

Thank you for taking the time to provide comment and feedback on our manuscript. Please see below for a list of changes made and responses to each of the comments made in the review.

	Comment	Adjustments	Section
	Introduction		
1	The first paragraph should be for agricultural drought monitoring. The second paragraph for machine learning models used to forecast agricultural droughts.	<p>Thanks for the feedback. The structure of the introduction is based on the comments made by reviewer 2 in the previous review. (see bullet point 1 of https://doi.org/10.5194/nhess-2024-60-RC2)</p> <p>We believe it would be wrong to make further changes to the introduction which would change it from which was originally requested by the second reviewer</p>	n/a
2	The second paragraph should begin with "Machine learning has been shown to outperform...", all the previous text should be in the first paragraph.	Begun second paragraph with this statement, prior text was moved to first paragraph	Page 2 section 1
3	<p>L20 "Drought are defined as an extended period in which a water deficit occurs, usually because precipitation is less than average resulting in water scarcity (Cunha et al., 2019)."</p> <p>What type of drought are you defining? In your manuscript, you are using VHI, which depends on temperature and vegetation. Thus, VHI is a drought index that does not depend on precipitation. A better definition of drought types is needed.</p>	<p>Specified that the initial definition of drought provided is that of meteorological drought</p> <p>Included a definition of agricultural drought at the end of the first paragraph</p>	End of 1 st paragraph of introduction
4	<p>L20 "Agricultural drought can have significant socio-economic impacts because they impact food security."</p> <p>However, you have not provided a definition for agricultural drought.</p>	Included a definition of agricultural drought at the end of the first paragraph	End of 1 st paragraph of introduction
5	<p>L25 ". This is because VHI, derived from AVHRR (Advanced very high resolution radiometer) data, responds cumulatively and quickly to changes in vegetation greenness."</p> <p>Remove the word "derived".</p>	Removed word derived	Paragraph 2 of introduction
6	L30 "Drought monitoring using vegetation indices such as VHI or	We have included a definition of VHI, VCI and TCI to clear up any confusion	Introduction paragraph 1

	<p>NDVI (Normalized difference vegetation index) or VCI (Vegetation condition index) has been developed in several locations using satellite imagery from products such as MODIS, and NOAA STAR (Sadiq et al., 2023; Kloos et al., 2021). "</p> <p>This sentence is confusing because $VHI = a \cdot VCI + (1-a) \cdot TCI$, and VCI uses NDVI and TCI uses BT or LST</p>		
7	<p>L30 "VHI is reported to improve on NDVI based monitoring as it provides a measure of vegetation condition relative to long term change (West et al., 2019)"</p> <p>VHI uses VCI, which in turn uses NDVI. Therefore, I am uncertain whether the authors have a thorough understanding of the VHI drought index at this point. Before these sentences, at least, we need a clear definition of VHI.</p>	<p>We have reworded this sentence to make clear that VHI is an improvement on using NDVI but is still based on NDVI.</p> <p>We have also included a definition of VHI, TCI and VCI</p>	Introduction paragraph 1
8	<p>L55 " For example, in 2020 drought in Rio Grande do Sul was estimated to have cost R\$ 36 billion in losses representing 7.36% of the states GDP (CNA, 2020). "</p> <p>For a wider audience, you should express numerical figures in dollars.</p>	Added US dollars in brackets	Section 1 page 3
Methods			
9	The study area should come first; the previous introductory paragraph is unnecessary.	Introductory methods paragraph was removed	Section 2
10	<p>L95 "...up of 9 different Köppen-Geiger climate zones from semi-arid in the northeast,..."</p> <p>For figures up to ten, should be preferred to use words rather than numbers.</p>	9 converted to words	Section 2
11	<p>L110 "The data was filtered using harvested areas from the crop grids dataset (Tang et al., 2023)"</p> <p>Which data are you talking about?</p>	Changed text to specify that all input and output data is filtered using harvested areas from crop grids	Section 2 paragraph 2
12	L160 "SPI indicators with longer accumulation periods were not	Included a new appendix E which provides a greater justification for not using SPEI 6 and 12,	Appendix E

	<p>tested because such accumulation periods would be longer than the growth periods of maize and soybean. SPI is a widely used index recommended by the world meteorological organisation (WMO). It is also used for operational drought monitoring at CEMADEN (Cunha et al., 2019)"</p> <p>The justification is insufficient to warrant the avoidance of using higher time scales. It will be worth analyzing up to a 12-month timescale.</p>	furthermore, I have also completed a correlation analysis to show that VHI is more strongly correlated with SPEI 2 and 3 than 6 and 12.	
13	<p>Why are soil moisture, total precipitation, and ERA5 included in the "Drought indices" section? These are not drought indices, but rather predictors.</p>	Changed title of section to be input variables & drought indices	Section 2.2
14	<p>L200 "Where spatial resolution has increased (spatial up sampling) this is calculated using a k-nearest neighbours algorithm</p> <p>This was nearest neighbor (k=1) or k-nearest neighbors? If the latter is the case, what was the value of k? Most GIS tools use nearest neighbors for resampling continuous data.</p>	Provided details on the value of K used for the spatial subsampling	Section 2.3 page 7
15	<p>L210 "user. random forest constructs a specified number of trees and then averages the result of each individual tree. Different trees</p> <p>After the dot (.) it should start with a capital letter?</p>	This has been corrected as part of a further proof read	Throughout
Results			
16	<p>In my previous review, I added the comment, "The study area includes crops of soybean and maize; it would be good to know how the ML model performs per crop type." But it wasn't addressed by the authors.</p>	<p>This comment was previously addressed in the reviewer comments response table point 14. We also included a justification for using maize and soybean growing area together rather than separately</p> <p>Included new appendix D which shows negligible differences between model performance when training the model on maize and soybean separately.</p>	Appendix D

17	L320 "Feature importance" You should choose to use "feature" or "predictor", but not both, because it is confusing.	Replaced all instances of the word “feature” with “input variable”	Throughout
18	In machine learning, you must calculate variable importance independently. For instance, a random forest utilizes out-of-bag (OOB) bootstrapping to estimate the importance of each feature (aka predictors). Typically, a single plot displays the importance of each feature, allowing you to compare them and determine which feature significantly influences the model's performance.	We opted for 2 complementary methods to calculate variable importance. The first is to determine the correlations between model performance and skill, (shown in Figures 7 and 8) and the other is to calculate Shapley values. The plot showing Shapley values was originally in the paper but it was described as ‘poor quality’ in the previous review. Without further explanation of what this meant the decision was taken to remove the Figure. For this latest version of the paper the Shapley values plot have been added back to the paper to provide a measure of variable importance in which the importance of each variable is calculated independently.	Section 3.5
Discussion			
19	The study lacks a comparison of the forecast model's performance against other comparable studies to provide context.	Added a new section in the discussion which describes some results of other work which forecast NDVI, VCI and other vegetation indices at various timescales and addresses how the results of our study fit within the wider context.	Section 4.5