JAAH 2016 No 19. Herschend
Log book

Frands Herschend. Towards a standardized discussion of priors in Bayesian analyses of 14C dated archaeological periods: A study based on the dates from Gjøsund.

Received:	2015-11-04
Length:	*c*. 8000 words
Editorial Comment:	The article complied with JAAH policy and the editors approach two referees 2016-01-25 2016-03-30 Report from Ulf Strucke 2016-04-13 Report from Andreas Hennius 2016-04-13 Reports sent to Herschend 2016-06-13 Final version w images
Language revision:	Yes, as part of the editorial process. Completed 2016-08-31
Copy rights and credits:	Must be fixed
Authors' and reviewers' comments:	Received and collated. Published 2016-09-23

Reviewer 1

Ulf Strucke

The paper is concluded with a comment from Hans Hildebrandt's Från äldre tider. Kulturvetenskapliga och historiska studier.
*Nevertheless, it is hazardous to refute a conclusion drawn from that which is known supported by conclusions that are unknown. [Det är dock vanskligt att söka häfva de slutsatser, man drager från det man vet, med hjälp av slutsatser man inte ve*t]".
The quote, which could very well have introduced the entire text, is indeed relevant as the discussion concerns a test of a chronological hypothesis based in Bayes' theorem. This theorem builds on the concept that it is possible to manipulate the statistical result using certain parameters. These parameters form the basis for a statistical process.
Herschend's paper is an educational and well-supported methodological study. The process has a great advantage in that the selected strategy and choices can be clearly followed to the selected hypothesis. Additionally, the validity of the parameters can be tested by repeated calculations and adjustments.

The introductory part of the paper is in my opinion unnecessarily extensive, describing thoughts about the situation of field archaeology. It is mentioned in passing that the accuracy of the laboratories may be equivocal. Within the frame for the quality studies continuously done by the laboratories (Scott 2010M Scott, M., Cook, G. Naysmith, P. 2010) it is established that with few exceptions, the quality is comparable. The part of the paper also touches on the number of datings that is considered suitable for an archaeological investigation; considerations are obviously made for the different conditions of the excavation. The number of datings is as decisive for the result as the criteria put up in selecting the material. This is then followed by a method chapter. Disregarding Fig. 1, which to me is more confusing than explanatory, the presented process flow is well thought through. However, I would still prefer that in a model, the chronological process is taken into account already when planning the investigation. The aim of creating a before-during-after may cause consequences in the fieldwork process; e.g. seemingly unimportant structures in upper soil layers may be the "after" that is sought.
Furthermore, I would like to reformulate theme ll:III to "The interpretation and relation of sub-contexts…", since analyses of finds and e.g. archaeobotanical material in the surrounding structures may provide information on distribution and origin of the material selected for ^{14}C-dating.

On to Gjösund. The introductory parts establish the site as well as the natural conditions, followed by a survey of the cultural characteristics of the building with comparisons. In this case, it is of course impossible to maintain a planned statistical process of the planning procedures. However, the building typology provides a clear guide to the highest possible age.
The charcoal was mainly analysed for wood species. Birch was mainly selected, hazel to a lesser extent and pine in two cases. The aim was to avoid as much as possible any wood with a high own age and any possible reused older building material. Re-use of older wood is a difficult issue to interpret. Using wood species analysis avoids an initial dating of rotting or decayed wood. A high own age should not only be seen as a problem. If this can be estimated, it is possible to make adjustments for the number when calibrating the ^{14}C-dating. To complement the analysis that was made it may be noted that hazel is generally distributed. One of the hazel datings is thus found outside the expected interval while two borders the earliest datings of the house. Thus it is not unequivocal to state that hazel would reflect an older phase. The problem of the wood selected for analysis is also indicated by posthole H-37, where the hazel is too old, while the pine

is well within the expected interval. This posthole contains a number of species which may be seen as an indication of disturbed layers. Studying the wider distribution of species creates a better understanding of the area. Arable land was found northwest of the house, and some of the constructions were cut through by plough marks. It is thus possible that some of the found charcoal originate in the older arable land. The fact that there are also plough marks across postholes in the house indicates that agriculture has continued even after the house was demolished. Hypothetically, we may assume that the nutritious soil of the cultural layer formed an attractive land for cultivation. Hazel wood functions well both for firewood and fodder, and the weak timber is used for fencing, building etc. The datings of hazel are undeniably problematic, but is seems likely that hazel poles have been useful throughout the lifetime of the house.

The Gjösund house borders the end of the Hallstatt plateau, 800 to 400 BC. It is not possible to define the time more precisely without contextual delimitations. The baseline's fluctuations create periods of accumulated events and apparent periodisations. It is dangerous to unreservedly interpret these variations in cultural terms. The transition from the Early Bronze Age to the Late Bronze Age is such an interval. It has not been completely established whether there is a connection between the variation of the baseline and prehistorical periods, for example the plateau that encompasses the Migration Period. Older ^{14}C-datings with greater standard deviations have hidden these delimitations, but an increased number of high-precision datings would further accentuate the boundaries. This is an additional argument for processing the datings with the theorem. In most cases, a simple congregation of the datings is merely based on the apparent periods. If not before, the problem appears in the final summary in fig 11.
The author admits that the chapter on Bayes' theorem is brief. Perhaps it is a little too short by comparison, since the discussion on the a priori parameters is decisive for the result. Whether one use BCal or OxCal, the most used software, is a matter of taste. Both softwares provide slightly different results as well as graphical presentations. In addition to providing a base for a hypothetical division, the analysis may also be used to check on the likelihood of the relation between find material, constructions and other datable material. An expanded discussion on this, as well as on the risk for circular evidence when used would be good. In the present case, an excerpt from the BCal software is presented. Unfortunately, datings showing the abandonment of the house are missing; thus, this part must be left out. Perhaps indications of cultivation and patina could have dated a subsequent period. This is one example of my initial suggestion to include more discussion on the Bayes' hypothesis already at the initiation of the project.
The statistical processing adheres to a protocol organised by BCal. The result is both presented numerically and as a series of graphs. The author elects to show the datings without rounding off, giving an apparent precision. Considering that all charcoal has an own age, a rounding off to the nearest 5 or 10 would hardly create any greater uncertainty. Unfortunately, such a rounding off means that the three groupings in Table 3 almost coincide. Nevertheless, the paper's aim to closer delimit the use-time of the house has been reached.

The paper is a welcome element of the method discussion on chronology within archaeology. Bayes' theorem has long been used within archaeology outside the Nordic countries. Hopefully, this article is only the first in line. It uses examples of datings in connection with the difficult period after the end of the Hallstatt plateau, but much can also be gained e.g. within the frame of medieval and early modern archaeology (cf. Bäck, M. & Strucke, U. 2003).

References

Bäck, M. & Strucke, U. 2003. Begränsningar inom den historiska arkeologin - metoderna eller vår fantasi? : C14-dateringar och 1600-talets arkeologiska lämningar. Nordiska stratigrafimötet (4 : 2001 : Viborg) Stratigrafiens mangfoldigheder. 2003 s. 19-25

Scott, M., Cook, G. & Naysmith, P. 2010. The Fifth International Radiocarbon Intercomparison (VIRI): An Assessment of Laboratory Performance in Stage 3. Radiocarbon, Vol 52 No 2-3 p. 859-865

Reviewer 2

Andreas Hennius

In the article "*Taking the Gjøsund case study forward – Towards a standardized discussion of priors in Bayesian analyses of ^{14}C-dated periods*" Frands Herschend discusses a model for how you might relate and process ^{14}C-samples, turn uncalibrated results of an analysis (Universal time) into calibrated years (contextual time) and how to interpret analysed samples using Bayesian statistics. Herschend's model is based on four interlinked steps (see Table 1 below).

Table. 1. Overview of the methodical precedure.

AREA:	**I) WHAT IS THE RELATION BETWEEN A ^{14}C-SAMPLE AND ITS ARCHAEOLOGICAL CONTEXT?**
THEMES:	
I:I	The site as a functioning environment
I:II	The prehistoric reality as decaying material phenomena
I:III	The archaeological context as duration
AREA:	**II) WHAT DOES A HUMANLY PRODUCED ARCHAEOLOGICAL ^{14}C-SAMPLE, ACTUALLY DATE?**
THEMES:	
II:I	The own age of the sample
II:II	The varying number of ^{14}C-sample traps
II:III	The interpretation of sub contexts containing ^{14}C-samples
AREA:	**III) HOW IS UNIVERSAL TIME DEPICTED BY ARCHAEOLOGICAL 14C-SAMPLE?**
THEMES:	
III:I	Sample production rate
III:II	Relative number of effective ^{14}C-sample traps
III:III	Sample preservation
AREA:	**IV) HOW, SHOULD AN ARCHAEOLOGICAL ^{14}C-SAMPLE PROBABILITY DISTRIBUTION BE TREATED?**
THEMES:	
IV:I	Scientific parameters
IV:II	Overview of probability distributions
IV:III	Bayesian-aided chronology

As case study Herschend uses an excavation from Gjösund in Norway, excavated some years ago. In one of the outer edges of the site, a small building was investigated. The house and nearby features were dated with 19 (or 20?) analysed samples on scattered coal taken from hearths and postholes from both wall-bearing and roof-bearing posts (Diinhoff & Slinning 2013), making it suitable for the Bayesian analysis in Herschend's article. There are many anomalies in the placing of posts and hearths in the house and the construction is not unproblematic; for example, Herschend discusses an asymmetrical roof construction. Some of the problems with understanding the remains might be owing to the fact that the house appeared near the edges of the trench, while other remains may be present outside the excavated areas.

Bayesian statistics forms a branch in statistics that aim to combine different collected sources of information. It also allows a personal judgement and earlier experiences cooperate with empirical facts to reach a statistically reliable statement. Most applications have been within jurisprudence and medicine, but during the past decades the method has been used by archaeologists processing and interpreting the results of ^{14}C analysis. Above all, the Bayesian statistics are useful in relation to a cryptic and non-intuitive calibration curve where the results from ^{14}C analyses may be combined with archaeological interpretation of certain phenomena and with help from Bayesian statistics receive shorter and firmer calibrated datings. The analysis can be done relatively simply using web based programs but as with many other, preferably mathematical/ scientific methods it takes time for archaeologists to understand and embrace new methods.

Frands Herschend's article *"Taking the Gjøsund case study forward – Towards a standardized discussion of priors in Bayesian analyses of ^{14}C-dated periods"* fills a great need in problematizing the interpretation of the results from the ^{14}C-analyse, but I have some doubt that the article in its present form will increase understanding or reach a standardized discussion.
There are certain aspects I would like to see a bit more elaborated, so that the discussion in the article is easier for the reader to follow. In the present version, there is no comprehensive presentation of the original interpretation of the site and the interpretation of the many ^{14}C-analysis from the Gjøsund site, made by Diinhoff & Slinning (2013). It is not clear in what way Herschend's reinterpretation of the Gjøsund material differs from the original interpretation; neither is the division clear on which is Herschend's interpretation and that of Diinhoff & Slinning, or the work behind.

Fig. from Diinhoff & Slinning (2013)

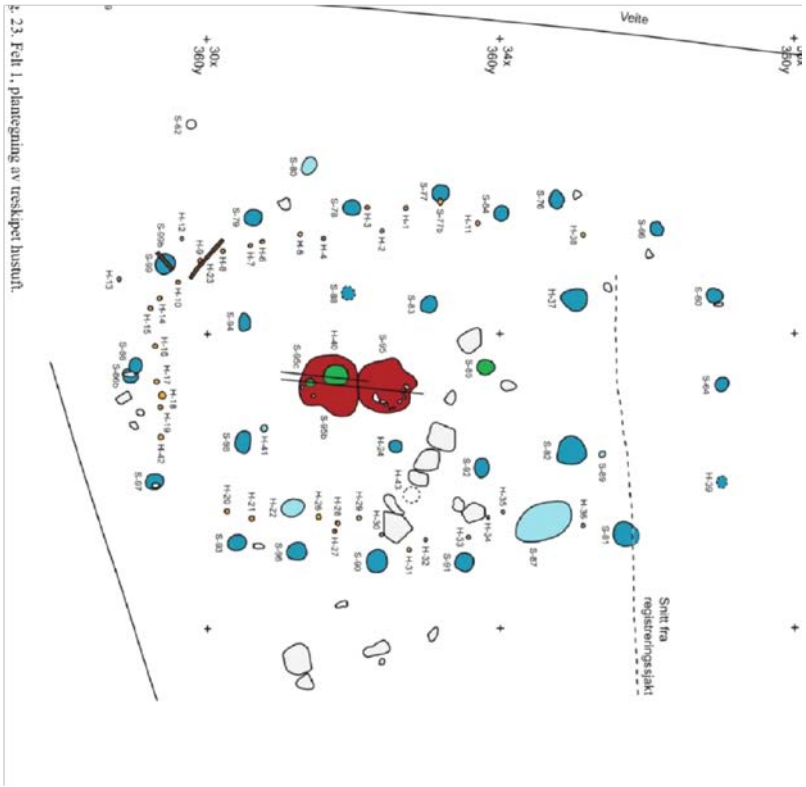
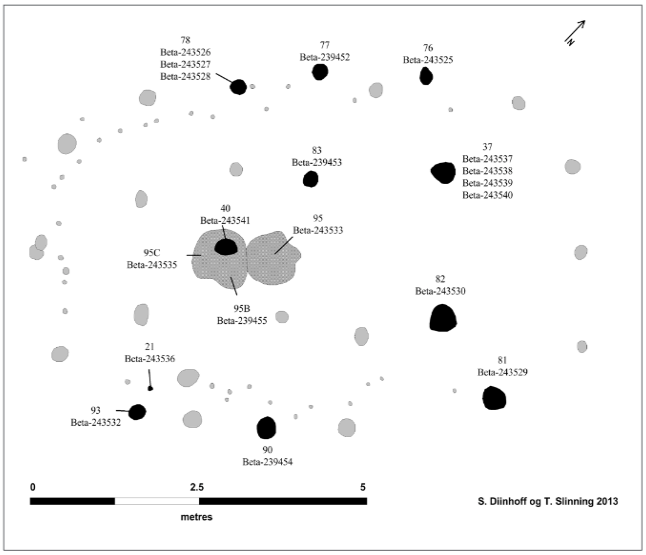


Fig. from Diinhoff & Slinning (2013)



Figur 4. *Fra langhuset ble det tatt ut 19 prøver til radiologisk datering. Strukturene markert med svart fyll viser hvor det ble tatt ut dateringsprøver, og tallene henviser til strukturnummer og derunder prøvenummer. Grafikk Diinhoff og Slinning.*

The second problematic point is also related to this and concerns the interpretation of the ^{14}C datings and the Bayesian analyse. The analysed ^{14}C-datings covers a period of 700 years (uncalibrated BP values) which on an intuitive basis could be divided into 3-4 occupational phases.

Fig. from Diinhoff & Slinning (2013)

Takstolper	2320	2340	2350	2350	2360	2720		
Veggstolper	2330	2340	2340	2340	2340	2360	2460	2470
Ildsteder	2350	2350	2640	2650				
Pælehul	3070							

Tabell 3. *De oppnådde dateringer er her angitt i kalibrert BP, og vises fordelt på de fire daterte anleggstypene. Grafikk Diinhoff og Slinning.*

Fig. from Diinhoff & Slinning (2013) showing the ^{14}C-samples

Dateringsprøve:	Prøve nr:	Struktur:	BP	±	Cal. BP	2-Sigma
Beta-239452	GJ066	S-77	2430	40	2460	2710-2630, 2620-2350
Beta-239453	GJ078	S-83	2260	40	2320	2350-2290, 2280-2150
Beta-239454	GJ092	S-90	2280	40	2330	2350-2300, 2260-2160
Beta-239455	GJ113	S-95b	2470	60	2680, 2640, 2500	2740-2350
Beta-243525	GJ064	S-76	2290	40	2340	2350-2300, 2240-2170
Beta-243526	GJ068	S-78	2290	40	2340	2350-2300, 2240-2170
Beta-243527	GJ068b	S-78	2440	40	2470	2710-2350
Beta-243528	GJ068c	S-78	2310	40	2340	2360-2310, 2230-2200
Beta-243529	GJ074	S-81	2380	40	2360	2670-2650, 2490-2340
Beta-243530	GJ076	S-82	2390	40	2360	2680-2640, 2500-2340
Beta-243532	GJ098	S-93	2330	40	2340	2360-2320
Beta-243533	GJ102	S-95	2460	40	2670, 2650, 2490	2720-2360
Beta-243535	GJ112	S-95c	2370	40	2350	2480-2340
Beta-243536	GJ120	H-21	2920	40	3070	3210-2950
Beta-243537	GJ121	H-37	2350	40	2350	2460-2330
Beta-243538	GJ121b	H-37	2360	40	2350	2470-2330
Beta-243539	GJ121c	H-37	2540	40	2720	2750-2670, 2650-2490
Beta-243540	GJ122	H-37	2290	40	2340	2350-2300, 2240-2170
Beta-243541	GJ124	H-40	2340	40	2350	2440-2410, 2370-2320

Tabell 1. *Radiologiske dateringer fra langhuset på Gjøsundneset. Grafikk Diinhoff og Slinning.*

With Herschend's method of analysing the datings and the subsequent Bayesian processing, the dating of the house is pinpointed to a 150 year long period, which is amazing. Unfortunately, the discussion is very hard to follow. From where are the datings and the diagram taken?
I would like to see a more comprehensive presentation of the analysed carbon datings used in the analysis. The Bayesian statistics helps us to interpret non-intuitive processes, but it is unclear how we get from the above diagram to the results presented by Herschend. Comparing with Diinhoff & Slinning do little to help. The sample numbers do not match and on the table above (Tabell 1 from Diinhoff & Slinning), only 19 samples are presented.

Figs 8 and 10 from Herschend's paper under review

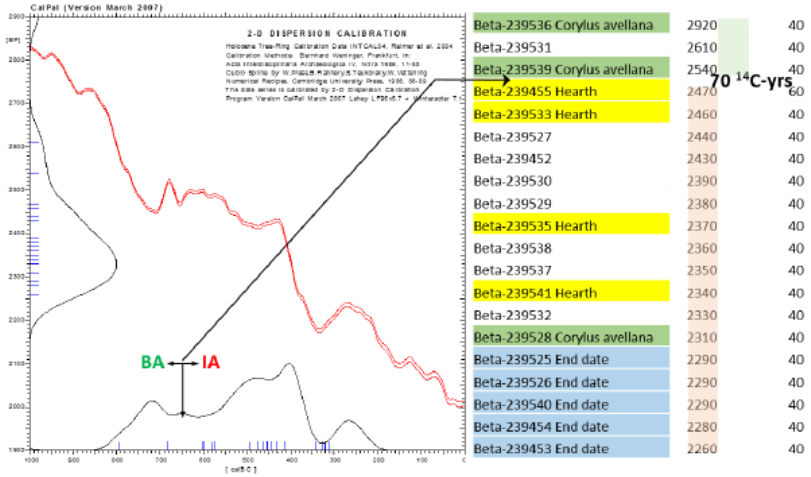


Fig. 8. A general overview of the 14C-dates from Gjøsund. The oldest date falls outside the diagram.

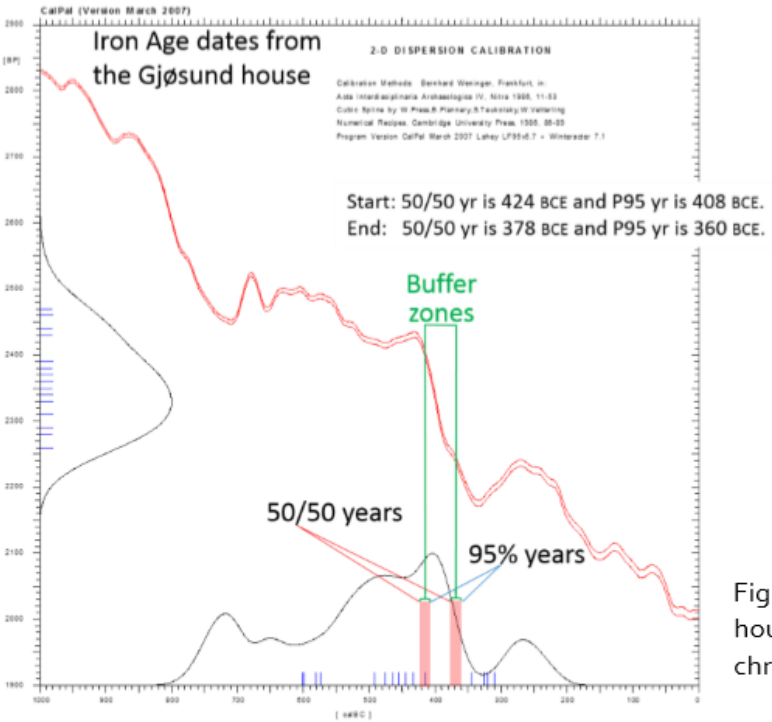
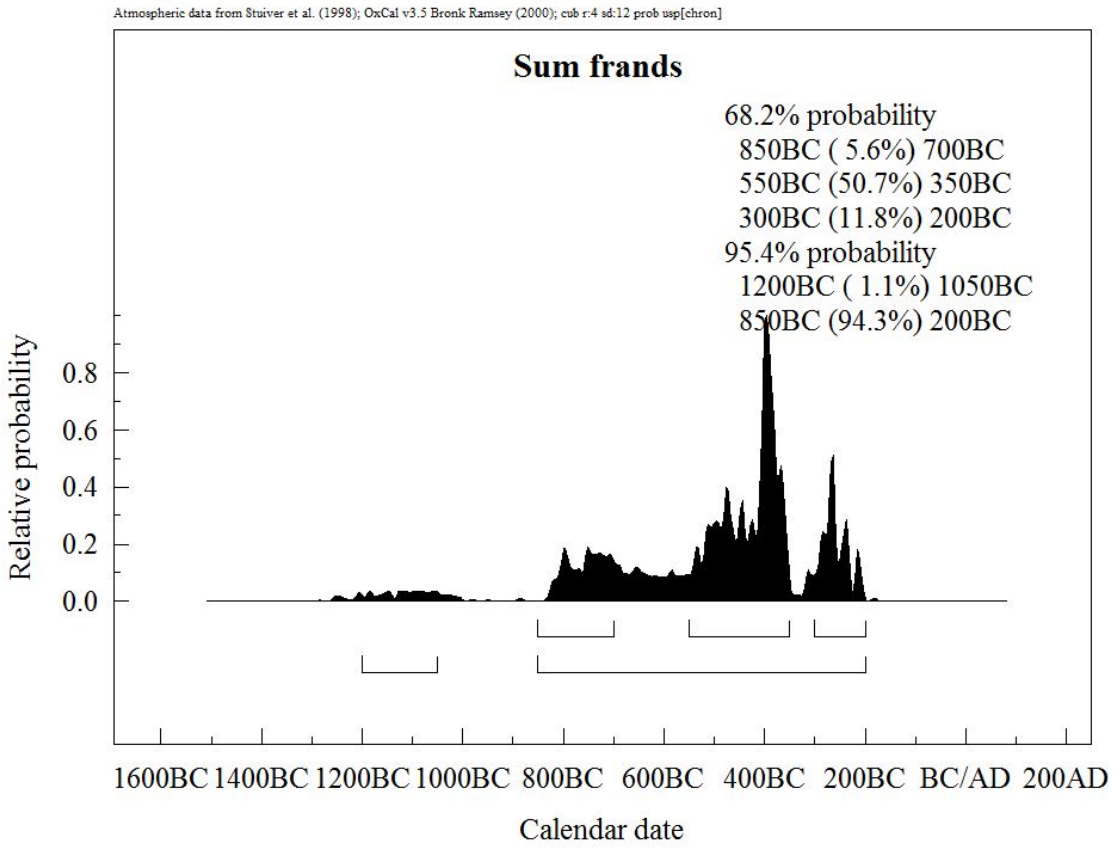


Fig. 10. Dating the Gjøsund house as a period within chronological buffers.



The above is a summarized diagram from Oxcal 3.10. The overall structure resembles the curve in Herschend's presented diagram but not exactly.

Despite the critique of the issues presented above, Herschend raises a very important question about how contract archaeology deals with the large number of ^{14}C analyses produced every year. A model for evaluating each ^{14}C-sample is a very helpful tool. Even if an evaluation is made intuitively, it is useful to have a formalized model. When it comes to the Bayesian analysis, I have some doubts. The normal way of dating houses, graves or other constructions in Swedish archaeology is often limited to one or two samples from each construction, i.e. far less than the Gjøsund case with 19 (or 20?) samples interpreted as belonging to more or less the same construction. The low amount of samples forms other statistical problems for Swedish archaeologists.

Author's comments

Frands Herschend

The reviewers' comments have been most helpful. Firstly, having read them, I have reformulated parts of the article, hopefully making the argumentation clearer, e.g. Table 1 and Fig. 1. I may not have succeeded. Secondly, I have benefitted from a number of important references, which I have worked into the text where they ought already to have been. Additions and corrections have affected 5 – 10% of the text.

I must stress two points and prefer to do this here in my comments rather than in the article:

(1) I do not intent to discuss Diinhoff & Slinning's article. My contribution is not meant to build up our understanding of any settlement in Western Norway. Instead, it is a matter of exemplified methodology and meant to be a contrast to Diinhoff & Slinning, not a critique. I find some of their arguments convincing, but not all of them and I have no intention of correcting or corroborating their argumentation. When it comes to my interpretation of the Gjøsund house, I do not aim at convincing anybody. In order to emphasize that the Gjøsund context is not under any serious discussion or interpretation I have changed the title of my article.

(2) I build no formal argumentation on the diagrams. Only the numbers resulting for the BCal calibration are important and if used, they must be brought into the ongoing methodological discussion of their probability. The diagrams are merely a way of proceeding from the intuitive to the formal. They are illustrations and can be used only in common-sense argumentations leading up to the actual statistical tests. The diagrams are not as detailed as an OxCal diagram, neither are they intended to be. It may be worthwhile to present detail as documentation, but detail in itself is not an interpretative quality as long as intuition discussion governs interpretation, since irrespective of detail, intuition can be no more than superficial.

Needless to say I find the way Reviewer One continues the methodological discussion rewarding and model.

This has stressed two issues, which I think are also reflected by the comments, although the reviewers would not necessarily share my opinions:

- Contract archaeology is trapped between rational research development and stale administration practices.
- It is necessary to make a sharp divide between formal results, in this case based on numbers, and intuitive and discursive results, based on interpretation and illustration.