Uncertainties, sensitivities and robustness of simulated water erosion in an EPIC based global-gridded crop model

By T. W. Carr et al.

Reply to Anonymous Referee #2

Dear reviewer,

Before we address each of your comments, we briefly clarify the main incentive of this study. Large-scale indicators about global-scale phenomena are needed to inform all major environmental and agricultural policies such as the European Union's Common Agricultural Policy (CAP), the United Nations Sustainable Development Goals (SDGs), the United Nations Convention to Combat Desertification (UNCCD) and the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). Water erosion will not be considered in any of these major environmental and agricultural policy programs without large-scale assessments (Alewell et al., 2019). Global-gridded crop models have the capacity to develop large-scale indicators and to inform about agricultural productivity in a transparent and consistent way across large areas (Mueller et al., 2017). This paper aims to address the gaps in the literature of the links between water erosion and crop cultivation in various large-scale and global impact assessments, as accurately as is currently feasible given data availability. Studies on large-scale and global climate change impacts in the agricultural sector lack representation of water erosion impacts on crops (Balkovič et al., 2018), studies on global terrestrial carbon fluxes do not account for carbon runoff from cropland through soil erosion (Chappell et al., 2016), and studies assessing large-scale and global market impacts of soil erosion rely on simple linear estimates of water erosion impacts on crop production (Panagos et al., 2018; Sartori et al., 2019). It is important, though, to understand the limitations of such assessments so that they can be improved in the future, and that is why we systematically test a number of approaches in our paper.

The model used in this study has been confirmed as a reliable tool for global crop yield projections and stands out against comparable global models due to its detailed representation of soil processes including water erosion and the impacts of tillage on soil properties. Therefore, a global-gridded EPIC model has the potential to deliver much needed indicators about relationships between erosion and crop productivity on large and global scales. This paper has the objective to support the ongoing development of large-scale and global model applications by analysing the robustness of average long-term water erosion estimates generated with global-gridded crop models. In other words, we do not attempt to reproduce soil loss rates measured in single fields but to analyse the robustness of large-scale water erosion estimates based on global-gridded crop model outputs. Moreover, we focus on the necessary improvements needed to account for water erosion in global models by analysing and discussing its robustness across global agro-environmental conditions, the importance of global input data on field management and the uncertainty resulting from different erosion equations.

Most of the criticism by the referee is directed against the lack of field data, model calibration, and general criticism on RUSLE-based erosion models and the ^{137}Cs-method used for erosion measurements. Each point of criticism is addressed in the following (Referee comments are printed in purple and our replies are listed below):

The EPIC model has been used to look at climate change impacts on crop yields and erosion rates e.g. Favis-Mortlock et al. (1991) and to model 7000 years of erosion under changing climates and land uses for a single field (Favis-Mortlock et al., 1997). It is stressed that EPIC needs calibration in order *to give reasonable results. This is the very firm conclusion of the GCTE erosion model testing exercise (Favis-Mortlock, 1998; Boardman and Favis-Mortlock, 1998).*

The first point of criticism is on the need to calibrate the EPIC model for reasonable results. The EPIC-IIASA model uses state-of-the-art global crop management and agro-environmental input data and has been positively evaluated for representing national average yields and inter-annual yield variability globally (Balkovič et al., 2014). It was used in several studies and its outputs have been compared to regional yield statistics and other global crop and land use models as a part of ISI-MIP and GGCMI model inter-comparison initiatives (Mueller et al., 2017). Global crop models are not calibrated to reproduce crop yields at field scale but rather to represent the crop yield patterns across regions and countries to address research questions that cannot be addressed through field scale studies. Following the same paradigm, we aim to analyse the robustness of EPIC to represent regional differences and regional spatial patterns in water erosion estimates rather than accurately reproduce erosion events occurring in the past in response to individual rain events of rainy seasons. We are aware that our approach would be inappropriate for the latter case. We are also aware that a proper model calibration is always needed to meet experimental data obtained from the field, foremost for a complex process like soil loss due to water erosion. At the same time, sound calibration for a wide range of global environments and crop management practices would require enormous capacity and work force while still facing a high degree of subjectivity in experimental data (e.g. Panagos et al. 2016). Given the current lack of consistent field measurements representing all global environments, it is not possible to produce plausible global erosion estimates using only bottom-up, field-scale modelling.

To further clarify the intention of this study, we add a reference clarifying the usefulness of large-scale models to line 57: "*Moreover, improving the representation of water erosion in large-scale models is urgently needed to inform major environmental and agricultural policy programs such as the European Union's Common Agricultural Policy (CAP), the United Nations Sustainable Development Goals (SDGs), the United Nations Convention to Combat Desertification (UNCCD) and the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) (Alewell et al., 2019).*"

We will further clarify the focus of this paper in the introduction, line 83: "*The overall aim of this study is: (i) to analyse the robustness of water erosion estimates in all global agro-environmental regions simulated with an EPIC-based global-gridded crop model; and, (ii) to discuss the main drivers affecting the robustness and the uncertainty of simulated water erosion rates on a global scale.*"

We will modify the last conclusion on line 474: "*The overlap of simulated and measured water erosion values in most environments used to produce wheat and maize underlines the robustness of an EPIC-based GGCM to simulate the differences in water erosion rates of major global crop production regions*"

The authors claim that they are evaluating their results against field-scale measures (lines 84 and 95). This is not the case: they use 137Cs and erosion plot data (line 219). Erosion plot data cannot be up-scaled to field scale: it is useful for relative assessments e.g Cerdan et al. (2010). Extrapolation from12 plots in central Belgium to give an average rate of erosion for Europe is a well-known (?) example of misuse of experimental plot data (Boardman, 1998): the current paper is heading in that direction!

The referee criticises the use of our erosion plot data. We do not extrapolate plot data to represent water erosion on continental scales as in the Belgium-example mentioned by the referee. As explained on line 235-239, we aggregate field data into groups with similar slope classes and precipitation classes and compare the values to model outputs with the same slope and precipitation classes. Thereby, we analyse the robustness of model outputs for different environments characterised by the most important parameters affecting water erosion on a global scale. Slope and precipitation are the most sensitive parameters influencing model outputs and are the parameters found previously to be most critical for the robustness of RUSLE-based models. We chose this method to illustrate the varying robustness of our model outputs around the globe and we identified regions where the model performance is not sufficient, which is communicated in the discussion (line 367-370) and the conclusion (line 472 – 473).

More generally, plot and field scale are not exactly defined. These two categories can be overlapping as large plots can have similar slope length as fields. Plots of 20 m are most common, and if they are equipped by multislot divisors, typing buckets they can be up to 100 m long, or even several hundred meters if equipped by Coshocton wheels. We know about 100 m erosion plots from Slovakia and Austria or 30 m plots in Zimbabwe. Fields can also have slope length about 20-100 meters, especially if they are on slopes, contour oriented or when they belong to small family farms in developing countries. We have seen 30-50 m long slopes in Uganda, Madagascar, or Slovakia.

Regarding the real scale of our data, most of them (both 137Cs and erosion plots) represent slopes of 10-100 m so they are at the margin of plot and field scale. Therefore, in the paper, we will name the range of spatial scales of the field data on line 218 (we increased the field data sample as explained below). "*We compared our simulated water erosion rates with 606 soil erosion measurements on arable land from 39 countries representing plot, hillslope and field scale; 315 records were derived by the ^{137}Cs method, 188 records from erosion plot experiments, and 103 from volumetric observations.* "

The term "*field-scale measurements*" on line 84 will be deleted as this sentence will be replaced with the sentence stating the overall aim of this study presented above.

The term "*field measurements*" on line 95 is not related to scale

RUSLE is an unvalidated model and its problems and poor performance are reviewed in Evans and Boardman (2016a and b). For a review of the general problems of using erosion models see Favis-Mortlock et al. (2001): in Harmon and Doe (ed) book.

Further criticism is focused on the validity of the RUSLE method in general, which is one out of seven similar erosion models included in our EPIC study. The RUSLE methodology is based on more than 10,000 plot-years of experiments and has been applied in more than 100 countries with varying robustness (Alewell et al., 2019; Renard et al., 1997; Wischmeier & Smith, 1978). As already disputed above, the limited availability of global input and experimental data requires simple erosion models for global studies. Therefore, RUSLE has been chosen by most studies focusing on global erosion (Borrelli et al., 2017; Doetterl et al., 2012; van Oost et al., 2007) and it is unfortunate that the referee disagrees with the approach adopted by most of the scientific community. We agree that the varying robustness of RUSLE-based methods around the world need to be considered, which is one of the main foci of this paper and has been addressed in the introduction (lines 66 – 68) and in the discussion (lines 360 – 373). Furthermore, we present the varying estimates of different water erosion equations and thereby demonstrate the uncertainty of relying on erosion estimates from a single erosion model (lines 268 – 271, 381 – 384, 409 – 417). The references used for criticising the RUSLE method by the referee promote field studies as alternatives to erosion models. However, this is not feasible for a range of research applications focusing on analysing scenarios at large and global scales (Alewell et al., 2019; Panagos et al., 2016), which is also a major purpose of global gridded crop models. Instead, focusing on the improvement of the application of simple erosion models such as RUSLE-based models as intended with this paper supports the ongoing development of global erosion impact assessments (Naipal et al., 2015).

137Cs has been seriously criticised recently (Parsons and Foster, 2011). The technique should not be used without dealing with these limitations. This problem is ignored in the paper.

The objections of Parsons and Foster (2011) were discussed by Mabit et al. (2013) published as a direct answer to Parsons and Foster in the same journal. One of the most important confirmation of the usefulness of the 137Cs method given by Mabit et al (2013) are the positive results of a comparison between erosion values obtained with various measurement methods and erosion values obtained by 137Cs method: "*Several studies at various scales (from plot to watershed scales) have been conducted to compare the 137Cs based erosion rates obtained with direct erosion measurement approaches such as erosion plots and catchment sediment yields (e.g. Schuller et al., 2003; Porto et al., 2003; Porto et al., 2004; Stankoviansky et al., 2006;Mabit et al., 2009; Parsons et al., 2010; Ceaglio et al., 2012; Porto and Walling, 2012a, 2012b). In northern California, rates of erosion from accumulated pond sediment and soil lost from hillsides assessed through 137Cs agreed well (i.e. O'Farrell et al., 2007). In Italy, the reliability of the mass balance model at slope and catchment scale was verified and confirmed by comparing the basin net soil erosion value obtained by 137Cs measurements against the mean annual value of sediment yield measured at the basin outlet (i.e. Di Stefano et al., 2005). For three small catchments located in Southern Italy, measurements of sediment output validated 137Cs theoretical conversion models to estimate soil redistribution rates (i.e. Porto et al., 2004)*".

Another paper by Parsons (2019) also criticises erosion plots, modelling (several models) and volumetric measurements of rills. As our paper is focused on global erosion modelling, we did not include the extensive literature on the advantages and disadvantages of each erosion measurement method. A comprehensive discussion of the objections against the 137Cs method and other measurement methods would but too extensive for our paper. However, we will address Parsons and Fosters objections in this response:

a. Regional and local heterogeneity of 137Cs fallout and its local redistribution by vegetation, infiltration, bioturbation, etc.: this is well known fact but it is considered in the methodology and the solution is the selection of reference sites in immediate vicinity of study sites and using statistical criteria (variation coefficient) criteria for microvariability of 137Cs. The microvariability of 137Cs is similar to most other soil properties and the potential error is similar or smaller than for other erosion measuring methods (as it will be demonstrated further). There is set a limit for variation coefficient which reference site should not exceed.
b. Mobility of 137Cs: the references presented by Mabit et al. (2013) clearly demonstrate that 137Cs is strongly bind to colloids (with details on particular clay minerals and organic matter) and its mobility (washing by runoff during the deposition, leaching and plant uptake are really negligible (representing less than one percentage to very few percentages of 137Cs fallout). The ideas about mobility, leaching and plant uptake of 137Cs presented by Parsons and Foster are based mainly on laboratory experiments with Caesium which do not represent its real behaviour in nature because the laboratory conditions are artificial, and the used doses of Caesium are too high. Parsons and Foster admit that in their paper. If these ideas would be correct, we would frequently see leached Caesium in deeper part of soil profile or the whole inventories at sites undisturbed by erosion would be depleted by plant uptake. But this is never the case in natural soil.

c. Selective removal and sorting of the particles: This is well known by all authors using the 137Cs method and it is mentioned in all handbooks and conversion procedure has a parameter for that. But it is true that this factor is difficult to quantify. But similar weak points are common in most methods of erosion measurements as will be demonstrated below.
d. Conversion models: Indeed, procedures to calculate soil loss require some parameters which are not always available. The accuracy of calculation depends on quality of input data and is different in individual studies. But this is the case for all erosion measurements and all methods as it will be demonstrated below.
e. Sample preparation: the problems with coarse fraction, dry or wet sieving can occur in specific soils, especially when having porous coarse fraction and concretions containing clay or organic matter so that the coarse fraction has electrical charges. This of course should be understood by staff who is expected to have basic pedological education. It is true that not all these details are mentioned in every methodological guidance, but they are discussed by some authors.
f. Gammaspectroscopy: Criticism of gammaspectroscopy is not relevant at all. Indeed, there are possible different geometries of detectors and detectors have to be calibrated. But this is task for laboratory staff. Cs-method obviously require staff having background in nuclear physics. Each laboratory method whether physical or chemical require staff with appropriate qualification.

Finally, we accept that we could have mentioned the limitations of the field data we collected and referred to some of the existing literature. We will briefly mention the main limitations of the field data we collected on line 449: "*The accuracy of the ^{137}Cs method has been criticised as well as the subjectivity of choosing reference sites and sample points (Parsons and Foster, 2011). Similarly, the accuracy of erosion measurements derived from runoff plots has been challenged due to the complexity in designing plot experiments and collecting soil runoff data (Boix-Fayos et al., 2006). Also, volumetric surveys have been criticised due to their high degree of subjectivity (Panagos et al., 2016). Nevertheless, the usefulness of the ^{137}Cs method, plot experiments and volumetric surveys has been confirmed in various studies (Boardman and Evans, 2020; Cerdan et al., 2010; Mabit et al., 2013).*"

It is simply not true to claim that there is a limited availability of field data and lack of long term measurements (lines 68-69). There are extensive data sets from Switzerland, north Germany and the UK. These could be used to validate the results of erosion models: see Boardman and Evans (2020: PPG) for a review of these methods of assessment of erosion at a field scale.

The need for slope and precipitation information accompanying erosion measurements in our evaluation method narrows down suitable field datasets, as meta data such as slope steepness is often not available in published datasets. Moreover, when we refer to a lack of field data, we are talking about the global scale and especially the imbalance in data availability among world regions. We have addressed the uneven distribution of global field data in the introduction (line 68) and the skewed focus of field data on the United States and Europe in the discussion (line 426). Also, the difficulty of gathering field data from a very heterogenous mix of measurement methods is comprehensively addressed in the discussion (line 419-449). We attempted to gather field data from as many continents as possible to represent different global environments. Insufficient field data representing all global regions and the lack of sufficient metadata in available datasets to further improve erosion modelling on large and global scale is an important conclusion of this paper (line 471).

We will further clarify these issues in the discussion after line 425 "*A variety of factors influencing water erosion such as climate, field topography, soil properties and field management need to be considered when modelling water erosion but are often not reported in available field measurements*

(García-Ruiz et al., 2015). This hampers a direct comparison between simulated and observed water erosion values. We demonstrated the varying match between measured and simulated water erosion using different tillage and cover crop scenarios. Metadata on field management often only provides the crop cultivated and therefore the conditions under which erosion was measured in the field are not known sufficiently to evaluate erosion values simulated under different field management scenarios. Similarly, information on field topography and soil properties is often not provided with recorded field measurements and thus their use is limited in an evaluation of water erosion estimates simulated in different global environments. Moreover, the geographical distribution of erosion data is unbalanced. Most data are concentrated in the United States, West Europe and the West Mediterranean (García-Ruiz et al., 2015). In summary, there is a lack of field data representing all needed regions, situations and scenarios (Alewell et al., 2019)."

We increased the field data sample to 606 records using publicly available datasets from Germany and the UK provided by Auerswald et al. (2009) and Benaud et al. (2020).

We add a description of the additional field data after line 231: "*We use also erosion rate measurements on arable land based on plot experiments and volumetric surveys collected by Auerswald et al. (2009), Benaud et al. (2020) and García-Ruiz et al. (2015). Measurements of soil runoff from plots are useful to compare the relative changes in erosion under different topographic, climate and controlled field management conditions (Auerswald et al., 2009; Cerdan et al., 2010). In volumetric surveys erosion rates are derived from measuring the volume of visible traces of erosion such as rills, gullies and fans, and have been considered as a good method to monitor erosion in large fields (Boardman and Evans, 2020)."*

An attempt to further increase the field data sample using the sources listed in Boardman & Evans (2020) was not possible as only aggregated values are published in the listed papers, which present large field datasets.

Some more general remarks: Although huge effort was spent by the erosion community to generate an enormous number of data, there is a serious lack of useful data to evaluate large-scale models. It was very challenging to gather the amount of field data used in this study for the following reasons:

1. As we are working with USLE-derived models, we were looking for data from certain spatial extend only. The USLE was developed at short slopes and represents sheet erosion and initial stages of rill erosion. Therefore, we preferred data from ca 10-100 m long slopes. This is the case for field data derived with the 137Cs method and erosion plots. Initially, we decided against using data from long slopes (several hundred of meters), which is usually the case of volumetric measurements of rill erosion and hydrological measurements in small watersheds.

2. At the beginning of this project we tried to focus only on data derived by 137Cs method as different methods represent different erosion processes and are subject to different systematic errors which are presented in the following:

   Erosion plots with total collection of sediment have problems to collect great volumes of sediment in case of extreme rain events (the sediment may exceed the capacity of containers) and it can be difficult to determine the weight of sediment (when it is wet, the whole volume cannot be carried to the laboratory for drying so the quantity of soil in the collected mud is just estimated by taking a sample of the mud to measure concentration).

Multislot divisors, tipping buckets and Coshocton wheels have many technical problems (multislot divisors may split the sediment unequally if they are not fixed exactly horizontally, the tipping buckets and Coshocton wheels loose part of the sediment when they are tipping or when the stream is strong water is splashing out, if the stream is week the soil material is sedimenting immediately in tipping buckets and the sample is not representative, data loggers can break, etc.).

Studies with replicated plots showed great variability for replicas. Nearing et al. (1999) report from almost 800 replicated plot pairs/year data a coefficient of variation ranging between 14% and 150%. Variability was decreasing with increasing soil loss. The rates of 10 tons/ha had coefficient of variability of ca 40%.

Geodetical method (erosion pins) has much bigger error than erosion plots because it has poor resolution. If one mm of soil is removed, the change of surface is hardly seen. But this represents already 10 tons of soil per hectare. On arable land the geodetic method has problems to distinguish between erosion and compaction.

Rill and gully volumetric measurements (preferred method in the reference provided by the referee (Boardman & Evans (2020)) neglects sheet erosion completely. The recalculation of obtained volumetric data to weight is problematic because of the limited information on soil bulk density and its vertical and horizontal variability. This is problematic as we need data in t/ha to compare with models. Usually it is not indicated whether rill measurement represent the whole year or only the vegetation season, whether they involve rills from snow melt or not, etc.

The measurement of rill volumes itself is a source of huge error. Authors who use this method know this (for example Evans, 2013, stated: “Mollenhauer notes (2002: 4) ‘The measurement of lengths and cross-section areas of linear forms can be extremely error ridden’; and quotes from Ruttiman and Prasuhn’s (1990) work in Switzerland that the ‘total error for soil loss volume can amount to between 20 and 40%’ (Mollenhauer, 2002: 4)” and further “The level of accuracy of field-based estimates depends on the amount of time spent in measuring/estimating the number, lengths and dimensions of rills and gullies and assessing volumes of depositional features such as fans. The larger an area surveyed inevitably means that cruder estimates of eroded amount will be made, for example, numbers, lengths and cross-sectional areas of channels will all have to be estimated rather than measured”).

Measurements have very few traverses (sometimes only 4, Boardman, 2003), which is a huge source of error. Boardman even says that sometimes one traverse is enough (Boardman, 2003: “The number of traverses is clearly subject to the time and resources available and also the purpose of the survey as to how much detail and accuracy is required. In many situations, it is reasonable to undertake one traverse across the mid-point of the eroded slope and estimate total erosion based on the mean rill length.”). Measurement based on one traverse compared to measurements based on 4 traverse revealed errors from -18,1% to +48,7%.

The method neglects interrill erosion which is an important portion of the whole erosion process. Estimates of the importance of interrill erosion differ significantly for different conditions (negligible amount of 0.3 m3/ha/y provided by Evans (1990, in Boardman 2003), few % of total erosion: 5-11% (Morgan et al., 1987, in Evans, 2013), up to few tens of %: 25% (Prasuhn, 2011) and 10-30% (Zachar, 1982)).

Parsons (2019) emphasize that the volumetric measurements of rills severely underestimate overall erosion because rills also involve large quantity of material which was delivered by sheet erosion to rills and further transported by rills. These proportions can be 40% for rill erosion and 60% for interrill. Luk et al. (1993, in Parsons, (2019)) determined the portion of rill erosion ranging between 0 (when only sheet erosion develops) and 56%. Our own experience from Central Europe with more heavy rainfalls from own unpublished measurements is that at steep slopes (ca 8-12 degrees) it can be in some years 10-40 tons per ha. Govers and Poesen found that the proportion between rill and interrill erosion can change significantly with time and according to changes in physical properties of top layer and deeper layers either proportion of rill erosion can rise or the proportion of interrill erosion can rise. In their case study the proportion of interrill erosion was decreasing with time from 46 to 22%, but other authors found opposite trends. Therefore, estimating interrill erosion from rill erosion using fixed ratios is wrong. They also find, that interrill erosion has higher proportion on short slopes than on long slopes.

Sometimes the presented rill measurements are not real measurements but just very rough and brief estimation. The rills and their lengths are estimated from photos, where smaller rills might be difficult to detect (for example Boardman, 2003: "Ground-level photographs of rills and gullies may be used as a record of length and size; subsequent analysis shows them to be a reliable means of estimating soil losses (Watson and Evans, 1991)" or Evans, 2013: "In this review paper, 'direct' assessment of water erosion is taken to mean the mapping of erosion and deposition as evident in fields (Figures 1–9) or on ground or aerial photographs and then when possible estimating eroded volumes based on lengths and cross-sections of rills and gullies and areas and depths of deposition (Evans, 1988; Herweg, 1996; Stocking & Murnaghan, 2001)." Watson and Evans (1991) estimate the sizes of rills on photos comparing it with the widths of crop rows and height of crop and thickness of colluvial fans they estimate according to their colour. They compared the results of volumetric measurements in field and on photos (12 photos) and found ratios from 0.67 to 2.12, so in some cases the estimations on photos are higher than in field and in other cases it is opposite.

Hydrological measurements in elementary watershed do not represent erosion only from agricultural land but also bank erosion and road erosion, and both these can be significant.

Sampling of suspended sediment is not well representative, and samplers or data loggers can break. The range of discharge in small catchments is so huge that it makes instrumentation of hydrological profiles difficult.

For these various possible sources of errors, we did not want to mix up different methods. The optimal case would have been to only use field data derived from one method, but to increase the amount of data we decided to take 137Cs data and some selected erosion plots (to further increase our sample we included suitable data from volumetric surveys).

3. Large amount of existing data is not accessible for various reasons:
    a. Many older publications are in national languages

b. Many older publications are not on internet
c. Many measurements were published in grey literature, local conference proceedings, national acta of scientific institutions, unpublished reports, etc.
d. Many published data are hardly interpretable because metadata are lacking (slope lengths, or inclinations or crop cover, period of measurement is not recorded, geographical position of the sites is not recorded, many measurements were running only during vegetation period of studied crop so they do not represent annual erosion but just few months, etc.).
e. International journals do not have interest to publish usual case studies which present raw data. To get paper published the authors need to present some special objectives to follow some special goals or developing methodological innovations. Therefore, also many new data sets cannot be found online.
f. Even if paper is published, journals have usually size limitations. To save space the primary data are not presented, only the results of interpretation, statistical processing, etc. are there. In publications using 137Cs it is more common to find primary data than in studies using other methods.

Please consider that all methods have a lot of weak points, methodological shortcomings and sources of errors, uncertainties and variability and there is lack of reliable comparisons and comprehensive assessment of all methods which would be widely accepted by the whole erosion community. There are different schools and groups of researchers who use predominantly one method. One group uses 137Cs method, other groups prefer erosion plots, the next one focuses on elementary watersheds using hydrological methods based on discharge and sediment concentration sampling (or combination of plots and watersheds) and other group focuses on volumetric measurements of rills and gullies. Some researchers using certain method are very critical about other methods, but they are very tolerant regarding the shortcomings of their favoured method. Although each method has advantages and limitations and each group has success and achievements as well as challenges and failures, it is normal if individual researchers or teams prefer one particular approach. But they should respect also other approaches.

Our collected data set represents a reasonable compromise to achieve the objectives of this study. It is far beyond the capacities of the team and the objectives of this study to collect all existing erosion data. Such task would require years lasting international project with participation of research teams from most countries, so that each team would be able to revise data sources in his country and provide summary of data including those published in national language and unpublished reports. Most existing data have limited accuracy and representativeness, but we cannot wait until dense coverage of perfect data will cover the Globe. Erosion is running, agriculture is in troubles and we should proceed under existing circumstances and using available tools.

The method of deriving a common slope within an area of 9x56 km is not clear and seems rather dubious (line 122). Averaging slope from a large cell (eg. 1km2) is a common failing of erosion modelling exercises (e.g. Evans and Brazier, 2005).

The most common slope is determined by a slope class covering the largest area in each simulation grid. Slope classes are taken from a global terrain slope database (IIASA/FAO, 2012) and are based on a high-resolution 90 m SRTM digital elevation model. We assume that the slope class representing the largest area in each simulation unit is most likely covered by the largest area of cropland. This builds on the idea that a spatially extensive and diverse landscape can be represented by a single "representative field" characterized by the prevailing combination of topography and soil condition found in the landscape. This method is designed to represent differences in large-scale global crop production with an emphasis on the most important global crop production regions.

We clarify the concept of the representative field on line 120: "*Each simulation unit is represented by a single field characterized by the prevailing combination of topography and soil conditions found in the landscape. Each representative field has a defined slope length (20 – 200 m) and field size (1- 10 ha) based on a set of rules for different slope classes (Table S1). The slope of each representative field is determined by the slope class covering the largest area in each simulation unit (Table S1). Slope classes are taken from a global terrain slope database (IIASA/FAO, 2012) and are based on a high-resolution 90-m SRTM digital elevation model.*"

A detailed discussion about the uncertainty in slope input data has been added in response to comments by the third referee (see response to referee #3).

One conclusion seems to be that wheat erodes are a greater average rate ((19t/ha) than maize (6t/ha) (line 244): this is contrary to all field evidence that I am aware of.

The criticised conclusion on falsely higher erosion rates in wheat fields compared to maize fields is based on a misunderstanding. The values represent a global average value (19 t/ha) and a median value (6 t/ha) of water erosion rates for both maize and wheat fields combined.

To avoid confusion, we will focus only on median values in the revised version as median values are less influenced by the skewed distribution of erosion values. In the discussion line 375 we mention both mean and median value to illustrate the skewed distribution of erosion rates due to very high values simulated on steep slopes.

We will present global median water erosion for both maize (7 t/ha) and wheat (5 t/ha) fields on lines 25,244 and 312, and will delete average values.

Global average water erosion values simulated under different management scenarios and different water erosion equations will be deleted on line 383 – 387 to focus only on median values.

References

Alewell, C., Borrelli, P., Meusburger, K., & Panagos, P. (2019). Using the USLE: Chances, challenges and limitations of soil erosion modelling. *International Soil and Water Conservation Research*, *7*(3), 203–225. https://doi.org/10.1016/j.iswcr.2019.05.004

Auerswald, K., Fiener, P., & Dikau, R. (2009). Rates of sheet and rill erosion in Germany - A meta-analysis. *Geomorphology*, *111*(3–4), 182–193. https://doi.org/10.1016/j.geomorph.2009.04.018

Balkovič, J., van der Velde, M., Skalský, R., Xiong, W., Folberth, C., Khabarov, N., et al. (2014). Global wheat production potentials and management flexibility under the representative concentration pathways. *Global and Planetary Change*, *122*, 107–121. https://doi.org/10.1016/j.gloplacha.2014.08.010

Balkovič, J., Skalský, R., Folberth, C., Khabarov, N., Schmid, E., Madaras, M., et al. (2018). Impacts and Uncertainties of +2°C of Climate Change and Soil Degradation on European Crop Calorie Supply. *Earth's Future*, *6*(3), 373–395. https://doi.org/10.1002/2017EF000629

Benaud, P., Anderson, K., Evans, M., Farrow, L., Glendell, M., James, M., et al. (2020). National-scale geodata describe widespread accelerated soil erosion. *Geoderma*, *371*(April), 114378. https://doi.org/10.1016/j.geoderma.2020.114378

Boardman J (1983) Soil erosion at Albourne, West Sussex, England. Applied Geography 3: 317–329.

Boardman, J., 2003. Soil erosion and flooding on the South Downs, southern England 1976–2001. Trans. Inst. Brit. Geogr. 28 (2), 176–196.

John Boardman, Robert Evans, 2019. The measurement, estimation and monitoring of soil erosion by runoff at the field scale: Challenges and possibilities with particular reference to Britain, Progress in Physical Geography, 1–19:

Boardman, J., & Evans, R. (2020). The measurement, estimation and monitoring of soil erosion by runoff at the field scale: Challenges and possibilities with particular reference to Britain. *Progress in Physical Geography*, *44*(1), 31–49. https://doi.org/10.1177/0309133319861833

Chappell, A., Baldock, J., & Sanderman, J. (2016). The global significance of omitting soil erosion from soil organic carbon cycling schemes. *Nature Climate Change*, *6*(2), 187–191. https://doi.org/10.1038/nclimate2829

Evans, R. 1988. Water Erosion in England and Wales 1982–1984.

Report for Soil Survey and Land Research Centre, Silsoe.

Evans, R., Brazier, R., 2005. Evaluation of modelled spatially distributed predictions of soil erosion by water versus field-based assessments. Environ. Sci. Pol. 8, 493–501.

Fischer FK, Kistler M, Brandhuber R, et al. (2017) Validation of official erosion modelling based on high resolution rain data by aerial photo erosion classification. Earth Surface Processes and Landforms. DOI: 10. 1002/esp.4216.

Fritz, S., See, L., Mccallum, I., You, L., Bun, A., Moltchanova, E., et al. (2015). Mapping global cropland and field size. *Global Change Biology*, *21*(5), 1980–1992. https://doi.org/10.1111/gcb.12838

Govers G and Poesen J (1988) Assessment of the interrill and rill contributions to a total soil loss from and upland field plot. Geomorphology 1: 343–354.

Hayward, J.A., 1968. The measurement of soil loss from fractional acre plots. Lincoln Papers in Water Resources. 5. New Zealand Agricultural Engineering Institute, Lincoln College, Canterbury, New Zealand. Hill, H.O., Mech, S.J., Pope, J.B., Progress report

Herweg, K. 1996. Assessment of Current Erosion Damage. Soil Conservation Research Programme, Ethiopia and Centre for Development and Environment, University of Berne, Berne, Switzerland.

Hudson, N.W., 1993. Field Measurement of Soil Erosion and Runoff. Soils Bulletin 68. FAO, Rome.

Lesiv, M., Laso Bayas, J. C., See, L., Duerauer, M., Dahlia, D., Durando, N., et al. (2019). Estimating the global distribution of field size using crowdsourcing. *Global Change Biology*, *25*(1), 174–186. https://doi.org/10.1111/gcb.14492

Luk, S.H., Abrahams, A.D., Parsons, A.J., 1993. Sediment sources and sediment transport by rill flow and interrill flow on a semi-arid piedmont slope southern Arizona. Catena 20, 93–111.

Morgan, R.P.C., Martin, L. & Noble, C.C. 1987. Soil Erosion in the United Kingdom: A Case Study from Mid-Bedfordshire. Occasional

paper No. 14, Silsoe College, Cranfield Institute of Technology, Silsoe.

Mueller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., et al. (2017). Global gridded crop model evaluation: Benchmarking, skills, deficiencies and implications. *Geoscientific Model Development*, *10*(4), 1403–1422. https://doi.org/10.5194/gmd-10-1403-2017

Naipal, V., Reick, C., Pongratz, J., & Van Oost, K. (2015). Improving the global applicability of the RUSLE model - Adjustment of the topographical and rainfall erosivity factors. *Geoscientific Model Development*, *8*(9), 2893–2913. https://doi.org/10.5194/gmd-8-2893-2015

Nearing, M.A., Govers, G., Norton, L.D., 1999. Variability in soil erosion data from replicated plots. Soil Sci. Soc. Am. J. 63, 1829–1835.

Prasuhn V (2011) Soil erosion in the Swiss midlands: Results of a 10-year field survey. Geomorphology 126: 32–41.

Panagos, P., Borrelli, P., Poesen, J., Meusburger, K., Ballabio, C., Lugato, E., et al. (2016). Reply to "The new assessment of soil loss by water erosion in Europe. Panagos P. et al., 2015 Environ. Sci. Policy 54, 438-447-A response" by Evans and Boardman [Environ. Sci. Policy 58, 11-15]. *Environmental Science and Policy*, *59*, 53–57. https://doi.org/10.1016/j.envsci.2016.02.010

Panagos, P., Standardi, G., Borrelli, P., Lugato, E., Montanarella, L., & Bosello, F. (2018). Cost of agricultural productivity loss due to soil erosion in the European Union: From direct cost evaluation approaches to the use of macroeconomic models. *Land Degradation and Development*, *29*(3), 471–484. https://doi.org/10.1002/ldr.2879

Renard, K., Foster, G., Weesies, G., McCool, D., & Yoder, D. (1997). Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE). *Agricultural Handbook No. 703*. https://doi.org/DC0-16-048938-5 65–100.

Sartori, M., Philippidis, G., Ferrari, E., Borrelli, P., Lugato, E., Montanarella, L., & Panagos, P. (2019). A linkage between the biophysical and the economic: Assessing the global market impacts of soil erosion. *Land Use Policy*, *86*(December 2018), 299–312. https://doi.org/10.1016/j.landusepol.2019.05.014

Stocking, M.A. & Murnaghan, N. 2001. Handbook for the Field Assessment of Land Degradation. Earthscan, London.

Stroosnijder, L., 2005. Measurement of erosion: Is it possible? Catena, Volume 64, Issues 2–3, 30 December 2005, Pages 162-173, https://doi.org/10.1016/j.catena.2005.08.004

Watson A and Evans R (1991) A comparison of estimates of soil erosion made in the field and from photographs. Soil & Tillage Research 19: 17–27.

Wischmeier, W. H., & Smith, D. D. (1978). Predicting rainfall erosion losses. *Agriculture Handbook No. 537*, (537), 285–291. https://doi.org/10.1029/TR039i002p00285