Answer to Reviewer #1

1 General Comment

This manuscript presents a statistical analysis of blocking in part of the Northern Hemisphere with a focus on the North-Atlantic European region. Using a set of reanalysis data, the authors investigate blocking frequencies and their trends during the period 1990-2019. The study applies a novel method to assess blocking based on the detection of centers of vorticity; this method allows one to distinguish two different types of blocking and to consider the transitions between them.
The paper is a welcome contribution to the discussion, since blocking and related trends are important topics in our science, and in the end it would be desirable that the results will be published. By the use of their specific methodology, the authors are able to quantify a few novel aspects which would be hard to address using other methods. At the same time there is a number of issues which I think should be sorted out before the manuscript can be published.
We thank the reviewer for carefully reading our manuscript, and for the constructive comments. In the following we will respond to the comments and point out any changes we intend to make. The line numbers and figure references in the reviewer's comments refer to the original manuscript. The reviewer's comments are in black italic; our responses are in blue.

2 Major issues

Statistical significance

In several of the figures I was missing a quantification of the statistical uncertainty (e.g., Fig. 6, Fig. 9). In my eyes all results must be tested with regard to their statistical significance.

Thank you for this hint. We agree that all results need to come with an uncertainty estimate. Figure 6 and 9 were meant to give an overview of what can be seen directly from the data (counts and duration), no conclusion was meant to be based on these figures. However, in the revised version, we use Fig. 6 to demonstrate the sensitivity of our blocking identification process, see example below in Fig. 6. To this end, we point out, that we aim to detect the blocks in a Lagrangian sense. Therefore we need to ensure, that blocks identified in consecutive time steps represent the same system, more precisely that the high is the same as in the previous time step. To ensure this coherence, we introduced a minimum distance criterion between the locations of the high center in two consecutive time steps. This distance criterion was set to about 1000 km in our initial submission. In the revised paper, we will estimate the uncertainty of the method based on 10 additional identification procedures that just differ in this distance criterion. We will describe the method more detailed in the revision.

With respect to Figure 9, we now show uncertainties associated with estimating the occurrence probabilities of a binomial/multinomial process, Fig. 9

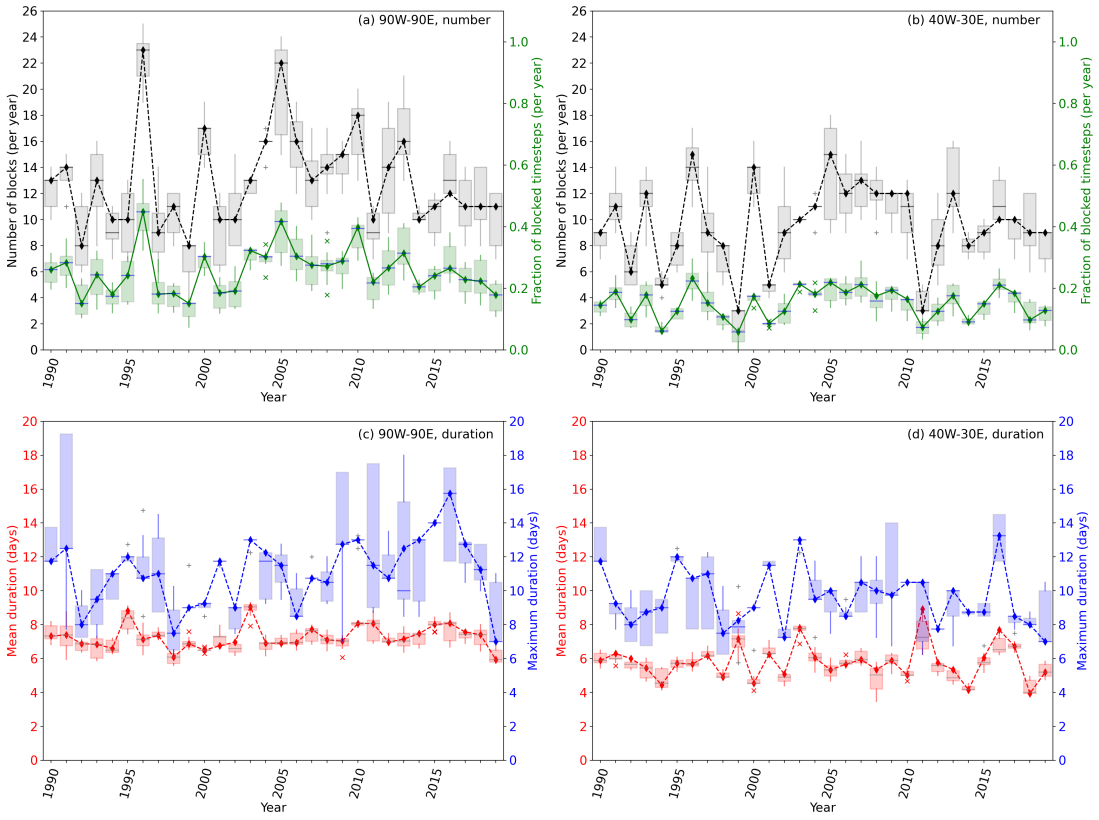
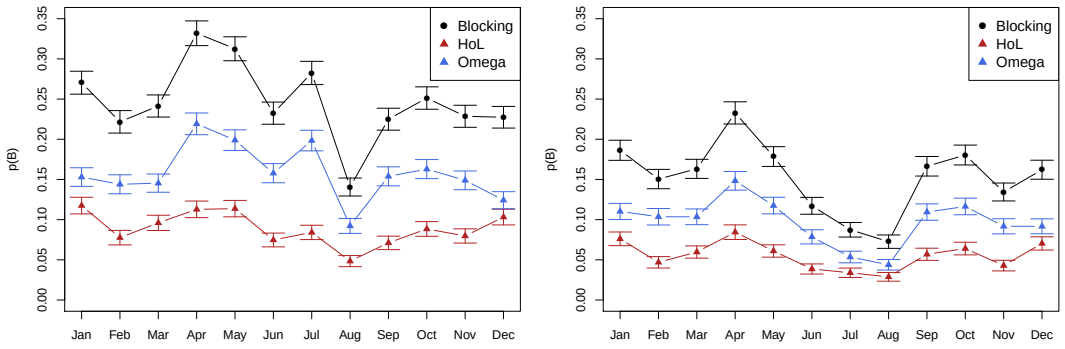


Figure 6: Sensitivity study of the blocking type detection method. Boxplots show the distribution of the annual number of blocking events and fraction of blocked time steps (top) for 10 different settings of the minimum distance criterion for the whole domain (90°W to 90°E) (left) and for Euro-Atlantic region (40° W to 30°E) (right). The lower two panels show the associated distribution of mean and maximum duration.



(a) 90° W to 90° E

(b) 40° W to 30° E

Figure 9: Blocking probability estimated for individual months for blocking in general, as well as separately for *High-over-Low* and *Omega*. (a) whole domain ($90° W$ to $90° E$) and (b) Euro-Atlantic subsection ($40° W$ to $30° E$). Whiskers shows 95% confidence intervals assuming Gaussian asymptotics for estimating binomial probabilities.

In some other plots the authors provide a range of statistical uncertainty, but I could not find out how this was determined and what assumptions were underlying this estimate. I think that the authors need to explicitly describe (in the methods section) how statistical uncertainty was determined.

Thanks for the hint! We estimate occurrence and transition probabilities using binomial and multinomial logistic regression realised in the framework of (vector) generalized linear models fitted with iteratively reweighted least squares Yee [2015] as mentioned in our section 3. Confidence intervals are derived based on the assumption of asymptotic normality of the likelihood-based estimator, i.e. $[\hat{\theta} \pm 1.96\,\sigma_{\hat{\theta}}]$. We state this now accordingly at the end of section 3.5.

In the main section (section 4) the authors discuss the results irrespective of whether they are significant or not. In several places they seem to draw firm conclusions from results which are (as the authors say themselves) not statistically significant. It think that this is not good scientific practice. Rather, only those results that are statistically significant can be considered as "results" and should be discussed and used (for instance) to test hypotheses etc. Marginal statistical significance occurs in the present case when breaking down the results to individual months or seasons in section 4 (e.g., Fig. 16). In particular, in the summary section the authors should only refer to those results which are statistically significant.

Thank you for this comment. Especially regarding the discussion, we now mainly consider significant results and if not, we marked that the results are not significant. We only show non-significant results, if they were our motivation for further studies (e.g. looking for seasonal dependence on blocks instead of years). We moved Fig. 13 and Fig. 14 with non-significant results into the supplementary material.

Better motivation

It would be desirable if the authors can provide a better or more explicit motivation for their work and, especially, for the specific method that they chose to use. What is the advantage of their method in comparison with the many other methods that have been used in the past? What specific questions can one address that previous authors were not able to address? Why are those questions important? One possible avenue for improvement into that direction would be to formulate an interesting hypothesis and test this hypothesis with the analysis.

Lacking a more explicit motivation, the reader is somewhat left in the limbo as to what one is supposed to learn. One can always invent a new method an apply that method to reanalysis data in order to produce "results". But without further comment it would not be clear to what extent these "new" results are important or relevant. To be sure, I believe that the authors are able to provide such an improved motivation. In fact, some material in this direction is scattered throughout the text. I just urge the authors to collect this information and illuminate it in a more explicit fashion.

These are very valid points. We now give a clearer motivation. The first central and innovative point in our analysis is the possibility to distinguish between High-over-Low and Omega blocking. Knowledge about trends in these two blocking types can have immediate impact on the trends of the associated weather phenomena such as heat waves, cold spells or extreme precipitation events. Although we do not study these impacts, our method and analysis clears the way for further studies regarding the associated or underlying processes. Therefore, we included in the discussion the following paragraph: *As a novel aspect we introduced a blocking type decision method, that identifies* High-over-Low *and* Omega *patterns for each blocked time step, separately. Blocked weather situations are usually analysed with respect to the persistent high pressure area, which might lead to droughts with devastating consequences. Additionally identifying the low pressure system allows for further studies on the impact of the steady low pressure systems such as heavy rainfall and floods. In this way, we could tackle the first question in the introduction, if we find a method to automatically distinguish between the different atmospheric blocking types,* High-over-Low *and* Omega *block.*

Second, we formulate appropriate statistical models to describe trends and tendencies of occurrence and transition probabilities. We assume binomial (blocking/no-blocking) and multinomial (HoL, Ω, no-blocking) processes and use vector generalized linear models with iteratively reweighted least-squares to estimate the associated probabilities and

trends together with their uncertainty.

The Markov-Chain as model for onset, decay (offset) and transitions between states are the third innovative point giving insight into the dynamics of blocking states. As last point, we consequently break down the analysis from annual over seasonal to monthly probabilities to demonstrate shifts in the seasonal cycle although overall annual probabilities do hardly change.

Moreover, we motivate our work with three main questions posed in the introduction and to be addressed in the discussion section:

1. Can we find a method to automatically distinguish between the two atmospheric blocking types High-over-Low and Omega blocks?

2. Do blocking occurrence probabilities undergo long-term changes? Do these changes depend on season or month?

3. Do onset, decay or transition probabilities from one blocking type to another undergo long-term changes? Do these changes depend on season or month?

What are the true results?

This issue is related to the previous one. In the discussion section (which is partly just a short version of section 4) the authors mention a few caveats and sensitivities, but the reader is not told whether and to what extent these have an impact on the results. In other words, which of the results are true results and which are only consequences of the specific method that was applied? For instance, transition probabilities depend on the temporal resolution of the underlying data, so the specific value of the probability cannot possibly be a "true result". Similarly, the discussion provides the statement that different methods yield different numbers. My question would be: which of the results survive a change in the method? More broadly speaking: what is this paper's unique contribution to the topic? The devil's advocate would argue: "Well, you are using a novel method to investigate a problem that has been studied often times before; your results differ to some extent from previous results and many of your results are statistically insignificant". I am sure the authors have a good reply to such a provocative statement.

Again, I think that there is a unique contribution to the topic from this paper, I only say that this must be worked out more clearly.

Thank you for this comment. You are right that the discussion has mainly been a short version of Section 4. We removed the summary in the first paragraph and reordered and rewrote the discussion. Thereby, we mainly discuss significant results and mention explicitly, whether a result is significant or not. As mentioned before, the main advantage of this study is the ability to automatically distinguish between the two blocking types *High-over-Low* and *Omega* block.

We further address your concerns by carrying out a sensitivity study on the method detecting and distinguishing the blocking types, see Fig. 6 above.

3 Minor issues

- *Quite a number of minor issues are added as comments to the pdf of the manuscript.*
  Thanks, we added the comments from the pdf-supplement below and answered them point-by-point.

- *Sections 5 and 6 are partly redundant. For instance, large parts of section 5 repeat what has been said in section 4. What's more important: the summary section should not simply be a shortened version of the results section; rather, the reader expects a summary (plus discussion) on a higher level of abstraction.*
  Thank you, we agree! For the revised manuscript we have already restructured the discussion and we will revise the conclusion, highlighting the main results and their possible explanations.

- *In some parts of the text the quality of the English could and should be improved. Please also note the supplement to this comment: https://wcd.copernicus.org/preprints/wcd-2020-62/wcd-2020-62-RC1-supplement.pdf*
  We will carefully reread the text and try to improve the English. We will answer your supplement comments, too, in the following.

Minor comments from the supplement

l. 14: not very idiomatic English ...
We rewrote the sentence: "A Markov model determines the probability of transitions between different states by taking into account the actual time step only while neglecting all previous ones."

l. 33: By the time this paper will be published, the years 2018 and 2019 will not be the past two years any longer.
Thank you for the hint, we have adjusted it accordingly.

l.45 ... but that's a very special form of "discretization". It may be misunderstood by some readers, because what comes to one mind in first place is that discretization means that the PDE is discretized by standard numerical methods ...which is probability not what you mean here.
Thanks for pointing this out! The point vortex formulation is based on a Langrange'ian view on a system of vortices. A vortex is represented completely by its location and

its circulation. The vorticity at these points goes to infinity, while the vorticity is zero elsewhere. Hence the term "discretized". We will clarify the statement in the text.

l.66 "Transitions between ..." This sentence does not connect well to the previous two sentences.
We tried to create a better connection to the sentence before.

l.76 .. but if the typical time scale of blocking formation etc. is much longer than 6h, then the novelty of this work (e.g. the 6 hourly time resolution) seems rather irrelevant...
The typical time scale of a long-lasting blocking event is about 5 to 7 days. However, we observe in our work, that within this time span, the blocks change between the High-over-Low and Omega blocking type. In our opinion it is very relevant to study these blocking type transitions as well as their trends since the weather associated with blocking can have a high impact. For example high precipitation events can occur that are typically associated with the lows of the blocking pattern. The location of the low(s) relative to the high determines the blocking type of course.

l.78 The science questions should be worth a little more elaboration, they are important but not very clear at this point
We have reconsidered and reformulated the scientific questions and tried to better integrate them into the introduction. We will then analyse our three main questions in detail in the discussion.

l.94 You are probably talking here about zero horizontal divergence, right? In this case this is NOT synonymous with "incompressible" flow, because the latter only implies zero 3D divergence.
Yes, your right. The point vortex model is a two-dimensional vortex model. Large-scale synoptic flows are quasi-two-dimensional and 2D flows are – due to mass continuity – equivalent to zero horizontal divergence. We will clarify this in the text!

l. 108 Do you mean "longitudinally-dependent"?
Yes, that is correct. The value of the central reference blocking latitude (CRBL) depends on the longitude. Thanks!

l. 114 unclear how this "shift" works in practice.
The variable *Delta*, which indicates the shift, is described in Richling et al. [2015] as a positive value [in deg_lat] for a possible latitudinal northward and southward shifting of the central reference blocking latitude (CRBL). A more detailed explanation can be found in Richling et al. [2015] in Section 2.1, where in equations (3) - (5) to calculate Φ_N, Φ_M and Φ_S *Delta* is applied. With the help of these latitudes, the geopotential height gradients on the northern (GHGN) and on the southern (GHGS) side of the CRBL can then be determined (Eq. (1) and (2) in Richling et al. [2015]) .

l. 128 resembles a box shape ...?
In a regular latitude-longitude projection or in a Mercator projection of the field, a high-over-low blocking can be approximated by a square (or a box), where the poleward side of the square surrounds the high and the equatorward side the low. We will rephrase the sentence to make it a bit clearer.

l. 134 v is the two-dimensional horizontal wind here, right? It would be good to say this explicitly.
Done.

l. 139 What does it mean to set a grid point to zero?
We will try to clarify this in the text. We now wrote in the revised manuscript: "On the other hand, all grid points with values of $W_k \leq 1$ will be set to zero to obtain a field of vortex patches. This field of vortex patches can then be multiplied with some field of interest, e.g. the vertical vorticty field. In this field, we search for the high that lies closest to the longest-blocked IBL."

l. 140 how exactly do you define "longest-blocked IBL"?
The IBL detection method gives a time- and longitudinally-dependent field $IBL = IBL(t, lon)$ of "1" and "0", where "1" stands for blocked and "0" for unblocked longitudes. A simply-connected "field" of "1" in this Hovmöller-like representation of the IBLs can then be studied regarding the duration each single IBL is blocked in a row. We count these time steps for each IBL. For example, if the IBL is blocked for 5 time steps in a row, it is labeled with a 5, etc. The IBL with the highest number is called the "longest-blocked IBL". We will clarify this in the text.

l. 141 Again, not clear how a dipole is enclosed by a box shape, neither how you maximize the trapezoid.
Thanks, we will add more text to clarify the procedure. However, the method is more detailed described in Hirt et al. [2018]. The main difference between Hirt et al. [2018] and the method used here is, that the decision which blocking type we have is done at every time step. Please take also a look at our example movie in the supplementary material, which shows nicely the transition between an *Omega* and a *High-over-Low* block.

l. 145 How did you determine the box?
The box surrounds the high. In a regular latitude-longitude block or a Mercator projection of the atmospheric fields, a *High-over-Low* can be approximated by a square that encases the poleward high and the equatorward low. We will add more meat to the text describing the method.

l. 146 What are "the circulations" of a pattern?
We will describe this in more detail in the text. In point vortex theory, each vortex has three important properties: its coordinates in the plane and its circulation $\Gamma = \int \zeta_d A$

with ζ: vertical vorticity and A: the vortex area. The circulation of a vortex can be interpreted as the global strength of the vortex and is a conserved quantity for 2D, inviscid flow. Each point vortex induces a circular velocity field around its location whose strength depends on its circulation and falls of with r^{-1} the distance r from the center. The motion of a point vortex is given by the sum of the velocities generated by all the other point vortices in the plane. Nummerically, the circulation can be determined for each grid point by multiplying the area associated with the grid point with the vertical vorticity at this grid point. The total circulation of a "real" vortex is then determined by summing up all grd points associated with the vortex. For example, for the high in our case, we sum up all grid point circulations that have a negative vertical vorticity value in the earlier described vortex patches field and that lie within the shape (trapezoid or box). We will add a clarifying picture to Figure 3. Moreover, for the principles of point vortex dynamics and its applications to atmospheric flows as well as details of the trapezoid method we refer to our previous works, e.g. Müller et al. [2015] and Hirt et al. [2018].

l. 155 Not clear how you arrived this criterion.
We will either try to add a clarifying picture to Figure 3 or add more text here.

Subcaption Figure 3: "The positive vorticity is calculated...." What is "the positive vorticity"?
We mean positive vertical vorticity, i.e. $\zeta > 0$. We will clarify this in the caption of the figure.

Subcaption Figure 3: The reader probably does not care what software you used to produce these plots.
Probably not. However, the journal usually wants to know if the there are any copyrights on the maps. That's why we added the software statement. We will ask the editor, if we can delete this information.

Figure 6: What is the level of statistical confidence of the data in these two plots?
For the two regions shown, we have only looked at the absolute numbers of blocking, mean and maximum durations in order to get a first impression of the identified blocking events with the help of our method and to recognise differences in the regions. Uncertainties can certainly be estimated for these numbers, but many steps are necessary in the method to obtain the blocking state. We will do some tests to estimate the uncertainty of the method.

l. 284 Are these (weak) increases statistically significant?
No, they are not. We decided to rewrite the results concerning Figure 6 and rather test for the uncertainty of the method explicitly.

Figure 7: How did you determine these confidence intervals?
We have added a short description in the method (Section 3.5) that explains in more

detail how we determined the confidence intervals for Fig. 7 and all subsequent figures.

l. 318 But not really statistically significant....
Yes, you are right. The small increase in the summer month (Fig. 8c) is not statistically significant, but the increase in Fig. 7b is significant for the summer.

l. 377 "offset" is very peculiar terminology in the present context, I would prefer "decay".
We can understand this argumentation, but will nevertheless stick with the term "offset". From a meteorological point of view, the use of the word "decay" describes the transformation from a blocked state to an unblocked state with the right words, as it is a process. In our work, we focus on the model view, in which there is only the state "on" or "off". Therefore we use the term "offset". However, when introducing the terms "onset" and "offset", we will add a sentence that addresses the underlying processes of the formation and the decay of a blocking.

l. 391 It seems that none of the trends in Figure 13 are statistically significant...
Yes, you are right. We have shown Fig.13 and Fig. 14 in our first draft to give an overview of all analysis and especially to visualize the results for the two transition matrices of Eq. 11 and Eq. 2. Now we decided to shift these two figures to the supplementary material and only show results for the 3x3 transition matrix.

l. 407 It seems somewhat problematic to discuss results that lack of statistical significance.
Thanks for pointing this out. We will try to only discuss statistically significant results. However, the matrix plots (Fig. 13 to 16 in the first manuscript) represent an entity. So we do not want to delete single subfigures. However, we decided to transfer the statistically insignificant figures (Fig. 13 and 14) to the supplementary material for the sake of completeness of the analysis.

l. 414 Again, if the result is not significant statistically, you should probably refrain from discussing it.
We will check everything in the results, discussion and conclusion sections and only focus on the statistically significant results.

Figure 14 Where is the color bar explaining the different colors?
Yes, we had forgotten the color bar in Fig. 14 and also in Fig. 16, which we have now added.

l. 424 investigations...are evaluated. [strange language]
The discussion section was completely rewritten. Hence, this sentence no longer exists in the revised script.

l. 430 Fine, you show the results, but some of these results are not statistically significant. I wonder what we learn from that.
We moved Fig. 13 and 14 to the supplementary material and will point out more clearly which the important and significant results are and what we can learn from them.

l. 432 Did you verify whether this increase is statistically significant?
No, we did not. It seems to be statistically insignificant and hence, we will remove this statement from the text. Thanks for pointing this out!

l. 443 Can you resolve this apparent contradiction between their results and your results?
One point is definitely that our methods are different. In our case, we search for a coherent blocking structure, that needs to have a minimum lifetime of 5 days (in the larger region) and should be composed of the same high (lows are allowed to change in time). Our initial IBL identification is moreover only a one-dimensional method. In Brunner et al. [2017] the identification of blocks is done with a two-dimensional blocking index. Blocks are long-lived (at least 5 days) and synoptically large ($\pm 7.5\circ$ longitudes). However, a blocking is counted whenever at least one of these blocked grid points is within the Euro-Atlantic region ($45°$–$72.5°N, 30°W$–$45°E$). In our case, at least half of the block, more precisely the circulation centroid, needs to be inside the Euro-Atlantic region (40W-30E). Summarized, there is no easy answer. We will do more tests and try to answer your question! Thanks a lot! This is also relevant for our discussion section.

l. 451 So to what extent do these caveats have an impact on your results?
We will try to do additional tests in the revised manuscript to estimate the impact!

l. 480 This number in itself does not make any interesting statement itself (because it depends on the temporal resolution of the underlying data)
Thanks, we added a time interval (6 hours) here.

l. 518 Where is the verb in this sentence?
We deleted this sentence.

l. 521 proportion of what?
Thanks. We rephrased the sentence: "While in July *Omega* blocks account for only about 25% of all observed blocks in 1990, we find on the one hand an increase in the number of blocks in general as well as a higher fraction of Omega blocks towards the end of the study period."

References

Lukas Brunner, Gabriele C Hegerl, and Andrea K Steiner. Connecting atmospheric blocking to european temperature extremes in spring. *Journal of Climate*, 30(2): 585–594, 2017.

Mirjam Hirt, Lisa Schielicke, Annette Müller, and Peter Névir. Statistics and dynamics of blockings with a point vortex model. *Tellus A: Dynamic Meteorology and Oceanography*, 70(1):1–20, 2018.

Annette Müller, Peter Névir, Lisa Schielicke, Mirjam Hirt, Joscha Pueltz, and Isabell Sonntag. Applications of point vortex equilibria: blocking events and the stability of the polar vortex. *Tellus A*, 67, dec 2015. ISSN 1600-0870. doi: 10.3402/tellusa.v67.29184.

Andy Richling, Chistopher Kadow, Sebastian Illing, and Oliver Kunst. Freie Universität Berlin evaluation system (Freva) - blocking. `https://freva.met.fu-berlin.de/about/blocking/`, 2015. (Documentation of the blocking plugin), Accessed: 2020-09-10.

Thomas W. Yee. *Vector generalized linear and additive models: With an implementation in R*. Springer, New York, USA, 2015.