3.6

YOU'VE CHANGED! INHOMOGENEITY DETECTION FOR COOP NETWORK PRECIPITATION MEASUREMENTS

Tressa L. Fowler[1], Edward I Tollerud[2], and Barbara G. Brown[1]
[1]National Center for Atmospheric Research, Boulder, Colorado
[2]NOAA Research - Forecast Systems Laboratory, Boulder, Colorado

1. INTRODUCTION

Detecting actual changes in the amount or frequency of precipitation received at a United States cooperative observer network (COOP) station requires eliminating apparent "changes" that are the result of instrument drift, alterations in method of measurement and/or reporting, modification of the station's surroundings, etc. Change point detection is very challenging even when the measurements possess nice statistical properties such as normality, continuity, and homogeneity of variance. However, precipitation data do not possess nice statistical properties. In fact, the occurrences of precipitation can be relatively infrequent and when precipitation does occur, the measurements tend to be skewed. For these sorts of measurements, use of standard change point methods is not recommended.

Fortunately, the COOP network is fairly dense. Each station has several neighbors also measuring precipitation. These neighbors are taken advantage of to develop an alternative approach for detecting inhomogeneities. In particular, the relationship between frequency and amount of precipitation at each station is compared to its neighbors for each month over the entire period of record. Thus the empirical distribution and time series of various measures of association between stations are obtained. New measurements can be compared to these to determine if the recent measures are "typical" or not. If not, the station can be flagged as possibly having experienced a change and further checks can be performed.

Preliminary results of these comparisons for Iowa are available in Tollerud et al. (2002). In this paper, the comparisons are performed on simulated data to determine the efficacy of the methods and the period of time required to detect changes of various types and magnitudes.

2. DATA

The data to be tested for inhomogeneity come from the TD3200 dataset observed by the COOP network. These data are available on CD-ROM from the National Climatic Data Center (NCDC). These rain gauge precipitation measurements are observed and recorded daily, then submitted, entered, and quality checked monthly.

Preliminary analyses of these data have been completed, and are detailed in Tollerud et al. (2002), hereafter T02. For the analysis in T02, the data from COOP stations in Iowa were selected because of the frequency of precipitation, dense coverage by the COOP network, and uniformity of terrain in that state.

On a daily time scale, precipitation measurements are highly variable. Thus, in T02 statistics based on daily measurements were aggregated to give seasonal statistics that are less variable.

In order to study the efficacy of the methods, we apply them here to simulated data. Use of simulated data facilitates variation of a single characteristic of the measurements to determine how each characteristic changes the statistics.

Precipitation observations were simulated in the following way:

For probabilities $p \in \{0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.01\}$, simulate a vector of Bernoulli trials (success or failure). These proportions represent the probability of precipitation over a season at a station. Thus, $p=0.5$ would represent a location with precipitation occurring frequently, typically on about half of the 90 days of each season. For $p=0.01$, precipitation is extremely rare. By using this variety of probabilities, we can determine how our ability to detect changes in precipitation depends on the frequency of precipitation.

For each success (e.g. day with precipitation), simulate a rainfall amount, T, as the sum of two gamma distributed random variables, x_1 and x_2, for the target station (e.g. $T = x_1 + x_2$).

For each of 10 neighboring stations, simulate a rainfall amount, N_i, as the sum of one of the gamma distributed random variables used to simulate the target station rainfall, x_1, and a different gamma distributed random variable, (e.g. $N_1 = x_1 + x_3$; $N_2 = x_1 + x_4$, etc.) (Brown, 1978).

In order to simulate some stations with no rain when neighbors are reporting rain, randomly select some small percentage of rainfall observations, in this case 5%, to be reset to zero. Thus, target events are not always concurrent with neighboring events, but there is good agreement.

Desired correlation of the target with the neighbors is achieved by setting the parameters of the gamma distributions appropriately. For the simulated data, the measurements at the neighboring stations were computed such that the correlation between the precipitation amount at the target station and each neighboring station is 0.7. This is a fairly good correlation, and was chosen to approximate the actual correlation between pairs of stations in Iowa.

Corresponding author address: Tressa L. Fowler, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000. e-mail: tressa@ucar.edu

For this paper, a single set of parameters was used for the gamma distributions and percent of observations failing to indicate rain when the neighboring stations indicate rain. However, a variety of parameters will be used in future work to determine if the measures work well on different types of rainfall measurements.

The simulated data almost certainly fail to represent some characteristics of real precipitation data. ("All models are wrong, some models are useful"). However, the simulated data can be thought of as an idealized representation of the real data. If the methods fail to work on the idealized, (i.e. less complex) simulated data, then they are unlikely to work on the more complex real data.

3. METHODS

The precipitation measures taken at each neighbor station are used as "truth", and then various scores are computed based on the resulting pair. The scores considered in T02 are the equitable threat score (ETS), the frequency bias (BIF) and the magnitude bias (BIM). The magnitude bias is ignored here, and will be analyzed in future work. The ETS and frequency bias are computed by comparing the binary measurements of precipitation vs. no precipitation at the two stations. Since these measures are only concerned with precipitation "events" rather than amounts, the data used to compute them are a series of Bernoulli trials, or a binomial. For these measures, the gamma distributed rainfall amounts (and therefore the associated parameters and correlation) are irrelevant. The agreement between events at the neighbors is relevant. For the simulations used in this paper, this agreement is very high.

		Observed		
Forecast		**Yes**	**No**	**Total**
	Yes	a	b	a+b
	No	c	d	c+d
	Total	a+c	b+d	n

$$ETS = \frac{ad - bc}{(a+c)(b+d)} \qquad (1)$$

$$BIF = \frac{a+b}{a+c} \qquad (2)$$

ETS and BIF near 1 indicate good agreement between the target station and its neighbors.

4. RESULTS

The results are presented here in time series graphs. Time series for both ETS and BIF scores are presented for the standard simulated data and several types of "bad" data. For each of the sets of "bad" data, some type of problem was constructed for the last of 33 seasons of precipitation observations. The effect of the "bad" data on the scores, if any, is made clear by the time series plots.

4.1 Standard Simulated Data

As shown in Figures 1 and 2, for typical station behavior, the scores fluctuate randomly.When rainfall events are more common, the scores for each 90 day season are less variable. When rainfall is rare, the scores sometimes exhibit erratic behavior. For the data with a probability of precipitation of only 1%, much of the time there is too little data to compute the scores. When there is enough data so the scores can be computed, the results can be extreme, as shown by the downward spike in both BIF and ETS during the 26th season in the graphs. For a probability of precipitation of 5%, the scores can usually be computed, but there are many spikes in the time series for this data as well. This high variability may make it more difficult to detect real differences when precipitation is relatively infrequent.

4.2 Shifted Data

For this analysis, two types of shifted data were simulated. For the first set of data, the first observation of the last season of data was deleted. By removing this single observation, the remaining observations for that season are "shifted" to the day before the corresponding observations at the neighboring stations. For the second set, the shifting occurred half way through the last season (e.g. on the 45th day). Shifting of this nature is a common occurrence among the COOP network observers (Peterson et al., 1998).

The graphs showing the ETS and BIF computed on the entire season of shifted data are presented in Figures 3 and 4. The BIF graph looks the same as the one for the standard data, as it should. BIF measures the frequency of precipitation over a season compared to the frequency at the neighboring stations. Shifting the data can only change the seasonal frequency by a single day. Thus, BIF does not give an indication that the data have been shifted.

ETS, however, drops dramatically due to the shift, as it should. When the precipitation at the target station is a day early, then the target station measurement is not a good "predictor" of precipitation at the neighboring stations. This causes the score to be very low.

For the second set of data, only the graph of ETS is shown (Figure 5), as the BIF should not change due to shifting. The ETS still drops off for the last season, though obviously not as dramatically as when the whole season was shifted. Fortunately, it appears that the change in ETS is large enough to be detected as "unusual" for most of the values of p.

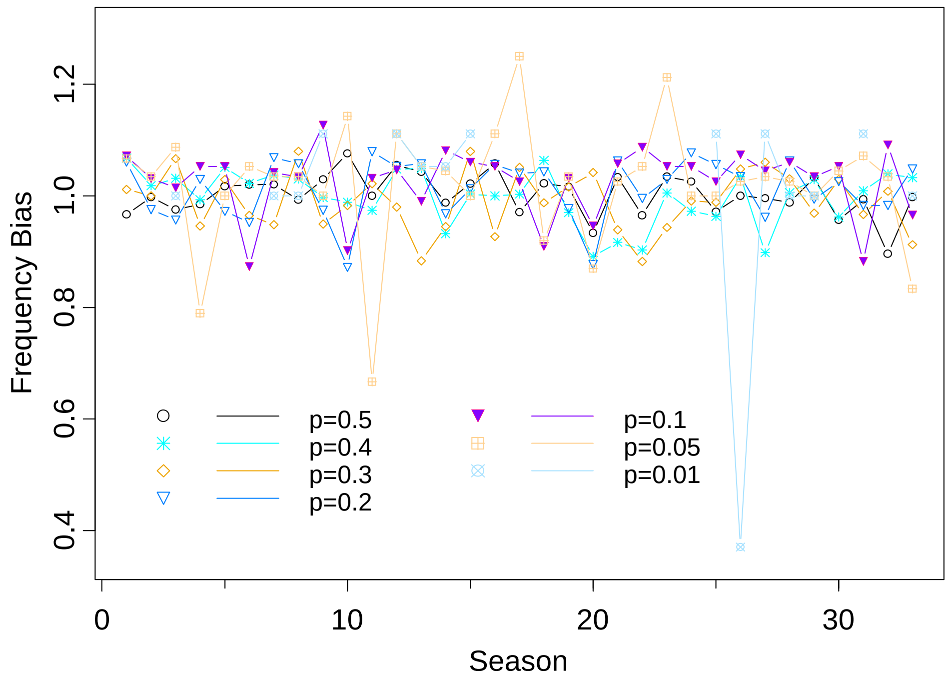

Figure 1: Time series of BIF for standard simulated data.

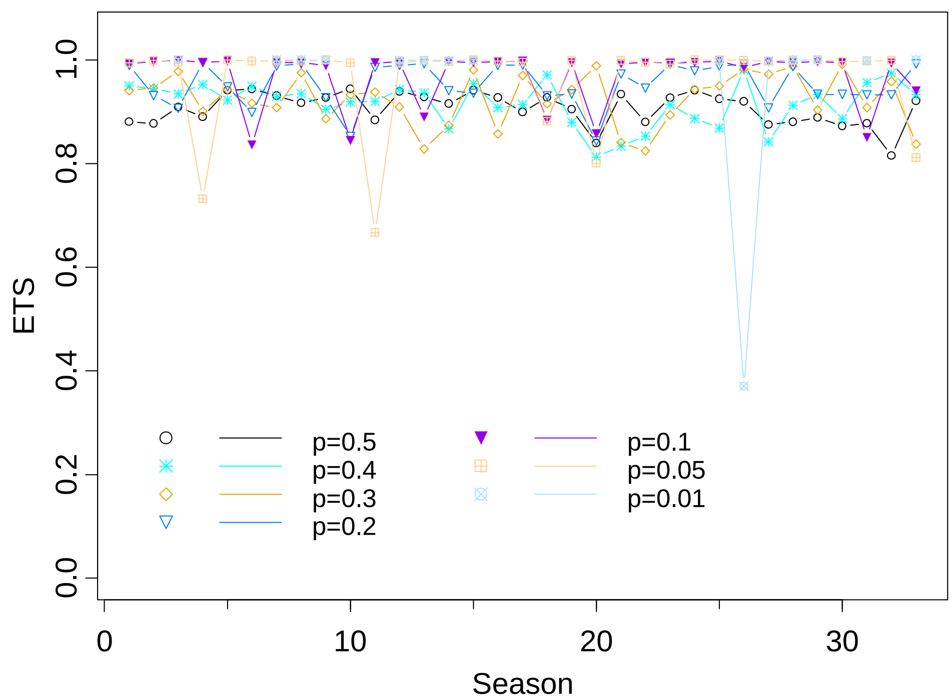

Figure 2: Time series of ETS for standard simulated data.

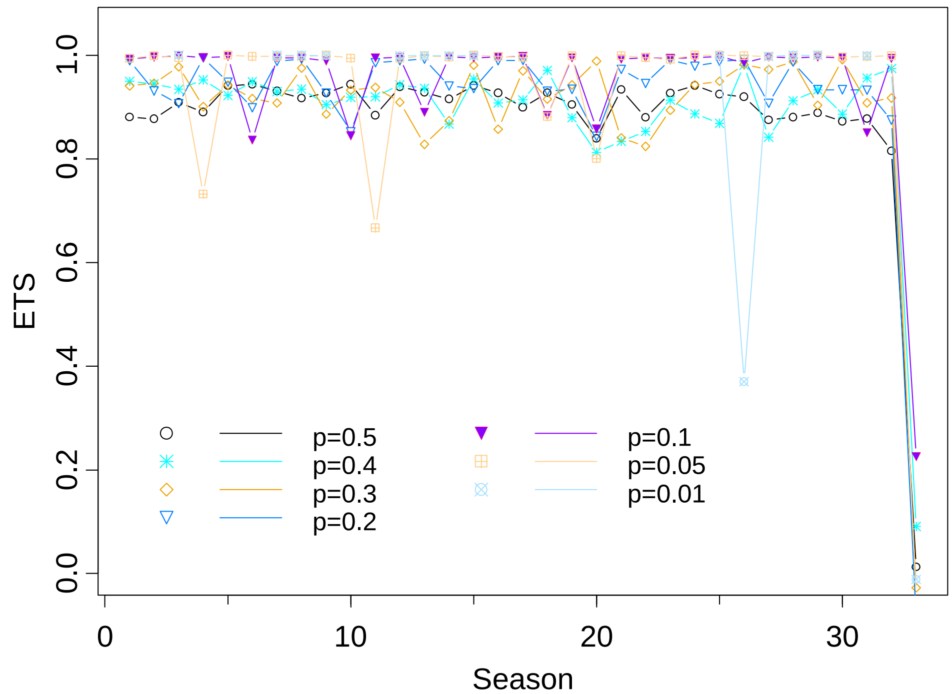

Figure 3: Time series of ETS for data with an entire season shifted.

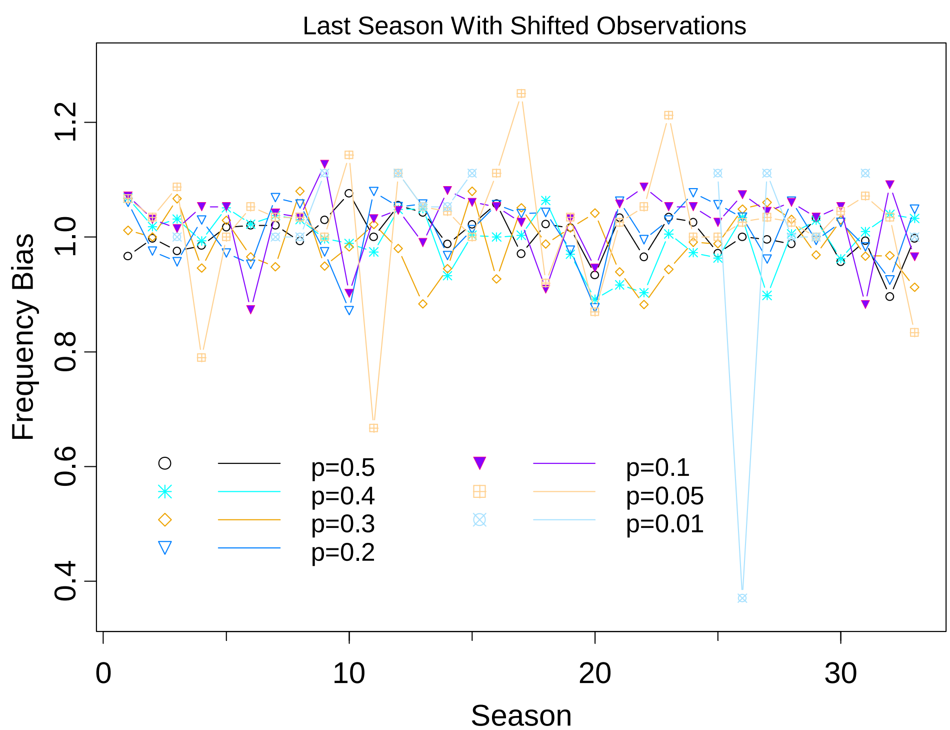

Figure 4: Time series of BIF for data with an entire season shifted.

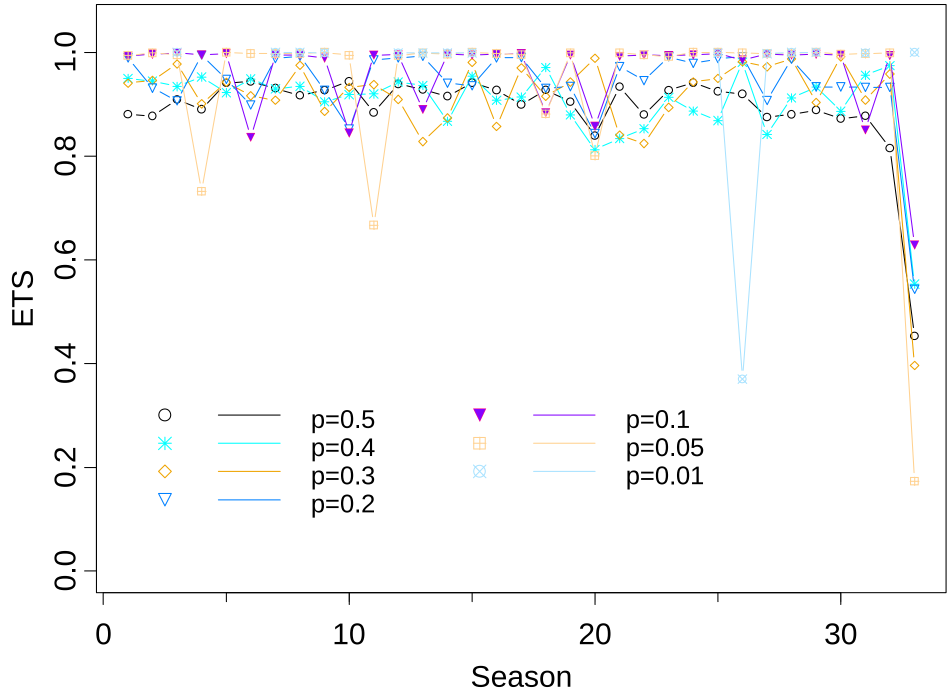

Figure 5: Time series of ETS for data with half of the last season shifted.

4.3 Missing Data

For this analysis, 20 and 40 observations, respectively, were randomly selected from the last season's 90 observations and replaced with observations of no precipitation. In some cases, the observation selected may have already indicated no precipitation so the replacement resulted in no change. The resulting scores for the resulting data sets are shown in Figures 6 through 9. For the data with 40 false no precipitation observations, the BIF and ETS drop off dramatically. Not surprisingly, the drop is less pronounced for the data with only 20 false observations. For *p=0.1* and *p=0.2*, there is no drop in the scores. Instead, they look fairly typical compared to scores for other years. The scores for some of the probabilities of precipitation do not appear to be unusual. Thus, the change would likely go undetected by the scores in these cases. When the probability of precipitation is low, the scores for the "bad" season may not be detectable as different from normal due to the high variability.

5. CONCLUSIONS AND FUTURE WORK

The ETS and BIF both show marked decreases when both 20 and 40 observations were replaced with observations of no precipitation. The ETS also decreases noticeably when date shifting occurs for both a whole season and a half season. The BIF does not change much due to date shifting, as the frequency of rainfall over the season is hardly changed. The drastic changes are more difficult to detect relative to scores from previous years for probabilities of precipitation of 1% and 5% due to the high variability of the statistics.

These analyses confirm that changes of various types are indicated by changes in BIF and ETS when the information from the neighboring stations agrees well with the information at the target station. The changes are easiest to detect with longer periods of record following the change and when precipitation is frequent. However, even changes followed by only a partial season of measurements may be indicated by a change in the scores.

These simulations assumed very good event agreement between the target station and its neighbors. Further research will include investigation of the scores computed on simulated data with less agreement between neighbors. It will likely be the case that less correlation among neighbors will result in these scores being less sensitive to inhomogeneities.

The ETS and BIF are not designed to detect changes in the amount of precipitation when the frequency is unchanged. Another measure, the magnitude bias, can be used to detect these types of changes. This measure is used in T02 on the Iowa data. Once more testing has been completed on the simulated data, the properties of this score will also be investigated.

All of the analyses in this paper and T02 focus on analyzing the seasons separately. Attempts are being made to homogenize the measurements from the different seasons, so that precipitation totals from all seasons may be considered together rather than separately, thus yielding a larger sample size in the same amount of time.

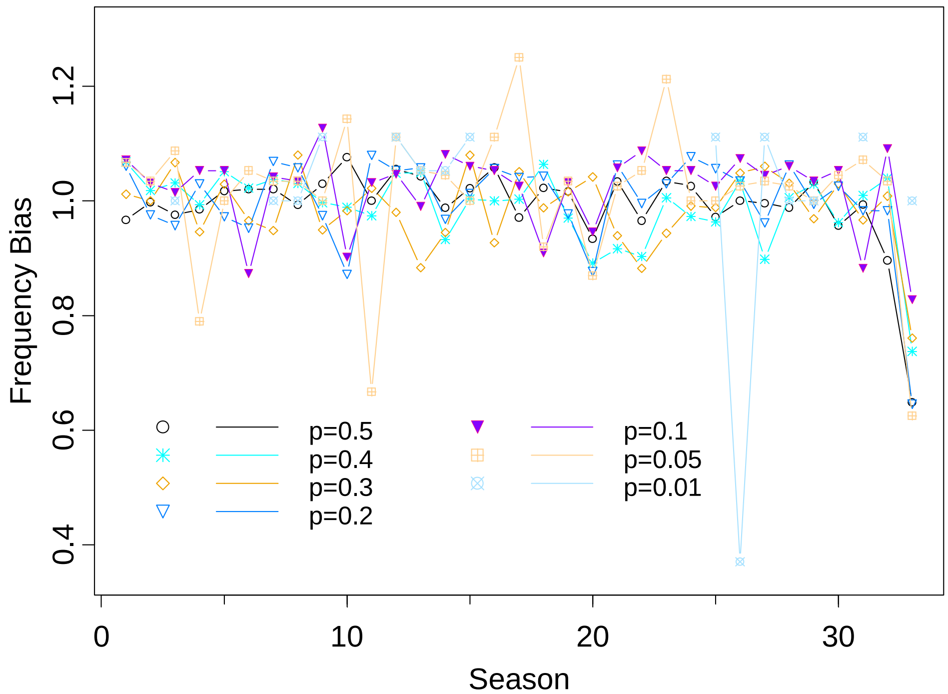

Figure 6: Time series of BIF for Data with 40 False Missing Observations.

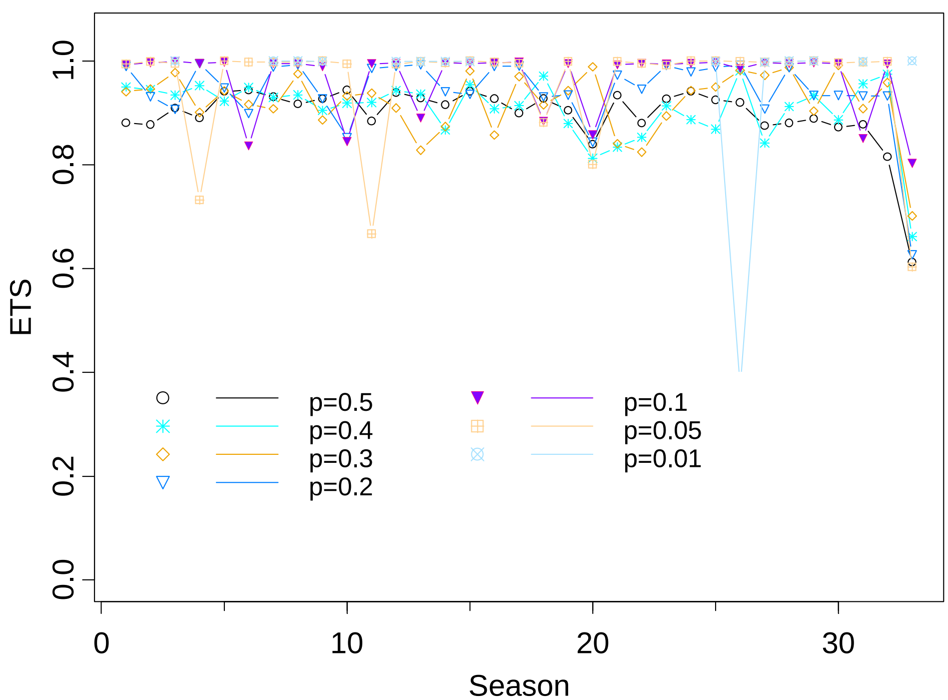

Figure 7: Time series of ETS for Data with 40 False Missing Observations.

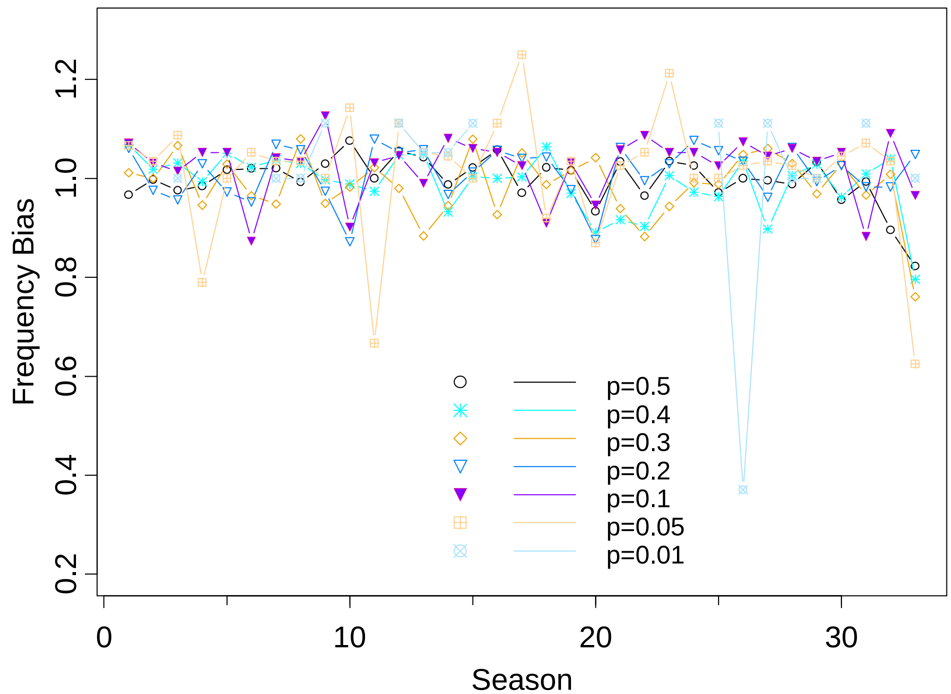

Figure 8: Time series of BIF for Data with 20 False Missing Observations.

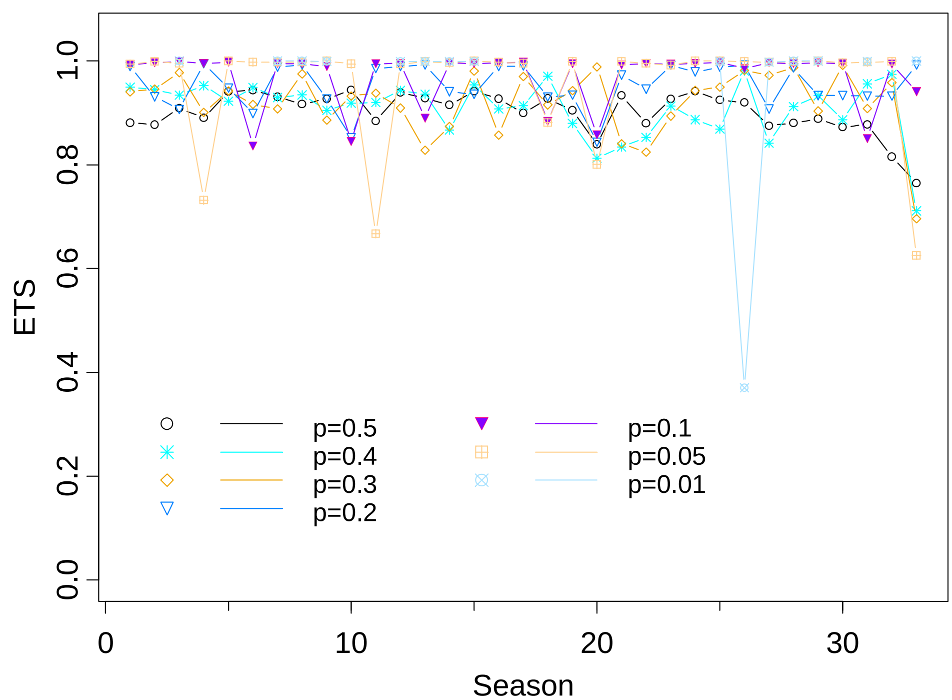

Figure 9: Time series of ETS for Data with 20 False Missing Observations.

ACKNOWLEDGEMENTS

This research was completed with funding from the NCDC Health of the Network Project. NCAR is sponsored by the National Science Foundation.

REFERENCES

Brown, B. G., 1978: Experimental Design in Weather Modification: The Value of Stratification. Charlottesville, University of Virginia, Department of Environmental Sciences, M.S. Thesis, 158 pp.

Peterson, T.C., D.R. Easterling, T.R. Karl, P. Groisman, N. Nicholls, N. Plummer, S. Torok, I. Auer, R. Boehm, D. Gullett, L. Vincent, R. Heino, H. Tuomenvirta, O. Mestre, T. Szemtimrey, J. Salinger, E. Forland, I. Hanssen-Bauer, H. Alexandersson, P. Jones and D. Parker, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. International Journal of Climatology, 18, 1493-1517.

Tollerud, E. I., B. G. Brown, and T. L. Fowler, 2002: Identifying Inhomogeneities in precipitation time series: 1. Diagnostic measures of spatial correlation. 13th Conference on Applied Climatology, Portland, OR, May 12-16.