXA extent to NSIDC average September extent is .197

The standard deviation in the change from 31 Aug weighted average of GSFC-JAXA extent and CT area to NSIDC average September extent is .199

RMSE of weighted average change predicted by area is .156

RMSE of weighted average change predicted by energy capture by ocean areas that may have been ice in previous years is .134

Note however that these RMSE numbers are likely to underestimate the likely error as they have the advantage of the method being tuned with data that cannot be available at the time of making a true prediction.

Removing that advantage:

Year	Prediction	Actual	Error
1991	6.861	6.55	0.311
1992	7.250	7.55	-0.300
1993	6.395	6.5	-0.105
1994	7.099	7.18	-0.081
1995	6.292	6.13	0.162
1996	7.556	7.88	-0.324

1997	6.745	6.74	0.005
1998	6.636	6.56	0.076
1999	6.156	6.24	-0.084
2000	6.253	6.32	-0.067
2001	6.822	6.75	0.072
2002	5.778	5.96	-0.182
2003	6.201	6.15	0.051
2004	5.918	6.05	-0.132
2005	5.569	5.57	-0.001
2006	5.840	5.92	-0.080
2007	4.412	4.3	0.112
2008	4.827	4.68	0.147
2009	5.288	5.36	-0.072
2010	5.026	4.9	0.126
2011	4.448		

Average absolute error 0.124 m
RMSE without tuning to unavailable data 0.154 m
The average of the absolute errors for the first 10 years is 0.151 m whereas the average in the last 10 years is only 0.097 m.

A major benefit of this method compared to my previous one is that the RMSE of .134 only shows a small rise to .154 when hindsight information is removed. (For comparison, my gompertz fit rmse of 0.145 rose to .191.)

In the format

	A	B	C	D	E	F
1	m_n	m_{n-1}	...	m_2	m_1	b
2	se_n	se_{n-1}	...	se_2	se_1	se_b
3	r_2	se_y				
4	F	d_f				
5	ss_{reg}	ss_{resid}				

The regression factors and data are

Regression Data - Change from weighted average predicted by energy captured by ocean that might otherwise be ice.	
0.003079557	-0.79059
0.000605136	0.113406
0.463310596	0.148322
25.89825285	30
0.569747071	0.659983

6. Invitation to discuss

Comments on this method or the error estimate or comparing different methods or error estimates between different contributions are welcome. I suggest such discussion could be useful be done at Neven's blog. The latest appropriate post being

http://neven1.typepad.com/blog/2011/07/september-search-outlook-contribution.html

Data Sources and References

1. NSIDC average September Extent
ftp://sidads.colorado.edu/DATASETS/NOAA/G02135/Sep/N_09_area.txt

2. Cryosphere Today daily area data referred to as CT area.
http://arctic.atmos.uiuc.edu/cryosphere/timeseries.anom.1979-2008

3. JAXA daily extent data http://www.ijis.iarc.uaf.edu/seaice/extent/plot.csv

4. GSFC daily extent data
http://polynya.gsfc.nasa.gov/datasets/nh_daily_observed_or_interpolated_sie_1972_2002.txt

5 GSFC-JAXA extended JAXA daily extent data
https://docs.google.com/spreadsheet/ccc?key=0AjpGniYbi4andElTczhVS2t2VW5Ka0sySnFrcndTTkE&hl=en_US#gid=0

6. Steele et al. 2009 Mechanisms of summertime upper Arctic Ocean warming and the effect on sea ice melt
http://psc.apl.washington.edu/zhang/Pubs/Steele_etal_2009JC005849.pdf