Reviewer Comments	Author's Response
General Comments	
Add a glossary or list of notation to define the many acronyms (CSS, CSO, NER, GLR, LTCP, VEMAP, NMC, IPCC, NCAR, EPA, GCM, WGEN, etc).	Agree: A list of abbreviations has been added to the document.
Double check the cited references….several of them are not listed in Section 5 Literature Cited (for example: page 5 line 36; page 6 line 33)	Agree: References have been checked and revised.
The procedure used to generate the 1-day and 4-day rainfall amounts is too sketchy. There is not enough detail to allow an independent researcher to apply the method and replicate the data. Hence, it is strongly recommended that the authors produce a flowchart showing two parallel vertical tracks: (a) Track 1 is a brief but clear written description of the key steps. (b) Track 2 shows graphically how the raw data are transformed into the final product along each step of the path. Hence, in the case of the 4-day moving average rainfall dataset, the first graph would presumably show hourly rainfall amounts over, say a period of a week; the last graph would show time-averaged rainfall intensities spread over a 4-day period; the intermediate graphs would illustrate all the data transformations made along the way. This could be done for both the historical data and the GCM outputs. When properly completed, the flowchart will contain enough information so that a class of engineering students could use it and could replicate the procedure and generate the same precipitation data set.	Agree: The presentation of methods in Section 2.2 has been substantially revised to be more detailed and clear. We have not added the figures suggested by the reviewer in order to keep the presentation brief. The approach is relatively straightforward, and although we agree with the reviewer that new figures would help to clarify the presentation of methods, we do not feel this is necessary to follow the basic steps taken in the analysis.
Whenever possible, all bar charts should be set up to have identical scales on their X axis and on their Y axis. This will help the reader to make quick effective comparisons among the different results. The 8 bar graphs have 7 different X axis scales (some are very busy and hence, difficult for the reader to decipher) and 3 different Y axis scales. A similar remark applies to the cumulative distributions, too.	Disagree: Although we agree in general that, if possible, scales should be kept consistent, because of differences in the range of values plotted we feel the value added by narrowing the x-axis range of each plot to only those values needed to show the data outweighs the value from keeping x-axis scales consistent. Due to differences in the future time periods evaluated in the 2 study regions, inter-regional comparison is also not a goal of the report. The figures have not been revised.
Finally, the authors should prepare a half page matrix to concisely summarize their findings (see Original Comments from Review for Table)	Disagree: A matrix of results has not been added. We agree a summary matrix would be helpful for comparing results between the regions. However, due to differences in the future time period considered in the different regions, comparison of results between the regions is not possible or a goal of this study. To help address any confusion in the presentation of results, Section 4 on results has been extensively re-written to be more clear.

...Thus, the general recommendation is to revise the report to stress a single theme of demonstration and not recommendation. In places the report implies the results may provide specific guidance for design or operation changes to respond to POTENTIAL climate change. The uncertainty of climate change in general and the methods in particular and the generality of the analysis restrict applicability to a general demonstration and should not be used for site-specific application. This caveat needs to be considered and the report revised to clearly state it (in all parts – Executive Summary, Introduction, Body, Conclusions). I believe the tone of the other reviewers is similar.	Agree: This study is a screening-level assessment of the potential order of magnitude of climate change impacts on combined sewer overflow mitigation efforts in the New England and Great Lakes Regions. The general intent was to assess whether the potential implications of climate change on CSOs in these regions warrant further consideration and study, and to assess the need for decision support tools and information that allow water managers to better incorporate consideration of climate change into their decision making process. As such, this study is only a first step towards understanding a complex issue, the implications of which will vary significantly for specific locations and systems. We have added text throughout the document (Exec Summary, Introduction, Methods, and Conclusion) to properly caveat results and be more clear about the limitations of this work. We also include a discussion of study limitations in Section 4.3.

To summarize my review, I feel that the overall concept of screening a time series of rainfall to estimate potential future Combined Sewer Overflows (CSOs) on a regional basis is sound and has technical merit. I also feel based on the evidence provided that the GCM model data used are an appropriate tool for developing the time series of rainfall. However I have some reservations related to the lack of information related to the rainfall disaggregation methods underlying the analysis, even while considering the gross regional, "screening-tool" philosophy employed. There should be a clear explanation of both the VEMAP GCM and the WGEN models that were used to generate the rainfall time series. Therefore I can not, at this point, endorse the results and conclusions of the report without a better understanding of the methods employed for the study. While I do not disagree with the results and conclusions, there is not enough information to make a decision. With additional information related to the methods employed I could better make a judgment on the results and possibly endorse them. The area that I am uncomfortable with is related to the disaggregation of monthly rainfall to daily rainfall totals. Specifically: a) I suggest that more detail be presented in the report regarding the stochastic weather generator WGEN. The tool is key to the analysis and without information regarding this tool it is difficult to assess the technical merit underlying the entire analysis, b) · Specifically, were the WGEN parameters estimated based on historical monthly-to-daily rainfall statistics? Or were they estimated based on some other estimate of future frequency and intensity statistics under a climate change scenario? This question must be answered prior to acceptance as a "final" deliverable. c) · The disaggregation method should be statistically "back-tested" on historical data that includes monthly/daily/hourly rainfall and CSO frequency (even if CSO frequency is estimated from a model) to verify its validity for the purposes employed. This would provide the evidence needed that the disaggregation technique is a meaningful way to estimate CSO frequency, thus adding credence to its use with GCM-generated data. d) · In general, I think a true rainfall "event" analysis would better predict CSO frequency than daily rainfall totals. An event analysis characterizes rainfall in a way that is more meaningful to CSO frequency prediction. More clarification of this point is provided in the answer to question #2. While I wholeheartedly encourage this type of screening analysis, I can not recommend that this report be published as "final" until the above cited questions are sorted out related to WGEN. I feel that the report could be published as final if the first three bullet points are addressed; however I feel the results of the verification of the dissagregation model would show that an event analysis would better characterize CSO frequency than a database of daily rainfall totals.	Agree: We agree the disaggregation of daily data is very important. The presentation of methods in Section 2.2 has been substantially revised to be more detailed and clear. For this simple analysis we chose to use the VEMAP daily data because it is readily available, and widely known and documented. Additional information has been added to the Methods section about the VEMAP data. References are also given where readers can go for more detailed discussion of how the VEMAP data was created, including use of the WGEN weather generator.
Charge Question 1	

The organization of the report is acceptable; however, the content and conception of the report suffer from several shortcomings: [1] Reproducibility of Approach – The description of the method is very vague. There is not enough detail to allow an independent researcher to apply the method, replicate the data and reproduce the analysis. The authors must provide more information on the method(s) used to generate the rainfall data. [2] Effective Communication – There are many instances where the authors could and should use figures, illustrations, or diagrams to reinforce their point and convey their meaning. For instance, while the report is devoted to the occurrence of combined sewer overflows during rainfall events, the document lacks even a simple definition sketch for (i) a CSO and (ii) a rainfall hyetograph. [3] Statistical Significance – It is not enough to point out differences in results produced by the competing models and by the time periods. The authors should apply simple hypothesis testing to confirm whether or not the observed differences are statistically significant. [4] Conflicting Message – In the end, there is no clear message. The conflicting and inconsistent outcomes will, unfortunately, greatly diminish the utility of the report and essentially neutralize it as a potential guide for planning or policy.	Agree [1]: The presentation of methods in section 2.2 has been substantially revised to be more clear. Disagree[2]: The suggestion to add a figure showing a typical CSS is a good one, but we have not added in order to keep the presentation brief; it is assumed that readers will already know the basic structure of a CSS. Disagree[3]: We agree in principle about the importance of documenting the statistical significance of results, e.g. projected changes. However, given the many other limitations of this study, and the screening-level analysis goal to identify the potential range of future changes, we have not conducted statistical hypothesis testing. Agree[4]: Disagreement in the direction of change projected by the two models in the New England region complicate interpretation of results. Because of the different future time periods used in the two regions, the study cannot determine whether this disagreement is due to the time period considered or projected regional differences in climate change. Results in the New England Region are thus inconclusive, and dont support or refute the liklihood of future impacts. Future study is required to address the questions this study did not. This is now acknowledged in the Results and Conclusions section.
The document is well organized into an introduction, description of methods, presentation of results, and conclusions. Additional sections describing limitations and future work are especially welcome. The level of detail provided for this screening-level report is appropriate, although the reference to the VEMAP document for description of the uncertainty associated with the GCM and downscaling should be expanded. The concept of the uncertainty associated with the methods is not clearly indicated by the report because this key discussion is not included. Additional content might be added to further emphasize the limitations of the methods and results in the introduction and executive summary. These parts of the report do not adequately present the considerable uncertainty of the methods and thus do not clearly indicate to the reader that the results are meant to demonstrate the general need to consider climate change, not a recommendation to include the results in their site-specific application (it is described in other parts of the report, just not in the early parts).	Agree: The presentation of methods in Section 2.2 has been substantially revised to be more clear. A new paragraph acknowledging several sources of uncertainty in the VEMAP data has been added.
The document organization is fine overall. However, the limitations of the report should be discussed in the executive summary.	Agree: The executive summary has been re-written, and study limitations are now mentioned.

Overall the effectiveness of the document could be strengthened by providing more background information on methods underlying the results. This is particularly true because the results show considerable variance; without detail on the background the variance of the results lack context and are not easily understood. Impacted stakeholders are less likely to trust the report without solid background from which to understand the variance depicted. The overall organization of this report is adequate. The level of detail is generally inadequate, especially with respect to the VEMAP GCM and the WGEN models. The completeness of the results is generally adequate; however the variance of the output should be more clearly stated in the executive summary. For example, figures 2 through 13 show considerable variance in the results. A geographic representation of the spatial variation of these results would be very useful to the reader. Additional information that would strengthen the report: - Additional information on the VEMAP GCM data and the differences between the Hadley and Canadian models. - Additional information on the WGEN stochastic weather generator. Because CSO response is very sensitive to rainfall intensity, I suspect the results of the analysis to be extremely sensitive to the methods used to disaggregate the monthly rainfall totals to daily totals. More information is needed so the reader can understand how the parameters underlying the stochastic weather generator were derived. - Spatial distribution of the results would be especially useful. For example, a GIS generated map that depicts how the variance shown in Figures 2-13 is spatially distributed would be very helpful to the reader.	Agree: The presentation of methods in Section 2.2 has been substantially revised to be more clear.
Charge Question 2	
The experimental design is marginal. Two different GCM models were applied to two different regions for two different time periods. It seems the authors have not carefully controlled the potential sources of variability and, consequently, they are unable to determine if the differences in outcomes are due to differences in the two time periods or to differences in the two geographic regions. This is a weakness which limits the usefulness of the final results. Further, the authors must better justify the use of daily time steps rather than hourly time steps for storm analysis. Numerous other suggestions to improve the study are provided under question [7].	Agree: The different time steps used in the 2 regions are problematic. This resulted from 2 studies conducted at different times and with different objectives being combined in this report. Agree: We agree that the use of hourly time step (event) data would be better, but daily data was much more easily available, and we feel is adequate to meet the goals of this study - a course, screening level analysis. We agree that more detailed study using hourly or greater temporal resolution is necessary to evaluate specific CSO mitigation design decisions at individual CSSs. These limitations are now mentioned in the Conclusions and Exec Summary.

The technical merit is OK. The study is limited by the use of the GCM and numerous methodological choices made to simplify the analysis. The goal of the report I am assuming is to initiate planners and engineers to the concept and potential impact of climate change on CSO control planning. The technical content is sufficient to make this point and it will successfully motivate CSO control planners to find ways to incorporate this consideration into their planning. The report also should motivate EPA and other federal agencies and research organizations to study this problem more rigorously and develop more accurate and substantiated guidance for communities to use. One technical area of concern is the choice to conduct the analysis using the GCM results given the considerable uncertainty. In addition to the limitations of the GCM, the limitations from scaling inaccuracies, lack of short duration (less than 1-day) events, missing snowmelt, and shift from snow to rain under certain climate change scenarios are also important and mentioned in the report. The magnitude of uncertainty associated with this analysis seems to point to the need for a more robust approach not reliant on specific climate scenarios. One way to remove the reliance on the GCM output based on one scenario (although two different models) would be to use a sensitivity analysis approach. The authors do this to some extent with the runoff analysis in that they use two possible conditions to set the range, but even that is limited in that both of those range limits may be incorrect. It seems a more general approach would be prudent, where incremental changes are analyzed and the limits of the ranges in incremental change are set based on some value obtained from a GCM or another study, but tempered with intuition and judgment. Flexibility and adaptability of system planning and design may then be better incorporated if ranges of potential impacts are considered. I would discourage providing specific numeric changes based on this analysis, but rather would recommend the report stress flexibility and adaptability in design that can account for future changes in a cost effective manner.	Agree: We agree with the reviewers comments concerning the many limitations of this study. Our intent is simply a course, screening-level analysis to "assess whether the implications of climate change in these regions warrant further consideration and study, and assess the need for decision support tools and information that allow water managers to incorporate consideration of climate change into their decision making process ". This is now mentioned in the Introduction, Conclusions and Exec Summary.

The study takes a very very simplified approach to assessing the potential impacts of global warming on CSO systems. I was wondering whether the use of a design storm is the typical approach for CSO system design? I would hope not and that continuous simulation approaches are applied. If that is the case, then there should have been some analysis of the implications of using design storm approach for this screening level analysis. There should have been some description of the typical design tools and methodologies used for designing and sizing CSO systems. Some survey of methods would have been appropriate so that the coarse method applied here could be compared. For example, continuous simulation models with about 50 year records? Factors of safety included? I would doubt that most systems would be designed without a more thurough analysis then using design storms. I would be concerned that a 24-hour rainfall analysis (as snow and snow melt were ignored) that is midnight-to-midnight and does not consider back-to-back storms could be very off from what one would design using continuous simulation approaches. I do think that the SYNOP program could have been used to separate storm events (vs. arbitrary time of midnight) to get a closer handle on storm events. I have not seen many storms that end at midnight. I have seen day after day of rainfall however. So, it may be very possible that systems design using continuous simulation that account for contiguous days of rainfall may not be that effected by there being more intense events, but fewer of them. The 4-day analysis would account for some of this. However, have the climate models also predicted more sustained days of rainfall along with increases in intensities? Not considering snow and snow-melt, including temperatures is also problematic. With warming, the frequency of snow may decrease. Melts between storms could increase, therefore making potentially less rain on snow events. A continuous simulation model could assess whether this could counter balance CSO design as some of the 4 events per year considered were rain on snow events. Some specific case studies should have been run as a way to “validate” the screening approach taken.	Agree: We agree that continous simulation would be an excellent approach, but modeling at the spatial scale considered here would be very costly and time consuming. The design storm approach used here and based on the "presumptive" strategy is a simple and straightforward informative way to screen for the general vulnerability of CSS in different regions. These limitations are now mentioned in the appropriate locations in the report.
Specific comments on the document: Page 1, lines 11-13. Should mention that with less frequent rainfall that irrigation levels would increase, which can counter-balance stormwater infiltration decreases.	Agree: We agree that changes in water use and management including irrigation could impact streamflow regimes. However, Section 1.3 on hydrologic impacts has been dropped completely from the report as it was not necessary to the issue of urban stormwater runoff. Thus this change was not made.
Page 1, line 24. The assumption that precipitation and hydrology are constant over time is not correct. If one uses a 50-year rainfall record in a continuous simulation approach to design a system, you are assuming that that is representative of typical 50-year sequences. However, you are NOT assuming that precipitation and hydrology are constant over time.	Agree: We agree. Precipitation and streamflow are random process. The statement was meant to refer to stationarity in a statistical sense; e.g. constant mean, variance, autocorrelation structure). Text in Section 1 has been revised to say that climate is typically assumed to be stationary.
Page 2, Line 7. Should acknowledge here that some CSOs are partially treated via screening or other methods (is later acknowledged on page 3, but mention of untreated or partially treated discharges would be good).	Agree: Text has been added in Section 1.1 mentioning that CSOs can include a mixture of raw and partially treated effluent.
Page 5, line 41. I would think that periods (e.g. time) would increase. Low flows may decrease. This effect would be somewhat off-set by increased irrigation in some cases	Agree: We agree use of text "increased low flows" is confusing. However, Section 1.3 on hydrologic impacts has been dropped completely from the report as it was not necessary to the issue of urban stormwater runoff. Thus this change was not made.

Page 6, Lines 4-6. Should acknowledge that increased runoff can dilute some pollutants. All depends on pollutant type and sources.	Agree: Section 1.3 on hydrologic impacts has been dropped completely from the report as it was not necessary to the issue of urban stormwater runoff. Thus this change was not made.
Page 6. Line 20.. Should acknowledge that "The research questions to be evaluated using a coarse screening level approach were:"	Agree: Text in Section 1.4 and throughout the document has been revised to refer to this study as a "screening-level analysis"
Page 9, Lines 1-3. I think you should have aggregated the historical data into monthly totals and then used the WGEN weather generator on that to see if the generator introduced some of the changes in storm sizes.	Agree: This is an excellent suggestion, however, in this study we used the VEMAP data "off the shelf", and cannot go back and rerun the version of WGEN originally used for VEMAP on the monthly historical data. The validity of WGEN simulations in the VEMAP data was also evaluated by the VEMAP team by comparing daily weather statistics from WGEN scenarios to daily historical data. This evaluation suggests WGEN does a reasonably good job of reproducing historical daily weather patterns, and can be found in the VEMAP documentation. Additional text has been added to Section 2.2 concerning the validity of WGEN daily weather scenarios.
Page 19-20. The fact that Chiew and FAO have opposite predictions shows the uncertainty in this one effect alonw.	Agree: Chiew et al. suggest the percent change in runoff is 2X the percent change in precipitation, and FAO suggests the percent change in runoff is 0.5X the percent change in precipitation. For this simple screening analysis, continous hydrologic modeling was not feasible. These values were selected to represent high and low endpoints of a broad range of potential runoff responses to changes in precipitation. The change in direction reflects the complexity of hydrologic response to changes in climate in different regions. Text has been added to Section 3.2.3 stating this.
Page 21. Lines 29-35. This seems like a serious caution, but not mentioned in Exec Summary and overall conclusions.	Agree: Text has been added to the Executive Summary and Conclusions referencing this caution.
Page 22. line 11. Add to this sentence. "…Water Management Model (SWMM) and especially if it was used in a continuous simulation mode that would address rainfall, snowfall, snowmelt, and rain on snow runoff events and system design." Using the model on an event basis would miss a lot for example.	Agree: Text has been added to Section 3.4.1 mentioning the the utility of continuous hydrologic modeling.
Page 22, lines 28-37. Why not use another approach to separate storm events? Use of a 6-hour dry period for example (could use SYNOP). You avoid the arbitrary cut-off (and joining together) of separate storm events.	Agree: However, the VEMAP data used in this study was available only at a daily time step. Hourly data would be needed to provide better handling of sub-daily events. Daily data is adequate for this coarse, screening analysis. More sophisticated ways for addressing sub-daily events will be necessary for a detailed analysis. Text has been added to Section 3.4.1 making this point.
Page 23. Lines 32-34. Case studies would also show how snow and snow/melt are addressed.	Agree: Text has been added referencing the importance of addressing snow and snowmelt.
Page 25. Line 2. Should add change in precipitation as snow, changes in snow melt rates.	Agree: Text has been added to Section 4 referencing snow and snowmelt as possible impacts of climate change.
Page 25. Should note in conclusion that the findings are based on the use of a design storm approach for CSO design and no consideration of changes in snow/rain on snow events. In general I think the conclusions are made a little strong for such a limited analysis.	Agree: Text has been added to Section 4 stating that conclusions are based on a screening-level analysis. Additional qualifiers have also been added concerning the methodology.

Page 25. Third Bullet. I think that one really need to look at snow/ snowmelt events to see if this trend would in fact result in more CSO events.	Agree: Text has been added to Section 4 mentioning that the methodology did not consider the effects of changes in snow and snow melt on CSO frequency.
It is difficult to properly judge the technical merit of the conclusions because there is missing information related to the disaggregation approach related to the stochastic weather generator. It appears that daily rainfall totals were statistically generated from the GCM model output in the form of monthly rainfall totals. It is unclear what the basis of this method is; is it based on estimated future conditions under a climate change scenario or is it based on historical monthly versus daily rainfall statistics? Has this relationship been adjusted to account for potential changes in the Intensity-Duration-Frequency statistics resulting from future climate change scenarios? The answers to these questions are important to the overall usefulness of the results. There are two other related areas that could be used to improve the usefulness of the screening results: 1) disaggregate to event totals rather than daily totals, and 2) develop a validation strategy to verify the disaggregation scheme. Each of these points are addressed below. CSO events tend to be triggered by high rates of rainfall intensity rather than large daily totals. This characteristic of urban hydrologic systems leads the analysis to a temporal domain that is clearly too detailed for a screening tool by requiring time scales at the hourly or sub-hourly level. However the daily totals clearly misrepresents rainfall and its relationship to rainfall frequency. For example, a 2 hour, 1 inch rainfall event that occurs between 11 pm and 1 am would register as two events in a daily rainfall database. And the implied intensity is misrepresented in a daily total as well, while the actual intensity is 0.5 in/hr the daily database would imply an intensity of 0.02 in/hr. A simple way around this problem is to use a stochastic disagreggation method that generates events from monthly totals rather than daily totals. A monthly total would disaggregate to some number of events, each with a total rainfall and duration. While this duration would be in hours or fractional days, it does not require an entire hourly time series and in fact is a compact way of representing long precipitation time series. A statistical back test of the stochastic weather generator should be performed to provide credence to the application of the method to future VEMAP GCM-generated data sets. Perhaps this verification has been done, however this information is not included in the report. A statistical back test would use a database of rainfall monthly, daily and event totals. The stochastic weather generator would be used to create daily and event totals, and the statistics of these synthetic data sets would be compared with the actual measured data. If this analysis has been performed, this verification process should be cited. If this analysis has not been done, it should be used to compare the accuracy of the daily and event based methods.	Agree: We agree with these concerns; they are all excellent suggestions. However, in this study we used the VEMAP data "off the shelf", and we do not have access to hourly data. For this screening level analysis, events that cross 12 midnight and are split into 2 are assumed negligible. Regarding validation of the WGEN data, we agree this would be an excellent idea but it is beyond the scope of this study to go back and rerun the version of WGEN originally used for VEMAP on the monthly historical data. Also, the validity of WGEN simulations in the VEMAP data has been evaluated by the VEMAP team by comparing daily weather statistics from WGEN scenarios to daily historical data. This evaluation suggests WGEN does a reasonably good job of reproducing historical daily weather patterns, and can be found in the VEMAP documentation. Additional text has been added to Section 2.2 concerning the validity of WGEN daily weather scenarios.
Charge Question 3	
I am not aware of any essential references that are missing.	OK

Focus of the report is on global climate change. No consideration is given to the effects of urban areas on weather patterns that may have a significant impact on temperature and precipitation, and in turn may impact CSO. Although urban-induced rainfall modification does not fit the classical definition of global climate change, it is climate change. There is a lengthy body of literature on the topic (including works by Bob Bornstein, Marshall Shepherd, and Stan Changnon). It is recommended that the authors consider reviewing some of this literature and at least mention it as another possible climate influence that necessitates appropriate CSO control planning and design resiliency. It may be that the effect of the urbanization and associated environmental modification (air pollution) may have a greater impact on localized rainfall patterns and thus a greater impact on CSO. Mentioning it in the report will at least provide planners another area of uncertainty to consider.	Agree: We agree this an interesting line of research, but did not attempt generate new climate change scenarios. Instead we relied on an existing data sets.
Charge Question 4	
Does the issue addressed by this analysis warrant further study? If so, what do you think the required next steps are to better, or more completely address the study goals?	
The authors should use concurrent study periods for both regions.	Agree: Future study, no edits required
Yes, this study identified one form of uncertainty that is not factored into the design of wastewater infrastructure. The authors outlined several useful extensions and revisions to the current study. There are also other areas of uncertainty to consider. Eventually, a study needs to be performed to identify and quantify the various uncertainty components in wastewater infrastructure design, operation, and management. Then the next step would be to develop new approaches to embed cost-effective adaptability and resiliency into system design (in retrofit and new design).	Agree: Future study, no edits required
It does warrant further study. I have addressed a number of items above that I think could be completed to more completely address the study goals. In summary: a. Use same rainfall disaggration method on historical monthly rainfall data to ascertain whether the technique used, b. Review typical CSO design methods and tools to see if the design storm approach utilized in this work is valid or not, c. Conduct some case studies that include continuous simulation modeling of precipitation and accounting for changes in temperatures to ascertain whether more intense precipitation results in increased CSO events with actual designs, d. Separate precipitation data into storm events by a selected number of dry hours (e.g. 6 hours dry and then later precipitation data is part of a new storm) to end the arbitrary cut-off of midnight to midnight. A clue that this may be a problem is that less increases occurs with 4-day events which would be less effected by this, e. Consider running simple continuous simulation models instead of rainfall analysis that would account for storage and treatment rate.	Agree: Future study, no edits required

Yes I believe the issues addressed here warrant additional study. In addition to the hydrologic improvements suggested above that are needed to improve the current work, the cost component of the subject matter should be explored in future work. Specifically, the areas of economies of scale and diminishing returns need to be explored with respect to CSO control mitigation measures. As noted in the review by Joel Smith, the linear assumptions of cost-to-volume are probably not warranted. Also storage should not be assumed to be the only treatment choice; high rate treatment options exist and would work in tandem with storage to reduce the cost of mitigation. The implications of these to the cost analysis should be addressed. Another area that should be explored is the effect of potential climate change on water quality impacts of CSOs. For example, if precipitation falls more sporadically but more intensely under future scenarios, and streamflows are therefore lower at the onset of an event due to lower baseflows, how might the receiving water respond to a CSO event? In other words, will the receiving water biota be more or less sensitive to the abrupt, episodic nature of CSO discharges under lower baseflow conditions with higher peak flows?	Agree: Future study, no edits required
Charge Question 5	
How important is the issue addressed by this study relative to other potential impacts of climate change on water quality and/or EPA water quality protection programs?	
The impact of climate change on CSOs deserves a high priority with the agency's water quality protection programs.	Agree: Future study, no edits required
I think the study is critical because of the massive costs associated with the typical storage mitigation measures associated with CSO control. And CSO mitigation projects are typically long-term and must project far into the future. Therefore, the magnitude of uncertainty is greater than shorter term planning and design projects. Furthermore, the time required to respond with system modifications is great and costly requiring careful planning now, once again stressing adaptation and flexibility to change.	Agree: Future study, no edits required
I think it is valid to be concerned that $ going for increased CSO programs may take away $ from urban runoff issues, etc.	Agree: Future study, no edits required
This area is important, however I believe a few other areas of research would rank as higher-priority. Namely, watershed approaches should be performed first to gain some understanding to the overall impact of climate change on the hydrologic water and pollution budget at the watershed scale. This would look at water use, infiltration, stream flow and stormwater runoff as well as SSO and CSO discharges. The importance of CSO in the overall watershed would then be apparent at the regional level and research efforts could be appropriately prioritized.	Agree: Future study, no edits required

Charge Question 6	
Please read and consider all public comments received by EPA on this report (comments were received from 5 individuals/groups). Please comment on the validity and appropriateness of comments submitted by each public entity. Are there specific comments you strongly agree or disagree with? If so, which ones?	
Overall, the public comments were very insightful and thoughtful…	Agree:
[b] MMSD – Four main points were made: (i) The focus should be on control of stormwater runoff because in some watersheds CSO and SSO pollutant loadings are small in comparison to stormwater inputs during wet weather events. This is a valid point and should be acknowledged in the EPA report. (ii) Results of the regional screening are highly variable; more credibility should be given to simulation exercises that are based on site-specific conditions. This point can be noted, but the challenge for the Agency would be to ensure rigorous uniform reproducible analyses prevailed from site to site. (iii) Results of the regional screening do not specify how or which GLR communities are impacted. This point can be noted, but it is beyond the scope of the EPA study. (iv) Results of the regional screening may be helpful to communities in the early planning stages. No need to comment.	Agree:
[c] AWWA – Three general points were made: (i) There needs to be more explanation of the methods used to get the benchmark storm events. Yes, this is a very important issue and I strongly agree with this recommendation. (ii) There were many, many concerns about cost implications. This is outside the scope of the current study; there should be a disclaimer about the economic impact associated with the CSO study. (iii) Case studies may be the next logical step to investigate the interplay between the "staggering uncertainty" and the "staggering cost". This is an excellent suggestion and it should be included as a recommendation for future work.	Agree:
[d] GL Waterkeeper – Two main points were made: (i) How is a CSO formally defined so that it is standardized across communities? This is a very good question and it should be addressed in the EPA report. (ii) The study does not consider projections for increased development. This is another good point. The study addresses climate change, but what about watershed change? This should be clarified in the EPA report.	Agree:
[e] NACWA – Four specific points were made: (i) The linear relationship between storage and increase intensity seems to be over-simplified. This is a reasonable assumption; the authors should elaborate and clarify this point. (ii) Report should acknowledge that LTCPs can be phased. This is out-of-scope, but could be addressed very easily. (iii) Report should acknowledge that storage is but one of several technologies considered by CSO communities. Agree and this is easy to incorporate. (iv) The use of full-day storm increments is questioned (there is a very different response from a 2-inch 24-hour storm and a 2-inch 2-hour storm). Agree. The authors should elaborate further on how the time interval (days –versus- hours) affects the sensitivity of their simulation results.	Agree:

Review of Comments by Marjorie Wall: These comments do not seem relevant as specific CSOs were not assessed, but rather the study was focused on precipitation analysis. It may be important in the context of case studies as suggested other reviewers.	Agree:
Review of Comments by MMSD: I. I agree with this comment – stormwater pollution is of greater importance from an annual perspective for many locations because on an annual basis the quantity of discharge and accompanying pollutants will be larger. However, I am not sure what the impacts of climate change would be on the annual stormwater pollutant load. Further, the focus of this report was not on climate change effects on seasonal or annual rainfall characteristics, which would be more important that the increase in single event intensity for the longer-term pollutant loading issues (e.g., separate stormwater discharges). I agree with the comment that the analysis is too simple to provide site-specific guidance (as I note above) and I also agree that if the EPA is taking a broader view of climate change impacts on water quality then stormwater would rank higher in priority than CSO (especially for cities that do not have CSS). But given CSO is the subject of the report I must assume the decision to prioritize CSO has already been made and is not relevant for the review of this report. But I would support the recommendation that EPA perform a similar assessment for stormwater and other higher priority water quality impacts from climate change at the national level. II. I agree entirely with the spirit of this comment. The analysis included in the report is too simplified and general to be used as guidance for specific entities where site-specific information and local judgment are necessary to effectively plan for climate change impacts. This is consistent with the point I made above that the report should contain language indicating that the document is not meant for site-specific guidance (designing for 10% increase in storage for example), but rather as an illustration of the need to consider climate change in local analyses leading to site-specific plan and design. III. I do not believe the report should include mention of site-specific locations. The analyses are too simplified and based on too uncertain data and methods to provide accurate estimates at particular locations. Providing a summary of specific systems that are higher or lower than the benchmark would likely result in incorrect responses. This ties back to the recommendation to caveat the report findings as not providing site-specific results, but providing a general recommendation based on a highly uncertain approach. IV. This comment is true – some communities are much farther along and have developed much better ideas of climate change impacts for their particular systems. This is why making this a guidance document would be inappropriate.	Agree:

Review of Comments by AWWA: Several statements were made in the cover letter recommending water quantity and water quality impacts of climate change be made a research priority by EPA is beyond the scope of the present report, but deserves serious consideration nonetheless. This is one of my recommendations above and is consistent with other reviewer comments. The recommendation to expand the VEMAP section has merit. It is consistent with my observation as described in review question 1 above. Upon further consideration I would revise my comment above to be consistent with the comment made in the AWWA contractor's review – to expand the section providing greater clarification of VEMAP and the precipitation data used in the analysis. The reference to the Lang and Balmforth (2005) study of climate change impacts on CSOs in the U.K. and the follow up about the inaccuracies of using a linear relationship precipitation, runoff volumes, and storage costs are important recommendations. The technical merit of the approach used in the study is questionable and must be explained/defended in more depth. The recommendation for case studies is a good one. The report lacks specific data of use to CSO planners, rather it stresses the potential of climate change impacts in a general sense. For the most part water managers are on cognizant of climate change and are aware of the need to consider it, but they lack the tools and knowledge to address it. This recommendation is consistent with the recommendation of the MMSD – to provide results for specific locations. But I disagree with that recommendation for this particular report unless the methods are substantially improved.	Agree:
Review of Comments by Great Lakes Waterkeepers: The question of CSO definition is important, but not relevant for the scope of the screening-level assessment. The screening-level assessment is based on defining a benchmark precipitation event, which is entirely driven by analysis of precipitation records. The definition of a CSO is not used. I do not see how it could be in a general sense for this report. It would be an absolute necessity if case studies were to be incorporated as recommended above, but for the screening analysis performed the assumption is that rainfall events of higher intensity/daily amount than the benchmark will cause a CSO. To define the occurrence of a CSO in detail would require a much finer resolution and scale of simulation on the order of what is mentioned in the MMSD comments (using MOUSE, SWMM, or other urban hydrology-hydraulic model).	Agree:

Review of Comments by NACWA: The criticism of the suggestion in the report that communities need to provide margins of safety in their CSO projects or to make other operational changes now is warranted and is consistent with comments by the other reviewers. The methods employed for the study are too uncertain to serve as recommendation for incorporating changes in current designs or operation to account for a very uncertain future. Rather the report should stress the need to consider climate change factors in long-term planning, which I believe it does for the most part. A subsequent report should then focus on case studies, specific guidance, and review of available technologies. I agree with the comment about the lack of mention of the limitations/uncertainty of the approach in the Executive Summary and Introduction – this comment was consistent with my observations reported above. This is a second criticism of the use of the linear relationship between precipitation, runoff, and costs. To enhance the technical merit of the approach, a more accurate approach should be selected or more stress placed on the fact that the report is meant more as a general recommendation to consider climate change in LTCPs rather than requiring it be factored into designs and operations based on the analysis contained in the report. The decision to incorporate climate change into a specific location's design or operation should be based on site-specific analysis and not the report. I agree with the comment to acknowledge other technologies besides storage as available to CSO control planners. However, I am not sure this needs to be emphasized in the report, rather a simple statement such as "storage or other technologies" should suffice or placing a listing of other technologies in parentheses. The comment about the lack of short-term rainfall potential by GCMs is correct. Spatiotemporal downscaling of GCM output is very crude when taking it to sub-daily time increments at a watershed scale. Essentially, it is not possible to accomplish with any degree of accuracy for site-specific analysis. Perhaps some day, but it should not be implied in the report that such an extension of the research is feasible at this time.	Agree:
American Water Works Association – Valid point that the precipitation record may already show some trends that are included in the historical record. Comments generally are fine.	Agree:
NACWA- Pointed out that agencies who design using historical record with design contingencies. Also valid point about meeting water quality goals vs. targeted number of CSO events.	Agree:

MMSD- Valid point that separated storm water is a big WQ issue. This would vary from situation to situation. Some assessment of the trade-offs on increased CSO control levels for global climate change vs. more $ for urban stormwater control could be assessed. However that is not the point of the EPA study. Agree with 2. that there is a lot of uncertainty in the predictions.	Agree
Milwaukee River Keeper, et. al. – Not sure that CSO systems would see a big increase in percent imperviousness as usually newly developed areas are in separated areas. Not sure that an action plan is needed yet until there is more science beyond this screening level look.	Agree
NACWA: Overall the MACWA comments are reasonable. I especially agree with the third bullet point on page 3 of their letter.	Agree
MMSD: Generally the point related to the relative importance of stormwater to CSO is important, however they make rather bold statements that might not be true for all jurisdictions related to the relative importance of stormwater to CSO. Generally they are correct in noting that the high variance in the results precludes the report as a basis for financial decision making, however I do not think this was the author's intent.	Agree
AWWA: The review by Joel Smith is very good and raises many important points. I agree with this assessment of the report.	Agree
Milwaukee Riverkeeper: Most of the concerns raised in this letter do not pertain to the report at hand, but rather on underlying assumptions such as problems with the presumptive approach, event definition, etc. The point regarding increased development and increased inflow and infiltration is important to specific systems, however these changes will occur irrespective of climate change and additional development will most likely not be served by a combined sewer. Therefore these issues should really cloud the results of the research in question.	Agree
Charge Question 7	
Do you have other comments or suggestions for improving the quality of this document?	
Page iv line 6, express 1 degree of latitude and longitude in terms of kilometers at the study site.	Agree: Executive summary has been re-written. VEMAP grid information in degrees no longer included. Thus change no longer required.
Page 1 Provide a definition sketch of a "typical" municipal wastewater collection system! Show the overall watershed (sewershed), collection points, the conveyance elements, the overflow points, the WWTP, the receiving streams, etc.	Disagree: It is assumed that the typical reader of this report knows how a typical CSS works. To keep the presentation brief, this figure has not been added.
Page 2 Under Figure 1, add a new heading called "History" (enjoyed the background material).	Agree: A new Section title, "History", has been added.
Page 3 line 9 is out of place and redundant with page 2 line 10.	Agree: The redundant text has been removed from this paragraph.
Page 6 line 3, "…result in increased high-flow stream conditions." Please clarify if you are referring to an increase in *frequency* of high flows or in *magnitude* of high flows or both?	Agree: Section 1.3 has been removed completely from the report because not relevant to urban stormwater runoff. Thus change no loger required.

Page 6 line 4-6, provide clarification and cite a reference for the statement that "Extreme precipitation events tend to be correlated to poorest water quality…"	Agree: Section 1.3 has been removed completely from the report because not relevant to urban stormwater runoff. Thus change no loger required.
Page 7 line 9, Indicate the total number of CSS communities that were mapped (or refer to results shown in Tables 1 and 2).	Agree: Text has been added giving the total number of CSS communities mapped.
Table 1, Suggest the States be listed in alphabetical or numerical order. Current listing appears to be random. Add columns showing population served and the percentage of the total study region.	Agree: States now ordered from highest to lowest number of CSSs. Information on percent of area and population served was not added. Data is not readily available and not directly relevant to study goals. Goals were not to identify variation within regions, but rather to estimate broad regionwide averages.
Table 2, Same comments as Table 1; Why is New York listed twice? Suggest you include a map that delineates the Great Lakes Region and the New England Region, as defined in this study.	Agree: States now ordered from highest to lowest number of CSSs. Information on percent of area and population served was not added. Data is not readily available and not directly relevant to study goals. Goals were not to identify variation within regions, but rather to estimate broad regionwide averages.
Page 8 line 23, Need to justify the selection of daily precipitation. Why was this time scale deemed appropriate for this analysis? Why not hourly rainfall values, for instance?	Agree: Text has been added to Section 2.2 better describing why the daily time step VEMAP data was used, and adding the caveat that daily data does not allow consideration of individual event characteristics.
Page 8 line 28, "…projected…" this term is undefined.	Agree: "projected" has been changed to "projected future"
Page 9 line 2, Who provided the "modified version" of WGEN? What modifications were performed and who did them?	Agree: Additional text has been added to Section 2.2 describing the use of WGEN. The VEMAP documentation is also cited as a source for a more detailed discussion of the VEMAP data.
Page 9 line 11, In an appendix, provide a cross-listing showing VEMAP sites and the assigned CSS communities.	Disagree: Figure not added. The goal of this study was not to focus on individual systems or variation with regions (different VEMAP points). Rather, the intent was to present broad, regionwide averages. This information is thus not critical to study goals.
Page 9 line 20, Strongly suggest that the authors confirm/demonstrate that the annual sum of the 1-day historical rainfall amounts recovers the published yearly rainfall amount. Similar QA step can be applied to the 4-day rainfall values.	Disagree: We agree this would be helpful, but it is beyond the scope of this study. This analysis was not added. We used the daily VEMAP data "off the shelf". For each VEMAP grid location, daily precipitation data from WGEN based on historical data were constrained by long-term monthly mean station values. Based on their quality control analysis, the VEMAP team concluded that "daily frequency distributions and extremes match <station data> well for a range of climates across the domain". A more detailed discussion of the VEMAP data can be found in the VEMAP documentation.
Page 9 line 32-34, In an appendix, provide a complete time series graph and a complete cumulative frequency histogram (empirical distribution function) of the historical one-day and historical 4-day precipitation values for at least one select site in both regions (40-year GLR and 25-year NER). On the time series graph, show the benchmark as a horizontal line. On the cumulative histogram clearly identify show the benchmark event on the upper tail.	Disagree: We agree this would be a nice figure to illustrate the methods employed. The Methods section has been extensively re-written to be more clear. Thus, we have not included this new figure.
Page 9 last line of footnote, Define "small event" and "large event". Does this refer to rainfall intensity? To rainfall duration?	Agree: Text has been added to the footnote in Section 2.2 to clarify this refers to small/large intensity events.
Page 10 line 7, Please explain why VEMAP grid locations were "weighted" according to the number of CSSs. What did this accomplish?	Agree: Text has been addedto Section 2.2 better describing the methodology; the weighting allowed estimation of the aggregated, regional regional impacts in each study area.

Page 11 line 10, Change "…percent change…" to "…percent increase…"	Agree: Text changed to "increase"
Page 11 line 27, What is the location of the 10 communities projects expected to decreases in CSO frequency? Are they in the same geographic region?	Disagree: The goal of this study was not to focus on individual systems or variation within regions. Rather, the goal was to present general, region-wide averages. Thus the change was not made.
Page 11 line 33: Figure 34 should be Figure 4.	Agree: Text revised to Figure 4.
Table 3 and 4, Make the titles consistent.	Agree. Title of Figure 4 was revised to be consistent with that of Figure 3.
Page 15 lines 6-19, Mention sooner that rainfall intensity will be used as a surrogate for rainfall volume (see suggestion about the recommended flowchart in the general comments section)	VEMAP data is daily depths, or mm/day. Text has been added to clarify what precipitation data was used in the study.
Tables 5,6 and Figures 10,13 Perform a statistical test (perhaps Kolmorgorov-Smirnov) to determine whether or not the differences between the 1-day and 4-day intensities is significant.	Disagree: Although an interesting point, comparison of 1-day and 4-day results was not a goal of this high-level screening study. Thus the additional analysis was not conducted.
Page 16 lines 15-20, The shifting of 1-day compared to 4-day distributions is really not very important compared to the change from the benchmark. Stay focused on what is relevant.	Agree. Text describing the subtle and non-quantitative differences in the skew of distributions for the 1-day versus 4-day events has been removed to better focus the discussion.
Page 18 line 13: Figure 12 should be Figure 13.	Agree: Figure reference has been revised
Page 19 Line 7, This should be a new section (say Section 3.3) and not a sub-section under CSO Benchmark Event Intensity.	Agree: The heading has been revised as suggested
Page 19 line 13, Define "runoff scaling factors".	Agree: Text now describes runoff scaling factors as multipliers on the changes in intensity used to estimate runoff volume.
Page 20 line 18, Believe "4 percent" should read "8 percent"	Agree: Text has been revised to say between "about 8 and 32 percent"
Page 21 lines 9-15, The experimental design is unable to determine if the differences in GCM behavior are due to differences in the two time periods or to differences in the two geographic regions. This is a weakness which limits the usefulness of the final results.	Agree: The inconsistency in planning horizons used for the two study regions is added to section 3.2.3 as a limitation of the study.
Page 22 lines 28-37, The points mentioned here can also be illustrated to some extent with the flowchart (see general comments section)	Agree: The text on page 22, lines 28-37, on limitations of the use of daily data in this analysis has been added to Section 2.2, the section describing the study methodology.
Page 23 lines 1-9, What does the historical record actually show about snowmelt events?	Disagree: The daily data we used in this analysis does not distinguish between rain and snow events. Similarly, no data is available regarding snowmelt events. It is clearly acknowledged that this is a limitation of the study. No additional text has been added.
Page 23, lines 15-20, The time periods should have been synchronized sooner.	Agree: This discrepancy results from these studies being conducted at different times and combined after-the-fact into this single report.
Page 24, Other issues that could have a large bearing on the CSS projection simulations include the adoption of BMPs that favor on-site or source control of runoff. For instance, rain gardens (virtually unknown a decade ago) are now a fashionably popular option for on-site stormwater management.	Agree: The assessment of BMP effectiveness in CSO mititgation is identified as an important future research need.

Page 25 lines 33-42, These comments are not “Conclusions” that can be drawn from this study; they seem more to be either “Observations” or “Recommendations”.	Agree: Text has been removed from "Conclusions" and added to Section 3.4.2, Future Research, where it is more appropriate.
Page 26 line 7, What is the time period for the NYC $12-$40 Billion price tag?	Agree: Text has been removed by request from NY DEP.
Preface, line 21-22 states “water treatment infrastructure” to me this implies potable water treatment, not the wastewater collection and treatment infrastructure associated with CSO.	Agree: Text revised to say "wastewater treatment infrastructure"
Page iii, line 1 – “quality of water available to meet…” (also on page 1 same line #)	Agree: Text revised; "to" added.
First paragraph of the executive summary. There is not a clear link from climate changes, impacts on ambient air temperature, and precipitation patterns, and the last sentence of the paragraph – “Future changes in climate could thus impact water quality management.” I think this key paragraph needs to be more carefully written to make the connection between altered climate and water quality more explicit. This could be as simple as a listing of examples – altered precip. may alter runoff impacting stormwater pollutant loading, CSO, SSO, etc. (same comment for the Introduction, page 1)	Agree: Executive summary and introduction have been extensively re-written.
Overall the first three paragraphs of the Executive Summary are not well integrated. It reads as a series of statements not well coordinated. (same comment for the Introduction, page 1)	Agree: Executive summary and introduction have been re-written.
Page 3, lines 4-6. The authors use the phrase “…results in significant health benefits…” this might be replaced with “…were linked to significant reduction in waterborne disease outbreaks…” I think this indicates disease reduction to be the specific health benefit.	Agree: Text in Section 1 revised to say "were linked to significant reduction in waterborne disease outbreaks"
Page 3, line 14 – recommend changing “…water treatment plant…” to “… wastewater treatment facility…”	Agree: Text "water treatment plant" changed to “…wastewater treatment facility…”
Page 3, lines 26-27 – reference for this study needs to be cited in the text and listed in bibliography	Agree: A citation has been added, (US EPA 2004)
Page 6, lines 4-6. I think the authors are missing part of the picture. They reference the changes in extreme precipitation in terms of water quality impacts from stormwater. The extreme events are likely the most important consideration for CSO and stormwater when considering the acute effects from pollutant inputs. But, many effects are chronic and are based on long-term loading of pollutants (e.g., suspended solids, nutrients, etc.) that may not produce immediate water quality degradation but over a period of years may produce reduce water quality. I think this concept is missing in this statement and in other parts of the report and analyses. I think it may be important and should be mentioned by the study authors as a sidebar related to water quality, yet not key for CSO. This is related to the suggestion by MMSD to consider stormwater pollution – long-term impacts.	Agree: We recognize the importance of longer-term, chronic water quality impairment associated with stormwater and other pollutants. The focus of this report, however, is on the specific case of episodic CSO events. Text has been added to Section 1.4 (Objectives of this research) stating that "it should be noted that although the focus of this report is on CSO events, other sources of water quality impairment also can impact aquatic systems. Understanding the potential impacts of changes in CSO frequency on aquatic systems will require watershed based approaches to determine the effects of multiple water quality stressors across a range of watershed spatial and temporal scales." Similar text has also been added to the Conclusions section.
Page 6, line 31 - recommend changing ”…investment in wastewater treatment.” to “…investment in wastewater collection, storage, and treatment infrastructure.” It is the collection and storage that needs expanding for most current CSO control technologies being implemented (tunnels, etc.)	Agree: Text added to Section 1.4 stating "...investment in wastewater collection, storage, and treatment infrastructure"

Page 9, lines 1-7 – recommend mentioning the spatial downscaling to accompany the description of the temporal downscaling. Both produce significant uncertainty. This is mentioned in other parts of the report, but not at this location.	Agree: Text revised to be more clear about development of VEMAP data set; reference given to location where detailed information about the VEMAP data set can be obtained.