CMIP6 captures the satellite-era jet slowdown and Arctic amplification - yet projects future jet speedup and tropical amplification

Rachel M. Robinson  
University of North Carolina at Charlotte

Jacob Scheff (jscheff@uncc.edu)  
University of North Carolina at Charlotte

Nicholas Golden  
University of North Carolina at Charlotte

Research Article

Keywords: climate change, jet stream, troposphere, temperature gradient, Arctic amplification

Posted Date: January 6th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2432307/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

The polar-to-subtropical temperature gradient in the free troposphere is a key driver of the mid-latitude jet stream response to climate change. Climate models tend to steepen this gradient in response to large greenhouse gas increases, due to very strong subtropical upper-level warming. This strengthens the simulated jets. However, multiple lines of observational evidence point to a slowing northern jet over the satellite era, driven by enhanced Arctic free-tropospheric warming and weakening of the gradient. Here, we reconcile these seemingly contradictory results by showing that the CMIP6 ensemble successfully simulates both the observed satellite-era slowdown/weakening, and the speedup/strengthening with strong global warming. Specifically, the observed gradient weakening from 1980–1997 to 1997–2014 appears inconsistent (p < 0.05) with the simulated gradient changes for just 6 of 45 models using Microwave Satellite Unit observations, and for just 5 of 45 models using reanalysis estimates. The observed jet slowdown appears inconsistent with the simulated jet changes for just 1 of 45 models. Yet a strong majority of the models reverse course under a high-emissions future-type scenario, simulating gradient strengthening and jet speedup. Future work will seek to clarify the cause(s) of this unexpected difference between past and future atmospheric responses.

1. Introduction

The Northern Hemisphere mid-latitude jet is driven in large part by the equator-pole temperature gradient. The strength and location of the jet are anticipated to change in a warmed climate, but how these changes will manifest is not well defined (Francis and Vavrus 2015; Francis 2017; Matsumura et al. 2019). Two responses of the jet to climate change are often found in the literature. One is a poleward shift of the jet and mid-latitude storm track, often associated with an increase in zonal wind speed, a less wavy jet pattern, and potentially less persistent weather patterns, although this is contested (Barnes et al. 2014; Butler et al. 2010; Ceppi and Hartmann 2016; Deser et al. 2015; Francis and Vavrus 2012; Yim et al. 2016; Shaw and Tan 2018; Peings et al. 2018). These features are also signatures of a positive North Atlantic Oscillation (NAO)/Arctic Oscillation (AO) (Feldstein and Lee 2014). In contrast, other studies show an equatorward shift of the jet and mid-latitude storm track, associated with a decrease in zonal wind speed, a wavier jet pattern, and possibly more persistent weather patterns (Barnes et al. 2014; Barnes and Polvani 2015; Butler et al. 2010; Ceppi and Hartmann 2016; Deser et al. 2015; Francis and Vavrus 2012; Yim et al. 2016; Shaw and Tan 2018; Peings et al. 2018). These are also features of a negative NAO/AO (Feldstein and Lee 2014).

The poleward shift and strengthening of the mid-latitude jet have been attributed to increased warming aloft in the tropical upper troposphere (Butler et al. 2010; Ceppi and Hartmann 2016; Deser et al. 2015; Yim et al. 2016; Shaw and Tan 2018; Peings et al. 2018), though Shaw and Tan (2018) found the subtropical upper troposphere plays a larger role than the tropics. Lower-stratospheric cooling in the Arctic has also been shown to cause a poleward shift of the jet (Butler et al. 2010; Deser et al. 2015; Yim et al. 2016; Shaw and Tan 2018; Peings et al. 2018). Conversely, the equatorward shift and weakening of the mid-latitude jet has been attributed to enhanced warming in the Arctic lower troposphere, known as
Arctic Amplification (Butler et al. 2010; Barnes and Polvani 2015; Deser et al. 2015; Ceppi and Hartmann 2016; Yim et al. 2016; Shaw and Tan 2018; Peings et al. 2018).

The balance between the warming of the tropical upper troposphere and the Arctic lower troposphere is the premise of the "tug-of-war" hypothesis: whichever one warms more in the future should determine the future jet response. Since both jet responses have opposing effects on weather, accurately projecting the future jet change in climate models is important for our understanding of the effects of climate change on mid-latitude weather.

However, there is evidence that the behavior of the jet in climate models is at odds with observations over the last several decades. Most climate models project a poleward jet shift and/or the tropics warming more (Yin 2005; Lorenz and DeWeaver 2007; Butler et al. 2010; Barnes and Polvani 2013; Delcambre et al. 2013; Lorenz 2014; Ceppi and Hartmann 2016; Lachmy and Shaw 2018; Screen et al. 2018; Li et al. 2019; Harvey et al. 2020), but studies based on observations and reanalysis of the satellite era indicate that the Arctic may be warming more than the tropics and/or that the jet has shifted equatorward (Graversen et al. 2008; Santer et al. 2013; Feldstein and Lee 2014; Francis and Vavrus 2015; Manney and Hegglin 2018; Santer et al. 2018).

Additionally, the discussion surrounding the competition between the Arctic and the tropics has focused specifically on the warming of the Arctic near the surface and the tropics in the upper troposphere. The Arctic upper troposphere is typically not considered, as most of the warming in the Arctic occurs near the surface (Lorenz 2014; Barnes and Screen 2015; Matsumura et al. 2019). However, some other reanalysis and observationally based studies indicate that the Arctic upper troposphere is also warming significantly, and that models are overestimating tropical warming and underestimating Arctic warming in the mid to upper troposphere (Graversen et al. 2008; Santer et al. 2013; Santer et al. 2018). Kim et al. (2021) showed that driving a GCM with Arctic free-tropospheric warming more effectively produces a jet shift than driving with surface-based Arctic warming.

While the multi-model means of the climate model-based studies tend to produce a poleward shift and strengthening of the jet, individual models have been shown to produce a wide variety of jet response patterns (Yim et al. 2016). Several studies have utilized the tug-of-war framework to understand differences in the jet response between different climate models and found a significant correlation between the Arctic-tropical warming contrast and changes in the jet (Barnes and Polvani 2015; Yim et al. 2016; Peings et al. 2018). Peings et al. (2018) found a strong correlation of 0.84 between the change in jet strength and the "ratio between upper-troposphere tropical and Arctic warming (RUTAW)" that is consistent with the tug-of-war hypothesis. The correlation between jet strength and RUTAW offers the benefit of being able to qualitatively predict the jet response in climate models.

Given the apparent disagreement between climate models and observations, as well as the disagreement between individual models, there is significant motivation to study the relationship between the Arctic-tropical warming contrast and the mid-latitude jet. Most studies have used older versions of CMIP (the Coupled Model Intercomparison Project). The most recent iteration, CMIP6 (Eyring et al. 2016), offers the
opportunity to use a suite of state-of-the-art climate models. This new set of models will help us answer a variety of questions. Does the significant relationship between Arctic and subtropical warming and zonal wind found by prior studies hold in CMIP6 future and historical climate projections, and if so, is it stronger or weaker? Are climate models accurately capturing the satellite era mid-latitude jet change and warming pattern? Finally, given the utility of the correlation between the Arctic minus subtropical warming and zonal wind change as a predictor of the future jet response, can models be “ruled out” by comparing their historical simulations to observations?

2. Data

a) High-Emission Future-Like Simulations

To represent the future under a high emissions scenario, 45 climate model simulations from CMIP6 were used. However, the idealized 1pctCO2 experiment was chosen instead of one of the more "realistic" SSP experiments, because relatively few modeling groups had made their SSP simulations available at the time of this study, while many modeling groups had made their 1pctCO2 simulations available (Table 1). While the SSP simulations have more realistic forcing, it is still mostly CO2 forcing (especially for the high-emission SSP5-8.5), so the 1pctCO2 simulations are a reasonable substitute. By the end of the 150-year 1pctCO2 simulations, CO2 concentration has more than quadrupled from 284 ppm to about 1263 ppm, resulting in an average global temperature that is several degrees Celsius warmer than today's average, analogous to a high-emission future. Only the first ensemble member of each model is used for the 1pctCO2 portion of the study because the other ensemble members behave nearly identically to the first in CMIP5 rcp8.5 (Golden 2020), and a cursory check of a few models confirmed this for CMIP6 1pctCO2 as well. This is expected due to the very high signal-to-noise ratio for this experiment.

Table 1 A list of all CMIP6 1pctCO2, historical, and synthetic MSU models used. The last three columns give the p-values from the two-sample t-tests for observational consistency with the model ensemble, described in section 3c. Red text highlights p-values less than 0.05
The variables analyzed in this study are the monthly air temperature (\(ta\)) and zonal wind (\(ua\)) on pressure levels, and a control and future-like period are identified for each variable. The control period is defined as model years 1 through 30 (average CO2 ~ 330 ppm) for each variable. The future-like period is defined as model years 121 through 150 (average CO2 ~ 1090 ppm) for each variable (Eyring et al. 2016). Note that these do not correspond to actual calendar years.

### b) Satellite Era Simulations

<table>
<thead>
<tr>
<th>Model Number</th>
<th>Model Name</th>
<th># of Historical Ensemble Members</th>
<th># of MSU Ensemble Members</th>
<th>Reanalysis Warming</th>
<th>Reanalysis Wind</th>
<th>MSU Warming</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACCESS-CM2</td>
<td>3</td>
<td>3</td>
<td>0.2043</td>
<td>0.1351</td>
<td>0.1043</td>
</tr>
<tr>
<td>2</td>
<td>ACCESS-ESM1-5</td>
<td>30</td>
<td>20</td>
<td>0.3994</td>
<td>0.8970</td>
<td>0.3678</td>
</tr>
<tr>
<td>3</td>
<td>AWI-CM-1-1-MR</td>
<td>5</td>
<td>5</td>
<td>0.1506</td>
<td>0.1478</td>
<td>0.1815</td>
</tr>
<tr>
<td>4</td>
<td>BCC-CSM2-MR</td>
<td>3</td>
<td>3</td>
<td>0.2712</td>
<td>0.2035</td>
<td><strong>0.0380</strong></td>
</tr>
<tr>
<td>5</td>
<td>BCC-ESM1</td>
<td>3</td>
<td>3</td>
<td>0.2559</td>
<td>0.4808</td>
<td>0.2035</td>
</tr>
<tr>
<td>6</td>
<td>CAM5-CSM1-0</td>
<td>2</td>
<td>2</td>
<td>0.2574</td>
<td>0.1625</td>
<td>0.0994</td>
</tr>
<tr>
<td>7</td>
<td>CESM2</td>
<td>11</td>
<td>11</td>
<td>0.3112</td>
<td>0.8517</td>
<td>0.1580</td>
</tr>
<tr>
<td>8</td>
<td>CESM2-FV2</td>
<td>3</td>
<td>3</td>
<td>0.6652</td>
<td>0.6629</td>
<td>0.1089</td>
</tr>
<tr>
<td>9</td>
<td>CESM2-WACCM</td>
<td>3</td>
<td>3</td>
<td>0.4332</td>
<td>0.1340</td>
<td>0.4415</td>
</tr>
<tr>
<td>10</td>
<td>CESM2-WACCM-FV2</td>
<td>3</td>
<td>2</td>
<td>0.5727</td>
<td>0.9116</td>
<td>0.3979</td>
</tr>
<tr>
<td>11</td>
<td>CIESM</td>
<td>3</td>
<td>3</td>
<td>0.0549</td>
<td><strong>0.0070</strong></td>
<td>0.0529</td>
</tr>
<tr>
<td>12</td>
<td>CNRM-CM6-1</td>
<td>29</td>
<td>29</td>
<td>0.1092</td>
<td>0.3131</td>
<td>0.0796</td>
</tr>
<tr>
<td>13</td>
<td>CNRM-CM6-1-HR</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>14</td>
<td>CNRM-ESM2-1</td>
<td>10</td>
<td>8</td>
<td>0.1367</td>
<td>0.3011</td>
<td>0.0544</td>
</tr>
<tr>
<td>15</td>
<td>CanESM5</td>
<td>40</td>
<td>40</td>
<td>0.4583</td>
<td>0.8598</td>
<td>0.4383</td>
</tr>
<tr>
<td>16</td>
<td>CanESM5-CanOE</td>
<td>3</td>
<td>3</td>
<td>0.0892</td>
<td>0.8714</td>
<td>0.3582</td>
</tr>
<tr>
<td>17</td>
<td>EC-Earth3</td>
<td>18</td>
<td>6</td>
<td>0.3586</td>
<td>0.6552</td>
<td>0.4459</td>
</tr>
<tr>
<td>18</td>
<td>EC-Earth3-Veg</td>
<td>9</td>
<td>5</td>
<td>0.0672</td>
<td>0.3369</td>
<td><strong>0.0392</strong></td>
</tr>
<tr>
<td>19</td>
<td>FG4ALS-F4-L</td>
<td>3</td>
<td>3</td>
<td>0.0942</td>
<td>0.9715</td>
<td>0.0874</td>
</tr>
<tr>
<td>20</td>
<td>FG4ALS-g3</td>
<td>6</td>
<td>3</td>
<td>0.2309</td>
<td>0.3536</td>
<td>0.1159</td>
</tr>
<tr>
<td>21</td>
<td>FIO-ESM2-0</td>
<td>3</td>
<td>3</td>
<td>0.1193</td>
<td>0.1281</td>
<td>0.1479</td>
</tr>
<tr>
<td>22</td>
<td>GFDL-CM4</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>23</td>
<td>GFDL-ESM4</td>
<td>3</td>
<td>3</td>
<td>0.0936</td>
<td>0.2927</td>
<td><strong>0.0276</strong></td>
</tr>
<tr>
<td>24</td>
<td>GISS-E2-1-G-p1</td>
<td>28</td>
<td>28</td>
<td>0.1598</td>
<td>0.5865</td>
<td>0.0990</td>
</tr>
<tr>
<td>25</td>
<td>GISS-E2-1-G-p3</td>
<td>9</td>
<td>9</td>
<td>0.1203</td>
<td>0.1712</td>
<td>0.0733</td>
</tr>
<tr>
<td>26</td>
<td>GISS-E2-1-G-p5</td>
<td>9</td>
<td>9</td>
<td>0.0473</td>
<td>0.0727</td>
<td><strong>0.0336</strong></td>
</tr>
<tr>
<td>27</td>
<td>GISS-E2-1-H</td>
<td>15</td>
<td>15</td>
<td>0.5056</td>
<td>0.6126</td>
<td>0.2675</td>
</tr>
<tr>
<td>28</td>
<td>HadGEM3-GC31-LL</td>
<td>5</td>
<td>4</td>
<td>0.9179</td>
<td>0.9825</td>
<td>0.8188</td>
</tr>
<tr>
<td>29</td>
<td>HadGEM3-GC31-MM</td>
<td>4</td>
<td>3</td>
<td><strong>0.0141</strong></td>
<td>0.2622</td>
<td><strong>0.0214</strong></td>
</tr>
<tr>
<td>30</td>
<td>IITM-ESM</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>31</td>
<td>INM-CM4-8</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>32</td>
<td>IPSL-CM5A-LR</td>
<td>32</td>
<td>32</td>
<td>0.1241</td>
<td>0.7405</td>
<td>0.1846</td>
</tr>
<tr>
<td>33</td>
<td>KACE-1-0-G</td>
<td>3</td>
<td>3</td>
<td>0.3402</td>
<td>0.5855</td>
<td>0.3779</td>
</tr>
<tr>
<td>34</td>
<td>MIROC5</td>
<td>50</td>
<td>50</td>
<td>0.1618</td>
<td>0.7681</td>
<td>0.0826</td>
</tr>
<tr>
<td>35</td>
<td>MIROC-ES2L</td>
<td>30</td>
<td>10</td>
<td>0.1390</td>
<td>0.9882</td>
<td>0.1840</td>
</tr>
<tr>
<td>36</td>
<td>MPI-ESM1-2-HR</td>
<td>10</td>
<td>10</td>
<td>0.1180</td>
<td>0.3849</td>
<td>0.0668</td>
</tr>
<tr>
<td>37</td>
<td>MPI-ESM1-2-LR</td>
<td>10</td>
<td>10</td>
<td><strong>0.0442</strong></td>
<td>0.1340</td>
<td><strong>0.0219</strong></td>
</tr>
<tr>
<td>38</td>
<td>MRI-EM52-O</td>
<td>5</td>
<td>5</td>
<td>0.2476</td>
<td>0.6962</td>
<td>0.0571</td>
</tr>
<tr>
<td>39</td>
<td>NESM3</td>
<td>5</td>
<td>5</td>
<td>0.4315</td>
<td>0.9204</td>
<td>0.2252</td>
</tr>
<tr>
<td>40</td>
<td>NorCPM1</td>
<td>30</td>
<td>30</td>
<td><strong>0.0181</strong></td>
<td>0.6893</td>
<td>0.0607</td>
</tr>
<tr>
<td>41</td>
<td>NorESM2-LM</td>
<td>3</td>
<td>3</td>
<td>0.4296</td>
<td>0.5569</td>
<td>0.3490</td>
</tr>
<tr>
<td>42</td>
<td>NorESM2-MM</td>
<td>2</td>
<td>2</td>
<td>0.1195</td>
<td>0.2003</td>
<td>0.2003</td>
</tr>
<tr>
<td>43</td>
<td>SAMO-UNICON</td>
<td>1</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>44</td>
<td>TaiESM1</td>
<td>2</td>
<td>1</td>
<td>0.6804</td>
<td>0.6925</td>
<td>N/A</td>
</tr>
<tr>
<td>45</td>
<td>UKESM1-0-LL</td>
<td>19</td>
<td>18</td>
<td>0.8398</td>
<td>0.8497</td>
<td>0.4543</td>
</tr>
</tbody>
</table>

The variables analyzed in this study are the monthly air temperature (\(ta\)) and zonal wind (\(ua\)) on pressure levels, and a control and future-like period are identified for each variable. The control period is defined as model years 1 through 30 (average CO2 ~ 330 ppm) for each variable. The future-like period is defined as model years 121 through 150 (average CO2 ~ 1090 ppm) for each variable (Eyring et al. 2016). Note that these do not correspond to actual calendar years.
Simulations of the satellite era from 1980–2014 are provided by the CMIP6 historical simulations from the same 45 models used above. While SSP output could be used to extend this up to the latest availability of observational data (e.g. 2021), there were few SSP models uploaded at the time of this study. Unlike the 1pctCO2 simulations, multiple ensemble members are included where available to help account for internal variability, due to the smaller signal-to-noise ratio over the satellite era. There are 469 total ensemble members across 45 models included in this part of the study (Table 1). Again, the variables analyzed are the monthly temperature and zonal wind, which are halved into two 18-year segments to obtain past (1980–1997) and present (1997–2014) periods. The start year is 1980 rather than the usual 1979 simply because one of the reanalysis datasets (below; Table 2) only begins in 1980. It should be noted that due to the odd number of years there is one year of overlap between the two time periods.

Table 2
A list of reanalysis products used

<table>
<thead>
<tr>
<th>Reanalysis</th>
<th>Years Available</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climate Forecast System Reanalysis (CFSR)</td>
<td>1979- Present</td>
<td>Saha et al. (2014)</td>
</tr>
<tr>
<td>ECMWF Reanalysis Version 5 (ERA5)</td>
<td>1979- Present</td>
<td>Hersbach et al. (2020)</td>
</tr>
<tr>
<td>ECMWF Interim Reanalysis (ERA Interim)</td>
<td>1979- Present</td>
<td>Berrisford et al. (2011)</td>
</tr>
</tbody>
</table>

c) Satellite Observations and Reanalyses

For the observational constraint, we use temperature and zonal wind output from five modern global reanalyses (Table 2), as well as satellite Microwave Sounding Unit (MSU) data on free tropospheric temperature. The analysis of MSU data requires two datasets for a direct comparison to CMIP6. The first is the actual MSU data, which consists of tropospheric temperature observations for 1979–2019 originating from over a dozen satellites that have been spliced together and interpreted by three different groups (Table 3). The second dataset is synthetic MSU data for the same time frame, developed by Po-Chedley et al. (2021), showing what a satellite would have observed in the atmosphere simulated by a CMIP6 climate model. There is synthetic MSU data for 44 of the 45 models (Table 1). For comparison to the CMIP6 analyses, only 1980–2014 will be considered.
Table 3
A list of the groups that interpreted the satellite MSU observations

<table>
<thead>
<tr>
<th>Group</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>The National Oceanic and Atmospheric Administration (NOAA) Center for Satellite Applications and Research</td>
<td>Spencer et al. (2017)</td>
</tr>
<tr>
<td>The University of Alabama at Huntsville</td>
<td>Mears and Wentz (2016)</td>
</tr>
</tbody>
</table>

Preliminary work by Golden (2020) determined that the warming throughout the depth of the free troposphere in both the Arctic and the subtropics is significant for the jet response (see methods section), so an MSU channel called the stratosphere-corrected TMT (temperature of the mid-troposphere), also known as “T24” or “TTT”, was chosen. The uncorrected TMT includes some influence from stratospheric cooling, which Fu et al. (2011) correct using a weighting function that removes stratospheric contamination. The version of the weighting function used for this study, tf1, employs the same weighting globally to avoid sharp cutoffs around 30°N.

3. Methods

a) High-Emission Future-Like Simulations

Zonal-mean, annual-mean temperature and wind change fields are computed for each CMIP6 model in Table 1 by differencing atmospheric temperature and wind between the control (years 1–30) and future-like (years 121–150) periods.

To understand the relationship between these variables, the extent of the tropics, Arctic, and mid-latitude jet must also be defined. A latitude-pressure box is defined for each of these three regions, chosen based on sensitivity tests conducted for CMIP5 in Golden (2020): the box defining the mid-latitude jet extends from 30°N to 70°N and 1000mb to 200mb, the box defining the Arctic extends from 60°N to 90°N and 850mb to 300mb, and the box defining the tropics extends from 20°N to 40°N and 850mb to 200mb (Fig. 1). Golden (2020) found that extending the Arctic box into the upper troposphere and the tropical box into the lower troposphere yields stronger correlations than using just the tropical upper troposphere or the Arctic lower troposphere. The chosen latitudinal extents also showed significantly better correlations than other tested combinations, except in the case of the 60°N to 90°N Arctic latitude bounds which were untested. The 20°N to 40°N extent for the tropics is also supported by Shaw and Tan (2018), which found that the jet is more sensitive to subtropical warming than tropical. Additionally, the choice to extend the Arctic box into the upper troposphere is supported by Kim et al. (2021), which showed that Arctic free-tropospheric warming more effectively produces a jet shift in modeling experiments compared to the Arctic near-surface alone.
The average is taken for each of the 3 boxes defined above to get the subtropical warming, the Arctic warming, and the zonal wind change for each model. The subtropical warming mean is subtracted from the Arctic warming mean to get the Arctic-subtropical warming difference. A simple linear regression is performed between the Arctic-subtropical warming difference and zonal wind change to test the tug-of-war hypothesis.

b) Satellite Era Simulations

The above process is repeated for the historical simulations of the 1980–2014 satellite era. The only differences are the use of the above-defined 18-year segments for the past and present instead of the 30-year segments used in the future period, and the inclusion of 469 ensemble members for the 45 models. The regression is initially performed using only the first ensemble member of each model. Using the first ensemble member provides a more direct comparison to the 1pctCO2 regression to examine how the strength of the relationship differs between the simulations. However, this is not representative of the full distribution for the satellite era, and the regression is repeated with all 469 ensemble members for comparison with the observations. The ensemble-mean warming difference for the satellite era for each model is also plotted against the 1pctCO2 warming difference, and the satellite era ensemble-mean zonal wind change is plotted against the 1pctCO2 zonal wind change. This allows for a comparison of how the models are portraying the temperature difference and the zonal wind change over the satellite era versus the future.

c) Observations – Reanalysis and Satellite MSU Data

The reanalyses are averaged over the same boxes as the climate models, then the reanalysis results are compared to the satellite era climate model results to determine which models are accurately capturing the patterns estimated over the satellite era.

In addition to reanalysis, satellite MSU data is also analyzed, though slightly differently. For both the observational MSU data and the synthetic climate model MSU data, the average warming between the first 18 years and the last 18 years is calculated. However, instead of using the full defined latitude-pressure boxes, the defined latitudes are used in tandem with the above-discussed stratosphere-corrected TMT channel to create the Arctic and subtropical domains. This channel corresponds most closely to the pressure levels used for the latitude-pressure boxes in the previous analyses. As before, the subtropical mean warming is subtracted from the Arctic mean warming. The synthetic climate model MSU temperature difference is correlated with the climate model zonal wind change over the satellite era for each model to determine the utility of the synthetic climate model MSU Arctic minus subtropical warming as a predictor of the modeled jet change. As with the reanalysis, the observational MSU Arctic minus subtropical warming is compared to the results of this correlation to determine if the models are successfully capturing the satellite era change.

To quantify which models are possibly consistent with the reanalysis or satellite MSU output, a two-sample t-test is used for both the warming difference and zonal wind change to compare each model to
reality as estimated by reanalysis and/or by the satellite MSU data. The average warming difference is calculated over all 5 reanalysis products in Table 2 and over all 3 MSU observational products in Table 3, resulting in a single estimate of reality for each that is used as one of the samples for the t-test. The collection of warming differences for all of the ensemble members of a given model is used as the other sample. For the satellite MSU, this is the collection of ensemble members of the model’s synthetic MSU output. While the sample based on reality will only have a single value, the sample for the models will have multiple values, depending on the number of ensemble members. This is repeated for the zonal wind change for the reanalysis only. The two-sample t-test evaluates the null hypothesis that the average observed warming difference or zonal wind change could be a reasonable member of each model’s full ensemble population. If the test produces a p-value less than 0.05, the null hypothesis can be rejected, and the model is not consistent with observations. Otherwise, the null hypothesis cannot be rejected and the model is tentatively consistent with observations. The two-sample t-test will not work for models that only have one ensemble member available, and these models are excluded from the analysis (Table 4).

4. Results And Discussion

The CMIP6 1pctCO2 and historical simulations both produced a variety of warming and jet change patterns that are consistent with the tug-of-war hypothesis. For example, the CESM2 historical simulation shows much stronger satellite-era warming in the Arctic than the subtropics (Fig. 2a), and a corresponding decrease in jet strength (Fig. 2b). The historical CIESM shows more warming in the subtropics, and an increase in the jet strength.

The regression between the Arctic-subtropical warming difference and the zonal wind change shows a very strong, statistically significant negative correlation in both the CMIP6 1pctCO2 simulations and historical simulations (Figs. 3a-b and 3d). In other words, when the Arctic warms more than the subtropics, the models tend to produce a decrease in the jet strength and when the subtropics warm more than the Arctic, the models tend to produce an increase in jet strength. This relationship is consistent with the tug-of-war hypothesis, and is much stronger in 1pctCO2 than previous studies have found for future-like simulations (Barnes and Polvani 2015; Yim et al. 2016; Peings et al. 2018). The relationship is also notably strong over the satellite era (though weaker than the relationship for 1pctCO2), which most other studies have not tested.

The reanalysis output and satellite MSU temperature data are also included on the satellite-era scatterplots in Figs. 3b and 3c. The reanalyses all fall in the lower right quadrant of the historical simulation scatterplot (Fig. 3b), showing stronger warming in the Arctic than in the tropics and a decrease in jet strength over the satellite era, consistent with prior literature (Graversen et al. 2008; Santer et al. 2013,2018; Feldstein and Lee 2014; Francis and Vavrus 2015; Manney and Hegglin 2018). The satellite MSU products also fall on the right side of the scatterplot, indicating increased warming in the Arctic (Fig. 3c). Additionally, both the reanalyses and the satellite MSU temperatures fall well within the model range. This was somewhat unexpected given that prior studies indicated the models were not capturing the warming observed over the satellite era (Santer et al. 2013,2018; Jansen et al. 2020). It should be
noted that the CFSR is an outlier compared to the other reanalyses. This is likely due to the transition of CFSR to CFSv2 in 2010, as noted by Wright et al. (2020), which may imply that the CFSR warming value is spurious, and should be excluded from the average in future work.

This ability of so many models to produce increased Arctic warming and a weaker jet over the satellite era is a new feature of CMIP6 compared to CMIP5. Golden (2020) performed a similar analysis of the Arctic-subtropical warming and zonal wind change for RCP8.5 and historical simulations in CMIP5, and found that models favored stronger warming in the subtropics and a stronger jet over both the historical simulations and the future-like simulations. This is noted in Figures S1-S2 in the supplemental information, reproduced from Golden (2020). Both the satellite-era simulations (Figure S1) and the future simulations (Figure S2) are weighted towards the top left quadrant, indicating a preference for the subtropics warming more and a strengthening of the jet, in agreement with other studies of the modeled future (Santer et al. 2013, 2018; Jansen et al. 2020).

The scatterplots of some individual models are shown in Fig. 4. The majority of the models unambiguously produce values that are reasonable based on the observations, e.g., Can-ESM5-p2 and UKESM-0-LL (Fig. 4a and Fig. 4d). Conversely, a handful of models do not produce values consistent with the observations, e.g., CIESM and GFDL-ESM4 (Fig. 4b and Fig. 4e). A few other models are a bit more ambiguous, and it's difficult to determine if they are consistent or not, e.g., FIO-ESM-2-0 and CNRM-ESM2-1 (Fig. 4c and Fig. 4f).

The results of the two-sample t-test confirm that the reanalyses and satellite observations could be plausible members of almost every model's ensemble (Table 1), leaving very few models that are clearly inconsistent with observations of both the warming difference and the zonal wind change. Of the models that fail the t-test on the basis of either the reanalysis warming difference or zonal wind change, or the satellite MSU warming difference, most of them have very few ensemble members (e.g., Fig. 4b and Fig. 4e). Such limited sampling means the relationship is likely not strong enough to definitively say these models could not be consistent with observations. Most of the models with a large number of ensemble members do pass the t-test when compared to the reanalysis, like Fig. 4a, and are clearly consistent with observations. Likewise, most models with a large number of ensemble members also pass the t-test when compared to the MSU, like Fig. 4d. However, similar to the models that failed the t-test, there are also many models that pass the t-test with very few ensemble members (e.g., Fig. 4c). Like those that failed the t-test, the sample size is too limited to say they are clearly consistent with the observations, as the t-test results would imply. Therefore, the models that pass or fail with very few ensemble members are inconclusive, and only the models that pass or fail with larger ensembles are convincingly consistent or inconsistent with the observations. With relatively few models that are convincingly inconsistent with the observations, this increases confidence that the models are better suited to simulating the satellite era than implied by Santer et al. (2013, 2018) and Jansen et al. (2020).

Interestingly, although many models are able to better capture the warming and jet change observed over the satellite era, they are largely failing to capture the full magnitude of the warming in the Arctic. This
can be seen in Figs. 3b and 3c, where the majority of the models fall to the left of the reanalysis or satellite MSU. This can also be seen for individual models. For example, in Fig. 4a and 4d, the reanalysis and satellite observations are weighted towards the right side of the model ensemble cloud. This difference in magnitude between observations and the models could be the result of internal variability (Rantanen et al. 2022), or flaws in the forced response.

Finally, despite the models’ general agreement with observations over the satellite era of stronger Arctic warming compared to subtropical warming, they surprisingly still tend to favor stronger subtropical warming and increased jet strength in the future-like simulations. This discrepancy can be seen in Fig. 3. On the 1pctCO2 scatterplot, a clear majority of the models fall in the upper left quadrant, favoring the subtropics warming more and the jet strength increasing. Far fewer points fall in the bottom right quadrant, favoring the Arctic warming more and the jet strength decreasing. Of those that do fall in the bottom right quadrant, most are clustered near the center, with only a few models showing significantly stronger warming in the Arctic and a more significant jet slowdown. However, on the historical scatterplots the majority of the models fall in the bottom right quadrant, favoring stronger Arctic warming and a decrease in jet strength.

This is shown more clearly in Fig. 5, which shows the ensemble mean historical versus the 1pctCO2 Arctic-subtropical warming difference and zonal wind change for each model. In the warming difference plot (Fig. 5a), the models are significantly weighted to the bottom right quadrant compared to the upper left, indicating that many models which produce increased warming in the Arctic over the satellite era still tend to produce increased warming in the subtropics in a future-like climate. Similarly, the models are significantly weighted to the upper left quadrant in the zonal wind change plot compared to the bottom right quadrant (Fig. 5b), indicating that many models which produce a jet slowdown over the satellite era tend to produce a jet speedup in a future-like climate. This shows that the warming difference in the future is systematically more tropical amplified than over the satellite era, indicating that a future where the tropics warm more and the jet speeds up may be consistent with a past where the Arctic warms more and the jet slows down, at least in these experiments. This difference in the warming and jet response to climate change could be influenced by the current and future state of the meridional temperature gradient, an inversion in the Arctic, and sea ice cover. These factors all significantly impact Arctic Amplification, and are expected to evolve due to climate change in a way that would reduce Arctic Amplification (Previdi et al. 2021). Alternatively, the enhanced Arctic Amplification in the satellite-era simulations could simply stem from their spurious high-latitude biomass-burning aerosol forcing as found by Fasullo et al. (2022), making those simulations “right for the wrong reasons” with regard to Arctic warming. Future work (below) will seek to clarify these competing explanations.

5. Conclusion

This study has shown that there is a strong correlation between the Arctic minus subtropical warming and zonal wind change in both CMIP6 historical and future-like simulations that is consistent with the tug-of-war hypothesis. The modeled future agrees with previous model consensus, tending to warm the
subtropics more and increase jet strength, while the modeled satellite-era is more consistent with recent studies that show the Arctic warming more than the tropics. This indicates that the climate models are capturing the historical Arctic warming better than before, and that the increased Arctic warming over the satellite era may actually be consistent with the future depicted by the models. However, while the models are performing better than earlier generations of models, they are still underestimating the magnitude of the warming in the Arctic over the satellite era. Comparing the modeled satellite era to reanalysis output and satellite MSU observations shows that most models are tentatively consistent with the observations over the satellite era, indicating that overall the models are performing better than expected given previous model output that tended to show the tropics winning the tug-of-war instead of the Arctic over the satellite era (Jansen et al. 2020; Yin 2005; Butler et al. 2010; Delcambre et al. 2013; Lorenz 2014; Lachmy and Shaw 2018; Screen et al. 2018; Santer et al. 2013, 2018; Golden 2020). This performance over the satellite era indicates that the past warming in the Arctic may actually be consistent with the future warming in the tropics and subtropics. It also shows that the satellite era may not necessarily be a straightforward predictor of the future response of the northern hemisphere mid-latitude jet to climate change.

Several key additional steps are planned for subsequent work on this topic. Most immediately, the endpoint of the satellite-era comparison will be extended from 2014 to the present, now that CMIP6 SSP simulations are widely available to continue the historical simulations beyond the year 2014. Similarly, the SSP future (i.e. to 2100) simulations will be examined in addition to the idealized 1pctCO2 simulations, to test whether the future-past discrepancy found here is specific to 1pctCO2, or more generalizable. The analyses will also be repeated for individual seasons and/or longitudinal sectors, to understand more precisely where and when the discrepancies arise. Most importantly, the key causes of the discrepancies will be investigated. Specifically, the potential role of forcing differences between the satellite era and the future, especially arising from the large and partly spurious satellite-era aerosol forcing (Fasullo et al. 2022), will be isolated using the CMIP6 single-forcing experiments (Gillett et al. 2016). The moist energy balance modeling approach of Bonan et al. (2018) will then be used to quantify and probe the results of feedback differences between the two eras. Combined, this work will yield more confident results about the future of the tropospheric temperature gradient and northern mid-latitude jet stream on a warming planet.

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for Publication**

Not applicable.

**Data Availability**
All CMIP5 data is publicly archived and available at https://esgf-node.llnl.gov/search/cmip5/. All CMIP6 data is publicly archived and available at https://esgf-node.llnl.gov/search/cmip6/. All reanalysis data used in this study is archived and available at https://esgf-node.llnl.gov/search/create-ip/. The synthetic MSU data from Po-Chedley et al. (2021) is available at https://pcmdi.llnl.gov/research/DandA/PNAS_2021/index.html.

**Competing Interests**

The authors have no relevant financial or non-financial interests to disclose.

**Funding**

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

**Authors’ Contributions**

Rachel M. Robinson carried out the study, produced the figures, and wrote most of the manuscript. Jacob Scheff conceived of the study and wrote small portions of the manuscript. Nicholas Golden carried out key preliminary work.

**Acknowledgements**

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. We would also like to acknowledge Stephen Po-Chedley of Lawrence Livermore National Laboratory for providing data for tf1 for both the observational MSU temperature and the synthetic MSU temperature.

**References**


Table

Table 4 is not available with this version.

Figures
Figure 1

Visualization of the latitude-pressure boxes used to define (a) the subtropics (green, 20°N-40°N, 850mb-200mb) and the Arctic (black, 60°N-90°N, 850mb-300mb), and (b) the mid-latitude jet (30°N-70°N, 1000mb-200mb)

Figure 2
Warming between the present (1997-2014) and past (1980-1997) period of the historical simulations for two example models (left column) and their corresponding zonal wind change (right column). The black contours are the zonal mean wind for the past period and are shown in 5 m/s intervals.

Figure 3

Scatterplot and regression lines for (a) the first ensemble member of each model for the historical simulation analysis, (b) all ensemble members for the historical simulation analysis with each reanalysis (blue circles) and the reanalysis average (red diamond) overlaid, (c) all ensemble members of synthetic MSU warming difference and CMIP6 historical zonal wind change, with each observed MSU product (blue dotted lines) and the average of the observed MSU temperature data (red dashed line) overlaid, and (d)
the 1pctCO2 simulation analysis. Pearson correlation and p-value as well as the equation for the regression line are included on each plot.

**Figure 4**

Scatterplot of (a,b,c) an individual model's ensemble with the reanalysis average (red triangle) and each reanalysis (red dots) overlaid and (d,e,f) each model's ensemble of synthetic MSU warming difference versus CMIP6 historical zonal wind change, with each observed MSU product (blue dotted lines) and the average of the observed MSU temperature data (red dashed line) overlaid.
Figure 5

Scatterplot of (a) the historical Arctic minus subtropical warming difference versus the 1pctCO2 Arctic minus subtropical warming difference and (b) the historical zonal wind change versus the 1pctCO2 zonal wind change. The ensemble mean is used for each historical model, while only one run is used for each 1pctCO2 model due to the very high signal to noise (see section 2a)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryText1231.docx