Constructing an Efficient Machine Learning Model for Tornado Prediction

In Aleskerov et al. (2018) the approach employing the threshold procedure and the superposition principle for tornado prediction implying detection of tornadic circulations from the set of all (tornadic and nontornadic) observations is proposed. The method utilized exploits sequentially (i.e., the result (output) of one function is the input of another choice function) several choice functions that allows to choose some subset of alternatives, that satisfy predefined conditions.

On the first step of data preparation correlational analysis was applied to reduce the number of predictors involved. Afterwards, only parameters with different distributions for the tornadic and nontornadic events were left. Eventually, features that establish the prediction when the disaster has already happened were abandoned preserving characteristics, which signalize for tornado occurrence beforehand.

Considering a finite set A of alternatives evaluated by retained n parameters, the choice function of the constructed model presentes the threshold procedure examined a weighted sum of two certain parameters to evaluate which alternatives may be tordadic events. The decreased set of observations is fed to the next choice function.

The main idea of the presented model is the building of a certain number of superposition sequences, the union of whose outputs will be the final prediction.

The results for the 50:50 training/testing ratio were the POD (Probability of detection) equal to 0.68, the FAR (False alarm ratio) equal to 0.16 and the CSI (Critical success index), which is division of tp on the summarization of tp, fp and fn, where tp is true positives or number of correctly identified objects, fp is false positives or number of incorrectly identified objects, and fn is false negatives or number of objects missed by the model, equal to 0.61. Herewith, in case of standard 70:30 data division the metrics' values are even higher.

To conclude, the accuracy of the superposition decision tree model surpasses the results of previous models. Moreover, the proposed method can be applicable in case of other natural disasters.

References:

1. Aleskerov, F., Baiborodov, N., Demin, S., Shvydun, S., Trafalis, T., Richman, M., & Yakuba, V. (2018). Constructing an efficient machine learning model for tornado prediction. *arXiv preprint arXiv:1806.05857*.