Interactive comment on “Modelling soil erosion at European scale: towards harmonization and reproducibility” by C. Bosco et al.

C. Bosco et al.

c.bosco@lboro.ac.uk

Received and published: 17 May 2014
Interactive comment on “Modelling soil erosion at European scale: towards harmonization and reproducibility” by Bosco et al: reply to Dino Torri

C. Bosco, D. de Rigo, O. Dewitte, J. Poesen, and P. Panagos

Torri (2014) provided a variety of insights on our work. We would like to thank him for the valuable comments. Below our replies.

The semantic array programming paradigm

[Comment] – “The main difference with respect to previous attempts is the programming approach which is based on freely available software and a “semantic array programming paradigm”. Judging from the frequent links to explanatory web pages, the software system looks powerful but I never used it. I feel that an extra paragraph explaining what this system does that others don’t would improve readability: this paradigm is certainly unknown to most of the potential readers”.

[Reply] – Although the impact of computational aspects in environmental modelling is steadily growing (Casagrandi and Guariso, 2009), they are not rarely undervalued
(Merali, 2010) and the mitigation of the software-driven component of uncertainty in complex modelling might surprisingly be understated while focusing on more traditional sources of uncertainty (Cerf, 2012; de Rigo, 2013). Indeed, part of the complication in computational models (affecting even their maintainability and readiness to constantly evolve) may be mitigated (McGregor, 2006). Compared to other computational approaches, array programming (AP) understands even large arrays of data as if they were a single logical piece of information. For example, a continental-scale gridded layer may be managed by AP languages as if it were a single variable instead of a large matrix of elements. As a consequence, a disciplined use of AP (Iverson, 1980) may allow nontrivial data-processing to be expressed with terse expressions (Taylor, 2003) within a simpler control flow. Following the suggestion, we will add an extra paragraph to the final version of the manuscript in order to better explain the Semantic Array Programming (SemAP) paradigm. Here, it is perhaps worthy recalling two main aspects which characterise SemAP as a specialisation of AP and which may be of use to better frame part of the topics briefly commented in the following:

• the modularisation of sub-models and autonomous tasks, paying attention to their concise generalization and the potential reusability in other contexts;

• the use of terse array-based constraints (SemAP semantic checks, de Rigo, 2011) to emphasize the focus on the coherent flow of the information and data among modules – which are often nontrivial in computational science. The SemAP semantic constraints natively apply to AP variables irrespective of their size (e.g. large arrays such as continental-scale geospatial layers). The semantic coherence of the information entered in and returned by each module is checked locally instead of relying on external assumptions. This may be essential especially when different modules rely on different expertise.

This way, even the essential implementation details within each module (for example, the implementation of the erosivity layer in the e-RUSLE as a climatic-driven composi-
tion of an array of local empirical relationships) may be at least partially decoupled from the overall modelling architecture. Ideally, atomic modules might easily be replaced by more complex compositions of arrays of sub-modules and data, without implying a major change in the modelling architecture. For example, the same methodology exploited for the erosivity layer – based on a climatic-driven composition of an array of atomic pieces of information (Relative Distance Similarity) – was also exploited in Bosco et al. (2013) for estimating landslide susceptibility. While it is impossible for the added semantic checks to catch all relevant sources of software uncertainty, the array-based semantic modularisation plays a twofold role in promoting good software engineering practices (often neglected in applied computational science, Joppa et al. 2013; Sanders and Kelly, 2008) such as information hiding between modules (“to minimise spread of change between system elements”, Lehman and Ramil, 2003) and at the same time in preserving the plain “readability” of key mathematical peculiarities and relationships among numerical variables. SemAP is also meant for non-experts in the particular domain of a given specialised module to be able to understand at least a subset of semantic requirements not to break when perturbing its input information from outside.

The detail and availability of harmonised data

[Comment] – “The model used in this exercise is the RUSLE to which an effect of rock fragments is added. My main objections to this paper are based on the choice of the model and its use (or misuse). It seems to me is that you did no efforts to represent a field scale model at a scale where cells may contain several fields: you did not mention cadastral maps among your data bases; it seems that you have not attributed a range of possible field sizes among which to choose the more correct one for any particular place using some criteria (e.g., fields nearby towns are smaller than far away fields). Maybe you calculated sediment accumulation flow. In this latter case, how? From
divide to permanent drainage lines? Which were the effects on the L factor? More or less the same comments, linked to the scale issue, can be done for the other RUSLE factor”.

**[Reply]** – The presented map has been calculated at 100x100 m resolution, using the most robust freely available public datasets, and aggregated at 1x1 km resolution. Due to our effort towards reproducibility, and in order to provide a new architecture applicable all over the world, it was not possible to collect sufficient detailed data to apply the suggested improvements all over Europe.

Europe shows a peculiar administrative heterogeneity with 28 member states in the European Union and a variety of official languages. Several countries are internally organised with a broad autonomy in their administrative units (which may result in non-centralised data collection even within a given country). Therefore, the use of more detailed local information for reducing the data uncertainty might easily rise as a drawback a cascade of problems in how to best harmonise uneven datasets which often may even differ in the definition of their categories. Therefore, we have chosen to exploit widely available and recognised datasets such as the Corine Land Cover (CLC, European Environment Agency, 2011) which has explicitly been designed to mitigate as much as possible these heterogeneities (Bossard et al, 2000). CLC has been exploited for USLE/RUSLE based approaches in different areas, such as Southern Italy (Terranova et al., 2009) and Slovakia (Šúri et al., 2002; Diodato, 2011). Undertaking a possible harmonisation effort at the pan-European scale by directly starting from uneven local data would be very challenging. Validating its performances would be even more challenging, in particular to demonstrate that the undertaken effort is able to outperform the overall accuracy of dedicated enterprises such as the CLC. The USLE/RUSLE family of models is also used extensively at national level. The recent assessment of data collected through a European Network (EIONET) (Panagos et al, 2014) showed the majority of the European member states to use RUSLE approaches for estimating soil loss rate by water erosion.
The purpose of the proposed modelling approach: the need for wide-scale (less accurate but harmonised) assessment of soil erosion by water for policy support

[Comment] – “Another important part is the definition of what we want to achieve: are you interested in present export of sediments? Or do you only want to know to the present rate of erosion/sedimentation on site? Or do you want an index of soil erosion (which is not to the real value)? The third one can be successfully approached by using some product of the USLE-family of models (once re-scaled). But is an erosion index the only goal? Or are you also interested in predicting what erosion will become in 10-20 or 50 years from now? Then USLE-derived models are useless unless they are re-written because USLE-like models do not isolate climatic factors (see further comments below) apart rain intensities and totals”.

[Reply] – The soil erosion indicator adopted in this paper is the estimated soil loss ($t \text{ ha}^{-1} y^{-1}$) as described in detail by Huber et al. (2008). As mentioned in the paper (page 2662, line 4) readers should be aware that the proposed map provides an overview of the soil erosion susceptibility at European level and not the actual rate of soil erosion on site. At the same time our effort for implementing a new technique for calculating the R factor within the model and for selecting the better and more robust approaches to calculate the other factors, jointly with our attempt for estimating the plausibility of the map, go in the direction to reduce the gap between our estimation and the real mean soil loss rate on site. Although parameterising the e-RUSLE model is not simple if good results are to be achieved in many different geographic locations, process-based models require considerable efforts to obtain appropriate parameter values in order to run them. This, and their failure to produce better results than achieved using the USLE/RUSLE family of models (Tiwari et al., 2000), encourages the use of the USLE/RUSLE model in applications for which it was not designed (Kinnell, 2010). Furthermore, the availability of a harmonised first level of approximation for estimating soil erosion by water at the pan-European scale may provide a unified
benchmark for the qualitative rapid assessment of erosion impacts. For example, de Rigo et al. (2013) and Di Leo et al. (2013) applied the methodology for the rapid support of wildfire related operational decision-making within a harmonised strategy for assessing many different sources of uncertainty.

The spatial resolution: the working resolution and the final aggregated one

[Comment] – “The modelling architecture: Is the USLE/RUSLE model applicable at 1x1km resolution? Personally I don’t think so, especially when the lower pixel size is 90x90 m. It seems to me that we are playing at producing colored maps unless the model has been changed enough to “average” the behaviour of the processes (already simplified and lumped inside the RUSLE), i.e. I believe that we need a rewriting of the RUSLE for the purpose/scale of application. This implies changing both the model and its input parameters”.

[Reply] – As mentioned before the e-RUSLE model has been calculated at 100x100m resolution and subsequently aggregated at 1x1km. Being the K factor derived by a 1:1,000,000 soil dataset and because of the need to provide a picture of the susceptibility to soil erosion by water at continental scale, we considered the 1x1km resolution as more appropriate for presenting the final results.

Topography, runoff, climate and the detail of reliable pan-European information

[Comment] – “Have you retained anything of the approximation made by Mitasova
and co-workers? And what about their modelling of the sediment fluxes which were both divergent and convergent, following the topography? What about DTM artifacts such as local minima where sediment can be trapped (but should not)? And what when the local minima are dams or karst or pseudo-karst sinks? Procedures for dealing with these two cases can be found in the GRASS software. Did you retain them? When along a slope you have a cascade of land uses, soils, slopes and slope length how do you operate? Do you use an average soil erodibility, S and C? do you use the total slope length or there is some sort of max admitted length (or max contributing area)? What when your unit cell is cut by roads? (asphalt or dirty, roads divert fluxes, and Europe is hyper-dissected by roads). What about property subdivisions, which call for canals, cumulated tillage erosion effects, and large differences in the timing of the agricultural operations?"

[Reply] – Concerning the approximation by Mitasova and colleagues (on which more information will be added in the revised version), in our approach the impact of flow convergence and divergence of the superficial runoff was considered by replacing the hillslope length factor with the Upslope Contributing Area (UCA) (Moore and Burch, 1986; Mitasova, 2002). L and S factors have been determined through GIS procedures already applied numerous times at large spatial scales. For considering local limits capable to affect our approach, we also assumed surface runoff concentrating in less than 300 meters. This value has been selected after analysing the available literature (Renard et al., 1997; Engel, 1999) and considering both the hyper-dissected characteristic of the majority of the territories in South and Central Europe and the more coarse dissection of Northern Europe. Regarding the unit cell cut by roads and other artificial obstacles, Panagos et al. (2012, 2014) have developed the G2 model which is a RUSLE family model incorporating the interception factor. This factor tries to consider features such as roads, paths between parcels, hedges, terrace steeps, cultivation and land use changes using the IMAGE 2006 (Soille, 2006). The effect of this factor on LS was less than 10%, both in Strymonas and in Crete (where G2 has been applied). So, the effect of all those features on soil erosion is relatively small and
the effort to apply the G2 model at the Pan-European scale is too high. The possibility to introduce within the C factor additional information regarding tillage erosion and the different timing of agricultural operations has been considered. Unfortunately, due to the lack of detailed information within the CLC regarding arable lands we considered that the additional level of explained uncertainty would have been negligible. One of our aims will be to consider these aspects in the future, extending the SemAp architecture applied for calculating the R factor to the calculation of the C factor as well. Local artefacts of the DEM as local minima have been filled using the ‘Fill’ tool of ArcGis (ESRI, 2011), it fills sinks in a DEM to remove small imperfections in the data. Concerning the presence of dams, a correct processing of their overall effects on sediment transport and storage would require detailed water reservoir management information. The impact of common management practices such as hydro-peaking, sediment sluicing and flushing is essential. While dam sediment deposition is a well-known phenomenon (Verstraeten et al., 2006; McCartney, 2009), Brandt (2000) underlines how “during sluicing, the sediment transport rates are equal to those of natural flows, and during flushing the rates are equal or higher than those of natural flows” (Wang and Hu, 2009, report a sediment releasing efficiency of 2,400–5,500% for empty/free-flow flushing). Unfortunately, taking into account a more realistic trade-off between forcing factors and feedbacks in the relationship between erosion and water reservoir management would have required unavailable management data. In order for the management history to be approximately reconstructed, the likewise unavailable detailed information would be needed on the local hydro-power energy market as well as on the other key water management criteria (irrigation, industry, flood protection, ...) and other policy-driven management constraints associated to each dam (Castelletti, A., Soncini-Sessa, 2006). Although the progresses in approximating the core patterns involved (de Rigo et al, 2001; Chenga et al., 2014), the complexity of this reconstruction remains prohibitive for a systematic assessment at the continental scale.
[Comment] – “Line 2649-23: Runoff is taken into account by R (rainfall energy and peak runoff as I30) and L (runoff accumulation), S (factor belonging to overland flow shear stresses and flow velocity), partly by C (raindrop interception and part of hydraulic friction) and K (soil profile permeability class). In other words, the whole structure of the RUSLE contributes in implicitly defining runoff, its shear stresses and transport capacity. It is not a limitation of R alone, it is a limitation of the whole model. RUSLE is an example of a model built following a purpose, that of evaluating soil loss at the field scale, (originally only East of the Rocky Mountains): I feel that using it at other scales means rewriting it. Furthermore: if climatic conditions change, also runoff changes and then all the mentioned parameters are to be tuned to the new situation”.

[Reply] – We agree with the considerations related to runoff. We will extend these few lines to better explain the concept. An empirical rainfall–runoff model is embedded in the RUSLE, in particular within the R factor an empirical relationship tends to exist between runoff amount and E, and between peak runoff rate and I30. However, the EI30 is not able to deal with the effect of runoff on soil loss at all well at some locations (Foster et al., 1982; Kinnell, 2010). In some cases the RUSLE estimates average event soil losses reasonably well but, in other cases, particularly when soils have a low runoff coefficient, it over-estimates low event soil losses and under-estimates high event soil losses (Nearing, 1998). Arguably, the failure to include direct consideration of runoff in the R factor is to a large extent responsible for this result (Kinnell, 2010).

[Comment] – “Soil erodibility: there are strong evidences pointing at a climatic effect differentiating erodibility values in at least 2 classes, (Borselli et al., 2012, Catena. 97, 85-94 – which already includes a rock fragment effect). Substantially semiarid or arid climate, or climates with hot dry periods have lower K values than temperate, cool climates”.

[Comment] – “Topographic factor: the effect of climate on it is shown by differences in the length effect on soil erosion. The majority of studies show an increase in soil
loss per unit surface with increasing slope length. Nevertheless, Borselli et al. (2008. Catena 75, 268–277) put an upper value to the max extent of the upslope contributing area in order to bring RUSLE predicted erosion into acceptable levels (in agreement with local observation). In central Sicily, where there is one of the best soil loss experimental stations presently active in Europe, the empirically determined L factor decreases with the plot length (Bagarello and Ferro, 2010, Biosyst. Eng., 105, 411–422). Yair and RazYassif (2004, Geomorphology, 61, 155-169) published similar findings. The quoted authors interpreted their data as an effect of how runoff is built and on how effective is runoff to transport sediments”.

[Reply] – The climate is an important parameter affecting several factors involved in soil erosion processes. We agree that soil erodibility and topographic factor are both influenced by climatic conditions. Borselli et al. (2012) proposed an interesting approach for considering the impact of climate on soil erodibility. The knowledge on the relationship between climate and erosion processes is rapidly evolving (Rempe and Dietrich, 2014). The promising novel techniques applied at local scale would require to be tested and validated at regional scale as an intermediate step before exploring their applicability at the European level. As already mentioned the proposed map provides an overview of the susceptibility to soil erosion by water at European scale. We are conscious of the limits and of all the uncertainties contained in the presented modelling architecture and the use of SemAP as architectural support for the e-RUSLE is meant to ease future continuous improvements. Extending the new technique for calculating the rainfall erosivity to the remaining factors of the e-RUSLE model will be considered in the future evolution of the model. It will be interesting to assess and potentially incorporate into the model promising techniques such as KUERY or the calculation of a connectivity index.

[Comment] – “Validation: Using Google Earth (GE) is a good idea but with some problems. Your model depicts an ideal situation with no interaction with anthropic features
such as roads and other infrastructures that instead tend to densely cover large part of the European landscape so your predictions cannot predict what is caused by these interactions. GE has a resolution, which is at the limit to see erosion features unless you simply look for bare soil as more eroded than vegetated spots. Using panchromatic you can see lines corresponding to rills (not always). Hence you actually need a validation of the erosion maps made using GE versus known (measured) averaged annual soil erosion totals. Moreover, you can count on several (not many) images. Every image is a single photogram that you compare with a mean annual value (i.e. an annual total averaged over many years). All this casts serious shadows over the validation, even if this validation is very much better than nothing”.

[Reply] – We agree with the reported considerations regarding the limits of our approach but some clarifications are needed. The common local-scale validation procedures are not technically and financially feasible at wide spatial extents. We applied a qualitative plausibility check based on expert judgement for assessing the validity of modelling estimates. The methodology is inspired on the categories for field validation of Warren et al. (2005) and also to Berry et al. (2003) and Kapalanga (2008). As stated by the reviewer details on the presence of sediments, presence of litter dams or rills are not always easy to discriminate when solely relying on pictures even considering the good quality of the pictures uploaded in Google Maps and Street View. When it was not possible to gather these details from the available images, additional information such as land cover or landscape characteristics (e.g. soil type, steepness, stoniness) have been considered for estimating the soil erosion rate. During the validation process we also tried to consider the interactions with anthropic features and activities. In the associated validation report (Bosco et al., 2014) the limits and considerations risen during the plausibility check have been reported, in order to better explain our decisions. Within the revised paper the term validation will be replaced with the already adopted and more appropriate ‘plausibility check’.

A revised version of the manuscript will be submitted at the end of the open discussion.
by integrating the changes, additional explanations and literature to meet the requirements of the reviewers.

References


Engel, B. and Mohtar, R.: Estimating soil erosion using RUSLE (Revised Universal Soil Loss Equation) and the Arcview GIS. Purdue University, 1999.


Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 2, 2639, 2014.