Experiments

Abhinav Tushar

March 14, 2018

This document contains some notes on the results.

1 Some findings

- Multiplicative weights does better than simple mean when started with equal weights. If started with weights from degenerate em, it does better than dem (3/4) on test data.

Target	Upper limit	Mean	DEM	k-DEM	MP
1-ahead	-1.8314	-2.7822	-2.6189	-2.6189 (k=1)	**-2.6182**
2-ahead	-1.8883	-2.9742	-2.8250	**-2.8225** (k=7)	-2.8241
3-ahead	-2.0386	-3.1339	**-2.9931**	**-2.9931** (k=1)	-2.9948
4-ahead	-2.1614	-3.2481	-3.0996	-3.09961 (k=1)	**-3.0986**

- Doesn't look like there are well defined regions (in the time axis) where specific models outperform some other consistently. *Deduced from the results using k partition dem models.*

- The dimension where we possibly can quantify the role of different models might actually be the y axis (wili). There might be more well defined discretizations here than on the time axis since the differences between statistical and mechanical models come into play here. For example in the high (and anomalous) time of this season, at one week we didn't have delphi and reich lab models (which are mostly statistical), the partial ensemble predicted better because of the CU's mechanistic models.

- Since dynamic weight models implicitly take into account this y axis variability, they should be more robust for our purpose.

- This season's results

Target	Upper limit	Mean	DEM	k-DEM	MP
1-ahead	-2.2603	-3.3267	-3.2496	-3.2496 (k=1)	**-3.2461**
2-ahead	-2.6580	-3.8180	-3.8424	-4.0286 (k=7)	-3.8546 (**-3.8060** with reset)
3-ahead	-2.9698	-4.1643	-4.2478	-4.2478 (k=1)	-4.2648 (**-4.1628** with reset)
4-ahead	NA	NA	NA	NA	NA

When started with weights from DEM model, MP doesn't do good. If instead started with equal weights (*with reset*), models with less weights get a chance to express and give results in anomalous season (where mean is better than dem).

To add this in MP without inspecting which season is weird and which not, we can use a fixed shared scheme [1] which adds another parameter α and avoids assigning near zero weights to models. Or simply put an interpolating parameter between equal weights and dem weights.

References

[1] Mark Herbster and Manfred K Warmuth. Tracking the best expert. *Machine learning*, 32(2):151–178, 1998.