January 22, 2019

Re: Resubmission of manuscript *Identifying ENSO Influences on Rainfall with Classification Models: Implications for Water Resource Management of Sri Lanka*, Ms. No. hess-2018-249

Dr. Wouter Buytaert
Editor, Hydrology and Earth System Sciences

Dear Dr. Buytaert,

Thank you for giving us the opportunity to revise our manuscript, *Identifying ENSO Influences on Rainfall with Classification Models: Implications for Water Resource Management of Sri Lanka.* We appreciate the careful review and constructive comments. We believe that the manuscript is improved after making the proposed edits to the introduction, methods, results, and discussion and after correcting writing mistakes.

Following this note are the reviewers comments in the blue color, followed by our comments explaining how and where the text was modified. Revisions were made in consultation with all the authors, and each author has approved the revised manuscript.

Thank you for your consideration.

Sincerely,

Thushara De Silva M.

Referee #1

General comments:

This study demonstrates application of different classification models to predict monthly rainfall using climate indices MEI and DMI. The findings of this study will be highly relevant for water managers of Sri Lanka. Below are my comments for Authors. Some of the similar comments I found that Authors have already addressed in the Discussion forum but please address all the points below.

Response: Thank you for identifying the importance of the study. Shortcomings of the paper identified by the reviewers will be rectified as suggested in the paper. Response to the detailed comments are given below.

Specific comments:

> 1) (a) Line 95 says that the river basin rainfall was calculated using the Thiessen polygon method. Why not divide the basin into sub-basins (using any GIS tool) based on digital elevation model and use sub-area averaged rainfall? Is this choice due to the fact that there are many reservoirs in the basin? Please clarify.

Response: We added the sentence below to clarify this point.

Since this study does not aim to explore rainfall across sub-basins, we do not use digital elevation maps to define the sub-basins.

(b) Out of 16 polygons in Mahaweli river basin and 11 in Kelani, what was the basis of selecting only 8 sub-basins?

Response: We revised the text as indicated below.

Considering the importance of sub-basins for the reservoir catchment and for water use, eight sub-basins are selected for analysis. Morape, Randenigala, Peradeniya, Manampitiya and Bowatenna represent the Mahaweli major reservoir catchments and irrigation tanks, and Norton Bridge, Norwood and Laxapana represent the Kelani basin reservoir catchments. The catchment of the major Mahaweli river reservoir cascade (Kotmale, Victoria, Randenigala, Rantambe, Bowatenna) is represented by Morape and Peradeniya located in the humid zone and by Randenigala and Bowatenna located in the intermediate zone. The arid zone major irrigation catchments of the Mahaweli are represented by Manampitiya. The catchment of the Kelaniya reservoir cascade (Norton Bridge & Moussakele) in the humid zone is represented by Laxapana, Norton Bridge and Norwood.

2) (a) Lines 104 to 109 describe how anomalies were calculated. Did you apply any of the transforms mentioned in line 108 to get normally distributed rainfall? Some plots/results can be included to clarify the rainfall anomaly classification.

Response: We added our response to this comment in the Appendix. We think that giving too many details in the manuscript decreases readability. Figure A.1 in the Appendix is referenced in the manuscript proper.

@ Appendix

Normality Testing:

The Shapiro-Wilk's method is used to identify the normality of rainfall anomaly distribution. The Manampitiya NEM normality test results are given below as an example.

Data 1: original data

W = 0.96675, p-value = 0.08185

Data 2: data transformed by square root

W = 0.98772, p-value = 0.7772

Data 3: data transformed by log

W = 0.91577, p-value = 0.0003325

Further, from data plots (Figure A. *1*) and the S-W statistic, we conclude that the square root transformed data is closer to being normally distributed than the other forms.

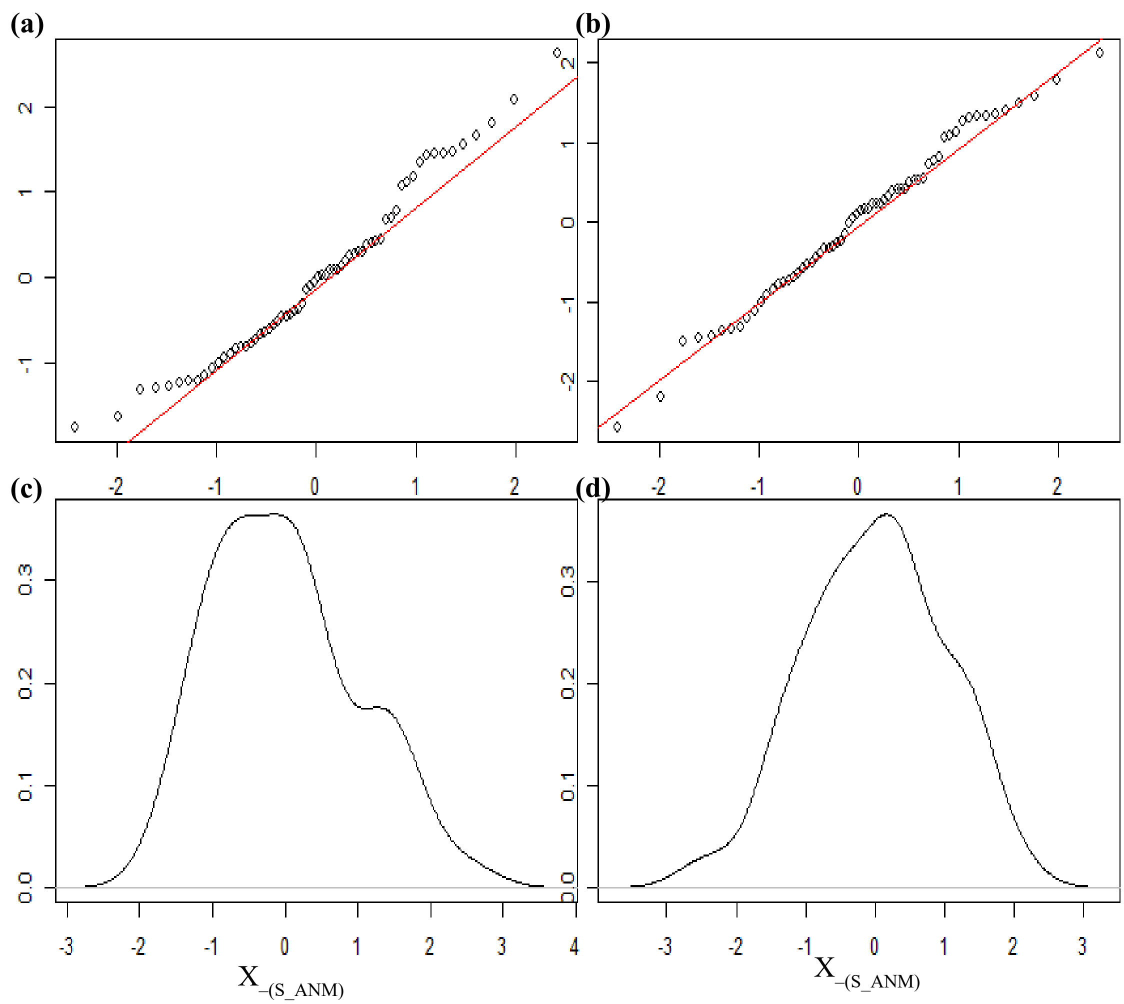

Figure A. 1: Manampitiya NEM standardized data (a) original form qqplot (b) square root form qqplot (c) original form density plot (d) square root form density plot

(b) In Table1, the use of 0.5 appears like a random choice. Please justify.

Response: We explained in the Appendix; Figure A.2 is referenced in the paper.

Extreme seasonal precipitation has been defined statistically in different ways using statistical thresholds (Easterling et al., 2000; Jentsch, Kreyling, & Beierkuhnlein, 2007; Knapp et al., 2015; Smith, 2011). We use 0.5 as a threshold to define three classes, which results in fairly evenly distributed data across the three classes (Figure A 2).

@ Appendix

Classification of data

Using 0.5 as a threshold for a normal distribution defines portions of the data that are fairly evenly distributed into three categories – about 31%, 38%, and 31% for a normal distribution (Figure A 2). We deemed this a reasonable choice for our analysis.

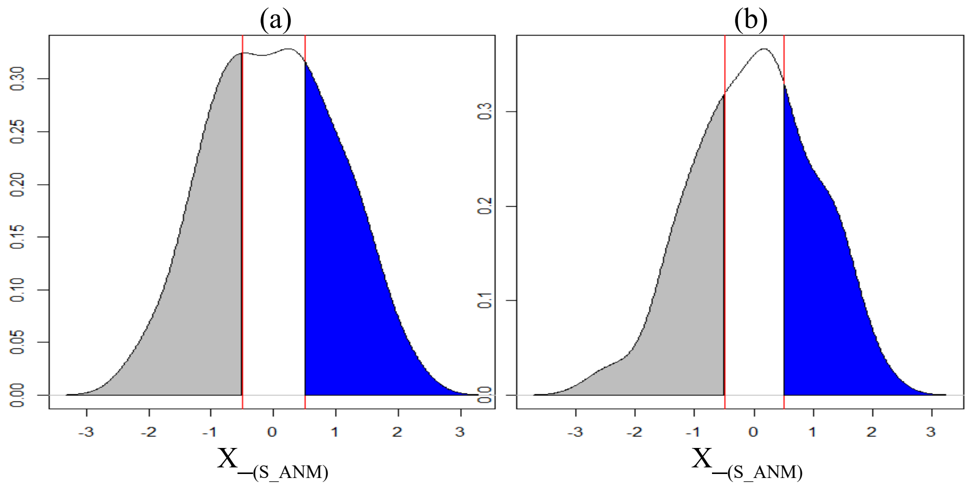

Figure A 2: (a) Norton Bridge SWM rainfall anomaly distribution (b) Manampitiya NEM rainfall anomaly distribution

3) (a) Line 122 says average of MEI and DMI were used but the figures 4 and 5 show that you have used them separately.

Response: We added below sentence to clarify the comment,

Because we analyzed the data in rainfall seasons, values of the climate indices over each season are averaged. For example for the NEM season, the MEI value is the average of January and February monthly values and for the SWM season, DMI is the average of May, June, July and September values.

(b) Authors should support the choice of MEI and DMI over several other climate indices which they could have used as predictors.

Response:

The paper was revised in several places and new results were added to the Appendix to support the choice of MEI and DMI climate indices.

2.2 ENSO & IOD Indices

The ENSO phenomenon is represented by MEI, NINO34, NINO3, NINO4 indices, and the Indian Ocean dipole phenomenon is represented by the DMI index. NINO34, NINO3, NINO4 indices are based on tropical sea surface temperature anomalies (National Center for Atmospheric Research, 2018) and the Multivariate ENSO Index (MEI) is based on sea-level pressure, zonal and meridional components of the surface wind, sea surface temperature, surface air temperature, and total cloudiness fraction of the sky (National Oceanic and Atmospheric Administration, 2017).

4 Results

Similar to other investigators, we observe several strong correlations between rainfall anomalies and the climate indices (Figure A 3, Table A 1, Appendix). Higher correlation values between MEI and rainfall anomalies can be seen compared to the correlation with other ENSO indices (Table A 1).

@ Appendix

Correlation analysis with multiple climate indices

We examined the correlation between rainfall anomalies and multiple climate indices to choose the two climate indices MEI and DMI ((Figure A 3, Table A 2). The ENSO phenomenon is represented by MEI, NINO34, NINO3, NINO4 indices. Correlation analysis indicates that MEI, which is estimated using several climate factors such as sea-level pressure, zonal and meridional components of the surface wind, sea surface temperature, surface air temperature, and total cloudiness fraction of the sky (National Oceanic and Atmospheric Administration, 2017), demonstrates higher correlation with rainfall anomalies in sub-basins for all rainfall seasons compared to the NINO34, NINO3 and NINO4. The Indian Ocean dipole phenomenon is represented by the DMI index, which represents the gradient of the sea surface temperature. Based on the correlation analysis and the content of the indices, we selected MEI as the indicator for ENSO and DMI as the indicator for IOD.

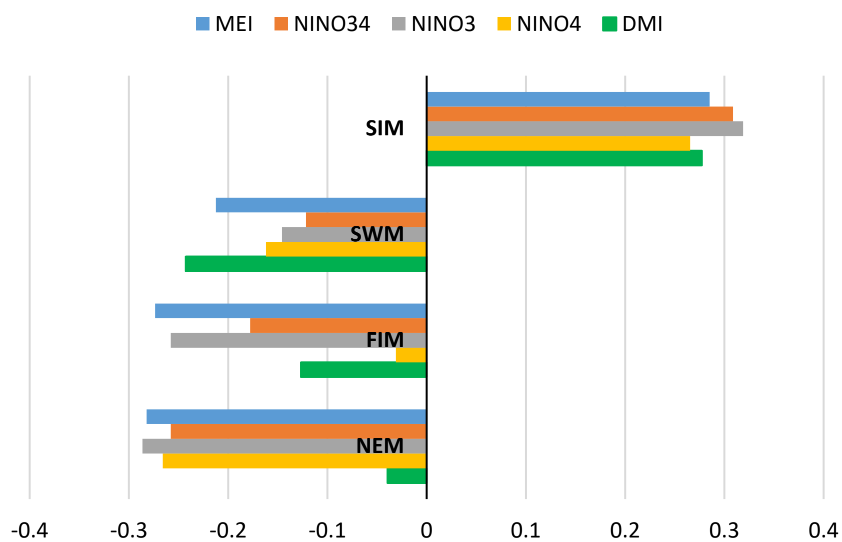

Figure A 3: Correlation between Norwood rainfall anomalies with multiple climate indices

Table A 1: Correlation analysis of rainfall anomalies and climate indices

Rainfall	**Morape**					**Peradeniya**				
Month	MEI	NINO34	NINO3	NINO4	*DMI*	MEI	NINO34	NINO3	NINO4	*DMI*
NEM	-0.35	-0.35	-0.34	-0.38	*-0.09*	-0.38	-0.40	-0.39	-0.42	*-0.11*
FIM	-0.28	-0.19	-0.28	-0.07	*-0.11*	-0.27	-0.18	-0.30	-0.06	*-0.06*
SWM	-0.35	-0.24	-0.23	-0.26	*-0.29*	-0.35	-0.26	-0.25	-0.27	*-0.31*
SIM	0.21	0.23	0.27	0.19	*0.12*	0.17	0.19	0.21	0.15	*0.09*
Rainfall	**Laxapana**					**Norwood**				
Month	MEI	NINO34	NINO3	NINO4	*DMI*	MEI	NINO34	NINO3	NINO4	*DMI*
NEM	-0.27	-0.26	-0.28	-0.27	*-0.01*	-0.28	-0.26	-0.29	-0.27	*-0.04*
FIM	-0.28	-0.16	-0.27	-0.03	*-0.07*	-0.27	-0.18	-0.26	-0.03	*-0.13*
SWM	-0.3	-0.23	-0.21	-0.25	*-0.31*	-0.21	-0.12	-0.15	-0.16	*-0.24*
SIM	0.1	0.10	0.14	0.06	*0.08*	0.29	0.31	0.32	0.27	*0.28*

Rainfall	Randenigala					Bowatenna				
Month	MEI	NINO34	NINO3	NINO4	*DMI*	MEI	NINO34	NINO3	NINO4	*DMI*
NEM	-0.30	-0.31	-0.29	-0.34	*-0.11*	-0.35	-0.36	-0.35	-0.38	*-0.2*
FIM	-0.29	-0.23	-0.33	-0.10	*-0.04*	-0.23	-0.17	-0.25	-0.09	*-0.02*
SWM	-0.17	-0.12	-0.09	-0.18	*-0.24*	-0.18	-0.09	-0.05	-0.11	*-0.12*
SIM	0.37	0.38	0.41	0.36	*0.35*	0.35	0.41	0.40	0.40	*0.36*
Rainfall	**Norton Bridge**					**Manampitiya**				
Month	MEI	NINO34	NINO3	NINO4	*DMI*	MEI	NINO34	NINO3	NINO4	*DMI*
NEM	-0.32	-0.30	-0.33	-0.33	*-0.01*	-0.26	-0.28	-0.26	-0.28	*-0.16*
FIM	-0.18	-0.12	-0.21	-0.01	*-0.08*	-0.2	-0.17	-0.31	-0.06	*-0.14*
SWM	-0.31	-0.22	-0.21	-0.22	*-0.37*	-0.07	0.08	0.08	-0.01	*-0.03*
SIM	0.02	-0.02	0.03	-0.04	*-0.15*	0.45	0.46	0.44	0.46	*0.51*

4) 64 years of historical data have been used, 75% of which are used for training and rest for testing model performance. If I understood properly, there is no demonstration of season-ahead forecast of rainfall and how those can be classified as dry or wet, the information useful for water managers. Authors write about forecast in Lines 240 to 246, but there is no assurance of enhancement in future skill using the three classification models used in this study.

Response: As indicated in the Lines 253-258; ENSO forecasts are available from the International Research Institute for Climate and Society (International Research Institute, 2017b) and IOD forecasts are available in the Bureau of Meteorology (BOM), Australian Government (Bureau of Meteorology, 2017). We do not know of the availability of archives of MEI and DMI past forecasts that could be used to evaluate the skill of the season-ahead forecasts of rainfall. Thus we could not evaluate the degree to which our analysis would improve overall skill.

5) Water managers will be mostly interested in extreme events. Would it be possible to obtain information about extreme dry or wet season/months from the three classification models used here?

Response: We added the sentence below and Table 5 to clarify.

The QDA method produces results that are promising with respect to identification of extreme dry events as indicated by seasonal rainfall (Table 5).

Table 5: Classification results for extreme dry (very low rainfall) and wet (very high rainfall) seasons.

Class	Range	NortonBridge SWM		Manampitiya NEM	
		tree	QDA	tree	QDA
Very dry	$X_{S\_ANM} < -1.0$	10/11	10/11	6/11	11/11
dry	-1.0 <= $X_{S\_ANM} < -0.5$	9/11	6/11	5/11	9/10
average	-0.5 <= X_{S_ANM}<0.5	8/22	9/22	9/25	11/25

wet	0.5 <= X_{S_ANM}<= 1.0	5/11	5/11	1/5	0/5
Very wet	1.0 <= X_{S_ANM}	6/11	6/11	7/11	1/11

Technical corrections:

(i) There are nomenclatures like dry and wet which are used for dividing the zones and also for classifying rainfall anomaly (see Lines 100, 160 to 177). It would be better if Authors can use different nomenclature.

Response: We changed our terminology for the three climatic zones to arid, intermediate and humid (instead of dry, intermediate and wet).

(ii) In Figure 4, caption of part (d) is missing.

Response: Caption of Figure 4 is corrected to include

(d) Manampitiya NEM rainfall classification by QDA

Referee #2

General comments:

This paper examines the use of climate indices to predict a high or low rainfall period. Classification tools are used for this. The results that are obtained are not very impressive, but the authors argue that, for the local farmers and water managers, this will still be of value, which is a fair point. So while the scientific interest of this paper is limited, it has some clear practical value. The paper in its current form suffers from: (i) an insufficiently detailed presentation of the methodology which would not enable a reader, even in principle, to understand how the methods work unless the reader had prior knowledge of them; (ii) a strange organisation of the material so that the presentation of the study area is given under a 'methods' section for instance. This may be because the authors seem to be wedded to organising the paper according to some standard headings: methods, results, discussion, etc. But this is not always helpful and here, as with the obligatory conclusion section which is just a repetition of material just above that section, I would urge the authors to feel free to adapt the structure to their needs.

Response: The objective of this study is to explore the climate teleconnections to monsoon and inter monsoon seasons and provide information for water resources managers and water users (farmers and hydropower producers) that indicate how forecasted climate indices may be useful to them. Specifically in Sri Lanka presently no seasonal precipitation forecast is available so. As the Reviewer acknowledges, even an imprecise forecast from our classification results may indeed be useful. In response to the Reviewer's comments, we made several modifications to the paper.

(i) The Methods section was revised. Explanations with new information, including additional references, were incorporated.

(ii) As suggested we changed the structure of the paper.

Detailed Comments:

1. something missing in this sentence, perhaps ‘to’ before ‘climate variability’ (without ‘the’) (page 1, line 21)

Response: This sentence was corrected.

Results from these models are not very accurate; however, the patterns recognized provide useful input to water resources managers as they plan for adaptation of agriculture and energy sectors in response to climate variability.

2. It is not clear here whether you are making a methodological point here. It seems that you are identifying two reasons for your methodological approach: (i) the weakness of the linear regression approach when the scatter is large and (ii) the nature of the forecasts available to water managers, which may just be of some broad category of rainfall rather than actual quantities. Based upon these two reasons, you are advocating a method based upon classification models. If that is the case, please spell this out as these are key issues for understanding your chosen approach. (page 2, line 56-61)

Response: Seasonal precipitation is generally forecasted in broad categories. For example, the US National Weather Service forecasts seasonal precipitation as above normal, below normal, and normal (National Oceanic and Atmospheric Administration, 2018). The International Research Institute for Climate and Society also forecasts seasonal precipitation as above, below and near normal (International Research Institute for Climate Society, 2018). We chose to follow a similar approach and present seasonal precipitation prediction in three classes based on ENSO and IOD. To clarify our approach we added the sentences below to the manuscript.

Seasonal precipitation is generally forecasted in broad categories. For example, the US National Weather Service forecasts seasonal precipitation as above normal, below normal, and normal (National Oceanic and Atmospheric Administration, 2018). The International Research Institute for Climate and Society also forecasts seasonal precipitation as above, below and near normal (International Research Institute for Climate Society, 2018). We chose to follow a similar approach and investigate river basin rainfall teleconnections to climate indices with classification models.

3. Secontion 2 until the middle of page 5 (the start of subsection 2.3) is not about methods. Please choose a more appropriate title for the section, such as ‘Hydrometeorology and climatology of the study area’. Subsection 2.3 can then become a section 3 entitled ‘Methods’. (page 2-5)

Response: Manuscript is rearranged as suggested by the comment. Sections of the manuscript now are as follows:

1. Introduction, 2. Hydrometeorology and climatology of the study area, 2.1 Sub-basin rainfall (Areal rainfall) 2.2 ENSO & IOD Indices 3. Methods, 3.1 Quadratic Discriminant Analysis (QDA), 3.2 Classification Tree model, 3.3 Random Forest, 4. Results, 5. Discussion, 6.Conclusion

4. I am not sure why you mention a minimum and a maximum in the table. Given that we have no idea what these might be, I suggest taking out any reference to them (so the first class is just defined for standardised anomalies below -0.5C, and similarly for the third class with standardised anomalies large then 0.5C) (page 5, line 117-118)

Response: The Table was revised. Originally, the terms maximum and minimum referred to the largest observed positive anomaly value and largest negative anomaly value respectively but the reviewer points out that these terms are misleading so they were dropped.

Table 1

Class	Range
dry	$X_{S\_ANM} < -0.5$
average	-0.5 <= X_{S_ANM}<0.5
wet	0.5 <= X_{S_ANM}

5. Capital letters are required here: 'Atmospheric' and 'Administration' as well as a comma after the latter word. (page 5, line 122)

Response: Words are corrected as "National Oceanic and Atmospheric Administration, 2017".

6. To make this presentation of the QDA clearer to someone who has, for instance, some idea of Bayesian statistics, but does not know this method, I suggest adding the quadratic discriminant function is therefore proportional to the logarithm of the a posteriori density function of class k conditional upon the value of the observed predictor x: this logarithm is the product of the prior probability of x and the density. (page 6, line 138-145)

Response: Method section was revised to include more details. References were added to direct a reader to more depth regarding methods.

In response to comments from Reviewers, Figure 3 was improved for the better readability.