

Spatiotemporal Data Mining to Anticipate Tornado Formation

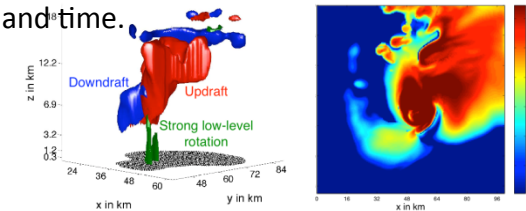
Faculty: Amy McGovern

Objectives

- Develop novel spatiotemporal relational data mining models to enable meteorologists to revolutionize their understanding of tornado development
- Apply the data mining models to multiple severe weather situations
- Bring the research into the classroom to provide authentic classroom examples of how CS can change the world. Improve retention and recruitment of underrepresented groups

Background

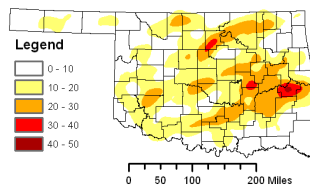
- Tornadoes are correctly warned 80% of the time BUT 80% of the warnings are false alarms. We want to improve prediction!
- Radars are limited in what they can observe. Prediction near limit with this system.
- We generate very high resolution simulations of storms and study these to develop a detailed model of how a storm evolves over space and time.



Preliminary Results

- Developed Spatiotemporal Relational Probability Trees
- Developed Spatiotemporal Relational Random Forests
- Applied SRPT & SRRF to multiple severe weather domains:
 - Convectively induced turbulence
 - Tornado development and fronts
 - Spread of drought

Tornadic Supercell Frequency 1994-2003



Relevant References

- McGovern, Amy; Supinie, Timothy; Gagne II, David John; Troutman, Nathaniel; Collier, Matthew; Brown, Rodger A.; Basara, Jeffrey; Williams, John. (under review) Augmenting Spatiotemporal Relational Random Forests for Use in Real-world Severe Weather Applications. Submitted to the Knowledge Discovery and Data Mining conference: KDD 2010.
- McGovern, Amy and Hiers, Nathan and Collier, Matthew and Gagne II, David J. and Brown, Rodger A. (2008). Spatiotemporal Relational Probability Trees. Proceedings of the 2008 IEEE International Conference on Data Mining, Pages 935-940. Pisa, Italy. 15-19 December 2008.
- Collier, Matthew and McGovern, Amy. (2008). Kernels for the Investigation of Localized Spatiotemporal Transitions of Drought with Support Vector Machines. Proceedings of ICDM 2008, the 8th IEEE International Conference on Data Mining Workshops. Pisa, Italy. 15-19 December 2008, pages 359-368.