Multivariate Non-stationary Stochastic Streamflow Models for Two Urban Watersheds

M. Ng¹, R. Vogel², C. Kroll³

¹Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155; email: Melissa.Ng@Tufts.edu
²Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155; PH (617) 627-4260; email:Richard.Vogel@Tufts.edu
³Department of Environmental Resources and Forest Engineering, State University of New York College Environmental Science and Forestry, Syracuse, NY 13210; PH (315)470-6500; email:CNKroll@ESF.edu

Abstract

It is now common knowledge that human activities including water diversions and land development, in addition to climate, have a significant impact on streamflow and other hydrologic processes. The effects of changes in land use, water withdrawals and climate are experienced simultaneously and may interact with each other. Increasingly, it is necessary for hydrologists to develop methods for predicting streamflow which account for human activities. Most previous research has investigated the effects of changes in climate or land use on streamflow, and rarely both simultaneously. There are even fewer studies that address the effects of water withdrawal on streamflow and even fewer on the impacts of all three. The primary objective of this study is to develop stochastic streamflow models for predicting monthly streamflows for two urbanizing watersheds in the vicinity of Boston. Streamflows in urban environments are impacted by climatic, land use and water use effects. In this initial study, a multivariate non-stationary monthly stochastic streamflow models are developed using multivariate linear regression.

Introduction

Hydrologic systems evolve due to a variety of both natural and anthropogenic influences: urbanization and changes in climate and water use are examples of such influences. The evolution of the watershed system in response to such influence at the scale of years to centuries has been termed its hydromorphological response (Dressler et al., 2006 and Vogel, 2011). In this study, we concentrate on the hydromorphological response of watersheds to anthropogenic influences associated with urbanization and climatic processes.

Over the past few decades a wide range of environmental damages have been linked to urbanization including, but not limited to: decreases in biodiversity, increased flooding, degradation of human health, decreases in evapotranspiration and decreased quality of air, water and soil resources. The hydrologic effects of urbanization are primarily a result of both continuous and abrupt land-use and infrastructure changes that lead to changes in the land and atmospheric component of the hydrologic cycle as well as changes in the water use cycle. Urbanization leads to construction of water distribution systems and storm water and sewage infrastructure. Essentially, watershed systems evolve due to changes in land use, climate, water use and an array of other anthropogenic influences.
The effects of climate change on streamflow have been studied extensively. It is well known that temperature and streamflow are inversely related, due to the fact that an increase in temperature results in an increase in evapotranspiration, whereas precipitation and streamflow have a positive correlation. In regions that do not experience extremes in climate, precipitation effects on streamflow generally overwhelm those of temperature (Claessens et al., 2006, Kirshen et al., 1995). Zhu et al., 2005, could not adequately explain the trends in streamflow at 47 streams in Pennsylvania with climate alone, emphasizing the need to incorporate land use effects. Additionally, the streams exhibited a downward trend in streamflow, which is opposite to the direction of streamflow trends projected by others, documenting that modeling streamflow is site-specific.

Urbanization replaces vegetative cover with impervious surfaces, decreasing infiltration capacity and groundwater recharge, causing an increase in storm runoff and a decrease in base flows (Brandes et al., 2005, Hejazi et al., 2007, Lee et al., 2006, Claessens et al., 2006). Maintaining baseflow is critical especially during dry seasons when it acts as the primary source of streamflow (Todd et al., 2007). Others believe urbanization may stabilize or increase baseflow due to leaks from water infrastructure and septic systems (Burns et al., 2005, Brandes et al., 2005) as well as decreases in evapotranspiration which result from replacing vegetative cover with impervious surfaces. Brandes et al., 2005 found that in the Delaware River basin’s low to moderately dense residential areas, impervious area has a negligible effect on low flows. The USGS, 2001, found a similar result in their study in the Ipswich River basin, in eastern Massachusetts.

Cuo et al., 2008, examined the effects of land use and climate change on streamflow for watersheds in differing elevations in the Puget Sound basin. It was found that in regions with pronounced snow cover, both climate and land use had an impact on streamflow, whereas in regions that had minimal snow, land use effects on streamflow were more pronounced than those of climate. These studies show that both land use and climate impact streamflow and it is not adequate to use land use or climate independently to model streamflow.

Due to studies that could not adequately explain streamflows based on land use and climate, researchers have begun to investigate the impact of water withdrawals on streamflows. The USGS, 2001, found that water withdrawal correlated more to low flows than land use in the Ipswich River basin; however, these effects were not significant in flows above the median flow. Brandes et al., 2005, had a similar finding in the Delaware River basin, as did Golladay et al., 2007, in Georgia.

Very few studies have considered the effects of climate change, land use change and water withdrawals simultaneously on streamflow. Claessens et al., 2006, assessed the sensitivity of the Ipswich River in eastern Massachusetts and found that for mean annual streamflow, land use effects were negligible while withdrawals decreased the flows. It was found that climate change proved to be the driving force, as it balanced out the effects of the withdrawals, resulting in no net trend in streamflow. More specifically, precipitation effects overwhelmed any temperature effects. Lee et al., 2006, studied the sensitivity of total runoff and low flows in a small watershed in Korea during the dry season. Land use in this case also did not
affect flow and once again climate was found to be the driving force. Unlike the findings from Claessens et al., 2006, temperature had a larger influence on streamflow than precipitation, most likely due to the fact that Korea experiences a more dramatic dry season than eastern Massachusetts. Similar to the USGS 2001 study, water withdrawal effects only exhibited significant influence on low flows.

While there appear to be more and more studies on the effects of human activities and precipitation on streamflows over the years, there are very few comprehensive studies. As development increases, it is imperative to understand the behavior of the hydrologic system in response to simultaneous changes in its natural surroundings as well as its human surroundings. This study will emphasize the importance of considering climate change, land use and water withdrawals, together, in an integrated fashion in streamflow modeling.

**Stochastic Streamflow Modeling**

Stochastic streamflow models are used widely for reservoir planning and design for a wide array of activities ranging from irrigation, water supply, recreation and hydropower to drought planning and river basin planning (Salas et al., 1980). Nearly all previous attempts to develop monthly stochastic streamflow models use historical streamflow alone to predict future monthly streamflows. By using historical streamflow, stochastic models account for persistence within the hydrologic system; meaning high flows tend to follow high flows and low flows tend to follow low flows. It is due to this feature that stochastic streamflow modeling was chosen for this study. A major assumption of stochastic streamflow modeling is that the hydrologic system is stationary. Stationary implies that the future will be statistically indistinguishable from historical records. In an urbanizing watershed it is necessary to account for the non-stationary terms which impact streamflow including climate, land use and water use. In this initial study we introduce a non-stationary monthly stochastic streamflow model, which as is customary, uses previous streamflow to predict current streamflow. However, our model also includes climate, land use and water use. Hence, the derived models are useful for planning for future anthropogenic changes with these exogenous variables.

**Study Areas**

The Neponset River basin (Figure 1) and the Aberjona River basin (Figure 2) are used as case studies. The watershed boundaries are defined by USGS HUCs. The Neponset River drains 34.7 sq miles, and the Aberjona drains 24.7 sq miles. The figures below display the watershed boundaries and stream gauge locations.
Figure 1. Neponset River basin in eastern Massachusetts. The USGS stream gage is indicated on the map by a triangle.
Figure 2. Abjerona River basin in eastern Massachusetts. USGS stream gage is indicated by the triangle.
Population changes were observed for Middlesex County and Norfolk County, which encompass the Aberjona River watershed and Neponset River Watershed respectively. From the figure below it can be seen that both counties experienced rapid increases in population up until about the 1970s at which the counties experienced a plateau in population.

![Population Changes 1940-2008](image)

Figure 3. Approximate Population changes in the Neponset River watershed (Norfolk County, MA) and the Aberjona River watershed (Middlesex County).

**Streamflow**

Streamflow behaviors for the Neponset River and Aberjona River can be described by their flow duration curves (figures 4 and 5 respectively). From 1940-2007, there is an increase in flood flows and a decrease in low flows in the Neponset River. In contrast, the Aberjona River exhibits overall increasing flows during the same time period. The flow duration curves are separated into 1940-1969 and 1970-2007 for two purposes: 1) The population in these watersheds have two distinct trends whose threshold occurs at 1970, and 2) ample climate change research suggest 1970 to be a turning point at which most noticeable changes can be observed in hydrologic systems, climate systems and other natural systems.
Land use

European settlement began within the Boston Basin in 1620. Agriculture was the dominant land use during this period until about the late 1800s with the onset of industrialization and westward expansion. With agriculture abandonment, forest cover was able to increase until about the 1920s when residential and commercial areas expanded. As a result, the amount of forest cover continuously decreased (Hall, et. al, 2002). As expected, with an increase in population there is an increase in development. Figure 6 shows the amount of developed land, in acres, for the Neponset River watershed and Aberjona River watershed. Land use is defined by total acres of developed land (impervious area), which includes: recreational areas, high density residential areas of less than ¼ acre, transportation, commercial and
industrial areas. The Neponset River watershed is characterized largely by forested areas, whereas the Aberjona River watershed is heavily urbanized.

![Land Use Change 1940-2007](image)

Figure 6. Acres of impervious area in the Aberjona and Neponset River basin from 1940 to 2007.

Water Use

In the models developed, only in-basin water withdrawals were included. Figures 7 and 8 show the total annual in-basin water withdrawals for Aberjona River watershed and Neponset River watershed. Walpole, Foxboro and Medfield contain wells within the Neponset River watershed. These towns have relied heavily on these groundwater supplies as the main drinking water sources from 1986-2007. Woburn, Winchester and Reading all have sources within the Aberjona watershed, but are also supplemented by Massachusetts Water Resources Authority (MWRA) resources. Woburn has used the MWRA as a source for drinking water from as far back as 1986, as has Winchester. Reading, however, like Walpole has relied on groundwater sources from 1986-2006. Starting in 2007 the town has relied solely on MWRA for drinking water.
Available Data

Streamflow, land use, and climate data were obtained for the years 1940-2007 and water withdrawal data was obtained for the years 1986-2007. Table 1 lists the sources of the datasets. The USGS gauges used in this study are Neponset River at Norwood, MA (USGS gauge 01105000) and Aberjona River at Winchester, MA (USGS gauge 01102500). Land use data was collected from the Harvard Forest School for 1920. The dataset only included a portion of the towns within each river basin. For Neponset River basin, land use data was only available for the towns of
Dover, Medfield and Norwood. For the Aberjona River basin, land use data was available for Burlington, Reading, Winchester and Woburn. MassGIS provided land use data for each town within the two river basins for the years 1971, 1985 and 1999. From each of these datasets, the amount of impervious land, in acres, for the years 1920, 1971, 1985 and 1999 were determined. The acres of impervious land spanning the years 1940-2007 were then interpolated between these years. Precipitation data was collected from PRISM for the years 1983-2007. Water withdrawal data was collected for in-basin withdrawals only as mentioned in the previous section.

<table>
<thead>
<tr>
<th>Table 1. Description of data available.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
</tr>
<tr>
<td>Streamflow</td>
</tr>
<tr>
<td>Land use</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Precipitation</td>
</tr>
<tr>
<td>Water Withdrawal</td>
</tr>
</tbody>
</table>

**Methodology for Non-stationary Monthly Stochastic Streamflow Modeling**

In this initial study, a multivariate regression methods was used to select and fit non-stationary models of monthly streamflow. For example, the Neponset River monthly stochastic streamflow model takes the form

\[
\ln[Q(m,y)] = a + b\cdot\ln[Q(m-1,y)] + c\cdot[Q(m,y-1)] + d\cdot\ln[P(m,y)] + e\cdot\ln[L(y)] + f\cdot\ln[W(m,y)] + \varepsilon
\]

where
- \( Q(m,y) \) = monthly streamflow in month m and year y (cfs)
- \( P(m,y) \) = monthly precipitation in month m and year y (inches)
- \( L(y) \) = land use in year y (acres)
- \( W(m,y) \) = water use in month m and year y (gallons)
- \( \varepsilon \) = normally distributed error with mean zero and constant variance

What is new about this study is that previous monthly stochastic streamflow models have not included the terms for precipitation, water use and land use. The above model can find use in planning applications where one’s interest is in determining how future traces of monthly streamflow will be impacted by changes in precipitation, land use and/or water use.

**Results and Discussion – Nonstationary Monthly Stochastic Streamflow Model**

Table 2 and 3 summarize the monthly nonstationary stochastic streamflow model for the Neponset River watershed and Aberjona River watershed. In all cases,
model parameters were highly significantly different from zero, as indicated by the very small p-values reported in parentheses in Table 2. For the Neponset River, the addition of landuse and in-basin water withdrawals into the monthly mean model not only improved the fit of the model, but also documents that both landuse and water withdrawals impact streamflow. The model residuals appear normally distributed and homoscedastic. Similar to the results from the Neponset River, the Aberjona River monthly mean model improved with the inclusion of human activities. The model residuals also appear well-behaved. In this study, landuse had no significant effect on the monthly mean streamflow in the Aberjona River. Figures 9 and 10 compare the model fits to the observed data for the Neponset River and Aberjona River.

| Table 2. Summary of Neponset River Monthly Mean Regression models: ln[Q(m,y)] = a + b*ln[Q_1(m-1,y)] + c*ln[Q_2(m,y-1)] + d*ln[P_1(m,y)] + e*ln[P_2(m,y-1)] + f*ln[W_1(m,y)] +... + g*ln[W_2(m,y-1)] + h*ln[LU(y)] |
|---|---|---|---|---|---|---|---|---|---|---|
| Model | a | b | c | d | e | f | g | h | R-sq (pred) | SER |
| Q_1, Q_2 | 0.740 (0.001)† | 0.582 (0.000) | 0.223 (0.000) | | | | | | 46.63% | 0.654 |
| Q_1, Q_2, P_1, P_2 | 0.539 (0.000) | 0.285 (0.000) | 0.776 (0.000) | -0.229 (0.000) | | | | | 79.23% | 0.406 |
| Q_1, Q_2, P_1, P_2, L | 0.577 (0.000) | 0.308 (0.000) | 0.815 (0.000) | -0.208 (0.000) | -0.0447 (0.045) | | | | 79.14% | 0.406 |
| Q_1, Q_2, P_1, P_2, W_1, W_2 | 22.994 (0.000) | 0.612 (0.000) | 0.208 (0.000) | 0.745 (0.000) | -0.186 (0.000) | -0.707 (0.009) | -0.537 (0.000) | | 82.58% | 0.370 |
| Q_1, Q_2, P_1, P_2, W_1, W_2, LU | 31.068 (0.000) | 0.614 (0.000) | 0.196 (0.000) | 0.747 (0.000) | -0.175 (0.000) | -0.761 (0.0013) | -0.576 (0.000) | -0.917 (0.022) | 82.76% | 0.367 |

†Models are derived for the years 1986-2004
*Three dots denote that the estimated values are not significantly different from zero based on its p-value (p>0.05)
†Terms in parenthesis are the p-values of the model parameters

| Table 3. Summary of Aberjona River Monthly Mean Regression models: ln[Q(m,y)] = a + b*ln[Q_1(m-1,y)] + c*ln[Q_2(m,y-1)] + d*ln[P_1(m,y)] + e*ln[P_2(m,y-1)] + f*ln[W(m,y)] |
|---|---|---|---|---|---|---|---|---|---|
| Model | a | b | c | d | e | f | R-sq (pred) | SE |
| Q_1, Q_2 | 1.052 (0.000)† | 0.505 (0.000) | 0.170 (0.000) | | | | 30.17% | 0.656 |
| Q_1, Q_2, P_1, P_2 | 0.490 (0.000) | 0.259 (0.000) | 0.873 (0.000) | | | | 78.23% | 0.363 |
| Q_1, Q_2, P_1, P_2, W_1, W_2 | 11.045 (0.005) | 0.521 (0.000) | 0.225 (0.000) | 0.872 (0.000) | -0.156 (0.000) | -0.578 (0.000) | 79.98% | 0.347 |

†Terms in parenthesis are the p-values of the model parameters
ΔModels are derived for the years 1986-2004
*Three dots denote that the estimated values are not significantly different from zero based on its p-value (p>0.05)
Figure 9. Model fits and observed data for the Aberjona River from the years 1986-2007. Model includes previous flows, monthly precipitation, land use and water withdrawals.

Figure 10. Model fits and observed data for the Aberjona River from the years 1986-2007. Model includes previous flows, monthly precipitation and water withdrawals.
Conclusion

The initial Non-stationary multivariate monthly streamflow models reported here indicate the importance of including human activities such as land use and water withdrawals into monthly streamflow models. Importantly, the monthly models for both the Neponset River and Aberjona River incorporate precipitation, land use and water withdrawals, and it was shown that their inclusion resulted in significant improvements in the overall goodness-of-fit of the resulting models over traditional stochastic streamflow models which do not include such exogenous variables. Traditional stochastic streamflow models tend to include only previous streamflows, since there is an assumption that previous climate, land use and water use conditions are stationary. It is no longer possible to make such stationarity assumptions, hence we envision a need for the types of models introduced here in future planning and design studies.

References


Claessens, L., Hopkinson, C., Rastetter, E., Vallino, J. 2006. Effect of historical changes in land use and climate on the water budget of an urbanizing watershed. Water Resources Research, 42


Singh, V.P., J.F. Cruise. 1990. Stochastic Streamflow Modelling for Reservoir Planning and Management. *Hydrology of Mountain Regions. II – Artificial Reservoirs; Water and Slopes*. 194
