Author Response to Anonymous Referee #3

Review comment	Author response
The Material and Methods chapter explains how the geospatial analysis is done and also the classification criteria for the LTFEs. However, there is no information on how the experimental design should be analyzed as stated as one of the two main objectives of this study.	We are going to write a section about the analysis of LTFE. There the analysis of individual LTFE shall be described as well as the analysis of several LTFE with similar treatments in the form of a meta-analysis. However, in our opinion, this section then will fit better in the introductory part.
Do statistical methods come to use? Which ones? The pure assignment of LTFEs to four different classes without further statistical analyses (e.g. various types of discriminant analysis, contingency and cross tabulation, factor analysis) is not very appealing. The same holds true for the analysis of the data for climate (CWB) and soil fertility (MSQR) given as number of cases and percentage of share of classes (tables 2 and 3).	It is important to stress that our database comprises a complete respository of all LTFE with a duration of more than 20 years conducted in Germany. As such, our database constitutes a complete enumeration of the whole population of LTFE in Germany. Due to the complete enumeration, we believe that descriptive statistics (cross-tabulations, contingency tables) provide the best means of analysing our data. We will write the two used methods down in the Material and Methods section. Methods of statistical inference, such as chi-squared tests for contingency tables, for example, are unecessary, precisely because of the complete enumeration. Such tests would only be helpful, if a random sample of LTFE were available out of a larger population. But such is not the structure of our data. The reviewer also suggests two multivariate methods, i.e., as factor analysis and discriminant analysis. Both methods would potentially use a large number of environmental covariates characterizing the LTFE. By contrast, our hypotheses relate to two clearly defined covariates that span a two-way classification, i.e. Müncheberg Soil Quality Rating and Climatic Water Balance. Moreover, we believe the two suggested multivariate techniques do not really match our objectives. The purpose of discriminant analysis it to provide a model-based decision rule that allows allocating new samples to known groups of units (LTFE in our case). This kind of application is clearly not what we need, as we already have a classification of all LTFE in our database. Moreover, there are no new LTFE to be classified. As regards factor-analysis, this is largely an exploratory method for a larger number of variates that allows exploring possible grouping in multivariate space. Again, this does not meet our needs; we already have the classification in hand that we are analysing,

	and this is based on just two well-defined covariates. Further on, we are going to include a further author, who is an expert i.a. in spatial methods for field trials, design of comparative experiments, and network meta-analysis.
(There are) five (classes of LTFE) in table 1 and eight in figure 3?	This is due to the fact that multiple nominations were possible in Table 1 and we wanted to deal with each individual LTFE in Figure 3. We will redesign Figure 3, also according to the comments of the other referees, and clarify what exactly we want to represent.
I am convinced that the manuscript would greatly benefit from a profound statistical analysis and that this would allow (i) a critical discussion of the value of the data that exist so far and (ii) to conclude how such laborious and expensive experiments could be designed in future.	See response to the 2nd statement.
A purely qualitative, merely descriptive analysis has certainly been carried out to a sufficient extent in the large number of papers already published on this subject, most of them mentioned generously.	Although various compilations of LTFE in Germany exist, this paper is new in the aspect, that it provides a carefully developed example on how a large number of long-term field experiments can be comprehensively characterized with meta-information. In addition, the geospatial analysis of LTFE sites is new. We will clarify this and further reveal the added value of knowledge gained in this paper.
A discussion of the results including international literature and experiences of long-term experiments, e.g. from England, China or the US, is missing to a large extent. I recommend that the discussion be significantly revised and expanded in these points.	We will include international literature in the discussion.
Appropriate quantitative methods for the analysis of the experimental design and the spatial distribution of the experiments with regard to climate and soil fertility should be added.	What is meant by “experimental design” here? We have chosen a descriptive approach to classify the total population of LTFE in Germany. We believe that contingency and cross tabulation are stringent methods for this. If instead e.g. a factor analysis would have been chosen, that would be a completely different approach.
Line 49-55: the enumeration of the number of LTFEs published over the years by Körschens seems unnecessary in this way. If the details here are important I would recommend to present it as a table. And Line 83: after the explanations in the introduction regarding the work on the German LTFEs prepared by Koerschens et al., it seems	The numbers show, that our work was needed. We had the opportunity to carry out an extremely extensive search, which led to more than twice as many LTFE (205) being known as in Körschens' most extensive study (97). In addition, the setup of new LTFE with a planned duration of at least 20 years goes on and we have also recorded LTFE that were setup after

incomprehensible why a new literature study should be made here and would require a corresponding justification. This should also explain why the work of Koerschens et al. is obviously not adequate to follow the objectives of this study.	Körschen's publications. In addition, we included grassland experiments. Also regarding the details to each of the experiments we provide much more information in our dataset (http://doi.org/10.20387/BonaRes-3tr6-mg8r). Although most of the details are not needed for the spatial analyses of this paper, the precise coordinates of the LTFE are needed and could only be found out through our extensive search. We will state that more clearly in the paper.
Lines 63-80: after the objectives of the work have been formulated in lines 61-63, the explanations given here seem like a description of material and methods. I recommend to shorten this part and to integrate it into the chapter Material and Methods.	Ok, we will shorten it and integrate it into Material and Methods.
Line 68: what is meant by research parameters? Please list.	By research parameter we mean everything that has ever been sampled and recorded in LTFE. Probably “measured parameters” is less misunderstanding. We will change that. An overview of the measured parameters known to us can be found on pages 9 to 11 of the fact sheet (Grosse, M., Heinrich, U., and Hierold, W.: Fact Sheet for the Description of Long-Term Field Experiments / Steckbrief zur Beschreibung von Dauerfeldversuchen, http://doi.org/10.20387/BonaRes-R56G-FGRW, 2019.). That is a very long list. Maybe we can simply refer to that?
Line 95: here, too, the technical justification for the selected research topics is missing. Especially with regard to the aspect of a meta-analysis of the research statements, which was prominently emphasized in the introduction, the research topics listed here appear incomplete.	We will insert a short explanation in line 95, why we clustered the LTFE according to these four themes. Furterhon there may be a misunderstanding here. We did not promise a meta-analysis in the introduction. It is a (descriptive) analysis of the LTFE in Germany with regard to land use, research themes and farming systems (lines 61-62). We will add "descriptive" and skip "experimental design", which probably lead to the misunderstanding.
Lines 200-206: the description of the methodology belongs in the corresponding chapter and is superfluous here, as are lines 208 and 209. Similar mixtures of results and material and methods are also shown in the following chapters. I would recommend to check the results part and to concentrate all methodical information at the appropriate place.	We will enhance the structure.
Figure 1 does not seem necessary to me, the content is very simple and directly repeats the	We would like to leave Figure 1, because we believe it improves the readability of the paper.

statements in the text without a gain in information.	
The core statements in figure 3 could certainly be presented much more clearly. At the moment most of the space is taken up by the legend. It also seems unusual to me that the figure itself contains a headline (‘Start of LTFE’).	We will enhance Figure 3, see above.