Black text: B. Gause's comments
Red text: S. Markstrom's response
I added numbers to the comments so I could more easily refer to them.

Major comments:
I encourage the authors to improve the readibility of the abstract to present the idea of this study in a clearer way.

Yes, I accept your specific comments below related to the abstract. In addition, I have rewritten much of the text.

1. Please think about the use of the notation "objective function" for mean, CV,... . In my understanding, these are statistical values describing different model outputs without giving information of the model performance. The use of the term "objective function" indicates an evaluation of the model performance according its common use in hydrological modelling. I propose to use "fundamental daily streamflow statistics (FDSS)" as mentioned in the text instead of "objective function".

Accepted. Yes, I agree that confusion may arise from the non-standard use of "objective function." I changed to the term "performance measure" as I want to empahsize that this is really a measure of how the model preforms in relation to the parameter values. Also, "preformance measure" seems to make the text flow better.

2. A table with the model parameters and their corresponding processes is missing. I see that you refer to another article. However, this manuscript would be more readable, if the reader has an idea of the parameter used for this study. When stating that a certain number of parameters is required "to account for 90% of the parameter sensitivity" is necessary to know how many parameters for this process are included in the model structure. For example, assuming that there are only two snow parameters, then it is not surprising when the number of required parameters is two. However, let's say that are eight parameters for the snow process then it is interesting to know that only two parameters are required.

Yes, I added a table (table 1) that lists all of the calibration parameters, description, and what I called "PRMS module type". This PRMS module type is what I believe you are asking for in your comment. I did not want to call this "process" because I did not want to confuse the reader with the sensitivity analysis based "process identification" the is performed on the PRMS output.

Now, table 2 does show the parameters and the corresponding identified processes, but this was determined by the sensitivity analysis. Processes identified here make no a priori assumptions about which parameters may affect any particular process. For instance, PRMS uses a potential evaportanspiration coefficinent parameter. Clearly, this parameter can be directly associated with the "transpiration process", but to what degree is this parameter associated with the "snowmelt process"? PRMS does simulate snow sublimation, but a priori, should the potential ET coefficint be considered a "snow melt parameter"? Because of the unknown relationships in model stucture, this must be determined with the global parameter sensitivity analysis, and that is the point of table 1.

3. Furthermore, in chapter 4.2, you should mention whether the parameters (accounting for 90%) are identical for a certain process or vary (P. 10, L.5-6).

That is the information I try to convey in table 2. The problem is that spatially (on an HRU-by-HRU basis), which specific parameters make up the 90% could vary. Table 2 sumarizes this across all HRUs for the CONUS. The idea that I was trying to get across is that the number of parameters needed to characterize a process is some measure of the complexity of that process, and that complexity varies by process and spatially by region of the CONUS. Table 2 summarizes this in a general way so that PRMS modelers could have some idea about which parameters to actually use in their models.

To address this, I added he percent of the CONUS HRUs in which that parameter is part of the set that accounts for 90 percent of the cumulated sensitivity on an HRU-by-HRU basisvalue to the parameter names listed in table 2. I hope this addresses comment 3 by showing which ones vary the most.

4. It is really interesting to see a systematic in the number of parameters as stated on P. 10, L.20-23. Could you explain it? At best in relation to the model structure? Are you expect a different result for different models (structures)?. While this result is reasonable for snowmelt, it is really surprising that you only need a small number of parameters to explain the soil moisture behaviour.

Yes, I added the sentence: “An analyis of theses parameter counts and how they relate to their respective process is beyond the scope of this article, but it could relate to the structure of PRMS and possibly indicate that some processes are overparameterized.”

5. I think that the article would benefit if you could relate the results (e.g. P.10, L.24-30) to the process heterogenity in the different parts of the CONUS. There are certainly regions with very complex process patterns and other with a clear dominance of a single process. Are there other studies looking at process dominance or process heterogeneity in the CONUS? Maybe you can make a comparison with these studies?

I am unaware of studies which classify watersheds (regions, HRUs, etc.) necessarily by process (e.g. “snowmelt watersheds”). Some studies that I am aware of tend to classify space by mappings of soil, geology, vegitation, etc. or properties of driving climate data. These tend to use a principle components type analysis, so there are distinct classifications, but these classifications can not necessarily be related to a dominate process. Other studies tend to be based on streamflow statistics for dendritic grouping. This method seems to be effective for clasification, but not necessarily classes that are associated with obviously identifable processes.

6. It is certainly required to discuss the relationship of model parameters and the corresponding processes. The stronger this relationship is, the more sensitive a parameter might be for this process. Could you mention how the parameter-process relationship affect your results?

Yes, the other reviewer suggested that I focus more on “parameter identification” and “process identification.” I think this is related to your comment here. I rewrote the Introduction with this in mind.

7. By summing up the first-order partial variance and using this value as indicator to estimate the dominant process, you do not consider the parameter interactions (second and higher order sensitivities). However, the parameter interaction depends (among others) on the parameter selection. Could you explain how this aspect affect you results?

Yes, to section 3, I added: “An important caveat is that these higher order variances are not accounted for in the analysis. It is assumed that first-order partial variance is sufficient to identify sensitive parameters. This same assumption, as applied to process identification, may be more problematic. If there are sets of interactive sensitive parameters that have not been identified, then the associated process(es) will not be identified as such.”

8. The interpretation of table 1 needs to be reworked. I do not agree at least with the sentence on P. 11, L.16-18 that a count of dominant parameters shows how important a parameter is. Assuming that a parameter is strongly related to a certain process, e.g. snowmelt, and is thus relevant for the three objective functions related to snowmelt, but not to the other processes (maybe except of runoff), it is still an important parameter for this specific process. This interpretation and also of the fig. 5 aggregates the results in my opinion in a strong way. It might be more interesting to look at the relationship of model parameters to the processes. To how many processes you can related a parameter? Are these results reasonable when looking at the model structure? An idea of how to relate model parameters and corresponding processes is given in the figures and tables in Pfannerstill et al. (2015).

Yes, I think the problem is my use of the word “important.” This is not the right word. I have rewritten these sentences. Hopefully it is clearer. Figure 5 does show how many processes are identified (related to) a parameter. I hope that my rewritten description makes this issue clearer.

9. Concerning the discussion of the spatial heterogeneity in parameter sensitivity (subchapter 5.1), it might worth looking at the expert knowledge on dominant processes in the CONUS. It is not surprising when a HRU with a complex hydrological situation with relevant contributions from different runoff components provides a different results as a HRU with a strong dominance of one hydrological component. Here, I think that a general discussion of process dominance is missing and a discussion in the context of former studies on dominant processes in the CONUS (if existing).

See response to comment 5 about other studies.

10. Maybe you can think about presenting the results in Tab. 1 and Figs. 4 and 5 in a different way, so that the most important outputs are more emphasized. It is rather difficult to extract information of the relationship of parameter and processes from Tab. 1 and a counting how often a parameter occurs is also time-consuming. But in my opinion this information is required to make Fig. 5 more informative.

I'm not sure how to do this. The most important outputs, in my opinion, are to give the modeler versions of the table and figures exclusively for the area that he is modeling. And I have been doing this for the people that I work with. For this article, the problem is that I have to keep it general for all of CONUS.

Fig. 4: Is it maybe relevant thinking about the variability, e.g. in the snowmelt subplot? It is stated that on average 2.25 parameters are required to explain 90%. The map (subplot 4M) shows that in most of the HRUs only 2 or 3 parameters are required. However in the snow-dominated northern parts up to 10 parameters are required. It might be worth thinking about extracting additional information from this idea. One way would be to add an additional line in the subplots 4A-4H which is only related to HRUs which have certain relevance of this process (kind of threshold exceedance approach or something similar).

I believe that I have addressed this issue of HRU parameter variability in table 2 and the text I added in relation to figure 5.

11. Fig. 6: Could you explain why infiltration is the inferior process in many HRUs. I cannot imagine a hydrological situation in which the infiltration process is less relevant than total runoff, all runoff components, ETP, soil moisture.

It's not that infiltration is not important, it's just that the sensitivity analysis indicates that there are no parameters that can be changed to affect the model output. Also, there are often multiple processes that are pretty much at the same level of "inferiorness" and one has to be the most. In a very preliminary draft I had version of these maps that showed, for each HRU, the two most infereior process, the three most, etc. These maps really confused my co-authors and in the end, I dropped them.

12. It might be interesting to think about the following results of the Fig 4-5: According to Fig. 4 only 4.15 parameters are required to explain soil moisture, which is a relative low value keeping in mind that the soil moisture interacts with almost all other processes. Furthermore, there are 7.05 parameters needed for infiltration. Then, it is stated in Fig. 5 that soil_moist_max is overall the most important parameter. Do this mean that the relationship between soil_moist_max and soil moisture is extremely high so that only a few additional parameters (about 3) are needed to reproduce the soil moisture conditions?

Yes, I think this interpretation is correct. A source of confusion could be my use of the word "important." In retrospect, that is a loaded word. See my response to your comments number 8 and 11.

Minor comments:
Abstract:
Page 2, Line 2: The first sentence of the abstract could be written more clearly. Why not only writing: "The Precipitation-Runoff Modeling System as a distributed-parameter

hydrologic model has been applied to the conterminous United States.

Yes, accepted.

P. 2, L. 4-5: Whilst it is certainly clear that the number of parameters is an aspect of model complexity, this is not fully clear for the "interpretation of the model output“. Is this really an aspect of complexity? Do you assume that the model which provides a higher number of model outputs is more complex?

Yes, rewritten. I’m trying to establish the point that by identifying the dominate processes (with respect to PRMS), users can focus on the output variables related to those processes.

P. 2, L. 5-8: To make the abstract more readable, I would suggest to subdivide this sentence into two separate ones. There are too many aspects in this sentence (parameter sensitivity for simplification, parameter identification and its relationship to dominant processes, spatial patterns)

Yes, accepted.

P. 2, L. 9-10: I do not think that this sentence is understandable when reading the abstract at first before knowning the whole article. What do you mean with "processes correspond to variables“? Which type of variables?

Yes, changed this sentence.

P. 2, L. 11: The notation "categories“ is not clearly described in the abstract.

Yes, changed.

P. 2, L. 12-13: How do you estimate the "model performance“ by visualizing categories? This part needs to be improved.

Yes, changed.

P. 2, L. 16: The benefit of a reduction of the dimensionality of output variables or objective functions is not clear.

Yes. changed

P. 2, L. 22: I would encourage the authors to add a final sentence to emphasise the general advantage of this study.

Yes, added.

Introduction:
P. 2, L. 28: The article would be benefit from a clear definition of "input parameters“. Is an input parameter related to a driver of the hydrologic cycle such as precipitation or solar radiation or more to a real model parameter? In all cases, it is better to avoid potential misunderstandings.

Yes, added.

P. 3, L. 1: References are missing such as for constraining parameter in models, e.g. Hrachowitz et al. (2014) and for stating that different parameter good have a comparable impacts on the model results.

Yes, added.

P. 3, L. 6: The three references are related to studies which investigate performance measures more precisely. It might be good to also have a reference to studies which are directly investigating the model output.

Yes, added.

P. 3, L. 11-12: Please also add the study from Reusser et al. (2009).

Yes, added.

P. 3, L. 14: Please indicate that you consider uncertainty in this study only on input parameter uncertainty and not on structural uncertainty in the model.

These lines were deleted in response to comments by another reviewer.

P. 3, L. 18-28: It might be good to mention here that it is at least at this scale impossible to support the results with adequate measurements in addition to the total discharge.

These lines were deleted in response to comments by another reviewer.

P. 4, L. 1: References are here missing, e.g. Wagener et al. (2003), Reusser et al. (2011), Guse et al. (2014).

Yes, added.

P. 4, L. 11: Reference of Reusser et al. (2011) is missing.

Yes, added.

P. 4, L. 20-22: As mentioned before, it is not clear why you aimed "to reduce the number of inputs and outputs“. I think the overall aim should be a clearer characterization of the model parameters and to focus on the dominant processes.

Yes, I reworded this sentence.

Methods:

1. P. 4, L.29- P. 6, L.7: Please check carefully if you could reduce the subchapter 2.1 in length. Do you really need this information for this article?

Yes, this section has been reorganized.

P. 6, L.8-25: The selection of the eight output variables is reasonable and seems to be representative for hydrological studies with distributed models. Maybe you can emphasize this to give the article a more general character.

Yes, added.

P. 7., L. 18: Please also add the reference of Guse et al., 2014, since it is the initial study for Pfannerstill et al. 2015.

Yes, added.

Results:

1. P. 8, L. 17: Please think about a more precise title for the subchapter 4.1.

Yes, changed it to “Parameter sensitivity by process and performance measure”

2. P. 8, L. 20-23: This sentence is not understandable. It is understandable that you have calculated the sum of the first-order partial variance. However, it is not clear how you can estimate an average value (average of what?).

Yes, the meaning of the text here is not clear to you. I have added several sentences here to make this clearer.

3. P. 8, L. 23: The total sensitivity is one, is it? Why do you need to scale the sum of the sensitivities to the total sensitivity?

The sum of the individual sensitivities is not necessarily one. If none of the parameters are sensitive then the sum of the parameter sensitivities will be closer to zero.

4. P. 8, L. 23: "category of modeled process“ instead of "category of process“.

Yes, accepted.

5. P. 8, L.28-30: I recommend to be more precisely here: You have calculated the sum of all partial sensitivities for a certain HRU for each process. Then, the process with the highest sum of the first-order sensitivity is indicated as "dominant process“. To make this clear, you should add that the dominant process is the process with the largest sum of all first-order partial variances (sensitivities). This is required since the sensitivity of a single parameter is not shown here.

Yes, reworded these sentences.

P. 9, L.17-18: Can you extract a systematic pattern in these results?

Yes, added ", and humid versus arid climates." to the previous sentence.

P. 10, L.24-25: Please add that this statement is not valid (or only to a low extent) to fig 4J and 4N.

Yes, added this.

P. 11, L. 6-9: Do you see a general systematic why the spatial patterns of parameter sensitivity are different for the different objective functions. It might be interesting to give further statements on this.

There are certainly patterns here and I very much agree that they are interesting. I have not had time to investigate this properly and would prefer to leave statements about this out of this article rather than speculate.

There is clearly a swath of sensitivity that goes through the Great Plains. Many hydrologic modelers in the US have noted that this area is notoriously difficult to model with physical, statistical, etc. models – and no one is realy sure why this is. Our group has a PhD student who is looking into this. Maybe a subsequent article can address this further.

P.11, L. 28-32: When stating that the parameter "soil_moist_max“ is the most important and a model calibration should be focused on it, then it is required to know for which process this parameter is relevant. Assuming that a typical calibration uses discharge as target variable, a focus on "soil_moist_max“ helpful in the case of a dominance of "soil_moist_max“ on runoff. However, to include this information in a calibration in the case of a dominance on other process but not on runoff?

Yes, I rewrote this paragraph based on comments from the other reviewer. I believe my revision addresses this comment as well.

P. 12, L.2-8: The part on the least sensitive parameter can be removed since the reader does not receive any details about the parameters. Or could you extract some further information from the fact that these parameters have a low sensitivity?

Yes, I now say that modelers should leave them at default values because there is limited information to calibrated them.

P. 12, L. 9-14: I think that the authors should add here some more details. It is really helpful if a parameter can be precisely characterized by saying that it is only dominant in a very specific case (e.g. for one process). But this information cannot currently not be extracted from article.

This varies by HRU/geographic region, so it is difficult to provide specific calibration instructions for the whole of the CONUS. I do provide exactly this type of information on an application site by application site basis to the modelers that I work with. I'm uncertain how to put this information into this article.

P. 13, L.8-12: I like this part. Maybe you can in addition relate it to the concept of vertical water redistribution (Yilmaz et al., 2008, Pfannerstill et al., 2015).

Yes, I added a sentence about this.

P. 14, L. 22-23, Step 1: Summed in time?

Yes, added.

P. 14, L. 24-25, Step 2: How to you obtain a score for each process? Do you assign each parameter to a certain process? If yes, then you have to mention somewhere which parameter is related to which process.

Please see my response to your comments 2 and 3 ("Major" comments section), and 2 and 3 in the "Results" comments section.

P. 16, L. 31: Spelling error: Mishra (2009)

On recommendation of other reviewer, I removed this paragraph.

Figures:
Fig. 1: Could be removed. I do not see an advantage of it. Maybe you can transfer it to the supplementary material.

Yes, removed.

Fig. 2: Does the last row and column present the average values along the row/column? Do you maybe have to change "process average“ and "objective function average“?

Please see response to "Major" coments 2 and 3.

I recommend to show the figure 3 before the figure 2, since fig. 3 provide a general map of the USA whilst, fig. 2 already show the distributed results.

Yes, moved figure 3 to figure 1 (after deleting old figure 1).

Figure 4 would benefit from knowning which parameters are within the 90% and how variable the parameters belonging to this 90% are?

Yes, see my response to comment 10.

Fig. 4: The legend needs to be graphically improved.

Yes.

I do not really see a real benefit of fig. 5. Maybe you can extract the results in a better way. One point might be that the model parameters are not explained and even the related processes are not highlighted in Fig. 5. In particular, it is not clear which information you can derive from the last place occurrence.

Please see my response to your comment 8.

It is not fully clear which information you can derived from investigating the most inferior process. It seems to be that this is either clear such as snowmelt parameter for California or related to the model structure.

The idea here is that modelers should not calibrate parameters associated with inferior processes in their watershed. If there are 35 calibration parameters, make sure to include the ones associated with the more dominate processes, and exclude the ones associated with the more inferior ones. I hope this idea comes across in the article.

Reference list:
Guse, B., Reusser, D. E., and Fohrer, N.: How to improve the representation of hydrological processes in SWAT for a lowland catchment – Temporal analysis of parameter sensitivity and model performance, Hydrol. Process., 28, 2651–2670, doi:10.1002/hyp.9777, 2014.

Hrachowitz, M., O. Fovet, L. Ruiz, T. Euser, S. Gharari, R. Nijzink, J. Freer, H. H. G. Savenije, and C. Gascuel-Odoux: Process consistency in models: The importance of system signatures, expert knowledge, and process complexity, Water Resour. Res., 50, doi:10.1002/2014WR015484, 2014

Pfannerstill, M., Guse, B., Reusser, D., and Fohrer, N.: Process verification of a hydrological model using a temporal parameter sensitivity analysis. Hydrology and Earth System Sciences 19: 4365–4376, 2015.

Reusser, D. E., Blume, T., Schaefli, B., and Zehe, E.: Analysing the temporal dynamics of model performance for hydrological models, Hydrol. Earth Syst. Sci., 13, 999–1018, doi:10.5194/hess-13-999-2009, 2009.

Reusser, D.E., and Zehe, E.: Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity. Water Resources Research 47(7): W07550. DOI:10.1029/2010WR009946, 2011.

Wagener, T., McIntyre, N., Lees, M.J., Wheater, H.S., Gupta, H.V.: Towards reduced uncertainty in conceptual rainfall–runoff modelling: dynamic identifiability analysis. Hydrological Processes 17: 455–476, 2003.

Yilmaz, K. K., Gupta, H. V., and Wagener, T.: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, Water Resour. Res., 44, W09417, doi:10.1029/2007WR006716, 2008.

Thank you for this reference list. I added citations to all of these references.