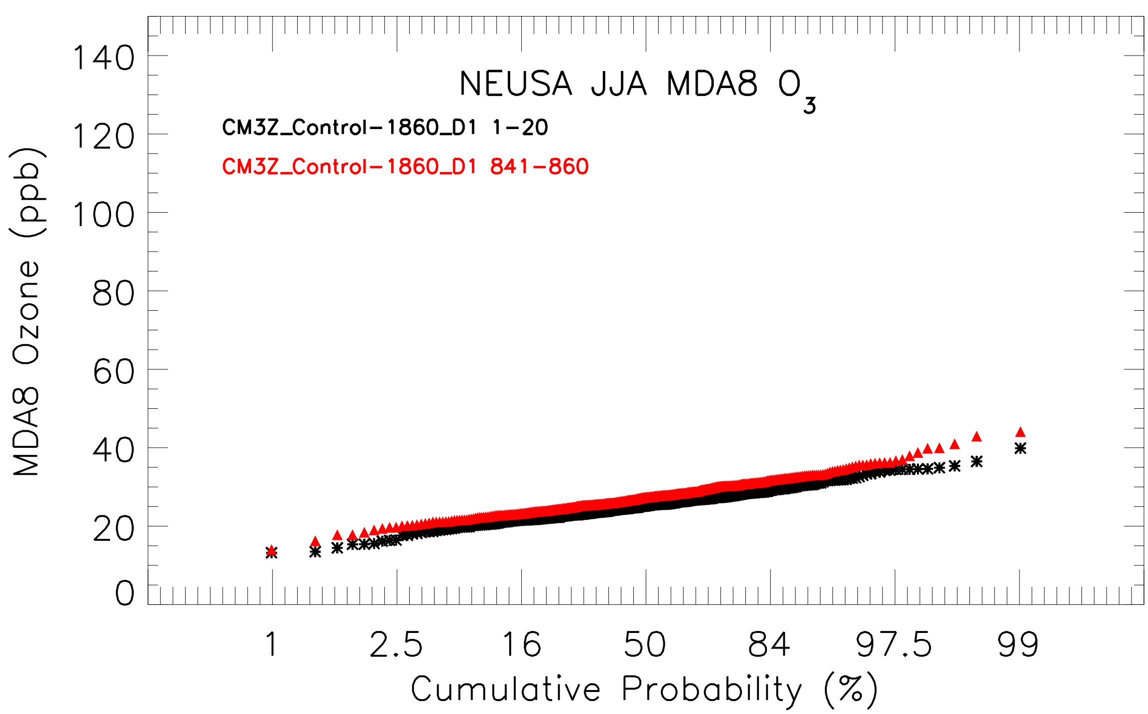
We appreciate the extremely insightful reviews on this manuscript and think they have greatly improved the quality of this manuscript. We have listed the Reviewer comments in gray and the responses are shown below in black. Relevant changes to the text are also listed (preceded by the line number of the changed text).

Reviewer #1 Comments:

1.) *The natural variability of cyclones is examined in the Pre-Industrial run, but there is no analysis of ozone (or high ozone events). Wouldn't this long simulation give a more robust definition of the relationship (correlation) between high-O3 events and cyclones? This would require redefining high-O3 events, but this can be done as for the historical simulation. I think a version of fig 4 showing time series of frequency O3 events, together with scatter plot of cyclones versus O3 events (e.g. Fig 8c) would be helpful additions to the paper.*

We did not use the long control simulation to characterize the relationship because the emissions are representative of pre-industrial conditions. This environment is fundamentally different than the present-day conditions since the background level is not substantially lower than "polluted" conditions (See Figure below showing the ozone MDA8 CDF for two 20-year periods in the control simulation). In this case, we expect the anti-correlation between cyclones and high-O_3 events will break down since the high events in the absence of anthropogenic emissions are likely driven by background.



We have added the following text to the manuscript, Line 230: "We do not conduct the same analysis to characterize variability in MDA8 ozone events because the distribution in the control simulation is fundamentally different from the present-day distribution due to the absence of anthropogenic emissions."

2.) *The main conclusions is that the sensitivity of high-O3 to cyclones does not explain much of the variability. This then raised the question of what are the factors that determine the variability. I think there needs to be some discussion / speculation of the possibly important factors.*

Agreed. We have added text to address this point:

Line 365: "Although we find no strong evidence of cyclone frequency explaining the variability of high-O_3 events, recent work by Barnes & Fiore (2012; submitted) suggests that the jet position in the model explains a substantial portion of surface ozone variability over the Eastern United States. Further investigation of the relationship between ozone variability (including the incidence of high-O3 events and storm counts) and their connection to jet position is warranted."

3.) *I understand the desire to use a comprehensive storm tracking algorithm, but how dependent are the results to this. Would the results be any different if you had used a simpler scheme (e.g., just use the statistics from first part of algorithm before tracking)? I am not sure how much is involved to do this, but it would be useful to know if tracking is really needed (especially as 6 holy data is not always available).*

We have provided a summary of statistics from center finding and the storm tracking below. Additionally, we have three plots showing the results of the center-finding, tracking, and the fraction kept during tracking. However, we feel this is somewhat disconnected from the rest of the paper, as the goal of the paper is not a comparison of different storm tracking methodology and prefer not to include it in the main text. Raible et al., *Mon. Wea. Rev.* (2007) provide a fairly thorough comparison of three different storm tracking algorithms: the University of Hamburg approach (HAM), ETH Zurich approach (ETH), and Australian scheme (AUS).

NRA2 1979--2010

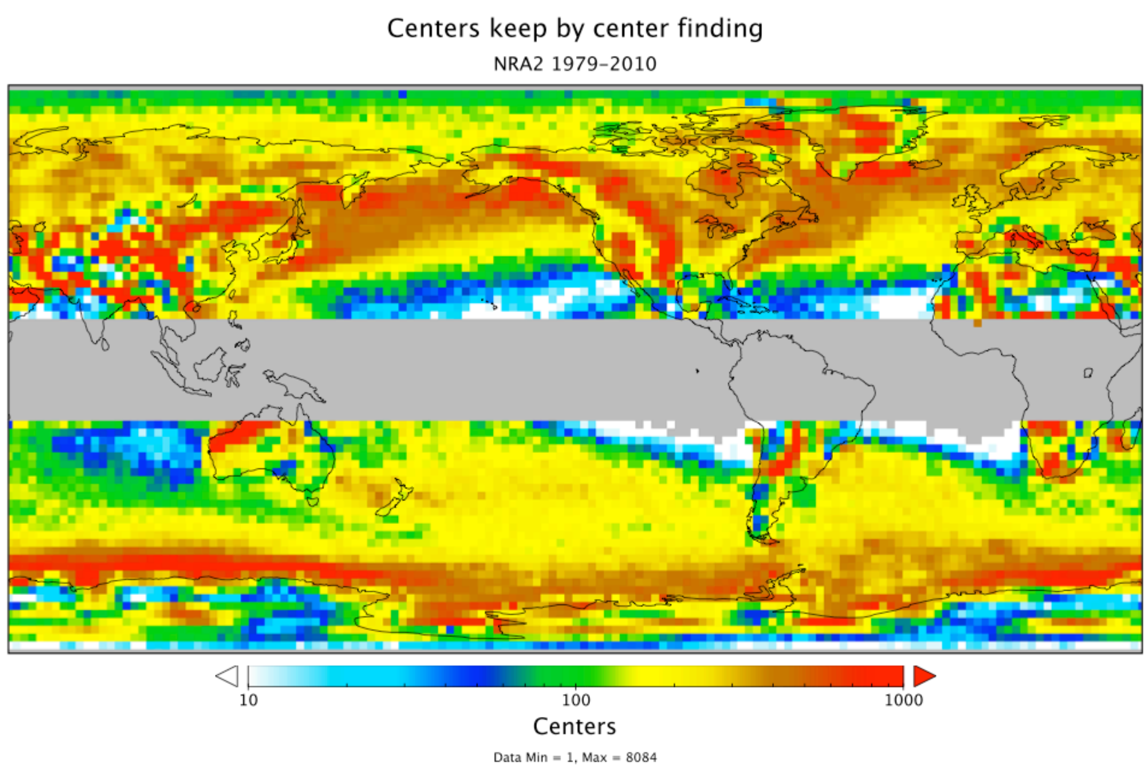
Results of just center finding:

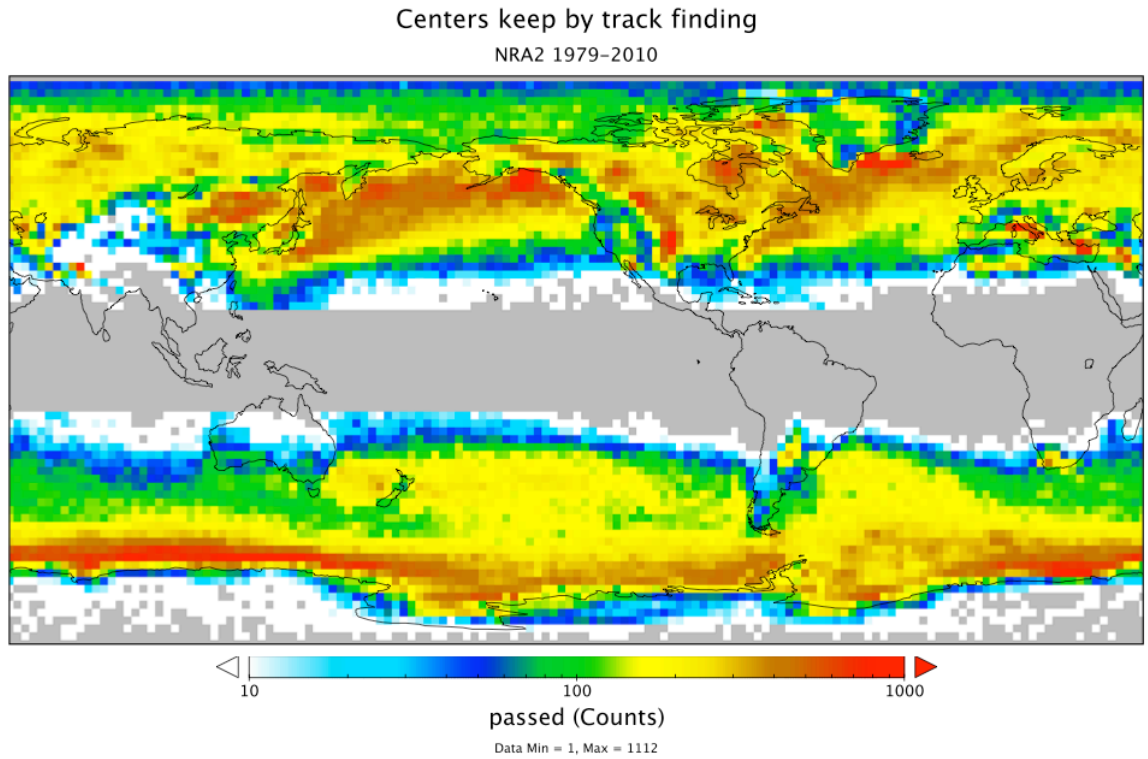
Final Center CNT: 2122891 (50.95%) from 4166585 candidates where
2122891 (50.95%) Failed all filters
1148282 (27.56%) Failed concavity/Laplacian filter
822970 (19.75%) Failed regional minimum filter
72442 (1.74%) Failed troubled center filter

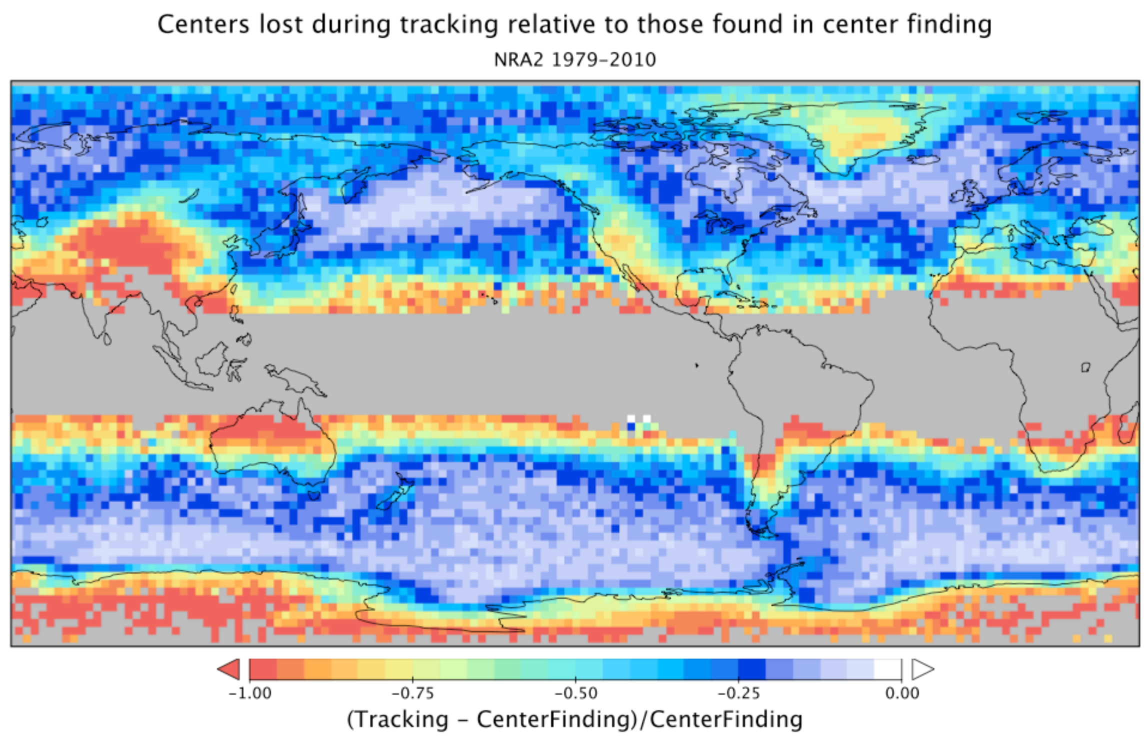
Post Tracking:

Total Centers Read: 2122891
Total Centers Saved: 1242986 (58.55%)
Total Tracks: 93956
Discards:
191645 (9.03%) Failed trackable center
308351 (14.53%) Failed track lifetime filter
127852 (6.02%) Failed track travel filter
63129 (2.97%) Failed track minimum SLP filter
188928 (8.90%) Failed extratropical track filter

Here you'll see that tracking mostly discards low latitude centers and likely noise over high topography.







Reviewer #2 Comments:

4) *The distinctions between this work and Leibensperger et al. (2008) should be clearly laid out in the Introduction and methods, since a number of figures are similar to those in Leibensperger et al.*

We have included a brief description of differences between this work and that of Leibensperger et al. (2008).

Line 65: "Building upon their (Leibensperger et al., 2008) work, which focused on the past few decades, we examine the spatial distribution, trends, and variability of mid-latitude cyclones in the Geophysical Fluid Dynamics Laboratory (GFDL) Climate Model version 3 (CM3) simulations of Pre-industrial, present, and future climate as well as in four reanalyses. We then examine the relationship between summertime mid-latitude cyclones and high-O_3 events in future climate projections."

5) *Table 2 presents results from a comparison of the reanalyses datasets over the full time period that the reanalyses is available for. This time period differs considerably amongst the datasets listed in Table 2. The comparison of the NCEP/NCAR reanalyses over the shorter period equivalent to that used in Liebensperger et al. is rather more insightful. Some further comparison of trends for all the reanalyses datasets for as similar periods as possible would be a useful addition to Table 2, as well as some discussion of differences in trends on decadal vs. centennial timescales.*

We have added trend analysis of the GFDL CM3 ensemble members, NCEP/NCAR Reanalysis 1, NCEP/NCAR Reanalysis 2 from 1980–2006 (the period examined by Leibensperger et al. 2008) as well as ERA-40 from 1964-1990 (the same record length as the 1980–2006 period). We did not include additional trend analysis for ERA-Interim because the period was already shorter than that of Leibensperger et al. 2008. We have reproduced the table below.

Table 2. Data used during the Historical time period (1860–2005). Mean values and standard deviations are in units of cyclones per summer (JJA), trend, p-value of an ordinary least-squares regression, and the variability ($\sigma/\mu \times 100$) is expressed as a percentage.

Dataset	Time Period	Mean μ	Standard Deviation σ	Trend m (p-value)	Variability RSD	Reference
GFDL CM3 Historical	1980–2005	14.45	3.41	-0.01 ($p = 0.80$)	23.6 %	Donner et al. (2011)
GFDL CM3 Historical (H1)	1980–2005	14.88	3.17	0.05 ($p = 0.58$)	21.3 %	
GFDL CM3 Historical (H2)	1980–2005	14.23	3.13	-0.07 ($p = 0.42$)	22.0 %	
GFDL CM3 Historical (H3)	1980–2005	14.69	4.07	-0.05 ($p = 0.65$)	27.7 %	
GFDL CM3 Historical (H4)	1980–2005	13.74	2.73	0.03 ($p = 0.66$)	19.9 %	
GFDL CM3 Historical (H5)	1980–2005	14.77	3.97	-0.02 ($p = 0.83$)	26.9 %	
NCEP/NCAR Reanalysis 1	1958–2010	14.49	3.52	0.02 ($p = 0.56$)	24.3 %	Kalnay et al. (1996)
NCEP/NCAR Reanalysis 1	1980–2006	14.31	3.67	-0.15 ($p = 0.04$)	25.7 %	
NCEP/NCAR Reanalysis 1	1989–2010	14.59	4.04	0.06 ($p = 0.65$)	27.7 %	
NCEP/NCAR Reanalysis 1	1961–1990	14.67	3.07	0.05 ($p = 0.43$)	20.9 %	
NCEP/DOE Reanalysis 2	1979–2010	13.56	3.37	0.05 ($p = 0.42$)	24.8 %	Kanamitsu et al. (2002)
NCEP/DOE Reanalysis 2	1980–2006	13.19	3.32	-0.00 ($p = 0.99$)	25.2 %	
NCEP/DOE Reanalysis 2	1989–2010	13.86	3.52	0.07 ($p = 0.57$)	25.4 %	
ERA-40 Reanalysis	1961–1990	13.50	2.60	-0.02 ($p = 0.67$)	19.2 %	Uppala et al. (2005)
ERA-40 Reanalysis	1964–1990	13.48	2.50	-0.03 ($p = 0.67$)	18.6 %	
ERA Interim Reanalysis	1989–2010	20.59	4.28	0.01 ($p = 0.93$)	20.8 %	Dee et al. (2011)

Including these shorter record length analyses does not impact our findings as none of these additional time periods show significant trends.

We have added the following text to the manuscript,

Line 196: "We sub-sampled the reanalysis datasets to compare trends over similar time periods, however only the NCEP/NCAR Reanalysis 1 (1980–2006) time period yielded a significant trend."

6) *The study makes use of a control simulation to quantify internal model variability in cyclone frequency but doesn't make use of the 5 ensembles for the historical period to quantify 20th century variability in cyclone frequency (ranges based on min and max ensemble members are depicted in Fig. 3 only the ensemble mean is discussed hereafter). Analysing the trends in cyclone frequency over the historical period would relate more directly to the findings of Liebensperger et al. who find a role of climate change between 1980-2006 on cyclone frequency. Hence, the analysis performed in Fig 4 performed for the 5 historical ensemble members (trends and their significance) may yield useful insights. If space is tight I think the analysis of the historical runs may be more useful that the analysis of the control run.*

We now include a brief discussion of the range of cyclone frequencies in the historical simulations (see response to Q7, below). We have also included trend analysis in the Historical period over the same time window that Leibensperger et al. (2008) analyzed (see table above). However, we feel that the control simulation still adds substantial value to the paper because we are able to more accurately assess the internal model variability in the absence of any time-varying forcings (which affect the Historical simulation). This analysis with the control simulation provides a baseline for the rest of the simulations with time-varying forcings and acts as a rudimentary test of the applicability of t-statistics to this dataset (do we reject the null hypothesis the correct fraction of the time). Performing a similar analysis with the historical period would be substantially more difficult because there are non-linear trends in emissions.

7) *These ensemble member would also be useful for assessing the importance of the changes in summer storm track frequency between 2006-2025 and 2081-2100. i.e. are the differences in cyclones /summer between 2081-2100 and 2006-2025 (Fig 6 , 3rd column) greater than the differences in cyclones/summer between the 5 ensemble members for present-day.*

This is an excellent question. We found a difference between the mean cyclone frequencies of 1.14 cyclones/summer between the maximum and minimum ensemble members averaged over 1980-2005. We have added text to the manuscript reflecting this point.

Line 198: "The variability ranges from 19.9% – 27.7% with a mean difference of 1.14 cyclones per summer."

8) *21685: line 14 "any storm tracking through the region"-How much of the storm track (number of points along its track) needs to fall in the bounded region for the "storm" to be included and how sensitive are the trend results to such assumptions?*

We confirmed with Dr. Eric Leibensperger that in Leibensperger et al. (2008) any storm that was inside the bounded region at the end of a timestep was counted as part of the GLST. For comparison with Leibensperger et al. (2008), where a strong relationship was found, we used this same metric. However, we performed some tests where each storm was weighted by the residence time in the bounded region and found that it didn't significantly change the frequency or relationships if we used this residence time weighting.

In this method (and the one from Leibensperger et al., 2008), a storm that is only in the bounded region for a single timestep would be weighted equally with a storm that travels the length of the region. However, it is actually the cold front that is ventilating the stagnant air mass. So once the boundary layer has been ventilated the additional residence time of the storm will not continue to be anti-correlated with extreme ozone concentrations. So the method of weighting by residence time will also have drawbacks. We chose to remain consistent with Leibensperger et al., (2008) for a more direct comparison.

We have clarified our method of counting storms in the text.

Line 140: "Following Leibensperger et al. (2008), we count any storm tracking through the region bounded by 70–90°W and 40–50°N as part of the GLST, depicted as the gray box in Fig. 1. For comparison with Leibensperger et al. (2008), the duration of the storm in the GLST is not taken into account and the results were found to be insensitive to this assumption."

9) *In a number of places numerical values in the text should be accompanied by the relevant period e.g. "a reduction of 5.7 cyclones in summer between 2006 and 2100".*

We thank the reviewer for pointing this out and have changed the manuscript to include this clarification as well as other places where the relevant period was not stated.

10) *In places numerical values are given in the text that are hard to see based on the figures scales as is. This applies in particular to Figs. 1 and 2. Make sure the figures scales and companion body text are compatible.*

Some of the figures in the ACPD paper are smaller than they will be when in print due to the reduced page length. The text on the figures seem quite reasonable in the draft form I have but we will be sure to check this at the proof stage.

11) *The abstract contains a lot of technical details that could be trimmed especially as these are given in the methods e.g. the details of the RCP 4.5* scenario. Some p values could be trimmed.*

We have trimmed some of the more technical details from the abstract.

12) *The authors state that their findings do not refute Yin (2005) yet the results in Fig. 5 do seem to differ from Yin (2005) in that there is no obvious dipole of change in behaviour. The argument given could be rejected or confirmed by considering percentage or normalised changes.*

We stated that our results do not refute Yin (2005) because all of the figures in Yin (2005) are zonally averaged. Our study focused on a single region in North America. So while our findings do seem to differ from Yin (2005) it could simply be due to differences at longitudes we did not consider. Our findings are consistent with those of Lambert and Fyfe (2006).

We have clarified the text to explain why our findings for a single region do not refute the zonally averaged findings of Yin (2005).

Line 249: "our findings do not necessarily refute the potential shift reported by Yin (2005) because they examined zonally averaged quantities whereas we focus on a single region. Additionally, Fig. 5c indicates a regional reduction in storm frequencies over the mid-latitudes with negligible changes at higher latitudes. This could indicate a shift in storm tracks that is masked by an overall reduction in storms."

13) *The authors also state that their findings are more consistent with Tai et al. (2012) who performed analysis for the period 1999-2010, yet the authors state above this text that on decadal rather than centennial periods that they find some periods in which cyclone frequency and high ozone events are strongly anti-correlated (r=-0.79).*

We have removed this line from our manuscript as it is not critical to our arguments.

14) *Storm track methodology description- (21684: line 21): It is unclear how this methodology (Bauer et al. 2102) differs from that used in Liebensperger et al. (2008) or from the other studies cited- how is it more comprehensive?*

The main difference is that the cyclone tracking method used by Liebensperger et al., (2008) is based mostly on the detection of SLP minima and a simple nearest neighbor method of associating them into tracks. The method used here (MCMS) is a much updated method that allows for example the inclusion of partial or weak minima that would have been missed with the earlier system in addition to better, more adaptive screening for noise over topography. The MCMS tracking algorithm is likewise more sophisticated in terms of judging whether to associate two cyclone centers over time, rather than simply selecting the closest spatial neighbor in the next time step. Ultimately MCMS finds on the order to 20% more cyclone centers than did the older system, and in particular for this paper it is better at following cyclones around the mountainous western US which is upstream of the study region.

15) *Can the authors check the p-values throughout the text and tables. In fig 4. panels (d-g) seem to show clearer trends than panel (i) yet the p-values vary widely.*

We have verified the p-values presented. With respect to Fig. 4, it is important to remember the t-statistic is proportional to the magnitude of the slope and inversely proportional to the variance,

$$t \propto \frac{\hat{m}}{\sigma}$$

The reviewer specifically points out panels (d-g) and (i). Panel (i) has the smallest variance (2.80) and largest slope ($0.02a^{-1}$) of the panels mentioned, thus the smallest p-value. We have provided a table of cyclone frequencies for the panels the reviewer mentioned.

Panel (d) 1286-1385		Panel (e) 1381-1480		Panel (f) 1476-1575		Panel (g) 1571-1670		Panel (i) 1761-1860	
1286	20.0	1381	15.0	1476	7.0	1571	11.0	1761	14.0
1287	17.0	1382	23.0	1477	11.0	1572	6.0	1762	18.0
1288	11.0	1383	17.0	1478	16.0	1573	10.0	1763	11.0
1289	11.0	1384	10.0	1479	15.0	1574	15.0	1764	15.0
1290	16.0	1385	10.0	1480	17.0	1575	11.0	1765	16.0
1291	14.0	1386	13.0	1481	16.0	1576	10.0	1766	13.0
1292	12.0	1387	7.0	1482	17.0	1577	12.0	1767	12.0
1293	19.0	1388	18.0	1483	16.0	1578	12.0	1768	10.0
1294	17.0	1389	20.0	1484	21.0	1579	13.0	1769	18.0
1295	15.0	1390	17.0	1485	18.0	1580	13.0	1770	13.0
1296	15.0	1391	16.0	1486	12.0	1581	15.0	1771	10.0
1297	14.0	1392	12.0	1487	10.0	1582	14.0	1772	20.0
1298	17.0	1393	18.0	1488	18.0	1583	17.0	1773	11.0
1299	15.0	1394	14.0	1489	16.0	1584	19.0	1774	12.0
1300	12.0	1395	14.0	1490	14.0	1585	14.0	1775	15.0
1301	17.0	1396	15.0	1491	9.0	1586	12.0	1776	18.0
1302	12.0	1397	14.0	1492	13.0	1587	13.0	1777	14.0
1303	11.0	1398	15.0	1493	9.0	1588	11.0	1778	11.0
1304	13.0	1399	12.0	1494	21.0	1589	15.0	1779	12.0
1305	21.0	1400	15.0	1495	12.0	1590	9.0	1780	13.0
1306	10.0	1401	14.0	1496	10.0	1591	14.0	1781	11.0
1307	15.0	1402	10.0	1497	17.0	1592	19.0	1782	14.0
1308	14.0	1403	14.0	1498	15.0	1593	16.0	1783	10.0
1309	11.0	1404	15.0	1499	15.0	1594	19.0	1784	9.0
1310	19.0	1405	11.0	1500	10.0	1595	13.0	1785	17.0
1311	10.0	1406	14.0	1501	11.0	1596	18.0	1786	10.0
1312	18.0	1407	18.0	1502	12.0	1597	11.0	1787	13.0
1313	19.0	1408	16.0	1503	17.0	1598	12.0	1788	11.0
1314	18.0	1409	16.0	1504	14.0	1599	14.0	1789	11.0
1315	16.0	1410	16.0	1505	13.0	1600	12.0	1790	10.0
1316	15.0	1411	11.0	1506	14.0	1601	13.0	1791	15.0
1317	21.0	1412	14.0	1507	17.0	1602	13.0	1792	12.0
1318	16.0	1413	14.0	1508	16.0	1603	13.0	1793	10.0
1319	19.0	1414	15.0	1509	14.0	1604	16.0	1794	16.0
1320	12.0	1415	14.0	1510	15.0	1605	16.0	1795	14.0
1321	13.0	1416	10.0	1511	16.0	1606	15.0	1796	17.0
1322	12.0	1417	17.0	1512	13.0	1607	21.0	1797	14.0
1323	11.0	1418	12.0	1513	15.0	1608	15.0	1798	15.0
1324	13.0	1419	16.0	1514	12.0	1609	8.0	1799	16.0
1325	11.0	1420	15.0	1515	15.0	1610	14.0	1800	15.0
1326	16.0	1421	17.0	1516	19.0	1611	13.0	1801	11.0

1327	15.0	1422	10.0	1517	17.0	1612	15.0	1802	12.0
1328	13.0	1423	13.0	1518	16.0	1613	13.0	1803	14.0
1329	15.0	1424	15.0	1519	15.0	1614	14.0	1804	14.0
1330	13.0	1425	16.0	1520	19.0	1615	17.0	1805	16.0
1331	19.0	1426	16.0	1521	17.0	1616	17.0	1806	20.0
1332	13.0	1427	14.0	1522	9.0	1617	17.0	1807	18.0
1333	15.0	1428	16.0	1523	9.0	1618	13.0	1808	13.0
1334	19.0	1429	16.0	1524	16.0	1619	11.0	1809	16.0
1335	16.0	1430	20.0	1525	18.0	1620	11.0	1810	20.0
1336	13.0	1431	15.0	1526	16.0	1621	13.0	1811	14.0
1337	8.0	1432	17.0	1527	15.0	1622	11.0	1812	11.0
1338	13.0	1433	18.0	1528	14.0	1623	12.0	1813	14.0
1339	10.0	1434	11.0	1529	16.0	1624	10.0	1814	14.0
1340	13.0	1435	13.0	1530	16.0	1625	15.0	1815	21.0
1341	18.0	1436	6.0	1531	11.0	1626	9.0	1816	16.0
1342	13.0	1437	18.0	1532	15.0	1627	19.0	1817	11.0
1343	15.0	1438	17.0	1533	13.0	1628	13.0	1818	16.0
1344	18.0	1439	16.0	1534	19.0	1629	12.0	1819	14.0
1345	18.0	1440	12.0	1535	14.0	1630	10.0	1820	17.0
1346	16.0	1441	14.0	1536	12.0	1631	17.0	1821	14.0
1347	9.0	1442	14.0	1537	9.0	1632	14.0	1822	11.0
1348	13.0	1443	11.0	1538	12.0	1633	16.0	1823	16.0
1349	14.0	1444	16.0	1539	12.0	1634	12.0	1824	14.0
1350	16.0	1445	12.0	1540	14.0	1635	9.0	1825	15.0
1351	12.0	1446	13.0	1541	22.0	1636	12.0	1826	13.0
1352	9.0	1447	10.0	1542	17.0	1637	15.0	1827	15.0
1353	13.0	1448	14.0	1543	12.0	1638	13.0	1828	16.0
1354	16.0	1449	15.0	1544	18.0	1639	14.0	1829	16.0
1355	10.0	1450	11.0	1545	13.0	1640	16.0	1830	12.0
1356	16.0	1451	19.0	1546	11.0	1641	10.0	1831	13.0
1357	20.0	1452	17.0	1547	16.0	1642	16.0	1832	13.0
1358	15.0	1453	16.0	1548	13.0	1643	14.0	1833	12.0
1359	12.0	1454	15.0	1549	14.0	1644	17.0	1834	11.0
1360	16.0	1455	20.0	1550	8.0	1645	20.0	1835	15.0
1361	11.0	1456	16.0	1551	10.0	1646	14.0	1836	16.0
1362	12.0	1457	14.0	1552	13.0	1647	10.0	1837	16.0
1363	14.0	1458	12.0	1553	10.0	1648	11.0	1838	16.0
1364	13.0	1459	16.0	1554	20.0	1649	13.0	1839	10.0
1365	9.0	1460	14.0	1555	16.0	1650	14.0	1840	14.0
1366	14.0	1461	15.0	1556	11.0	1651	10.0	1841	14.0
1367	16.0	1462	15.0	1557	16.0	1652	10.0	1842	15.0
1368	15.0	1463	18.0	1558	10.0	1653	17.0	1843	13.0
1369	19.0	1464	11.0	1559	18.0	1654	16.0	1844	12.0

1370	17.0	1465	14.0	1560	12.0	1655	18.0	1845	15.0
1371	12.0	1466	13.0	1561	14.0	1656	13.0	1846	18.0
1372	14.0	1467	11.0	1562	11.0	1657	15.0	1847	9.0
1373	17.0	1468	16.0	1563	14.0	1658	16.0	1848	19.0
1374	11.0	1469	12.0	1564	18.0	1659	5.0	1849	11.0
1375	16.0	1470	13.0	1565	10.0	1660	21.0	1850	15.0
1376	15.0	1471	10.0	1566	18.0	1661	14.0	1851	20.0
1377	15.0	1472	11.0	1567	10.0	1662	16.0	1852	13.0
1378	16.0	1473	15.0	1568	15.0	1663	14.0	1853	18.0
1379	13.0	1474	13.0	1569	18.0	1664	14.0	1854	18.0
1380	8.0	1475	14.0	1570	18.0	1665	20.0	1855	18.0
1381	15.0	1476	7.0	1571	11.0	1666	11.0	1856	16.0
1382	23.0	1477	11.0	1572	6.0	1667	12.0	1857	17.0
1383	17.0	1478	16.0	1573	10.0	1668	16.0	1858	10.0
1384	10.0	1479	15.0	1574	15.0	1669	17.0	1859	15.0
1385	10.0	1480	17.0	1575	11.0	1670	17.0	1860	14.0

16) *The use of normalised cyclone frequencies is advocated to account for offsets or biases in Fig 2a. Hence, a comparison of Fig 6 and Fig 2a would be useful to see if there are any biases when using the starting years of the RCP scenario. An extra column in Fig 6, showing the normalized cyclones/summer for the base 2006-2025 period would achieve this.*

This is an excellent point. We have now included an additional column in Fig. 6 (left column) that shows the normalized cyclones/summer for the base period (2006-2025). This column can now be used to compare the starting year for the three different scenarios (RCP 4.5, RCP 4.5*, and RCP 8.5). The differences between these normalized cyclone frequencies are small, so we conclude that there are not substantial spatial differences between the three difference scenarios. We also added a brief discussion of this in the text

Line 265: "Comparison of normalized cyclone frequencies for the base period in the three scenarios (RCP 8.5, RCP 4.5, and RCP 4.5*) does not show any major discrepancies. We conclude that the starting conditions do not impact the resulting cyclone distribution in the future period."

17) *The smaller variability in the future period in Fig. 7 is an interesting result. Has this been reported elsewhere? It may show up more clearly in the standard deviation than in the RSD (which only shows differences of ~2%).*

We did not find this reported elsewhere and appreciate the reviewer's suggestion to use the standard deviation instead of RSD. The standard deviation does show a more pronounced decrease but it is important to consider the standard deviation in the context of the mean value. This is particularly important for a quantity like cyclone frequencies because they are a bounded quantity (cannot be negative). So we used the RSD in this work to minimize this impact on our presentation of the results.

18) *Fig 1: (21686: line 15): the reductions of 30 and 40 ppb are hard to ascertain from Fig 1. First, the units on the scale (increments of "14") could be modified to make for an easier comparison with the text. It might be useful to highlight the points on the track that are observed for each day (e.g. July 26 and 27 show identical northern storm tracks). As a minor point - give the year of July storm event in Fig 1 caption and body text.*

We appreciate the reviewer's suggestion for this figure and have added a colored dot, marking the current position of the storm, for clarity on Fig. 1. We have also clarified that this is a GCM so not representative of any particular year.

Specific comments:

21688: line 27: The results from the control simulations may differ depending on period

This is true, however as we state in the text, "Only the 1761–1860 time period shows a statistically significant trend ($p < 0.10$), however this is not surprising as a normally distributed dataset would be expected to return one significant trend at the 10 % significance level given 10 samplings." The findings shown here fall in line with what would be expected of a dataset with no trend and provide a testbed to analyze the natural variability. Additionally, the 100-year trends in the RCP scenarios are all larger than those found here in the control period, giving confidence that the trends in the future period are forced by anthropogenic warming.

21685: line 10: Why is the "maximum travel distance" needed if you impose a maximum speed criteria? The use of the parabolic fit is similar to that used by Murray and Simmonds (1991) and related schemes. Murray, R. J., and I. Simmonds, 1991: A numerical scheme for tracking cyclone centres from digital data. Part I: Development and operation of the scheme. Aust. Meteorol. Mag., 39, 155-166.

As you say, the maximum travel distance is set by the maximum speed criteria. This value is important in the tracking scheme as a way to limiting the search for candidate centers and is also used to normalize selection criteria by closeness. For example, the tracking algorithm favors tracks without large changes in direction from center to center. However, it seems reasonable that a large apparent change of direction between two closely located centers (i.e., a near stationary period in the track) is not really the same thing as the same change in direction but also between two centers relatively far apart (i.e., possibly a unrelated cyclone or a secondary system). The parabolic fit is only used to refine the location of the center within the grid cell it was located within. In this sense it is only used superficially unlike Murray and Simmonds (1991), who use the fit as a way of detecting a center rather than just to fine-tune its location.

All of the other reviewer's specific comments have been addressed in the manuscript.

21680: Line 6: expand “yr” to “years” in all places.
21681: line 21: “Mid-latitude cyclones ..synoptic and climatic scales .. regional scale.” This sentence could be written more clearly.
21683: very briefly outline how H1 to H5 differ- e.g. initial conditions
21683: line 25: The temperature values given depends on the GCM. Make sure this is noted e.g. ”corresponds to an .. warming of 4.5K in the GFDL CM3 GCM” .
21684: line 16: mention some of the well-known storm tracks algorithms (often used operationally): e.g.: Serreze et al. http://www.esrl.noaa.gov/psd/map/docs/ stormtracks.maproom.html Hodges et al. A Comparison of Extratropical Cyclones in Recent Reanalyses ERA-Interim, NASA MERRA, NCEP CFSR, and JRA-25 , 2011, K. I. Hodges, R. W. Lee and L. Bengtsson, J. Clim. , 24, 4888-4906
21685: line 15: explain why the southern storm track is “southern”.
21685: line 18 rephrase “as major storm track”.
21685: line 22 defining GLST here is rather repetitive (c.f. line 24 on page 21682).
21687: line 6: the “relative standard deviation” is commonly referred to as the coefficient of variation.
21687: line 10: Explain what “following Liebensperger et al.” means in this context.
21687: line 17: clarify what “the GFDL CM3 model cyclone frequency is within 10% throughout ” means, make sure the scale in Fig 2c is somewhat compatible in its labelling.
21687: line 25: explain the trend fitting procedure -“p-value of a trend”.
21688: line 2: re-word “interannual correlation coefficient” (ditto Fig 8. caption)
21688: line 24: Control or unforced simulations are commonly performed to quantify internal model variability within a model hence the reference to ENSO in unnecessary.