Assessing methods to disaggregate daily precipitation for hydrological simulation

Peng Gao[1], Gregory J. Carbone[1], Daniel L. Tufford[2], Aashka Patel[1], Lauren. F. Rouen[1]

AUTHORS: [1]Department of Geography, University of South Carolina 709 Bull Street, Columbia, SC 29208 USA
[2] Department of Biological Sciences, University of South Carolina, 715 Sumter Street Columbia, SC 29208 USA
REFERENCE Proceedings of the 2012 South Carolina Water Resources Conference, held October 10-11, 2012 at the Columbia Metropolitan Convention Center

ABSTRACTAs researchers explore the impacts of climate change on water resources, they often will use downscaled data from General Circulation Models (GCMs). While such output provides our best estimates of climate response to anthropogenic change, it lacks the spatial and temporal resolution that may affect stream flow. For example, most hydrologic models require hourly inputs, but many statistical downscaling methods provide only daily data. In this study, we evaluate three different disaggregation methods to construct hourly precipitation time series from daily precipitation. These disaggregated data are used as input to the Hydrologic Simulation Program-Fortran (HSPF) watershed model distributed by the US Environmental Protection Agency (EPA) to compare model performance against observed stream flow in two watersheds representing two physiographic settings- South Yadkin (Piedmont), North Carolina and Black (Coastal Plain), South Carolina. The first method disaggregates daily precipitation using a triangular distribution centered around the middle of the day. The second method estimates the number and volume of rain events each day and selects measured intensity patterns from a nearby station to simulate hourly rainfall. The third method is the same as the second, but it conserves daily rainfall volume. Model performance was assessed by comparing observed to simulated stream flow from 2001 to 2003. Six commonly used performance statistics were derived for each simulation: index of agreement, mean absolute error, Nash-Sutcliff efficiency, percent bias, root mean squared error, and coefficient of determination. Though the performance of the three methods was similar, the second and the third methods were more robust than the first one.

INTRODUCTION

Continuous simulation modeling for watershed hydrology is a principle tool to investigate the impacts of climate change on water resources. The use of these models requires high spatial and temporal resolution (e.g. hourly or subdaily) rainfall data and faces two challenges due to the constraint of data availability. First, the model requires many years of hourly data for calibration and verification. However, precipitation data are often available only at coarser levels (i.e., daily). In the United States there are over 25,000 daily recording stations while only 8,000 hourly stations exist (Booner, 1998). Second, despite the wide use of downscaled data from General Circulation Models (GCMs) to estimate future climate, meteorological variables from the GCMs such as precipitation and temperature needed for hydrological simulation are typically at monthly or daily scales. Various methods have been developed to disaggregate daily precipitation to hourly time series to make them usable for continuous hydrologic simulation. These methods fall into two categories that are not necessarily mutually exclusive. The first group is represented by Hershenhorn and Woolhiser (1987) and Connolly, Schirmer and Dunn (1998). This group of methods assumes the number of events in a day, the starting and peaking time, duration of each event, and the amount of rainfall in each event follow certain distributions (e.g. Weibull distribution or mixed beta distribution). Different distributions were adopted in different study areas. Historic data were used to parameterize the distribution curves. The second group is represented by Choi, Socolofsky and Olivera (2008), Socolofsky, Adams and Entekhabi (2001) and Knoesen and Smithers (2009). This group of methods estimates the rainfall hours or events each day by stochastically selecting rainfall hours or events from the existing hourly precipitation records while forcing or maintaining certain statistical characters of the existing hourly rainfalls (e.g. cumulative distribution functions or the ratio of maximum hourly rainfall to daily rainfall). Though many disaggregation methods have been developed, few tests have been conducted to assess the performance of these methods on hydrologic simulations. In this study we bridge the disaggregation methods and hydrologic simulations. We hypothesized that the disaggregation method could have implications for how well the model performed compared to observed stream flow. We studied this by examining three different disaggregation methods to construct hourly precipitation time series from daily precipitation, then using those time series as input to previously developed models of two subwatersheds in the greater Winyah Bay Watershed in North and South Carolina.

STUDY AREA AND METHODS

This study used streamflow simulation models of the South Yadkin River (Piedmont in North Carolina) and Black River (Coastal Plain in South Carolina) Watersheds. These watersheds are part of a Winyah Bay Watershed model that was developed using the Better Assessment Science Integrating point and Non-point Sources/Hydrologic Simulation Program Fortran (BASINS/HSPF) suite of programs distributed by the USEPA. The Parameter Estimation (PEST) program was also used to assist with calibration (Rouen and Tufford, in prep). These two watersheds were selected watersheds for their different physiographic settings, which affects the stream flow generation. A virtual hourly precipitation station is created by taking the sum of the selected hourly stations weighted by the area of Thiessen polygons intersecting the watersheds. Precipitation in the model was derived using these polygons constructed using ESRI ArcGIS (Rouen and Tufford, in prep).

This first disaggregation method (hereafter triangular distribution method) distributes the daily precipitation into hourly precipitation following a triangular distribution among 24 hours using a look-up table (Table 1). The algorithm searches for the daily total closest to but larger than the measured daily precipitation. Then daily precipitation is distributed proportional to the hourly ratios in the table. The peak of the daily precipitation is always centered around the middle of the day.

The second disaggregation method (hereafter, the Socolofsky method) was first introduced by Socolofsky, Adams and Entekhabi (2001). The Socolofsky methodfirst builds a database of precipitation events from hourlyprecipitation stations selected by the user. A precipitation event is defined as a continous sequence of hourly precipitation, spearated from the next event by at least one hour with no precipitation. The daily precipitation is then disaggregated into hourly precipitation by borrowing the eventcharacteristics from the precipitation event database.

For this study, we used the virtual hourly precipitation station discussed above to build the database of precipitation events. The Cumulative Density Functions (CDFs) of event depth are constructed for each month in the verification time period 2001-2003. To dissagregate a depth of daily rainfall D_T, the Socolofsky method first findsthe ordinate a of the D_T from the corresponding monthly CDF. A random number is selected between 0 and $a$subject to a random distribution, the corresponding event is read from the CDF, and its depth $D_i$is subtracted from D_T. This process iteratively searches the new $D_i$and D_T, unitl the new D_T lower than an assigned minimum threshold ε. The $\varepsilon$is associated with the expected depth of the smallest rainfall evetns.The calibration of ε is described by Choi, Socolofsky and Olivera (2008).An excessive number of trace events can be avoided by setting ε and enforcing the $D_T$larger than ε. The remaining daily precipitation (i.e., the rest amount of rainfall which is belowε) is modeled as a single, one-hour event with exponentially distributed intensity. Once individual events D_i selected from the

Table 1. Daily total and corresponding hourly ratios of triangular distribution(part of the table is shown due to the linit of space)

	daily total	0	0.01	0.02	0.04	0.08	0.16
ratio for each hour	1	0	0	0	0	0	0
	2	0	0	0	0	0	0
	3	0	0	0	0	0	0
	4	0	0	0	0	0	0
	5	0	0	0	0	0	0
	6	0	0	0	0	0	0
	7	0	0	0	0	0	0
	8	0	0	0	0	0	0
	9	0	0	0	0	0	0
	10	0	0	0	0	0	0.01
	11	0	0	0	0.01	0.01	0.04
	12	0	0.01	0.01	0.02	0.03	0.06
	13	0	0	0.01	0.01	0.03	0.04
	14	0	0	0	0	0.01	0.01
	15	0	0	0	0	0	0
	16	0	0	0	0	0	0
	17	0	0	0	0	0	0
	18	0	0	0	0	0	0
	19	0	0	0	0	0	0
	20	0	0	0	0	0	0
	21	0	0	0	0	0	0
	22	0	0	0	0	0	0
	23	0	0	0	0	0	0
	24	0	0	0	0	0	0

CDF, the depths of hourly rainfall of these events are assigned to each hour of dissaggregated daily rainfall with the starting time taken from the actually starting time of these events. Depth of rainfall in the hours when the events overlap each are added together. The third disaggregation method (hereafter, the adjusted Socolofsky method) follows all steps of the Socolofsky method except that the remaining daily precipitation is directly placed into a one-hour storm event.

Using the hourly precipitation generated by the three methods as input, each simulated stream flow is compared against the observed stream flow in the verification time period 2001-2003 to evaluate the performance of the three methods in terms of hydrological simulation. Since the Socolofsky method and the adjusted Socolofsky method involve randomness when they select the precipitation evnets, ten hourly precipitation time series are created by each method to make the comparison capture the general trend. The comparion is also referred to the simulated stream flow using the precipitation input from the virtual hourly station. To sum up, twenty two hourly precipitation time series are input to the HSPF (i.e., one from the virtual hourly station, one created by the triangular distribution method, and ten each generated by the Socolofsky method and the adjusted Socolofsky method respectively). Each simulated stream flow is compared to the observed stream flow using six commonly used performance statistic derived for each simulation: index of agreement (d), mean absolute error (MAE), Nash-Sutcliff efficiency (NS), percent bias (p-bias), root mean squared error (RMSE), and coefficient of determination (R^2).

RESULTS AND DISCUSSION

Table 2 and 3 show daily and monthly average stream flow in the South Yadkin watershed, North Carolina. The daily stream flow simulated by using the precipitation input from the virtual hourly station (model verification) had the best fit based on the performance statistics, followed by some simulations using the precipitation from the adjusted Socolofsky method (i.e., AS5 and AS7 in table 2), the Socolofsky method (i.e., A10 and A1 in table 2), and the triangular distribution method. When the daily stream flow aggregated to monthly scale, all the simulations from the adjusted Socolofsky method outperformed the virtual hourly station, and one of the simulations from the Socolofsky method had the best fit (i.e., S5 in table 3).

Table 4 and 5 show daily and monthly average stream flow in the Black River watershed, South Carolina. The triangular distribution method performed best in terms of both daily stream flow and monthly average stream flow. The adjusted Socolofsky method and the Socolofsky method had similar performance to the stream flow simulated using the precipitation input from the virtual hourly station.

Table 2 Statistic of daily stream flow in the South Yadkin River Watershed, North Carolina

	d	MAE	NS	p-bias	RMSE	R^2
V	0.91	67.63	0.71	-15.58	81.53	0.74
T	0.92	69.20	0.70	-15.81	83.29	0.74
S1	0.93	63.95	0.75	-15.58	76.17	0.78
S2	0.91	68.24	0.71	-18.21	82.02	0.75
S3	0.91	66.93	0.70	-10.40	82.62	0.72
S4	0.93	60.89	0.76	-9.04	74.22	0.77
S5	0.93	59.25	0.79	-11.47	69.74	0.81
S6	0.87	84.07	0.51	-26.72	105.51	0.62
S7	0.90	70.13	0.68	-10.38	85.44	0.70
S8	0.90	72.95	0.62	-25.55	92.82	0.71
S9	0.91	68.43	0.68	-15.60	86.25	0.71
S10	0.92	66.13	0.73	-17.99	78.47	0.77
AS1	0.93	62.72	0.75	-14.31	75.37	0.78
AS2	0.92	64.92	0.75	-13.58	76.35	0.77
AS3	0.93	62.79	0.76	-13.84	74.44	0.78
AS4	0.93	63.66	0.75	-14.57	75.16	0.78
AS5	0.93	63.34	0.76	-12.96	74.88	0.78
AS6	0.93	63.14	0.76	-12.85	74.25	0.78
AS7	0.93	63.09	0.76	-13.46	74.66	0.78
AS8	0.92	65.93	0.74	-13.84	77.93	0.76
AS9	0.93	63.14	0.75	-14.44	75.71	0.78
AS10	0.93	62.64	0.76	-13.29	74.31	0.78

Abbreviations: the simulated stream flow using the precipitation input from the virtual hourly station (V), the triangular distribution method (T), the Socolofsky method (S) and the adjusted Socolofsky method (AS). The numbers (1-10) indicate ten simulations for S and AS respectively

By examining the performance of the three methods in the two watersheds (table 2-5), we found the following general trends. Though the adjusted Socolofsky method and the Socolofsky method had similar performance to the stream flow simulated using the precipitation input from the virtual hourly station, the Socolofsky method was more variable than the adjusted Socolofsky i.e. the statistics of the Socolofsky method in table 2 have a wider range than those of the adjusted Socolofsky method. Similar trends can also be seen in tables 3-5. The Socolofsky method assigns the residual daily rainfall which belowε as a single, one-hour event with exponentially distributed intensity. Therefore, it cannot conserve the total depth of daily precipitation. This is the likely explanaiton for the higher variation of the Socolofsky method. The triangular distribution method was the best one in the Black River Watershed but no better than other methods in the South Yadkin River Watershed. The robustness of the triangular distribution method needs to be tested in additional watersheds.

Table 3 Statistic of monthly average stream flow in the South Yadkin River Watershed, North Carolina (The meaning of the abbreviations is the same as these in table 2.)

	d	MAE	NS	p-bias	RMSE	R^2
V	0.91	67.63	0.71	-15.58	81.53	0.74
T	0.92	69.20	0.70	-15.81	83.29	0.74
S1	0.93	63.95	0.75	-15.58	76.17	0.78
S2	0.91	68.24	0.71	-18.21	82.02	0.75
S3	0.91	66.93	0.70	-10.40	82.62	0.72
S4	0.93	60.89	0.76	-9.04	74.22	0.77
S5	0.93	59.25	0.79	-11.47	69.74	0.81
S6	0.87	84.07	0.51	-26.72	105.51	0.62
S7	0.90	70.13	0.68	-10.38	85.44	0.70
S8	0.90	72.95	0.62	-25.55	92.82	0.71
S9	0.91	68.43	0.68	-15.60	86.25	0.71
S10	0.92	66.13	0.73	-17.99	78.47	0.77
AS1	0.93	62.72	0.75	-14.31	75.37	0.78
AS2	0.92	64.92	0.75	-13.58	76.35	0.77
AS3	0.93	62.79	0.76	-13.84	74.44	0.78
AS4	0.93	63.66	0.75	-14.57	75.16	0.78
AS5	0.93	63.34	0.76	-12.96	74.88	0.78
AS6	0.93	63.14	0.76	-12.85	74.25	0.78
AS7	0.93	63.09	0.76	-13.46	74.66	0.78
AS8	0.92	65.93	0.74	-13.84	77.93	0.76
AS9	0.93	63.14	0.75	-14.44	75.71	0.78
AS10	0.93	62.64	0.76	-13.29	74.31	0.78

Table 4 Statistic of daily stream flow in the Black River Watershed, South Carolina (The meaning of the abbreviations is the same as these in table 2.)

	d	MAE	NS	p-bias	RMSE	R^2
V	0.89	387.22	0.69	-18.73	656.19	0.72
T	0.91	332.88	0.75	-9.14	586.33	0.77
S1	0.88	406.48	0.65	-19.96	696.45	0.67
S2	0.89	371.88	0.70	-17.11	642.77	0.73
S3	0.88	385.38	0.67	-14.89	670.64	0.71
S4	0.88	404.92	0.65	-19.66	695.13	0.67
S5	0.89	386.18	0.69	-15.66	651.73	0.71
S6	0.90	386.77	0.70	-18.89	638.21	0.73
S7	0.89	380.69	0.69	-16.58	654.67	0.71
S8	0.89	353.51	0.70	-9.96	648.65	0.72
S9	0.88	420.55	0.66	-22.75	688.25	0.69
S10	0.88	407.98	0.67	-21.13	678.30	0.70
AS1	0.89	377.54	0.69	-17.85	651.55	0.72
AS2	0.89	377.80	0.69	-16.82	653.17	0.72
AS3	0.89	377.67	0.69	-18.08	654.14	0.72
AS4	0.89	374.14	0.69	-17.65	652.59	0.72
AS5	0.89	377.88	0.69	-16.15	656.10	0.72
AS6	0.89	375.29	0.69	-15.35	658.06	0.71
AS7	0.89	377.09	0.69	-17.49	656.39	0.72
AS8	0.89	379.17	0.69	-17.89	655.83	0.72
AS9	0.89	378.42	0.69	-16.95	656.40	0.72
AS10	0.89	373.90	0.69	-17.23	651.41	0.72

Table 5 Statistic of monthly average stream flow in the Black River Watershed, South Carolina (The meaning of the abbreviations is the same as these in table 2.)

	d	MAE	NS	p-bias	RMSE	R^2
V	0.95	290.37	0.83	-18.89	394.34	0.86
T	0.96	233.90	0.87	-9.31	348.16	0.88
S1	0.94	310.69	0.80	-20.06	434.42	0.82
S2	0.96	265.48	0.85	-17.27	367.70	0.88
S3	0.94	296.39	0.82	-15.04	413.00	0.85
S4	0.94	316.20	0.79	-19.75	445.17	0.81
S5	0.94	288.67	0.81	-15.86	415.59	0.83
S6	0.95	275.97	0.84	-19.05	389.87	0.86
S7	0.95	295.59	0.83	-16.73	395.75	0.86
S8	0.95	265.25	0.84	-10.08	382.30	0.87
S9	0.93	326.57	0.78	-22.89	447.22	0.82
S10	0.94	321.99	0.79	-21.27	436.30	0.83
AS1	0.95	286.71	0.83	-17.99	396.28	0.86
AS2	0.95	284.88	0.83	-16.97	393.99	0.86
AS3	0.95	287.04	0.83	-18.21	391.09	0.87
AS4	0.95	282.49	0.84	-17.79	389.81	0.87
AS5	0.95	282.54	0.84	-16.32	390.14	0.87
AS6	0.95	282.69	0.83	-15.49	397.54	0.86
AS7	0.95	286.87	0.83	-17.63	391.60	0.86
AS8	0.95	291.76	0.82	-18.06	404.81	0.85
AS9	0.95	291.47	0.83	-17.10	400.18	0.86
AS10	0.95	275.47	0.84	-17.37	382.07	0.88

CONCLUSION

In this study, we examined the performance of three methods that dissaggregate daily precipitation to hourly time series for contiuous hydrological simulation in the South Yadkin River Watershed in North Carolina and the Black River Watershed in South Carolina. The results suggest that the adjusted Socolofsky method is the most robust in terms of performance when compared to the model verification run using the observed hourly precipitation as input. The adjusted Socolofsky method can be a useful means of disaggregating the daily precipitation from GCMs under different scenarios, when these GCMs are adopted to investigate the impact of future climate change on water resources. Additional comparisons should be carried out in more watersheds to test the robustness and consistency of these methods. More completed testing should lead to better decision making in water resource management.

LITERATURE CITED

Bonner, W. D. (1998) Future of the national weather service cooperative observer network.National Research Council, National Academy Press, Washington, D.C.

Choi, J., S. A. Socolofsky and F. Olivera (2008) Hourly disaggregation of daily rainfall in Texas using measured hourly precipitation at other locations. *Journal of Hydrologic Engineering*, 13**,** 476-487.

Connolly, R. D., J. Schirmer and P. K. Dunn (1998) A daily rainfall disaggregation model. *Agricultural and Forest Meteorology,* 92**,** 105-117.

Knoesen, D. and J. Smithers (2009) The development and assessment of a daily rainfall disaggregation model for South Africa. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques,* 54**,** 217-233.

Hershenhorn, J. and D. A. Woolhiser (1987) Dissaggregation of daily rainfall. *Journal of Hydrology,* 95**,** 299-322.

Hummel, P., J. Kittle, Jr., and M. Gray. WDMUtil Version 2.0, A Tool for Managing Watershed Modeling Time-Series Data User's Manual, Contract No. 68-C-98-010, Work Assignment No. 2-05, Aqua Terra Consultants, Decatur, GA, 2001.

Socolofsky, S., E. Adams and D. Entekhabi (2001) Disaggregation of daily rainfall for continuous watershed modeling. *Journal of Hydrologic Engineering,* 6**,** 300-309.

Rouen, L.F. and D. Tufford, in prep.Watershed modeling of mixed-physiography basins in the Carolinas.