Comments from referee #3

This ambitious paper sets out to assess the exposure of the Swedish power network to an extreme geomagnetic disturbance (GMD). The paper describes an approach to making the assessment and discusses the outcomes.

The analysis is based on a single storm model, a 50-min substorm of the Halloween storm of 2003. Variability within the class of extreme events has been modelled by two cases generated from this one event. The first case is generated by scaling upwards the magnitudes of the disturbance. The second case is generated by extending the most intense sub-region over the whole geographic region of study. Neglecting the practical limitation of the physical disturbance that these approaches represent, the results of the simulations are presented and analysed in detail that might not be evident in a real power system response.

We thank the reviewer for a critical review and helpful comments. We have done our best to reply to each concern. The original review is written in black and our replies in red. The line numbers refer to the original version of the manuscript.

In my opinion, the paper falls short of its objective because of the lack of rigour in matching the assumptions to reality. Maybe, it is impossible to do so. Nevertheless, the paper exposes an important aspect underlying the approach – how to model the physical form and characteristics of a 1-in-100-year extreme GMD. Although this issue is not addressed directly, the paper has value in its clear description of the approach to the study of a large region and the patterns and discussion of some simulation results.

Specific comments: Individual scientific questions/issues

Extreme space weather events are characterised by the Dst index (-400 to beyond -1000 nT) or the maximum amplitude of EH at some time resolution, and the authors recognise that whether an event occurs during the day or night makes a difference. In this paper, a 1-in-100-years extreme event has been characterised using sparse data (one event), apparently without reference to the practical effects of the disturbances on power systems.

We thank the reviewer for this comment. We emphasise that a “worst-case scenario” in our study refers only to the 1-in-100-year geoelectric-field amplitude, which is the physical driver of GICs, and not to a worst-case scenario for the power grid itself. We would like to point out that although we use one event to characterize the spatial distribution of the perturbations for Case 1, the thresholds that we used are based on decades of data.

We recognise a vast number of extreme geoelectric field scenarios exist and we only use one event. However, the goal of the study is to understand the implications of the impact on the Swedish power grid if a previously recorded event would have reached a 1-in-100-year geoelectric field level. In our approach we made the logical choice to use the event with the strongest geoelectric field (Halloween event).

Using a simplified power-grid representation, we convert this extreme E_H value into line-average geoelectric fields (voltage per km) to map physical exposure, not operational impact. Although the 1-in-100-year geoelectric field is not tied to power systems disturbances, such large geoelectric fields typically occur during strong geomagnetic activity, which also increases the likelihood of power-grid disturbances. This is consistent with Rosenqvist et al. (2025), where most recorded grid disturbances occur during elevated Kp conditions.

We have included additional text at the beginning of the Method section, in line 105.

[...] Our motivation for scaling the geoelectric field was that it is the physical driver of GICs. Using real ground magnetic perturbations as input preserves a realistic ionospheric forcing (rather than relying on a synthetic magnetic field time series), while the scaling ensures that the resulting geoelectric field reaches the level associated with a 1-in-100-year event from the perspective of the physical driver of GICs. This allows us to generate as many worst-case scenarios as magnetic field inputs are available. However, because ground magnetic perturbations are driven by ionospheric current systems that activate during geomagnetic storms and substorms, we recommend using magnetic field inputs from previously recorded events that produced the largest geoelectric field response.

It has been postulated by others that 'failure' due to GMDs occurs by a process of transformer insulation degradation (such as at Salem in 1989, though even that damage did not lead to an immediate fault), by power system collapse (such as in Quebec in 1989, though that collapse was initiated by protection relay operation removing SVCs from service), or by protection relay maloperation not traced to equipment damage (such as at Malmo in 2003 and Bandsjö in 2017 and 2023). The effect of GMDs on power systems is not a simple relationship between peak measurements of geomagnetic or geoelectrical parameters and power system degradation. The relationship to simulated parameters of artificial GMDs is probably even weaker.

We thank the reviewer for this very interesting point. We agree with the reviewer that the mechanisms leading to power-system failures during geomagnetic disturbances are complex and that the effect of the physical driver on power systems is not a simple relationship between the peak measurements of geoelectric fields and power system degradation.

We included the following text in line 330:

[...] Understanding which combinations of geoelectric field, induced voltage, network configuration, and system state lead to real incidents remains an open problem and requires further investigation. Such studies will soon be possible with the recently installed GIC monitoring device in Karlshamn, connected to the transformer neutral-to-earth point and deployed by the Swedish transmission system operator, Svenska Kraftnät.

Therefore, my first question is whether the authors can justify their decision to base an 'extreme event' on a 10 s EH threshold (line 118) and apparently arbitrary line voltage thresholds of 100 or 200 V (line 255) using the base parameters of a single event? Even with the power utilities' addition of line resistances to the calculated line voltages, is the resultant peak GIC the most useful and reliable metric of system stress likely to initiate outages?

We thank the reviewer for raising this question. Our worst-case scenario event (Case 1) represents an extreme geoelectric-field event, specifically illustrating how the geoelectric field during the Halloween storm in Sweden would have appeared if it had reached a 1-in-100-year level. This is only one possible scenario, reconstructed from the past event that has resulted in the largest geoelectric field.

The 10-s EH threshold represents the 1-in-100-year extreme geoelectric-field amplitude, which is the physical driver of GICs, not a proxy for power-system failure. We do not use line voltage as a proxy for GIC either; we simply integrate the extreme electric field along each line to show how such an external driver would map into line-average geoelectric fields. This represents an extreme event from the external forcing, not an extreme event for the power grid itself.

The authors generate two 'plausible' cases of GIC events using different assumptions and scaling the parameters of the Halloween storm.

We are not generating two plausible cases of GIC events, we are generating two plausible extreme geoelectric field events.

An important issue is summed up at line 108: "The EH reached during this event has been classified as having an estimated recurrence of 1-in-100-year event in northern Sweden, but only a 1-in-10- to 1-in-50-year event in southern Sweden". This leads to a question that is not identified in the paper: Would the ionospheric structures retain their physical shape with scaling to higher intensities: case 1: "the actual observed magnetic field perturbation extrapolated to the whole Fennoscandian region", or case 2: "the magnetic field time series is spatially uniform across all magnetometer stations"?

We thank the referee for this comment. The purpose of our two cases is not to reproduce an extreme ionospheric current system or to simulate how ionospheric structures would scale under higher intensities. Instead, the cases illustrate two different ways in which an extreme geoelectric-field amplitude (the 1-in-100-year E_H value) could map onto Sweden, depending on the assumed spatial pattern.

To make it clearer we have added a few sentences at the beginning of the Method section, and after the response to your previous comment in line 105:

[...] To illustrate the range of possible worst-case conditions, we explore two complementary approaches. The first worst-case scenario (Case 1) is intended to remain as realistic as possible, with each geographic location experiencing a different magnetic field perturbation. This is achieved by using as many magnetometers as are available in Sweden, providing a well-resolved ionospheric driver. The second worst-case scenario (Case 2) is intended to represent an idealised case in which every location in Sweden experiences the same magnetic field perturbation (i.e., the frequency content remains identical). This is achieved by using data from a single magnetometer. Such an idealised situation could be produced by a large-scale westward electrojet (WEJ) extending across Fennoscandia.

The results of the simulations are described in detail. The basis of the scaling is consistent with approaches by others to similar modelling of the Swedish E-field profiles, so it is not surprising that the results are generally consistent with other published results. The observations are detailed but not especially novel. For example, at line 195, reference is made to ‘the transformer trip event’, a detail that has not yet been introduced and, relevant to line 196, several papers have shown the relationships between frequency components of the B-field, E-field, and GICs. The difference between the results of the two cases, as depicted in Figure 5(c) surely depends on the way the cases were constructed. To what extent do the simulation results represent plausible GMD events of 1-in-100-years severity?

Thank you for this question. Our answer is that we do not know, and determining how 1-in-100-year geoelectric-field values relate to plausible 1-in-100 year GMD is outside the scope of this study. Analyses of 1-in-100 year GMD and 1-in-100 year GIC scenarios are planned for future work, where comparison between the different worst-case scenarios could be carried out.

The discussion of the modelled results, and the explanations in section 4, appear to support and validate the simulations. However, the focus on voltages and voltage increase with line length (line 253) fails to mention that the line resistance also increases with length, such that longer lines might not contribute significantly higher induced currents. Further, the comments about Malmö (line 237) and Karlshamn (line 241) suggest high susceptibility to GICs even in the south.

Thank you for pointing this out. We have updated the text in line 253. not it reads:

[...] In general, the voltage in the line increases with length as shown in Figure 8c; however, because line resistance also increases with length, longer lines might not contribute significantly higher induced currents.

In relation to the high susceptibility in Malmö and Kalshamn, we have included the following text in Line 242:

[...] As with Malmö, this behaviour may be influenced by sparse magnetometer coverage in southern Sweden and by limitations in the conductivity model. In addition, the most intense GICs are often observed at substations located near the edges of a power-grid network (Viljanen et al., 2012). Both Malmö and Karlshamn lie along the southern boundary of the Swedish transmission system, which may enhance their vulnerability. These examples highlights how local grid topology and substation characteristics may influence vulnerability independently of regional E_H strength.

Figure 9 is interesting. A related paper [Rosenqvist et al., 2025] provides details of three GIC incidents during the Halloween storm period used in these simulations. They occurred at 19:55, 20:03, and 20:07. Two of the three did not coincide with the peaks of the number of lines exceeding the various voltage thresholds. From these records, it appears that the GIC effect that caused incidents was not necessarily a peak line voltage. This information raises questions about the validity of the presumed relationship between voltage stress and power system response.

We thank the reviewer for this important observation. We agree that line-voltage peaks do not necessarily coincide with GIC peaks or with the timing of power-grid disturbances. Our results are consistent with this: extreme geoelectric-field values tend to occur during very active geomagnetic conditions, but the resulting GIC impacts do not align in a simple or direct way with the maximum induced voltages. This has also been observed in other recent events, such as the May 2024 storm (see Figure 5 in Wallner et al., 2025; Figure 9 in

Dimmock et al., 2024), where disturbances did not occur at the time of the largest E_H or line-voltage values.

Understanding the detailed relationship between voltage stress, GICs, and power-system response is a key point, but outside the scope of this paper. At present, we do not know the exact relationship between line-voltage exceedances and the timing of power-system incidents. A more complete assessment would require modelling of the full power-system response, including line resistances, grounding, and network topology. We are actively exploring how to better connect geoelectric field driven voltages with observed disturbances.

In the revised manuscript, Lines 263-265 have been removed and we included the following text in Line 318:

[...] We have shown that during intervals of very active ionospheric current variations, particularly during substorm activity, large dB/dt and E_H values were estimated at latitudes below 60° MLAT. As a consequence, several power lines exceeded high voltage thresholds (e.g., approximately 100 power lines experienced induced voltages exceeding 50 V at several times during the substorm period). For comparison, Lucas et al., (2018) reported that 62 transmission lines exceeded 100 V during the March 1989 geomagnetic storm, which triggered a major blackout in the Hydro-Québec power system. Although power grid disturbances do not always coincide with the time of peak voltage or geoelectric field (Dimmock et al, 2024, Wallner et. al., 2025), or with peak GICs (Pulkkinen et al., 2005), intervals of high induced voltages still imply larger voltage variations.

Have the authors adequately considered the possibility that the differences between the two constructed simulations may be less significant than the types and settings of the relays that initiated tripping in the several GIC-related incidents identified in Rosenqvist's paper? Perhaps the detailed comparison of the two case studies is misleading and could be reduced, with advantage to the overall clarity of the paper. What truly significant conclusions can be drawn from the simulations?

We thank the reviewer for these thoughtful questions. We hope that the justification for including the two worst-case scenarios is now clearer based on our previous responses.

From the two constructed scenarios, we have obtained an estimation of the expected geoelectric field values across Sweden in a worst-case scenario. We have identified where the strongest geoelectric fields occur, and validated these magnitudes against previously recorded or estimated extreme geoelectric field values. This allowed us to determine which power lines are most likely to experience the largest induced voltages per kilometre. From the perspective of the Swedish Power Grid operators, the identification of vulnerable regions due to the external driver (geoelectric field) is very valuable.These results provide essential boundary conditions for future work that will explicitly assess worst-case impacts from a power-grid perspective.

The discussion of power grid implications (section 5.2, starting at line 308) could have been used to define the extreme event parameters before the modelling started. It appears that the argument follows the structure of:

> Extreme space weather events are linked with transformer damage. No references are given, but the examples are generally the Salem, USA transformers in March

and November 1989; several Eskom generator transformers at different power stations following the GMDs of 2001-2003, and one or more transformers in New Zealand.

In the introduction we had included examples of extreme space weather events linked to problems for power systems. The included references were Bolduc 2002 (13 March 1989 event) and Pulkkinen et al., 2005 (29-31 October 2003 event). In the revised version, we have also included the reference Girgis et al., 2012 where the authors present the GIC current at Salem Generating station in 1989.

R. Girgis and K. Vedante, "Effects of GIC on power transformers and power systems," PES T&D 2012, Orlando, FL, USA, 2012, pp. 1-8, doi: 10.1109/TDC.2012.6281595.

The records indicate that only one New Zealand transformer failed during a severe or extreme GMD, all the remainder failed or were removed from service several days, weeks or months after the GMD although techniques like oil analysis can show that the degradation was linked to GMD events. Therefore, simultaneous transformer tripping due to physical damage appears unlikely during a 1-in-100-year event. Transformer tripping by relays responding to waveform distortion or other parameters could cause multiple unit tripping – this is a relay/protection problem. Apart from the ambiguity in the reported new specification for Swedish transformers to withstand GIC, it is not evident that the thermal stress is directly linked to peak exposure to line voltages.
High geoelectric voltages on lines are given significance in this paper. However, it was not a line voltage that initiated collapse (blackout) of the Hydro Quebec system in 1989; it was a relay response to harmonics. It is unclear whether any voltage collapse blackouts have been initiated purely by geoelectric voltages or GICs.

We agree that transformer tripping due to physical damage is unlikely during a 1-in-100-year event. Our analysis of transformer GIC-withstand requirements led us to the same conclusion. However, previous studies have shown that power grid disturbances can coincide with intervals of peak geoelectric field (see Fig. 4d in Rosenqvist et al., 2025), and therefore we do not dismiss the relevance of analysing peak induced voltages.

The authors suggest high geoelectric voltages could “pose a significant risk to protection relays” (line 326). However, the mechanism of risk is not identified. Considering the timing of actual relay events (see above) the problem is possibly that inappropriate relays (such as one measuring a non-physical quantity or lacking immunity to harmonics), or inappropriate relay settings, or, possibly, faulty relays or relay systems might cause ‘unnecessary’ tripping. Collapse might be initiated by one or multiple coincident relay maloperations, each of which might remove a transformer or line circuit for several hours while its integrity is checked. Based on Figure 9, such relay events might not occur as a direct result of high or maximum voltage but at high or changing rates of change of voltage.

We thank the reviewer for this very important point. We modified the text in line 326, now reads:

[...] The large number of power lines subjected to high induced voltages in our scenario could lead to elevated GICs, which in turn may affect protection relays. Relay misoperations are often linked to harmonics, inappropriate relay types or settings, or faulty relay systems, rather than to the voltage magnitude itself. Collapse can be initiated by one or several coincident relay misoperations, each removing a transformer or line from service while its integrity is verified. Based on this, relay actions may occur during intervals of high or rapidly changing rates of change voltage rather than strictly at maximum voltage.

Understanding which combinations of geoelectric field, induced voltage, network configuration, and system state lead to real incidents remains an open problem and requires further investigation. Such studies will soon be possible with the recently installed GIC monitoring device in Karlshamn, connected to the transformer neutral-to-earth point and deployed by the Swedish transmission system operator, Svenska Kraftnät.

What is it, then, that defines an extreme event from the perspective of the electricity system owner and operator? Does it cause threshold maximum conditions throughout the region, could there be 'hot spots' of intense stress, and how big is a region? For electricity utilities, an extreme event might focus on geomagnetic and geoelectric conditions that stress different items of the utilities' equipment, rather than on peak (10 s) values. (Other infrastructure, such as pipelines, might see extreme events differently.) Would simulations directed towards identifying such conditions be more consequential?

Thank you for these insightful questions. They raise exactly the kinds of considerations that are essential for advancing the perspective of the electricity system owner and operator. Defining what constitutes an "extreme event" for the power-grid operator has been included as an open question in Line 349.

Though this paper does not answer these basic questions, it makes an important contribution to the field simply by its clear description of the approach that exposes these other aspects.

We thank the reviewer for raising these fundamental questions, which are indeed highly relevant for future studies that aim to adopt the perspective of electricity-system owners and operators.

Addressing an extreme event from that operational point of view lies beyond the scope of the present study. Our aim here is to generate worst-case scenarios from the perspective of the external physical driver, independent of the internal power-system response.

We appreciate that the approach presented here helps expose these broader issues, and we see this study as a necessary step toward future investigations that integrate both the physical driver and the power-system response.

Technical corrections: spelling, grammar, etc

In line 204, correct 'factos' to factors.

Done

At line 243, the word 'drive' suggests the vulnerability is directly caused by the topology or substation independently from the geoelectric voltage. I suggest instead that '… grid topology and substation characteristics may influence vulnerability as much as does the regional EH strength' – or similar.

Done

At line 252, would 'dominant' be more suitable than 'preferred'?

Done

At line 267, it appears MLT should be MLAT.

We refer to MLT sector (nightside) as the region where the largest magnetic activity typically occurs at high latitudes. We changed the sentence to:

This study presents the first assessment of a worst-case scenario for the entire Swedish power grid using real data during the development of a premidnight substorm event.

At line 311, the units of line resistance are probably ohm/km values, not simply ohms.

Done

The details of many references are incorrect. Many take the form of https://doi.org/https://doi.org/10.1029/2022SW003304 (line 395). Another does not return a page.

All references have been reviewed