The Third International Workshop on Climate Informatics
National Center for Atmospheric Research, Boulder, Colorado, September 26-27, 2013
https://www2.image.ucar.edu/event/ci2013

MRF-Based Spatial Expert Tracking of the Multi-Model Ensemble

Scott McQuade and Claire Monteleoni
Department of Computer Science, The George Washington University
mcquade@gwmail.gwu.edu, cmontel@gwu.edu

1 Introduction

We consider the problem of adaptively combining the "multi-model ensemble" of General Circulation Models (GCMs) that inform the Intergovernmental Panel on Climate Change (IPCC), drawn from major laboratories around the world. This problem can be treated as an expert tracking problem in the online setting as in [5], where an algorithm maintains a set of weights over the experts (here the GCMs are the experts). At each time interval these weights are used to make a combined prediction, and then the weights can be updated based on the performance of experts. In this work we focus on tracking the GCMs at different geographic locations and effectively incorporating spatial influence and correlations between these locations. In [4], we proposed a simple method that allowed regions to influence their immediate neighbors. Here we extend [4], providing a more rigorous derivation of the spatial influence and allowing globally coherent influence by using a Markov Random Field (MRF).

2 Technical Approach

We approach this multi-model ensemble problem using a pairwise MRF, where the state of each hidden variable is the identity of the best GCM at a specific location. Similar MRF-based methods have been used recently to analyze climate data. In [2] an MRF-based approach was used to spatially and temporally detect drought states throughout the twentieth century. Our proposed method differs from [2] in that (1) we estimate the full marginal distributions of the hidden variables rather than their most likely state and (2) we apply our method in an online setting where at each time iteration we fit a new MRF and then perform inference to obtain the marginals.

We utilize a pairwise MRF, where every neighboring pair (x_i and x_j) of variables has an associated "energy" function $E(x_i, x_j)$, with lower energy indicating a more likely state. The joint distribution of all variables is $p(x_1, x_2, \ldots x_{\max}) = \prod_{(i,j)} e^{-E(x_i,x_j)}$, where (i, j) is the set of all neighboring variables [6].

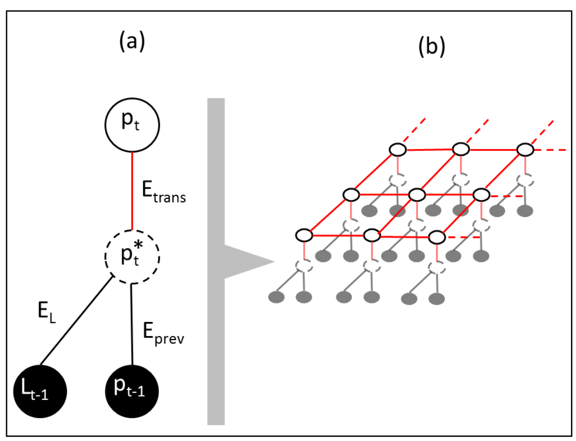

Figure 1: Construction of the MRF

To establish reasonable energy functions for the MRF, we first show that the Fixed-Share algorithm of [3] can be expressed as a simple MRF. Fixed-Share is one of several common "share update" algorithms for expert tracking that incorporate switching dynamics, to model situations where the best expert may switch over time. By expressing Fixed-Share as an MRF, we identify the energy function that corresponds to the switching dynamics. Since an MRF is an undirected graph, this "switching" energy function can be naturally applied to spatial links between variables as well.

Figure 1(a) shows the MRF corresponding to Fixed-Share. The black nodes are observed variables or evidence: p_{t-1} was our "belief" (represented as a fixed probability distribution over the GCMs) from the previous time iteration of which GCM would be the best predictor, and L_t represents the losses we observed for each GCM (typically the squared difference between the prediction and the observed value) at the previous time iteration. The white nodes are the hidden variables: p_t represents the identity of the GCM that is currently the best predictor, and p_t^* is an intermediate node, indicated by the dotted border. We define the energy functions as:

$$E_{\text{prev}}(p_t^* = i, \boldsymbol{p_{t-1}}) = -\log \boldsymbol{p_{t-1}}(i) \tag{1}$$

$$E_L(p_t^* = i, \boldsymbol{L_{t-1}}) = \boldsymbol{L_{t-1}}(i) \tag{2}$$

$$E_{\text{time}}(p_t = i, p_t^* = j) = \begin{cases} -\log(1 - \alpha_{\text{time}}) & \text{if } i = j \\ -\log(\frac{\alpha_{\text{time}}}{n-1}) & \text{if } i \neq j \end{cases} \tag{3}$$

Where n is the number of experts, and α_{time} is a parameter of Fixed-Share that captures how frequently the best expert switches. A simple calculation of the marginal probability of p_t confirms that this MRF is indeed equivalent to the Fixed-Share update rule.

Figure 1(b) illustrates how we construct a spatial lattice MRF by using a Fixed-Share MRF at each location. The p_t node for each region is linked to each of its adjacent spatial neighbors, with the energy of each link (E_{space}) taking the same form as the energy of the temporal switching dynamics (E_{time}) but with a different switching rate parameter α_{space}.

At each time iteration we consider a new MRF, with the inferred marginal distributions from the previous round (now held fixed) and the GCM losses serving as the observed variables. To calculate the marginal probabilities of the hidden variables (i.e. our new beliefs over GCMs), we apply Loopy Belief Propagation (LBP) to the MRF. In LBP, each node sends messages to neighboring nodes about the sender's "belief" of the neighbor's state. On tree graphs, the belief propagation is guaranteed to quickly determine the correct marginals. However on graphs with loops, such as our lattice MRF, there is no universal guarantee of convergence or the accuracy of the resulting marginals. Nevertheless, LBP has been shown to empirically perform well for a number a different applications with loopy graphs [6]. In our experiments, LBP also converged quickly, with the messages converging in less than 10 message passing iterations for all reasonable α_{space} values.

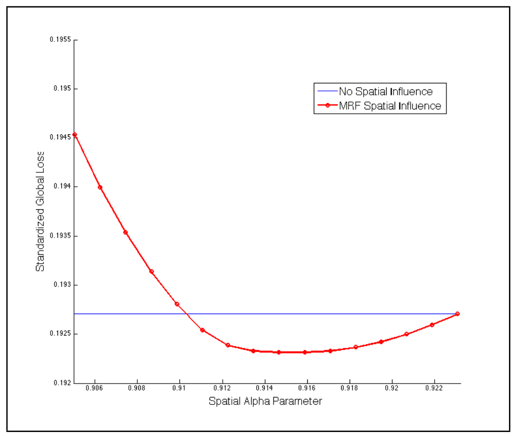

Figure 2: The performance (mean annual squared loss for global predictions) of our method versus α_{space} with $\alpha_{\text{time}} = 0.05$. The red line represents our proposed approach. The blue line is equivalent to TCM [5] and NTCM with no neighborhood influence [4].

Figure 2 show our initial results from an online evaluation of historical GCM temperature predictions from the IPCC Phase 3 Coupled Model Intercomparison Project (CMIP3) archive [1]. The right-most point on the graph corresponds to $\alpha_{\text{space}} = \frac{n-1}{n}$ (n is the GCM ensemble size) for an algorithm variant with no spatial influence. Spatial influence increases with subsequent smaller α_{space} values. When $\alpha_{\text{space}} = 0$ the spatial influence is at a maximum, and all the hidden variables must have the same state (this case is equivalent to tracking a single set of experts over all locations without modeling any spatial variation). Figure 2 indicates that the optimal value of α_{space} is between these two extremes, as the performance initially improves with increasing spatial influence (decreasing α_{space}), but eventually diminishes with α_{space} values that are too small.

In conclusion, we proposed a rich method for incorporating spatial influence into the multi-model ensemble problem. We showed that this method performs well when an appropriate value for α_{space} is used. Future contributions could include methods to learn the α_{space} parameter from the data or methods to use climate science domain knowledge help set a value for α_{space}.

References

[1] CMIP3. The World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset. `http://www-pcmdi.llnl.gov/ipcc/about_ipcc.php`, 2007.

[2] Q. Fu, A. Banerjee, S. Liess, and P. K. Snyder. Drought detection of the last century: An mrf-based approach. In *SDM*, pages 24–34. SIAM / Omnipress, 2012.

[3] M. Herbster and M. K. Warmuth. Tracking the best expert. *Machine Learning*, 32:151–178, 1998.

[4] S. McQuade and C. Monteleoni. Global climate model tracking using geospatial neighborhoods. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, July 22-26, 2012, Toronto, Ontario, Canada*, pages 335–341, 2012.

[5] C. Monteleoni, G. Schmidt, S. Saroha, and E. Asplund. Tracking climate models. *Statistical Analysis and Data Mining: Special Issue on Best of CIDU*, 4(4):72–392, 2011.

[6] J. S. Yedidia, W. T. Freeman, and Y. Weiss. Understanding belief propagation and its generalizations. *Exploring artificial intelligence in the new millennium*, 8:236–239, 2003.