Review of "A statistical examination of the effects of stratospheric sulphate geoengineering on tropical storm genesis" by Wang *et al*

Tropical cyclones (TCs) are high-impact climate phenomena which are poorly represented by the current crop of low resolution climate models. Instead, many studies have utilised implicit statistical relationships between atmospheric conditions and TC activity to investigate changes to TC features under global warming. In this study, Wang *et al* employ the oft-used Genesis Potential Index (GPI) and Ventilation Index (VI) to investigate TC changes under stratospheric aerosol injection (SAI). They find that GPI is lower under SAI than standalone global warming while VI increases, suggesting fewer storms under SAI. They then use a variety of statistical techniques to explore which climatic features contribute to changes in GPI and VI, finding that changes to relative humidity and potential intensity (and furthermore surface temperature) describe most of the variance in the GPI.

I applaud the authors for taking on this hefty subject, and I think they do a commendable job at including many useful analyses. However, I found the manuscript to be poorly written and with a confusing and non-linear structure that reduces readability. In particular, Section 3.5 (TCs from TRACK with HadGEM2-ES) adds nothing to the manuscript and should be removed – especially as it uses only one model (when the manuscript highlights significant inter-model disparities) and a completely different methodology to the rest of the study. I urge the authors to instead focus on the GPI and VI results in Sections 3.1-3.3 which are sufficiently interesting. I'd also suggest adding more justification for the inclusion of Section 3.4, which seems incongruent to the rest of the results section, despite being an interesting study in its own right. Finally, I urge the authors to restructure the Introduction and Conclusion so that key points are not repeated.

I also take issue with fundamental inferences in the manuscript that do not seem to be supported by the tables or figures. Namely, that the GPI (VI) are significantly lower (higher) in G4 than RCP4.5 (L217). In fact, the GPI and VI seem to be very similar between the G4 and RCP4.5 simulations on a model-per-model basis (Fig. 1), and this is also reflected in columns 5 and 10 of Table 3 where only one value is in bold font, suggesting that almost all of the hemispheric differences are not significant. Perhaps add the global-mean GPI and VI differences to Table 3 to confirm that global changes are significant? As this inference (i.e. GPI reduced under geoengineering) is a key result of the paper, I feel this issue must be rectified before I can recommend publication. I have added *many* further comments below.

Major Comments

1. I found the abstract to be long and confusing, it should grab the reader with the most significant results. There's lots of irrelevant detail such as "vertical wind shear and vorticity is insignificant" which does not need to be here. The last sentence (Line 52) adds nothing to the abstract and does not follow from the results of the study

2. Restructure the introduction (Lines 83-121). It currently bounces between the various ways of measuring TC activity, when it only needs to say that there are implicit (e.g. GPI), semi-explicit (e.g. dynamical downscaling) and explicit (e.g. feature tracking) methods for measuring storm activity (with references). Then go on to describe GPI and VI theory.
3. I think you need more justification in the text as to why Section 3.4 belongs in the manuscript – it seems to me to be a rather unnecessary accessory that detracts from your GPI and VI results. Also, are there any references that have identified a clear physical relationship between ENSO and storms outside the WNP and NA basins?
4. Remove Section 3.5 which is confusing and does not anything to the study
5. Please establish that the GPI and VI differences between G4 and RCP4.5 are significant as Table 3 and Figure 1 do not currently support this argument
6. There are *many* grammatical and spelling issues that need addressing. Please have the study checked for grammar.

General Comments

1. Use acronyms throughout the manuscript – please stop jumping between using acronyms and full terminology (e.g. potential intensity and PI are used interchangeably) which is confusing. Define acronyms on first usage (e.g. Greenhouse Gas GHG) and revert to using them exclusively
2. Use either Stratospheric Sulphate Geoengineering (SSG) or Stratospheric Aerosol Injection (SAI) throughout instead of generic terms such as 'geoengineering' or 'SRM' which comprise a variety of other methods that may have completely different impacts on storms
3. I'd suggest either using Tropical Storms (TS) or Tropical Cyclones (TC) as the terminology throughout. I see you use typhoon or hurricane at some points, which are basin-exclusive terms and I think not be used
4. Occasionally you refer to the use of climate indices such as GPI or VI as the 'direct' way of measuring TC activity (e.g. L85) or you say storm tracking is 'indirect' (L489). I disagree completely! Rather, tracking storms is more direct or explicit, whereas indices are implicit or as you say empirical. Please change this throughout the manuscript.
5. You often give p-values – note within the text which tests were used to derive these p-values (I assume 2-sided t-test but this should be specified)
6. Please check references throughout. My particular gripes are that all papers with 2 authors should be labelled as such, for instance 'Tang and Emanuel (2012)' not 'Tang et al. (2012)' (see L100). I found *many* such instances. Some references are misused and do not contain pertinent detail to the text (see specific comments for details), while the Thomas et al. (2015) reference (Line 120) is missing.
7. If you do keep Section 3.5 in the manuscript, please acknowledge the Hadley Centre for providing you with the 6 hourly data. Please also acknowledge Kevin Hodges if he assisted you in running TRACK.

Specific Comments

1. L33 – ‘a complete description of TC variability requires much more dynamical data than models can provide at present’ – I don’t think this is true, I think the issue is the coarse spatiotemporal resolutions, the models have sufficient dynamics. Please rephrase.
2. L35 – do you need to list all the individual components of the GPI and VI in the abstract? Surely this is more for the Introduction or Methods section
3. L41 – ‘Globally, GPI under G4 is lower than under RCP4.5, though both have a slight decreasing trend’. I am concerned that people might read from this that SAI is not able to counteract GPI changes under global-warming. Rather, the slight decreasing trend in G4 simply relates to the experimental design (i.e. a constant forcing). I would remove the ‘slight decreasing trend’ line and add a caveat in the conclusions saying if a different SAI approach were taken (e.g. Jones et al (2018) stabilizing global warming at 1.5K) then the GPI trends may be different
4. L42 – ‘spatial patterns in the effectiveness of geoengineering show reductions in TC’ – I’m not sure what this means, please clarify
5. L47 – ‘genesis potential’ -> ‘GPI’
6. L52 – final line – again I’m not sure what you mean by this final sentence. Do you mean that simple statistical models based on surface temperature or relative humidity changes are appropriate for examining TC changes? I don’t think you really show this though as you don’t explicitly link GPI or VI to modelled TCs in this study
7. L58-L62 – numerous grammatical issues
8. L63 – replace ‘retard’ with suitable word such as ‘counteract’
9. L67 – consider replacing ‘facilitate’ with ‘homogenize’
10. L68 – ‘and is supported by about 12 model groups’ replace with ‘and is currently supported by 12 model groups’ – be specific on the number
11. L69 – ‘Climate system thermodynamics will certainly change under SRM’ – this is a strong statement, change compared to what by the way? If you mean compared to business-as-usual then change (of a kind) may welcome! Please reword
12. L82 – ‘methods that rely on the statistical links between the thermodynamics of the ocean and atmosphere with cyclone dynamics have been the topic of studies’. This is not entirely true, Jones et al (2017) do explicitly model storms. Add predominantly between have and been
13. L83 – as mentioned, replace typhoons with TCs
14. L83 – as mentioned in the major comments, this paragraph is very confusing, please revise.
15. L96 – ‘factors influence genesis’ -> ‘factors influence TC genesis’ or cyclogenesis
16. L96 – ‘a quantitative theory is lacking’ – please add a suitable reference
17. L99 – What is the definition of potential intensity, which is rather an abstract concept? Define on first use
18. L108 – Dynamical potential intensity is more about ocean feedbacks (i.e. storms stir up cold water, which in turn reduces the potential intensity) than general ocean impacts
19. L109 – ‘These indices represent the climatological thermodynamic spatial and seasonal control -> please simplify, superfluous language

20. L111 – ‘more or less beyond the abilities of contemporary climate models’. What about the high-resolution models though, which are able to model storm intensities capably (see Murakami *et al* (2016) and Roberts *et al* (2015))
21. L112 – Wang et al (2012) only consider one basin – please replace with a suitable reference comparing different basins
22. L119 – What do you mean by severe TCs? Perhaps give windspeed constraints
23. L120 – Thomas et al (2015) reference missing from bibliography
24. L120 - Kang et al (2012) reference makes no predictions about future changes in TC activity, please change to a relevant reference
25. L126 – The sentence describing Jones et al (2017)’s results is confusing. All you need to say is the SAI in the north reduces North Atlantic TC frequency, while SAI in the south enhances NA TC frequency. Their results were inconclusive for the G4 scenario, as investigated here
26. L131 – Please sell the merits of your study. No other study has looked at GPI and VI in the context of SAI. No other study has investigated storm changes under SAI in basins outside the North Atlantic! No other study has attempted to attribute changes to storms under SAI to thermodynamic changes. This is good work, an important scientific development, and should be highlighted.
27. L138 – ‘We quantify the contribution of each variable to TC genesis using two statistical methods’. This is a rather weak statement, which variables do you study and which statistical methods do you utilize?
28. L139 – ‘Finally we study the effect of ENSO on TC and TC track of HadGEM2-ES’, what justification have you for including these studies here, they don’t fit with the GPI and VI work that you have set up to assess
29. L150-L156 – You present the results from Yu et al (2015) and Moore et al (2015), but these will necessarily differ to your results as you use 6 models (you include NorESM-1) and they use 7 models (CSIRO-MK3L and GISS-E2-R). Please state the temperature changes in *your* ensemble of models, and preferably include ranges. Consider also moving this to the results section of the manuscript
30. L162 – Sentence beginning ‘It is, however,’ should have a suitable reference
31. L170 – replace ‘vector’ with ‘wind’
32. L172 – define Cp in Equation (2)
33. L174 – why do you use the 100 hPa level for the outflow temperature? Do you have a suitable reference?
34. L177 – please define what the GPI is, i.e. the theoretical maximum intensity, and what increases/decreases to GPI signify (with a suitable reference)
35. L188 - please define what the VI is and what increases/decreases to VI signify (with a suitable reference)
36. L190 – ‘greenhouse gas’ - > RCP4.5
37. L193 – ‘air temperature’ on levels or near surface air temperature?
38. L200 – ‘researchers’ -> ‘studies’
39. L200 – Emanuel 2010 is not a suitable reference as it makes no predictions for the future, only studying observations from 1908-1958
40. L200 – ‘find’ -> ‘predict’

41. L201 – ‘in the Southern Hemisphere’ … under global warming
42. L201 – ‘but increasing frequency in the northern hemisphere’ – Knutson et al (2015) find no such thing! Sure, they find an increase in the East North Pacific and North Indian basins, but they also predict a decrease in the North Atlantic and West North Pacific basins!
43. L203 – ‘The observed TC annual-mean numbers for the period 1980-2008 for each basin are also listed in Table 2’ – where did these numbers come from? I can only find the basin boundaries in Emanuel 2010. Are these numbers consistent with your basin boundaries? Please provide a suitable reference
44. L209 – ‘annual’ -> ‘annual-mean’
45. L210 – you use August to October for the Northern Hemisphere, but the North Indian basin has two peaks in activity (one in May) (Li et al., 2013). How does GPI and VI change in this second peak in the Indian basin. Please comment on this.
46. L210 – what percentage of total annual storms in each basin occur during your chosen timeframes? These 3 month timeframes seem very narrow to me
47. L217 – ‘Furthermore, the G4 means for all models were significantly lower than their RCP4.5 values’ – Table 3 seems to say the opposite, that none of the changes to GPI are significant! Please give annual-mean values, standard deviations and p-values for G4 and RCP4.5, perhaps in Table 1?
48. L221 – ‘The time series indicate that tropical storms will become more frequent with time and that G4 significantly reduces numbers’ – *Fig. 1 does not indicate this at all to me!* There are many years, for each model, where the GPI is higher for G4 than for RCP4.5. Figure 1 will need rethinking as it does not support the central tenet of your paper. How for instance, can a difference of -0.3 % in CanESM2 be significant?
49. L226 – Sentence starting ‘During most years from 2020 to 2069…’ – this is hardly a sufficient statistical test for significance, simply saying VI looks higher for G4 than RCP4.5! Please perform significance tests and identify which models show significant VI changes and which ones don’t. Table 3 suggests that no VI changes are significant!
50. L231 – ‘As with GPI, there is about a factor of 2-3 range in absolute values between the models’ – perhaps plot normalised anomalies relative to 2020-2030 in Figure 1 instead?
51. L245 – ‘All models except NorESM1-M show negative differences in the North Indian basin’ – this may be true on a basin-wide basis, but BNU-ESM shows GPI increases in the Bay of Bengal. This might indicate a change in the spatial distribution of storms in the North Indian basin. This is also apparent in the ensemble-mean
52. L269 – consider changing ‘item’ to ‘component’ or ‘term’ throughout this paragraph
53. L274 – has this decomposition (Eq. 5) been used before? If so, provide a relevant reference
54. L301 – ‘Hence, these are the factors that *primarily* enable *solar* geoengineering’
55. L304 – do you have any idea as to why MIROC-ESM-CHEM is so different?
56. L312 – Sentence starting ‘Fig. 4 shows that the HadGEM2 values tend to be smaller’ – CanESM2 is similarly muted and seems to have the same signs as HadGEM2
57. L356 – ‘The key factors affecting TCs’ – consider adding a more informative title, possibly, ‘Primary factors that control GPI and VI changes’?

58. L357 – You seem to have found in Section 3.2 that relative humidity is the most important factor for GPI, and then ignore this finding in Section 3.3, which I found curious
59. L368 – 'The model ensemble' -> 'The ensemble-mean'
60. L370 – 'Fig S3 shows that correlations for both models under RCP4.5 and G4 separately are not atypical, simply that their G4-RCP4.5 differences are small' – I'm not sure what you mean by this sentence, please rephrase
61. L373 – Similarly the last sentence is unclear. Do you mean that all models excepts CanEMS2 and NorESM1 exhibit significant correlation between Ts and Vpot in all basins?
62. L376 – Change 'variability' to 'cycle' throughout this paragraph
63. L384 – Remove 'Comparing'
64. L388 – The last sentence – can you also plot the seasonal cycle of Ts in ERA-interim just to confirm that all the models are doing reasonably well here?
65. L390 – Consider splitting this paragraph into two. It is too long and unwieldy as it is
66. L391 – Sentence beginning 'In Figs 7d and 7e we plot' – please reword this sentence to something like 'we plot correlations between H / Vshear and Ts.
67. L399 – 'there is generally an anti-correlation *between Vshear and* Ts'
68. L402 – Vecchi *and Soden* (2007) found that wind shear increases in both the North Atlantic *and the East Pacific* under global warming.
69. L404 – 'If the models *assessed* here'
70. L407 to L415 – I'm not sure what you are trying to prove here, it seems peripheral and needs to be reworded
71. L441 – 'The analysis for individual basins indicates most models have significant correlations with ENSO in the WNP' – This is not true! Only 4/7 of the models have significant correlation in the WNP in RCP4.5
72. L448 – is there any previous studies that suggest a link between ENSO and tropical cyclone activity in basins outside the North Atlantic and the Pacific. If so, please cite
73. L449 – 'is most consistently felt in the Pacific Ocean' –particularly the South Pacific
74. L473 – why are the TRACK results so much lower in your Table 4 than in Jones et al. (2017)? For instance, you get 1.2 storms per year in the North Atlantic basin in G4 compared to ~11 per year in their work (their Fig. 4). Their reasoning behind the use of the (4.5,3.5,4) configuration was to attain ~10 storms per year on average in the historical period. Please check these numbers, they seem wrong.
75. L487 – Change 'typical' to 'current'
76. L489 – 'The storms that may be counted using indirect methods such as the TRACK algorithm include the whole climate condition' – This doesn't make sense to me. Consider replacing with 'Simulated storms that may be counted using methods such as the TRACK algorithm allow for feedbacks with the climate system'
77. L490 – 'Statistical methods (Moore et al., 2015) also implicitly include feedbacks between storm and climate conditions' – in what way do they include feedbacks? I don't understand this. They are simply diagnostics
78. L492 – 'but dynamical downscaling methods (Emanuel, 2013) cannot include them' – I disagree, Emanuel employs a simple ocean model which can be adjusted to provide

climate feedback. In fact, I think the semi-explicit scheme offers more opportunity to incorporate feedbacks than the statistical methods

79. L493 – change 'apply' to 'utilize'
80. L495 – change 'relatively little data' to 'coarse temporal-resolution data'
81. L502 – Change 'diagnose tropical storms in climate models' to 'relate tropical storm activity to ambient meteorology'
82. L507 – 'Thus stratospheric sulphate aerosol injection could lead to fewer TCs in the North Atlantic …' – note that this is one solar geoengineering scenario (a uniform one). Injecting aerosol preferentially into one hemisphere may increase the amount of storms in the North Atlantic (Jones et al (2017)) with unknown effects in other basins
83. L510 – 'The impact of ENSO on TCs can be detected in the GPI' – this is poorly worded, you have not explicitly looked at ENSO and TCs, only at ENSO and GPI. Rephrase in such a way: 'ENSO is found to be correlated with GPI'. Are there any implications specifically in terms of solar geoengineering from your results? I mean, is there a decrease in El Nino years in the G4 simulations?
84. L515 – remove 'such as'
85. L521 – 'a simplified representation of TCs depending on fewer variables is possible' -> 'a simplified representation of the GPI depending on fewer variables may be possible'
86. L523 – sentence running from 'it is encouraging that the thermodynamic state …' I don't understand what you mean here?
87. L529 – '(the 100hPa level)' -> (evaluated at 100 hPa)'
88. L529 – Replace 'note that' with 'find that changes to' and add 'changes' after GPI
89. L542 – rather than using temperature changes from Pitari et al (2014), can you give the ensemble mean upper-tropospheric temperature changes from your 6-member ensemble please
90. L542 – 'This is about half the range of the G4-RCP4.5 difference in static stability (Fig. 7)' – Figure 7 does not show that changes to static stability…
91. L544 – remove 'significant'
92. L545 – why does T0 not warm with most models under G4? Do you have a reason that you can offer? It is that the aerosol particles are small?
93. L578 – 'Many models, owing to their low resolutions, produce much weaker and larger TC' – this statement has been repeated a few times (e.g. L487). Please do not repeat statements
94. L582 – change 'would be' to 'this is'

References

Jones, AC, *et al.* (2017), Impacts of hemispheric solar geoengineering on tropical cyclone frequency, *Nature Communications*, 8, 1382

Jones, AC, *et al.* (2018), Regional Climate Impacts of Stabilizing Global Warming at 1.5 K Using Solar Geoengineering. *Earth's Future*

Li, M (2013), Bimodal Character of Cyclone Climatology in the Bay of Bengal Modulated by Monsoon Seasonal Cycle, *Journal of Climate*

Murakami, H, *et al.* (2016), Seasonal Forecasts of Major Hurricanes and Landfalling Tropical Cyclones using a High-Resolution GFDL Coupled Climate Model, *Journal of Climate*

Roberts, M, *et al.*, (2015), Tropical Cyclones in the UPSCALE Ensemble of High-Resolution Global Climate Models, *Journal of Climate*