Author's final response to referees

We thank Zoran Perić, an anonymous reviewer, and the editor for their comments.

Reviews below are shown in normal typeface with comments in blue. We have made all changes following the online discussion. We have also corrected a few typographical errors in the manuscript as suggested by Referee 1. For a better overview, we have copied-in the previously given responses to the reviewers below.

Review by Referee#1:

In the last decade or two, a large amount of new absolute chronological data from European loess sediments has been published, mainly for the last glacial. Most of these are luminescence ages, a smaller part radiocarbon dates. Although some reviews of mass accumulation rates calculated from previously published age data have been published in recent years, the authors have recognised the need to create a single database, which can further be developed in a uniform way. ChronoLoess is a great initiative and the authors have obviously invested a lot of time and effort in the creation of this database of over 1400 entries. They have also partially revamped the previously accepted methodology of extracting useful information from the published absolute ages and in this very exciting manuscript they also present and discuss the spatial and temporal variations of loess accumulation focussing on the North European Loess Belt (NELB) and the Perialpine Loess (PL).

The language of the manuscript is excellent, the structure is logical and the arguments are easy to follow. The methodology is well summarized and seems scientifically correct, with only two comments from my side concerning the 1) calculation of Rb concentrations from the K contents and the 2) treatment of obvious age inversions in chronologies of some of the sites (e.g. Bialy Kosciol). Regarding the first issue: the major problem here is that the equation published by Mejdahl (1987) systemtically overestimates Rb from K and the correlation between these two elements are rather weak in reality (see Fig 3 and discussion in Buylaert et al., 2018, Radiation Measurements 120, 181-187). Please comment on this and the second question on age inversion.

Citing the online discussion:

(1) We had pragmatically selected the Mejdahl (1987) calculation of the Rb concentration (if not determined by the study authors in the first place), because this was the option available in DRAC (Durcan et al., 2015) and it appeared more meaningful to include Rb than leaving it out. This follows Mejdahl (1987), who concluded that in case of an absent of a measured Rb concentration, the dose rate is larger if Rb is omitted compared to situations where only a rough estimate is used.

Buylaert et al. (2018), who investigated the dependency of the equivalent dose on the potassium concentration on single grains, showed, as Huntley and Hancock (2001) and Mejdahl (1987), that the true Rb concentration is likely to vary tremendously between samples when measured. Undeniable, the data in Buylaert et al. (2018) provides evidence for their samples that the method by Mejdahl (1987) would systematically overestimate the true Rb concentration compared to their measurement of single grains. This might be true for some multi-grain values in the ChronoLoess database, while it might not match other situations.

Therefore, we suggest adding a few lines to the revised paper to address the concerns but stick to the Mejdahl (1987) approach. Suppose better data (e.g., a combination of Mejdahl (1987), Huntley and Hancock (2001) and Buylaert et al., 2018) becomes available. In that case, this can be implemented in an updated version of DRAC and hence used in the ChronoLoess database (Durcan et al., 2015).

(2) To address the equally justified concerns of age inversion, we believe this should not generally matter for our analysis but reflects results typical for loess dating studies. The age inversion

might be true or just an extreme 2value. Regardless, the weight for the overall interpretation of the European loess deposit is negligible.

The main conclusions of the authors are correct and I can agree with most of them, although there are some comments that I would suggest for reflection or improvement. One critical issue is how to determine in a statistically exact way when the time-activity curve crosses the uniform random distribution (URD), as the authors have used this as a basis for inferring the initiation of significant dust accumulation and have correlated this with certain Heinrich events (HE). In my opinion, this has considerable uncertainties on the order of 1000-3000 years that need to be taken into account, however, I believe that the dust accumulation peaks are well defined (NELB: 21.8 ka, PL: 23.9 ka) and may indeed be causally related to HE2.

The three activity curves in Figure 4 provide the answer to the question posed: the error envelope intersected by the uniform distribution indicates a time uncertainty that is effectively of the order of 1000 to 3000 years, depending on the curve. The exact statistical determination of when the activity curve intersects the uniform distribution is therefore done graphically, once the error envelope has been calculated exactly according to the protocol described in the article (sections 2.1.3 and 2.1.4), at a confidence level of 95%.

I have further suggestions, especially for certain interpretations, such as the potential causes of the time lag in the aforementioned peak accumulation (NELB vs. PL). This may indeed depend on several factors, as the authors rightly state, but these factors need to be reconsidered/refined. However, these are rather pedantic points and do not affect the main conclusions. I believe that if the authors improve these points in a small revision and the other peer reviews end up favourably, this paper can be published in ESSD.

Below we list our responses along with changes.

Line 141: Specify what kind of events are you talking about.

We suggest modifying to '*the number of events (here, the number of dated dust accumulation events) that have occurred per unit of time.*'

Line 219: picture of past dust deposition?

We changed by '*a reliable picture of past dust deposition'*

Line 249: Geographical regions and countries are mixed up here. For example, instead of England, you could write British Isles, or instead of Poland, you could find a suitable definition of a geographical region.

We changed by 'Southern England', 'Southern Poland' to be consistent with the other regions…

Line 274: Add Újvári et al. (2008): Origin, weathering, and geochemical composition of loess in southwestern Hungary. Quaternary Research 69, 421-427, which actually first proved this relative homogeneity for European loess.

OK. We added the reference Ujvari et al. (2008).

Line 290 and following: The fundamental question here is how the authors define the starting date of increased dust accumulation. For the NELB, the time-activity curve leaves the URD around 32.5 ka, correctly identified by the authors. If we take this as a basis, then the same happens for PL at around 42 ka (GS-11, based on the most likely scenario: bold line), in contrast to the authors' determination (40 ka, GS-9).

Indeed, a further issue here is the uncertainty of the chronologies, which applies not only to the Heinrich events, but also to the time-activity curve based on the luminescence ages. If we consider the most likely scenario represented by the thick curve and its error, it is clear that it is not possible to say unambiguously (on a decadal/centennial scale) exactly when the dust accumulation crossed the URD. The uncertainty of this is 1-3 thousand years. From this point of view, any discussion of a correlation between Heinrich events and a significant increase in dust accumulation can only be theoretical and cannot be proven.

However, I should note here that the age of this dust accumulation peak is much better defined (even if I don't see any error bounds), so I can basically accept the observation that it is correlated with Heinrich Event 2.

Thank you for this comment. We suggest the following rephrasing: '*In the former area, dust accumulation increased from about 32 ka b2k, i.e. during GS-5.2 (Rasmussen et al., 2014) and rose steeply after 30 ka b2k, i.e. during GS-5.1 (Fig. 3B). In contrast, deposition in the perialpine area started earlier and increased as early as 42 ka b2k (GS-11), with a further rise after 40 ka b2k (GS-9) (Fig. 3C). Although the chronological limits of Heinrich Events (HEs) remain relatively imprecise, GS-5.1 and GS-9 correlate with HE-3 and HE-4 respectively in marine records (Sanchez Goñi and Harrison, 2010).*'

Line 326 and following: Do you suggest that the identified time lag in accumulation (~on the order of thousands of years) may have resulted from significant differences in wind transport capacity between the NELB and PL?

In my opinion, sediment is available even if the characteristics of the eolian surface do not allow its removal. So soil moisture or vegetation cover affects the properties of the eolian surface and has nothing to do with sediment availability. Low moisture content or temperature of the eolian surface, sparse vegetation and/or no crusts on the surface may all contribute to an actively emitting source (see Újvári et al., 2016: The physics of wind-blown loess: Implications for grain size proxy interpretations in Quaternary paleoclimate studies. Earth-Science Reviews 154, 247-278). I suggest changing 3) to the "nature of the eolian surface"...which is actually dependent on paleoclimate.

The factors listed according to Kocurek and Lancaster (1999) are those conventionally evoked when working on aeolian deposits in the geological record, although not all of them are necessarily relevant here. We suggest a minor modification of the phrase: "*In agreement with Kocurek and Lancaster (1999), several factors may have influenced this time lag, such as (i) the amount of fine particles released by their respective sedimentary sources; (ii) the wind transport capacity; (iii) the local availability of sediments (role of vegetation and soil moisture). Among the various potential factors, fluctuations in the amount of particles available for deflation due to changes in ice sheet pattern appear to be a pivotal point that explains the chronological disparities between aeolian systems*."

Line 330: Are the authors referring here to the MIS2/3 boundary or just in general to an age at the end of MIS 3? Because the boundary is around 28,000 years.

OK. We will change to '*During late MIS 3'*.

Line 336: (Fig. 4)

Yes, we have added the reference to Figure 4.

Lines 421-422: Note that not only the height, but the density, shape, arrangement and flexibility of roughness elements (e.g. vegetation) are influencing the surface roughness, which will eventually define particle mobilization and accumulation. On top of that, particle settling and accumulaton is grain size dependent (see e.g. Újvári et al., 2016 ESR and references therein).

We agree with this remark. But we don't think it's necessary to expand on this point. Moreover, we have developed this idea in another paper (Bosq et al., 2018) cited in the text.

Line 437: Relative to what?

We have removed “small” as it wasn’t relevant.

Figure 3: Write MIS instead of SIM in this uppermost panel of Figure 3.

Changed

Review by Referee#2 Zoran Períc:

The study “Last Glacial loess in Europe: luminescence database and chronology of deposition” by Bosq et al., aims to create a unique database of European loess osl dated chronologies and accumulation rates during MIS 1-MIS 3 period. The data appears robust, and the methods used are sound and appropriately applied.

In my opinion this is a comprehensive study which will be a significant contribution to the loess, geochronological and palaeoclimatological community. The manuscript itself is well-structured, language is appropriate, the methods ant data are well presented and the conclusions are sound. The figures are also well designed and complement the data accurately. I would like to highlight that the authors did an excellent job with the manuscript and I very much enjoyed reading it.

Thank you very much for your positive and very supportive comments; we appreciate it.

I recommend publication; however, I would only like for the authors to take into account a few of my questions.

- Have you considered the dating resolution as a parameter when calculating MARs? In my experience, dating resolution can have a significant impact on the accurate calculation of MAR distribution over the investigated time period. Maybe you should mention that.

Indeed, the dating resolution has an impact on the calculation of MARs. Notably, this is the case for extreme MARs, i.e., the highest values for a given sequence. Long intervals between dates mean the MAR is averaged over the period. For instance, high resolution (as is possible using radiocarbon on molluscan shells and earthworm granules) tends to provide high extreme MARs over short periods.

In our case, where possible, we used both OSL ages and radiocarbon ages for each sequence analyzed to obtain the most accurate age-depth model. With this, the obtained values are close to those published by other authors, e.g., Ujvari et al. (2017). This is likely caused by the number of dates available for the most significant dust accumulation periods. For example, nine levels have been dated for the Balta Alba Kurgan LPS between 35 ka b2k and 24 ka b2k, and up to 13 levels for the Bialy Kosciól LPS between 28 ka b2k and 20 ka b2k.

On the other hand, it should be kept in mind that while the individual sampling strategy is always a trade between what is economically feasible and what is required to answer the research question, study authors usually sampled at spots where changes were visible in the stratigraphy. This, in turn, means that some short extreme MARs might have been overlooked. Still, since the sampling usually bases on field observations, sampling was also not unjustified, reflecting the actual situation on site.

We added the following sentence in subsection 2.2.3:

“*The dating resolution has an impact on the calculated MARs. This is particularly true for extreme MARS, i.e. the highest values for a given sequence. Long intervals between dates necessarily result in averaging MARs over the considered period. In this study, the chronological models consider the recalculated luminescence ages (extracted from the ChronoLoess database) and AMS ^{14}C ages where available. Since many dates are available for the periods of greatest dust accumulation in most* *sequences, for example, nine levels have been dated for the Balta Alba Kurgan LPS between 35 ka b2k and 24 ka b2k and up to 13 levels for the Bialy Kościól LPS between 28 ka b2k and 20 ka b2k. We consider the extreme MAR values obtained representative, although we acknowledge that we cannot exclude their exact timing might partly suffer from a dating resolution-related inaccuracy*."

- Have you thought on maybe applying an additional age-depth modelling method and compare the results? I acknowledge that this would be a large amount of additional work, but often different modelling software produces different results. It would be interesting to see the discrepancies. This is just a thought, not a critique point.

We did not use different methods for calculating age-depth models. Nevertheless, we could compare our results with other authors' results, at least for some LPS. The comparisons showed good agreement between the data outlined in the text; therefore, we do not expect to see much different results from this comparison. However, it might be an exciting follow-up study, perhaps with an updated ChornoLoess dataset.

- Have you thought of expanding the investigation beyond MIS 3? I do know that these chronologies are very scares, but there are some studies from the Carpathian Basin where LPS were dated beyond MIS 5.

We had restricted ourselves to the period younger than 60 ka b2k for three reasons: (1) Most of the data are available for this period and (2) considered more reliable given the limits of luminescence dating than for earlier periods, and (3) it would have taken a considerable amount of extra time to extend our database at this time beyond the MIS 3. Given the technical difficulties involved with the different luminescence dating methods, we would have ended up debating the accuracy of specific dated sites. At this stage, this was simply beyond the scope of our approach. Nevertheless, as written in the manuscript, we are planning to gradually extend the ChronoLoess database and add new and more ages as they become available, and this will also include dates older than 60 ka b2k.

- I think it would have been interesting to compare these results with MARs from LPS outside of Europe (e.g. Chinese Loess Plateau). We might get a better insight into the atmospheric dust activity on an intercontinental scale. However, I do recognize that this might be beyond the scope of this paper. Please take this just as an idea and/or a suggestion.

We agree, but it would first require a similar dataset for other sites, and then that information can be combined. Perhaps this becomes possible if others pick up the chosen approach for different regions so that, at some point, that information can be combined and compared. Unfortunately, at the current state, this goes well beyond our study's scope, but we are very thankful for the suggestions.