

# Intelligent Outlier Detection Algorithm

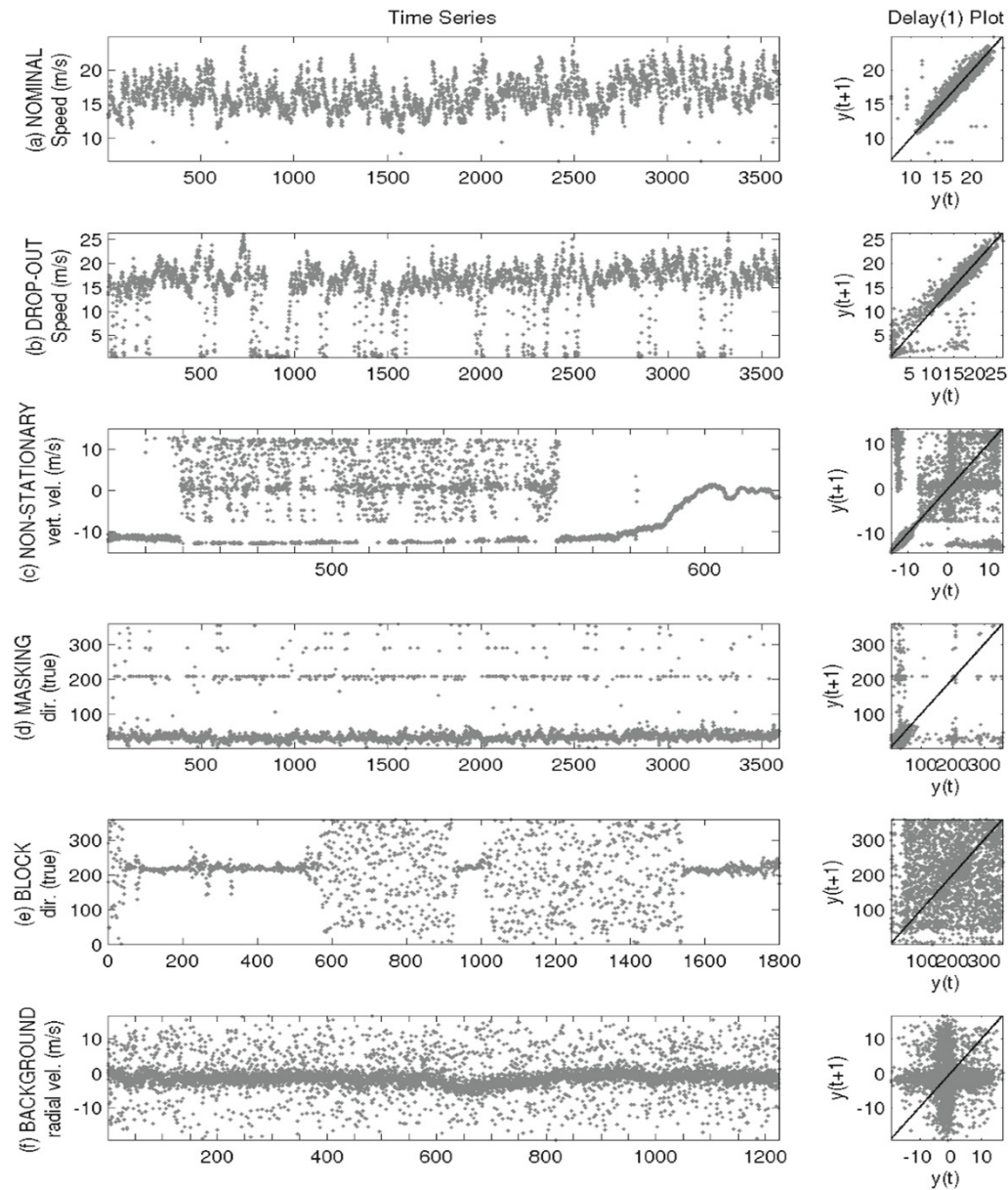


Andrew Weekley  
National Renewable Energy Lab  
[Andrew.weekley@nrel.gov](mailto:Andrew.weekley@nrel.gov)

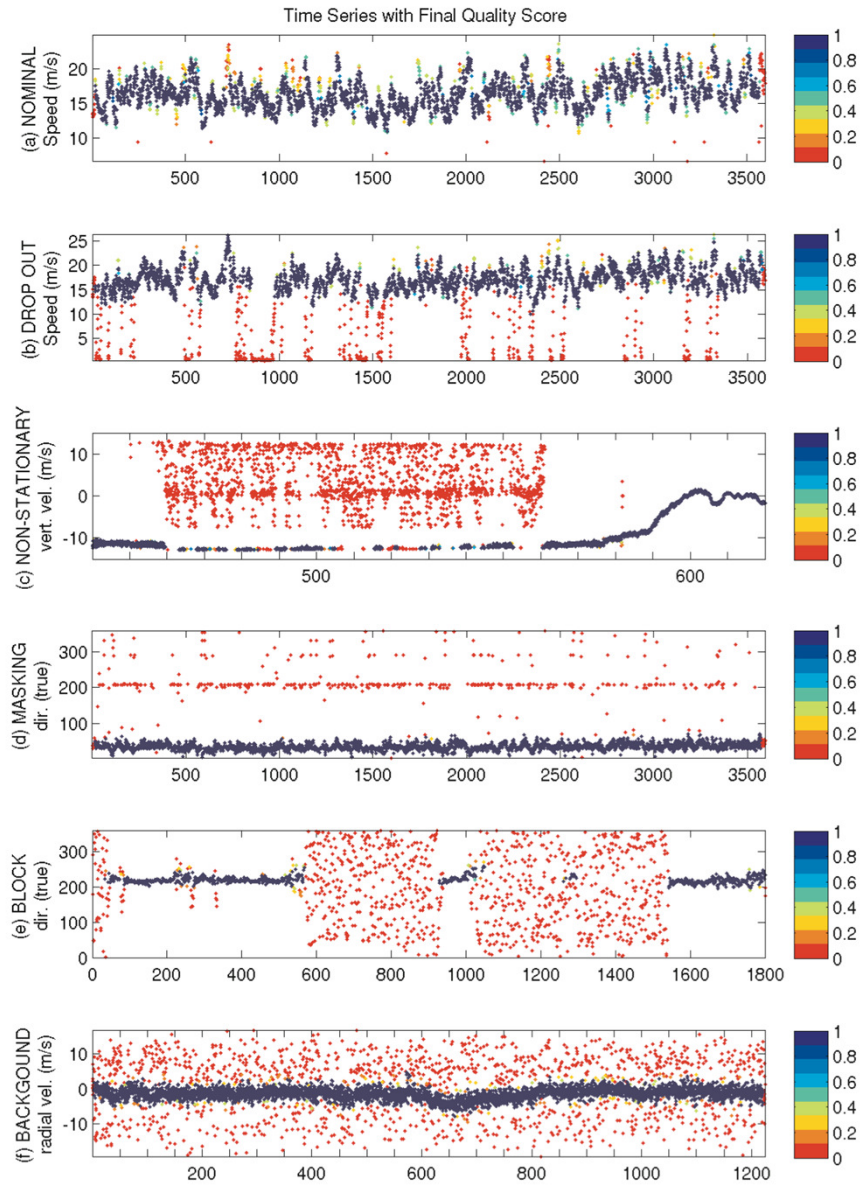
# Acknowledgments

- Co-investigators
  - Kent Goodrich (CU Boulder Dep. of Math, NCAR)
  - Larry Cornman (NCAR)
- Work originally funded by the National Center for Atmospheric Research
- Work performed while employed by NCAR

# Cases studied



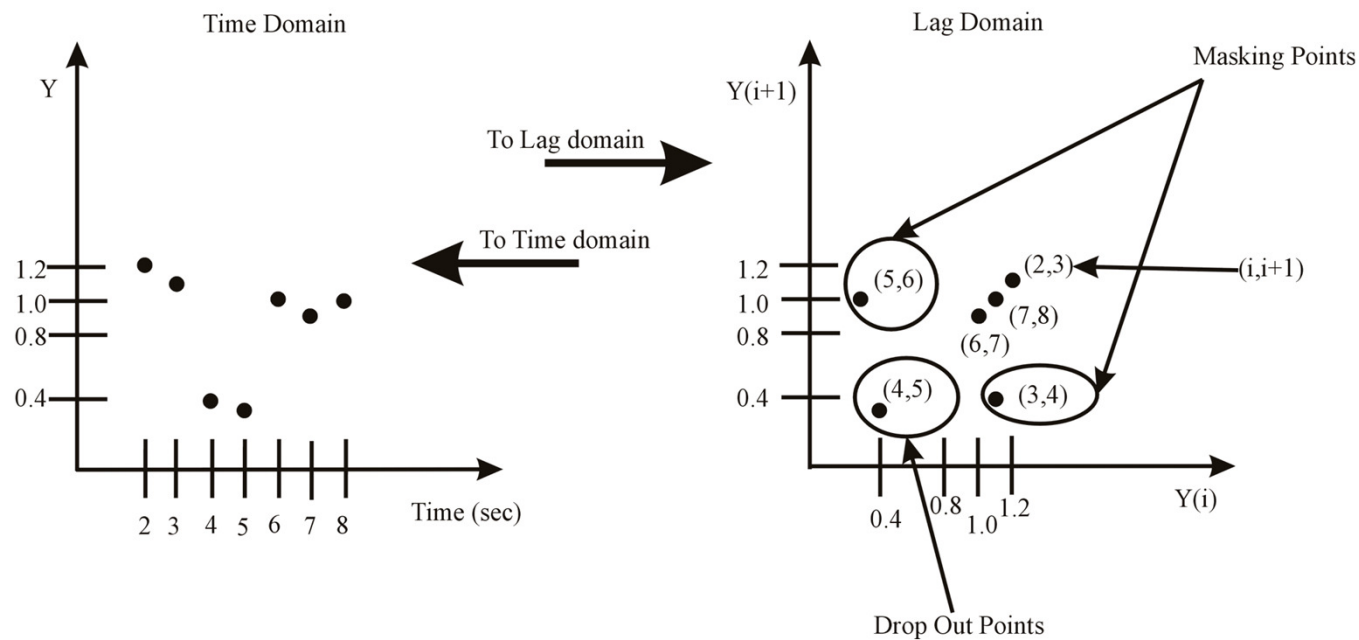
# Final Confidence



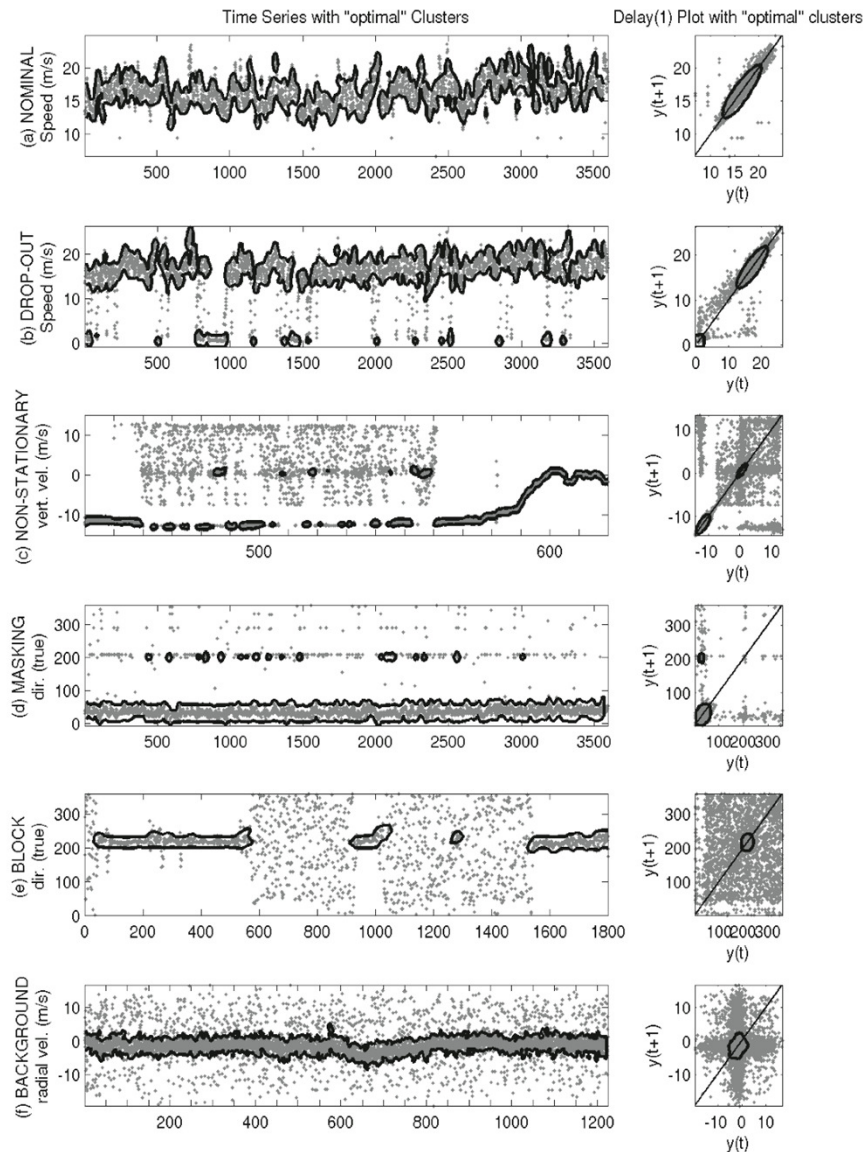
# Finding Suspect Data

- Humans can identify suspect data in a time series with little or no information
- Finding suspect data with a computer may involve calculating a statistic
- The statistic may be used to identify the suspect data
- But the suspect data corrupts the statistic
- And depending on the situation the statistic may not identify the suspect data

# Scatter Plot of $Y(t_2)$ vs. $Y(t_1)$



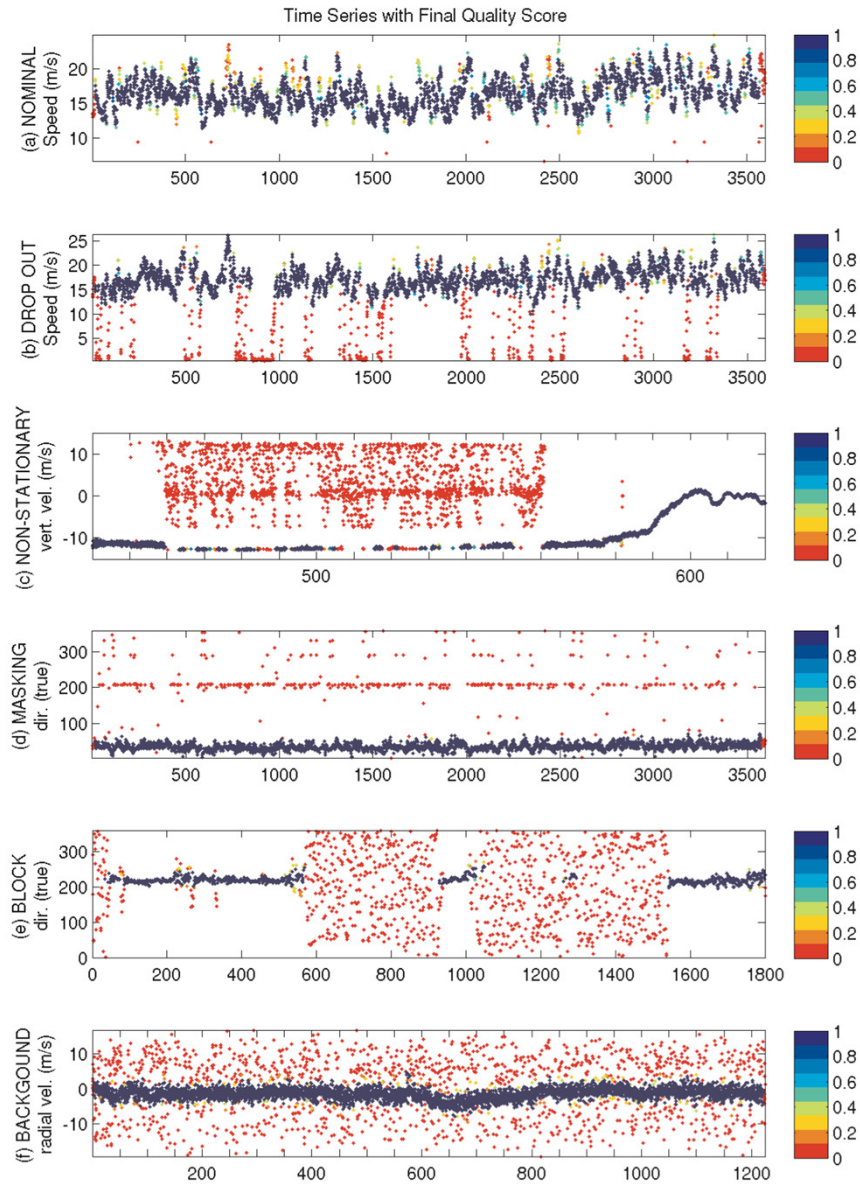
# Optimal clusters



- optimal clusters in time domain and delay space
- build a “feature” in the time domain from optimal clusters in the time domain
- drop-out and non-stationary cases have two optimal clusters in delay space, but distinct representations in the time domain.
- nominal, block and uniform cases have single clusters in delay space but distinct representations in the time domain



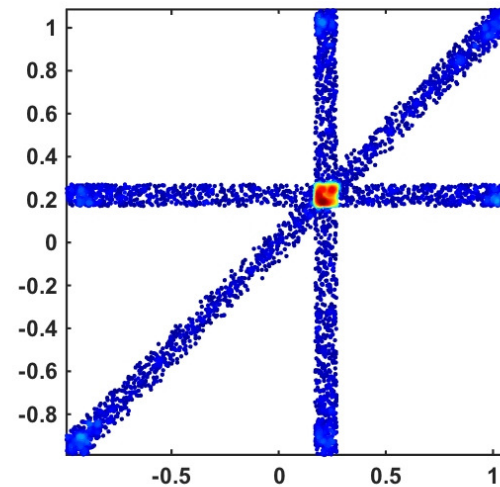
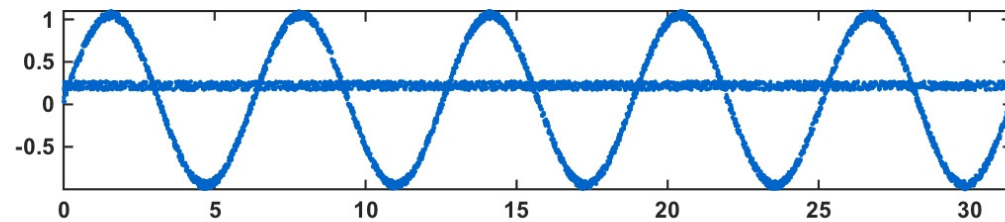
# Final Confidence





# Simulated example

- Identified a set of feature vectors which could be used in a machine learning algorithm such as SVM
- Build a training data set using simulated data



# References

- Ma, J., and S. Perkins, 2003: Time-series novelty detection using one-class support vector machines. *Proceedings of the International Joint Conference on Neural Networks, 2003.*, Vol. 3 of, IEEE, 1741–1745
- Weekley, R. A., R. K. Goodrich, and L. B. Cornman, 2003: Fuzzy Image Processing Applied to Time Series Analysis. *3rd Conference on Artificial Intelligence Applications to the Environmental Science*, Vol. 54 of, Citeseer, 258
- Weekley, R. A., R. K. Goodrich, and L. B. Cornman, 2004a: Feature Classification for time series data. *US Patent 6,735,550*.
- Weekley, R. A., R. K. Goodrich, and L. B. Cornman, 2010: An Algorithm for Classification and Outlier Detection of Time-Series Data. *Journal of Atmospheric and Oceanic Technology*, **27**, 94–107, doi:10.1175/2009JTECHA1299.1.