Answers to reviewer 1 comments:

Small scale spatial heterogeneity of soil respiration in an old growth temperate deciduous forest

General comments

The authors decribe an attempt to characterise the spatial pattern in soil respiration using repeated measurements with a closed chamber system. Generally, the mauscript is well-structured. However, there are many deficiencies. The overall scientific construction of the manuscript is weak. The two hyptheses mentioned in the introduction are rather common and are basically already known from literature.

Relevant literature like e.g. the paper of Kosugi et al. is missing. There are major methodological concerns. Basic principles of geostatistics are not fully understood by the authors and are thus applied in a strange way. Terminology is used falsly. Many relevant steps in the data handling are only roughly described or completely missing. Results are discussed, which were not presented properly in the results section, see 'specific comments'. The written English does clearly not meet international publication standards. I made some suggestions. Since I am not a native speaker, too, I suggest some serious editing by a native speaker.

We greatly appreciate the constructive and elaborative comments of reviewer 1. Every single issue raised has been carefully examined. This resulted in a considerable overhaul of the manuscript. It clearly profited and we feel that it is much better now. In the following we answer all specific comments.

Specific comments

9978 2 the abstract should not begin with "large scale spatial heterogeneity" when the focus is on small scale heterogeneity

We only mentioned the amount of studies focusing on soil respiration between different sites compared to studies focusing on micro scale heterogeneity of soil respiration. However, now the abstract starts with the fact: "Soil respiration has been studied intensively in various ecosystems throughout the last decades."

9978 3 replace "comparably little" with "less"

Because of extensive changes to the abstract the whole sentence has been skipped.

9978 4-6 but all this factors are not focus of this manuscript.

ditto

9978 7 how could "heterogeneity" be "evaluated", better use "investigated"; skip "To do so,"

The sentence has been changed according to your recommendations.

9978 21 the term "extrapolation" is falsly used throughout the whole manuscript. "interpolation" would have been the correct term.

The terms extrapolation and interpolation are not used consistently in all fields and in geostatistics. However, we followed your advice and use "interpolation" now throughout the manuscript

9978 26 replace "permits" with "allows"

Has been replaced according to the suggestion.

9978 27 replace "reliable" with "reliably"
Has been replaced according to the suggestion.

9979 12 there are much more and newer references like e.g. Rayment and Jarvis 2000, Soil Biol Biochem., Herbst et al. 2009, Vadose Zone Journal
We now considered the suggested recent literature as well as additional new publications.

9979 13 replace "high" with "large"; replace "total annual values of site scale" with "annual averages of site-specific"
The text has been changed accordingly.

9979 18 give a reference to soil organic matter content, for example Fang et al. 1998, Plant and Soil
We followed your suggestion.

9979 21-22 I do not agree to this statement. "Spatial heterogeneity" describes any variability in space not just between sites.
We used the expression "spatial heterogeneity" to comply with some publications that understood the expression in the way that you criticised. However, we agree with your criticism and now apply the expression correctly throughout the manuscript.

9979 27 "extremely rare" I strongly disagree. There is a number of relevant publications, which are obviously unknown to the authors.: Kosugi et al. 2007, Agricultural and Forest Meteorology, Fang et al., 1998, Plant and Soil, Xu & Qi 2001, Global Change Biology.... Further, there are references cited within this manuscript explicitely dealing with the spatial heterogeneity in old forests like Saiz et al. (2006) and Soe & Buchmann (2005).

We are aware of the mentioned publications. Our statement "extremely rare" only referred to studies dealing with "micro scale spatial heterogeneity within old growth forest with otherwise relatively homogenous site conditions". Xu & Qi 2001, (Global Change Biology) worked in a young ponderosa pine plantation dominated by 7-8 year old trees. Kosugi et.al. 2007 worked in a tropical rainforest in Malaysia within a mixed dipterocarp forest and Fang et.al 1998 worked in a "second-rotation slash pine planted after clear-cutting (stem-only harvest) of the previous stand in 1972" in Florida. Whereas methods depicting "spatial heterogeneity" can be possibly transferred to our study site, their site conditions are not comparable. Nevertheless we will clear out the expression "extremely"and we will add the term "temperate deciduous", because this states clearly what we mean.

9980 5 skip "Having said this,"
We skipped the words according to the suggestion.

9980 5-6 but this is not investigated within this study. You just work univariate, only for respiration you try to investige spatial autocorrelation.

That is true. We are aware that soil respiration is influenced by various confounding factors which cannot all be considered within the scope of one paper. For example, spatial heterogeneity of soil temperature is small, but its seasonal variation is large. Here, we just state, that it is not to expect that different variables change on exactly the same scales. We wanted to underpin the need for a "spanning scales approach". However, we might raise expectations with this sentence that the paper does not come up with. We skipped the sentence and changed further parts of the paragraph. It now reads:
"An approach spanning various spatial scales seems to be adequate to identify the scales on which the variation occurs. Therefore,..."

9980 this is impossible. If I get it correctly, you want to improve the average by using kriging? Kriging always reproduces the average of the sampling values. That is one of the main features of kriging, best linear unbiased estimator (Blue).... Look for it in a text book on geostatistics, e.g. the one written by M. Armstrong, Basic Linear Geostatistics

You are right. Kriging reproduces the average of the sampling values at each sampling location. But we compare the total average respiration rate of all sampling locations with the average respiration value of all grid cells of the kriging results. The results cannot be the same especially in our case where the sampling locations are distributed randomly. Thus, there can be a large difference between averaging soil respiration measurement values of the sampling locations, or averaging the values of all raster cells obtained from an ordinary kriging map of soil respiration data.
We now explain that in more detail in the geostatistical method section at page 9985.

9980 14-19 This are rather weak hypotheses, which are known from literature
We now further specified our hypotheses. This paragraph now reads:
(1) The environmental controls for soil respiration vary on different spatial scales within old growth temperate deciduous forest sites. Therefore, a multiple scale sampling approach improves the reliability of site scale soil efflux rate estimates extrapolated from point measurements.

(2) The consideration of seasonal parameters like soil temperature and soil moisture further improves the interpolation results.

9980 23-25 please skip this sentence, not relevant
The information about the location of our study site next to an eddy covariance tower is necessary, because the size of the fetch defined the size of the study site. When this information is not given, the irregular shape of the investigation area that is obvious on the provided maps, might provoke questions. This approach enabled us to compare our model results with the tower data but this is beyond the scope of this paper. We state this now in the text.

9981 measurement locations should not be called "plots". This is simply inappropriate wording.
We changed the terminology according to your suggestion throughout the text.

9981 6-23 I do not understand why you used 43 random locactions (in a grid) and additionanally 81 completely randomized locations. There is a bunch of literature how sampling schemes should be designed, especially for geostatistical purposes. am pretty sure the nested approach of Oliver andWebster 1987, J. Soil Sci. would have been much more efficient. In particular against the background that autocorrelation at different spatial scales should be investigated. Small sampling distances are only given for two cells of the grid. And you can never tell whether their small-scale variability is representative for the whole site.

You are right in pointing to the nested approach. However, a completely nested setup is simply not manageable when measuring respiration over time in such a relatively large site. It is possible when one follows a one or few times measuring scheme, but it is not feasible when measuring bi-weekly or weekly for several years. Due to labour restrictions we were forced to adapt a sampling scheme that accounts for the problem whilst still providing the possibility to look at the scale issue. The two grid cells with higher sampling densities were chosen randomly. However, you are right in the last point you raise. We tackle that now within the discussion.

9982 9 skip that sentence
The sentence has been skipped according to your suggestions.

9982 14 "were seen as reliable..." was that a question? Why do you compare measurement devices here?
There is no doubt about the reliability of our equipment; therefore we skip that.

9982 21 But those sensors are only available at one point in space!
Unlike soil moisture, the CV of soil temperature is very low in space within our data. Therefore, using this one spot data for the missing reads seemed to be the better choice compared to having no data.

9983 1 How did you convert the gravimetric water content into volumetric, required in Eq. 3? You would have needed bulk density (g/cm3) measurements at every sampling location. Was that the case?
Yes that is right, we obtained undisturbed soil (core) specimens next to the soil respiration measurement locations at the beginning of the measurement campaign. This information has been added to the text:
"At the beginning of both measurement series, soil samples were also taken singularly for determining the soil bulk density."

9983 17 please give the units of k and a
The constants k and a are dimensionless model parameters determining the shape and scale of the function. Our procedure follows general practise considering temperature dependence of soil respiration.

9983 24 In Eq.3 you assume an exponential relationship between respiration and soil water content, similar to the arrhenius approach. I understand, that makes it easier. However, first you have to prove this functional relation. Plot the residuals of the fit of Eq. 1 against water content...

We took the formula from Knohl et al. 2008. They established this relationship with a much smaller data set from the same investigation site. We added the plot that you asked for (Fig 5). It shows that the residuals of the regression respiration against temperature have a non-linear relation to the soil moisture content. Under very dry conditions (below soil moisture content of 20%) respiration is underestimated, whereas under medium moisture conditions the fit is generally good but can, under specific conditions driven by other variables not measured here, also be overestimated. However the residuals are generally small compared to Knohl et al. 2008.

9984 21 was the standard deviation computed from R_standard or from R_deltaTdelta_theta? However, was no skewness detectable in the data? Kolmogoroff-Smirnoff...., see above in the manuscript. That would significantly hamper the usefullness of Eq. 6 as it is. Usually a log-transform solves that problem.

The standard deviation has been calculated from R_standard. There is skewness in the data (as you correctly assume). However, as one can see in eq.3 and as we explain in 9984 1-3 we took the logarithm of the respiration (we logarithmized the whole formula) to account for the skewness.

9985 3 replace "construction" with "determination"
The wording has been changed according to your suggestions.

9985 4-6 What was the criterion for this optimisation?
The criterion has been the maximising of the range.

9985 13 "Fig. 6" The figure numbering should be in accordance to the appearance in the text. So, this should be Fig. 2 I guess.
You are right, we will change that.

9985 19-21 Show plots of the regressions, R^2 and so on....
Due to space restrictions, we cannot present plots for each measurement location.

9985 25-26 Again, kriging always reproduces the average of the sampling points....
See comment above

9986 8 skip the"x" in the units
We skipped the "x" in the units according to your suggestion

9986 14 Q_{10}=2.3, Q_{10} amounts to 3.9???? What is true now?

Thank you for pointing out that mistake. Q_{10} of 3.9 is calculated using all data. As you can see in figure 4 we split the dataset depending on soil moisture content. Q_{10} =2.3 is true for soil respiration data between >60-90% soil moisture content. The paragraph now reads:
"The seasonal variation of soil respiration is clearly temperature dependent (all data: r^2= 0.60, Q_{10}=3.9). The total average soil respiration value is 0.28 g CO_2 m^{-2} h^{-1}. However, the relation between soil respiration and soil temperature was much closer at collars with lower soil moisture content (0-30 % Vol., r^2= 0.69, Q_{10}=2.3) compared to wetter plots (>60-90 % Vol., r^2=0.32, Q_{10}=4.1) (Fig.4). The highest average respiration rates (0.30 g CO_2 m^{-1} h^{-1}) were observed at medium soil moisture between >30-60 [% Vol.]."

9986 20-21 "at a given measurement date"? I suggest to compute coefficents of variation (=mean/standard deviation) and to investigate this systematically for every single measurement campaign.
The low variability of soil temperature can be seen in figure 3. However, we added an additional figure (Fig. 6) in which we plot CV of soil temperature against day of year.

9986 25 "R^2 increased" but what did you use for R^2=0.6? Temperature only?
Yes that is right. We added that information now.

9987 3-5 That is called first order stationarity. However, this does not need to be mentioned explicitely. If that would have not been the case geostatistics could have not been applied.

We express that we do not need to care about trend removal procedures that means removing trend before kriging and adding it back later. Nevertheless we skipped that paragraph.

9987 17-24

????? You could not infer kriging parameters from varying sampling density. This is completely strange.

That is a misunderstanding. We do not say that we could not infer kriging parameters from varying sampling density. Instead, we write here, that we actually did exactly this. We optimized the kriging parameters by varying sampling densities. The point is, that we did with subsets of the data that are centred around the nests to test for many different densities. We tested how the average soil respiration rate within the two nests was affected by different sampling densities, as well as by outliers. Outliers were therefore removed during the optimization of the semivariograms. We reformulated that paragraph. It now reads:
Data point density was varied to optimize the kriging parameters (Fig. 7). We did this on subsets of the data centered around the nested cells. That allowed us to test a large range of different densities. Whereas the calculated average soil respiration rates around one of the grid cells changed little when altering sampling densities by subsampling from the available data, soil respiration rates decreased significantly with higher sampling densities. Both cells are characterized by artificial block structures at low and medium sampling point densities, indicating a distance between the sampling points wider than the range of the semivariograms. Much of the block structures can be smoothed by using a coarser scale for soil respiration, but not all.

9988 2-3 because average sampling distance is so much larger than the range! What makes all this ambiguous...

That is right. But it was not foreseeable. As we explained earlier a completely nested approach as proposed by Webster is simply not possible when measuring bi-weekly on more than 100 locations. Therefore we choose the standardisation approach with which a larger range is obtained due to the removal of temperature and moisture effects. This allows us to compute a sound estimate of the spatial variation in standardised respirations. The variation that is left is due to the influence of other not measured variables. We added a sentence now, explaining the larger range of the standardisation approach.

9988 10-22 trivial
We will skip that paragraph as described above.

9989 2 show regression and R2
We show the regression including R^2 now, according to your suggestion

9989 3-6 but the predited standard error is only smaller because you removed the variability from the data. This tells you nothing! The average computed from the non-standardized respiration has a much higher variability. This approach is not valid!

Figure 10e shows an idle map where the influences of different soil temperature and soil moisture values were eliminated. It is true that this is an abstraction and will never be the case in nature. But the map expresses the remaining variability of soil respiration due to other parameters like different soil properties, root distribution patterns etc.

9989 4 Which RMSE?
The RMSE is 0.04497, this information has been added to the text.

9992 1-11 this paragraph complete lacks from a link to the results of this study
We skipped it.

9992 15 If outlier stection is that relevant, why was it not applied?

It was applied during the semivariogram calculations. See 3.2.4

9993 25 30 cm!!
The area covered by their soil cores was extremely small, compared to our soil respiration chambers. See: Stoyan et. al. (2000): Spatial heterogeneity of soil respiration and related properties at the plant scale, Plant and Soil

9994 7-8 "show little autocorrelation" This result is not presented at all within this manuscript.
We added example semivariograms, two for temperature and two for soil moisture to the results section.

9994 3-5 I strongly disagree. There is a suite of widely applied geostatistical methods that allows multivariate co-regionalisation, like external drift kriging (Ahmed & DeMarsily 1987, Water Resources Research) or regression kriging (Odeh et al. 1995, Geoderma). Something like that should have been applied instead of all the analyses presented here.

We only have point data measured at the soil respiration measurement locations and no continuous maps that could serve as auxiliary data. Therefore, the presented approach was followed. Anyway, we think that our postulated aim, the augmentation of the kriging results, can be obtained by the approach that we follow.