Large Uncertainty in Observed Estimates of Tropical Width From the Meridional Stream Function

Daniel Baldassare[1], Thomas Reichler[1], Piret Plink-Björklund[2], and Jacob Slawson[2]

- [1]Department of Atmospheric Sciences, University of Utah, Salt Lake City, UT 84112, USA
- [2]Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO 80401, USA

Correspondence: Daniel Baldassare (daniel.baldassare@utah.edu)

Abstract

Recent Hadley cell expansion rate estimates vary substantially, as a multitude of methods and reanalysis datasets yield conflicting results. Among the many methods of estimating the Hadley cell width, the meridional stream function 500 hPa zero-crossing is the most widely used, as it is directly related to the poleward edge of the Hadley Cell (HC). Other common metrics use atmospheric phenomena associated with the HC as a proxy, for instance the zonal surface wind zero-crossing. As each of these metrics require different reanalysis data, each with varying error, the level of data-driven uncertainty differs between each metric. While previous work has analyzed the statistical and dynamical relationships between metrics, to date no study has quantified and compared the uncertainty due to reanalysis data error in different HC metrics. In this study, we use ERA5 ensemble members, which include small perturbations in atmospheric variables based on the data error, to quantify the uncertainty associated with six commonly used HC metrics as well as the range of their trend estimates. In the Northern Hemisphere, the tropical expansion rate calculated by the stream function is roughly 0.05 degrees per decade, while the Southern Hemisphere rate is 0.2 degrees per decade over the period from 1979-2022. Of the six metrics, only the meridional stream function and precipitation minus evaporation have substantial uncertainties. The stream function errors are large due to uncertainty in the underlying meridional wind data and the

presence of large regions of near-neutral circulation at the poleward edge of the tropics. These errors have decreased in recent decades because of improvements in the assimilated observations. Despite these

30 improvements, metrics which use well-observed and constrained quantities such as the zonal surface wind zero crossing have lower uncertainty, particularly in summer and fall in the Northern Hemisphere.

Plain Text Summary

Using ensemble members from the ERA5 reanalysis, the meridional stream function, the most widely

35 used method for estimating trends in the extent of tropical circulation, was found to have large error, particularly in the Northern Hemisphere and in the summer, because of weak gradients at the tropical edge and poor data quality. Other methods using better observed data were found to have smaller error.

> **Deleted:** tropical width
>
> **Deleted:** the meridional stream function,
>
> **Deleted:** Another method, using the latitude where the surface wind switches from westerly to easterly, was found to have smaller error and is recommended for future studies.

1   Introduction

40 Observed estimates of Hadley Cell (HC) meridional extent change in the last few decades vary substantially in the literature, as a variety of methods and reanalysis datasets, as well as time periods, yield conflicting results (Staten et al., 2018; Xian et al., 2021). HC expansion has been a well-studied phenomenon, as the poleward edge of the HC is associated with decreased precipitation, particularly over the ocean (Schmidt and Grise, 2017). During the 2000s, HC expansion rates were estimated across a wide

45 range of positive values as high as 3 degrees per decade, while recent annual mean trend estimates have decreased to 0.2-0.4 degrees per decade (Grise et al., 2019). The strong expansion estimates from previous studies likely resulted from natural variability and reanalysis data error, indicating the impact of data quality on Hadley cell extent estimates (Staten et al., 2020). Regional variation in tropical expansion has been found to vary substantially by season (Grise et al., 2018) possibly due to natural variability or

50 forcings other than CO_2 (Staten et al., 2019).

Many methods have been created to determine the latitude of the poleward edge of the HC (Chen et al., 2002; Fu et al., 2006; Hudson et al., 2006; Hu and Fu, 2007; Lu et al., 2007; Seidel and Randel, 2007; Previdi and Liepert, 2007; Seidel et al., 2008; Hu et al., 2011; Staten et al., 2011; Zhou et al., 2011; Choi et al., 2014; Karnauskas & Ummenhofer, 2014). Of these methods, the meridional stream function 500

60 hPa zero-crossing (SF) is the most frequently used due to its direct representation of the zonal-mean HC. Other commonly used metrics such as the sea-level pressure maximum (PSL), the zonal surface wind zero-crossing (UAS), the subtropical jet maximum (STJ), the eddy driven jet maximum (EDJ), and the precipitation minus evaporation zero-crossing (P-E) are thought to measure the HC extent more indirectly (Waugh et al., 2018). Previous research has studied the trends associated with the many HC extent

65 metrics as well as the correlation and physical links between metrics in reanalyses and climate simulation (Davis and Birner, 2017; Seviour et al., 2018; Waugh et al., 2018). However, to date no study has analyzed the impact of reanalysis data error on HC width uncertainty. Without consideration of the reliability of the data used for each metric it is not possible to determine the uncertainty in the trends or thoroughly analyze the disagreement between metrics.

70 HC extent is typically studied using zonal mean data, which is then temporally averaged over either a season or a year, removing the impact of longitudinal variation and short-lived storm systems (Staten et al., 2019). Reanalysis data is often used to study observed circulation changes as it uses observational data to estimate historical atmospheric conditions, spans multiple decades, and is spatially and temporally continuous. Climate model data is commonly used as well, though some natural variation is not present,

75 and important forcings such as the Pacific decadal oscillation and aerosols are often not accurately modeled (Allen et al., 2014).

In the present study we use data from ERA5, a modern high-quality, high-resolution reanalysis dataset (Hersbach et al., 2020). The ERA5 reanalysis provides continuous data from 1950, though here we will only use data from 1979 onwards to align with previous studies. ERA5 is one of the only reanalysis

80 products which is nearly mass conserving, avoiding the questionable meridional circulation found in other

85   reanalyses which are not mass-conserving (Davis & Davis, 2018). Since reanalyses combine observations

and modeling to produce estimates of atmospheric variables, observation density and quality impact the

reliability of reanalysis estimates. As a result, the reliability of a particular tropical extent metric depends

partially on the input data error at the relevant location. Moreover, each metric has a specific sensitivity to

the input data error which depends on the meridional gradient in the underlying data near the position

90   where the metric is defined. Weak gradients around the region of interest result in less robust estimates, as

small errors in the underlying data can result in large uncertainties in the position of the tropical edge.

Unlike most other reanalysis products, ERA5 includes nine ensemble members to allow for a

quantification of uncertainty. Comparing the ensemble members to the standard ERA5 product showed

that these two products produce similar tropical extent time series. The ensemble members use the same

95   observations and data assimilation scheme, but with slightly less precision (Hersbach et al., 2020). These

ensemble members are produced by introducing slight perturbations in observations and model

parameters within their respective error ranges (Isaksen et al., 2010), meaning that the ensemble spread

does not represent other sources of error such as structural uncertainty, and is therefore only a portion of

the actual uncertainty (Tebaldi and Knutti, 2007). Because reanalysis products use different observations

100   and model configurations, the inter-reanalysis spread which previous studies have analyzed (Davis and

Rosenlof, 2012) is much larger than the ERA5 ensemble spread and may overstate the uncertainty,

particularly when older reanalyses are included. While the ERA5 ensemble underestimates the

uncertainty, it does allow for a systematic analysis of the uncertainty in a single reanalysis product which

may represent the relative uncertainties in each metric more accurately.

105   To date no study has quantified the impact of the data uncertainty in a single reanalysis product on HC

extent trend estimates, resulting in a lack of information about this source of error in HC extent metrics.

Due to this knowledge gap, previous studies have not been able to consider the reanalysis data error

differences between metrics, or the sensitivity to this error. In the present study, we provide this

information by analyzing the ERA5 ensemble members, as this allows for a systematic analysis of data

110   uncertainty and its impact on HC estimates. In utilizing these ensemble members, we provide estimates of the uncertainty in the reanalysis data used by each metric and the sensitivity of each metric to these data errors. This study further aims to provide a range of observed HC extent trends for the various metrics in each hemisphere and season using the modern ERA5 reanalysis dataset.

The structure of the paper is as follows. In Section 2 we describe the methodology, including the HC

115   extent metrics to be analyzed. In Section 3, we start by measuring the HC trend uncertainty of all metrics in each season. Next, we determine the annual variation in HC extent for all six metrics in each season. Subsequently, we attempt to determine the causes of uncertainty in the SF estimates. Following these results, in Section 4 we provide suggestions for determining the correct HC extent metrics to use in future studies.

120

2   Data and methods

2.1   Data

ERA5 monthly averaged ensemble member data from March 1979 through February 2022 were acquired for use in this study. March is selected as the starting month as meteorological winter spans multiple

125   calendar years, meaning that March 1979 is the start of the first complete season, while February 2022 is chosen as the final month to allow for the maximum number of complete seasons. ERA5 contains 9 ensemble members and one control member. Because the control member is produced through a more thorough data assimilation process, involving slightly more precision, the control was excluded from the ensemble analysis. The ensemble members are created by perturbing the observations and model

130   tendencies. As these ensemble members only represent a portion of the actual uncertainty, compared to the interannual variation and inter-reanalysis variation these differences are small. Because the ensemble spread in ERA5 is smaller than the actual uncertainty, we will primarily focus on the relative uncertainties between the different metrics. Acquired pressure level data includes the meridional and zonal wind fields,

135   while surface level data included 10-m zonal wind, sea-level-pressure, precipitation, and evaporation.

This data is used to compute the tropical extent for the six metrics described in the next sub-section. Data

from these 9 ensemble members are regridded using a first-order conservative regridding from a 0.5-

degree resolution to a 1-degree resolution using the Climate Data Operator (CDO) 'remapcon' function

(Schulzweida, 2022). The 1-degree and 0.5-degree resolution data produced similar tropical extents. All

140   ensemble members are then zonally and seasonally averaged into March-April-May (MAM), June-July-

August (JJA), September-October-November (SON), and December-January-February (DJF). Separately

annual means are calculated for each year from March-February.

2.2   Methods

Six different tropical width estimation methods are used, each covering different atmospheric regions near

145   the HC edge or using different atmospheric variables. Each method is computed using the Python version

of TropD, a module which estimates HC width from zonally averaged atmospheric data, yielding a

latitude estimate for the poleward extent of the tropics in each hemisphere (Adam et al., 2018). PyTropD

uses spline interpolation to produce more latitudinally precise estimates of tropical extent than the 1-

degree resolution. All tropical width metrics use the standard configuration of TropD.

2.2.1   Meridional Stream Function

150   The meridional stream function 500-hPa zero-crossing (SF) is chosen as it is the most commonly used

metric. This method uses the meridional wind to estimate the edge of the zonal-mean meridional

circulation. The stream function values are calculated by integrating the zonal-mean meridional wind at

each latitude from the top of the atmosphere to the level of interest. From these stream function values,

155   the NH (SH) 500 hPa zero-crossing North (South) of the minimum (30 degrees) and South (North) of the

maximum (60 degrees) is determined as the poleward edge of the tropics in each hemisphere.

2.2.2   Subtropical Jet

The subtropical jet adjusted peak between 100 and 400 hPa (STJ) is presented here as it requires upper-troposphere zonal wind, which directly contrasts with the upper-tropospheric meridional wind used by the SF method. Substantial issues exist with the STJ metric, and it has been shown to poorly correlate with other tropical extent metrics (Waugh et al., 2018).

2.2.3 Eddy Driven Jet

The eddy driven jet maximum at 850 hPa (EDJ) uses lower-troposphere zonal wind data and is located substantially poleward of the other metrics, but has been shown to be closely linked to other metrics of tropical expansion (Solomon et al., 2016; Davis and Birner, 2017).

2.2.4 Precipitation Minus Evaporation

The precipitation minus evaporation zero-crossing (P-E) utilizes surface-level precipitation and evaporation, which are both known to be poorly estimated by reanalyses. Furthermore, the P-E metric is not well connected to other expansion metrics (Seviour et al., 2018). Regardless, this metric is still occasionally used in the literature as it attempts to directly measure the most impactful societal effects of tropical widening.

2.2.5 Surface Zonal Wind Zero-Crossing

The surface zonal wind zero-crossing (UAS) uses 10-m zonal wind data and has been shown to be well-correlated with and dynamically linked to the SF metric (Davis and Birner, 2017).

2.2.6 Sea-Level-Pressure Maximum

The subtropical sea-level-pressure maximum (PSL) reflects the surface pressure anomaly generated at the descending branch of the HC. This metric was found to be well correlated to SF, moderately correlated to P-E, and very highly correlated to UAS (Waugh et al., 2018).

2.2.7 Additional Methods

180   Notably absent from these metrics are the tropopause break height, which is avoided because it is closely related to the subtropical jet (Davis and Birner, 2017), and outgoing longwave radiation-based metrics, which are avoided due to infrequent use in recent studies. As a result of these decisions, SF, STJ, EDJ, P-E, UAS, and PSL will be analyzed in subsequent sections, with a particular emphasis on SF as it is the most frequently used metric.

185   To quantify differences between ensemble members multiple statistical methods are used. The ensemble standard deviation (STD) is calculated by taking the STD of the nine ERA5 ensemble members for a given quantity. In Section 3.3, a normalized ensemble STD is calculated by normalizing this quantity by the interannual STD of the ensemble average, and then multiplying by 100 to present the ratio as a percent. This normalized ensemble STD reflects the ensemble spread as a percentage of its interannual

190 variability and is unitless, allowing for comparison between variables.

    Kernel density estimates are used to estimate the probability of different Hadley cell expansion rates for each metric. This method, which is described in detail in Silverman (2018), uses kernels to produce smooth nonparametric density estimates. The kernel density estimates are implemented in Python using the Gaussian kernel density estimate function in SciPy (Virtanen et al., 2020) with the standard

195 configuration. The standard deviation of the smoothing kernel is determined from the number of data points and dimensions following Scott's Rule (Scott, 1992).

    Because of the limited number of ensemble members present in ERA5 there is substantial uncertainty in ensemble quantities such as the STD. Following the equation for the fractional uncertainty in the STD: $\frac{\partial \sigma}{\sigma} = \frac{1}{\sqrt{2(N-1)}}$ where σ is the STD and N is the number of ensemble members, the relative uncertainty in

200 the ensemble STD is 0.25 (Harding et al., 2014).

3   Results

In the first two results subsections, we analyze uncertainties in HC trends among the six metrics with the goal of determining the reliability of each metric and the range of trend estimates. In the subsequent sections, we examine the SF metric more in depth as it is the most commonly used metric in the literature.

3.1   HC trend uncertainty

To quantify the HC trend uncertainty, trends are calculated for each ensemble member using all six metrics over the period from 1979-2022. From these trends, kernel density estimates are computed in each season and the annual mean, for all six metrics in the Northern Hemisphere (NH) (Fig. 1) and Southern Hemisphere (SH) (Fig. 2). A brief description of kernel density estimates is presented in Section 2.2.7.

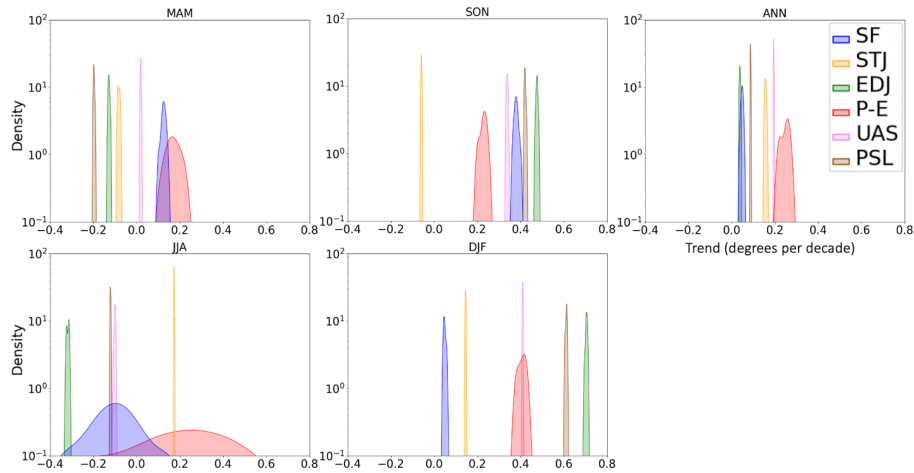



Figure 1 Tropical widening trends over the NH. Shown are kernel density estimate (y-axis) of tropical extent trend (1979-2022) in degrees per decade (x-axis), computed from the nine ERA5 ensemble members for all seasons and the annual mean. Density indicates the relative proportion of trends equal to a particular value, with the area under the curve equal to 1. Note the logarithmic scale for density on the

y-axis due to the large differences between metrics. Positive x-axis values represent northward trend in tropical extent.

230   While in individual seasons the trends vary substantially between metrics, the annual mean trends are typically more similar. The NH annual mean trend for SF is roughly 0.05 degrees per decade, while other metrics estimate 0.05 to 0.3 degrees per decade. The SH annual mean trend ranges between 0.1 and 0.3 degrees per decade, with most metrics (including SF) estimating 0.2 degrees per decade. The near-zero NH tropical expansion rate and weak SH expansion are similar to the generally weak expansion found

235   across multiple reanalysis datasets in Grise et al. (2019). It is also of note that the rate of expansion for the annual mean of P-E in the NH is larger than all other metrics.

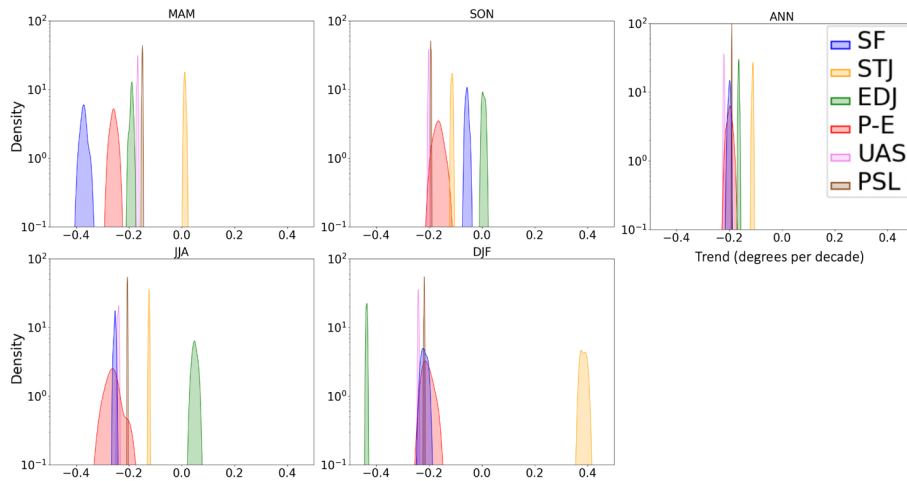



Figure 2 Same as Fig. 1 for SH. Note the different x-axis values, with negative values representing southward expansion.

240

On seasonal time scales, the trends and their uncertainty vary substantially between the different metrics, as shown by the kernel density estimates and the ensemble trend STD; a table of the latter is presented in the Supplement (Table S1). P-E features the largest uncertainty, particularly In the NH and in JJA where the uncertainty is roughly 2 orders of magnitude greater than other metrics. Over the NH, the SF exhibits the second least certain trend in all seasons, particularly in JJA, where the range of estimates include both poleward and equatorward trends and is nearly one order of magnitude less certain than other metrics. Considering that the ensemble spread underestimates the actual uncertainty, the presence of both positive and negative trends indicates that in JJA in the NH SF is not a reliable metric. Over the SH, SF estimated tropical expansion is more robust, as shown by the consistency of the negative sign in all seasons as well as the smaller uncertainty. While the uncertainty over the SH is small for STJ and EDJ in most seasons, STJ in DJF and EDJ in JJA are substantial outliers in trend, estimating tropical contraction while all other metrics record expansion.

We hypothesize that data uncertainty is a major contributor to the larger uncertainties in SF and P-E. SF is based on poorly observed and poorly constrained meridional wind, and P-E uses imperfectly modeled evaporation and precipitation data. On the other hand, STJ and EDJ are based on better observed and, constrained (through the thermal wind relationship) zonal wind data, and UAS and PSL are also well-observed and constrained quantities.

3.2 HC extent uncertainty

We next examine the HC extent uncertainty in individual years. In doing so, the impact of improved observations over time can be seen along with the potential influences of internal climate variability on tropical width estimates. The yearly ensemble STD of tropical extent is calculated for each metric in the NH and SH (Fig. 3). P-E has the greatest STD in nearly all seasons and is particularly unreliable in the NH JJA, where the uncertainty is roughly 2 orders of magnitude greater than the four reliable metrics. SF

uncertainty is typically one order of magnitude larger than the four most reliable metrics and is

particularly unreliable in the NH JJA. Over the SH, the SF extent uncertainty is largest relative to the

other metrics. We also note that there is not much interannual variation in STD for the six metrics in

270   either hemisphere, indicating that internal climate variability does not have a substantial impact on

uncertainty. Counter to this observation are SF and P-E in the NH in JJA, which are even undefined in

some years, resulting from P-E not recording a zero-crossing near the poleward edge of the tropics, and

the SF not recording a closed circulation cell. Examples of the absence of a closed HC in 2019 and 2020

are included in the Supplement (Fig. S1). As a result, in these years the STD is calculated excluding

275   undefined ensemble member values, resulting in imperfect estimates of variation. It is interesting that

2019 and 2020 are the only two years with some members not reproducing closed NH summer circulation

cells, and it remains to be seen whether this is part of a systematic longer-term trend.

Figure 3 Interannual variation in tropical extent uncertainty. Shown are annual ensemble STD of HC

280   extent (degrees latitude) for NH (top) and SH (bottom) by season for all six metrics. Note the logarithmic

scale for the y-axis. The start year for each season is the first available year.

285 Most notably, Fig. 3 shows a reduction in uncertainty over the 40-year period in most seasons and for

nearly every metric, likely due to the increased quantity and quality of observations assimilated into the

reanalysis. Because the ensemble members are created by perturbing the model parameters along with

observations, both a lack of observations and the presence of lower quality observations increase the data

uncertainty. The metrics based on u-wind and sea-level pressure (STJ, EDJ, UAS, PSL) have consistently

290 small uncertainty in all seasons, years, and hemispheres, though the variation for PSL especially in the SH

is the smallest of all metrics. Despite the presence of more conventional observations in the NH,

particularly in the earlier decades, the extent uncertainty is smallest in the SH for all metrics, indicating

that both observations and regional dynamics are relevant. The greater uncertainty in the NH may be due

to the abundance of ocean in the SH and topographic variability in the NH, both of which impact the

295 atmospheric processes which are measured by the HC extent metrics.

3.3 Spatial variability of SF data uncertainty

Because the SF is the most widely used metric, the following sections will take a closer look at the

magnitude and impact of data error for just the SF. To quantify the spatial uncertainty of SF data, the

300 ensemble STD of SF is calculated for all years using zonally and annually averaged meridional wind data.

These annual values are then averaged over two 20-year periods and normalized as described in the

Methods section. This results in the annual mean SF normalized ensemble STD over two time periods

(Fig. 4). The two time periods are 1979-2000 and 2001-2021, comparing the period with fewer

observations to the better observed period following the deployment of many remote sensors (Hersbach et

305 al., 2020). The normalized ensemble STD during the earlier period is clearly larger than during the later

period, indicating the role of improved observations in the reduced SF extent uncertainty in recent

decades seen in Fig. 3. As the SF metric used here calculates the tropical extent using the 500 hPa SF

values between the Hadley and Ferrel cells, only the improvements in these regions are impactful, while

the large decrease in normalized STD in other regions is not directly relevant. However, the

310 improvements in observations in these regions, particularly the inner tropics, constrain the meridional

circulation, likely contributing to reduced uncertainties in other regions of the Hadley circulation in recent

decades. As can be seen from Fig. 4, the normalized STD improves from roughly 6% to 4% at 500 hPa at

the tropical edges in both hemispheres between the two time periods, though these changes are not

statistically significant at the 95% confidence level due to the 0.25 fractional uncertainty in STD as

315 described in Section 2.2.7. While the ensemble variation is generally less than 10% of the interannual

variation, the meridional wind uncertainty is nearly one order of magnitude larger than in the zonal wind

used by STJ and EDJ (Fig. S2). Larger normalized uncertainty in the data underlying the SF metric causes

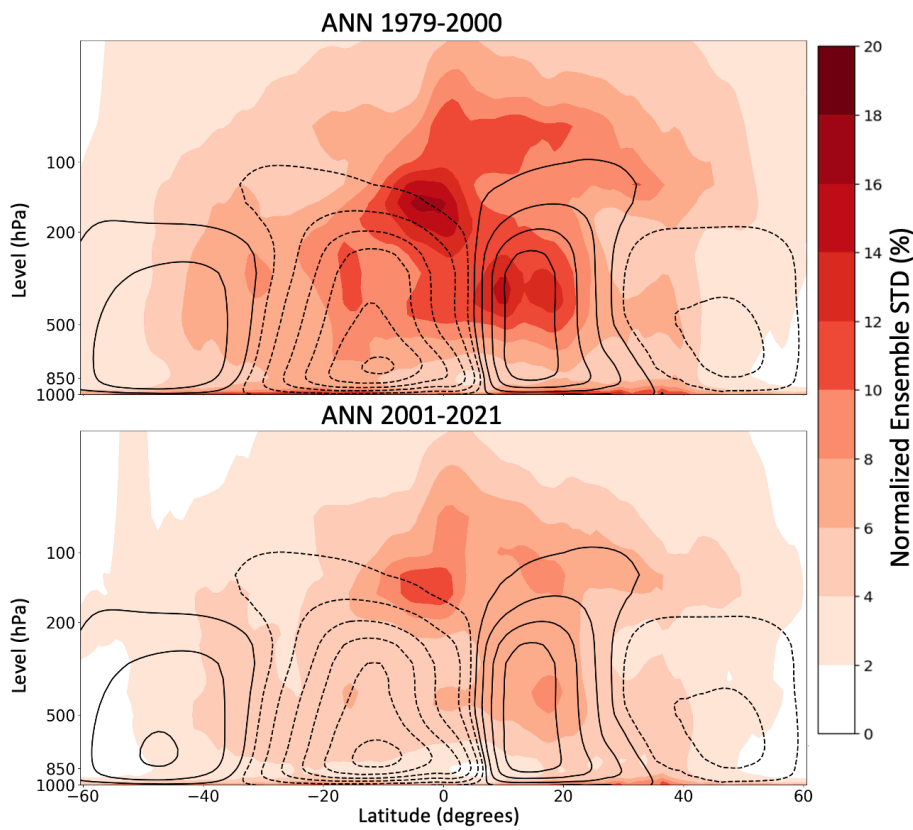
larger error in SF-derived HC expansion rates.

320

Figure 4 Meridional cross sections of SF uncertainty for 1979-2000 (top) and 2001-2021 (bottom),

demonstrating improvement in SF estimates in recent decades. Shading shows the 20-year average of the

annual mean SF normalized ensemble STD (%). Dashed lines indicate negative (counterclockwise)

circulation while solid lines indicate positive (clockwise) circulation.

325

3.4 HC extent errors due to weak SF gradients

As shown earlier in this study, in Section 3.2, the SF extent uncertainty is greater in the NH than SH and is particularly pronounced in the NH in JJA. Here, we aim to investigate the impact of the HC structure on SF extent uncertainty. In order to calculate the tropical extent, the latitude where the circulation changes from clockwise to counterclockwise at 500 hPa must be determined. Observational errors will have a larger impact on SF extent estimates if the region of near-neutral circulation between the Hadley and the Ferrel cell is large, as small variations between the ensemble members can cause the zero-crossing to occur over a larger latitude range. We define Δ in each hemisphere as the width in degrees latitude of the region in the vicinity of the HC edge (poleward of 20 degrees and equatorward of 50 degrees) at 500 hPa where the SF is very weak ($\pm 5*10^9$ kg s^{-1}). This Δ is larger in the NH than SH and is particularly large in NH JJA, coinciding with the seasons and hemispheres where SF trend and extent uncertainty are greatest (Fig. 5).

As shown by Fig. 5, the HC over the NH has a smaller latitudinal extent and features a distinctive narrowing in the middle and upper troposphere. This narrowing coincides with a large Δ, which is not present in other seasons or in the SH. The wide Δ likely represents the impacts of large land masses and greater longitudinal variation in meridional circulation in the NH (Hoskins et al., 2020). In both hemispheres, however, the region of near-zero circulation in summer is larger than winter, suggesting a possible relationship between the weak summer HC and meridional gradients at the poleward edge of the HC. We speculate that Δ is impacted by both the zonal variation in circulation and the strength of the meridional circulation. Hence, in JJA in the NH, the weak meridional circulation and large zonal variation in circulation combine to create a persistent large Δ which causes the SF metric to perform poorly. It is also worth noting that the latitudinal width of the near-zero circulation region does not change significantly during the 40-year period, suggesting that this is a typical and quasi-permanent feature of the atmosphere, and is unrelated to measurement errors.

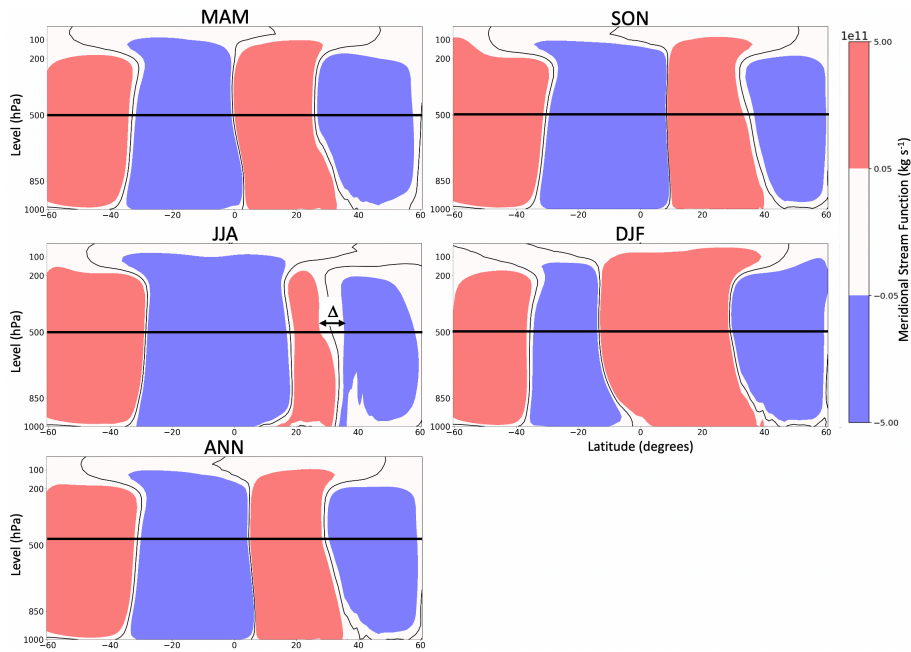



Figure 5 Climatological mean stream function, highlighting regions with values near zero. Red shading represents clockwise circulation stronger than $5*10^9$ kg/s, blue shading indicates counterclockwise motion stronger than $5*10^9$ kg/s, and white shading the region of near-zero stream function values. Thin black contours denote a zero SF value, while the thick horizontal black line indicates the 500 hPa level where the tropical extent and Δ are calculated. In JJA an illustration of Δ is shown in the NH near 500 hPa.

3.5 Impacts of data error and weak gradients on SF uncertainty

In the previous sections, the SF extent uncertainty was shown to be related to both data error and weak meridional gradients. We next determine the impact of these two factors on seasonal and annual HC estimates. The large NH JJA and SON tropical extent uncertainty coincides with large Δ values, suggesting that Δ plays a significant role when using the SF to estimate tropical width (Fig. 6a). Here, Δ is

computed from the 1-degree regridded ERA5 data, resulting in relatively coarse latitudinal estimates. Although data uncertainty, as shown earlier (Fig. 4), partially explains the poor performance of SF relative to other metrics (Figs. 1 and 2) as well as the improvement over time (Fig. 3), no significant relationship was found between the climatological averages of uncertainties in SF data and SF-derived

375   HC extent; a scatter plot showing the poor correlation is included in the Supplement (Fig. S3).

However, when looking at individual annual means, the uncertainty of the SF-based HC extent is well-approximated as a linear function of the average of the SF ensemble STD in the vicinity (± 2 degrees latitude) of the HC edge at 500 hPa (Fig. 6b). In the NH, the extent uncertainty is larger than in the SH, likely due to the aforementioned complicating influences of the NH land masses. The extent uncertainty is

380   well correlated with the SF uncertainty and poorly correlated with Δ, indicating that the reduction in data error is the main reason for the decrease in SF extent uncertainty seen before (Fig. 3); a scatter plot showing the poor fit between Δ and the annual mean SF extent uncertainty is included in the Supplement (Fig. S4). The improved observations in recent decades thus result in more precise tropical width estimates relative to earlier decades.

385   The uncertainty of the SF-derived HC trend has no discernable relationship with the SF data uncertainty when comparing seasons (Fig. S5). However, the trend uncertainty is greatest in seasons and hemispheres where Δ is largest, though this is mostly noticeable in NH JJA (Fig. S6).

Following this analysis, we conclude that Δ represents a persistent factor for the uncertainty in the SF metric, which varies primarily by season and hemisphere, and is nearly constant over the years. As

390   discussed earlier, Δ can be seen as a sensitivity of the HC extent to the data uncertainty, as a larger Δ results in a larger range of latitudes where data error could cause SF to be zero. Δ in individual seasons and hemispheres does not vary much from year-to-year as it is mostly a consequence of relatively fixed meteorological and geographical factors that control the structure of the HC. As a result, Δ represents a

400 persistent amplification of data error, resulting in greater extent uncertainty in certain seasons and

hemispheres, in particular JJA in the NH.

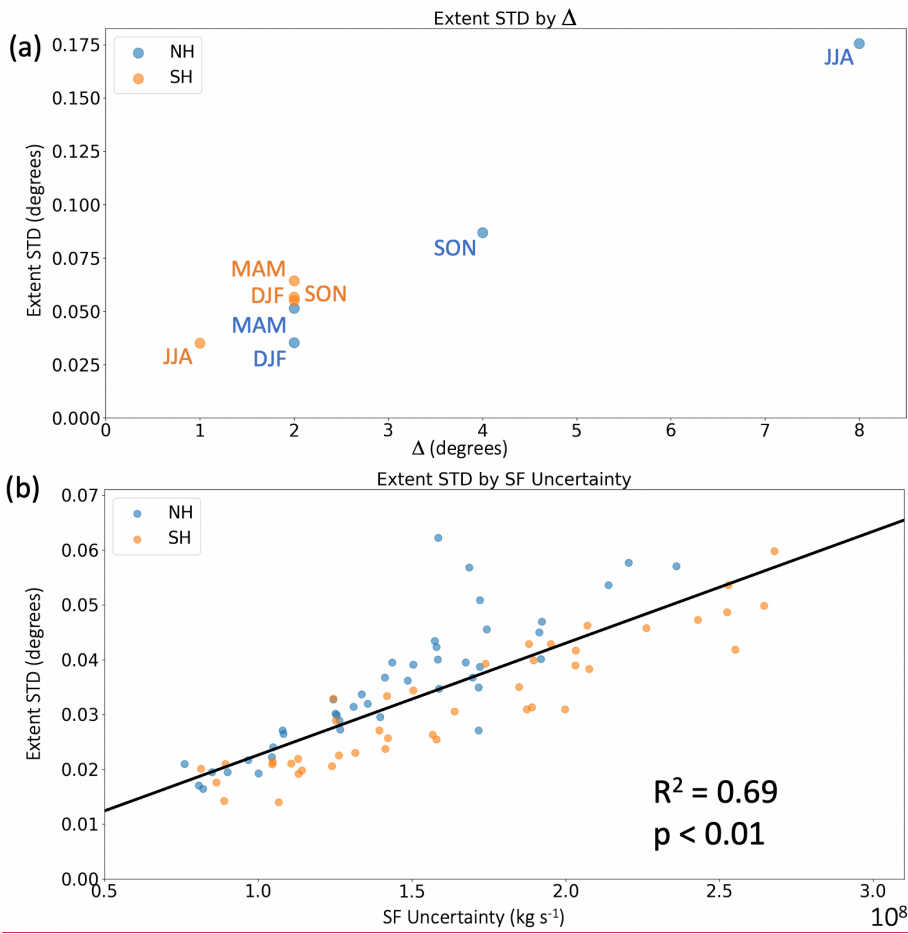



Figure 6 Uncertainty in the SF-based HC extent. (a) climatological mean of HC extent uncertainty

405 (degrees latitude) by Δ (degrees) in all four seasons in both hemispheres. Δ is defined as the width of the

near-zero circulation region at 500 hPa in the vicinity of the HC edge, as shown in Fig. 5. (b) Annual

mean HC extent uncertainty (degrees latitude) by the average of the SF ensemble STD at 500 hPa

averaged within 2 degrees of the HC edge (kg s^{-1}). A linear best fit line is presented in black.

4   Conclusions and discussion

415   In this study, we used ERA5 ensemble members from 1979-2022 to quantify the uncertainty and long-

term trends of different Hadley cell (HC) extent metrics. The annual mean SF HC expansion rate in the

NH was found to be approximately 0.05 degrees per decade in the NH, and 0.2 degrees per decade in the

SH. The HC extent trend uncertainty analysis showed the high uncertainty of SF and P-E trends relative

to the other metrics, particularly in the NH, and most prominently in JJA in the NH. Over the roughly 40-

420   year reanalysis record, we found substantial improvements in HC extent error for all metrics (Table S2),

especially in the SH, likely due to increases in the number and quality of observations assimilated by the

ERA5 reanalysis system (Figs. 3 and 4). The seasonal and hemispheric differences in SF-derived HC

extent uncertainties were well described by variations in the strength of meridional gradients in SF at the

poleward edge of the HC (Figs. 5 and 6a). Differences in annual mean SF extent uncertainty were well

425   correlated with the SF uncertainty near the poleward edge of the HC (Fig. 6b). The high sensitivity to data

uncertainty combined with substantial SF data error suggests that the SF metric is not the most reliable

method of determining the width of the HC in reanalyses.

In the following paragraphs, we discuss our findings to provide suggestions for future studies. For each

metric, the data-driven trend and extent uncertainty, agreement with other metrics, and findings from

430   previous studies will be considered.

Despite being the most widely used metric of tropical extent, SF was found to have much higher

uncertainty in both trend and extent relative to STJ, EDJ, UAS and PSL. The meridional wind used in the

SF metric is generally weak and likely less well-observed than the zonal wind used in other metrics.

Zonal-mean meridional flow is also not dynamically constrained by the temperature field through the

thermal wind relationship as is the case for the zonal wind field. As a result, meridional wind is less accurately represented in reanalysis than zonal wind or surface variables, leading to larger uncertainties relative to the variables used by other metrics. This issue is compounded by the presence of wide regions of near-zero meridional overturning, resulting in a greater sensitivity of the HC edge latitude to the already large data uncertainty. As a result of these two issues, and despite the near-ubiquitous usage in the literature, the SF is not the most useful metric when analyzing tropical extent trends in reanalyses, particularly if the time-period of interest is prior to 2000 or the focus is the NH summer. Over recent decades, when analyzing annual averages, SF becomes a more reliable metric, particularly in the SH.

P-E has the greatest variation between ensemble members in both HC trend and extent, particularly in the NH during JJA. P-E has high uncertainty mostly due to the well-documented issues in modeling precipitation and evaporation in reanalyses (e.g., Simmons et al., 2010). This metric is further challenged by somewhat poor correlations between meridional circulation and vertical moisture flux at the descending branch of the HC, particularly over land masses (Schmidt and Grise, 2017). Despite P-E and related metrics such as the precipitation minimum providing the most societally impactful information, the uncertainty in these metrics is too high to be useful for HC trend analysis with reanalysis data, particularly when other more reliable metrics exist which can be used as proxies.

While STJ featured very small uncertainty, it is not well correlated with SF or P-E, the most direct and relevant metrics of tropical extent, which causes issues when analyzing tropical expansion (Davis and Birner, 2017).

PSL was shown in Davis and Birner (2017) to be well correlated with SF and moderately correlated with P-E, suggesting that it functions as a useful proxy for the societally impactful effects of HC extent change. In our study the uncertainty for PSL tropical extent was found to be very small both in trend and extent. In light of these results, we conclude that PSL is reliable and potentially useful.

EDJ featured little uncertainty in most seasons and hemispheres in both trend and extent but resulted in
trends outside of the range of the other metrics over most seasons. This indicates that the EDJ, though
reliable and well correlated with SF, is also impacted by other processes that are not closely related to the
width of the HC, particularly over individual seasons. As a result, EDJ is most useful for analyzing the
mid-latitudes where changes to the EDJ are most relevant and for studying annual mean HC change.

For UAS, both the trend and extent uncertainties were found to be small in all seasons and hemispheres,
and the trends were within the range estimated by other metrics. UAS has also been shown to be well
correlated to SF and P-E (Davis and Birner, 2017). Because of these factors, we find UAS to be a reliable
and useful metric for analyzing tropical extent trends in reanalyses.

This study focused on errors within a single reanalysis dataset, while many previous studies were based
on data from multiple reanalyses, climate models, or observations. When analyzing multiple reanalyses,
uncertainty arising from measurement error is likely to exist, as many of the issues highlighted in this
study are present in all reanalyses. However, additional uncertainties may be present in the inter-
reanalysis spread due to variation in observations and assimilation schemes. Our results from the ERA5
ensemble may translate poorly to inter-reanalysis ensembles containing older reanalyses, leading to some
metrics which were found to have low uncertainty in ERA5 being less reliable across reanalyses. Some of
the issues seen in this analysis, such as the weak meridional gradients in SF are likely to be present in
climate model data as well, while others such as observational uncertainty are instead replaced by
modeling error. Future work expanding on this topic should quantify the tropical extent uncertainty in
climate models, and in addition determine the most reliable metrics.

Code availability

The code for analysis is available at https://zenodo.org/record/7430530.

Data availability

This study uses monthly averaged ensemble members from the ERA5 reanalysis, which can be downloaded at https://cds.climate.copernicus.eu (Hersbach et al., 2020)

485 **Supplement**

Author Contributions

DB and TR designed the study. DB performed the analyses and wrote the manuscript with feedback from all authors.

490 **Competing Interests**

The authors declare that they have no conflict of interest.

Disclaimer

Acknowledgements

495 We acknowledge the European Center for Medium-Range Weather Forecasts (ECMWF) for producing the ERA5 reanalysis dataset. We also thank Hans Hersbach and Paul Berrisford at ECMWF for their communication on the ERA5 reanalysis dataset, and the Center for High Performance Computing at the University of Utah for providing computing resources.

Financial Support

500 Thomas Reichler acknowledges funding from the National Science Foundation under award no. 2103013.

References

Adam, O., Grise, K. M., Staten, P., Simpson, I. R., Davis, S. M., Davis, N. A., Waugh, D. W., Birner, T. and Ming, A.: The TROPD software package (V1): Standardized methods for calculating tropical-width diagnostics, Geoscientific Model Development, 11(10), 4339–4357, doi:10.5194/gmd-11-4339-2018, 2018.

Allen, R. J., Norris, J. R. and Kovilakam, M.: Influence of anthropogenic aerosols and the Pacific decadal oscillation on tropical belt width, Nature Geoscience, 7(4), 270–274, doi:10.1038/ngeo2091, 2014.

Chemke, R. and Polvani, L. M.: Opposite tropical circulation trends in climate models and in Reanalyses, Nature Geoscience, 12(7), 528–532, doi:10.1038/s41561-019-0383-x, 2019.

Chen, J., Carlson, B. E. and Del Genio, A. D.: Evidence for strengthening of the tropical general circulation in the 1990s, Science, 295(5556), 838–841, doi:10.1126/science.1065835, 2002.

Choi, J., Son, S.-W., Lu, J. and Min, S.-K.: Further observational evidence of Hadley cell widening in the Southern Hemisphere, Geophysical Research Letters, 41(7), 2590–2597, doi:10.1002/2014gl059426, 2014.

Davis, N. A. and Davis, S. M.: Reconciling hadley cell expansion trend estimates in Reanalyses, Geophysical Research Letters, 45(20), doi:10.1029/2018gl079593, 2018.

Davis, N. and Birner, T.: On the discrepancies in tropical belt expansion between reanalyses and climate models and among tropical belt width metrics, Journal of Climate, 30(4), 1211–1231, doi:10.1175/jcli-d-16-0371.1, 2017.

Davis, S. M. and Rosenlof, K. H.: A multidiagnostic intercomparison of tropical-width time series using reanalyses and satellite observations, Journal of Climate, 25(4), 1061–1078, doi:10.1175/jcli-d-11-00127.1, 2012.

Fu, Q., Johanson, C. M., Wallace, J. M. and Reichler, T.: Enhanced mid-latitude tropospheric warming in satellite measurements, Science, 312(5777), 1179–1179, doi:10.1126/science.1125566, 2006.

Grise, K. M., Davis, S. M., Simpson, I. R., Waugh, D. W., Fu, Q., Allen, R. J., Rosenlof, K. H., Ummenhofer, C. C., Karnauskas, K. B., Maycock, A. C., Quan, X.-W., Birner, T. and Staten, P. W.: Recent tropical expansion: Natural variability or forced response?, Journal of Climate, 32(5), 1551–1571, doi:10.1175/jcli-d-18-0444.1, 2019.

Grise, K. M., Davis, S. M., Staten, P. W. and Adam, O.: Regional and seasonal characteristics of the recent expansion of the Tropics, Journal of Climate, 31(17), 6839–6856, doi:10.1175/jcli-d-18-0060.1, 2018.

Harding, B., Tremblay, C. and Cousineau, D.: Standard errors: A review and evaluation of standard error estimators using Monte Carlo simulations, The Quantitative Methods for Psychology, 10(2), 107–123, doi:10.20982/tqmp.10.2.p107, 2014.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.

540 J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146(730), 1999–2049, doi:10.1002/qj.3803, 2020 (data available at: https://cds.climate.copernicus.eu, last access: 1 June 2022).

Hoskins, B. J., Yang, G. Y. and Fonseca, R. M.: The detailed dynamics of the June–August Hadley cell, Quarterly Journal of the Royal Meteorological Society, 146(727), 557–575, doi:10.1002/qj.3702, 545 2020.

Hu, Y. and Fu, Q.: Observed poleward expansion of the Hadley circulation since 1979, Atmospheric Chemistry and Physics, 7(19), 5229–5236, doi:10.5194/acp-7-5229-2007, 2007.

Hu, Y., Zhou, C. and Liu, J.: Observational evidence for poleward expansion of the Hadley circulation, Advances in Atmospheric Sciences, 28(1), 33–44, doi:10.1007/s00376-010-0032-1, 2010.

550 Hudson, R. D., Andrade, M. F., Follette, M. B. and Frolov, A. D.: The total ozone field separated into meteorological regimes – part II: Northern Hemisphere mid-latitude total ozone trends, Atmospheric Chemistry and Physics, 6(12), 5183–5191, doi:10.5194/acp-6-5183-2006, 2006.

Isaksen, L., Bonavita, M., Buizza, R., Fisher, M., Haseler, J., Leutbecher, M., and Raynaud, L.: Ensemble of data assimilations at ECMWF, Technical Memorandum No. 636, ECMWF, 1–48, 555 https://www.ecmwf.int/node/10125 (last access: 3 March 2023), 2010.

Karnauskas, K. B. and Ummenhofer, C. C.: On the dynamics of the Hadley circulation and subtropical drying, Climate Dynamics, 42(9-10), 2259–2269, doi:10.1007/s00382-014-2129-1, 2014.

Lu, J., Vecchi, G. A. and Reichler, T.: Expansion of the Hadley cell under Global Warming, Geophysical Research Letters, 34(6), doi:10.1029/2006gl028443, 2007.

560 Previdi, M. and Liepert, B. G.: Annular modes and Hadley cell expansion under Global Warming, Geophysical Research Letters, 34(22), doi:10.1029/2007gl031243, 2007.

Schmidt, D. F. and Grise, K. M.: The response of local precipitation and sea level pressure to Hadley cell expansion, Geophysical Research Letters, 44(20), doi:10.1002/2017gl075380, 2017.

Schulzweida, Uwe: CDO User Guide, https://doi.org/10.5281/zenodo.7112925, 4 October 2022.

565 Scott, D. W.: Multivariate density estimation: Theory, practice, and visualization, J. Wiley., 1992.

Seidel, D. J. and Randel, W. J.: Recent widening of the Tropical Belt: Evidence from tropopause observations, Journal of Geophysical Research, 112(D20), doi:10.1029/2007jd008861, 2007.

Seidel, D. J., Fu, Q., Randel, W. J. and Reichler, T. J.: Widening of the Tropical Belt in a changing climate, Nature Geoscience, 1(1), 21–24, doi:10.1038/ngeo.2007.38, 2007.

570 Seviour, W. J., Davis, S. M., Grise, K. M. and Waugh, D. W.: Large uncertainty in the relative rates of dynamical and hydrological tropical expansion, Geophysical Research Letters, 45(2), 1106–1113, doi:10.1002/2017gl076335, 2018.

Silverman, B. W.: Density Estimation for statistics and data analysis, Routledge., 2018.

575 Simmons, A. J., Willett, K. M., Jones, P. D., Thorne, P. W. and Dee, D. P.: Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from Reanalyses and monthly gridded observational data sets, Journal of Geophysical Research, 115(D1), doi:10.1029/2009jd012442, 2010.

Solomon, A., Polvani, L. M., Waugh, D. W., & Davis, S. M. (2016). Contrasting upper and lower
580 atmospheric metrics of tropical expansion in the Southern Hemisphere. *Geophysical Research Letters*, *43*(19). https://doi.org/10.1002/2016gl070917

Staten, P. W., Grise, K. M., Davis, S. M., Karnauskas, K. B., Waugh, D. W., Maycock, A. C., Fu, Q., Cook, K., Adam, O., Simpson, I. R., Allen, R. J., Rosenlof, K., Chen, G., Ummenhofer, C. C., Quan, X.-W., Kossin, J. P., Davis, N. A. and Son, S.-W.: Tropical widening: From global
585 variations to regional impacts, Bulletin of the American Meteorological Society, 101(6), doi:10.1175/bams-d-19-0047.1, 2020.

Staten, P. W., Grise, K. M., Davis, S. M., Karnauskas, K. and Davis, N.: Regional widening of tropical overturning: Forced change, natural variability, and recent trends, Journal of Geophysical Research: Atmospheres, 124(12), 6104–6119, doi:10.1029/2018jd030100, 2019.

590 Staten, P. W., Lu, J., Grise, K. M., Davis, S. M. and Birner, T.: Re-examining tropical expansion, Nature Climate Change, 8(9), 768–775, doi:10.1038/s41558-018-0246-2, 2018.

Staten, P. W., Rutz, J. J., Reichler, T. and Lu, J.: Breaking down the tropospheric circulation response by forcing, Climate Dynamics, 39(9-10), 2361–2375, doi:10.1007/s00382-011-1267-y, 2011.

Tebaldi, C. and Knutti, R.: The use of the multi-model ensemble in Probabilistic Climate Projections,
595 Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1857), 2053–2075, doi:10.1098/rsta.2007.2076, 2007.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore,
600 E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A. P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C. N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D. A., Hagen, D. R., Pasechnik, D. V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young,
605 G., Price, G. A., Ingold, G.-L., Allen, G. E., Lee, G. R., Audren, H., Probst, I., Dietrich, J. P., Silterra, J., Webber, J. T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J. L., de Miranda Cardoso, J. V., Reimer, J., Harrington, J., Rodríguez, J. L., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N. J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P. A., Lee, P., McGibbon, R. T.,
610 Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., et al.: SciPy 1.0: Fundamental algorithms for scientific computing in python, Nature Methods, 17(3), 261–272, doi:10.1038/s41592-019-0686-2, 2020.

Waugh, D. W., Grise, K. M., Seviour, W. J., Davis, S. M., Davis, N., Adam, O., Son, S.-W., Simpson, I. R., Staten, P. W., Maycock, A. C., Ummenhofer, C. C., Birner, T. and Ming, A.: Revisiting the

615   relationship among metrics of tropical expansion, Journal of Climate, 31(18), 7565–7581, doi:10.1175/jcli-d-18-0108.1, 2018.

Xian, T., Xia, J., Wei, W., Zhang, Z., Wang, R., Wang, L.-P. and Ma, Y.-F.: Is Hadley cell expanding?, Atmosphere, 12(12), 1699, doi:10.3390/atmos12121699, 2021.

Zhou, Y. P., Xu, K.-M., Sud, Y. C. and Betts, A. K.: Recent trends of the tropical hydrological cycle
620   inferred from global precipitation climatology project and International Satellite Cloud Climatology Project Data, Journal of Geophysical Research, 116(D9), doi:10.1029/2010jd015197, 2011.

625

630

635

Supplement of

Large Uncertainty in Observed Estimates of Tropical Width From the Meridional Stream Function

Daniel Baldassare[1], Thomas Reichler[1], Piret Plink-Björklund[2], and Jacob Slawson[2]

- [1]Department of Atmospheric Sciences, University of Utah, Salt Lake City, UT 84112, USA
- [2]Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO 80401, USA

Correspondence: Daniel Baldassare (daniel.baldassare@utah.edu)

640

645

650

655

Table S1: Ensemble Trend STD

Ensemble Trend STD

		MAM	JJA	SON	DJF	ANN
NH	**SF**	0.010	0.118	0.009	0.005	0.005
	STJ	0.005	0.001	0.002	0.002	0.004
	EDJ	0.004	0.005	0.004	0.004	0.003
	P-E	0.030	0.263	0.015	0.017	0.019
	UAS	0.002	0.003	0.003	0.002	0.001
	PSL	0.003	0.002	0.003	0.004	0.001
SH	**SF**	0.011	0.004	0.006	0.011	0.004
	STJ	0.003	0.002	0.003	0.011	0.002
	EDJ	0.005	0.009	0.006	0.002	0.002
	P-E	0.011	0.027	0.017	0.018	0.009
	UAS	0.002	0.003	0.002	0.002	0.002
	PSL	0.002	0.001	0.001	0.001	0.001

660   Table S1: HC trend ensemble standard deviation (degrees per decade) in each hemisphere, for all seasons and the annual mean, of all six metrics across the nine ensemble members. Results are qualitatively similar to the kernel density estimates (Fig. 1, Fig. 2) but instead provide a single uncertainty value rather than a range of probable values.

665

Text S1: Unenclosed stream function northern hemisphere summer 2019 and 2020

The NH JJA HC in 2019 and 2020 was found to be unenclosed in some ensemble members, with no positive circulation near 500 hPa and 20 degrees North. In 2019 a majority of ensemble members showed these features, while in 2020 only one member did not have a single enclosed cell in the NH. These were the only two years where this behavior occurs. A variety of climatic oscillations were checked in the attempt to find an explanation for why this only occurs in these two years, but no satisfactory explanation was found. Examples of the unenclosed circulation are shown in Fig. S1. The SF metric is not able to calculate a HC extent in this situation, leading to imperfect estimates of SF uncertainty.

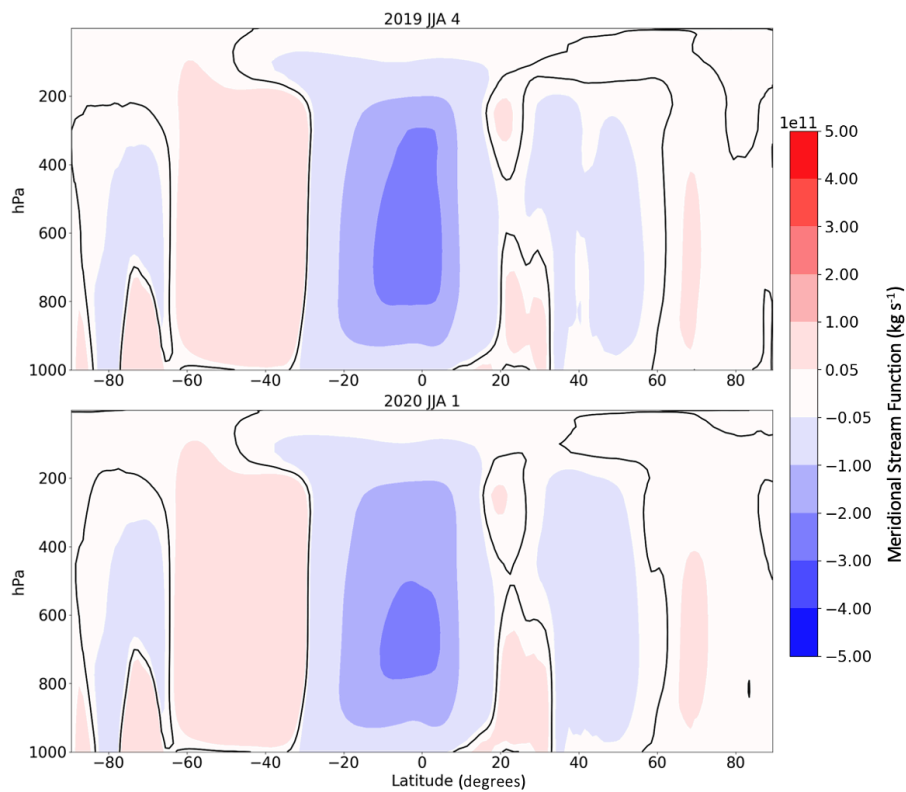



Fig. S1: SF values (kg s⁻¹) for ensemble member 4 in JJA 2019, and ensemble member 1 in JJA 2020, showing an unenclosed NH HC. The black contours indicate the zero-crossing, while shading signifies SF intensity. Note the region near 500 hPa and 20 degrees in both years where there is no positive circulation.

Text S2: Meridional and zonal wind errors

680   The normalized error in meridional wind was found to be much larger than zonal wind, as shown in Fig. S2. The larger normalized differences between ensemble members in meridional wind compared to zonal wind indicate that the greater uncertainty in SF compared to STJ, EDJ or UAS may be the result of larger relative errors in the underlying data. The non-normalized ensemble STD for meridional wind was found to be of similar magnitude to zonal wind (not shown), with the low absolute values and subsequent smaller time variation resulting in larger
685   normalized errors.

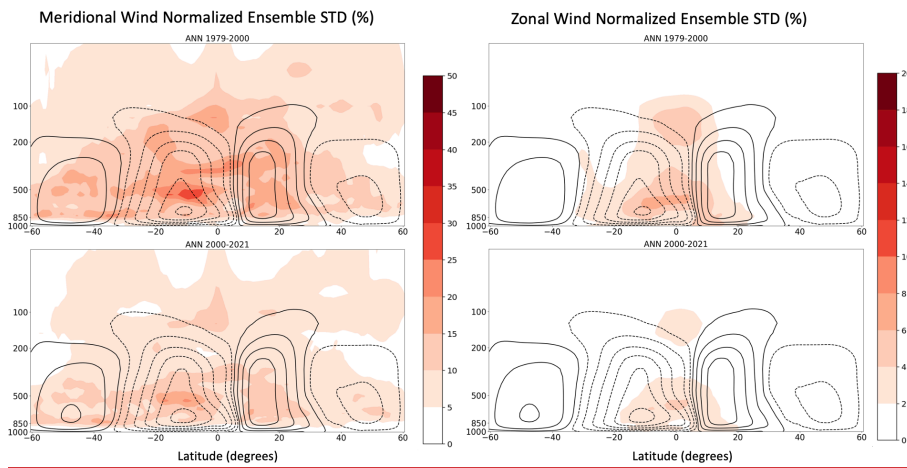



Fig. S2: Annual mean meridional wind (left) and zonal wind (right) normalized ensemble STD (%) from 1979-2000 (top) and 2001-2021 (bottom). Note the different scales for meridional and zonal wind due to the much larger errors in meridional wind.

690

695

Text S3: Poor correlation between climatological mean SF uncertainty and SF extent

The climatological mean SF uncertainty and climatological mean of SF extent STD were not found to be well correlated in this analysis, leading to the interpretation that climatological mean SF uncertainty is best explained by Δ. The scatter plot of these two variables is shown in Fig. S3. The SF uncertainty near the tropical edge at 500 hPa does not vary noticeably by season or hemisphere, while the extent STD does.

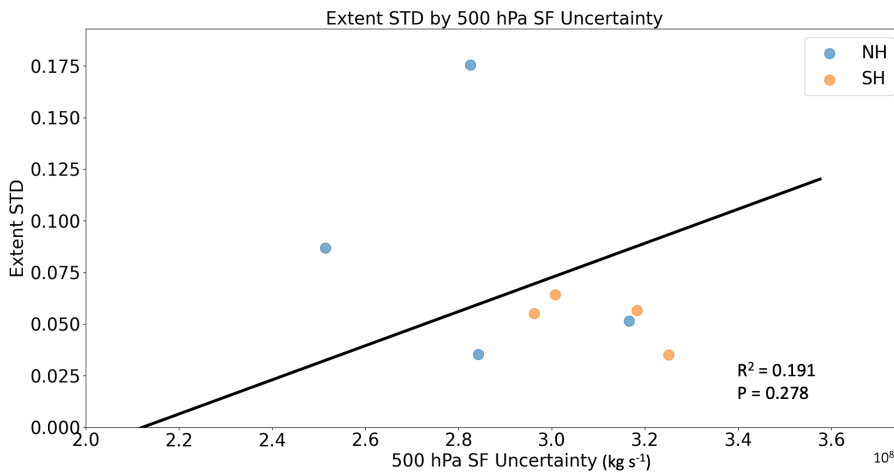



Fig. S3: Climatological mean SF extent STD (degrees latitude) by 500 hPa SF uncertainty (kg s^{-1}) near the tropical edge, showing no significant relationship.

Text S4: Annual mean SF extent STD and Δ

When analyzing the annual mean SF uncertainty over individual years, Δ is a poor predictor. The relationship between these two variables is shown in Fig. S4. The width of Δ does not change noticeably during the time period of this study, while the extent STD decreases. Further, Δ is most noticeable in the difference between seasons and hemispheres, and is less important in the annual mean.

720

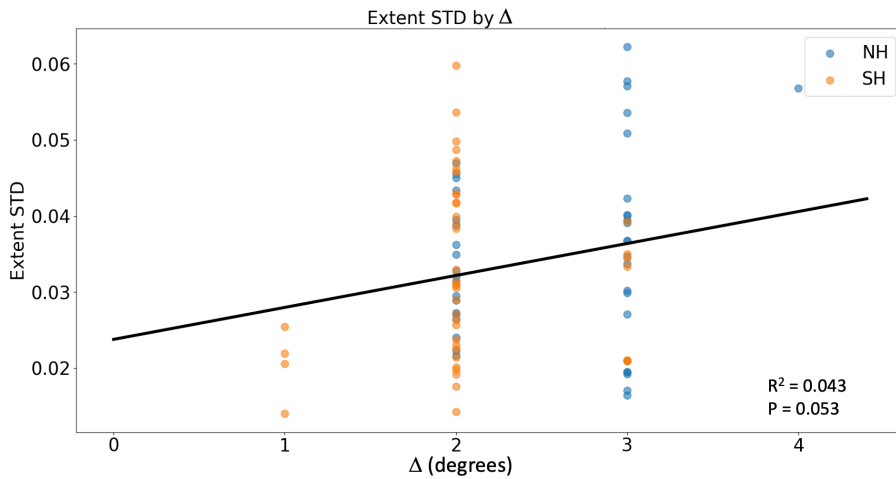



Fig. S4: Annual mean extent STD (degrees latitude) by Δ (degrees latitude), demonstrating that Δ is not a useful predictor of SF extent STD in the annual mean.

725

730

Text S5: SF trend STD by 500 hPa SF uncertainty

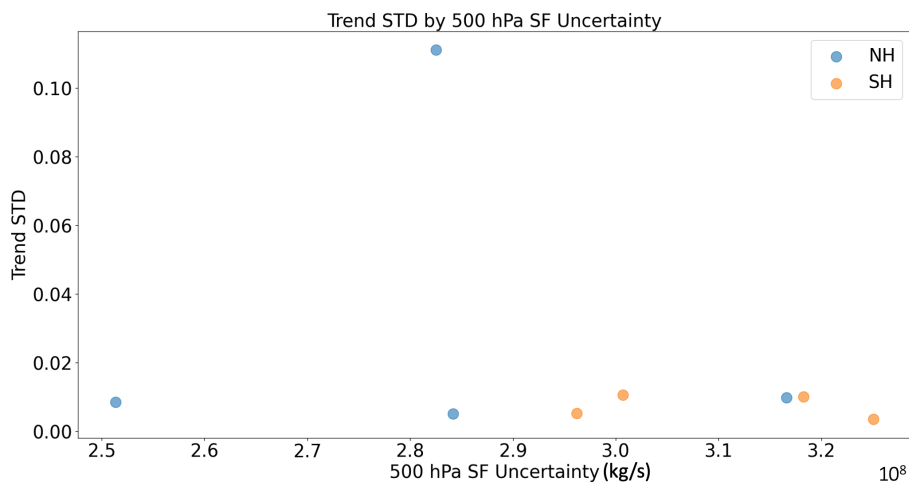
Similar to the climatological mean SF extent STD, no relationship was found between the SF trend STD and the 500 hPa SF uncertainty. The trend uncertainty differences between seasons are not well-described by the SF error differences, indicating that the particularly poor performance of SF in NH JJA is unrelated to data error.

735

Fig. S5: SF trend STD (degrees per decade) by 500 hPa SF uncertainty (kg s^{-1}) for all four seasons in both hemispheres.

740

745

Text S6: SF trend STD by Δ

Aside from NH JJA, which features both large trend STD and large Δ, the relationship between trend STD and Δ among seasons is unclear. However, it is shown here that the large trend STD in NH JJA is caused by the large Δ values during this season as the small perturbations in reanalysis data are able to cause large differences in trend.
750   Here we can see once again the Δ functioning as a sensitivity to data error.

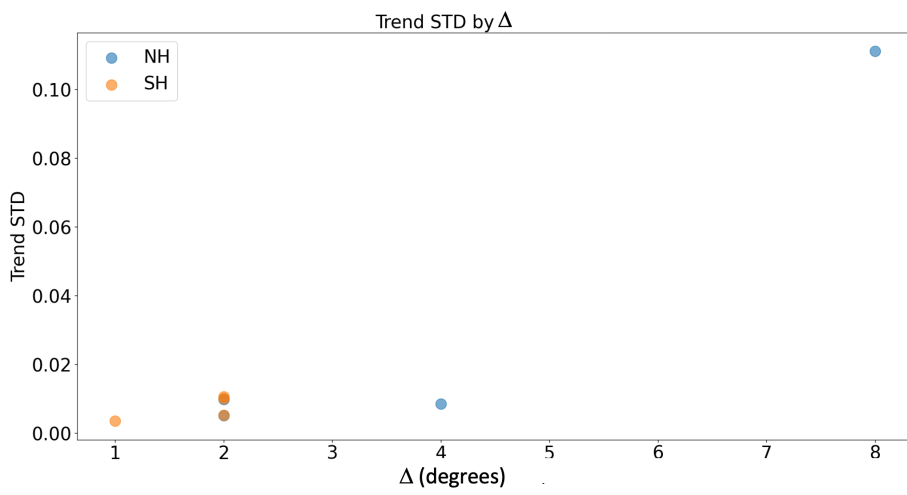



Fig. S6: SF trend STD (degrees per decade) by Δ (degrees) for all four seasons in both hemispheres.

755

760

Table S2: Improvements in HC extent uncertainty

As shown in Fig. 3, the HC extent uncertainty improves from 1979-2022 for almost all metrics in every season and the annual mean. This coincides with increasing quantity and quality of observations. Here, we quantify the improvement in the annual mean uncertainty by applying a linear best fit to the time series of ensemble extent STD for each metric in the NH and SH. This results in larger improvements for the least certain metrics, P-E and SF, though improvements are substantial for all metrics in both hemispheres.

765

ANN Extent Ensemble STD Change (degrees per decade)

	SF	STJ	EDJ	P-E	UAS	PSL
NH	-0.0042	-0.0033	-0.0011	-0.0065	-0.0010	-0.0014
SH	-0.0074	-0.0037	-0.0041	-0.0021	-0.0020	-0.0012

Table S2: Annual mean HC extent ensemble STD change (degrees per decade) for all six metrics in both hemispheres.

770