Response to referee comments

We thank the referee for taking the time to review this manuscript again. We have repeated the comments below and our responses are indented.

Comments

I have reviewed the changes that the authors made, based on my original comments. I believe that the changes to the methodology and the explanation of the methodology are well done. There are improvements to the description of data and metadata used in the study, but I still have a major problem with how some of the data/metadata are used.

I have extensive personal experience with USGS data collection techniques and don't believe the authors have a good understanding of the way USGS collects and reports data and metadata, based on the use and interpretation of some of the metadata in their study.

-1- Change of base discharge and change of gauge datum do not indicate any change in flows or in measurement techniques. As I mentioned in my original review, "the 'base discharge' is a level set to allow 3-4 peaks per year on average to exceed this level. Instantaneous peaks above this level are then recorded. A 'change of base discharge' does not indicate any change in actual flows recorded or any anthropogenic change in watersheds above a gauge. " This may have been unclear though, so I'll add some more information. The base discharge level doesn't affect USGS computation of flows, it only affects what gets reported as a peak above base discharge at a streamflow gauge. These peaks are sometimes used in peak-over-threshold studies (rather than using annual peaks). Change in base discharge should not be used in this study. It doesn't indicate anything of use for the study and takes away from the other metadata, such as indicated regulation, that do mean something.

-2- Change in gauge datum also does not indicate any change in flows or change in measurement techniques. As I mentioned in my original review "The 'change of gauge datum' also does not indicate any change in flows or anthropogenic influence. It indicates only that the arbitrary zero gauge height for the rating curve has been changed, normally because of changes in the gauge control point on a river (the riffle or channel section that controls the relation between river height and flow at a gauge)."

To add more information, change in gauge datum does not represent a discontinuity in measurement techniques or flows. Just the opposite, it means that USGS is continually making flow measurements to verify the rating curve at a gauge. USGS has done this throughout its long history. Many control sections for gauges change because the riffle or channel controlling the relation between river height and flow can change over time. This is due primarily to channel/riffle changes, often caused by high flows. Sand channels are much less stable than bedrock channels. When a gauge datum is changed, it's because the channel at the gauge (not necessarily other channel sections nearby) changed enough to bring the gauge height below the arbitrary zero point that was established when the gauge was created. The arbitrary zero point is then changed so there won't be negative gauge heights, necessitating a change in the rating curve. There is no discontinuity of flow magnitudes on either side of this change. There is a discontinuity in the arbitrary height of flow that corresponds to the flows. This is not relevant change for your study. It would be relevant if trends over time in river heights were being studied.

There is a new statement in the report (p. 8, line 19) “Changes in the rating curve used to estimate streamflow from measured water levels are not recorded in the USGS notes but may be a significant source of variation in low flow values that is not accounted for”. This statement is not true. Changes in the rating curve are made to reflect the current relation between stream height and flow. As explained above, USGS regularly checks the rating curve to make sure the current relation is accurate. If the rating curve has changed based on coincident measurements of streamflow and river height, the rating curve is changed. Again this does not represent a discontinuity in the flow data. There would be a discontinuity in the flow data if this was not done regularly.

Change in gauge datum should not be used to classify trends. It is not a meaningful code for the purposes of the current study. I looked at the percentage of gauges noted as having change in datum, for the different categories in Table 1. Most of the categories (no trend, decreasing trend, increasing trend, etc.) have about 13% to 18% of gauges that had changed datums. In other words, all these groups have about the same percentage of gauges with changed datum. The exception is the “Step Change” category with only about 5% of gauges with a change in datum. This is likely explained by the fact that a lot of the change-of-datum gauges are in the mid-Atlantic region where there are a lot of sand channels (Figure 1), but there aren’t many step changes in this area (Figure 2).

This paper would be strengthened by using the metadata in the interpretation of results that has meaning (such as indicators of regulation) and removing ones with no meaning (change in base discharge and change in gauge datum) in all locations in the article. All text, tables, and figures that reference these two metadata codes should be changed to remove them from the study.

> We have changed the manuscript to address these comments. Most importantly we have removed these categories from the list of codes and have updated the related text in sections 2.2, 3.2 and 5.2, figures 4 and 7 and table 1. Importantly, this does not change our results very much. The number of sites with step changes and flagged regulation remains high (57%) whilst the sites without step that are flagged decreases.

Response to referee comments

We thank the referee for taking the time to review this manuscript again. We have repeated the comments below and our responses are indented. Where we have added/edited text, we have highlighted this in italic font. Please note that some of the responses to Part 1 of the referee comments were addressed in the last round of review, and so the current manuscript already incorporates those changes.

Part 1. Omitted Points 15-25

From the previous review, there is no trace of the authors' responses to the points 15 to 25. I urge the authors to reply/address them (pasted below).

> We apologize for this omission. This was an unintended mistake on our part. We provide the responses below.

```
point-15
Page 2770 Lines 22-24:
You introduce the index ``day of the year of low flows'', but there is
no indication on how it is obtained. You provide a description at the
beginning of section 4.2 – that you could improve (e.g. clarifying how
you identify the 4- month periods) and move to this section.
```

> We removed the details of this from an earlier version of the manuscript to reduce the page length. However, we agree that some minimum level of detail is needed here and so we now provided that at the end of this section (now called "2.3. Low Flow Indices" in response to other comments) and moved the explanation of the identification of the low flow seasons from section 4.2. to here:
>
> "We also calculate the day of the year of low flows and use this to identify the primary (and in some regions the secondary) low flow season, as well as any long-term changes in timing. *The primary season is defined as the 4-month period that contains the majority of the low flow occurrences, and the secondary season as the 4-month period that contains the majority of the remaining low flows. If the onset time of the low flow season for a site occurs 70% to 100% in a specific month, that site is assumed to have only one low flow season. The sites that have low flow events occurring 40-70% of the time in one month and 20-40% of the time in a different month are characterized as having two low flows seasons*."
>
> Section 4.2 is now:
> "4.2. Variability in low flow timing
> Figure 8 summarizes the distribution of the onset of the low flow season for Q7, *for the primary season* (top panels) and *the second season* (bottom panels). The left panels show the onset month of the season and the right panels show the probability of the onset season in that month. For Q1, 353 sites out of 395 (almost 90%) sites *have a single low flow season, and the onset* of the season changes from north to south. Most of the sites north of North Carolina have low flow seasons starting in July, which is generally driven by the slight decline in precipitation during the autumn as well as the increased evaporation during the

summer (Small et al., 2006). In Florida the season starts in April-May. For coastal sites, the season starts earlier (mostly in June), and for sites in the southwestern part of the domain, the season starts mostly in September-October.
The sites with two low flows seasons are mostly in Florida, and along the coastline of Georgia, South and North Carolina, New York, New Jersey, and Maine and their second season occurs mostly in fall. For New York, New Jersey, and some sites along the west coastline of Florida, the second low flow season mostly starts in November and December. Sites near the Gulf of Mexico and some sites in North Carolina have second low flow seasons starting in April. The second low flow season for the far northeast sites begins in December or January and can be related to freezing conditions that may store water as snow and river ice."

point-16
Page 2771 Lines 13-14:
Visual inspection simply provides an indication. I suggest to either delete this phrase or replace ``can be very helpful in determining'' with something like ``can provide indication in the attempt to assess stationarity''.

We have rewritten this sentence to incorporate the reviewer's suggestions and to better link to the hydrological literature highlighting the benefits of visual inspection (or more broadly exploratory data analysis):

"Visual inspection of the time series and the changes therein *can provide an indication in the attempt to assess stationarity*, *in that a change in the underlying process leads to changes in values that are obvious (Lins and Cohen, 2011; Koutsoyiannis, 2011; Serinaldi and Kilsby, 2015).*"

point-16
Page 2771 Lines 21-24:
You should clarify the following: 1) Provided that autocorrelation is an issue for both MK and Pettitt tests, if autocorrelation is present the Pettitt test is applied, but the same is not valid for the MK test, why? Also: for MK there are adaptations of the test proposed by Hamed and Rao (1998) and Yue and Wang (2002,2004) to account for autocorrelation, did you consider this option?

We have updated the algorithm to better reflect our intended analysis, including accounting for autocorrelation in both tests, and this is addressed in response to point 18 later. Here we have clarified this sentence to note that autocorrelation will affect both tests and that analysis of the autocorrelation is carried out before applying either test:

"An identified change in the mean by either of the first two tests would rule out stationarity, except in the case of autocorrelated data, for which the *Pettit and* Mann–Kendall tests will characterize too many sequences of the time series as having a *step or trend and therefore increase the rejection rate of the null hypothesis of no change* (Douglas et al., 2000; *Serinaldi and Kilsby, 2015*). Therefore, analysis of autocorrelation is carried out before conducting the Mann–Kendall *and Pettit* tests."

point-17
Page 2771 Line 27:
``We assume that the change year corresponds to human intervention'' I find this assumption questionable. As written in point-6, a change point could result from climate variability.

In response to other reviewers' comments, we have added a sentence in the introduction paragraph to note that step changes can also be because of natural causes, but also note that our assumption is based on identifying abrupt and visually obvious step changes, which are likely to be due to anthropogenic influences. Visual inspection of the time series indicates that there are obvious abrupt shifts in many of the time series that are unlikely to be of natural origin, and these are identified by the combination of the Ljung-Box and Pettitt tests as described below in the updated description of the algorithm.

"Low flow time series (and flows in general) can show two general types of non-stationarity: gradually increasing or decreasing trends, and abrupt changes (Villarini et al., 2009) in the mean and/or variability. As McCabe and Wolock (2002) observe, the distinction between a gradual trend and a step change is important, particularly for climate-change impact studies, since climate change usually manifests as a trend and not a step change. We therefore assume that step changes (*abrupt and visually obvious*) in the time series are indicative of an anthropogenic effect, and that gradual trends reflect a climate effect, *which may be due to anthropogenic climate change or long-term persistence (Cohn and Lins, 2005). As it is possible that step changes may be driven by natural variability (e.g. McCabe and Wolock, 2008) our assumption is based on identifying abrupt and visually obvious step changes.*"

point-18
Page 2772 Lines 3:
In light of the previous observations I find this algorithm should be reconsidered. Also, a visual (flow chart) of the algorithm would be useful to guide the reader through the different steps.

The description of the algorithm in the manuscript is not completely consistent with the overall approach and we have updated it to better reflect our intention and respond to this and other reviewer comments, including accounting for lag-1 autocorrelation in the Pettit and M-K tests, after testing the overall autocorrelation structure of the time series. We removed a flow chart from an earlier version of the manuscript, and believe that the updated description of the algorithm is sufficient.

"The three statistical tests (Ljung-Box, Pettitt and Mann-Kendall) were combined into a recursive algorithm to identify non-stationarity in the low flow time series and decompose the series into *potentially* stationary sub-series. In the first step of the algorithm, a Ljung-Box test with 20 *lags* was applied to the entire time series of each site, and sites with significant overall autocorrelation (*5% significance level*) were identified. *The Ljung-Box test identifies sites that are non-stationary* *and is able to identify sites with abrupt changes because the series of values before the change appear to be autocorrelated relative to the values after the change, and vice-versa. This was confirmed by visual inspection of the time series.* For the sites with significant *overall* autocorrelation, we *then* applied the Pettitt test (*5% significance level*) to *confirm the existence of any step change and identify its timing. The series were pre-whitened to remove lag-1 autocorrelation following Kumar et al. (2009). It is necessary to identify sites with potential step changes using the Ljung-Box test first because the Pettitt test will identify step changes in time series with gradual trends. Similarly the MK test will identify gradual trends in series with step changes*. If a significant change is found *by the Pettitt test*, the series is split into two parts either side of the step change. Each part is assumed to be a new series at the same location, and if it has a record length of 30 years or more, the decomposition algorithm is applied again. If the length is less than 30 years, the site is removed from further consideration. If a statistically significant step change is not identified, we *note that the series is autocorrelated overall. We then applied the Mann-Kendall (MK) test (5% significance level) on the remaining sites to identify statistically significant trends in the data. Again, the series were pre-whitened to remove lag-1 autocorrelation*. The series and sub-series are assigned categories as follows:

1. Category 1: Non-autocorrelated site with no trend (MK=0);
2. Category 2: Non-autocorrelated site with a statistically significant decreasing trend (MK=-1);
3. Category 3: Non-autocorrelated site with a statistically significant increasing trend (MK=1);
4. Category 4: Autocorrelated site with statistically significant step change, time series split and the sub-series re-categorized recursively;
5. Category 5: Autocorrelated site with no step change. "

```
point-19
Page 2774/L20-2775/L10:
''From table 1 we observe that:``. See point-2.
```

This section has been updated to reflect the comments from point-2 and other reviewer comments. This section notes the occurrence of flags for sites with identified non-stationary behavior but does not specifically attribute the changes to the flagged influence.

```
point-20
Page 2775 Line 11:
See point-6 on causes of abrupt changes neglected in this study.
```

We have addressed this in response to point-6 and other reviewer comments, and have updated the text in various places to make it clear that the identified shifts are abrupt and visually obvious.

```
point-21
Page 2776 Line 21:
Figure 7b: there must be something wrong with the counting - 55
```

decreasing trends seems like too much compared to Figure 7a (same number?). Also dots overlap a lot, might be a good idea to reduce the size.

There was an error in Figure 7b in how the decreasing trends were plotted. In any case, Figure 7 has been updated to show the results of the updated algorithm, including pre-whitening of the data. The dots have been reduced in size to aid visualization.

point-22
Page 2778 Lines 6-7:
''applied within the 4 month season of Q1 and Q7 low flows``. It is not clear to which series the MK and Pettitt tests have been applied.

We updated the analysis to focus only on Q7 and Q30. The text has been substantially updated to reflect the results from the updated algorithm and in response to other comments.

point-23
Page 2778 Lines 11-:
''Out of the remaining 335 sites``, should numbers add up in e.g. Fig. 9A (17+13+1)?

These numbers refer to sites with identified changes (step or trends) out of the total of 335 sites of which the rest have no identified changes. These numbers have been updated with the new algorithm. The figures and text have also been updated to reflect the new numbers.

point-24
Page 2779 Line 4
As you write in the Conclusions: ''However, definitive attribution will require detailed analysis of these competing factors and possibly carefully crafted modeling studies.`` I would not call section 5.1 Attribution.., maybe Towards the attribution of trends in low flows, or similar.

We have changed the section title to “Potential Drivers of Trends in Low Flows”

There should also be mention, either in this section or in the introduction, of the distinction between trend detection and attribution and on the difficulties of performing the latter (e.g. Merz et al. (2012)) [Merz, B., Vorogushyn, S., Uhlemann, S., Delgado, J., Hundecha, Y., 2012. Hess opinions ‘‘more efforts and scientific rigour are needed to attribute trends in flood time series. Hydrol. Earth Syst. Sci. Discuss. 9, 1345–1365, HESSD.]

Agreed. We discuss the question of low-flow generating mechanisms in the context of attribution of changes in the conclusions section. We have added the reference in the following context at the end of the paragraph:

“However, definitive attribution will require detailed analysis of these competing factors and possibly carefully crafted modeling studies. *In parallel with calls for* *more rigorous efforts at attributing changes in flood time series (Merz et al., 2012), increased effort is also needed for understanding and attributing changes in low flows.*"

>> [Added Oct 25 – complement to point-24] >> there should be mention of novel approaches to attribute changes to external drivers in line with the ideas outlined in Merz et al (2012). For example Harringan et al. (2014) and Prosdocimi et al. (2015), in which specific effort is made to identify whether some anthropogenic variable can be related to some evolving behaviour in hydorlogical series, keeping the natural climatic variability in control either via the Multiple working hypothesis approach or via a complex statistical model. [Harrigan, S., Murphy, C., Hall, J., Wilby, R. L., and Sweeney, J.: Attribution of detected changes in streamflow using multiple working hypotheses, Hydrol. Earth Syst. Sci., 18, 1935-1952, doi:10.5194/hess-18-1935-2014, 2014; *Prosdocimi, I., T. R. Kjeldsen, and J. D. Miller (2015), Detection and attribution of urbanization effect on flood extremes using nonstationary flood- frequency models, Water Resour. Res., 51, 4244–4262, doi:10.1002/2015WR017065*]. >>

Agreed. We have added this to the end of the same paragraph in the conclusions:

"*Several new approaches have been put forward recently that show promise for detecting and attributing changes in hydrological time series, including extremes, based on multiple working hypotheses (Harrigan et al., 2014) and complex statistical modeling (Prosdocimi et al., 2015).*"

point-25
Page 2779 Lines 20-28:
No reference to antecedent precipitation is found in the results, I think this block belongs to the results section, to be later discussed in this section.

We prefer to keep this with the discussion on potential reasons for the changes, rather than as a separately presented result, which may be confusing.

Part 2. Updated manuscript.
In general, I find it inappropriate to talk about attribution under this study's framework: the authors have changed the title of Section 5.1 from "Attribution" to "Potential Drivers", they should also change terminology in the remainder of the paper accordingly: i.e. L. 19, L. 180, L. 433, L. 498, etc.

We have updated the text as follows:

"The goal of this paper is to examine non-stationarity in low flow generation across the eastern U.S. and explore the *potential* anthropogenic influences or climate *drivers.*"

"The results on the variability and trends in are given in Section 4. Finally, we discuss the results, the *potential drivers of changes and their* implications, and present conclusions in Section 5.

> "The *drivers* of trends at these sites *are* therefore likely related to climate variability/change and/or land use change, rather than management of, or influence on, flows."
>
> "To understand the *potential drivers* of these trends more comprehensively, Figure 10 shows the Q_7 trend magnitude and the antecedent precipitation for the previous 180 days."

With regards to the pre-whitening, the authors have cited Kumar (2009) (please add to the references list), but have not specified which method they used of the four proposed by the reference.

> The method used is the trend-free pre-whitening, which was proposed by Yue et al. (2002). We have updated the text as follows:
>
> "The series were pre-whitened to remove lag-1 autocorrelation *using the trend-free pre-whitening method of Yue et al. (2002) as implemented by* Kumar et al. (2009)."
> The Kumar et al. (2009) reference has been added.

Moreover, the hypothesis that step changes are human induced and that slow changes are related to e.g. long range dependence is questionable or, at the very least, a huge approximation. If this stays a disclaimer should be put in the discussions.

> We have added two sentences at the end of the discussion that highlights that our approach is not perfect and that our assumption (or simplification) is subject to the complexities of the influences on low flow generation:
>
> "The results of this study can help in understanding changes in low flows across the eastern U.S., and the impact of anthropogenic and natural changes. It can therefore provide information for water management, and restoration of stream flows and aquatic habitats. *Although we do not claim to make a definitive judgment on whether low flows at a particular site are influenced by human activities or are completely free of influences because of the complexities of low flow generation, our approach shows promise for systematically identifying sites for further investigation, especially where supporting information (such as site notes) are available to support the statistical results. Our approach may be especially useful for exploring large-scale, climate-driven changes in the low flow regime where pooling of results across sites increases confidence in the robustness of any identified changes.* The methods are readily transferable to other parts of the U.S. and globally, given long enough time series of daily streamflow data, although further work is required to understand their universal application."

Regarding the updated manuscript:

Line 151-154: ``The goal of this paper is to examine non-stationarity in low flow generation across the eastern U.S. and attempt to systematically identifiy time series that are potentially free of the effects of human intervention and examine these in terms of the impact of climate variability and change.''
I think there are too many claims in this paragraph, I also suggest ``attempt to systematically identify [..]'' comes first.

We do not claim to make a definitive judgment on the influence of human activities and/or climate variability, but to examine the data for signs (statistical and documented) of influence. We have updated this sentence as follows, changing the wording to be more cautious. We have also removed "climate change" because we do not know whether the changes in precipitation are due to climate change or variability:

"The goal of this paper is to examine non-stationarity in low flow generation across the eastern U.S. *by attempting* to ~~systematically~~ identify time series that are potentially free of the effects of human intervention and examine these in terms of the impact of climate variability ~~and change~~."

Line 169-170: ``our assumption is based on identifying abrupt and visually obvious step changes''. I don't consider this an assumption that can hold, but a simplification with two inherent shortcomings: how arbitrary is the judgment of an abrupt change? More importantly, natural variability can produce abrupt changes too. This issue was raised in point 17 too.

We applied the Pettitt test with a significance level that identified step changes that are visually abrupt. This ensures that only large and abrupt changes, likely associated with some form of human influence are detected. At the same time, we agree that this approach will not detect human effects that are gradual in nature, or for step changes that are small relative to the variability. We also compare the results of the step change test with the site information, which shows that most of these sites with step changes are indeed influenced by management, increasing our confidence that are approach, and therefore out assumption or simplification, has potential for identifying sites with low human influence in situations where no site information is available. We have updated these sentences as follows to note that this is a simplification and not an assumption.

"We therefore make the *simplification* that step changes (abrupt and visually obvious) in the time series are indicative of an anthropogenic effect, and that gradual trends reflect a climate effect, which may be due to anthropogenic climate change or long-term persistence (Cohn and Lins, 2005). As it is possible that step changes may be driven by natural variability (e.g. McCabe and Wolock, 2008) this *simplification* is based on identifying abrupt and visually obvious step changes."

We should also note that the McCabe and Wolock (2002) judged the changes in annual minimum streamflow to be a step change based on visual examination of the time series of normalized departures averaged over all 400 sites in the earlier version of the HCDN. Although the step change is apparent in their figure, it could just as easily be interpreted as a gradual trend if plotted differently (see Serinaldi and Kilsby, 2015 for examples of alternative interpretations of changes in streamflow statistics). In fact we have no idea whether a step change model or a gradual trend model fits the McCabe data better, and how the type of change (step or gradual) manifests spatially for individual sites. Furthermore, the evidence for a step change via attribution from precipitation is not provided in the McCabe paper. They report high correlations between annual mean precipitation and standardized annual median streamflow across the U.S. but only refer to previous studies on increases in precipitation. No evidence is provided of a step change in precipitation, and no mention is given of links to climate variability in the form of climate indices. Although we do not dismiss the idea that step changes could occur because of the step changes in precipitation or large-scale climate indices, the evidence for this is not apparent in this particular paper.

Serinaldi F, Kilsby CG. Stationarity is undead: Uncertainty dominates the distribution of extremes. Advances in Water Resources 2015, 77, 17-36.

`Line 231-233: Both referee 1 and myself had suggested to check your results on HCDN stations. The authors added that 64 of the sites are in the HCDN-2009 database. I strongly suggest the authors to go beyond listing the number of HCDN stations and actually report on how their method performs on those stations.`

We applied the same methods to the HCDN dataset for all 64 sites in our domain that had data for the common time period (1951-2005). We find that 82% and 86% of the sites were found to be in category 1 (stationary) for Q7 and Q30, respectively, with most of the remaining sites identified in category 3 (increasing trend; 9% and 8%) or category 6 (autocorrelated; 5% and 4%). This confirms that our method is capable of identifying sites without management (step changes).

We have added some discussion of this at the end of section 3.2:

"*We also applied the algorithm to the HCDN-2009 sites within the domain, to confirm that the algorithm can identify sites that have been independently determined as unaffected by human influences. We found that 82% and 86% of these sites were placed in category 1 (stationary) for Q7 and Q30, respectively, with most of the remaining sites in category 3 (increasing trend; 9% and 8%) or category 6 (autocorrelated; 5% and 4%).* "

Nonstationarity of low flows and their timing in the eastern United States

S. Sadri, J. Kam, and J. Sheffield

Department of Civil and Environmental Engineering,
Princeton University,
Princeton,
New Jersey,
08544
U.S.A.

Corresponding author: Justin Sheffield, Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, U.S.A, 08544 (justin@princeton.edu)

Abstract

The analysis of the spatial and temporal patterns of low flows as well as their generation mechanisms over large geographic regions can provide valuable insights and understanding for climate change impacts, regional frequency analysis, risk assessment of extreme events, and decision-making regarding allowable withdrawals. The goal of this paper is to examine non-stationarity in low flow generation across the eastern U.S. and explore the potential anthropogenic influences or climate drivers. We use nonparametric tests to identify abrupt and gradual changes in time series of low flows and their timing for 508 USGS streamflow gauging sites in the eastern US with more than 50 years of daily data, to systematically distinguish the effects of human intervention from those of climate variability. A time series decomposition algorithm was applied to 1-day, 7-day, 30-day, and 90-day annual low flow time series that combines the Box-Ljung test for detection of autocorrelation, the Pettitt test for abrupt step changes and the Mann-Kendall test for monotonic trends. Examination of the USGS notes for each site showed that many of the sites with step changes and around half of the sites with an increasing trend have been documented as having some kind of regulation. Sites with decreasing or no trend are less likely to have documented influences on flows. Overall, a general pattern of increasing low flows in the northeast and decreasing low flows in the southeast is evident over a common time period (1951-2005), even when discarding sites with significant autocorrelation, documented regulation or other human impacts. The north-south pattern of trends is consistent with changes in antecedent precipitation. The main exception is along the mid-Atlantic coastal aquifer system from eastern Virginia northwards, where low flows have decreased despite increasing precipitation, and suggests that declining groundwater levels due to pumping may have contributed to decreased low flows. For most sites, the majority of low flows occur in one

Deleted: attribution

Deleted: to

Deleted: or changes in measurement characteristics

season in the late summer to autumn, as driven by the lower precipitation and higher evaporative demand in this season, but this is complicated in many regions because of the presence of a secondary low flow season in the winter for sites in the extreme northeast and in the spring for sites in Florida. Trends in low flow timing are generally undetectable, although abrupt step changes appear to be associated with regulation.

Keywords: Eastern U.S.; Low flows; Non-stationarity; Abrupt change; Gradual trends; Autocorrelation; Ljung-box test; Mann-Kendall test; Pettitt test

1. Introduction

Low flows - the minimum flow in a river during the dry periods of the year--- are an important part of the streamflow regime that have direct impacts on water supply, water quality, and ecosystem health (Bradford and Heinonen, 2008). Knowledge of low flow characteristics and generation mechanisms over large geographic regions is important for regional frequency analysis, risk assessment of extreme events, decision-making regarding allowable basin withdrawals and water quality, and understanding climate change impacts (Tallaksen and van Lanen, 2004). For example, in every state of the U.S., estimates of low flow statistics are needed for issuing and/or renewing of National Pollution Discharge Elimination System permits, as required by provisions in the Clean Water Act of 1977 (U.S. Senate, 2002). Furthermore, low flow periods are critical to aquatic habitats due to potentially low dissolved oxygen concentrations and/or high pollutant concentration (U.S. Senate, 2002). However, the study of low flow statistics and patterns have received little attention in comparison to droughts and floods (Kroll et al., 2004). Poff et al. (1997) emphasize the need of paying particular attention to low flows because they present critical stresses and opportunities for a wide array of riverine projects.

Low flows are generally controlled by subsurface flows sourced from groundwater that maintain flows during the dry periods of the year, such that low flow volumes are related to the physiological and geological make up of the area. In some regions, where precipitation is significant in the warm season, surface flows also play a role in maintaining low flows. However, our understanding of these low flow generating mechanisms is limited (Smakhtin, 2001), and is further compounded by the sensitivity of low flows to changes in climate, land use and human impacts on stream flow (Rolls et al., 2012). For example, large-scale teleconnections may play an important role in driving inter-annual to multi-decadal changes in streamflow (e.g. Mauget, 2003) and low flows (e.g. Giuntoli et al., 2013). Regulation generally introduces non-stationarity into low flow time series that impedes the development of regional or at-site frequency analysis models. In most instances, such models show a high standard error between modeled and observed quantiles (Kroll et al., 2004).

In the eastern United States, (defined as the area covering the 20 ecoregions of the eastern US (USGS, 2012)), both direct anthropogenic and climate influences may have impacted low flows, including land use change impacts via changes in sub-surface flow and groundwater recharge, direct impacts on flows via reservoirs and other streamflow management, and changes in precipitation and evaporation that have altered recharge. In particular:

1. In the U.S., more than 85% of the surface runoff is artificially controlled and nearly 1 million km of rivers are affected by dams (Poff et al., 1997). Surface water covers 4.5% of the eastern U.S., and the majority of streams have been flagged by the U.S. Geological Survey (USGS) as regulated. The USGS estimates that the spatial extent of surface water increased by 1.3% during 1973-2000, with most of this increase in the southern coastal plain and southern Florida coastal plain (USGS, 2012) and associated with reservoir developments required to meet the needs of the expanding population. Figure 1a shows the location of major dams in the eastern U.S. (defined as those 50 feet or more in height, or with a normal storage capacity of 5,000 acre feet (~6,200,000 m^3) or more, or with a maximum storage capacity of 25,000 acre feet (~30,800,000 m^3) or more (USACE, 2012)). Generally dams and reservoirs are considered the largest man-made regulations on streamflow, but other sources include farm ponds, surface water extraction, inter-basin transfers, and wastewater treatment

plant discharge (e.g. Walker and Thoms, 1993; Acreman et al., 2000; Brandes et al., 2005; Thomas, 2006; Deitch et al., 2009; Kustu et al. 2010).

2. The eastern U.S. has gone through significant land use change over the past several decades. For example, between 1973 and 2000, 8.2% of the 23,620,000 km^2 of the northeast ecoregion and 8.9% of the 30,000,000 km^2 of the southeast ecoregion experienced changes associated with active timber harvesting and replanting, which may have impacted low flows and related environmental and ecosystem well-being (USGS, 2012). Furthermore, in the expanding urbanized areas of the region with high levels of impervious ground, infiltration has decreased, which may have led to a decrease groundwater recharge and low flow volumes (USGS, 2013). On the other hand, urbanization can lead to increase in low flows because of leakages from water supply and wasterwater pipes, direct wastewater discharge, reduced evapotranspiration, and water imports that can offset groundwater pumping (e.g. Brandes et al., 2005).
3. The region is one of the wettest parts of the U.S. receiving 700-1600 mm of precipitation per year. However, due to population growth and associated increased use of surface and groundwater resources, the future is expected to bring water stress for this area (Averyt et al., 2013). Some of these changes are already being observed. For example, USGS (2013) reports on 3-10 km^3 of depletion of unconsolidated and semi-consolidated sand and gravel aquifers of the east coast between 1900 and 2008. Overuse of surface water in turn does not allow recharge of groundwater leading to groundwater depletion. In parts of the eastern U.S., groundwater resources have become limited and hence municipal and industrial water users are increasingly relying on surface waters (e.g. Daniel and Dahlen, 2002). Changes in both surface water and groundwater use have impacts on low flows.

4. Precipitation has likely changed over the past several decades (Karl and Knight, 1998; Small et al., 2006). Evaporation may have changed due to increasing atmospheric demand from higher temperatures (e.g. Walter et al., 2004), although direct measurements of evaporation are limited in spatial and temporal coverage. Each of these changes may impact on low flows and in some cases may combine to exacerbate or counteract changes in low flows. Warmer temperatures may have also impacted winter-time low flows, via changes in snow (Burakowski et al., 2008) and river ice (Hodgkins et al., 2005).

Past evaluations of changes in low flows over the eastern U.S. have mainly been within studies on the entire U.S. and often with respect to mean and high flows. Douglas et al. (2000) estimated trends in both flood and 7-day low flows for three major geographic regions in the U.S. (East, Midwest, and West) over two time periods: 1959-1988 and 1939-1988, and found evidence of upward trends in low flows across the Midwest, but not in the eastern U.S. Other studies have attempted to explain the general patterns of low flow trends. For example, Small et al. (2006) analyzed trends in annual 7-day low flow, average, and high flows along with seasonal precipitation over individual basins in the U.S. for 1948-1997. The number of sites shown to have statistically significant trends in low flows and fall precipitation in the eastern U.S. was small and restricted to the south of Maine, western Pennsylvania, coastal areas of South Carolina, and western Florida. In the northeast and west of Pennsylvania, precipitation showed an increasing trend during the fall but not during the spring and the increase in fall precipitation appeared to result in an increase in low flows in the northeast areas. The only statistically significant decrease in the low flows was found in the south Atlantic-Gulf region, west of Florida, consistent with the findings from Lins and Slack (1999). However, no specific reason for this decreasing trend was given. McCabe and Wolock (2002) examined historic changes in streamflow, using the annual minimum, median, and maximum daily streamflow at 400 sites across the U.S. during 1941-1999. They found an increase in annual minimum and median daily streamflow around 1970 that primarily occurred in the eastern U.S. as a step change, rather than a gradual trend. Andreadis et al. (2006) used model simulations to examine trends in soil moisture, runoff, and drought characteristics over the U.S. for the period 1915-2003. They found increasing runoff over parts of the northeast, which was most evident during winter months, with decreases in hydrological and agricultural drought, and drying trends in the summer in the southeast, with increases in drought. These changes were attributed to changes in precipitation, and they speculated that increasing drought in the southeast was associated with higher atmospheric demand due to warming. Although these studies are generally consistent for the eastern U.S. they tend to focus on the spatial pattern of trends in 7-day low flows only, and were limited to earlier periods available at the time of the study. Furthermore, these studies focused on sites that were deemed to have minimal anthropogenic influence, and so did not explore the role of anthropogenic influences, such as land cover change or water withdrawals (Brown et al., 2013).

The goal of this paper is to examine non-stationarity in low flow generation across the eastern U.S. by attempting to identify time series that are potentially free of the effects of human intervention and examine these in terms of the impact of climate variability. A way to determine whether a river has been subject to anthropogenic influences, at least in terms of regulation, is to examine the site notes for the gauging station. However, site notes might not be available, complete, or accurate, and examining the notes for multiple sites can be unwieldy. Furthermore, whether a site is determined to be regulated or not is often based on high flows and not on low flows. Here, we develop an approach that makes the simplification that the impact of human

Deleted: and

Deleted: systematically

Deleted: and change

activities can be detected in the streamflow data in a systematic way. This is generally more efficient and can complement site notes or compensate for errors in them. Low flow time series (and flows in general) can show two general types of non-stationarity: gradually increasing or decreasing trends, and abrupt changes (Villarini et al., 2009) in the mean and/or variability. As McCabe and Wolock (2002) observe, the distinction between a gradual trend and a step change is important, particularly for climate-change impact studies, since climate change usually manifests as a trend and not a step change. We therefore make the simplification that step changes (abrupt and visually obvious) in the time series are indicative of an anthropogenic effect, and that gradual trends reflect a climate effect, which may be due to anthropogenic climate change or long-term persistence (Cohn and Lins, 2005). As it is possible that step changes may be driven by natural variability (e.g. McCabe and Wolock, 2008) this simplification is based on identifying abrupt and visually obvious step changes.

Deleted: assume

Deleted: our assumption

Our overall approach is to use nonparametric statistical tests to identify abrupt and gradual changes in the value and timing of *n*-day low flows, and identify stationary segments of the time series. Furthermore we analyze the co-variability of low flows with antecedent precipitation to understand the influence of changes in precipitation and atmospheric demand (as quantified by potential evapotranspiration) on changes in low flows. The paper is organized as follows: Section 2 describes the streamflow data and the methodology, including the use of three straightforward and already-established statistical methods, for identifying non-stationarity in annual low flow time series. The results on the systematic identification and characterization of abrupt changes in low flow volumes and timing are presented in Section 3. The results on the variability and trends in are given in Section 4. Finally, we discuss the results, the potential drivers of changes and their implications, and present conclusions in Section 5.

Formatted: Font:Not Italic

Deleted: ir attribution and

2. Data and Methods

2.1. Study area

Our study area covers the eastern U.S. from Maine in the northeast to Florida in the southeast and westwards to the Appalachian Mountains and the Mississippi River in the south, and is based on the 20 ecoregions of the eastern U.S. (USGS, 2012). According to the USGS (2012), 52.4% of the eastern ecoregion in 2000 was forest. However, both forests and agriculture have been in decline since 1973 and instead, urbanization has increased and continues to increase. Most land cover change has occurred in the southeast and is associated with forest harvesting, agricultural abandonment, and development (USGS, 2012). Changes in the northeast have been mostly associated with timber harvesting. Changes in the north Central Appalachian region have been more heterogeneous and include examples of non-mechanical transitional change. Unlike the northeastern Coastal Plain, the southern Florida Coastal Plain has not experienced loss of agricultural land, but the largest decrease in surface water and significant loss of wetlands (-2.4%). Changes in surface water in the southern Coastal Plain have primarily been due to urbanization (USGS, 2012).

The eastern U.S. is one of the wettest parts of the country (Small et al., 2006), with average precipitation of about 1100 mm per year, with maxima along the coastal plain and the mountains of the Appalachians. Part of the precipitation in the northeast falls as snow in the wintertime (Hayhoe et al., 2007). The eastern seaboard is susceptible to tropical storms and hurricanes during the Atlantic hurricane season, normally running from June to end of November, which enhance precipitation across southern and eastern parts, and play a role in alleviating drought (Kam et al., 2013). The El Niño-southern Oscillation (ENSO) alters precipitation patterns across the southeast (Colby, 2008). Coastal extra-tropical cyclones bring the bulk of the wintertime precipitation to that region, forming along the natural temperature gradient of the Gulf stream before moving up the coastline (Gurka et al., 1995). Seasonally, there are slight changes in the precipitation distribution through the year. For example, Burlington, Vermont has a summer maximum and a winter minimum while Portland, Maine has a fall and winter maximum, with a summer minimum in precipitation. The water supply in the northeast is mainly derived from surface waters, which are heavily regulated to meet the water supply demand of urbanized areas such as New York City, although there has been an increase in groundwater sources in recent years. In contrast, the southeast, including Florida, lies on active aquifers (USGS, 2009). Projections of future climate indicate an increase in precipitation over the eastern U.S. (Hayhoe et al., 2007; EPA, 2008) with consequences for changes in low flows across the region.

2.2. Streamflow data

Initially, 4878 sites with daily streamflow records were retrieved from the USGS National Water Information System (NWIS) (USGS, 2014) for the eastern U.S. as defined by Hydrological Unit Codes (HUC) of 01, 02, or 03. Previous studies on low flows (e.g. Kroll et al., 2002, 2004; Douglas et al., 2000) have used the USGS Hydro-Climatic Data Network (HCDN; now updated to HCDN-2009; Lins, 2012), in part because anthropogenic influences at these sites are deemed to be negligible, but as such, is limited to 204 sites across the domain. Of the original 4878 sites, 2811 were active in the 2000's or later. Among these, 1092 sites had at least 30 years worth of daily data, 740 sites had 50 years or more, and 324 sites had 75 years or more. We used sites with at least 50 years of data as a balance between having enough of data at each site to identify long-term changes and the need to have many sites to characterize the spatial pattern of changes. We included only sites that did not have any missing years of daily data. This reduced the number of sites to 508 (Figure 1b). Only 64 of these sites are in the HCDN-2009 database and have data for the common time period (1951-2005) that is used for analyzing trends across the domain (see section 4). The drainage area of the candidate sites ranges from very small (5-100km^2) to large (38,000-67,000km^2), with the majority of areas between 200-500 km^2 and these are spread fairly uniformly across the study area. The majority of the 508 sites are clustered on the eastern flank of the Appalachians and the northeast from eastern Virginia to New Hampshire. There is also a cluster of smaller catchments in central Florida. The mean, median, minimum and maximum record lengths are 74, 72, 50, and 120, respectively.

Based on the USGS site notes (available on the NWIS website), we identified sites that are flagged as: regulated, partially regulated, flow below the rating curve limit, dam failure, affected by urbanization, change of base discharge, and change of gauge datum. It should be noted that the USGS flags are developed for instantaneous peak flows and while it is uncertain whether these are directly applicable to low flows, it is likely that low flows are more sensitive to regulation. Some of the flags are unrelated to anthropogenic influences and are unlikely to have impacted the continuity of flow magnitudes, such as "change of base discharge", which is a level above which peak flows are recorded, or "change of gauge datum", which is the arbitrary zero gauge height for the rating curve. Figure 1c shows the location, flag type, and the number of the sites under each flag. Almost half of the sites have no flag and these are located throughout the domain. A few sites have more than one type of flag and we show the flag associated with a higher likelihood of the flows being affected (e.g. regulated). The majority of regulated or

Deleted: ~~Changes in the rating curve used to estimate streamflow from measured water levels are not recorded in the USGS notes but may be a significant source of variation in low flow values that is not accounted for.~~

partially regulated sites are concentrated in the northeast, but this is also where the majority of all sites are located. The sites in the mid-Atlantic states are generally more affected by urbanization or have experienced a change of gauge datum. Overall, 198 sites out of 508 sites are flagged as affected in terms of anthropogenic influences. In the results section, we show how the results of our statistical methods compare with the USGS site flags that are related to regulation or some other human influence.

Deleted: 271

Formatted: Highlight

Deleted: or changes in measurement method

2.3. Low Flow Indices

We analyze four variants of low flows based on different time scales, to understand how non-stationarity is dependent on the time scale as the data become smoother, with implications for the detection of non-stationarity. The 1-day minimum low flow, Q_1, is the annual minimum daily streamflow. The other three variants, Q_7, Q_{30}, Q_{90}, are obtained by applying the same analysis to 7-day, 30-day, and 90-day moving average versions of the time series. Together, we refer to the four low flow variables as the *n*-day minimum flows. Q_7 (dry weather flow) is the most widely used low flow statistic in the U.S. (Kroll et al., 2004; Smakhtin, 2001), but the others are important for different applications, such as Q_1 for ecological assessments and Q_{90} for reservoir operations. We also calculate the day of the year of low flows and use this to identify the primary (and in some regions the secondary) low flow season, as well as any long-term changes in timing. The primary season is defined as the 4-month period that contains the majority of the low flow occurrences, and the secondary season as the 4-month period that contains the majority of the remaining low flows. If the onset time of the low flow season for a site occurs 70% to 100% in a specific month, that site is assumed to have only one low flow season. The sites that have low flow events occurring 40-70% of the time in one month and 20-

Formatted: Font:Not Italic

Formatted: Font:Not Italic

40% of the time in a different month are characterized as having two low flows seasons. The timing results are shown based on Q7 and Q30 flows.

2.4 Identification of stationary time series

A sequence of realizations of random variables, Y, is stationary if the distribution of the sequence is independent of the choice of starting point (Kendall et al., 1983; Ruppert, 2011). Determining stationarity of a time series is not straightforward (Lins and Cohen, 2011) and in practice, it is common to look at restricted measures of stationarity. A time series is defined as weakly stationary if it satisfies three criteria:

$$\mathrm{E}(Y_i) = \mu, \quad (\forall i) \tag{1a}$$

$$Var(Y_i) = \sigma^2, \quad (\forall i) \tag{1b}$$

$$Corr(Y_i, Y_j) = \rho(|i - j|), \quad (\forall i, \forall j) \tag{1c}$$

where μ is the sample mean, σ is the standard deviation and ρ is the correlation, with i representing one realization of a time series. This means that for a weakly stationary variable, the mean and variance do not change with time and the correlation between two values depends only on the lag (the time between values). Visual inspection of the time series and the changes therein can provide an indication in the attempt to assess stationarity, in that a change in the underlying process leads to changes in values that are obvious (Lins and Cohen, 2011; Koutsoyiannis, 2011; Serinaldi and Kilsby, 2015).

Deleted: can be very helpful in determining

We apply three tests to identify weak stationarity: (1) the Mann-Kendall test (Mann, 1945; Kendall, 1975), which tests for increasing or decreasing trends; (2) the Pettitt test (Pettitt, 1979), which tests for abrupt changes or change points; and (3) the Ljung-Box test (Ljung and

Box, 1978), which tests for autocorrelation. An identified change in the mean by either of the first two tests would rule out stationarity, except in the case of autocorrelated data, for which the Pettitt and Mann-Kendall tests will characterize too many sequences of the time series as having a step or trend and therefore increase the rejection rate of the null hypothesis of no change (Douglas et al., 2000; Serinaldi and Kilsby, 2015). Therefore, analysis of autocorrelation is carried out before conducting the Mann-Kendall and Pettitt tests. Even when a site is identified as non-stationary, further analysis is required to understand the overall regime of the data at such a site. For example, the time series may have two separate stationary regimes with one change point in between or an overall trend. We then assume that the change year corresponds to human intervention, which is generally borne out by investigating the site notes.

2.5. Decomposition algorithm

The three statistical tests (Ljung-Box, Pettitt and Mann-Kendall) were combined into a recursive algorithm to identify non-stationarity in the low flow time series and decompose the series into potentially stationary sub-series. In the first step of the algorithm, a Ljung-Box test with 20 lags was applied to the entire time series of each site, and sites with significant overall autocorrelation (5% significance level) were identified. The Ljung-Box test identifies sites that are non-stationary and is able to identify sites with abrupt changes because the series of values before the change appear to be autocorrelated relative to the values after the change, and vice-versa. This was confirmed by visual inspection of the time series. For the sites with significant overall autocorrelation, we then applied the Pettitt test (5% significance level) to confirm the existence of any step change and identify its timing. The series were pre-whitened to remove lag-1 autocorrelation using the trend-free pre-whitening method of Yue et al. (2002) and

Deleted: using the trend-free pre-whitening method of Yue et al. (2002) and implemented by

implemented by Kumar et al. (2009). It is necessary to identify sites with potential step changes using the Ljung-Box test first because the Pettitt test will identify step changes in time series with gradual trends. Similarly the MK test will identify gradual trends in series with step changes. If a significant change is found by the Pettitt test, the series is split into two parts either side of the step change. Each part is assumed to be a new series at the same location, and if it has a record length of 30 years or more, the decomposition algorithm is applied again. If the length is less than 30 years, the site is removed from further consideration. If a statistically significant step change is not identified, we note that the series is autocorrelated overall. We then applied the Mann-Kendall (MK) test (5% significance level) on the remaining sites to identify statistically significant trends in the data. Again, the series were pre-whitened to remove lag-1 autocorrelation. The series and sub-series are assigned categories as follows:

1. Category 1: Non-autocorrelated site with no trend (MK=0);
2. Category 2: Non-autocorrelated site with a statistically significant decreasing trend (MK=-1);
3. Category 3: Non-autocorrelated site with a statistically significant increasing trend (MK=1);
4. Category 4: Autocorrelated site with statistically significant step change, time series split and the sub-series re-categorized recursively;
5. Category 5: Autocorrelated site with no step change.

3. Stationarity Results

3.1. Categorization of sites

Figure 2 shows the spatial distribution and the number of sites in each category after the first recursive level of the decomposition algorithm. The results for all *n*-day low flow metrics are presented for the available length of record at each site, which ranges between 1891 and 2011. No site has a record length less than 50 years and no site has any gap in the *n*-day low flow series. As we move from Q_1 to Q_{90}, a larger number of sites appear stationary (category 1) and the number of sites identified using the Pettitt test as having an abrupt shift in the time series (category 4) decreases. The algorithm re-applies the Pettitt test to category 4 sites to identify useable sub-series. For example, the Q_1 time series of 155 sites are split into two parts, which are subjected to further categorization.

Figure 3 summarizes the time periods that were identified as useable at each step of the recursive algorithm for all sites for Q_1. The light blue lines represent the original record length for each site. The vertical axis shows the site number from 1 to 508 ordered from the lowest to highest latitude. Therefore, site 1 is the most southerly and site 508 is the most northerly. The left panel of Figure 3 shows the record length of sites, which, in the first step of categorization, had no significant autocorrelation. These sites are colored according to their MK trend value: 0 (no significant trend), -1 (significant negative trend), or 1 (significant positive trend). The middle panel again shows the original record length for each site in light blue, but highlights the sites that were identified with an abrupt step change by the Pettitt test and were split into two parts. For each part that exhibits no autocorrelation, the trend values were calculated. The right panel shows the parts of the time series that were recovered in the next step of the decomposition algorithm. As long as the record length is greater than or equal to 30 years the algorithm is applied recursively on the remaining parts of the time series. The number of sites shown in the right panel is small but their data are still useful for subsequent analysis.

3.2. Comparison with USGS flags

Table 1 shows the breakdown of the number of sites in each category and the relation to USGS flags for Q_7 and Q_{30}, and indicates that in every category, anthropogenic influences are documented by the USGS. For Q_7, the majority of sites in categories 4 (57%; step change) are flagged by the USGS as somehow affected. This suggests that the algorithm has some skill in identifying managed or altered flow series. However, there are also many sites in category 1 (36%; no trend), 2 (16%; decreasing trend) and 3 (42%; increasing trend) that are also flagged (see Figure 4) suggesting that anthropogenic impacts for these sites are minimal and/or are overwhelmed by any climate or land use induced changes. The fact that the majority of stationary sites (category 1) are not flagged is encouraging. Figure 4 shows all the sites from each of the 5 categories that have no anthropogenic flag for Q_7: 310 out of 508 sites are not flagged but only 153 of these 310 sites show absolute stationarity behavior (category 1) and the rest exhibit some form of non-stationary.

From Table 1 we observe that:

1. If a site is flagged and its low flow series has a trend, the flags are mostly for regulation of partial regulation; sites with increasing trends are more likely to be flagged as regulated.
2. If a site is flagged and it exhibits a step change, the flag is mostly associated with regulation, or possibly urbanization;
3. If a site is in category 5 (not considered further due to significant autocorrelation), it may be flagged as regulated;

Deleted: or measurement changes

Deleted: 64

Deleted:), and 5 (58%; significant autocorrelation)

Deleted: 50

Deleted: 35

Deleted: 60

Deleted: or changes in measurement characteristics

Deleted: 237

Deleted: 119

Deleted: 237

Deleted: decreasing

Deleted: a change of gauge datum

Deleted: <#>If a site is flagged and its low flow series has an increasing trend, the flags are mostly related to regulation or a change of gauge datum;

Formatted: Bullets and Numbering

Deleted: , a change of gauge datum

Deleted: or its gauge datum has changed

4. If a site shows no trend but is still flagged, the flag relates to regulation. This suggests that the impact of the flagged change was either minimal or good management practices have been put in place. The majority of these sites are located in the upper Mid-Atlantic in the states of New York, New Jersey, and Virginia.

Deleted: or a gauge datum change

We also applied the algorithm to the HCDN-2009 sites within the domain, to confirm that the algorithm can identify sites that have been independently determined as unaffected by human influences. We found that 82% and 86% of these sites were placed in category 1 (stationary) for Q7 and Q30, respectively, with most of the remaining sites in category 3 (increasing trend; 9% and 8%) or category 6 (autocorrelated; 5% and 4%).

Deleted: Dam failure, flow below the rating curve limit, or change of base discharge and even urbanization do not appear to have a significant impact on low flows or their measurement for the sites considered.

3.3. Variability in year of abrupt change

For sites that were identified by the Pettitt test as having an abrupt change, Figure 5a shows the variability of the year of change for Q_n. Most of the changes occurred between 1962 and 1986, and as discussed above, most of these are flagged as having regulation. The spatial distribution of changes indicates that stream regulation began in the northeast before spreading to the southeast. The Pettit test tends to identify significant changes away from the either ends of the time series, and so may not identify changes in the earlier or later part of the record. However, earlier or later step changes are identified in the second recursion of the decomposition algorithm.

Deleted: or a change of gauge datum

We further examined the consistency of the change year among the Q_n series, with the expectation that abrupt changes would be identified for the same year across all or most Q_n time series. Figure 5b shows the spatial distribution and the number of sites with a consistent year of change among the Q_n. Out of 176 sites whose time series were identified as having a step change by the Pettitt test, 82 (almost half) showed the same change year for 3 out of 4 Q_n series. Only 7 sites showed the same change year for all Q_n. Although we have identified the change year for all Q_n, the results for Q_7 may be the most appropriate for identifying a change since the data are close to the original values, but are less affected by measurement errors than Q_1 (WMO, 2008).

4. Variability and Trends in Low Flows and Timing

4.1. Trends in low flows

We identified a time period (1951-2005) common to all sites for which they have useable data, and calculated statistics of Q_n, including the trend, and the consistency of trends among Q_n values. The MK trends for Q_n for the sites that were categorized as 1, 2, or 3 by the decomposition algorithm are shown in Figure 6a. The sites with significant trends tend to occur in all Q_n (e.g. the sites in Florida). Sites with lower trend magnitudes tend to become non-significant (MK=0) as we move from Q_1 to Q_{90} (e.g. the two sites in the northeast in Maine). Some sites to the east of the Mississippi River do not have significant trends for Q_1 but show a significant decreasing trend for Q_{90}. Overall, the northeastern sites show increasing trends in low flows and the southeast sites show decreasing trends.

A summary of the consistency of trends across *n*-day low flows is shown in Figure 6b. 208 sites (41% of the sites) have the same trend, such that the Q_n series are all increasing, decreasing, or not changing. 162 sites (32%) agree on the sign of trend for three out of four of the Q_n trends, and 87 sites (17%) agree for 2 out of 4 of the Q_n trends. Overall, the consistency in trends among the Q_n series is generally uniformly distributed across the domain.

Figure 7 (top left) shows the spatial pattern of the MK trend test values for Q_7 for all sites (without testing for step changes or autocorrelation), and when we only consider sites without step changes (top right). In both cases, the pattern of increasing trend in low flows in the northeast and a decreasing trend in the southeast is apparent. However, ignoring the effect of autocorrelation may give rise to misleading results by showing a denser pattern of significant trends. The bottom left panel shows the results removing sites with step changes and pre-whitening the data for the remaining sites. The bottom right panel show the trends when sites that have USGS flags are also excluded, e.g. for sites without documented anthropogenic impacts. The drivers of trends at these sites are therefore likely related to climate variability/change and/or land use change, rather than management of, or influence on, flows.

Deleted: attribution

Deleted: is

4.2. Variability in low flow timing

Figure 8 summarizes the distribution of the onset of the low flow season for Q_7, for the primary season (top panels) and the second season (bottom panels). The left panels show the onset month of the season and the right panels show the probability of the onset season in that month. If the onset time of the low flow season for a site occurs 70% to 100% in a specific month, that site is assumed to have only one low flow season. For Q_7, 353 sites out of 395 (almost 90%) sites have a single low flow season, and the onset of the season changes from north to south. Most of the sites north of North Carolina have low flow seasons starting in July, which is generally driven by the slight decline in precipitation during the autumn as well as the increased evaporation during the summer (Small et al., 2006). In Florida the season starts in April-May. For coastal sites, the season starts earlier (mostly in June), and for sites in the southwestern part of the domain, the season starts mostly in September-October.

Deleted: ,

Deleted: where the first season is defined as the 4-month period that contains the majority of low flow occurrences

Deleted: as the 4-month period that contains the majority of the remaining low flows

Deleted: ,

Deleted: fit in this category. For sites with one low flow season, the onset

Formatted: Font:Not Italic

The sites with two low flows seasons are mostly in Florida, and along the coastline of Georgia, South and North Carolina, New York, New Jersey, and Maine and their second season occurs mostly in fall. For New York, New Jersey, and some sites along the west coastline of Florida, the second low flow season mostly starts in November and December. Sites near the Gulf of Mexico and some sites in North Carolina have second low flow seasons starting in April. The second low flow season for the far northeast sites begins in December or January and can be related to freezing conditions that may store water as snow and river ice.

Deleted: The sites that have low flow events occurring 40-70% of the time in one month and 20-40% of the time in a different month are characterized as having two low flows seasons. These sites

4.3. Changes in low flow timing

To determine whether low flow timing has changed over time, we examined sites with one low flow season as defined as 70% of low flow occurrences in the same season, again for the common time period of 1951-2005. Analysis of changes in timing irrespective of the season (not shown) did not show evidence of shifts in timing from one season to another. For Q_7, for example, 47 sites out of the total 508 were removed because their low flow season occurs less than 70% of the time in one season. Out of the remaining 467 sites, 20 sites showed a decreasing (earlier) trend in timing and were mostly in Pennsylvania and the Carolinas (Figure 9) and 14 showed an increasing (later) trend with most of these in the northeast. The MK test for Q_{30} timings showed mainly decreasing (earlier) trends (26 sites), with most overlap with the Q_7 results in Pennsylvania. These sites have low flow seasons starting in July, and half of them are regulated or partially regulated. Only a few sites were identified by the Pettitt test (5% significance) to have a significant step change in either direction.

The tendency for low flows (Q_7 and Q_{30}) to occur earlier in the season in recent years may be because of a shift of low precipitation from the late to mid summer, but given the small number of sites with significant trends and their low spatial coherence, this is speculative. Although the sites in Pennsylvania did not show a trend in low flow volumes, the overall trend for the northeast is an increasing trend in low flow volumes suggesting that early summer low precipitation might also be increasing. More investigation is required to confirm whether low precipitation is happening earlier in summer, for example during May and June, and whether the amount is increasing.

5. Discussion and Conclusions

5.1. Potential Drivers of Trends in Low Flows

We found spatially coherent patterns of increases in low flows in the northeast and decreases in the southeast, which was robust to the presence of USGS flags and autocorrelation in the time series, despite the smaller number of sites. The pattern of increasing low flows in the northeast is consistent with regional scale studies (e.g. Hodgkins and Dudley, 2011) and are consistent with the increases in 7-day low flows and fall precipitation shown in Small et al. (2006) that focused on a smaller set of sites across the eastern U.S. from the HCDN. Several other studies (e.g. Douglas et al., 2000; McCabe and Wolock, 2002; Hayhoe et al., 2007; Andreadis and Lettenmaier, 2006) have identified an overall increasing trend in precipitation over the past 50 years, and a decreasing pattern in soil moisture drought over the much of the U.S. including the northeast (Andreadis and Lettenmaier, 2006). Therefore, an increase in low flow volumes in the northeast is consistent with the overall shift to wetter conditions. The generally decreasing trends in the southeast are also consistent with the results from Small et al. (2006) and Lins and Slack (1999), which is despite an overall increase in precipitation in the region.

To understand the potential drivers of these trends more comprehensively, Figure 10 shows the Q_7 trend magnitude and the antecedent precipitation for the previous 180 days. This period was chosen as it provides the highest correlation with low flow volumes (Kam et al, 2015), although the results with 150 and 90 days are similar. The precipitation data are taken from the long-term precipitation dataset of Livneh et al. (2013) and are averaged over the basin corresponding to each site. The similarity between the trends in low flows and antecedent precipitation is striking with a clear increasing trend in the north and decrease in the south, although many of the trends are not statistically significant.

Formatted: Font:Not Italic

Deleted: attribution

The main disparity is in coastal plains of eastern Virginia, Maryland and northwards to Maine, where Q_7 low flows have decreased but antecedent precipitation is increasing (both often statistically significant). The reason for this is unclear, but groundwater is likely playing a role across the coastal plain aquifer of the mid-Atlantic states and up into New England (Dudley and Hodgkins, 2013) either via changes in recharge or indirectly through anthropogenic impacts. Groundwater pumping has reduced levels in the north Atlantic Coastal Plain aquifer system by tens of meters (e.g. Konikow, 2013, USGS, 2006) and has likely reduced discharge to streams in the northeast (e.g. Pucci and Pope, 1995; Brutsaert, 2010; Barlow and Leake, 2012). Similarly, overuse of groundwater resources in the southeast (Konikow, 2013) may be contributing to decreases in low flows across the region (e.g. Bosch et al., 2003; Opsahl et al., 2007; Brutsaert, 2010).

Increases in evaporation (Walter et al., 2004; Nolan et al., 2007; Huntington and Billmire, 2014) may have also led to declines in groundwater recharge and streamflow (Hodgkins and Dudley, 2011), and potentially cancelled out the overall increases in precipitation across much of the U.S. (Andreadis and Lettenmaier, 2006). Figure 10 also shows an estimate of the trend in late summer/early fall potential evaporation based on the NLDAS2 dataset of Xia et al. (2012). Potential evaporation has increased over the eastern U.S. with statistically significant trends over much of the mid-Atlantic states and the southeast. This suggests that increasing atmospheric demand in the southeast may have exacerbated declines in low flows, and this may have offset increasing precipitation somewhat in the northeast. Changes in land use may also explain trends in both regions, whereby land abandonment in the northeast and forest harvesting and urban development in the southeast may have contributed to the respective trends in each region (Cho et al., 2009; Payne et al., 2005; USGS, 2012), although attribution is difficult.

The analysis of trends in timing of low flows showed one cluster of sites with a trend to earlier timing. These sites are mostly in central and west Pennsylvania, and central southern New York. The reasons for the changes are unclear, but may be related to regulation and possibly a shift in the low precipitation season to earlier in the summer. The timing of low flows in the other parts of the domain has not changed based on a 5% significance level.

5.2. Conclusions

This study has examined the presence of non-stationarity in low flows across the eastern U.S. in terms of volumes and timing. We focused on the full period of available data at each site to identify abrupt shifts that may be associated with management, in particular dam construction, and gradual trends that may be an impact of climate change, land use change or surface/ground water withdrawals. A decomposition algorithm was used to identify useable sub-series of the data that could then be further analyzed for trends. Comparison with USGS site flags indicates that the majority of sites with identified step changes and increasing trends are noted to be regulated in some way, and some are documented as having undergone urbanization. For sites

Deleted: a change of stream gauge datum or with decreasing and increasing trends, about one sixth and one half, respectively, have USGS flags and these are almost all for regulation. Furthermore, about one third of sites with no trend are also flagged as being regulated or partially regulated, Our approach is therefore generally capable of identifying sites with documented regulation, and confirmed by the evaluation of the HCDN-2009 sites, but that changes do not always manifest in a detectable change in the low flow time series. This may be because the documented regulation or other change may not have an impact or that the signal is small compared to the variability in the time series. This is particularly the case for higher low flow metrics such as Q_{90}, for which the regulation is generally less detectable. For sites with documented regulation but no detectable signal, the fact that the USGS flags relate to high flows rather than low flows may help explain this, or that the sites are well managed in terms of low flows. For example, flows are often artificially elevated above the natural levels of low flow to create ``anti-droughts" to manage the restoration of river systems (Bunn et al., 2006). Although we do not claim to make a definitive judgement on whether

Several outstanding questions remain, most importantly what are the low flow generating mechanisms across the eastern U.S. and what are the drivers of long-term changes in the volumes and timing. Potential mechanisms include, but are not limited to: changes in antecedent precipitation and teleconnections with large-scale climate (e.g. the North Atlantic Oscillation; Kam et al., 2015), land use change, surface and groundwater abstraction, and streamflow regulation. The results of this study suggest that low flow variability in the eastern U.S. is driven by a mixture of climatic and anthropogenic effects, with suggestions that changes in climate have played a role in both the northeast and southeast. However, definitive attribution will require detailed analysis of these competing factors and possibly carefully crafted modeling studies. In

Deleted: third

Deleted: tend

Deleted: to be

Deleted: a change in gauge datum height

Deleted: ;

Deleted: a similar proportion

Deleted: for a change in gauge datum

Deleted: , and so it is unclear whether this type of change has influenced the low flow time series

Deleted: This

Deleted: implies that o

Deleted: or a change in measurement method

Deleted: these

Deleted: or a change in measurement characteristics parallel with calls for more rigorous efforts at attributing changes in flood time series (Merz et al., 2012), increased effort is also needed for understanding and attributing changes in low flows. Several new approaches have been put forward recently that show promise for detecting and attributing changes in hydrological time series, including extremes, based on multiple working hypotheses (Harrigan et al., 2014) and complex statistical modeling (Prosdocimi et al., 2015).

The results of this study can help in understanding changes in low flows across the eastern U.S., and the impact of anthropogenic and natural changes. It can therefore provide information for water management, and restoration of stream flows and aquatic habitats. Although we do not claim to make a definitive judgment on whether low flows at a particular site are influenced by human activities or are completely free of influences because of the complexities of low flow generation, our approach shows promise for systematically identifying sites for further investigation, especially where supporting information (such as site notes) are available to support the statistical results. Our approach may be especially useful for exploring large-scale, climate-driven changes in the low flow regime where pooling of results across sites increases confidence in the robustness of any identified changes. The methods are readily transferable to other parts of the U.S. and globally, given long enough time series of daily streamflow data, although further work is required to understand their universal application.

Author Contribution

S. S. and J. S. conceived the study. S. S. performed the analysis with help from J. K. S. S. prepared the manuscript with contributions from the other authors.

Acknowledgements

This work was supported by the USGS (G11AP20215) and NOAA (NA14OAR4310130 and NA14OAR4310218)

References

Acreman, M. C., B. Adams, and B. Connorton, 2000: Does groundwater abstraction cause degradation of rivers and wetlands? *Water and Environment Journal*, 14, 200–206.

Andreadis, K. M., and D. P. Lettenmaier, 2006: Trends in 20th century drought over the continental united state. *Geophys. Res. Letts.*, 33, 1-4.

Averyt, K., J. Meldrum, P. Caldwell, G. Sun, S. McNulty, A. Huber-Lee, and N. Madden, 2013: Sectoral contributions to surface water stress in the coterminous United States. *Environ. Res. Lett.*, 8, 035046 (9pp). doi:10.1088/1748-9326/8/3/035046

Barlow, P. M., and S. A. Leake, 2012: Streamflow depletion by wells—Understanding and managing the effects of groundwater pumping on streamflow. *U.S. Geological Survey Circular*, 1376, 84p.

Bosch, D. D., Lowrance, R. R., Sheridan, J. M. and Williams, R. G., 2003: Ground water storage effect on streamflow for a Southeastern Coastal Plain watershed. *Ground Water*, 41, 903–912. doi: 10.1111/j.1745-6584.2003.tb02433.x

Bradford, M., and J. Heinonen, 2008: Low flows, instream flow needs and fish ecology in small streams. *Canadian Water Resources Journal*, 33, 165–180.

Brandes, D., G. J. Cavallo, and M. L. Nilson, 2005: Base flow trends in urbanizing watersheds of the delaware river basin. *Journal of the American Water Resources Association*, 41: 1377–1391. doi: 10.1111/j.1752-1688.2005.tb03806.x

Brown, T. C., R. Foti, and J. A. Ramirez, 2013: Projected freshwater withdrawals in the United States under a changing climate. *Water Resour. Res.*, 49, 1259–1276, doi:10.1002/wrcr.20076.

Brutsaert, W., 2010: Annual drought flow and groundwater storage trends in the eastern half of the United States during the past two-third century. *Theoretical and Applied Climatology,* 100 (1-2), 93-103.

Bunn, S. E., M. C. Thoms, S. K. Hamilton, and S. J. Capon, 2006: Flow variability in dryland rivers: boom, bust and the bits in between. *River Research and Applications*, 22 (2), 179186.

Cho, J., V. A. Barone, and S. Mostaghimi, 2009: Simulation of land use impacts on groundwater levels and streamflow in a Virginia watershed. *Agricultural Water Management*, 96 (1), 1-11

Cohn, T. A., and H. F. Lins, 2005: Nature's style: Naturally trendy. *Geophys. Res. Lett.*, 32, L23402, doi:10.1029/2005GL024476

Colby, F. P., 2008: Mesoscale snow bands in an ocean-effect snowstorm. *Tech. rep., University of Massachusetts Lowell*, Lowell, MA 01854.

Cox, D. R., and A. Stuart, 1955: Some quick sign tests for trend in location and dispersion. *Biometrika*, 42(1/2), 80–95.

Daniel, C., and P. Dahlen, 2002: Preliminary hydrogeologic assessment and study plan for a regional ground-water resource investigation of the Blue Ridge and Piedmont Provinces of North Carolina, Investigations Report 02-4105, USGS.

Deitch, M. J., G. M. Kondolf, and A. M. Merenlender, 2009: Hydrologic impacts of small-scale instream diversions for frost and heat protection in the California wine country. *River Research and Applications,* 25, 118-134.

Douglas, E. M., R. M. Vogel, and C. N. Kroll, 2000: Trends in floods and low flows in the United States: impact of spatial correlation. *J. Hydrology*, 240, 90–105.

Dudley, R.W., and G. A. Hodgkins, 2013: Historical groundwater trends in northern New England and relations with streamflow and climatic variables. *J. American Water Resources Association*, 49 (5), 1198-1212.

EPA Climate Change Division, 2008: Precipitation and storm changes, url: www.epa.gov/climatechange/science/indicators/weather-climate/index.html.

Giuntoli, I., B. Renard, J.-P. Vidal, and A. Bard, 2013: Low flows in France and their relationship to large-scale climate indices. *J. Hydrology*, 482, 105–118.

Gurka, J. J., E. P. Auciello, A. F. Gigi, J. S. Waldstreicher, K. K. Keeter, S. Businger, and L. G. Lee, 1995: Winter weather forecasting throughout the eastern United States. Part II: An operational perspective of cyclogenesis. *American Meteo.*, 10, 21–41.

Harrigan, S., C. Murphy, J. Hall, R. L. Wilby, and J. Sweeney, 2014: Attribution of detected changes in streamflow using multiple working hypotheses, *Hydrol. Earth Syst. Sci.*, 18, 1935-1952, doi:10.5194/hess-18-1935-2014, 2014.

Hayhoe, K., C. P. Wake, T. G. Huntington, L. Luo, M. D. Schwartz, J. Sheffield, E. Wood, B. Anderson, J. Bradbury, A. Degaetano, T. J. Troy, and D. Wolfe, 2007: Past and future changes in climate and hydrological indicators in the U.S. Northeast. *Climate Dynamics*, 28, 381–407.

Hodgkins, G. A., R. W. Dudley, and T. G. Huntington, 2005: Changes in the number and timing of days of ice-affected flow on northern New England rivers. *Climatic Change*, 71 (3), 319-340. doi: 10.1007/s10584-005-5926-z

Hodgkins, G. A., and R. W. Dudley, 2011: Historical summer base flow and stormflow trends for New England rivers. *Water Resources Research*, 47 (7), W07528

Formatted: Font:Italic

Huntington, T. G., and M. Billmire, 2014: Trends in precipitation, runoff, and evapotranspiration for rivers draining to the Gulf of Maine in the United States. *J. Hydrometeor.,* 15(2), 726-743.

Kam, J., J. Sheffield, X. Yuan, and E. F. Wood, 2013: The influence of Atlantic tropical cyclones on drought over the Eastern U.S. (1980-2007). *J. Climate*, 26 (10), 3067-3086. http://dx.doi.org/10.1175/JCLI-D-12-00244.1.

Kam, J., S. Sadri, and J. Sheffield, 2014: Changes in the low flow regime over the Eastern United States (1962-2011): Variability, trends, and attributions, submitted to *Climatic Change*.

Karl, T. R., and R. W. Knight, 1998: Secular trends of precipitation amount, frequency, and intensity in the United States. *Bull. Am. Meteorol. Soc.*, 79, 231–241

Kendall, M. G., 1975: Rank Correlation Methods. Charles Griffin, 202 pp.

Kendall, M., A. Stuart, and J. K. Ord, 1983: The Advanced Theory of Statistics, Vol. 3, Design and Analysis, and Time Series, 4th ed., 780 pp., Oxford Univ. Press, New York.

Konikow, L. F., 2013: Groundwater depletion in the United States (1900–2008): U.S. Geological Survey Scientific Investigations Report 2013–5079, 63 p., http://pubs.usgs.gov/sir/2013/5079

Koutsoyiannis, D., 2011: Hurst-Kolmogorov dynamics and uncertainty. *J. American Water Resources Association*, 47 (3), 481-495

Kroll, C., J. Luz, B. Allen, and R. M. Vogel, 2004: Developing a watershed characterisitics database to improve low streamflow prediction. *J. Hydrologic Engineering*, 9 (2),116–125.

Kroll, C. N., and R. M. Vogel, 2002: Probability distribution of low streamflow series in the United States. *J. Hydrologic Engineering*, 7 (2), 137–146.

Kumar S,, V. Merwade, J. Kam, and K. Thurner, 2009: Streamflow trends in Indiana: effects of long term persistence, precipitation and subsurface drains, *J. Hydrol.*, **374**, 171–183.

Kustu, M. D., Y. Fan, and A, Robock, 2010: Large-scale water cycle perturbation due to irrigation pumping in the US High Plains: A synthesis of observed streamflow changes. *J. Hydrol.*, 390 (3-4), 222-244. doi:10.1016/j.jhydrol.2010.06.045

Lins, H. F. and T. A. Cohn, 2011: Stationarity: Wanted Dead or Alive?. *J. American Water Resources Association*, 47: 475-480. doi: 10.1111/j.1752-1688.2011.00542.x

Lins, H., and J. R. Slack, 1999: Streamow trends in the United States. *Geophys. Res. Letts.*, 26, 227–230.

Lins, H. F., 2012: USGS Hydro-Climatic Data Network 2009 (HCDN–2009): U.S. Geological Survey Fact Sheet 2012–3047, 4 p., available at http://pubs.usgs.gov/fs/2012/3047/.

Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier, 2013: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *J. Climate*, 26, 9384–9392.

Ljung, G. and G. Box, 1978: On a Measure of Lack of Fit in Time Series Models. *Biometrika*, 65, 297-303.

Mann, H. B., 1945: Nonparametric tests against trend. *Econometrica*, 13, 245–259.

Mauget, S. A., 2003: Multidecadal regime shifts in U.S. streamflow, precipitation, and temperature at the end of the twentieth century. *J. Climate*, 16, 3905–3916.

McCabe, G. J., and D. M. Wolock, 2002: A step increase in streamflow in the conterminous United States. *Geophys. Res. Letts.*, 29 (24), 1–4.

Merz, B., S. Vorogushyn, S. Uhlemann, J. Delgado, and Y. Hundecha, 2012. Hess opinions: More efforts and scientific rigour are needed to attribute trends in flood time series. Hydrol. Earth Syst. Sci. Discuss. 9, 1345–1365

Milly, P. C. D., J. Betancourt, M. Falkenmart, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer, 2008: Stationarity is dead: Whither water management? *Science*, 319 (5863), 573–574.

Opsahl, S. P., S. E. Chapal, D. W. Hicks, and C. K. Wheeler, 2007: Evaluation of ground-water and surface-water exchanges using streamflow difference analyses. *J. American Water Resources Association*, 43, 1132–1141. doi: 10.1111/j.1752-1688.2007.00093.x

Payne, D. F., M. A. Rumman, and J. S. Clarke, 2005: Simulation of ground-water flow in coastal Georgia and adjacent parts of South Carolina and Florida—Predevelopment, 1980, and 2000: U.S. Geological Survey, Scientific Investigations Report 2005-5089, 91p.

Pettitt, A. N., 1979: A non-parametric approach to the change-point problem. *Applied Statistics*, 28, 126–135.

Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegaard, B. D. Richter, R. E. Sparks, and J. C. Stromberg, 1997: The natural flow regime: a paradigm for river conservation and restoration. *Bioscience*, 47, 769–784.

Prosdocimi, I., T. R. Kjeldsen, and J. D. Miller, 2015: Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models, *Water Resour. Res.*, 51, 4244–4262, doi:10.1002/2015WR017065].

Formatted: Font:Times New Roman, 12 pt

Formatted: Font:Times New Roman, 12 pt

Formatted: Font:Times New Roman, 12 pt

Formatted: Font:Times New Roman, 12 pt

Formatted: Font:Italic

Pucci, A. A., and D. A. Pope, 1995: Simulated effects of development on regional ground-water/surface-water interactions in the northern Coastal Plain of New Jersey. *J. Hydrology*, 167 (1–4), 241–262.

Rolls, R. J., C. Leigh, and F. Sheldon, 2012: Mechanistic effects of low-flow hydrology on riverine ecosystems: ecological principles and consequences of alteration. *Freshwater Science*, 31(4), 1163–1186.

Ruppert, D., 2011: Statistics and data analysis for financial engineering, Springer Texts in Statistics, Springer.

Sadri, S., and D. H. Burn, 2011: A Fuzzy C-Means approach for regionalization using a bivariate homogeneity and discordancy approach. *J. Hydrology*, 401, 231–239.

Serinaldi, F., and C. G. Kilsby, 2015: Stationarity is undead: Uncertainty dominates the distribution of extremes. *Advances in Water Resources*, 77, 17-36

Smakhtin, V. U., 2001: Low flow hydrology: a review. *J. Hydrology*, 240, 147–186.

Small, D., S. Islam, and R. M. Vogel, 2006: Trends in precipitation and streamflow in the eastern U.S.: Paradox or preception?. *Geophys. Res. Letts.*, 33, 1–4.

Tallaksen, T., and H. A. Van Lanen, 2004: Hydrological drought, processes and estimation methods for streamflow and groundwater. Vol. 48, 580 pp., Elsevier Science.

Thomas, B., 2006. Trends in Streamflow of the San Pedro River, Southeastern Arizona, U.S. Geological Survey Fact Sheet 2006-3004, 4 pp., http://pubs.usgs.gov/fs/2006/3004/.

U.S. Army Corps of Engineers, 2012: url: http://geo.usace.army.mil/pgis/f?p=397:12

U.S. Senate, 2002: Federal Water Pollution Control Act, url: www.epw.senate.gov/water.pdf.

USDA Forest Service, 2010: Climate Change and Water Issues for the Eastern US, Tech. rep., USDA.

USGS: Sustainability of the Ground-Water Resources in the Atlantic Coastal Plain of Maryland, USGS Fact Sheet, 2006-3009, 2 p., 2006.

USGS, 2009: Groundwater atlas of the United States, url: www.pubs.usgs.gov/ha/ha730/.

USGS, 2012: Land cover change in the eastern United States, url: www.landcovertrends.usgs.gov/east/regionalSummary.html.

USGS, 2013: Groundwater depletion in the United States (1900-2008). Scientific Investigation Report 2013-5079, U.S. Department of the Interior, U.S. Geological Survey, Reston, Virginia.

USGS, 2014: National Water Information System data available on the World Wide Web (Water Data for the Nation), url: http://waterdata.usgs.gov/nwis/

Villarini, G., J. A. Smith, F. Serinaldi, J. Bales, P. D. Bates, and W. F. Krajewski, 2009: Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Advances in Water Resources*, 32, 1255–1266.

Walker, K. F. and M. C. Thoms, 1993: Environmental effects of flow regulation on the lower river Murray. *Australia. Regul. Rivers: Res. Mgmt.*, 8, 103-119. doi: 10.1002/rrr.3450080114

Walter, M. T., D. S. Wilks, J.-Y. Parlange, and R. L. Schneider, 2004: Increasing evapotranspiration from the conterminous United States. *J. Hydrometeor.*, 5, 405–408. doi: http://dx.doi.org/10.1175/1525-7541(2004)005<0405:IEFTCU>2.0.CO;2

WMO, 2008: Manual on low-flow estimation and prediction. Operational Hydrology Report 30, WMO-N 1029. Geneva, Switzerland

Yue, S., P. Pilon, B. Phinney, and G. Cavadias, 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes,* 16 (9), 1807–1829.

Table 1. Comparison of the number of streamflow gauging sites in each category of the decomposition algorithm and their USGS flags for Q_7. DamFail: dam failure; RegPar: partially regulated; Reg: regulated; Urban: affected by urbanization.

Category	Q_7	Q_{30}	Flag	Q_7	Q_{30}	Flag type	Q_7	Q_{30}
No Trend	240	260	Flagged	87	91	DamFail	1	2
						RegPar	33	37
						Reg	51	48
						Urban	2	4
			Not flagged	153	169			
Decreasing Trend	62	61	Flagged	10	6	DamFail	0	0
						RegPar	3	1
						Reg	5	5
						Urban	2	0
			Not Flagged	52	55			
Increasing Trend	55	70	Flagged	23	37	DamFail	0	0
						RegPar	8	13
						Reg	15	24
						Urban	0	0
			Not Flagged	32	33			
Step Change	112	89	Flagged	64	53	DamFail	1	0
						RegPar	21	16
						Reg	38	32
						Urban	4	5
			Not Flagged	48	36			
Autocorrelated	38	27	Flagged	13	10	DamFail	1	1
						RegPar	4	2
						Reg	7	7
						Urban	1	0
			Not Flagged	25	17			

Deleted: Below: flow below rating curve limit;

Deleted: ; ChangeDis: Change base discharge; ChangeDatH: change gauge datum

Deleted: 121

Deleted: 126

... [1]

... [2]

Deleted: 119

Deleted: 134

Deleted: 22

Deleted: 19

... [3]

Deleted: 0

... [4]

Deleted: 40

Deleted: 42

Deleted: 33

Deleted: 48

Deleted: 8

... [5]

... [6]

Deleted: 22

Deleted: 22

Deleted: 1

Deleted: 72

Deleted: 60

... [7]

... [8]

Deleted: 40

Deleted: 29

Deleted: 22

Deleted: 17

... [9]

... [10]

Deleted: 16

Deleted: 10

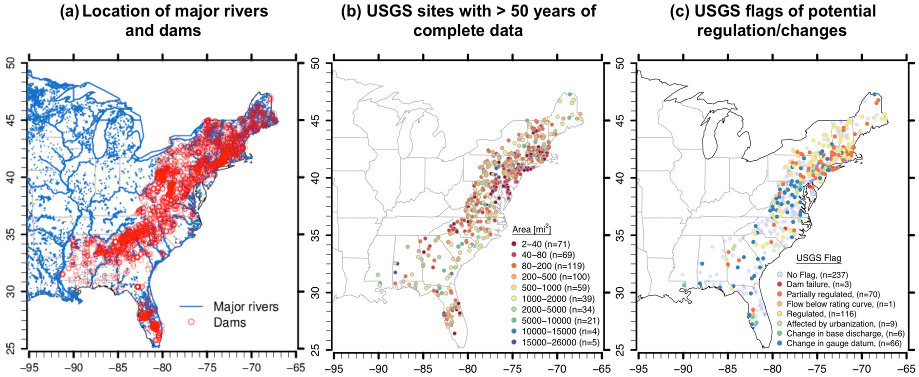


Figure 1. (a) Location of 2,352 major dams in the eastern U.S. (b) Location of the 508 streamflow sites with 50 years or more of complete daily data. (c) Flagged sites according to the USGS.

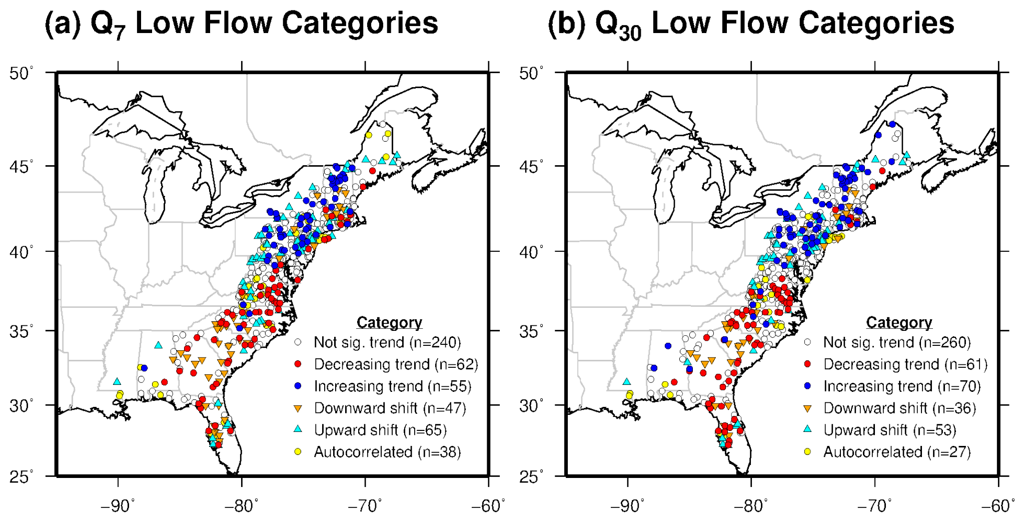


Figure 2. Categorization of non-stationarity of sites for Q_7 and Q_{30}.

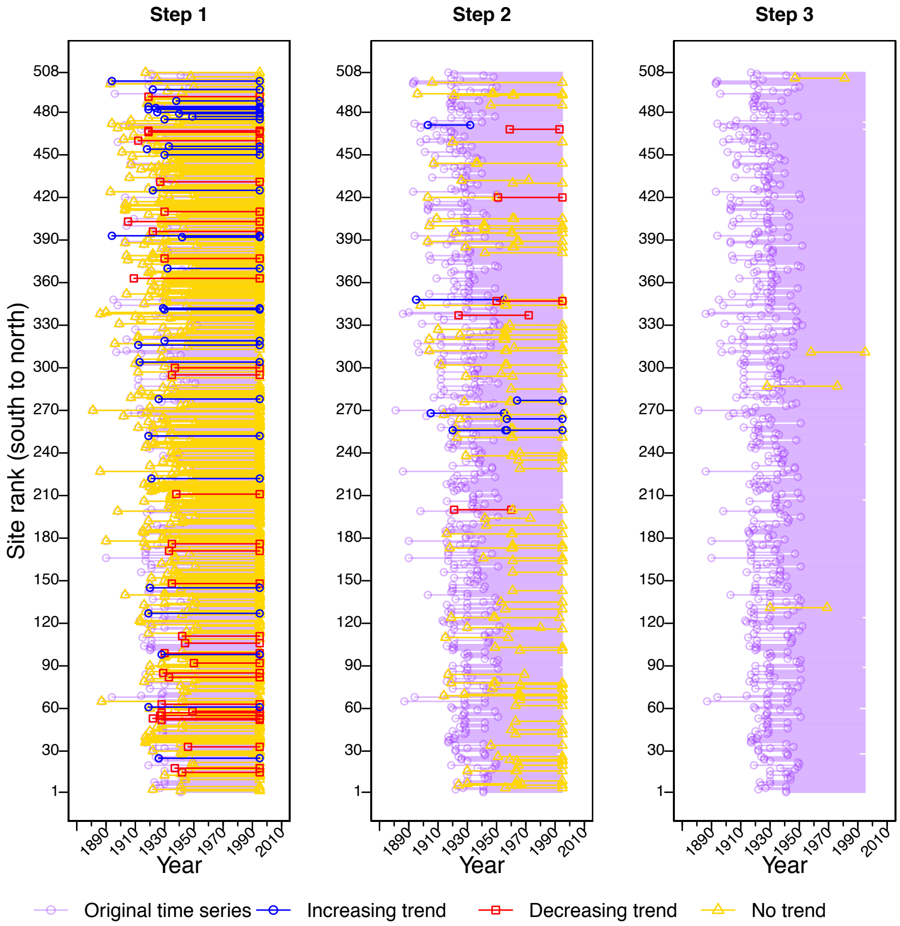


Figure 3. Range of years for each site that are stationary or show a trend, for each step of the decomposition algorithm.

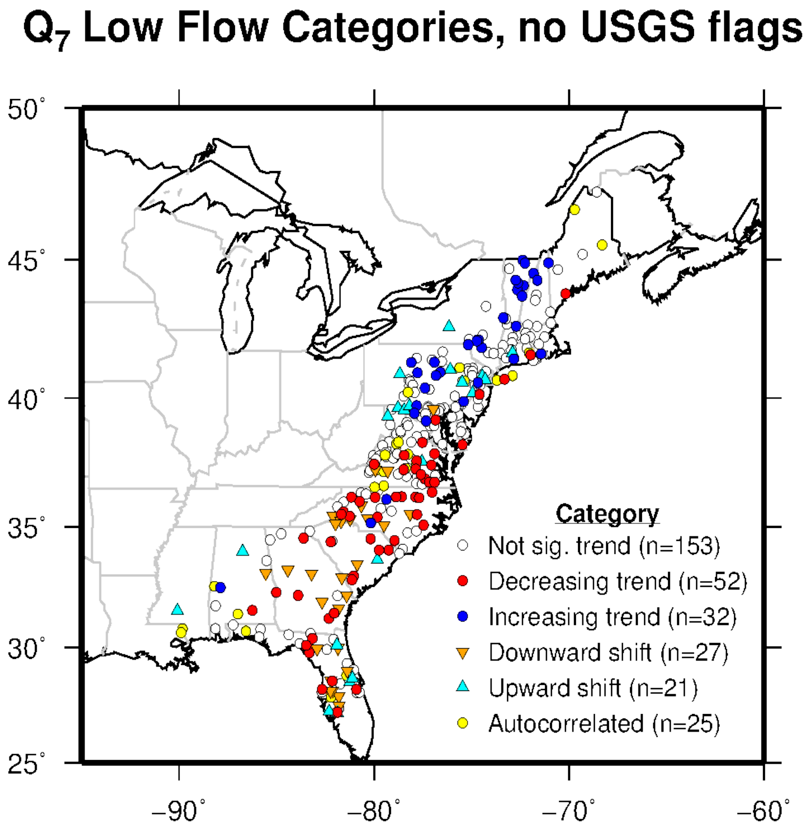


Figure 4. Categorization of non-stationarity of sites for Q_7 with no USGS flags from the first step of the decomposition algorithm.

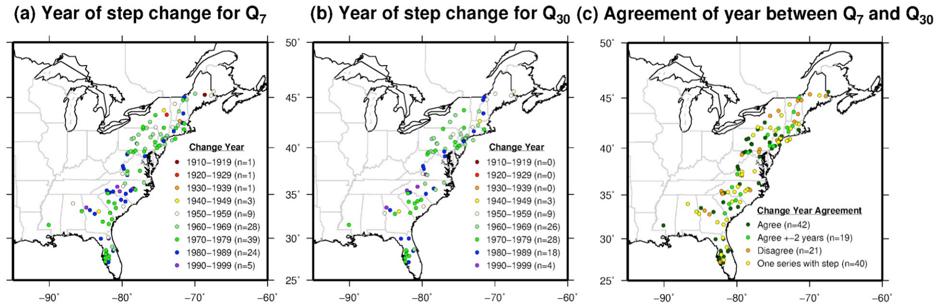


Figure 5. Year of step change for (a) Q_7 and (b) Q_{30}. (c) Agreement in year of step change between Q_7 *and* Q_{30} time series.

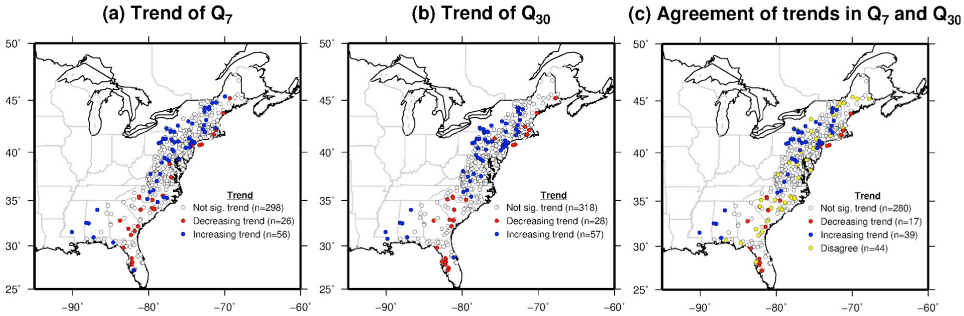


Figure 6. Trends in (a) Q_7 and (b) Q_{30} for 1951-2005 and (c) their agreement.

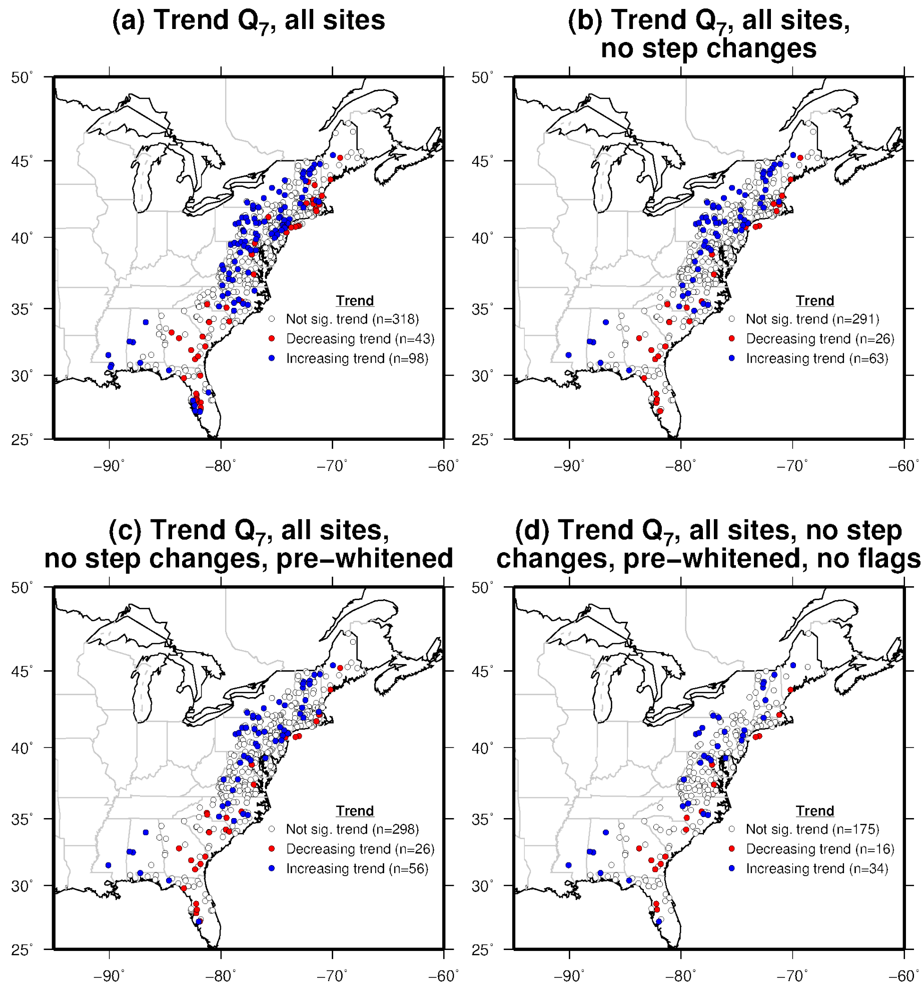


Figure 7. Trends in Q_1 for 1951-2005 for (a) all sites, (b) excluding sites with step changes or overall autocorrelation, (c) as (b) but with pre-whitened data, and (d) as (b) but without USGS flags.

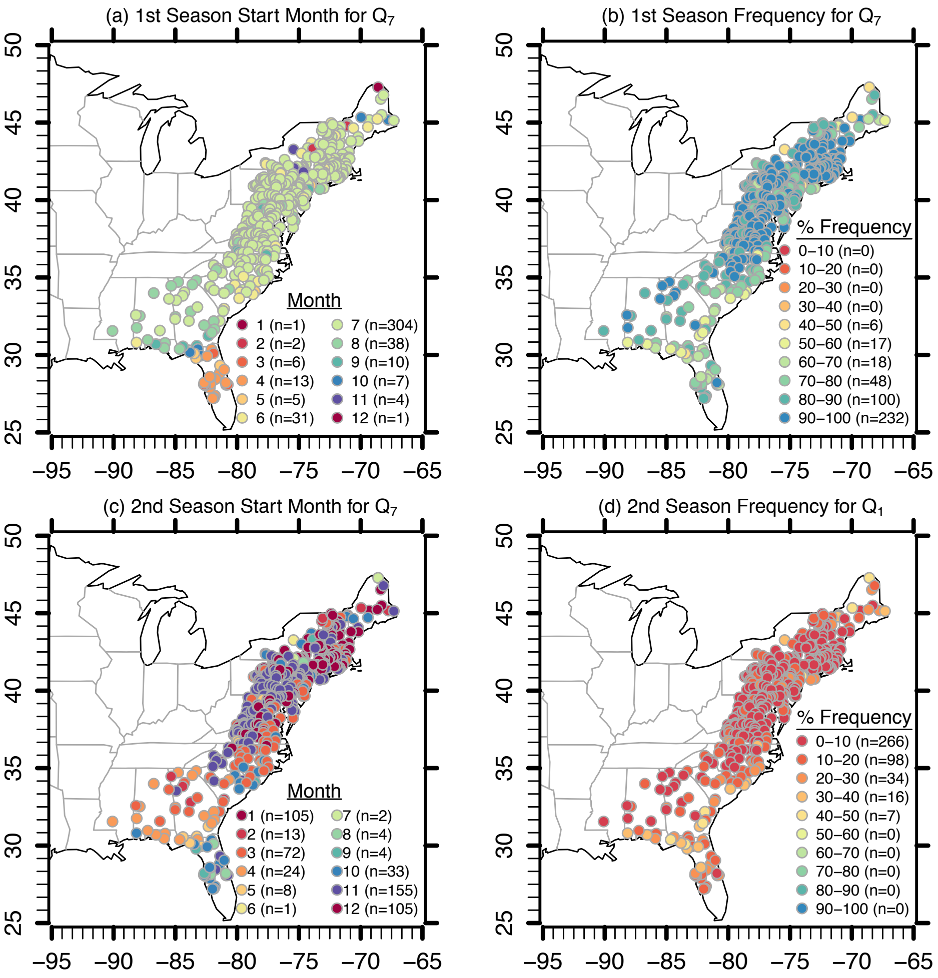
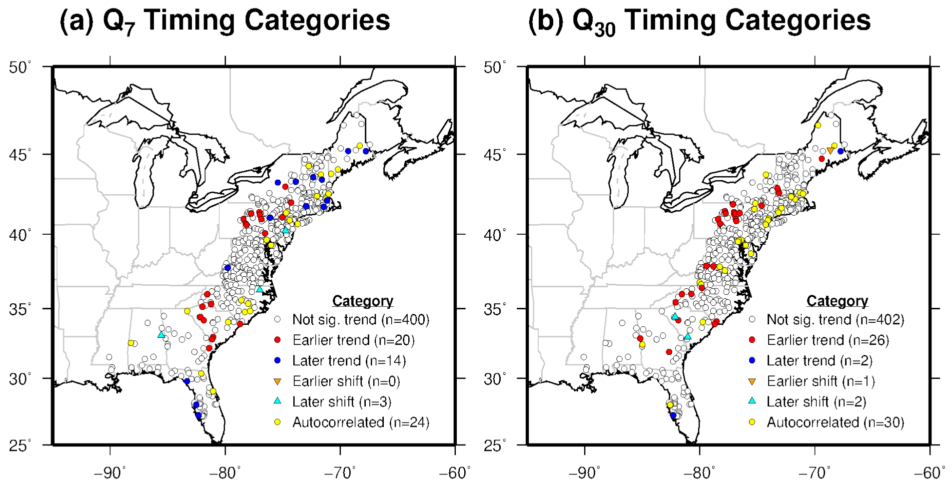


Figure 8. Primary and secondary seasons of occurrence of Q_7 low flows and their frequencies.



(a) Q_7 Timing Categories

(b) Q_{30} Timing Categories

Figure 9. Categorization of non-stationarity of sites for timing of (a) Q_7 and (b) Q_{30}.

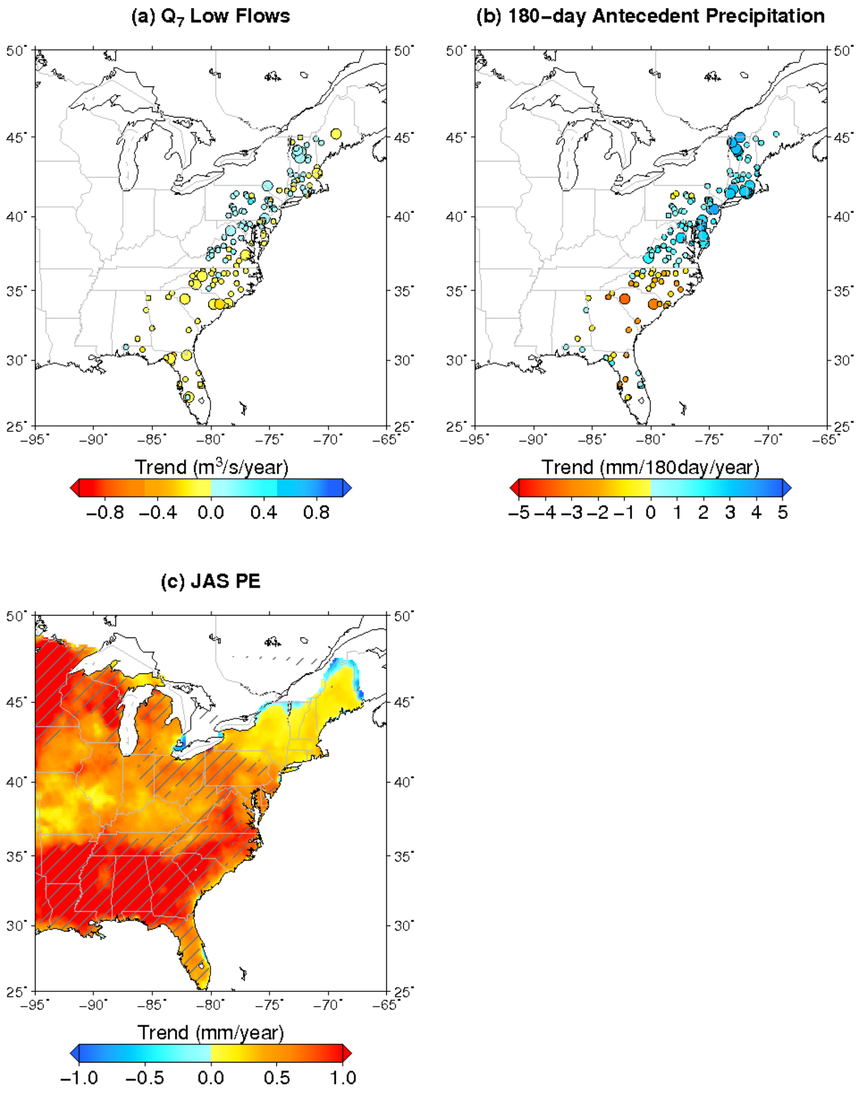


Figure 10. (a) Trend in *Q_7* low flows for 1951-2005 for the warm season. (b) Corresponding trend in 180-day antecedent precipitation. For (a) and (b), trends that are statistically significant at the 0.05 level are shown in large symbols. (c) Trend in July-August-September (JAS) potential evaporation for 1979-2012. Statistically significant trends are shown by hatching.