USING LONG-TERM WEATHER RECORDS TO GENERATE COMPLETE DRAINMOD PRECIPITATION INPUT FILES[1]

J.J. Prenger, K.R. Mankin, B.C. Atherton, L.C. Brown[2]

ABSTRACT

Missing data from long-term weather station records can be a major impediment to the use of climatic data in simulation models. It is especially difficult to find long-term hourly rainfall records necessary for the DRAINMOD water management model. Therefore, a method is being developed to replace missing precipitation data using information from other weather stations. First, the precipitation files at several weather stations were compared to determine whether their rainfall characteristics correspond. Next, DRAINMOD's relative yield responses using those comparison stations versus response with a base station were examined. Results of the precipitation characteristic comparison suggest that selecting replacement stations in the same microclimatic region, or within 30 miles had significantly similar coincidence of rain events; in addition stations having within one inch/year annual average precipitation had statistically similar characteristics. However, DRAINMOD simulation relative yield comparisons did not show increased correlation between base and comparison site overall average relative yields when the stations were near in proximity or equivalent in annual average precipitation. Thus, further analysis is being conducted to analyze the component relative yields, evaluate the model's sensitivity to replacements at different times of the year, and analyze DRAINMOD's response when files with replaced missing data are used in the model.

INTRODUCTION

Computer simulation models have been widely used to study water and chemical movement in agricultural systems. The DRAINMOD water management model developed at North Carolina State University by Wayne Skaggs (1989) was used in this study. This model requires hourly rainfall and maximum and minimum temperature inputs. These inputs are used to quantify runoff, infiltration, evapotranspiration, and subsurface drainage requirements. Accurate and complete climatic data are essential to achieve representative crop yield results at the site being modeled. A major limitation of this model, and others, is often the lack of complete, long-term precipitation data for the specific region of concern. Hence, when trying to select rainfall

---

[1] Contribution from the Ohio Agricultural Research and Development Center (OARDC) and Ohio State University Extension, Dept. of Food, Agricultural, and Biological Engineering, The Ohio State University, in cooperation with the Department of Biological and Agricultural Engineering, Kansas State University. Project supported, in part, by OARDC, Overholt Drainage Education and Research Program, and OSU Extension; Department of Food, Agricultural, and Biological Engineering, The Ohio State University.

[2] Research Assistant, Dept. of Food, Agric., and Biol. Engr., The Ohio State University, Columbus; Assistant Professor, Dept. Biol. and Agric. Engr., Kansas State University, Manhattan; Graduate Research Associate and Associate Professor, Extension Agricultural Engineer, Dept. of Food, Agric., and Biol. Engr., The Ohio State University, Columbus.

inputs for the model, several problems can arise: these data are (i) unavailable for the region of concern; (ii) available, but at the wrong time scale; e.g., daily maximum and minimum instead of hourly; or (iii) available, but incomplete; i.e. missing days or months of data. Often weather stations have not been operational long enough to compile the long-term data needed to allow appropriate generalization for future conditions. Efforts to use computer generated data to fill these gaps in data have met with limited success (Elliot et al., 1992; Robbins and Skaggs, 1988). Thus despite the limitations, detailed weather station records remain the best source of precipitation data.

An extensive weather station network has been maintained throughout Ohio for over 40 years (NOAA, 1992). Hourly precipitation has been recorded at over 100 stations. When these data are used in hydrologic simulation models, it is important to use climatic data which are representative of each microclimatic region in the state, since local climates may vary over short distances.

Hence, the objectives of this preliminary study were to (1) evaluate the use of annual average precipitation as a guideline for selecting nearby weather stations with similar rainfall characteristics, and (2) test how closely overall average relative yields correspond between DRAINMOD simulations using precipitation inputs at a base and a comparison weather station. The overall goal is to produce representative, long-term precipitation and temperature records for each of Ohio's microclimatic regions for use in agricultural water management modeling.

METHODS

Several tests were used to determine an appropriate technique for assembling complete data sets. An initial hypothesis proposed that using annual average precipitation can be a guideline for selecting weather stations with similar rainfall characteristics. Data from these replacement stations will be as close to the actual rainfall characteristics at a base station as possible. This premise was examined by selecting five base stations which are shown in figure 1 (Owenby and Ezell, 1992). Seven to fifteen comparison stations were chosen for each base station. Each comparison station was then characterized in three ways: 1) by its proximity to the base station--proximity was described as *near* if the station was in the same microclimatic region as defined by ODNR (see figure 2; ODNR, 1997) or within 30 miles, and far if outside of these boundaries, 2) its average annual precipitation taken from an ODNR isopleth map (see figure 3; Harstine, 1992), and 3) whether its average annual precipitation was greater or smaller than that at the base station.

In order to examine the similarities in the precipitation events between each specific base and comparison station pair, the digital rainfall files were formatted to list each hour in the period of record with its rainfall (including zero values). A FORTRAN program was created to sum a 24-hour storm beginning each hour and produce a file with the largest storm for each day in the record. In turn, a spreadsheet program was developed to tabulate the percentage of times a rainfall event of specified value occurred at both the base and comparison stations. For example, each time a rainfall in the amount of 0.50 inches in 24 hours occurred at the base station, it was recorded, as well as each time it occurred at both the base and comparison stations. Finally a ratio was formulated demonstrating the percentage of time the events coincided.

A second comparison assessed the average overall relative yield response in DRAINMOD at the base stations versus the yield response at the comparison stations to test how closely their results corresponded. The climatic record for each station was input into DRAINMOD using selected poorly drained and well-drained soil series; other model inputs were kept the same for each site (*Note*: these inputs are still being reviewed; see Patterson et al., 1997). Only periods of record in which neither station had important missing information (i.e., missing months from the planting date through the harvest season) were included in the overall average yields.

An analysis of variance (ANOVA) procedure (SAS, 1989) was used to analyze these two data sets. In each, a 4 × 2 × 2 treatment structure was used to evaluate three variables: (1) difference in average annual precipitation (DIFF—4 levels: 0, 1, 2, and 5 in./yr difference), (2) proximity to the base weather station (PROX—2 levels: *near* which includes stations in the same Ohio microclimatic region or within 30 miles, and *far* or outside those boundaries), and (3) whether the difference in annual average precipitation was greater (higher) or smaller (lower) than the base weather station (HILO—2 levels: higher and lower). A probability level of $p<0.05$ was set as the test of significance.

In the first analysis, the percentage of times a selected weather station had precipitation events above a specified cutoff level within a 1- or 2-day window of separation compared with the base station was assessed. The minimum precipitation event cutoffs used for comparison were a 24-hour 0.25-in. event, and a 24-hour 0.50-in. event. Thus, four event combinations were possible: 0.25-in. event +/-1 day (Q1), 0.25-in. event +/-2 days (Q2), 0.50-in. event +/-1 day (H1), and 0.50-in. event +/-2 days (H2). Only periods of record in which neither station had missing data were used.

The second analysis tested the percentage difference between the average overall relative yield produced using precipitation inputs at the base station versus that from the comparison station--each characterized by the same DIFF, PROX, and HILO variables.

RESULTS AND DISCUSSION

Weather Station Precipitation Comparisons

From the ANOVA analysis, DIFF and PROX were found to be significant at the $p<0.0001$ level for all precipitation event combinations, while HILO was uniformly not significant as a main effect or in any interactions. Subsequently, the linear model procedure was run again using only the DIFF and PROX classes. The main effects DIFF and PROX were significant for all precipitation event combinations, while the DIFF×PROX interaction was significant ($p<0.05$) in only one instance (Q1) (Table 1). It is interesting to note that PROX was significant ($p<0.0001$) in all cases, while the significance of DIFF decreased (from $p<0.004$ to $p<0.038$) with increased event intensity or greater allowable days of separation between acceptable events in the base and comparison station. Although both proximity and difference in average annual rainfall are important, proximity may be the more critical of the two. This point is further demonstrated in Table 2, showing least squares means for each main effect. In all cases, the 0 and 1 in./yr levels were not significantly different from each other, but both were significantly different from the 5 in./yr level. Interestingly, the 1 in./yr treatment consistently had the largest least squares mean value even compared to stations with records having 0 in./yr difference from the base station.

These data indicate that similar average annual rainfall (within 1 in./yr) and near proximity are important variables in predicting whether precipitation records from two given weather stations will be similar. Thus, both are important criteria in selecting appropriate stations from which to take data to replace missing rainfall records at other locations.

Table 1. Analysis of variance for main effects and interaction (p<0.05)

	Q1 0.25 in./day +/- 1 day	Q2 025 in./day +/- 2 days	H1 0.50 in./day +/- 1 day	H2 0.50 in./day +/- 2 days
DIFF	0.0040	0.0135	0.0186	0.0382
PROX	0.0001	0.0001	0.0001	0.0001
DIFF×PROX	0.0438	0.0850	0.1352	0.1649

Table 2. Analysis of variance, least squares means

	Q1 0.25 in./day +/- 1 day	Q2 025 in./day +/- 2 days	H1 0.50 in./day +/- 1 day	H2 0.50 in./day +/- 2 days
		DIFF (in./yr)		
1	0.785 a†	0.825 a	0.691a	0.727 a
0	0.765 ab	0.810 ab	0.674 a	0.714 a
2	0.748 bc	0.795 bc	0.653 ab	0.693 ab
5	0.723 c	0.775 c	0.617 b	0.663 b
		PROX (Near or Far)		
N	0.787 ‡	0.827	0.712	0.745
F	0.723	0.776	0.606	0.654

† Values in the same column followed by the same letter are not significantly different (for p < 0.05).
‡ All column-wise comparisons are significantly different (for p < 0.0001).

<u>DRAINMOD Overall Relative Yield Comparisons</u>

The average overall relative yield over the period of record using a base station for the precipitation input in DRAINMOD was compared to the yield using the precipitation record at a comparison station. No significant differences were found by ANOVA analysis for main effects or interactions of DIFF, PROX, and HILO on either of the well-drained or poorly drained soils evaluated. Whereas, in the precipitation comparison discussed earlier, the PROX and DIFF variables had a significant effect on the rainfall events corresponding between the base and comparison station; in case of overall relative yield responses using a base and comparison station, there was no increased similarity between the responses if the stations were near in proximity or equivalent in annual average precipitation. This lack of significance was surprising. It is unclear at this time whether the results were caused by a real lack of sensitivity in the model to precipitation record changes or some other factor. Furthermore, the overall average relative yield from DRAINMOD is only one yield consideration: drought, excess, and planting delay reduction factors all have contributions to the overall yield. However, it was evident from these preliminary analyses of DRAINMOD results that the model is very sensitive to timing during the day in which the rainfall event occurs (see also Robbins and Skaggs, 1988). Precipitation events that lead to wet soil conditions in the planting and/or early part of the growing season should have substantial effects on the excess stress relative yield, while similar events in the summer, as ET increases, should have less of an effect. If the non-significance is real, it may suggest that over time, differences in annual rainfall amount are not as significant as the timing of the events at a site. This result would have a substantial bearing on replacement of missing precipitation data: missing data in the spring, and possibly the autumn, would be much more critical to be replaced accurately than in the winter or summer.

SUMMARY AND RECOMMENDATIONS

Analyses in progress include study of DRAINMOD's sensitivity to replacement of rainfall data for specific months. Systematic replacement of periods of data with information from nearby stations will help identify the causes of variability in DRAINMOD response test. Additional statistical analyses are necessary to examine the effects of proximity and average annual precipitation of comparison weather stations on each of DRAINMOD's relative yield components: drought, excess, and planting delay stress. Furthermore, precipitation data files completed using the considerations outlined above will be created for several Ohio microclimatic regions. These files will be tested for their achievement of DRAINMOD relative yields corresponding to the base site being modeled. Ultimately, these efforts will provide a source for reliable, long-term regional climatic records for use in modeling long-term effects of agricultural water management strategies.

REFERENCES

DRAINMOD User's Manual. 1989. Biological and Agricultural Engineering Department, North Carolina State Univ., Raleigh, NC.

Elliot, W.J., Qiong, W. and A.V. Elliot. 1992. Suitability of CLIGEN for generating rainfall data for DRAINMOD. *Transactions of the ASAE* 8(6):807-812.

ODNR. 1997. Monthly Water Inventory Report for Ohio. D.H. Cashell (compiler). ODNR Division of Water, Water Inventory Unit.

Owenby, J.R. and D.S. Ezell. 1992. Ohio Monthly Station Normals of Temperature, Precipitation, and Heating and Cooling Degree Days - 1961-90. Climatography of the United States No. 81. NOAA/NCDC. Asheville, NC.

Patterson, J.M., Atherton, B.C., Tornes, L., Gehring, R.M., Steiger, J.R., Prenger, J.J., Brown, L.C., Zimmerman, T.L., and G. Hall. 1997. Modeling drainage response and recommendations for a modern agricultural water management guide. ASAE Paper No. 97-2166. ASAE Summer Meeting. Minneapolis, Aug. 1997.

Robbins, K.D. and R.W. Skaggs. 1988. A regional rainfall disaggregation method for DRAINMOD. ASAE Paper No. 88-2567. ASAE Winter Meeting. Chicago, Dec. 1988.

SAS Institute. 1989. SAS/STAT user's guide, Version 6, 4th ed., Vol. 1. SAS Inst., Cary, N.C.

Acknowledgments. Many thanks to Jane Patterson, Brian Doan, Leslie Zucker, Kate Weber, and Kate Brown, The Ohio State University; John Boyer, Kansas State University; James Marron, Portland, Oregon, USDA-NRCS; Jeff Rogers, Ohio State Climatologist; and Bob Parkinson, Ohio USDA-NRCS for their help with data acquisition, analysis, graphic design and production**.**