Dear Referee,

Thank you very much for your helpful, insightful and detailed suggestions regarding our submitted paper, ‘How many active regions are necessary to predict the solar dipole moment?’. We have taken these into account and have adjusted the paper accordingly (changes made in the paper are in bold). Our responses to each recommendation are below. Figures from the paper are referred to as ‘Fig. n’ whilst Figures presented in this document are referred to as ‘Figure n’.

1. Page 1, last sentence: ‘After emergence, the magnetic flux then diffuses across the surface by being pushed to the edges of convection cells (Leighton 1964). This can lead to cross-equatorial flux cancellation between the leading polarities from opposite hemispheres, and any remaining trailing flux is transported poleward by a combination of diffusion and a meridional flow (Howard 1979) of the order $\sim$10–20 $\mathrm{m\,s^{-1}}$ (e.g. Komm et al. 1993; Jackiewicz et al. 2015).’ The wording here is awkward and insinuates that only leading polarity flux can be transported across the equator and that only trailing flux can be transported poleward, when in fact both polarities can cross the equator or both can be transported poleward. I suggest rephrasing for clarity.

- This paragraph now reads: ‘*After emergence, the magnetic flux diffuses across the surface by being pushed to the edges of convection cells (Leighton 1964), is advected poleward by meridional circulation, and sheared by differential rotation. Due to the combined effects of Hale’s and Joy’s laws, the net result of this process is the cancellation of leading polarity flux across the equator and the accumulation of trailing polarity flux at the poles. This cancels the polar flux of the previous cycle and builds up new polar flux of the opposite polarity.*’

2. Page 2, Section 2, first paragraph: Please include the resolution of the NSO Kitt Peak and SOLIS synoptic magnetograms and the cadence at which the data are assimilated.

- These details have now been included.

3. Equation 2.5: The axial dipole of an active region is expected to change significantly during its emergence and subsequent evolution. You might want to mention that this measurement is a snapshot of highly variable parameter. I assume you are only assimilating once per rotation, therefore you are sampling the initial $D(t)$ at different phases of in the life of the Active Region for each Active Regions (e.g., emergence, peak, decay) and this should be mentioned as well.

- This is a good point. We have added the following to the penultimate paragraph of Section 2, page 5: ‘*Here the initial axial dipole moment of an active region is measured at the time of assimilation, that is, on the day it crosses the central meridian*’. This means that we are capturing the properties of each active region at different stages of their evolution.

4. Page 4, end of Section 2: Details about the Active Region detection and assimilation are extremely lacking. The authors need to include a paragraph that describes this more thoroughly. In particular, the following questions should be addressed: How is an Active Region defined (e.g., all contiguous pixels above B_{par}, a certain area surrounding B_{par}, etc.)? How frequently are Active Regions included (e.g., once per rotation, multiple times per Active Region)? Are the returning Active Regions from the next map considered a new Active Region or a continuation of the old Active Region? If the later, how does $D(t)$ for specific Active Regions change in time? Can you show plots of this for the nine Active Regions shown in Fig. 2? It could easily be included as an additional panel to Fig. 2.

- We agree that there needs to be a more thorough description of the assimilation process, and the first half of Section 2 (page 3) has been changed accordingly. In short, the answers to the questions are: The synoptic magnetograms are corrected for flux imbalance and smoothed, then the active regions are defined as all areas of connected pixels greater than B_{par}. Further details of the model and assimilation algorithm can be found in Yeates et al. (2015), which we have now highlighted in the first half of Section 2. Active regions are included once per Carrington rotation, but returning regions on the next map are considered new regions and replace the pre-existing ones (rather than being superimposed) so that the axial dipole moment contribution from a returning region is not counted twice. We would like to have been able to count a repeated region as a single region, but it is not always easy to define whether two regions from different synoptic maps are the same because of the flux emergence or cancellation that may have occurred between the two observations (top right paragraph on page 3, see also point 10 below). This is particularly hard to do using an automated procedure as we would have to in the model. Two such regions appear in Fig. 2, which we have discussed in Section 2 (bottom of page 4): '*Among these are two regions that share similar features (left and centre panels of the middle row), and are likely to have been the same region appearing in two consecutive rotations, having undergone some sort of interaction in the interim.*'

5. Page 5, Section 3.1, paragraph 2: 'There is then a sharp decrease when the two biggest contributions are included...'. I assume you mean biggest contributors in terms of dipole (rather than flux), but you should state this explicitly.

- This is true, and this detail has been added.

6. Page 5, Section 3.1, paragraph 2: 'Note that more than 25% of $D_{\mathrm{tot}}(T)$ is attained by the largest few regions alone. This is the effect of exponential decay, which decreases the axial dipole moment contribution of a small number of regions so that it is closer to D_{tot} at the end of a cycle (see Fig. 3 and compare with Fig. 10 in Appendix A). This effect is even stronger for the other two cycles.' This is not clear, please rephrase for clarity. Also, I'm not convinced this is an effect of the exponential decay - it may be 'enhanced' by the decay, but I don't think it is 'caused' by the decay. This can be easily determined by including lines for the non-decay cases in your Fig. 4. If the offsets are still present in those lines as well, then it is not cause by the decay. Adding these lines to the figure may also help to provide additional clarity to your description of the effect.

- Admittedly this was poorly written; the intention was just to explain why Final D_{rel} starts at 25-40% when only a small percentage of regions are added, as shown in Fig. 4 in the paper. It is more a side-effect of the measure we use. When decay is not present (Fig. 10) and e.g. 10 regions are included, the end-of-cycle dipole moment is far away from the original end-of-cycle dipole moment (thick black line). However when we include decay (Fig. 3), these profiles both go closer to zero, thereby reducing the difference between the two end-of-cycle dipole moments and hence increasing the *relative* dipole moment obtained by the 10 regions. This paragraph has been re-worded along the same lines (page 7). As suggested, we have plotted the equivalent profiles where decay wasn't used in Figure 1 below, in the same plot as the profiles where decay was used. We hope it demonstrates that there are no changes in the small variations of each profile, just a different starting point. It is also a cluttered image, and so for these reasons we opt to exclude it from the paper, sticking with the original Fig. 4.

7. Page 5, Section 3.1, throughout: I expect that the flux by percentage of Active Regions varies significantly per cycle. It would be helpful to see a plot of this. It could easily be included as an additional panel to Fig. 4.

- Figure 2 below shows the requested plot (assuming we have interpreted the request correctly). Whilst there is a difference between cycles as predicted, we do not consider this to be significant enough to include in the paper (perhaps a future study?). If you disagree then we can include it as part of Fig. 4 as suggested.

8. Page 5, Section 4: This section feels out of order, it makes more sense to see the details about the Active Regions before the simulations. However, this is merely a note to the authors and they may keep the present order if it is preferred.

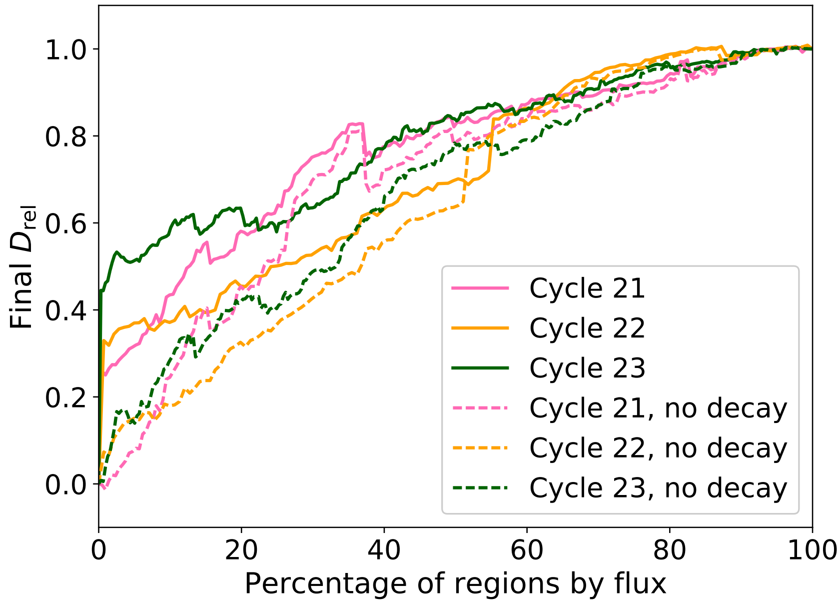

Figure 1: Final $D_{\rm rel}$ against percentage of regions included for Cycles 21 (pink), 22 (yellow) and 23 (dark green). Simulations performed with a decay term are denoted by solid lines, and simulations without decay are denoted by dashed lines. Regions are ordered by flux and the top x% of the strongest regions are incorporated.

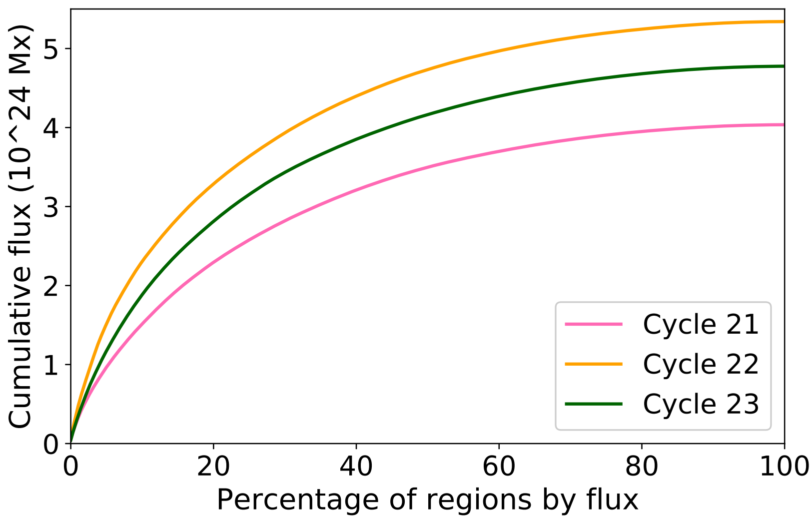

Figure 2: Cumulative against percentage of regions included for Cycles 21 (pink), 22 (yellow) and 23 (dark green). Regions are ordered by flux and the top x% of the strongest regions are incorporated.

- The paper was originally ordered as you suggested, but we decided to swap the two main sections around so that our main result about removing/including active regions appeared first. We consider the active region properties to be additional analysis and so we have designated this to be secondary. For these reasons we opt to keep the same structure.

9. Page 5, Section 4, paragraph 1: When is an Active Region determined to have reached the time of its final $D_{\mathrm{rel}}(t)$? Is this the time step before the active region falls below B_{par}? Earlier, when you ordered the ARs by $D_{\mathrm{rel}}(t)$, which $D_{\mathrm{rel}}(t)$ was used (e.g., initial, final, something in between)? This should be explicitly stated.

- It seems that there is some confusion surrounding the definition of the relative dipole moment $D_{\mathrm{rel}}(t)$, particularly the final dipole moment which we denoted by the time $t = T$. We now denote the start and end times of each cycle as t_{start} and t_{end} respectively throughout the paper, and the specific dates of these are now given after the (updated and more thorough) definition of $D_{\mathrm{rel}}(t)$ in the second half of Section 2, page 4. Hopefully this clarifies that we are ordering the ARs by final $D_{\mathrm{rel}}(t)$ when plotting Figs. 3 and 10. This does also mean that we are taking the end of the cycle to be the same for all active regions, and so some late-emerging regions will not have completed their evolution and subsequent contribution to the axial dipole moment (see also point 11). However, we are only interested (in this study) in the end-of-cycle dipole moment, because this is the proxy used for future cycle amplitude predictions. The effect of late-emerging regions on future cycles is an interesting topic for future study (see penultimate paragraph of conclusions, page 12).

10. Page 5, Section 4, paragraph 1-2: 'We find that most significant contributors to the axial dipole moment emerge below $\pm 20°$, the very largest of which emerge below $\pm 10°$. We also find that these regions do not necessarily have strong levels of magnetic flux... We discover that the relationship between initial and final D_{rel} is largely determined by the emergence latitude:' I am very concerned that these conclusions may include selection bias based on how the active regions are defined or how the final $D_{\mathrm{rel}}(t)$ is determined. For instance, this could be due to in part to: 1) Active regions in some latitude may be more like to undergo interactions with neighbouring active regions and 2) active regions at high latitudes are more likely to have flux leak out of the selection bounding window, due to the shearing effect of the differential rotation. Once these two things begin to happen, I'm not sure can still consider the $D_{\mathrm{rel}}(T)$ the contribution of that Active Regions alone. Similarly in paragraph 2 with the discussion of the cross-equator cancellation - once this occurs, the contribution to the dipole is no longer due to that AR alone. In other words, I am not convinced that the final $D_{\mathrm{rel}}(t)$ is an accurate representation of the active region's contribution to the axial dipole. Please take a closer look at the evolution of $D_{\mathrm{rel}}(t)$ for some of the Active Region to ensure that these conclusions are robust. It may also be beneficial to see the Fig. 5 plots as a function of the initial D_{rel}. Most importantly, please demonstrate how the final $D_{\mathrm{rel}}(t)$ is defined and that it is an accurate (or at least a reasonable) representation of the active region's contribution to the axial dipole. For example, can you show that the cumulative sum of final $D_{\mathrm{rel}}(t)$ for the 10 (100, 250..) excluded active regions is consistent with D_{tot} that you obtained when you ran the simulation without those active regions?

- Note that $D(t)$ is calculated over the whole solar surface, not just in the original assimilation region (Equation 2.5). Moreover, t_{end} is the same for all regions from the same cycle. In the absence of repeated regions, interactions with other regions are linear, so the cumulative sum of 10 regions would match the simulation of the same 10. However, summing the final contributions with returning regions included would overestimate the final dipole moment. This is why we cannot simply sum the contributions in our analysis. Instead we must perform the simulations with 10, 100, etc. regions with replacement included, so that the correct dipole moment is produced. This method ensures that the repeated regions don't affect our conclusions. We have outlined this in the text (page 4): '*[...the magnetic field of a region is] computed after its initial insertion by solving Equation 2.1 with no other field present. Isolating the evolution of a single region like this is meaningful because Equations 2.1 and*

2.5 are approximately linear, so that the contributions $D^{(i)}(t)$ may be added together to give the overall dipole moment $D_{\rm tot}(t)$. The linearity is only approximate because our newly inserted regions replace pre-existing flux, and strong returning regions from the previous rotation are treated as new regions, as discussed above. Nevertheless, the evolution of the strongest of a set of repeated regions is a good approximation, and it is therefore useful to isolate them'.

We have also plotted the Fig. 5 plots as functions of initial $D_{\rm rel}$ instead (Figure 3 here) as requested. We acknowledge that this also depends on emergence latitude, flux and shape/orientation. However, since in this paper we are primarily interested in the final dipole moment contribution, we would prefer to leave this figure out (especially since we already have a lot of figures). Furthermore, we can find the initial $D_{\rm rel}$ of a region given its latitude and final $D_{\rm rel}$ based on the right-hand column of Fig. 5 or on Fig. 6.

11. Page 7, last paragraph: 'While the cross-equatorial group is important for reasons discussed above, the majority of regions in the late-emerging cluster might not have had as significant an effect on the current cycle as if they had instead emerged earlier in the cycle, as discussed by Nagy et al. (2017).' While you to cite a reference, without reading that reference, it is completely unclear why you think the timing of the emergence would change the contribution. Please include a summary of explanation in addition to the reference.

- The paragraph now reads (page 9): '*While the cross-equatorial group is important for reasons discussed above, the majority of regions in the late-emerging cluster might not have had as significant an effect on the current cycle as if they had instead emerged earlier in the cycle, as discussed by Nagy et al. (2017), who inserted an extreme active region into a dynamo model simulation at different times throughout a cycle and found that late-emerging regions had the smallest effect. This is because any poleward-advected flux would not have had enough time to reach the pole and cancel with the polar field before the end of the cycle.*'

12. Page 8, Fig. 5: I found myself drawing the unity line on the third panels of Fig. 3. It would be useful if you did this for your readers.

- This is a sensible suggestion and the identity line has been added to the third panels of Fig. 5.

13. Page 9, Fig. 6: Again I am concerned about how the final $D_{\rm rel}(t)$ is defined. Is it consistent for all active regions? Moreover, is it consistent across all cycles? If not, this figure and the conclusions from it could be misleading.

- See 9.

14. Page 9, end of 2nd paragraph: '...we concluded that emergence latitude is the dominant parameter controlling the amplification or suppression of the initial dipole moment of a region.' If this does turn out to be a robust conclusion, please comment on why this latitude dependence exists, i.e., what are the mechanisms that cause these changes in the $D_{\rm rel}(T)$.

- We have included the explanation of latitudinal dependence in the same paragraph (page 11): '*This latitude dependence exists because a large dipole moment arises from hemispherical polarity separation, which occurs most effectively when regions emerge tilted and at low latitudes so that cross-equatorial transport of flux can occur (Wang & Sheeley 1991; Yeates et al. 2015)*'.

15. Page 9, 3rd paragraph: '...most large contributors had fluxes of less than 2×10^{22} Mx'. I expect that the smallest (in flux) active regions don't contribute much either, might be more beneficial to list this as a range.

- After plotting the regions by flux on a log-scale, it appears that there are no large contributions

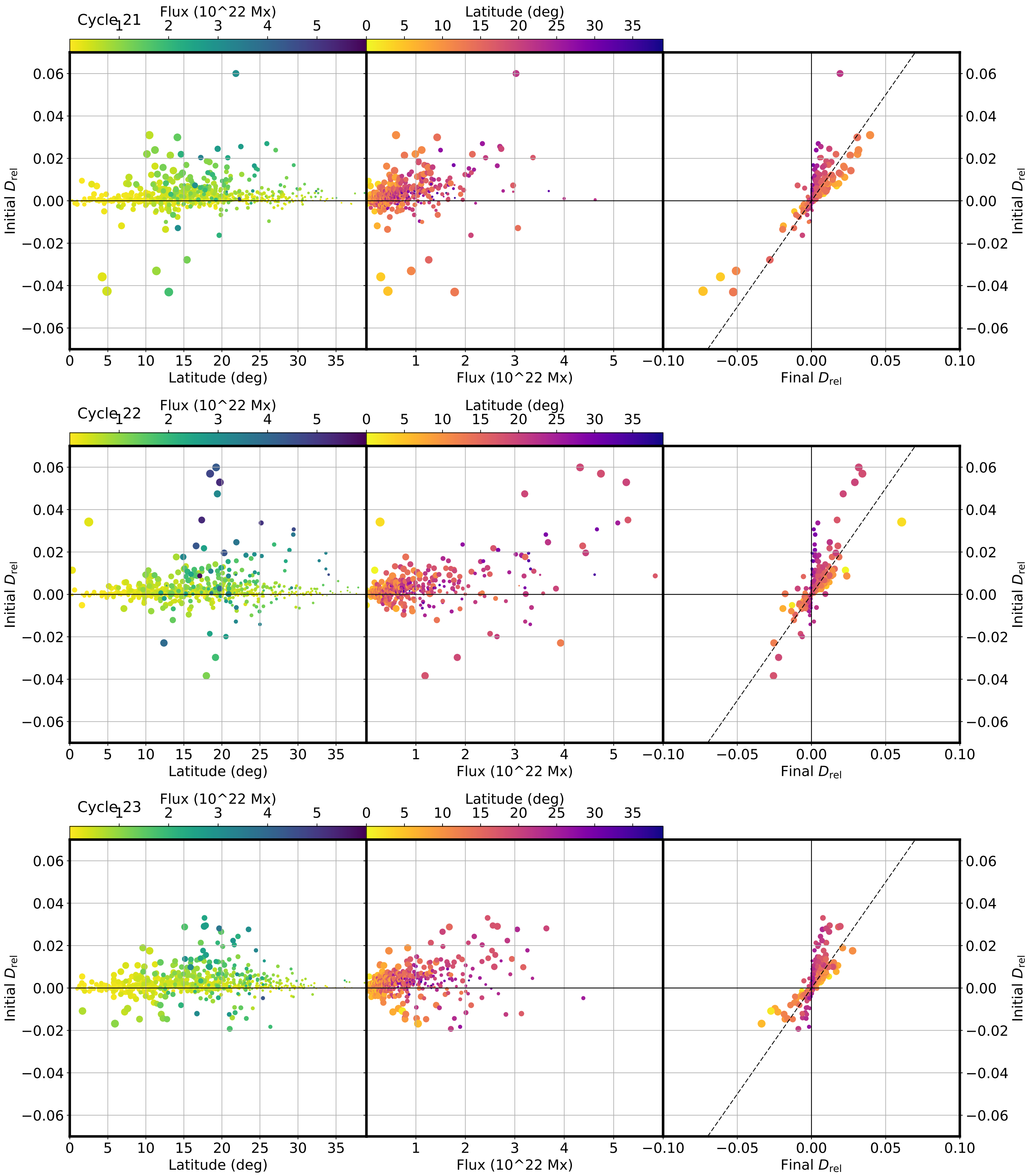

Figure 3: Initial $D_{\rm rel}$ for each region against absolute latitude (left panels), flux (middle panels) and final $D_{\rm rel}$ (right panels). Markers are sized by absolute final $D_{\rm rel}$, and coloured by flux (left panels) and absolute latitude (middle and right panels).

to the dipole moment below 1×10^{21} Mx. This information has been added to the text (page 11): '*Incidentally, across all cycles there were no significant contributors with fluxes less than* 1×10^{21} *Mx, indicating that the smallest regions are not able to drastically alter the axial dipole moment, regardless of emergence latitude.*'