University of Reading

Clear-air turbulence in reanalysis data

A thesis submitted for the degree of Doctor of Philosophy

Department of Meteorology

Mark Prosser

March 2024

Table of Contents

University of Reading

Declaration of Authorship

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Mark Prosser

Chapter 3: Prosser, M.C., Williams, P.D., Marlton, G.J. and Harrison, R.G., 2024. Using aircraft accidents and incidents to evaluate the skill of clear-air turbulence diagnostics. . **In preparation.**

Chapter 4: Prosser, M.C., Williams, P.D., Marlton, G.J. and Harrison, R.G., 2023. Evidence for large increases in clear-air turbulence over the past four decades. ***Geophysical Research Letters***, *50*(11), e2023GL103814.

Chapter 5: Prosser, M.C., Williams, P.D., Marlton, G.J. and Harrison, R.G., 2024. Towards seasonal predictions of global clear-air turbulence. **In preparation.**

In all 3 results chapters, under the supervision of Williams and Marlton, Prosser conducted the analysis and ECMWF Reanalysis 5 (ERA5) data retrieval. The initial formulation of the research idea for Chapter 4 was Prosser and for 3 and 5, was largely Williams, with Prosser and Marlton contributing to the research design as the projects unfolded. The drafts of all 3 results chapters were first written by Prosser then feedback given by Williams and Marlton. Harrison provided a final check prior to submission.

Thesis Abstract

Turbulence is a significant problem for the aviation industry, costing hundreds of millions of dollars each year. Clear-air turbulence (CAT) is particularly troublesome as it is invisible and tends to occur in the vicinity of the jet stream, where eastbound aircraft fly to take advantage of the westerly tailwinds. With climate change, CAT is projected to increase.
Because CAT is sub-grid scale in numerical weather prediction models and therefore cannot be explicitly simulated, diagnostics are used to indicate regions of CAT.

In this thesis, using 80 serious aircraft turbulence encounters, the skill of 21 such diagnostics at forecasting injury-causing CAT is evaluated and substantial variation in diagnostic skill is found. This knowledge is then retrospectively applied to two previous future CAT projections.
Using the same 21-diagnostic average, the global distribution of 197hPa CAT is calculated for the years 1979 and 2020 as well as the relative and absolute change over this period.
Diagnosed CAT occurs mostly over the midlatitudes (more so in the northern hemisphere than the southern) and is more frequent over the oceans than the land with hotspots to the west of ocean basins.
Since 1979, the biggest changes have been a 54% increase in severe-or-greater CAT over the north Atlantic and 41% over the continental USA.
The global distribution of 197hPa CAT is examined in different phases of ENSO and the NAO. Both sources of climate variability have a substantial impact on the distribution of CAT, although the effect of ENSO is global whereas the impact of the NAO is more local to the north Atlantic and Europe.
This information can be used to make a seasonal CAT forecast which should help airlines better optimise their fuel requisitioning.

The work in this thesis indicates that aviation-affecting CAT is sensitive to climate conditions, both variability and trends.
If the aviation industry is to successfully adapt to the changing climate, advanced preparations as well as continued progress in accurately forecasting CAT will be essential.

Acknowledgements

A PhD is a huge undertaking, and this one in particular spanned 5.5 years. There are a large number of people to thank, people without whom I may not have finished.

First and foremost my main supervisor Paul Williams and Graeme Marlton, who kept me on track and gave me lots of useful feedback throughout the PhD, I really enjoyed many of our conversations. It was particularly fun dealing with the media splash that occurred following the publication of our first paper. Thank you also for the flexibility you showed when mental health issues occurred, that certainly helped a lot in over coming them. Thanks here also to Giles Harrison for coming up with a very helpful strategy for dealing with literature reviews.

Many thanks also to my monitoring Committee (Prof Tom Frame and Prof Danny Feltham) who provided lots of useful input to the PhD at various junctures.

Although not formally involved in my PhD, I would also like to thank Denis Sergeev from the University of Exeter who provided enormous help over the years with Python related issues and enabled me to become a confident Python user. I hope I didn't abuse his generosity too much. James Delph who provided a lot of useful advice on terminal/bash related queries. Paul Dando from the ECMWF also provided extensive help with retrieving and processing ERA5 data, help that was crucial to enabling my PhD project. Paul Dando was also a very helpful and generous man, so thank you Paul.

On a slightly different note my enormous gratitude to Dan Shipley a fellow PhD student who agreed to tutor me throughout much of my 1st year on maths/fluid dynamic issues. It is largely due to him that I managed to get through the fluid dynamics exam in the first year. Dan's love of his subject was infectious and throughout the PhD I would often share a pint with him and discuss this and that in the world of Physics.

Throughout my PhD I lived with a number of very cool housemates. Tim, Will, Sudeep, Wilf and Mike, but special thanks goes to Mike, who I particularly enjoyed living due to the wide-ranging and often hilarious conversation we would have. We also shared the fact that we were Scenario cohort 5 PhD students together so shared the trials and tribulations of the PhD journey! Thanks also to NERC, and of course Scenario doctoral training partnership’s Wendy Neale!

While also discussing home-life, thanks also to our wonderful landlady Catherine who despite living next door, let us live our own lives, kept the rent from going up for the entire 5 years I lived there, and was always very helpful if there were any issues!

Another big part of a PhD experience is ones officemates and I was blessed with a number of great characters in 2U07. In the 1st year, Davi Mignac, an ever cheerful Brazilian along with Will Morrison and Tom Hall. Even after they graduated we came together for the occasional 'ladz' holiday, which were all very memorable. Then after they had graduated I was lucky enough to share an office with Gwyneth Matthews and Dan Ayres, who are both cool cats and with whom I shared many an interesting conversation.

And of course a very big thank you extends to my family, my mother Fionnuala and brother Donal and cousin Stephen. To my many friends who supported me throughout the PhD, to my ex Millie with whom I enjoyed many Saturdays in 2022/23, to my old friend Charlotte who would always act as a valuable sounding board to this or that.It would also be remiss of me not to mention the South Reading Leisure Centre squat rack. On rare occasions when PhD satisfaction was low, this was somewhere I could always turn.

No man is an island and all of those above help act as a solid foundation that ensured I managed to successfully complete, and occasionally thrived in this PhD.

ACRONYMS

AO	**Arctic Oscillation**
ATOVS	**Advanced TIROS Operational Vertical Sounder**
BLT	**Boundary-layer Turbulence**
BUFR	**Binary Universal Form for Representation of Meteorological Data**
CAT	**Clear-air turbulence**
CIT	**Convectively induced turbulence**
DEF	**deformation**
DEVG	**derived equivalent vertical gust**
DJF	**December/January/February**
EA	**East Atlantic**
EB	**Eastbound**
ECMWF	**European Centre for Medium-Range Weather Forecasts**
EDA	**ensemble data assimilation**
EDJ	**Eddy-driven jet**
EDR	**Eddy Dissipation Rate**
ENSO	**El Niño-Southern Oscillation**
EOF	**Empirical orthogonal functions**
ERA5	**ECMWF Reanalysis 5**
FP	**False Positives**
FN	**False Negative**
GCM	**General Circulation Model**
GHG	**Greenhouse gases**
GNSS-RO	**Global Navigation Satellite System – Radio Occultation**
GTG	**Graphical Turbulence Guidance**
GW	**Gravity waves**
HNL	**Honolulu airport**
IGW	**inertia-gravity waves**
JFK	**John F. Kennedy airport**
JFMAMJJASOND	**Jan/Feb/Mar/Apr/May/Jun/Jul/Aug/Sep/Oct/Nov/Dec**
JJA	**June/July/August**
KHI	**Kelvin-Helmholtz instability**
LHR	**London Heathrow**
LIDAR	**Light detection and ranging**
LMOG	**light-to-moderate-or-greater (turbulence)**
LOG	**light-or-greater (turbulence)**
MAM	**March/April/May**
MLR	**multiple linear regression**

MOG	**moderate-or-greater (turbulence)**
MSOG	**moderate-to-severe-or-greater (turbulence)**
MWT	**Mountain-wave Turbulence**
N-S	**North-South**
NA	**North Atlantic**
NAO	**North Atlantic Oscillation**
NCAR	**National Center for Atmospheric Research**
NCEP	**National Center for Environmental Prediction**
NCT	**Near-cloud Turbulence**
NDJFM	**Nov/Dec/Jan/Feb/Mar**
NH	**Northern hemisphere**
NTSB	**National Transportation Safety Board**
NWP	**Numerical Weather Prediction**
ONI	**Oceanic Niño Index**
PIREPS	**Pilot Reports**
PODn	**Probability of Detection (no)**
PODy	**Probability of Detection (yes)**
PV	**Potential Vorticity**
PVU	**Potential Vorticity Units**
ROC	**Receiver Operating Characteristic**
SFO	**San Francisco airport**
SH	**Southern hemisphere**
SIT	**Shear induced turbulence**
SOG	**severe-or-greater (turbulence)**
SON	**September/October/November**
SP	**sub-polar**
SPJ	**sub-polar jet**
ST	**sub-tropical**
STJ	**sub-tropical jet**
TI1/TI2	**Turbulence index**
TIROS	**Television Infrared Observation Satellite**
TN	**True Negatives**
TOVS	**TIROS Operational Vertical Sounder**
TP	**True Positives**
ULTURB	**Upper Level TURBulence**
VWS	**Vertical wind shear**
WB	**westbound**
WVT	**Wake Vortex Turbulence**

Chapter 1

Introduction

1.1 Turbulence and its impacts on aviation

Turbulence is already a problem for the aviation industry.
Every year encounters with the more severe strength categories causes injuries to passengers and crew and damage to the air frame that then need to undergo inspection resulting in delays and lost productivity costs.

Certain forms of turbulence are location specific or occur in easily observable meteorological phenomena such as clouds and are therefore manageable. Convectively induced turbulence tends to occur inside clouds and Mountain wave turbulence tends to occur above mountain ranges for instance. Clear-air turbulence (CAT) on the other hand tends to occur in the vicinity of the jet streams and because aircraft makes ample use of the jet stream to get a tail-wind when flying eastwards, CAT encounters can occur at any moment. Because CAT happens in clear-air, it cannot be remotely detected with radar, this combined with the fact that the seat belt signs tend to be off at cruise altitude leads to a number of hospitalisations each year. While light detection and ranging (LIDAR) attached to the front of aircraft shows promise as an early warning system capable of giving passengers a few seconds to get strapped in, it is still considered too expensive (Kauffmann, 2002).

CAT (and turbulence more generally) is an inherently complicated phenomenon; “One of the last great challenges of numerical weather prediction” (Sharman et al., 2012) and while our ability to forecast it continues to improve, it is still far from perfect with many false positives and false negatives. Because processes which give rise to turbulence occur at scales far smaller than the grid size of climate and weather numerical models, much of the research aimed at improving the forecasting of CAT tries to develop statistical relationships between larger weather patterns that do get resolved in models and the occurrence of CAT. These statistical relationships are known as ‘diagnostics’ which can be applied to the output of such models to diagnose the likelihood of turbulence. The Graphical Turbulence Guidance (GTG) system (Sharman et al., 2006) is one such system applying different weights to different CAT diagnostics to diagnose CAT of different strength categories. GTG has been evaluated using ROC curves in Sharman et al, (2006). However, such an evaluation hasn't been performed on the injury causing events which by their nature are likely to be CAT of severe or extreme strength.

As discussed above CAT is already an issue but research by Williams & Joshi (2013), Williams (2017) and more recently Storer et al. (2017) and Lee et al. (2019), suggest that climate change is likely to increase the volume of CAT in the atmosphere making the problem even worse. The work by Storer et al. (2017) compared a heavy climate change scenario (RCP8.5) out to 2050-2080 with a pre-industrial model run and found a doubling to tripling of CAT in the midlatitudes. As of 2024, humanity has managed to increase the concentration of carbon dioxide in the atmosphere by approximately 50% relative to the pre-industrial, which leads to the question: How much has cruise altitude CAT increased already?

The satellite era is now longer than 40 years making it sufficiently long enough to detect significant trends from noise (E.g. Lee et al., 2023). Reanalysis (such as ERA5) can now be interrogated to find the answer to this question.

Both flying through or around turbulence requires addition fuel, therefore, if a seasonal forecast of CAT were possible then it would enable the aviation industry to buy in the fuel in advance when it is potentially cheaper serving as a means of climate change adaptation. The ERA5 reanalysis dataset can also be analysed with reference to sources of climate variability to see if any useful patterns that might enable a seasonal forecast might emerge.

1.2 Thesis aims

Aims related to the evaluation of CAT diagnostics

Work by Williams and Joshi (2013) have utilised an ensemble of 21 CAT diagnostics (See Appendix A for equations) to diagnose aviation affecting CAT both regionally and globally in climate models, but in conducting their analysis have implicitly assumed that each diagnostic has equal CAT diagnosing skill. Sharman et al. (2006) in their work developing the Graphical Turbulence Guidance (GTG) system do attach different weightings to different CAT diagnostics in different circumstances to optimise the overall CAT diagnosing ability for different strength thresholds, however GTG is based on pilot reports (PIREPs) which have known limitations (Schwartz, 1996; See section 2.1.3 for a discussion of these), and not on injury causing turbulence encounters. Here we aim to evaluate these 21 CAT diagnostic's ability to diagnose severe-or-greater (SOG) CAT and then on the basis of this, aim to explore what implications this has for previous CAT projections.

1.2.1 Aims related to the global distribution and changes in cruise-level CAT

Work by Jaegar and Sprenger (2007) produced maps of 4 different CAT diagnostics in both December/January/February (DJF) and June/July/August (JJA) at the 2 Potential Vorticity Units (PVU) tropopause level in the northern hemisphere using ERA-40. While interesting, global maps of diagnostic

averaged annual CAT at cruise altitude based on more modern reanalysis such as ERA5 have so far not been produced although Lee et al. (2023) have examined NH midlattitude CAT using ERA5. This thesis aims to produce such maps and once produced, aims to examine how this diagnostic averaged CAT has changed over the 40 year satellite era globally, in the midlatitudes and in any particularly note-worthy regions. Once these broad pictures have been ascertained, a breakdown by diagnostic, season and CAT strength level will be conducted.

1.2.2 Aims related to the impact of climate variability on CAT

Work by Cheyne (2020) amply demonstrated that ENSO has a dramatic impact on the global distribution of vertical wind shear (VWS), a diagnostic of CAT. This begs the question of whether other sources of climate variability might impact on CAT. This thesis aims to extend the analysis of Cheyne 2020 to examine the diagnostic average impact of ENSO (as well as other relevant sources of climate variability such as the North Atlantic Oscillation) on cruise altitude CAT in order to understand which areas are affected the most and in what direction the relationship is in. This work builds on Cheyne (2020) by considering 21 diagnostics instead of just VWS, uses the more modern ERA5 instead of NCEP-DOE Reanalysis 2 and examines the top 3 % of the CAT diagnostic distributions at 5 turbulence intensities as opposed to Cheyne (2020)'s diagnostic average at a single turbulence intensity (MOG). Once this analysis has been conducted, a multiple linear regression (MLR) will be performed using the relevant sources of climate variability in order to predict CAT prevalence.

1.3 Thesis structure

This thesis will be structured as follows:
Chapter 2 will present the current literature related to the following aspects: the preferred locations and times CAT tends to occur in the atmosphere along with the mechanisms that generate it; the changes in circulation that have occurred over the last 40 years that are relevant to CAT generation; how CAT is forecast; the ERA5 reanalysis dataset; the impact of climate change on CAT and aviation; and finally the impact of climate variability on CAT. Then chapter 3 will use commercial aircraft CAT encounters to evaluate the skill of 21 CAT diagnostics at diagnosing injury causing CAT. Chapter 4 will analyse global changes in 197 hPa CAT over the past 40 years and chapter 5 will examine the effect of ENSO and the NAO on global aviation altitude CAT. Chapter 6 will offer a conclusion to the thesis, suggest limitations and opportunities for future work.

Chapter 2
Literature Review

2.1 Where and why does Turbulence occur in the Earth's atmosphere?

2.1.1 What is turbulence?

Chaotic turbulent flow is often contrasted with ordered laminar flow and occurs at all length scales in Earth's atmosphere (Williams et al., 2017).

The best succinct, elucidating discussion of a phenomenon as complex as turbulence in my opinion is given on p207 of Salmon (1998), summarised below:

1. Presence of vorticity
2. Contains a complex structure, broad range of time/space scales.
3. Contains a high degree of apparent randomness and disorder, but also contains ordered embedded flow structures
4. 3-dimensional (unless constrained by strong rotation or stratification) and a high rate of viscous energy dissipation
5. Rapid mixing of advected tracers
6. Often exhibits high levels of intermittency

Salmon adds to these 6 points a final property that is arguably the most important, namely that turbulence is chaotic (sensitive to initial conditions) and thus requires a statistical rather than a deterministic treatment.

2.1.2 Types of atmospheric turbulence

Atmospheric turbulence can be delineated into different types (Clear-Air Turbulence [CAT], Convectively induced Turbulence [CIT], Near-cloud Turbulence [NCT], Mountain-wave Turbulence [MWT], Boundary-layer Turbulence [BLT]) based predominantly on occurrence locations (such as in clear air, clouds, the vicinity of clouds, mountains and the boundary layer).

However another method of classification would be to focus solely on the causative mechanisms in the nomenclature. For example, types of turbulence could be: Shear induced turbulence (SIT), convectively induced turbulence (CIT) and wave induced turbulence (WIT). The advantage of this system is that the somewhat artificial divide between NCT and MWT disappears as they are often both examples of WIT.

Another term that bears introducing is 'aviation-affecting turbulence'. This is important as turbulence of a scale either much larger or smaller than the length scale of an aircraft (~100m), won't substantially impact on it (Sharman et al., 2012).

When classifying 'aviation affecting turbulence events', any of the above labels may also be a simplification as the observed turbulence isn't always generated by a single distinct cause. A so-called CAT event might be the product of both convection and jet stream wind shear together, or so-called MWT might be the result of both wind shear and gravity waves.

CAT (or SIT) is a type of turbulence that is strongly associated with jet streams and frontal systems, although it can also occur in and around the tropopause (Jaeger and Sprenger, 2007).

2.1.3 Is it even known where the turbulence in the atmosphere actually occurs?

Currently, a lot of knowledge about where and when turbulence occurs in the atmosphere is provided by 'Pilot Reports' (PIREPs). Pilot Reports is a report of actual flight conditions encountered by an aircraft.

While these have undoubtedly contributed much to the understanding of the temporal, geospatial and intensity of turbulence, there are known biases associated with them as discussed by Schwartz (1996).
One of the largest problems is that PIREP locations are strongly tied to the aviation flight routes and as such are not providing representative samples of the entire atmosphere both in the horizontal (at cruise altitude) but particularly in the vertical (at non-cruise level altitudes). This tends to mean that aviation meteorologists are often more confident about turbulence at cruise level over the US, Europe, and eastern East Asia than they are over the rest of the atmosphere. However, put another way the fact that PIREPs are produced in the regions aircraft fly is useful to the aviation industry and can therefore be seen as an advantage.

Apart from the aforementioned geographical sampling bias, there are other sampling biases too. Pilots don't tend to produce many Null Turbulence PIREPs, i.e. a PIREP which explicitly states that no turbulence was observed and therefore the assumption is often made that a PIREP which doesn't specifically mention turbulence, counts as an implicit Null Turbulence PIREP. Whilst Null turbulence is likely to be under reported in the areas of high flight density, areas of heavy turbulence may also be under-represented in the PIREP data as pilots where possible will attempt to avoid areas of the sky suspected of being highly

turbulent. This could be because of a previous turbulence forecast warning, or because earlier aircraft on the same route encountered turbulence. This will result in under-sampling, especially if the turbulence forecasts the pilots use continue to become increasingly skilled at predicting turbulence. To illustrate this, imagine having a perfect turbulence forecast. PIREPs data would suggest the atmosphere is turbulence free, when this needn't be the case at all. Schwartz also suggests that there is a disincentive for aircraft to report severe turbulence or report it as moderate instead as severe turbulence requires a mandatory grounding and inspection.

It is difficult to ascertain the extent to which the various problems Schwartz point out add up to a problem as significant as Schwartz appears to be at one point quoting Reap (1996) "Thus, the forecasts are not probabilities of detecting an event but are probabilities of detecting a report of an event". This problem is perhaps greater with turbulence forecasting that is limited to PIREPs for its verification since other data sources such as weather balloons, aircraft and satellite observations aren’t suitable.

If all this wasn't enough, there is some degree of uncertainty in both the spatial and temporal coordinates of PIREPs which limits their usefulness in certain contexts where high precision is important (e.g. with Near-cloud Turbulence [NCT] and Convectively induced turbulence [CIT] studies) and the reported severity of turbulence is not only a function of aircraft size but also of the individual pilot making the PIREP.

Newer sensors attached to the aircraft hull which directly measure a more objective acceleration (Eddy Dissipation Rate [EDR] and derived equivalent vertical gust [DVG]) rather than ‘subjective bumpiness’ may eliminate some of the aforementioned problems (Sharman et al., 2012). For one thing such EDR estimates are, at least in principle, independent of both the aircraft size/model and pilot. The data collected is continuous rather than ordinal and under-reporting of null turbulence would be less of an issue. The 'avoidance of potentially highly turbulent areas' issue would still remain as would the aircraft route selection bias issue. Since the processes which generate EDR are sub-grid scale, it cannot be calculated from ERA5.

While it is important to bear all these issues in mind, and very easy to point out all the flaws, PIREPs have clearly taught us a great deal about where and when turbulence occurs at least along common flight routes, if not the entire atmosphere. New observational techniques (such as those discussed above) will hopefully produce more accurate, precise, and therefore useful data which will hopefully further deepen the scientific understanding.

2.1.4 Where CAT occurs geographically

When asking where CAT occurs within the atmosphere more generally, Kaplan et al. (2005) draw our attention to convection. 86% of the 44 severe turbulence events they studied were within 100km of

convection. This could suggest that CIT rather than CAT was the predominant type of turbulence across the US, however the proportion, verging on 9/10 is quite high, perhaps suggesting that CIT contributes to turbulence rather than is the sole cause. In addition, they note that the combination of convection and a curved jet stream (upward vertical motions, low relative vorticity, horizontal cold air advection, and leftward directed ageostrophic flow) was a good predictor combination for a severe turbulence event. All this adds weight to the idea that convection may, at least to some extent, be behind turbulence labelled CAT occurring in the vicinity of the jet stream. Not everyone agrees with the aforementioned statistic of Kaplan et al. (2005), however, Meneguz et al. (2016) find that only 14% of aircraft encounters with turbulence occur in the vicinity of convection. Apart from the different methodologies employed (Unlike Kaplan et al., Meneguez et al. Use 4s automated aircraft data) the two studies look at different regions (Meneguez et al. examine the north eastern Atlantic and Europe and Kaplan et al. Examine the US) which may well differ in relation to the prevalence of convection.

2.1.5 The location of CAT in a global latitude-longitude slice

When talking about a curved jet stream it is perhaps sensible to divide it into 4 regimes: Ridge Pole-side (RP), Ridge Equator-side (RE), Trough Pole-side (TP) and Trough Equator-side (TE) as well as 4 additional nodal regions e.g. Ridge to Trough Pole-side (R-TP) etc. The cyclonic/anticyclonic side of the jet stream is synonymous with the poleward/equatorward side.

Curved regions of the jet stream tend to be more associated with CAT (e.g. Ellrod and Knapp, 1992).
Kaplan et al. (2005) who analysed a number of turbulence events writes a lot about turbulence being more prevalent just downstream of flow curvature, although they are slightly non-specific on whether this is more likely to be curvature associated with a ridge or trough. Alternatively, Roach and Bysouth (2002) associate the generation of severe CAT and 'sharp/disrupting troughs' without mentioning whether they were associated with the pole or the equator side.
Kaplan et al. (2005) found a fairly strong correlation between cold air advection and CAT.

Stang (2020) suggests that in a curvy jet, the flow is in gradient wind balance (The Pressure Gradient Force [PGF] has to be balanced with the coriolis and centrifugal) not geostrophic balance. Hence the flow is at maximum speed at the crests (super-geostrophic) and a minimum at the troughs (sub-geostrophic). This means at the descending nodes, the flow is decelerating (hence convergence/sinking) and accelerating (divergence/rising motion) on the ascending nodes. Since divergence/convergence is a diagnostic of CAT, these regions may contain CAT.

Jet streaks, regions of particularly fast flowing air within the jet stream, are often associated with curved regions of the jet stream (e.g. Kaplan et al., 2005). Sharman et al. (2012) declares jet exit regions as preferred regions for inertia-gravity waves (IGW) generation (See also Uccellini and Koch, 1987 and Guest et al., 2000) which can cause CAT either in-situ or further afield and does not mention 'entrance regions' in

their review paper at all. However, Kaplan et al. (2005) focuses on jet entrance regions (34 mentions for 'entrance' vs just 3 for 'exit') which are connected with leftward directed ageostrophic flow and the consequent convergent/descending on the poleward side and divergence/ascending on the equatorward side that goes with this. For Kaplan et al. (2005), convection coinciding with a jet entrance region on a curved part of the jet stream are a combination of factors that make turbulence likely in their estimation.

Koch et al. (2005) note that moderate-or-greater (MOG) turbulence occurring on the cyclonic/poleward side of a jet streak possibly in relation to a south-eastwards jet-streak. They suggest that if dynamic (shear) instability is the more important source of turbulence then you might expect said turbulence to occur at the troughs and ridges, whereas if static instability was more important then turbulence should occur at the nodes midway between the crests and troughs.

Knox (1997) comment that CAT in anticyclonic/equatorward flow is rarer and consequently hasn't received the same research attention as cyclonic/poleward flow. In a later paper (Knox et al., 2008), they link cyclonic/poleward flow with Kelvin-Helmholtz instability (KHI) via frontogenesis and deformation (See also Dutton and Panofsky, 1970 and Ellrod and Knapp, 1992) and anticyclonic/equatorward flow (regions of zero or negative potential vorticity) with strong gravity wave (GW) generation. Adding 'divergence tendency' to diagnostics may improve diagnostic power in such anticyclonic/equatorward regions (Knox et al., 2008).

Koch et al. (2005) mention that aircraft measurements have revealed GW (with wavelengths between 2-40km) emanating transverse to the flow at jet stream levels (See also Shapiro 1978, 1980; Gultepe and Starr 1995 and Demoz et al. 1998), with implications for where CAT gets generated.

Ellrod and Knapp (1992) suggest that the confluence zone between the sub-tropical and sub-polar jet is a favourable location for CAT (See also Reiter and Nania, 1964). Kaplan et al. (2005) echo this by their observation that the polar and subtropical jet entrance region both coincided over the launch site of the ill-fated Challenger rocket disaster of 1986 (See also Uccellini et al., 1986). Such conditions produce very large wind shears conducive to wave breaking and KHI. Roach and Bysouth (2002) examining severe CAT over the tropical Pacific Ocean note that while unusual for CAT to occur closer to the equator than the latitude of the typically very stable sub-tropical jet (~30 degrees), CAT does occur as close as 10 degrees to the equator in instances where the subtropical jet stream buckles and folds. A scenario they give for this to occur is when the subtropical and polar jets merge over Japan in winter. The resulting diffluent flow means that CAT is even more likely downstream over the Pacific Ocean.

2.1.6 Looking at the oncoming jet stream in a meridional slice

Roach and Bysouth (2002) and Wolff and Sharman (2008) suggest that turbulence is prevalent both above and below the jet core. This is echoed by Fig 17 in Shapiro (1976) shown in Figure 2.1

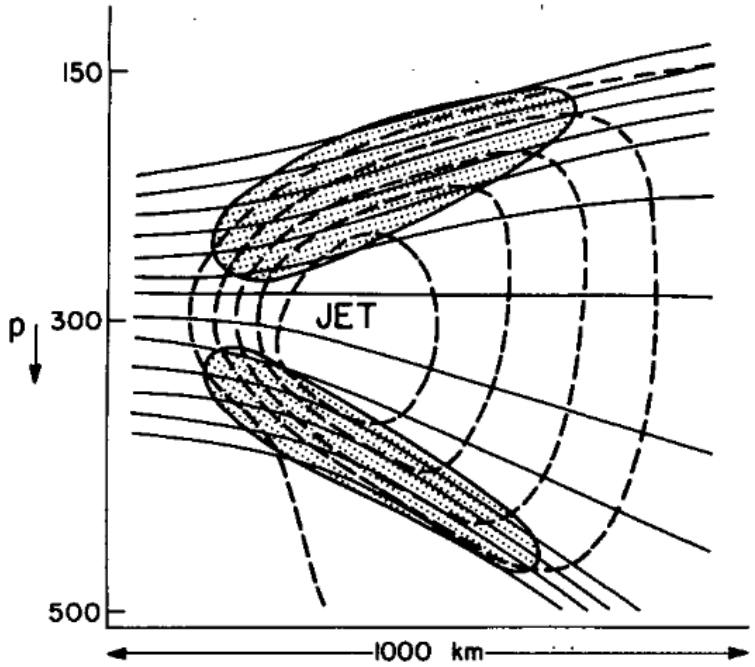


Figure 2.1: Regions of clear-air turbulence, stippled, in the vicinity of a jet stream/frontal zone system. Potential temperature, solid lines; wind speed, dashed lines. Taken from Shapiro (1976).

which shows regions of CAT above and below the jet core, slightly tilted towards the poleward side. Sharman et al. (2012) talks about an instance where CAT occurs above the jet core due to IGW activity (See also Lane et al. 2004 and Koch et al., 2005). CAT occurring beneath the core may also be connected with tropopause folds as these are found beneath the jet cores (Stefan et al., 2020) and contain lots of shear and mixing.

Jaeger and Sprenger (2007) and Wolff and Sharman (2008) associate CAT with the tropopause (both in and nearby). The latter reports that the altitude of 9km has 1.7 times the background rate of turbulence and the rate within 1 km of the tropopause (median height of 11.6 km) is even higher than this. Stefan et al. (2020) hints that it is tropopause folding rather than the tropopause that could be behind the aviation affecting CAT (See also Shapiro, 1980 and Kim and Chun, 2010). Tropopause folds are intrusions of stratospheric air that occur under both the polar and tropical jets. If tropopause folds are a source of CAT, then they have the advantage that they can be observed using water vapour satellite imagery.

Kaplan et al. (2005) looking at 44 severe turbulence encounters across the US conclude that turbulence (not CAT per se) was most common in the 9-12km altitude range.

2.1.7 CAT in different geographic regions

One way to get around the global sampling errors of PIREPs is to use reanalysis data which is precisely what Jaeger and Sprenger (2007) did when constructing their 44-year (1958-2001) Northern hemisphere (NH) climatology of 4 CAT indices (Richardson number, Potential Vorticity, N^2 and Turbulence Index 1 [TI1]) using ERA40. When looking at their 2 PVU (tropopause altitude) figures it is important to bear in mind that:

1/ They are showing the geographical distribution of turbulence diagnostics rather than turbulence itself (e.g. EDR).

2/ even if a more objective measure of turbulence such as EDR were shown, EDR is not necessarily the same as aviation affecting turbulence as it is the rate at which turbulent energy dissipates and is a function of all turbulent length scales not just those that affect aviation, and

3/ Aircraft may fly above or below the tropopause depending on latitude.

Even though there is often disagreement between the diagnostics themselves over the turbulence hot spots, there are areas of overlap in the spatial distribution of diagnosed turbulence in their figures 2 and 3 (see figure 2.2 and 2.3 respectively).

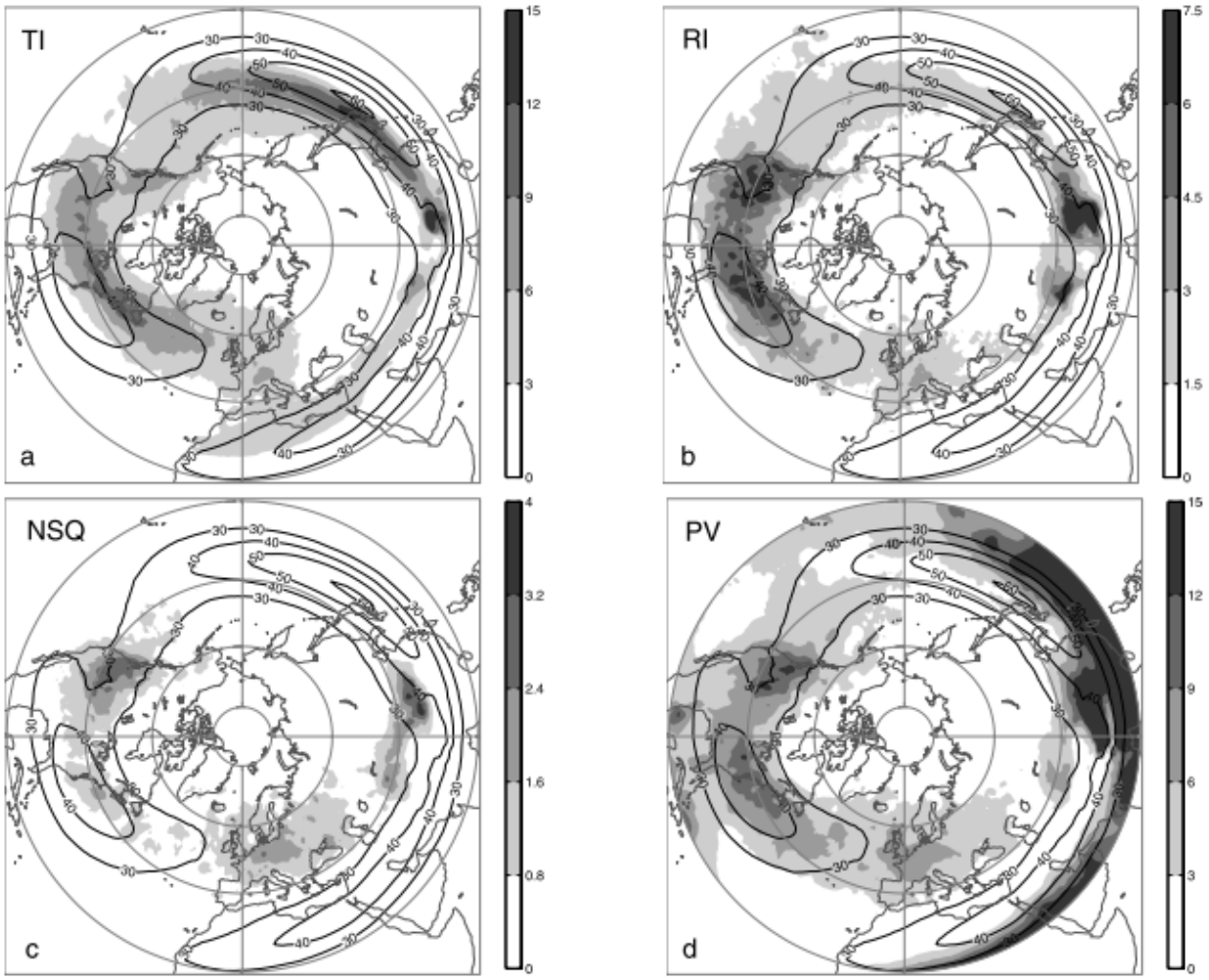


Figure 2.2: Winter means (December, January, and February) of frequencies (%) of high TI and low Ri, N2, and PV near the tropopause. Mean wind speed (m s^-1) is plotted as contour lines (every 10 m s^-1 for

wind speeds >30 m s^-1). Mean covers the winters from 1958 to 2001. Note the different grey scales (Caption copied from Jaeger and Sprenger (2007).

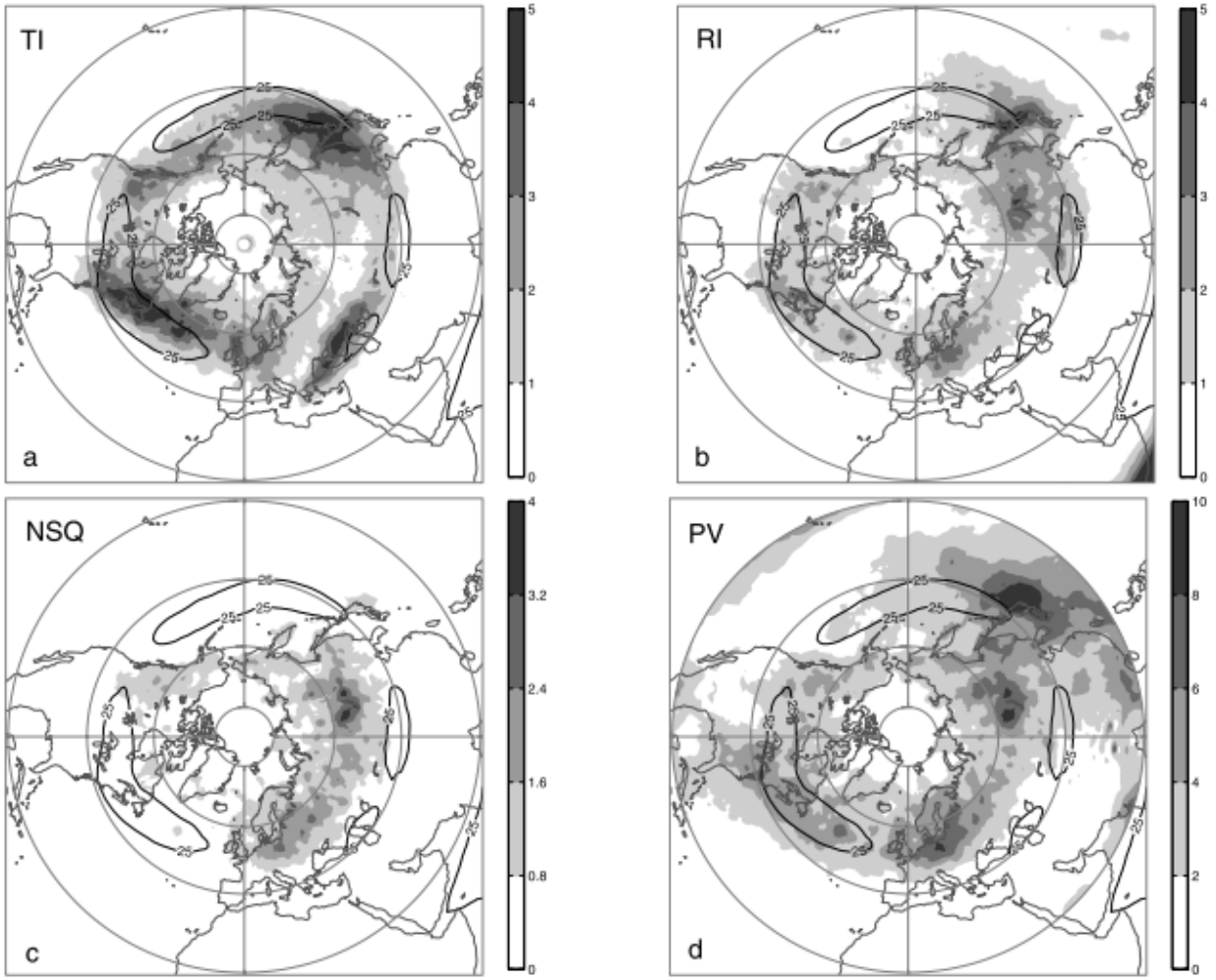


Figure 2.3: As in Figure 2.2, but for summer (June, July, and August). Wind speeds larger than 25 m s^-1 are shown. Note the different grey scales compared to Figure 2. (Caption copied from Jaeger and Sprenger (2007).

In winter (DJF) the diagnostics agree that NH turbulence is generally a ring of turbulence that circles the Earth straddling a latitude of ~40 degrees (though Potential Vorticity [PV] also indicates a band at 20 degrees that stretches from the west coast of Africa all the way to Japan). There is also agreement that western Europe, northern China, and the northern USA are hotspots. In summer (JJA) the aforementioned 40-degree band moves northward to 50 degrees and apart from western Europe, there is less agreement on the location of turbulence between the 4 diagnostics. The bands of turbulence are incidentally, roughly 7 degrees north of the location of the subtropical jet (~33 degrees in DJF and 42 degrees in JJA).

From a longitudinal, pressure level-latitude slice perspective (Fig 4 in their paper), Ri, TI and N^2 maxima seemed to occur to the groundward, poleward side of the subtropical jet independent of season and longitude slice, which could potentially indicate that a polar trough dipping down far enough to merge with the subtropical jet might be a significant cause of the higher diagnostic values and possibly real world CAT too.

Although the 40–50-degree latitude band of CAT is a somewhat simplified version of the more longitudinally varying CAT diagnostic distributions shown in figures 2 and 3 of Jaeger and Sprenger (2007), assuming for a moment that this is where the bulk of the CAT exists, how does it compare with aviation flight routes? Is the overlap large or minimal?

The 3 big hubs for aviation activity are the continental US (~30-40-degree latitude), Europe (~37-57 degrees latitude) and East Asia (~22-42 degrees latitude)
If we draw such a band of CAT spanning the NH globe between ~37-53 degrees latitude, then the US and East Asia routes appear to be half in and half out at the lower end and similarly with the North Atlantic flight corridor (~46-56 degrees latitude) poking out at the upper end. Europe more or less straddles this CAT band. While only 4 diagnostics are being applied, and diagnostics are not necessarily the same thing as turbulence, the suggestion is that the Venn circle of 'CAT in the cruise level atmosphere' and that of 'aviation flight routes' only half overlap, at best. It might also be the case that the US and East Asia might experience more CAT in the NH winter as the diagnosed turbulence band shifts south, whereas the North Atlantic (NA) flight corridor might be more prone to turbulence in the summer when this band shifts north. However substantially larger datasets would be needed to test this hypothesis than we currently possess.

As mentioned before we don't have good global coverage of severe turbulence aviation encounters, however Kaplan et al. (2005) managed to find and examine 44 severe turbulence events (1990-1996) over the United States. Although these events aren't limited to CAT the average latitude of events that occurred in summer is ~35.1 degrees latitude (N=26) and 32.8 for winter (N=18) which although only significant at the p=0.25 level, is consistent with the CAT band moving northward in summer as shown in Figs 2 & 3 in Jaeger and Sprenger.

Jaeger and Sprenger note a maximum over Greece/Turkey in the NH summer potentially related to tropopause folding. This also chimes with the work of Stefan et al. (2020).

2.1.8 The causes of CAT/turbulence

The main mechanism by which CAT is generated is thought to be Kelvin-Helmholtz instability (KHI) (Ellrod and Knapp, 1992) which in experimental settings has been found to occur in regions where the Richardson number (Ri) is less than 1/4.

This occurs when either the degree of vertical stratification (numerator) is low or the vertical wind shear (denominator) is high or both, conditions prevalent near jet streams. Indeed, near the jet stream the probability of CAT is about three times larger than in areas of weak winds (Reiter 1963c).

Gravity waves (GWs) and Inertia-gravity waves (IGWs) caused by processes which disturb the equilibrium of air masses (e.g. topographic forcing, convection etc etc) can also induce CAT either by

breaking/overturning in low Richardson number environments or by locally (for example, in crests and troughs) reducing an otherwise stable environmental Richardson number (Knox et al., 2008 and McCann et al., 2012).

McCann (2001) explains that low (but >¼) Richardson number environments need minimal GWs to result in turbulence, while even frenetic GW activity might not be enough to precipitate turbulence if the Richardson number is much higher than the critical threshold.

2.1.9 Richardson Number (Ri)

Although described above as a favourable predictor for CAT, in practice the Richardson number has a number of problems.

Ellrod and Knapp (1992) point out 3 in particular. Firstly, sparse vertical wind-shear data to calculate Ri from can make the final value inaccurate in practice and although vertical resolutions have since increased since, there is still further to go in this regard. Secondly Ri values can often be >0.25 during actual CAT events and thirdly, low Ri numbers seem to occur in both smooth and turbulent regions. Kaplan et al. (2005) also point out that for severe turbulence in particular, Ri values and the strength of the jet stream are not well correlated.

Deformation zones (which can be found along certain cloud boundaries and regions of hyperbolic flows e.g. jet exit regions and cols, regions of slightly elevated pressure between two anticyclones) are thought to maintain regions of low Ri number (Ellrod and Knapp, 1992).

2.1.10 Vertical wind shear (VWS)

The connection between CAT and VWS was emphasised in a 1970 Dutton and Panofsky paper.
Even in contrast to stratification, VWS is perhaps up on a pedestal in regards to the generation of CAT.
This is because whatever value you have for the stratification numerator of the Ri number equation, turbulence can always happen if you increase the VWS high enough. However if the VWS is zero, then no matter what happens to the stratification numerator (unless it is zero), turbulence cannot happen. Therefore although stratification plays a role in how easily turbulence takes place, it is the VWS that is of critical importance.

Indeed, out of the ensemble of 21 diagnostics first used in Williams and Joshi (2013), 9 have the vertical shear visible in their equations (CAT-d01: Negative Richardson number, CAT-d02: Magnitude of vertical shear of horizontal wind, CAT-d03: Colson--Panofsky index, CAT-d06: Brown energy dissipation rate, CAT-d07: Variant 1 of Ellrod's turbulence index, CAT-d08: Variant 2 of Ellrod's turbulence index, CAT-

d19: Version 1 of North Carolina State University Index and arguably CAT-d04: Frontogenesis function and CAT-d14: Wind speed times directional shear; See Appendix A for diagnostic equations).

Knox (1997) considers the 'enhancement of vertical shear' essential (the desideratum) for CAT production and argues that whether it is frontogenetical/deformation based or from geostrophic imbalance or inertial instability, it is vertical wind shear that leads to the KHI.

McCann (2001) also point out that VWS, apart from causing KHI can also create or amplify GWs that then go on to cause turbulence when breaking.

2.1.11 Gravity waves/Inertia Gravity waves (GW/IGWs)

As mentioned above, depending on how low the environmental Ri number is, GWs/IGWs can precipitate/trigger turbulence either by locally reducing the Ri number prior to breaking/overturning or upon breaking/overturning itself (McCann, 2001). In the mind of Koch et al. (2005), IGWs are generated 'transverse' to the flow at jet stream levels from regions of flow imbalance via geostrophic adjustment. McCann (2001) agrees that GWs are generated when a loss of geostrophic balance occurs. In this instance, flow imbalance involves frontogenesis (Wolff and Shaman also link frontogenesis to IGW generation) and is associated with a jet streak (Kaplan et al. (2005) also link GWs with jet streaks). Koch et al. (2005) note that it is only GWs of a certain horizontal wavelength (1–20 km) that are correlated with aviation affecting CAT/KHI.

McCann (2001) reminds us that unbalanced flow are not the only cause of GWs (convection can also cause GWs) and GWs are not the only cause of CAT. Koch et al. (2005) comment that the relationship between gravity wave instability and turbulence is not well established in the troposphere. However, Knox et al. (2008) went on to find a more well-established link and associate the advection of relative vorticity with IGW generation.

GWs may also be the product of KHI which may generate a feedback loop helping to maintain the CAT (Roach and Bysouth, 2002).

2.1.12 Convection

Although convection causes atmospheric turbulence in its own right (often experienced by aircraft when flying over the ITCZ), a number of references suggest that Shear Induced Turbulence (SIT) and Convectively Induced Turbulence (CIT) may be less delineable than commonly thought. Kaplan et al. (2005) notes that ~86% of the 44 severe turbulence events occurring across the US occurred within 100km of moist convection (any type of convective cloud). If true/representative, then this would suggest that either the bulk of severe turbulence across the US might be CIT, not Shear induced turbulence (SIT) related

or that CIT is more commonly playing a role in SIT than previously thought. As the Richardson number (See Appendix A – D01 for equation) is made up of both a vertical shear component as well as a buoyancy component it stands to reason that any moist convection encountering vertical shear would be particularly susceptible to turbulence. Kaplan et al. (2005) note that in particular, convection occurring at a jet stream entrance region was the most important synoptic signal for a severe turbulence event. Schwartz (1996) from a study looking at 454,324 PIREPs, split the US up into 6 parts (NW N-Central, NE, SW, S-Central and SE) and finds that a greater proportion of all MOG and severe events occur in the south than the north of the country (12.7%, 10.8%, 13.2%, 18.5%, 23.7% 15.4% respectively for MOG and 8.1%, 7.1%, 12.5%, 20.9%, 22.5% and 23.3% respectively for severe). These results seem to indicate that aviation affecting turbulence within the US is more prevalent in the south suggesting that convection is a bigger player than CAT/SIT. Schwarz (1996) seems to then conclude that CAT must be being under sampled yet doesn't really seem to demonstrate this conclusion adequately. This along with the '~86% within 100km of moist convection' assertion of Kaplan et al. (2005) suggests CIT is the bigger player here of the two. Storer et al. (2019a) reminds us that the polar jet gets nearly all of its magnitude from the anticyclone generated from latent heat release (See also Trier et al., 2012 and Trier and Sharman, 2009). Storer et al. (2019a) comments that this shouldn't be viewed as CIT though as the turbulence is an indirect rather than a direct result of convection. Sharman et al. (2012) points to the same Trier et al. (2012) study again indicating that convection may be behind supposed CAT events (See also Knox et al., 2008).

2.1.13 Frontogenesis, frontolysis, deformation and CAT

A highly deformed flow pattern is one where either the shearing and/or stretching deformation is large and is a function of zonal wind velocity (u) and meridional wind velocity (v).

Flow deformation =

$$\left(\left(\frac{\partial \mathrm{v}}{\partial \mathrm{x}}+\frac{\partial \mathrm{u}}{\partial \mathrm{y}}\right)^{2}+\left(\frac{\partial \mathrm{u}}{\partial \mathrm{x}}-\frac{\partial \mathrm{v}}{\partial \mathrm{y}}\right)^{2}\right)^{0.5} \tag{2.1}$$

Frontogenesis (Fr) is the process whereby a front between a cold and warm air mass is intensified and is a function of zonal wind velocity (u), meridional wind velocity (v), vertical wind velocity (w) and potential temperature (θ).

$$\frac{\partial}{\partial t}(Fr)=\frac{\partial}{\partial \mathrm{x}}\frac{\partial \theta}{\partial \mathrm{t}}-\frac{\partial \theta}{\partial \mathrm{x}}\frac{\partial \mathrm{u}}{\partial \mathrm{x}}-\frac{\partial \theta}{\partial \mathrm{y}}\frac{\partial \mathrm{v}}{\partial \mathrm{x}}-\frac{\partial \theta}{\partial \mathrm{z}}\frac{\partial \mathrm{w}}{\partial \mathrm{x}} \tag{2.2}$$

Greater diabatic heating on the warm side of the front than the cold, convergent, cyclonic horizontal vorticity and poleward rotating vertical vorticity 'tilting' contribute to frontogenesis. Frontolysis is the opposite process whereby the temperature gradient between two air masses weakens.

Mancuso and Endlich (1966) talked about a relationship between turbulence and the destruction or creation of wind and temperature gradients i.e. frontogenesis and frontolysis.

CAT seems to occur along the boundaries of certain clouds which happen to be deformation zones. Such deformation zones are found in regions of hyperbolic flow (e.g. cols and jet exit regions).

A number of authors see CAT as a result of the deformation regions of frontogenesis (e.g. Mancuso and Endlich (1966); Haman et al. (2011)) but Jaeger and Sprenger (2007) comment that it is deformation zones rather than frontogenetical processes that are key here. This distinction is worth making because deformation fields can also be the result of frontolysis.

Knox (1997) and Jaeger and Sprenger (2007) also note the frontogenesis-deformation-CAT pathway but think it is only appropriate in cyclonic flows and not in anticyclonic flows because deformation is small in such instances yet the VWS can be very large, even larger than in straight or cyclonic flows.

So why is deformation itself linked with CAT? Ellrod and Knapp (1992) suggest it is because large scale stretching and shearing deformation is important for maintaining the Ri number below the critical value (deformation strengthens gradients which are positively associated with turbulence. Roach and Bysouth (2002) echo this saying that CAT is maintained by stretching/deformation of the horizontal flow (and shear advection). Knox (2001) tie frontogenesis with vertical wind shear.

Apart from frontogenetic deformation's role in lowering the Ri number, Koch et al. (2005) and Wolff and Sharman (2008) note a connection between frontogenesis and GWs or IGWs with the former noting the GWs emanating transverse to the regions of frontogenetical flow imbalance via geostrophic balance adjustment.

Frontogenesis via cold air advection (Kaplan et al., 2005) at a jet stream trough may be part of the reason why CAT is associated with a curvy jet stream.

As to the link between frontogenesis and GWs, Roach and Bysouth (2002) comment that frontogenesis increases wind shear and wind shear, if large enough, can create or amplify GWs. This combination of wind shear and GWs results in low Ri number and often, turbulence.

2.1.14 Horizontal Cold air advection

Horizontal cold air advection is talked about a lot in Kaplan et al. (2005) and seems to be quite linked with frontogenesis. In fact they talk about cold air advection along with curviness in particular being key to turbulence.

"The inference being that since cold air advection follows a cold front, particularly if flow curvature is significant, hence, some combination of curvature and solenoidal/cold frontal structure is the key to understanding what establishes an environment predisposed to turbulence."

2.1.15 Horizontal wind shear

Even though Colson and Panofsky (1965) say that horizontal wind shear correlates badly with CAT and much of the emphasis in CAT theory is on vertical wind shear, 14/21 of the CAT diagnostics of Williams and Joshi (2013; See Appendix A for diagnostic equations), include at least one horizontal wind shear term (du/dy or dv/dx). 7 of these 14 are through having flow deformation (the equation of which is composed entirely of horizontal shear terms) in the equation (d05: Brown index, d06: Brown energy dissipation rate, d07: Variant 1 of Ellrod's turbulence index, d08: Variant 2 of Ellrod's turbulence index, d09: Flow deformation, d15: Flow deformation times wind speed and d16: Flow deformation times vertical temperature gradient), five vorticity based ones (d10: Magnitude of potential vorticity, d11: Relative vorticity squared, d19: Version 1 of North Carolina State University Index, d20: Negative absolute vorticity advection and d21: Magnitude of relative vorticity advection) as well d04: Frontogenesis function and d17: Magnitude of residual of nonlinear balance equation.

2.1.16 Geostrophic departure

Mancuso and Endlich (1966) first noticed that divergence and geostrophic departure may be correlated with turbulence.

Knox (2001) tries to delineate CAT which occurs in cyclonic/poleward flow which is linked with frontogenesis/deformation from CAT generated from the 'gradient balanced' anticyclonic flow caused by strong vertical shear resulting from geostrophic adjustment/inertial instability. McCann (2001) once again remind us that loss of geostrophic balance is not necessarily the only cause of GW and the process by which loss of geostrophic balance produces GWs isn't well understood. In any case there must be validity in this delineation/theory as the Upper Level TURBulence (ULTURB) turbulence diagnostic based on it seems to be more skilled than GTG2 despite being only a single CAT diagnostic (McCann et al., 2012). Sharman et al. (2012) comments that in some instances CAT thought to be caused by spontaneous adjustment might be in fact due to moist convection (See also Trier et al., 2012 and Knox et al., 2008).

2.1.17 Vorticity

Knox et al. (2008) expect GW generation in regions of relative vorticity advection because this is the leading order term in a scale analysis of the Lighthill-Ford radiation terms.
In anticyclonic flows, deformation is small, the VWS (mainly the ageostrophic component) can (especially at the limit of the gradient balance e.g. upper level ridges) greatly exceed that for straight or cyclonic flows. These conditions coincide with weak or negative absolute vorticity, which itself is associated with inertial instability.

Shapiro (1978) related gradients of potential vorticity to CAT and Kaplan et al. (2005) has done the same with relative vorticity. A relationship is believed to exist between CAT and strong negative absolute vorticity advection (e.g. Sharman et al. (2006) appendix A). Knox et al. (2008) suggests a parabolic nonlinear relationship between CAT and absolute vorticity with a minimum at intermediate values. Knox et al. (2008) note that the emphasis on CAT forecasting has been on vertical shear rather than horizontal vorticity.

Kaplan et al. (2005) links large horizontal curvature (vorticity) and low vertical vorticity to CAT.

2.2 Changes in circulation/the Jet stream (strength, latitude, altitude, sheer) at 200hPa

As discussed in the previous section a large proportion of CAT in the atmosphere occurs in the vicinity of the jet stream, so understanding its position and intensity are vital for understanding the distribution of CAT.

2.2.1 Latitudinal trends in the jet stream with climate change

The question of whether the jets have been changing latitude with climate change is often best served by breaking down the question by jet stream (e.g. sub-tropical jet [STJ], sub-polar jet [SPJ]), by hemisphere and also by longitude and season.

Seasonality is certainly very important as climatologically both the sub-polar (SP) and sub-tropical (ST) jet streams tend to be weaker (and shift poleward) in the summer than the winter due to a reduced meridional temperature gradient. Sherwood and Nishant (2015) using a radiosonde dataset between 1979-2012 to compute a zonal and annual mean (see their figure 5) of zonal wind, finding both Northern Hemisphere (NH) and Southern Hemisphere (SH) STJs have increased in altitude and moved pole-wards, especially in the NH over this time period.

Starting with the NH and going from west to east, examining trends in ERA5 between 1979-2022, Simmons (2022) finds northward movement of the jet stream region over the USA, an equator-wards shift in DJF, March/April/May (MAM) and JJA of the strongest mean flow region of the eastern North Atlantic and western Europe, an equatorward shift of the maximum mean flow in DJF and MAM in Eurasia, a shift northwards of the STJ in the Middle East and central Asia in DJF (but equator-wards in MAM), a shift northwards in DJF of the jet-stream region of the central and eastern North Pacific and finally a small poleward shift in the jet around Japan. Strong and Davis (2007) examining a different reanalysis data set (National Center for Environmental Prediction- National Center for Atmospheric Research [NCEP-NCAR] reanalysis between 1958-2007) come to a similar conclusion regarding the poleward shift of the DJF STJ over the East Pacific and Middle East and equatorward shift over the Atlantic, increasing its separation from the Eddy-driven jet (EDJ) there.

For the SH, the main jet regions shift south, though by less than 1 degree, generally considered to be caused by ozone destruction over Antarctica (Stendel et al., 2021). Moving from west to east again, the region of strongest flow over South America moves equator-wards. In the South Atlantic, the jet-stream region moves south in DJF but equatorward in MAM, JJA and September/October/November (SON). The annual mean wind maxima shifts equator-wards west of 22W but south east of 18W and to the south of Australia and New Zealand, the flow maximum has shifted equator-wards slightly.

2.2.2 Intensity trends in the jet stream with climate change

As discussed in the previous section, trends in jet stream intensity depends on geographic position and season.

Sherwood and Nishant (2015) (mentioned above) find radiosonde evidence of both subtropical jets increasing in speed by around 1ms^-1 and the Austral SPJ increasing in speed by up to 2ms^1. They find the middle-to-upper tropical jet has an easterly trend (strengthening) of just under 1ms-1 over the time period.

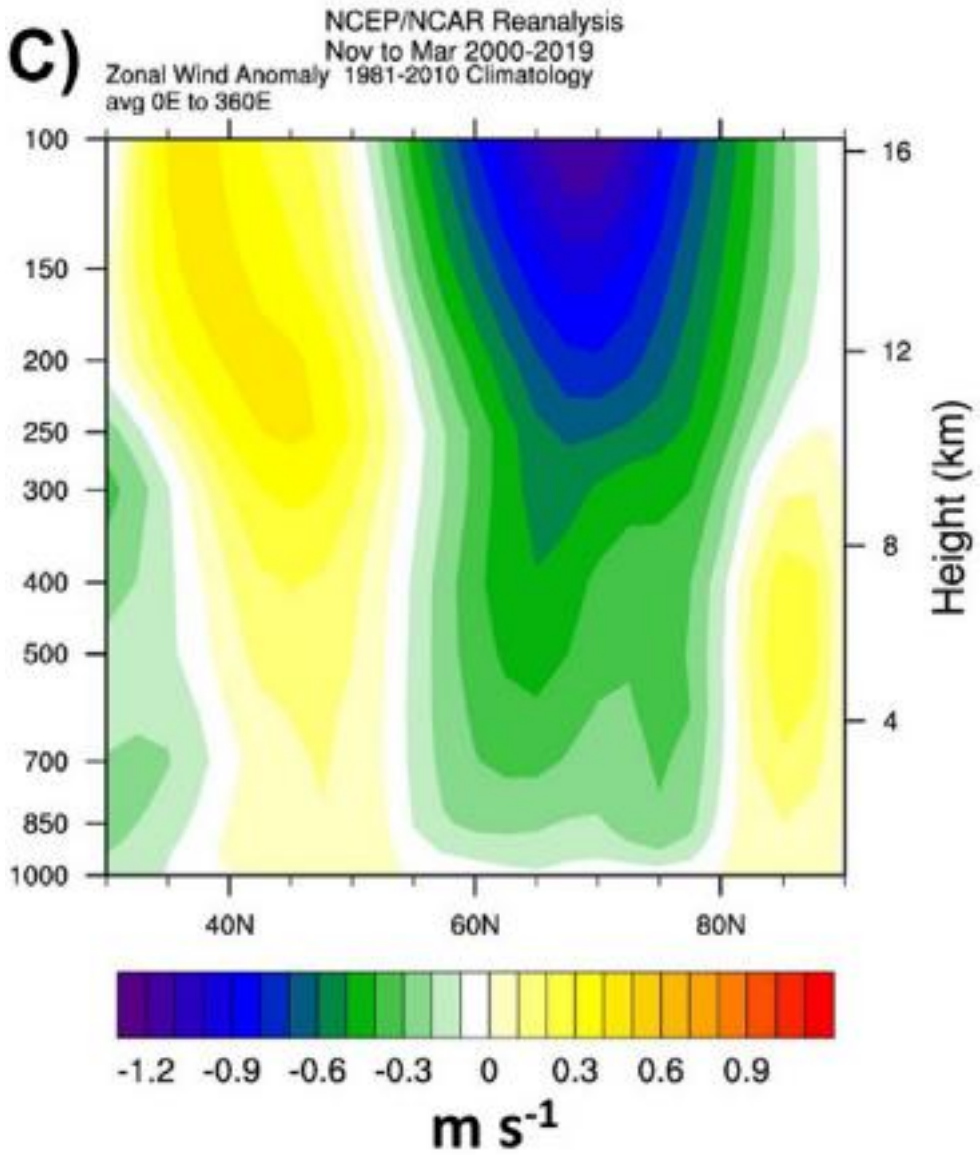


Figure 2.4: Figure 15.10c in Stendel et al. (2021): Zonal winds (C, m/s) for the cold season (Nov/Dec/Jan/Feb/Mar [NDJFM]) during 2000-2019. Anomalies relative to 1980-1999. Data from NOAA/ESRL/PSD.

An interesting figure is Figure 15.10c in Stendel et al. (2021; See Figure 2.4).
They show a latitude by pressure zonal slice of the zonal wind anomalies (2000-2019 mean vs 1980-1999mean) for NDJFM for the NH extra-tropics.

It either shows a strengthening of the westerlies between 40-55N and a weakening of westerlies 55-80N at all altitudes or an equatorward movement of the jet or a combination of the two.

Coumou et al. (2015; in Stendel et al. 2021) finds a slowing of the NH summertime jets whereas Strong and Davis (2007) find increases of up to 15% in the core speeds of the Eddy-driven jet (EDJ) and an intensification of the Atlantic STJ.

For the NH, going from (roughly) north to south, Simmons (2022) finds a strengthening of the westerlies along a zone to the south of the Arctic in DJF between 50-60 degrees latitude, a weakening of the mean flow over Eurasia and north-western Canada in DJF (but a strengthening of the mean flow over Eurasia in

MAM), a strengthening of the region of strongest flow over the central and eastern North Pacific, a strengthening of the wind in the region of strongest flow over the eastern seaboard of North America, a weakening of the subtropical jet in JJA over the Middle East and central Asia (see Dong et al., 2022) and finally a weakening of the strong mean flow close to Japan in SON. Lee et al. (2019) examining the zonal wind speed between 1979-2019 in the North Atlantic (80W-10W and 30N-70N) find no trend in speed over that period (see their figure 3b).

For the tropics, there has been a strengthening of the Pacific Walker circulation 1979-2012 [see Ma and Zhou (2016)] and the westerlies/easterlies of the tropical Pacific as well as the Hadley circulation since the 1980s.

For the SH, in general the 200 hPa westerlies strengthen at most longitudes, including over South America, over and to the south of South Africa (in JJA) and the polar-jet-stream region eastward from south and south-eastwards of Australia. Over and downstream of Australia however there is a weakening of the subtropical jet and weakening westerlies from the Pacific eastwards to the SH Atlantic.

Simmons (2022) summarises the changes in 200 hPa ERA5 over the past 40 years and similarly Vallis et al. (2015) attempts to summarise robust changes seen in over 40x 1% CO2 models and 40x RCP8.5 scenario climate change model runs. Vallis et al. (2015) make the point that it is not merely a large amount of agreement between both comprehensive and idealised models that assure robustness of the result, but also whether there is a well understood physical argument to explain the phenomena. Because of this, large scale thermodynamical/radiative changes such as enhanced warming aloft in the troposphere, surface polar amplification in the Arctic, an increase in the height of the Tropopause (See Figure 2.5), upper stratospheric cooling and tropical expansion are more robust than dynamical changes as there is usually a more straightforward and agreed upon physical explanation. For dynamical changes even when the models generally agree, clear unambiguous physical explanations are often harder to come by. Models however don't always agree on dynamical trends as often the large natural variability acts to mask these.

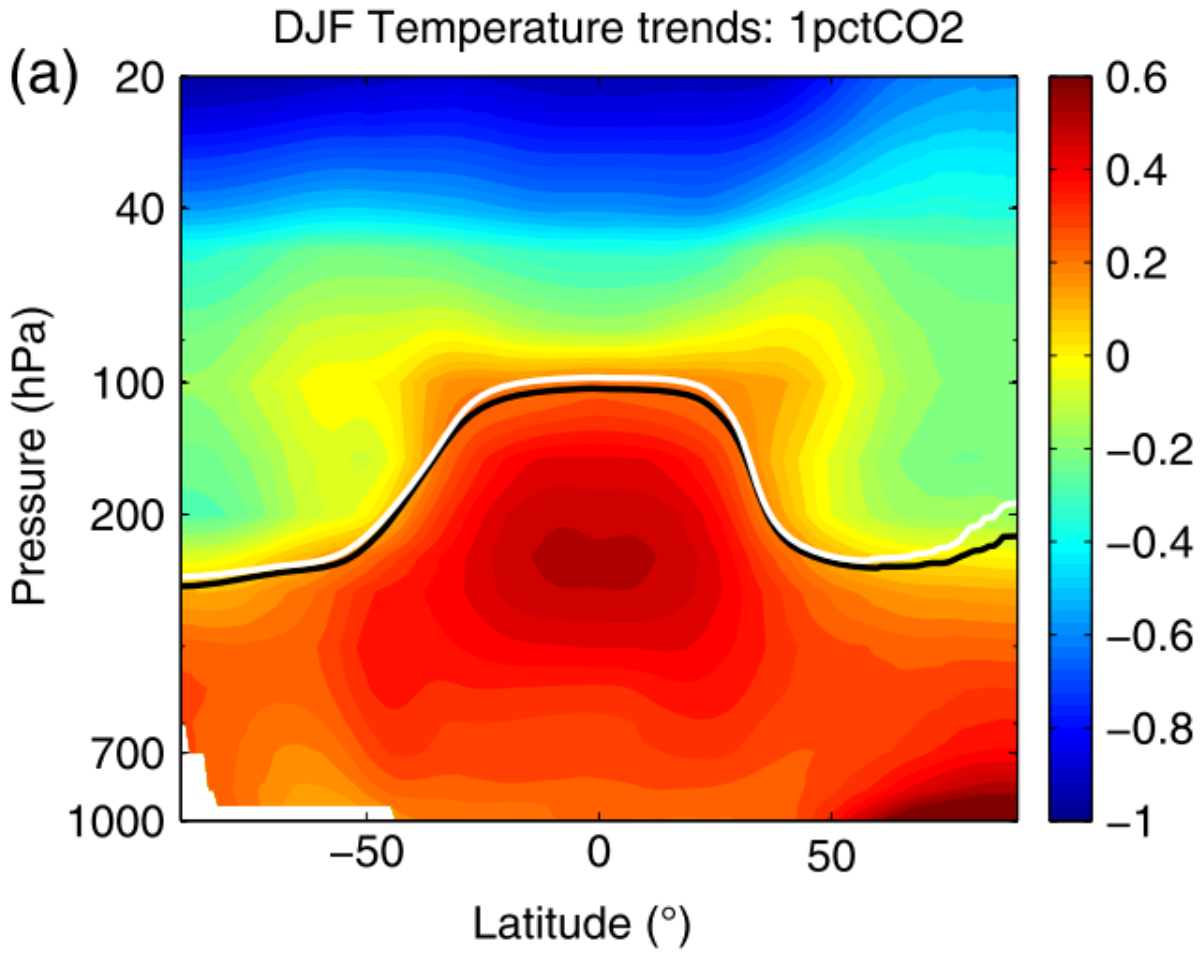


Figure 2.5: Figure 6a in Vallis et al. (2015): The simulated ensemble-mean zonal-mean temperature trend (K/decade) in NH DJF, as calculated over a period of 70 years in a 1%/year CO2 increase experiment. The black and white lines denote the tropopause height at the start and end.

The general expectation (according to Stendel et al., 2021) is for tropical upper tropospheric warming to strengthen and shift pole-wards the subtropical jets and for Arctic polar amplification to weaken them and push them equator-wards with the former of these 2 effects perhaps eventually winning this tug-of-war. There seems to be more consensus for the jets to eventually move poleward.

The trend in speed however may be less clear cut with Stendel et al. (2021) being of the view that despite the eventual poleward consensus of the jet, there isn't a consensus on whether the jets as a whole will speed up, slow down or remain the same. However, Delcambre et al. (2013) in their figure (See Figure 2.6), appear to contradict this by suggesting that both the NDJFM subtropical and polar jets above 400/500 hPa in both the Atlantic and the Pacific are likely to speed up by the end of the 21st century, another paper (Wills et al., 2019 in Stendel et al., 2021) suggests that the jets may weaken in summer.

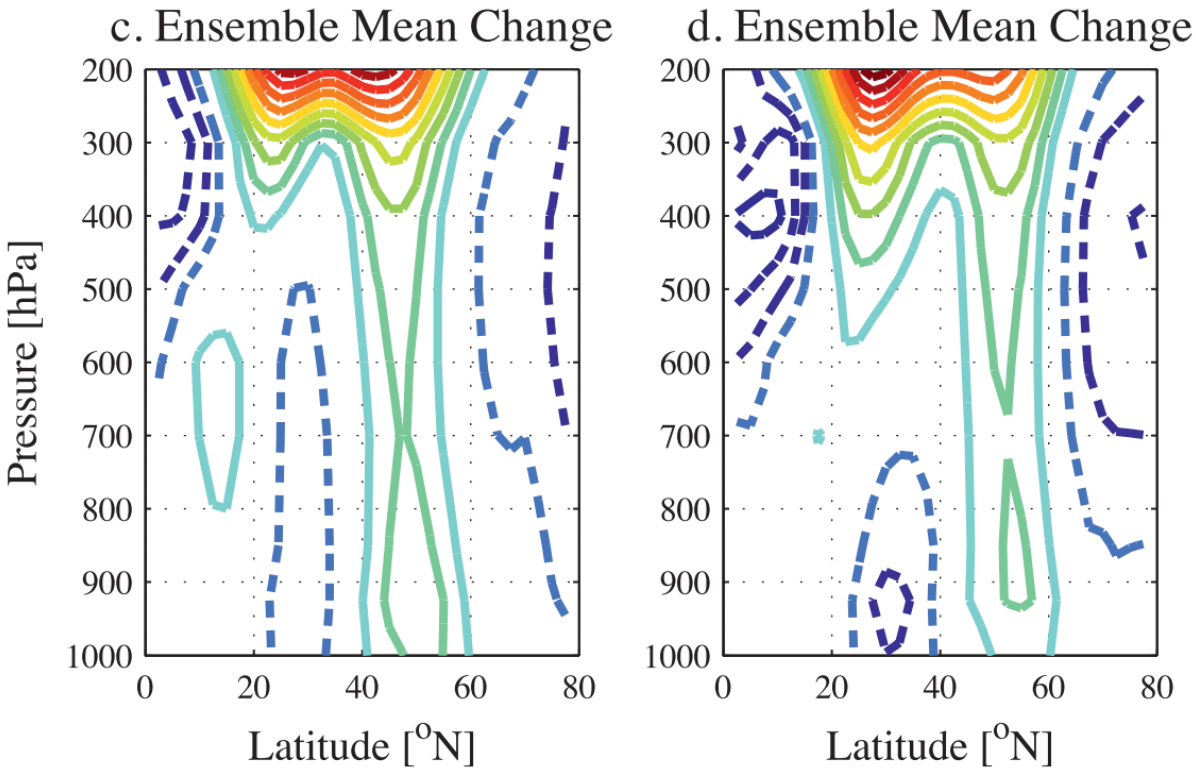


Figure 2.6: Figure 1c/d in Delcambre et al. (2013). The simulated ensemble-mean, zonal-mean, zonal wind change in the NH DJF from the end of the 20th century to the end of the 21st century in the CMIP3 project, in the Pacific (left) and Atlantic (right). The contour interval is 0.25 m/s, with positive contour lines solid, negative contour lines dashed, and the zero line removed.

For the midlatitude circulations in Vallis et al. (2015), and for zonal means - Barnes et al. (2013) in Stendel et al. (2021) discuss zonal asymmetries in the Pacific and Atlantic jets - except for the NH JJA in the 1% integrations, there is a poleward shift in surface westerlies in both hemispheres of about 1% over 70 years. Although Vallis et al. (2015) don't talk specifically about higher up, it is reasonable to assume that this pattern filters up to 200hPa to some degree. Splitting the results apart by hemisphere and JJA/DJF (see their figures 16 and 19), the NH DJF has a 1 degree poleward shift of surface westerlies with a weakening of surface winds. NH-JJA also shows a weakening of surface westerlies but presents no latitude shift. SH-DJF (SH-Summer) has a 1.5 degree poleward shift and a strengthening of surface winds which is also the case for SH-JJA, but with just a 1 degree poleward shift of surface westerlies. In terms of explanation, Vallis et al. (2015) proposes a warming of the tropical troposphere leading to a poleward shift of the meridional temperature gradient pushing pole-wards the region of baroclinicity and storm tracks. They mention that the Southern hemisphere (SH) midlatitude jet could trend equatorward (Polvani and Kushner, 2002; Polvani et al. 2011), if the Brewer Dobson circulation of the stratosphere strengthened.

Other perhaps less relevant changes in circulation examined by Vallis et al. (2015) are the extent and strength of the Hadley cell and the increased height of the tropopause. Not totally unrelated however because they suggest that there is a correlation between the expansion of the Hadley cell and the latitudinal shift of the jet. The Hadley cell determines the extent of the tropics and the position of winds at 200 hPa. Vallis et al. (2015) find an expansion of 0.5 degrees for the NH and 0.8 degrees for the SH for RCP8.5 with a 2.5K temperature rise (see their figure 20). In terms of strength, they find a weakening of the strength in

both JJA and DJF in the NH. The SH was largely unaffected (see their figure 24). In addition, by altering the height of the tropopause in a simple General Circulation Model (GCM), Lorenz and DeWeaver (2007) come to the conclusion that an increase in tropopause height might be the main driver of a poleward shift in the tropospheric jets and storm tracks.

One stakeholder who might be in a position to know whether the jet streams have become faster over the past 40 years is the aviation industry that make use of the jets to reduce fuel costs. Indeed anecdotally perhaps the first subsonic sub 5 hour crossing from New York to London was achieved in February 2020. Tenenbaum et al. (2022) examined the speed of the jet exit region of the North Atlantic jet stream (47N-57N, 40W-10W) using ERA5 and NCEP-NCAR reanalyses as well observational datasets. Although they find a statistically significant upward trend from 2002-2020 onwards, there is no statistically significant trend when the whole 1979-2021 period is considered (See their figure 1). This also fits with the aforementioned Lee et al. (2019).

Looking forward into the world of climate simulations, Williams (2016) compares trans-Atlantic flight times between a double CO2 climate change simulation and a pre-industrial simulation. He finds that the jet streams do indeed intensify resulting in round trips getting longer, an annual average increase of 1 minute 6 seconds which is slightly longer in the summer and shorter in the autumn.

2.2.3 Vertical wind shear trends in the jet stream with climate change

In addition to the question of whether the jets have intensified, changed latitude or altitude there is also the question of whether they have become more sheared and indeed this is what Lee et al. (2019) found when they examined a 250hPa box over the NA (30–70°N 10–80°W) using 3 different reanalyses over the period 1979-2017. At this altitude they found the jet had become 15% more sheared over the period (see their figure 3a). One main explanation for this effect (another involves albedo change from sea ice loss) is as follows: At the equator more of the heat from global warming goes into the upper troposphere (latent heat from convection) than the surface. This keeps the surface cooler than it would otherwise be. In the Arctic however, due to greater stratification, more of the global warming heating goes into the lower troposphere making it warmer than it would otherwise be. These competing effects serve to increase the meridional temperature gradient aloft but decrease it nearer to the surface. These 2 competing effects (via the thermal wind equation) largely cancel out in terms of the jet speed at 250 hPa (see their figure 3b) but manifest as a greater vertical wind shear.

2.3 How CAT is forecast

How aviation affecting turbulence is currently forecast and how skilful these forecasts are

Any turbulence an aircraft encounters in the atmosphere can be considered aviation affecting turbulence. However, aviation affecting turbulence can be split into Clear air turbulence (CAT) which is generated from shears in jet streams and from breaking mountain and gravity waves. Near cloud and in-cloud turbulence (NCT) is caused by in-cloud processes predominantly driven by convection (See Figure 2.7). NCT is easier to predict and is less hazardous to aircraft as convective parameterizations within Numerical Weather Prediction (NWP) models can predict locations of deep convection. Furthermore, deep convection is easily observable meaning aircraft are able to avoid it. CAT on the other hand is hard to observe and can only be observed by aircraft already transiting a CAT region and vertical wind profilers and doppler lidars which are sparsely located. Thus, most research has focussed on predicting CAT, However as shown in the diagram below, the CAT and NCT processes are not independent meaning some CAT diagnostics could also predict aspects of NCT turbulence. Therefore in this review on aviation affecting turbulence forecasting we discuss both.

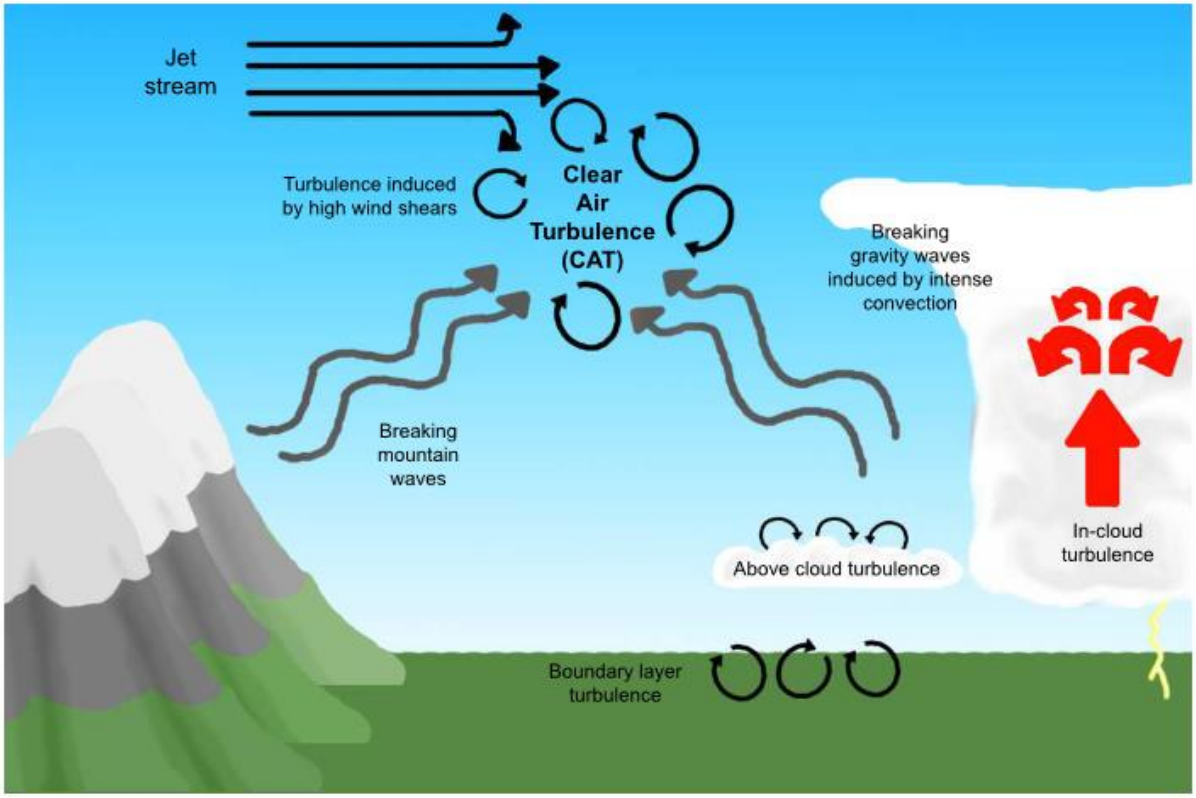


Figure 2.7: A diagram showing the different sources of turbulence in the atmosphere. See text for more in-depth descriptions for each turbulence source (From Marlton, 2016)

2.3.1 CAT Prediction

In essence, CAT is forecast using CAT diagnostics. Diagnostics are used because the turbulence is too small scale to be explicitly modelled using the current resolution of NWP models (Sharman et al., 2006) The diagnostics are computed from the NWP output. The diagnostics are derived from theoretical considerations such as the Richardson number and relative vorticity advection or from empirical relationships such as wind speed or deformation. Either way the diagnostics are designed to take the synoptic and mesoscale features and translate them into a metric indicative of turbulence.

2.3.2 Turbulence diagnostics

Williams and Joshi (2013) use an ensemble of 21 CAT diagnostics (See Appendix A for diagnostic equations). Because different processes generate CAT it makes sense that not every diagnostic will diagnose every instance of CAT based on different synoptic scale features. Therefore, an ensemble approach is likely to be better than any given diagnostic.

Among the 21 diagnostics there are those that contain the Richardson number (The ratio of the buoyancy of the atmosphere to the vertical wind shear), those based around deformation (a deformation flow in the atmosphere can result in the sharpening of the gradients in other variables such as temperature) and vorticity (e.g. the curviness of the jet stream).

Although used to diagnose CAT, some may also diagnose other forms of aviation affecting turbulence. For example the Richardson number may diagnose Convectively Induced Turbulence as well as CAT, similarly Flow Deformation multiplied by Wind speed may diagnose wave related turbulence e.g. Near-cloud Turbulence or Mountain-wave turbulence.

2.3.3 MWT AND CIT Diagnostics

It is likely that a combination of CAT, CIT and MWT diagnostics used together would improve the overall success of any system attempting to diagnose aviation affecting turbulence but as of 2018, GTG doesn't contain many CIT diagnostics (only the Richardson number) although it does contain MWT diagnostics. This is likely to change with GTG4.

Lane et al. (2012) comment that CIT has low predictability which requires the use of ensemble. On top of that PIREPS do not have the spatial and temporal accuracy with which to verify. Pinto et al. (2015) and Sharman and Pearson (2017; in Gultepe et al., 2019), also concur that because of their transient nature and small spatial scale, neither current NWP models nor their diagnostics are particularly skilled. However Gill and Stirling (2013) appear to disagree with this pessimistic assessment using convective precipitation rate, convective precipitation accumulation, and Convective Available Potential Energy (CAPE) to successfully forecast convective turbulence events.

For MWT Elvidge et al. (2017) tested a new diagnostic that for the first time explicitly resolves mountain wave fields (as opposed to a parameterisation) and seems to verify reasonably well (PODy of 80% and a False detection rate of 40%).

More recently Sharman & Pearson (2017) have combined CAT diagnostics with low level winds to improve predictive power in areas at risk of mountain waves.

2.3.4 The Graphical Turbulence Guidance System

As discussed in the previous section there are multiple turbulence diagnostics which focus on different turbulence generation mechanisms. Therefore turbulence forecast systems such as the GTG have been developed which use a combination of turbulence diagnostics to produce a forecast.

In short the GTG uses a cycle of turbulence observations in the last few hours to days to assign a weighting to each turbulence diagnostic to allow the best combination of diagnostics for particular situations and altitudes.

Given the success of GTG it is now implemented by the Washington and London World Area Forecast Centres (WAFCs)

2.3.5 Verifying the performance of turbulence diagnostics and forecasting systems

Typically observations of turbulence (typically in the form of PIREPs, though more recently Eddy Dissipation Rate [EDR] and Derived Equivalent Vertical Gust) are compared against the turbulence predictions from GTG / Diagnostics.

This results in 4 possible outcomes for every observation-grid point pair under consideration. True Positives (TP) when both the observations and predictions both indicate turbulence, True Negatives (TN) when both indicate null (no turbulence), False Positives (FP) when the prediction is turbulent while the observation in null and finally False Negative (FN) when the prediction is null but the observation is turbulent. FP in a forecast might result in aircraft being diverted needlessly and wasting time and fuel as a result, but a FN, depending on the severity of the turbulence could at worst result in hospitalisations and damage to the aircraft's airframe.

If our ability to forecast CAT was perfect then we would have only True Positives (TP) and True Negatives (TN), however Sharman et al. (2006) are of the opinion that we haven't yet reached even an 'acceptable' standard.

In order to quantify the respective number of TP, TN, FP and FN, two statistics in particular are used. These are Probability of Detection yes (PODy) and Probability of Detection no (PODn). In a given set of observation, model pairs, PODy is the number of TP as a proportion of all turbulent (as opposed to null) observations. PODn on the other hand is the number of True Negatives (TN) as a proportion of all null observations. Following the aforementioned statement by Sharman et al. (2006), they quote one definition of an 'acceptable' benchmark as having a PODy of at least 0.8 and a PODn of at least 0.85.
Figure 2.8 shows what such a standard would look like in practice:

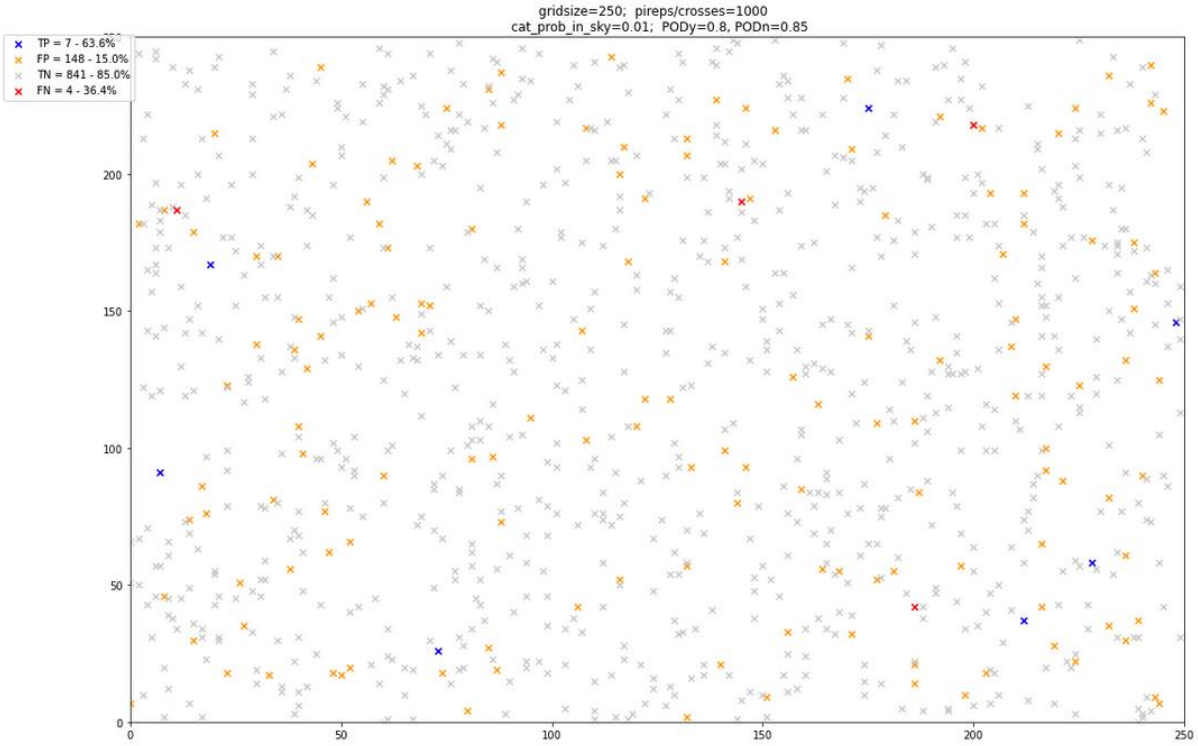


Figure 2.8: A demonstration of the number of TP, TN, FP and FN that would be expected from the 'acceptable' benchmark of Sharman et al. (2006) of PODy=0.8 and PODn=0.85 for a spatially random selection of 1000 PIREPs given a 1% underlying prevalence of CAT.

There is an inherent trade-off between PODy and PODn. If ones goal was to maximise the TPs, then one would lower the threshold for a turbulent prediction. This would likely result in a greater PODy score, however as a consequence the number of FPs would increase and the PODn score would drop. Likewise if the goal was to minimise FP, this could be achieved by decreasing your turbulence threshold sensitivity, but this would result in a drop in the number of TPs. This trade-off is best illustrated in the form of Receiver operating characteristic (ROC) curves where PODy is plotted against 1 - PODN for varying sensitivities of detection. ROC curves introduce another statistic relevant for turbulence forecast evaluation, the 'Area under the curve' (AUC) this varies between 0.5 (diagonal line - no skill) and 1 a perfect skill score.

So how do different turbulence diagnostics fare against one another in terms of PODy, PODn and AUC? Table 2.1 summarises some of the verification statistics given in the literature:

Table 2.1: Verification statistics (e.g. PODy and PODn) of various diagnostics found within the CAT literature.

	Assessed Diagnostic/Forecast System		PODy	PODn	1-FARatio	
		Observation No	TPR	TNR	PPV	AUC
ELLROD&KNAPP_1992		~1100/1109	.7-.84			
MCCANN_2001	Divergence tendency (d17)				0.858	
	Anticyclonic instability				0.7925	
	Wind shear squared (d02)				0.725	
	Richardson number (d01)				0.7175	
	Inertial advection wind				0.688	
	Unbalanced ageostrophic wind				0.558	
SHARMANETAL_2006	GTGClimatology (FL200+)	49703	0.614	0.926		
	GTGDynamical (FL200+)	49703	0.623	0.926		
	GTGClimatology (FL100-200)	8063	0.53	0.813		
	GTGDynamical (FL100-200)	8063	0.527	0.816		
KNOXETAL_2008		~35	0.829	0.785		
KIMETAL_2009	Colson--Panofsky (d03)					0.722
	Negative Richardson number (d01)					0.693
	Variant 1 of Ellrod's turbulence index (d07)					0.687
	Frontogenesis function					0.682
	NGM2 - Flow deformation times vertical temperature gradient - (d16)					0.623
	Horizontal temperature gradient (d12)					0.605
ELLROD&KNOX_2010	DT1		0.421	0.692		
	Variant 1 of Ellrod's turbulence index (d07)		0.284	73.6		
KIMETAL_2011	GTGeastasia	1370	0.61	0.75		0.795
MCCANNETAL_2012	ULTURB		~.71	~.81		
SHARMAN&PEARSON_2017	GTG CAT high (ensemble)	693379	0.221	0.965		0.788
	GTG CAT midlevel (ensemble)	84910	0.326	0.907		0.74
	GTG CAT low (ensemble)	91318	0.174	0.963		0.721
	GTG MWT high (ensemble)	649722	0.435	0.989		0.933
	GTG MWT midlevel (ensemble)	81147	0.42	0.919		0.835
	GTG MWT low (ensemble)	87113	0.422	0.909		0.791
	GTG CAT high (ensemble) - 2015	176062	0.246	0.957		0.766
	GTG CAT high (ensemble) - 2015 (no CONUS)	38693	0.346	0.947		0.808
KIMETAL_2018	G-GTG					0.85
	Ellrod3					0.81
	WAFS					0.71

As can be seen from table 2.1, although some of the diagnostics/forecasting systems achieve an impressive PODn score, none manage to do so while keeping PODy above 0.8.

2.3.6 Where are we better/worse with our forecasts?

GTG gets better verification scores (Sharman et al., 2006) at above FL200 (20,000ft) than it does between FL100-200 (10,000-20,000ft), which is perhaps unsurprising as aircraft spend more time at cruise level than they do below. The ULTURB algorithm of McCann et al. (2012; a modification of an algorithm developed by Knox et al., 2008) whose leading order term is the advection of relative vorticity) can also skilfully forecast at below Fl100 and is better at forecasting severe turbulence possibly because it takes into account turbulence generated by gravity waves.

2.3.7 Issues with underlying PIREPs and NWP accuracy

When we are assessing the skill of a given diagnostic or forecasting system, chiefly two assumptions are made. The first is that inaccurate predictions are due to the diagnostics lack of skill as opposed to inaccuracies in the NWP forecast or analysis on which the diagnostics were applied. The second assumption is that the turbulence observations, often PIREPs, are accurate and unbiased which is not always the case. The various issues with PIREPs are best summarised by Schwartz (1996) and include issues such as the under reporting of null turbulence results, observational data highly biased towards the common flight routes and the subjective assessment of the severity of turbulence due to pilot experience as well as aircraft size as well as spatial and temporal uncertainties. Some of the issues with PIREPs will likely be mitigated

in future when automatic quantitative recording of more objective quantities EDR and DEVG become more common place. While such data still won't solve the problem of the observations being over concentrated along flight routes, they will provide a more balanced coverage of those areas with a better balance between turbulent and null observations.

2.3.8 Uncertainty in forecasting turbulence

As stated briefly above, if a diagnostic/diagnostics are applied to a deterministic model, then the diagnostic is subject to any errors in the underlying forecast (for example the position and speed of the jet stream). Such a number conveyed to policy makers while better than nothing does not convey forecast uncertainty so is less useful than it might be. Storer et al. (2019b) investigated whether a forecast from a multi model ensemble (Met Office forecast model with 12 members + European Centre for Medium-Range Weather Forecasts (ECMWF) forecast model also with 12 members) using a single diagnostic (TI1) was superior to either model's ensembles by itself when verified against DEVG observations from a fleet of Boeing 747/777 aircraft. Although they found that the multi-model ensemble wasn't significantly better (greater area under the curve) than either of the individuals models ensembles, they found an improvement in forecast value at low cost/loss ratios a variable important to end users of such forecasts. Storer et al. (2020) followed up by comparing single-diagnostic, multi-model forecasts with multi diagnostic (3x CAT, 1x MWT and 1x CIT), multi model forecasts as another approach for generating probabilistic forecasts that include uncertainty. Here, they found multi diagnostic forecasts superior in skill to single diagnostic forecasts and that reducing the number of ensembles members from 51 to 12, didn't reduce the forecast skill appreciably.

2.3.9 Applying turbulence diagnostics to climate models

The turbulence diagnostics used to make the forecasts can equally be applied to the output of climate models. In Storer et al. (2017), an ensemble of 21 turbulence diagnostics are applied to the output of a PI control and a RCP8.5 scenario run to 2050-2080 on the HadGEM2 CMIP5 climate model in order to calculate changes in turbulence. The authors find a 200%-500% average global increase in the number of times a 'moderate or greater' (MOG) threshold is exceeded globally. While such a forecast uses a similar methodology, the aim is to demonstrate the changes in the prevalence of CAT with climate change as opposed to helping flight routers decide where and where not to route flights.

2.4 ERA5 at 200 hPa

2.4.1 Introduction

ERA5 is the ECMWF's successor to ERA Interim (Dee et al., 2011) and is an uncoupled global atmospheric reanalysis dataset with much higher resolution than its predecessor: A quarter of a degree horizontal resolution and 137 vertical model levels with hourly output.

Unlike ERA-Interim, ERA5 also has a 3 hourly 10-member ensemble data assimilation (EDA), which helps provide an estimate of uncertainty.

Upon first release, ERA5 ran from 1979 when the first TOVS (TIROS [Television Infrared Observation Satellite] Operational Vertical Sounder) satellite observational data began being assimilated - including observations from the Stratospheric Sounding Unit (SSU) from the Global Atmospheric Research Programme's 1979 global observing experiment (Simmons, 2022). Subsequently ERA5 has been extended back to 1950 (Bell et al., 2021).

2.4.2 Quality over time

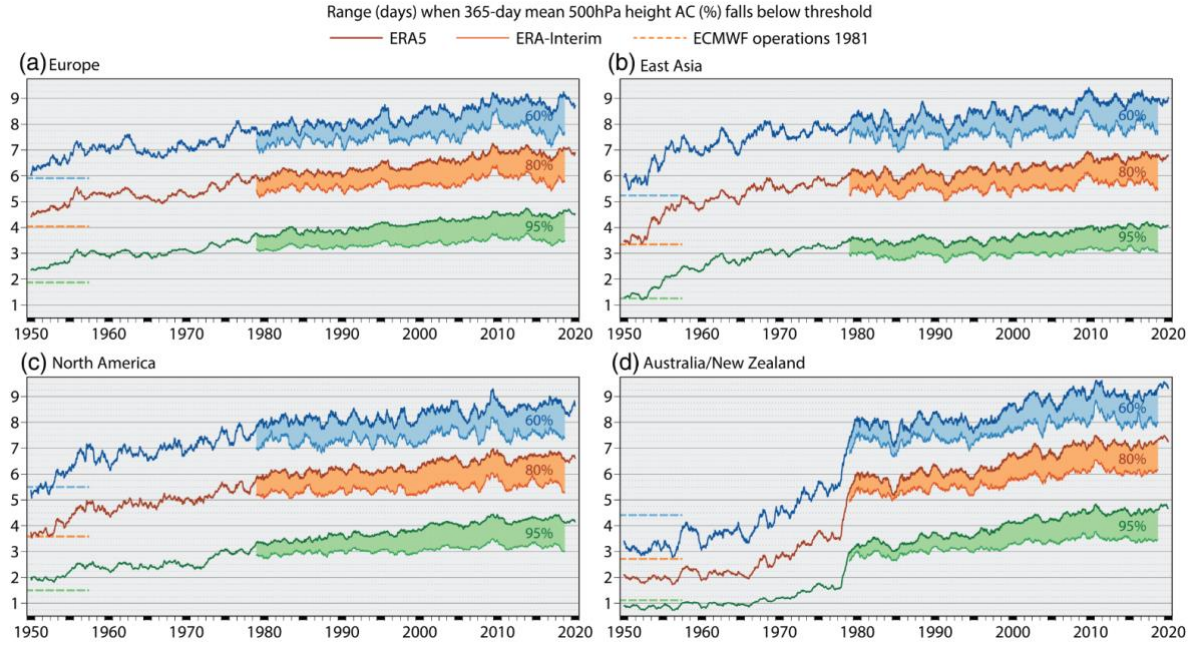


Figure 2.9: Range (days) at which running 365-day mean anomaly correlations of 500 hPa height forecasts at 0000 and 1200 UTC from 1950 to 2020 reaches 95% (green), 80% (orange) and 60% (blue) for (a) Europe, (b) East Asia, (c) North America and (d) Australia/New Zealand. Also shown (dashed) is the average skill of ECMWF operational forecasts for 1981. Heavy lines denote ERA5; thin lines denote ERA-Interim. Shading denotes the difference between ERA5 and ERA-Interim during the period for which both are available (1979–2019). Taken from Bell et al. (2021).

Figure 2.9 is taken from Bell et al. (2021) and looks at how far in the future a re-forecast of 500 hPa geopotential height initialised with ERA5 can be accurate. In 1978, a significant performance jump is evident for Australia and New Zealand with the influx of satellite data. No such sudden jump however can

be seen for Europe, East Asia, and North America. Figure 25(f) in Bell et al. (2021) also shows evidence of an improvement in quality with the new satellite data for radiosonde temperature at 50hpa.

Since its release ERA5 has had its performance evaluated in the Arctic (Graham et al., 2019) and in the Antarctic (Tetzner et al., 2019) and has been found to perform well in representing winds, temperature, and humidity at the surface. Here we are specifically interested in how well it performs at cruise altitude at 200 hPa.

2.4.3 Quality diagnostics

According to Bell et al. (2021), there are a number of diagnostics that can be used to evaluate the quality of ERA5 and whether or not it has been improving over time: These are:

1 - The size of the 10-member ensemble spread (the smaller the better)

2 - The size of the first-guess departure variances. This is the model first guess equivalent relative to the observed value (the smaller the better).
See reduction in the standard deviation of upper air temperature in Figure 14(a) in Hersbach et al. (2020) in 2015 after Binary Universal Form for Representation of Meteorological Data (BUFR) radiosonde data are assimilated and wind (b) first-guess departures between 1979-2020 (The first half of this figure covering the period 1959-1980 can also be seen in Bell et al., 2021, figure 11). They both show a steady decline with a value of 3.7m/s in 1950 falling to 3.3 in 1979 then 2.5 in 2020.

3 - The coherence of analysis increments
See Figure25d in Bell et al. (2021) for fairly consistent temperature analysis increments at 50hpa for ERA5 between 1970-2020

4 - The accuracy of re-forecasts initialised by ERA5
ERA5 in general can be shown to be an improvement on ERA-Interim as it is able to accurately forecast roughly a day further in re-forecasts (see figure 2.9)

According to the ERA5: uncertainty estimation part of the ECMWF website: https://confluence.ecmwf.int/display/CKB/ERA5%3A+uncertainty+estimation
"The EDA (Ensemble of Data Assimilations) takes into account mostly random uncertainties in the observations, sea surface temperature (SST) and the physical parametrizations of the model. In principle, as long as these uncertainties are properly described and there are no additional sources of uncertainty (other sources of omitted uncertainties include: Radiative forcing due to greenhouse gases, or systematic errors in

the model or the way in which observations are used), then the EDA will properly describe the reanalysis uncertainties. However, systematic model errors are not taken into account by the EDA and the errors (uncertainties) as defined by the EDA are uncorrelated"

This is useful as it helps determine where and when ERA5 is more/less reliable. For example the EDA can be used to show that reanalysis reliability has improved considerably over the last 40 years (2017 vs 1980) as the number of assimilated observations has increased by roughly 30x (See Figure 11 in Hersbach et al., 2020).

2.4.4 ERA5 at 200 hPa

This PhD makes substantial use of the horizontal wind and temperature fields (u, v and t) at ~200 hPa.

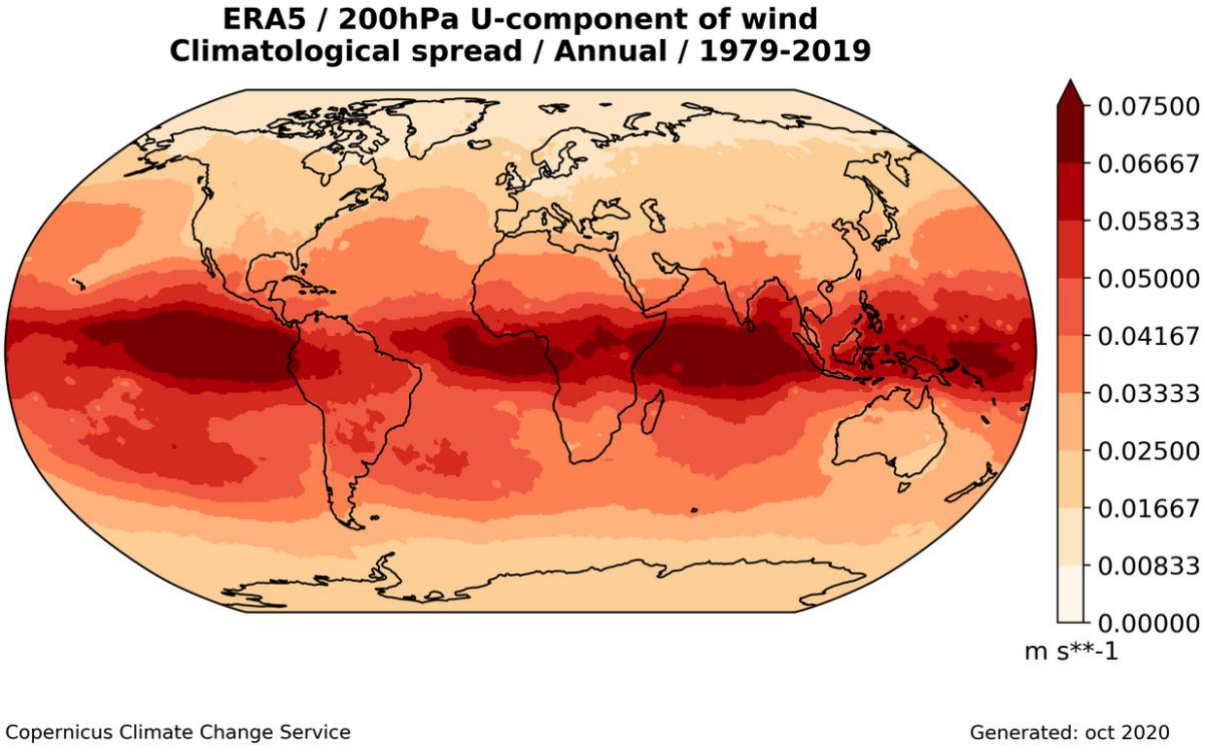


Figure 2.10: The ERA5 200hPa U-component of wind climatological spread (Annual, 1979-2019). Generated from https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=overview click on 'Quality Assessment', select 'U-component of wind'. Then under 'Independent assessment' click 'Expert evaluation' then choose 'climatological spread' under 'Diagnostics'.

Figure 2.10 shows the climatological spread (1979-2019) of the ensemble members for u at 200 hPa from which it is evident that there is more uncertainty at the tropics than the midlatitudes.

In a relative sense, the tropics are the least reliable but Figure 24(f) in Simmons et al. (2020) seems to show extremely good agreement between monthly mean observations of 100hPa wind speed for the tropics between 1979-2020 which seems to verify even the least certain area of the globe for the u wind.

The same graph for T however (See figure 2.11), Indicates that the greatest uncertainty exists over the Southern Ocean.

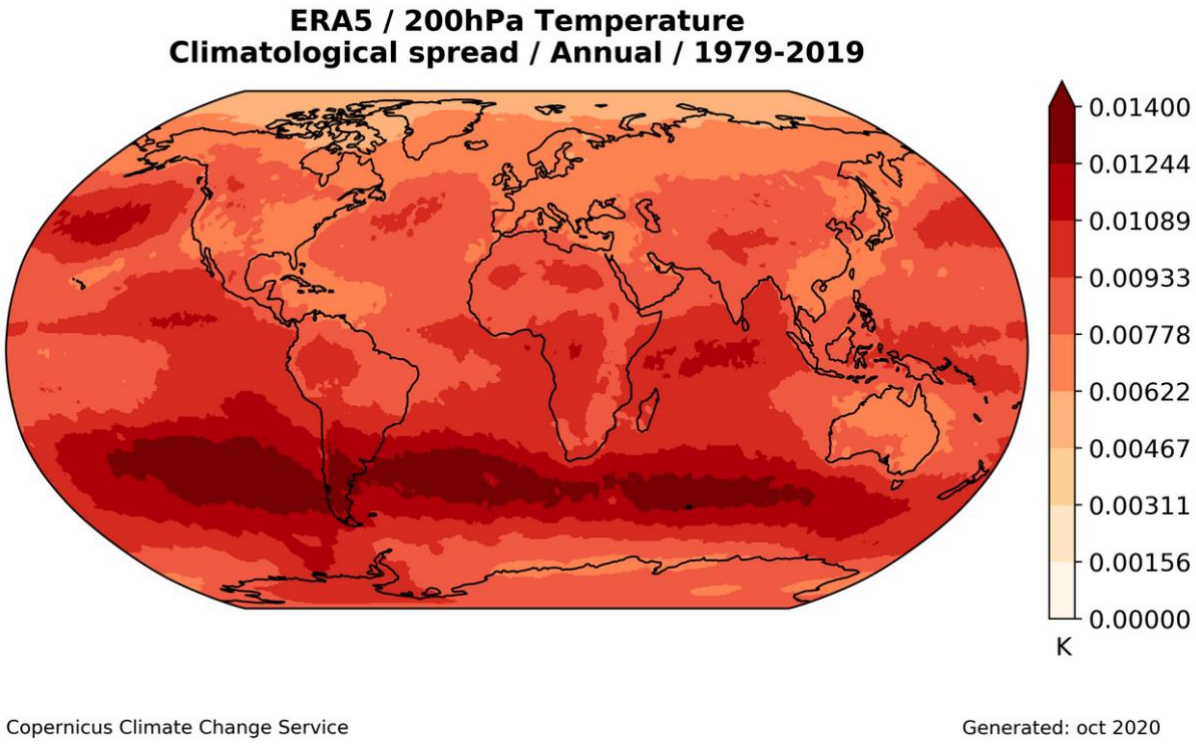


Figure 2.11: The ERA5 200hPa Temperature climatological spread (Annual, 1979-2019). See figure 2.10 caption for generation instructions.

In the 1980s, observational data assimilated at tropopause altitudes in ERA5 came mainly from TOVS satellites and radiosondes which goes someway to explaining the smaller ensemble spread in the northern hemisphere where radiosonde data is/was more available.

Over the 1980-2020 period the ensemble spread for both t and u (See Figure 2.12) has halved with significant improvements in 1998 following the introduction of ATOVS (Advanced TIROS [Television Infrared Observation Satellite] Operational Vertical Sounder) data and an increase in the amount of GNSS-RO (Global Navigation Satellite System – Radio Occultation) data in 2006.

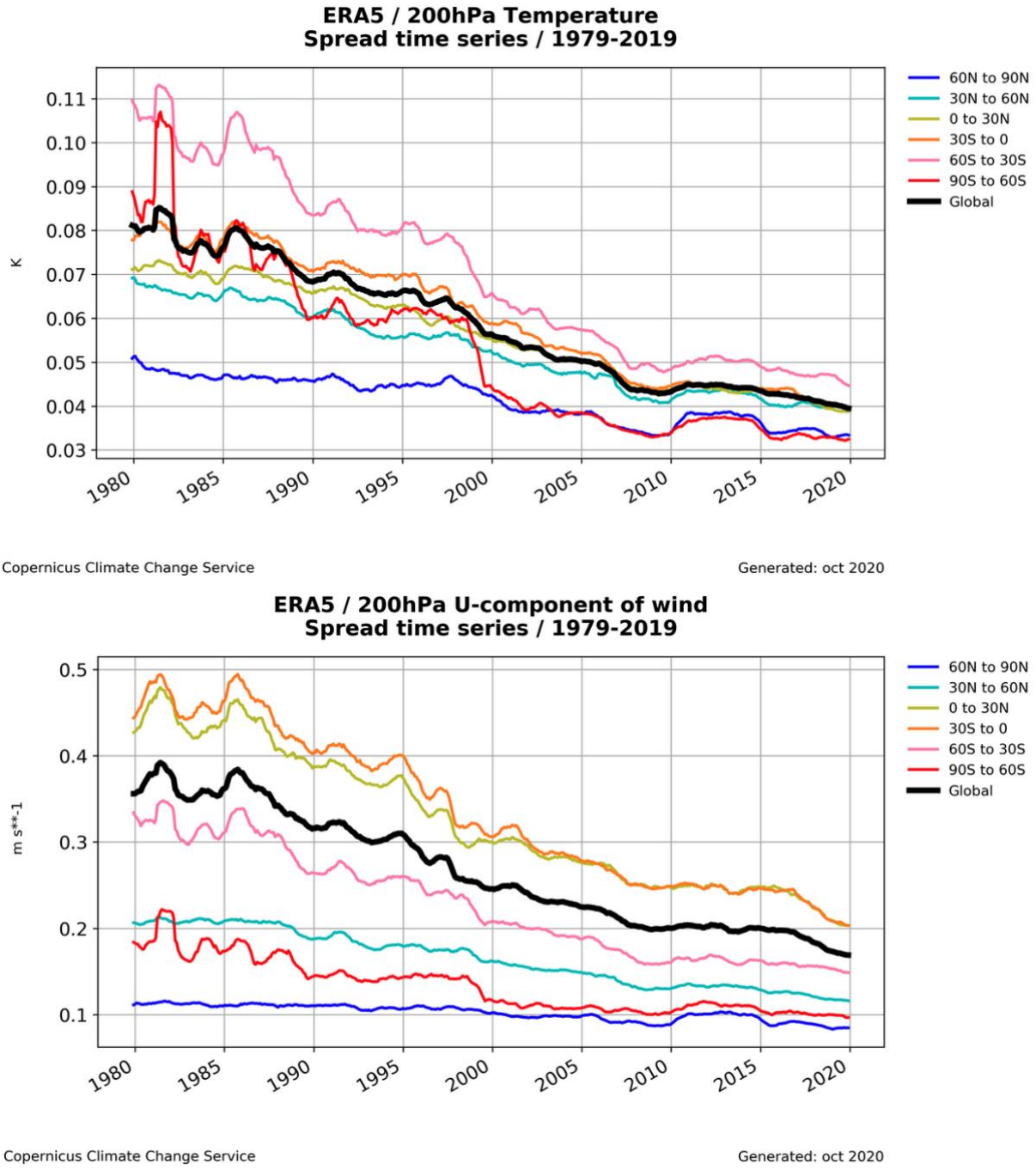


Figure 2.12: The ERA5 200hPa Temperature (upper panel) and U-component of wind (lower panel) ensemble spread (1979-2019). See figure 2.10 caption for generation instructions.

Is the quality of ERA5 in the early satellite era good enough? Simmons (2022) states that ERA5 has good multidecadal consistency with plentiful wind observations (e.g. Atmospheric motion vectors [AMVs] when winds throughout the troposphere can be inferred from the movement of water vapour in satellite images) from aircraft and satellites near the tropopause. There is however a paucity of upper-air observations over the Southern Ocean but ERA5 compensates here using surface and thermal wind information from satellites. Figure 12 of Simmons (2022) shows the number of radiosonde, aircraft and satellite data assimilated in a 240 degrees of longitude region of the tropics (10S-10N) between 175 and 225 hPa. Figure 14 shows the same information for 2x 30 degrees of longitude regions in the extra-tropics at the same altitude. In all cases the number of assimilated observations increases over time. Another source of

confidence for Simmons in ERA5 at 200 hPa is the broad agreement with another Japanese reanalysis dataset called JRA55 in both the tropics and extra tropics, but especially so in the latter.

2.4.5 Problems with ERA5

A list of known issues with ERA5 can be found here:
https://confluence.ecmwf.int/pages/viewpage.action?pageId=82870405#ERA5:datadocumentation-Knownissues

With respect to the Upper Troposphere and Lower Stratosphere (UTLS), the chief among these is the cold bias between 2000 and 2006. This bias can be seen for 40-60hPa in Figure 15 in Hersbach et al. (2020). This was later remedied by the creation of ERA5.1.

Globally averaged temperature values higher than 10hpa are a bit spurious and inconsistent with respect to other reanalyses but at altitudes lower than 30hpa there is good agreement between ERA5, ERA-Interim, JRA-55 and MERRA2 (See Figure 25c in Hersbach et al., 2020) at least after 1980 and especially after 2006. Prior to that, JRA-55 and ERA-Interim and ERA5 diverge in the 1970s and 80s (See Bell et al., 2021, figure 25b).

2.5 Impact of climate change on CAT and aviation

Aviation both impacts on and is impacted by the changing climate.

The former is well appreciated (see Lee et al., 2021 for the latest assessment), but the latter is perhaps less so and is the subject of this literature review.

Understanding how aviation is impacted by climate change is important not only as a subject in its own right but also because, by doing so, ways in which aviation's climate impact may be mitigated. As an example, if severe CAT was likely to increase in a certain geographical location, then perhaps aircraft flying in this route could be fitted with LIDAR that could detect CAT with a few seconds lead time and reduce injury among passengers and crew.

To break down the title slightly, the 'impact of climate change' refers to both impacts up until the present day as well as potential future impacts and 'aviation', refers to both aircraft in flight as well as infrastructure on the ground such as airports.

Broadly speaking the literature points to 4 ways in which climate change has and will continue to impact aviation.
These are:
1. An increase in the intensity and frequency of clear air turbulence (CAT)
2. Changes in flight times on certain routes
3. A more viscous atmosphere for aircraft to fly through
4. A reduction in aircraft take-off performance

In addition to these 4, some airports, particularly those close to sea-level are at risk from sea-level rise, storm surges and storm activity in general. Lightning activity also looks set to increase (e.g. Romps et al., 2014)

2.5.1 An increase in the intensity and frequency of clear air turbulence (CAT)

CAT is already an issue for the aviation industry causing hospitalisations to both passengers and crew as well as costing the industry an estimated $150-500 million each year in the US.

At most 1% of the atmosphere at upper levels is likely to contain moderate or greater (MOG) levels of CAT (Sharman et al., 2006) and despite decades of research effort, the ability to forecast it remains limited with many false positives and negatives (forecasting based on diagnostics; see Sharman et al. (2006). CAT cannot be remote sensed with current technology.

There are 2 main reasons why we are focusing on CAT specifically here as opposed to turbulence more generally.

The first is that many aircraft make use of the jet stream in order to increase their ground speed and reduce the fuel consumption on their eastward bound flights in the midlatitudes.

The downside of this relationship is that an estimated three times as much CAT occurs near the jet stream compared with areas of weaker winds (Reiter 1963c) and aviation is thus particularly prone to it.

The second reason is that, unlike other forms of turbulence, CAT is not linked to some specific feature of the environment (e.g. Mountain wave turbulence occurs near mountains, Convectively induced turbulence

occurs in or near clouds) that might enable anticipation and therefore mitigation (e.g. avoidance or seat-belt fastening).

Why might climate change have an impact on CAT?

The Arctic lower-tropospheric and tropical upper-tropospheric amplification of global warming (via the lapse-rate feedback) is causing the Earth's northern hemisphere average North-South (N-S) temperature gradient to reduce/equilibrate in the lower troposphere (1000-600mb - See Lee et al., 2019, figure 2), to stay constant at ~500mb and to increase/diverge in the upper-troposphere lower stratosphere (UTLS) (400-200mb). From the thermal wind equation, the vertical shear of the NH westerly zonal wind at any x, y and p in the atmosphere is proportional to the latitudinal temperature gradient at that same x, y and p coordinate. Hence, any pressure altitude experiencing a growing N-S temperature gradient (i.e. 400-200mb) is likely to exhibit increased wind shear. Conversely, any pressure altitude experiencing a reducing N-S temperature gradient (i.e. 1000-600mb) is likely to exhibit decreased wind shear.

CAT is liable to occur in regions of the atmosphere with a Richardson number (See Appendix A – D01 for equation) less than 1/4 (Miles and Howard, 1964).

This occurs when the numerator is low (high buoyancy, low stability) and the denominator (high vertical wind shear) is high. Given the previous thermal wind argument, we might expect a higher vertical shear (and therefore lower Richardson number) at any pressure altitude experiencing a growing N-S temperature gradient.

Lee et al. (2019) looking at 3 reanalysis datasets over the period 1979-2017, found growing N-S temperature gradients at pressure altitudes between 200-400hPa with the largest growth occurring at 250hPa. Correspondingly, they (see their Extended Figure 1) find the jet-stream to have become more vertically sheared between 200-300hPa (statistically significant at at least 250hPa with a 15% increase), and, in general, less vertically sheared below (though below the threshold for significance). As ~200/250hPa is the average flight cruise altitude for commercial aircraft, it should therefore be wholly unsurprising if climate change were to make for bumpier flights.

Looking at the relationship between climate change and CAT more directly, Williams and Joshi (2013 - hereafter referred to as 'W&J13'), examine how the DJF climatological distribution of 21 CAT diagnostics change between a pre-industrial control run and an equilibrated 2xCO2 simulation in the northern north Atlantic (50-75 degrees latitude at 200hPa). They found an increase in both the strength (the median diagnostic value shifted right on average 25.7%) and the frequency of MOG CAT (the mean increase over the 21 diagnostics was 83.9%).

One limitation of the 'MOG' approach of W&J13 was that it doesn't distinguish moderate CAT (unpleasant, but not dangerous) from severe CAT (dangerous with vertical acceleration exceeding 1g). Williams (2017

- hereafter referred to as 'W17') performed a further study repeating the setup and methodology of W&J13 but this time examining 5 different CAT strengths (light, light-to-moderate, moderate, moderate-to-severe, and severe). On average, across the same 21 diagnostics, they found frequency increases of 58.7%, 72.0%, 83.9%, 106.2% and 141.1% respectively. While increased diagnostic values are likely to be related to future increases in bumpiness (vertical acceleration) felt by passengers (a point also emphasised in W&J13), it is also important to bear in mind that they aren't necessarily the same thing and other factors such as improvements in CAT forecasting and remote sensing (e.g. Lidar) may mitigate this to some extent.

While W&J13 and W17 had given clear indications about how climate change was likely to impact CAT, they had only done so only in the northern north Atlantic winter (DJF) at 200hPa.

Storer et al. (2017, hereafter referred to as 'SEA17'), proceeded to fill this gap by examining CAT in all 4 seasons, globally and at both 200hPa and 250hPa.

In the DJF northern north Atlantic at 200hPa, SEA17 found increases of 75.4%, 124.1%, 143%, 148.9% and 181.4% in the frequency of the 5 strength categories of CAT respectively.

These numbers are some 50% higher than the equivalent ones of W17 and may be due to the different climate change scenario (RCP8.5 by 2050-2080) being used by SEA17. RCP8.5 is arguably a stronger climate change scenario than 2xCO2 as by the 2050-2080 period, atmospheric concentrations of greenhouse gases (GHGs) are projected to be between 675-1000ppm CO2e which is significantly higher than that of a 2xCO2 simulation (~315ppm CO2e for pre-industrial control vs ~595ppm CO2e for 2xCO2). On the other hand however, unlike the equilibrated 2xCO2 simulations, global temperatures in RCP8.5 by this period haven't yet caught up with the atmospheric GHG load (i.e. the Transient Climate Response or TCR).

In SEA17, global CAT diagnostic threshold (e.g. MOG) frequency increases were somewhat less than the northern north Atlantic and more uniform across the CAT thresholds at +39.8%, +35.9%, +30.8, +27.9% and +34.7% respectively (versus +75.4%, +124.1%, +143.3%, +148.9%, +181.4% for the north Atlantic) which is unsurprising given that away from the northern and southern hemisphere jet streams, diagnostic increases were typically much less. However the work of SEA17 (like the 2 previous studies), were only diagnosing CAT and CAT isn't the only form of aviation affecting turbulence and therefore just because the low latitudes look relatively unaffected by CAT increases between 2050-2080, doesn't mean they won't be affected by other forms of turbulence (e.g. Convective turbulence). Clearly there is more work to be done here.

Another point to emphasise about these studies is that they show the changes in CAT relative to the pre-industrial and not the present day. While we cannot yet say how much the frequency of CAT has changed in the atmosphere since the pre-industrial, as we have so far tracked relatively closely to the RCP8.5 high emissions scenario (and are close to 1.5xCO2 at the time of writing) it is likely that some of the projected increases in CAT in these 3 papers are likely to have already materialised and recent work by Williams and

Storer (2022) suggest that climate models may underestimate the change relative to reanalyses over the same period.

SEA17 note that planning for such a turbulent future is worth doing now as "Many of the aircraft that will be flying in the second half of the present century are currently in the design phase."

In the opening paragraph I argued that a secondary reason for studying the impact of climate change on aviation is that climate change can impact on aviation in such a way that aviation's impact on the climate is modified. A standard recourse for aircraft encountering turbulence is to adjust altitude, but in so doing, increase their fuel consumption. Williams (2021) estimate that such additional altitude changes require up to an additional 160 million gallons of kerosene per year (in just the US) resulting in both additional CO2 (1.5 million extra tonnes) and the associated nonCO2 climate impact.

2.5.2 Changes in flight times on certain routes

Because climate change is likely to affect the jet stream (strengthen and shift it northward), it is natural to wonder if this will result in changes to journey times on routes such as New York to London that make significant use of it.

One might think that a faster jet stream would result in faster eastward flights at the expense of westward flights leading to no appreciable difference in the round-trip journey time. However, this intuition (based on the arithmetic mean) isn't correct and the average round trip journey speed falls (see harmonic mean) as the speed of the jet stream increases. To illustrate this, consider that an aircraft travelling at 500 mph both eastwards and westwards would complete the round-trip journey far faster than another aircraft travelling at 999 mph eastwards and 1 mph westwards (average speed of just 2 mph).

2 papers that looked at this question (and the New York to London route) were Williams (2016 - hereafter referred to as 'W16') and Irvine et al. (2016 - hereafter referred to as 'IEA16') who both examine how the distributions of the 'minimum time route' changes with climate change on both legs of the New York to London journey. The 2 studies reach slightly different conclusions.

W16 finds eastbound (EB) journeys (5 hours 38 minutes in the pre-industrial run) shorten on average by 4 minutes, whereas westbound (WB) journeys (6 hours 40 minutes in the pre-industrial) lengthen by on average by 5 minutes 18 resulting in the mean round trip journey lengthening by 1 minute 18 (his time reduces to 33 seconds in JJA and increases to 1 minute 51 in SON at 200hPa). The time increase at DJF at 150 and 250hPa is 1 minute 18 and 1 minute 33 respectively.

IEA16 find on average less than a one-minute shortening of eastbound flights, and average of a one minute lengthening of westbound flights resulting in negligible? changes to the round trip journey time.

Much of the methodology used between the two studies was similar (same New York to London route, the 'minimum time' parameter, 250m/s aircraft speeds, all season etc) however there were also significant differences. W16 used a single CMIP3 model where a preindustrial control run was compared with 2xCO2 climate change simulation (1% increase over 70 years) that was then allowed to equilibrate. IEA16 used 5x CMIP5 models (Selected as they showed a range of jet stream behaviours and realistic NAOs) that compared the historical climate of 1979-2005 with an RCP8.5 climate scenario to the period 2073-2099. Of note is that RCP8.5 is an unequilibrated near tripling of CO2e (~380ppmCO2e to ~1050ppmCO2e) occurring over 94 years whereas W16 doubling of CO2/CO2e is equilibrated. There are other key differences too. Unlike W16 who conduct their analysis at the 150, 200 and 250hPa level, IEA16 choose the 250hPa level and in addition the following: "It is important to note that the wind speed values for the jet speed shown here will be considerably lower than the maximum wind speeds observed in the jet stream, since we average the wind speeds over a sector of longitude to find the jet stream.". Some or all of these factors may help to explain the different conclusions reached by these 2 studies.

In terms of projected impacts, W16 quotes an extra 2000 flight hours burning an additional 7.2 million US gallons of fuel emitting an extra 70 million kg of CO2 from all transatlantic traffic per year (assuming 2016 levels of traffic), whereas IEA16 do not.

While large it is worth comparing this to the amount of fuel wasted annually in the US from aircraft adjusting altitude to avoid turbulence, an amount some 200x smaller.

Both authors note that their analysis does not factor in changes to turbulence and the 'minimum time route' is unlikely to be the one that minimises the climate impact, a consideration in flight-route planning that may become more important into the future.

2.5.3 A more viscous atmosphere for aircraft to fly through.

So a jet stream modified by climate change may have an impact on flight times, but what about the rest of the atmosphere away from the jet streams?
Ren et al. (2020) examine this question by comparing an ensemble of 34x CMIP5 climate models using the RCP8.5 and RCP4.5 scenarios runs (simulated through to 2050-2099) with historical runs (1950-1999).

They note 4 variables that are likely to impact on the fuel efficiency of aircraft:

Tropopause height (Their figures 3&4). As the globe warms, the height of the tropopause is expected to rise due to the thermal expansion of the troposphere. In the RCP8.5 scenario, the tropopause is some 50-260m higher than in the historical period with the greatest rise being in the tropics (~240m), followed by the Arctic (~170m) and then the Antarctic (~75m). The ascent of an aircraft is fuel intensive so a longer ascent will inevitably require more fuel (~0.04% for a >1000km flight).

Thermal efficiency changes at 200hPa (Their figures 5&6). As the ambient air gets warmer, the thermal efficiency drops. This drop in efficiency between the RCP8.5 scenario and historical runs is greatest at low latitudes (~-0.5%) and smallest in the Antarctic (~-0.03%).

Mechanical efficiency changes at 200hPa (Their figures 7&8). In contrast to the previous 2 variables, a warmer ambient air temperature is likely to increase the engine mechanical efficiency. For the RCP8.5 this mechanical efficiency was maximum at the low latitudes (~+0.3%) and at a minimum in the Antarctic (~+0.01%).

To summarise so far, the mechanical and thermal efficiencies largely cancel out and the increase in fuel use due to the tropopause height increase is small. However, an increase of up to 2% dynamic viscosity (and consequently skin frictional drag) due to the increase in temperature looks likely to increase total drag and thus fuel consumption in 2100 relative to 2010 by ~0.22% resulting in an estimated +~160million additional gallons of fuel being used globally each year.

2.5.4 A reduction in aircraft take-off performance.

So increased CAT and a more viscous atmosphere look likely to impact on flights at cruise altitude, but what about the take-off phase itself?

According to the equation of state for a gas (rho=pM/RT), a gas held at constant pressure will increase in volume and decrease in density if heated.

This holds true for any layer of the atmosphere that warms and consequently the density of the troposphere will drop will decrease as it warms over the course of the 21st century. Now as the amount of lift generated on an aircraft wing is directly proportional to both the air density and the square of the velocity of the aircraft, a decrease in air density will reduce the amount of lift generated at a given velocity.

On particularly hot days on airport runways, this can be compensated for by increasing the take-off distance provided the runways are long enough to accommodate this. If not, then the take-off mass of the aircraft itself must be reduced (typically by unloading passengers and cargo) in order for the aircraft to achieve enough lift for take-off by the end of the runway.

This phenomenon has already been observed in the increasing number of both 10k and 15k lbs weight restricted days for a Boeing 737-800 observed in Summer (MJJAS) over the 1980-2010 period at 4 US airports in a study by Coffel and Horten (2015 - hereafter referred to as C&H15). Ren et al. (2019 - hereafter referred to as REA19) followed this study up by calculating the near surface air density (and consequent reduction in maximum take-off weight MTOW) at 6 international airports over the 1950-2015 period from

the NCEP/NCAR reanalysis dataset and a noticeable reduction/increase in average air density can already be observed in some/none of them. Similarly, Graton et al. (2020) looking at 10 Greek airports in the period 1950-2020 observe a general trend of increasing take-off distances and/or reductions in maximum take-off mass (e.g. for a larger turbofan). Averaging over multiple Greek airports, there has been an increase of 2.7m(0.15%)/year at full weight, and at runway-limited Chios & Skiathos, a drop in payload per take-off by 133kg/year and 18kg/year respectively.

Both C&H15 and REA19 use climate models (17 and 24 CMIP5 models respectively) to project how these impacts might play out into the future (to 2070 and 2100 respectively) given an RCP8.5 high emissions scenario and both studies find the trends are set to continue and become increasingly significant into the future. As well as the temporal trend, REA19 note the non-uniform spatial nature of the impact on air density/MTOW with the high latitudes, and the NH high latitudes in particular being most adversely affected by the changes. This makes sense given both the current and projected non-uniform spatial heating pattern (e.g. Arctic amplification) associated with climate change (See Figure 5 in REA19).

As climate change may also impact the winds (global stilling) which in turn have impacts on take-off performance (a reduced headwind increases take-off distance), Gratton et al. (2020) extends the analyses of the previous 2 studies to also take into account changes in the wind vectors. While changes in the wind direction and strength could potentially affect the take-off performance in either an adverse or proverse manner, in this study, it usually worked to further reduce take-off performance though wind effects were less significant than temperature effects.

C&H15 provides the most detailed discussion of the degree to which the industry can potentially adapt to this impact. Options include extending runways or building new ones, substituting in higher performing aircraft, whilst poorer performing ones can be rescheduled out of the hottest parts of the day. However, not all these options are available to every airport and reduced operations and profits may be the inevitable result, especially as the climate becomes hotter and the changes to MTOW, larger. It is difficult to evaluate how big the economic impact from this issue appears from the perspective of the industry. Perhaps it is a serious concern, or maybe a minor annoyance in the grand scheme of things at the moment, but any impacts are likely to be reduced with forward planning, so the issue is certainly worth their attention especially as it is projected to worsen.

2.5.5 In summary

Climate change appears to be already impacting aviation in a number of ways, 4 of which (CAT, flight-times, viscousness of the atmosphere and take-off performance) have been discussed above. The literature often compares pre-industrial climate simulations with the higher end climate change scenarios simulated out to the 2nd half of the 21st century. This is useful, as it gives a (hopefully) upper limit on what can be expected, information that should ideally help guide government and industry policy makers and planning

for the future. In addition to modelling studies, there is also observational evidence that the impact on aviation by climate change is already beginning to be felt (e.g. REA19, REA20, C&H15 and GEA20) particularly in the literature on take-off performance. Unfortunately, without radical global emissions cuts, these impacts look set to grow. From the above literature it is somewhat difficult to get a sense of the degree to which the industry is cognisant of climate impacts, where it ranks in their list of 'most pressing problems' and the extent/significance of their plans for adaptation.

2.6 Impact of climate variability on CAT

A literature review on how climate oscillations/variability impact on aviation (via turbulence/CAT).

Climate oscillations are recurring pattern of variation in the Earth's climate system over periods from months to decades, and they affect variables such as temperature, precipitation, atmospheric pressure and ocean currents. The main sources of variation that will be considered here are ENSO and the NAO

The reason the 'via turbulence/CAT' has been added in brackets above is because much of the literature to date focuses on how climate oscillations/variability have affected aviation via flight routing/travel time. The NAO for example affects the strength of the north Atlantic jet which trans-Atlantic commercial aircraft either use (eastbound) or try to avoid (westbound). If journey time increases because of the variation in these jets, then costs go up as does fuel consumption and the consequent climate impacts of that fuel consumption. So although this literature will touch on this aspect, the main focus of it will remain the turbulence impacts on aviation from climate oscillations/variability.

Figure 1 of Wolff and Sharman (2008, Figure 2.13), shows the monthly averaged PIREP count per day for total, MOG and SOG CAT between 1994-2005 over the continental US. There is obvious seasonal variability in the figure with winter counts being much higher than summer. However, apart from this high frequency variability, there is a peak of over 300 counts/day during the winter of 1998, just after the large El-Nino event that had just occurred. The hypothesis is that perhaps El-Nino contributed to this peak in CAT encounters.

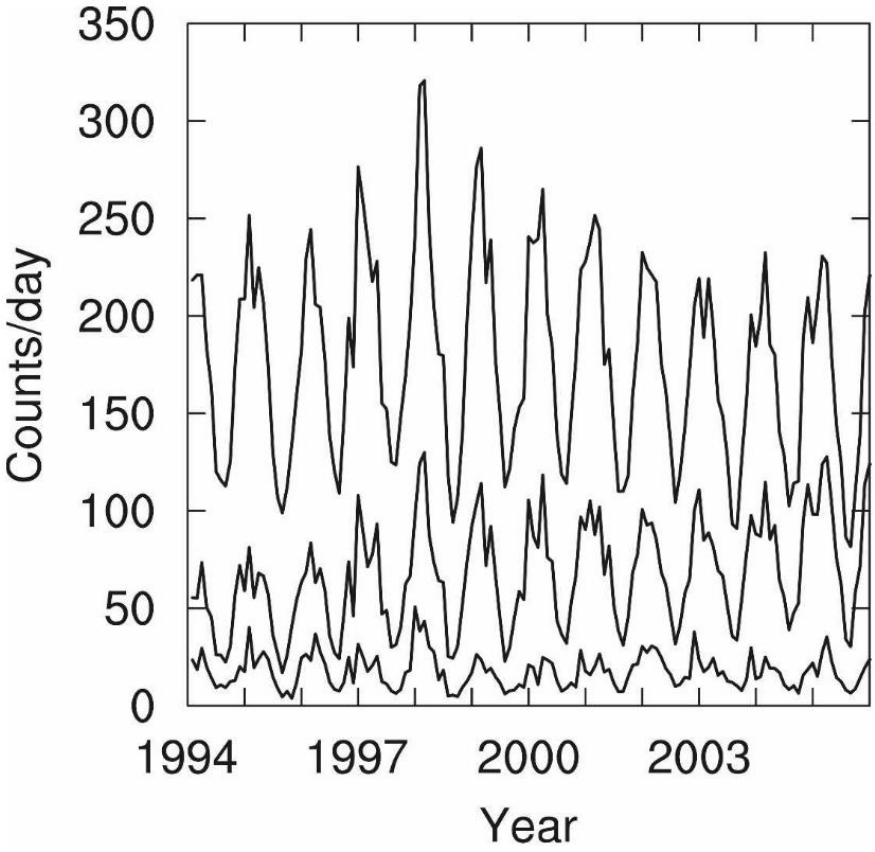


Figure 2.13: Monthly averaged counts per day from January 1994 through December 2005 of total (top line), MOG (middle line), and SOG (multiplied by 4 to enhance the trace; bottom line) PIREPs. Taken from Wolff and Sharman (2008).

On the supporting side, if one looks at Figure 7 (see Figure 2.14 for a reproduction) in the thesis of Cheyne (2020), it is evident that the wind-shear anomaly at 250hPa is positive (more sheared than usual) in a composite of strong El-Nino years between 1979-2019. As vertical shear is a key ingredient in the development of CAT, such a vertical wind shear anomaly over the US does indeed suggest a greater frequency of turbulence and thus a greater likelihood of CAT encounters.

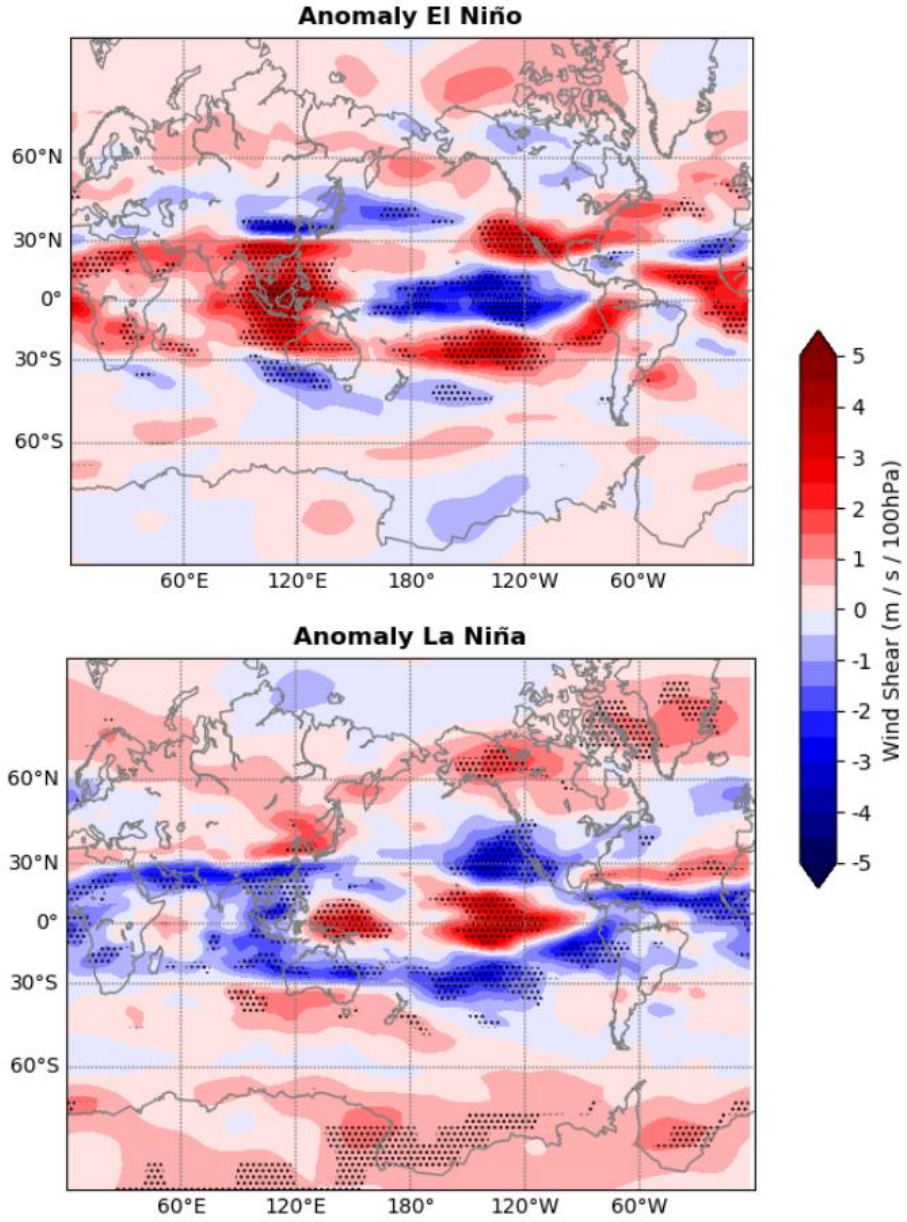


Figure 2.14: Anomaly 250 hPa vertical wind shear in extreme El Niño and La Niña composite events showing deviation from average DJF conditions (1979- 2019). Red and blue signalling wind shears higher and lower than average respectively. Stippling indicates deviations significant at the 95% confidence level. Taken from Cheyne (2020).

However this hypothesis is somewhat challenged by the fact that after the strong El-Nino of 1997/1998, comes an equally strong La-Nina in 1999/2000. The work of Cheyne (2020) suggests that a La-Nina produces a reduction in the vertical wind shear anomaly over the US, which might suggest less frequent CAT and fewer encounters and yet no such rapid drop off is evident in the Wolff and Sharman figure 2.13 above.

It also bears mentioning that there are numerous problems with PIREPS themselves which means that they may not always be reflective of the underlying CAT-frequency distribution over the US 250hpa atmosphere (see Schwartz, 1996).

Despite this, the hypothesis remains intriguing and given the evident widespread effect on the global wind-shear anomaly, a path for ENSO to affect CAT is clearly there.

Coming back to the NAO, work by Kim et al. (2016) examines the difference in flight routing and likelihood of encountering CAT between a winter with a positive NAO (04/05) versus a winter with a negative NAO (09/10), this of course impacts the flight time (+NAO eastbound is fastest, the -NAO eastbound, -NAO westbound then +NAO WB being slower). In terms of proportion of MOG CAT, the eastbound flights have more CAT than the westbound flight, the explanation being that eastbound flights fly nearer the jet and regions of horizontal and vertical windshear, with -NAO winter eastbound giving the greatest proportion of CAT (1.31% vs 1.02% MOG for the +NAO eastbound). This is explained by the jet stream shifting to the south in -NAO meaning that planes travelling on the shortest distance 'great circle route' will be travelling on the cyclonic (poleward) side of the jet, which is thought to be more prone to CAT than on the anticyclonic side.

In regards to how the NAO affects the mean latitude of the DJF jet (the effect in JJA is not so great), figure 1a (see Figure 2.15) in Irvine et al. (2013) shows a useful timeseries of this between the period 1989-2009 in ERA-Interim.

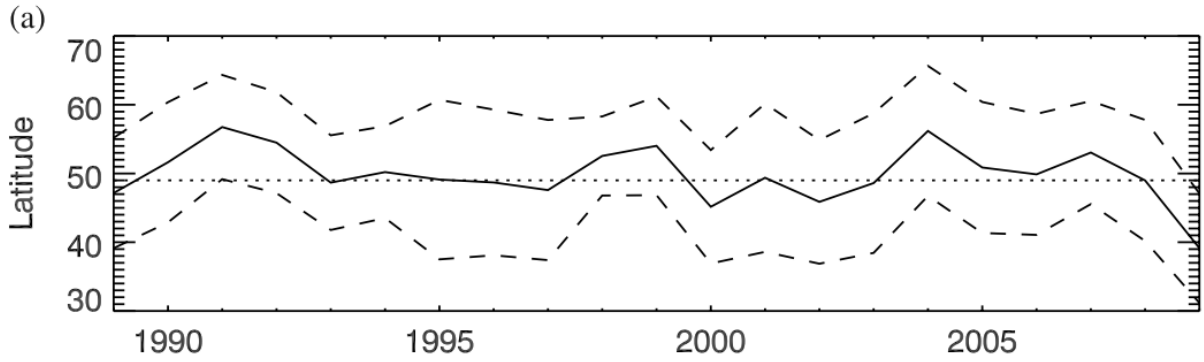


Figure 2.15: Time series of the winter (DJF) mean jet latitude in ERA-Interim (solid line) and the intra-seasonal standard deviation (dashed lines). The 1989–2009 winter mean jet latitude is overlaid on (a) (dotted line). The year on the axes corresponds to the December. Taken from Irvine et al. (2013).

It confirms a high degree of correlation, and an even higher degree if both the NAO and the East Atlantic (EA) index are used (Woollings et al., 2010).
As most (2/3rds) of the CAT in the cruise level atmosphere is located within the vicinity of the jet stream, so its position is key to understanding the geographical distribution of the turbulence.

The rest of the literature reviewed looks at how flight times and routes are affected by climate oscillations and doesn't particularly look at turbulence. It is however still worth mentioning as where climate oscillations cause jets to become stronger/weaker, shear can be affected (A faster jet will naturally cause greater horizontal and vertical shear with the stiller air surrounding it), which increases the likelihood of turbulence. For example, Kim et al. (2020) following up on Kim et al. (2016) examining another flight route between HNL (Honolulu airport) and SFO (San Francisco airport) in the subtropical Pacific as well as the JFK (John F. Kennedy airport) and LHR (London Heathrow). They choose 3x +NAO, -NAO, +ENSO, -ENSO winters

and examine how the flight routes are affected. They find a 4.24 ~ 9.35 min increase in round trip flight time for NAO+ relative to NAO- and a 5.92 ~ 8.73 min increase for +ENSO winters relative to -ENSO in the subtropical Pacific. How this affects turbulence is not examined.

Karnauskas et al. (2015) (prior to Kim et al., 2020) also examine the upper level (300hPa) winds between Hawaii and the continental US. Rather than looking at round trip times as Kim et al. (2020) did, they define a variable deltaT which is the difference between eastbound (faster) and westbound (slower) flights. The interannual variability of delta T correlates well (0.85, 0.91 with a 2-3 month lag) with ENSO and also with the Arctic Oscillation (AO; -0.48). They also don't look at turbulence, but given the finding of Kim et al. (2016), it would be surprising if turbulence wasn't affected by the changes in wind speed.

Chapter 3

Using aircraft accidents and incidents to evaluate the skill of clear-air turbulence diagnostics

3.1 Introduction

In the US alone, Pilot Reports (PIREPs) indicate that 63,000 aircraft encounter moderate–or–greater turbulence each year. In addition to that number, 5000 aircraft encounter severe or greater turbulence each year (Sharman et al., 2006). This costs the aviation industry an estimated $150-$500 million annually in injuries to passengers and crew as well as damage to cabins and air frames. For aircraft in the general aviation (GA) category, meaning aircraft with neither commercial or military use, around 40 lives are lost every year due to turbulence (NCAR-RAL, n.d.).

Clear-air turbulence (CAT), a common type of turbulence, is particularly problematic because it cannot yet be detected remotely in a way that might enable aircraft to reliably avoid it (Kopeć et al,, 2016). CAT forecasts, while improving, are still not good enough to prevent encounters, some which involve multiple hospitalisations from occurring each year (Storer et al., 2019a). Sharman et al. (2012) states that CAT prediction is, “one of the last great challenges of numerical weather prediction”.

With perhaps 23% of aircraft accidents and incidents being caused by turbulence in clear air (Eick, 2014), CAT is clearly already an issue for the aviation industry. However, recent work by (Williams and Joshi, 2013; Williams 2017; Storer et al., 2017 and Lee et al., 2019) suggest that it is likely to become an even bigger problem with increased levels of global warming. This, along with a projected increase in aviation passengers means that if the industry is to stand a chance of successfully adapting, it needs to have as clear a picture as possible of how much, where, and by when, CAT is likely to change throughout this century.

One tool the aviation industry uses to avoid turbulence is to use a turbulence forecast issued by a World Area Forecast Centre. One of the leading turbulence forecast systems was developed by Sharman et al. (2006). Named the Graphical Turbulence Guidance (GTG) forecasting system. The system uses a suite of different CAT diagnostics to achieve this. The majority of CAT diagnostics are derived from empirical relationships which are calculated from the standard outputs (T, u, v, P) of a Numerical Weather Prediction (NWP) model. Based on altitude, geographic region, and performance in the previous forecasting window each of the diagnostics are given weightings before being combined them to give a singular forecast output. GTG achieves this by using aircraft turbulence encounters from the local forecasting area to evaluate and verify the multiple CAT diagnostics that the system uses. For example, the version of GTG covering East Asia (Kim and Chun, 2011) has different weightings to its US counterpart. Additionally Sharman et al. (2006) showed how the CAT diagnostic configuration changed based on altitude.

The turbulence we care most strongly about forecasting is turbulence that causes injuries. However, to the best of our knowledge, no previous study has used a database of injury-causing turbulence encounters to test the ability of CAT indices to diagnose these events. That is therefore the goal of this chapter.
In addition to this, knowing what the best performing CAT diagnostics for injury causing turbulence are will enable improved estimates to the frequency of injury causing CAT events for various climate projection scenarios.

In this paper aircraft accidents and incidents involving CAT that were serious enough to warrant an National Transportation Safety Board (NTSB) report (NTSB, n.d.) or a media report (AvHerald, n.d.) are collected and then 21 CAT diagnostics are applied to their space and time coordinates using the ERA5 reanalysis dataset ERA5 (n.d.). The percentiles of these diagnostic values relative to a global distribution are then ascertained for each of the 80 events and then plotted as histograms to see which of the CAT diagnostics have the most skill in diagnosing aviation-affecting CAT.

Once the 80-event means are calculated for each diagnostic, their skill scores will then be used to recalculate the 21-diagnostic mean increases in north Atlantic CAT between a pre-industrial control and doubled-CO2 simulation produced by Williams (2017) and to refine a previously produced global CAT projection by Storer et al. (2017) in terms of injury-producing aviation affecting CAT. Section 3.2 discusses CAT in more detail including the severity classifications. Section 3.3 discusses the collection of the datasets used. Section 3.4 evaluates the skill of the 21 CAT diagnostics calculated from ERA5 during the 80 severe encounters. Section 3.5 takes the derived skill to refine a projections of increased North Atlantic CAT in a doubled-CO2 scenario and global CAT in a RCP8.5 climate change scenario. A conclusions and summary are given in section 3.6

3.2 Clear-air Turbulence (CAT)

CAT is an invisible type of turbulence, due to it occurring in clear air often distant from thunderstorms and mountains, two of the other main sources of atmospheric turbulence. CAT is predominantly found

in the vicinity of jet streams, which have regions of strong wind speeds and are surrounded by areas of strong vertical and horizontal shears. Near the jet stream the probability of CAT is about three times larger than in areas of weak winds (Reiter 1963c). The main mechanism by which CAT is generated is thought to be Kelvin-Helmholtz instability (Ellrod and Knapp, 1992), which occurs in regions of the atmosphere where the Richardson number (*Ri*) <0.25. *Ri* is given by equation 3.1.

$$\mathrm{Ri} = -\frac{g}{\rho}\frac{\frac{\partial \rho}{\partial z}}{\left(\frac{\partial u}{\partial z}\right)^2} \tag{3.1}$$

where g is the acceleration due to gravity and ρ is the density of the fluid, u is the velocity of the fluid, and z is the height co-ordinate. *Ri* effectively compares the strength of the vertical stratification (numerator) to the wind shear (denominator). At sub critical values (Ri<0.25) the intensity of the wind shear exceeds that which can be restored through stratification and Kelvin-Helmholtz instability occurs leading to turbulence. Gravity waves (GWs) and inertia-gravity waves (IGWs), which cause perturbations in pressure, temperature and wind velocities, can also induce CAT either by breaking or reducing an otherwise stable environment's *Ri* to a sub critical value (e.g. Knox et al., 2008; McCann et al., 2012 and Fritts 2015).

3.2.1 Other types of turbulence in clear air

As discussed earlier there are two other sources that form turbulence in clear air which are classified differently due to their proximity to their sources.

The first is Mountain-Wave Turbulence (MWT) which typically occurs above and around mountains. Pilots are well aware of it and try to avoid flying in the downstream regions if the winds are strong. Nevertheless, MWT can sometimes be mistakenly reported as CAT, as aircraft are often flying in clear air when they experience it. MWT can also occur some distance away from the mountains that generate it, depending on the local meteorological conditions. Sometimes CAT is defined such that it includes MWT (e.g. Kim and Chun, 2010), but here the two are separated due to the nature of the turbulence diagnostics used.

Second, Near-Cloud Turbulence (NCT) which is clear air turbulence that occurs in the vicinity of intense convective cells (Lane et al., 2012). The strong vertical currents of air within the the convective cell are highly turbulent and that turbulence is called Convective Induced Turbulence (CIT). For this reason, pilots typically avoid regions of obvious convection (e.g., towers of cumulonimbus clouds). Near-Cloud Turbulence (NCT) occur when convective updrafts collide with the tropopause and generate GWs (Fovell et al., 1992). These GWs can travel outwards and cause turbulence due to wave breaking at significant

distances (several hundreds km), depending on the synoptic situation, Because of this, NCT is often misidentified as CAT.

Third, Wake Vortex Turbulence (WVT) is generated whenever aircraft fly through the fluid atmosphere. The wake can quickly collapse into turbulent eddies (Lester, 1993). Any aircraft flying too close behind such a wake can quickly be affected by the turbulence. For this reason, there are minimum required separation distances between planes throughout the whole take-off, cruise, and landing cycle. Despite these measures, WVT events still occur due to their unpredictability (for recent examples, see AvHerald, n.d.).

Finally, Boundary-Layer Turbulence (BLT) occurs when wind blowing over complicated topography breaks down into turbulence. This often occurs in the lowest few km of the atmosphere.

3.2.2 CAT thresholds

CAT is graded in its intensity, with the severe category causing vastly more impact on the industry than the light category. Williams (2017) defined five different strength categories for aviation-affecting CAT based on observations, which are shown in Table 3.1. As the 80 CAT events that are to be examined in this paper have resulted in media reports and (typically) hospitalisations of passengers and crew, they are highly likely to have been examples of severe-or-greater CAT.

Table 3.1: The defining characteristics of six turbulence strength categories for a large commercial aircraft (from Lane et al., 2012; Williams, 2017).

Threshold	**Percentile range**	**Magnitude of aircraft vertical acceleration (g)**	**EDR$^{1/3}$ range ($m^{2/3}$ s^{-1})**	**Passenger experience**
Null	0-97.0	0–0.2	0-0.1	-
Light	97.0-99.1	0.2–0.4	0.1-0.2	A slight strain against seat belts; unsecured objects may be displaced slightly; food service may be conducted with little difficulty walking
Light-to-moderate	99.1-99.6	0.4-0.6	0.2-0.3	-
Moderate	99.6-99.8	0.6–0.8	0.3-0.4	Definite strain against seat belts; unsecured objects are dislodged; food service and walking are difficult
Moderate-to-severe	99.8-99.9	0.8–1.0	0.4-0.5	-
Severe	99.9-100	>1.0	>0.5	Occupants are forced violently against seat belts; unsecured objects are tossed about; food service and walking are impossible

Much of the previous work on evaluating the skill of CAT diagnostics has been done using Pilot Reports (PIREPs; e.g. Sharman et al., 2006; Kim and Chun, 2011; Kim et al., 2011) in combination with Receiver Operating Characteristic (ROC) curves (Gill, 2016). Reported aircraft incidents which mention CAT are

used here, instead of PIREPs, as these are likely to be the result of severe or extreme turbulence and therefore should present a strong signal for the 21 diagnostics to diagnose.

3.2.3 CAT diagnostics

As CAT is a sub-grid scale phenomenon, it cannot be explicitly simulated in numerical weather prediction models. However, most of the energy of the eddies that cause aviation-affecting CAT (on length scales of the order of 100~m) cascade down from larger resolvable scales of atmospheric motion. If these meteorological features can be accurately identified, in principle, improvements in the forecasting of CAT are possible (Sharman et al., 2006). On this basis, a number of CAT diagnostics based on the physical mechanisms known to generate CAT have been developed and are applied to output data from numerical models to indicate regions of high CAT probability.

Table 3.2 lists the 21 CAT diagnostics used here, as first used in Williams and Joshi, (2013) along with the reference to the paper that first introduced each diagnostic (For diagnostic equations, see Appendix A). Marlton, (2016) reviewed the diagnostics shown here and found that the majority of diagnostics included one or more of the following physical mechanisms: Richardson number (Diagnostics 1, 2 & 19), deformation (Diagnostics 4, 5, 6, 7, 8, 9, 15 & 16) and vorticity (Diagnostics 5, 6, 10, 19, 20 & 21). Seven of the 21 CAT diagnostics used in the present paper are used in the US Graphical Turbulence Guidance (GTG2) product (Sharman et al., 2006) to forecast CAT over the US and are marked with an asterisk.

Table 3.2: The 21 CAT diagnostics calculated in this paper. $*$ and $\vee$ indicate GTG2 upper-level (FL200+) and mid-level (FL100-FL200) diagnostics, respectively, as given in Sharman et al. (2006). See Appendix A for diagnostic equations.

GTG2?	No.	Diagnostic name	Units	Reference
$*$	01:	Negative Richardson number	—	Endlich (1964) Kronebach (1964) Dutton and Panofsky (1970)
	02:	Magnitude of vertical shear of horizontal wind	s^{-1}	—
$*$	03:	Colson-Panofsky index	kt^2	Colson and Panofsky (1965)
$*\vee$	04:	Frontogenesis function	$m^2 \cdot s^{-3} \cdot K^{-2}$	Bluestein (1993)
	05:	Brown index	s^{-1}	Brown (1973)
	06:	Brown energy dissipation rate	$J \cdot kg^{-1} \cdot s^{-1}$	Brown (1973)
$*\vee$	07:	Variant 1 of Ellrod's turbulence index	s^{-2}	Ellrod and Knapp (1992)
	08:	Variant 2 of Ellrod's turbulence index	s^{-2}	Ellrod and Knapp (1992)
	09:	Flow deformation	s^{-1}	Ellrod and Knapp (1992)
	10:	Magnitude of potential vorticity	PVU	Knox (2001)
	11:	Relative vorticity squared	s^{-2}	Sharman et al. (2006)
$*\vee$	12:	Magnitude of horizontal temperature gradient	$K \cdot m^{-1}$	Buldovskii et al. (1976)
$\vee$	13:	Wind speed	$m \cdot s^{-1}$	Endlich (1964)
	14:	Wind speed times directional shear	$rad \cdot s^{-1}$	Sharman et al. (2006)
$\vee$	15:	Flow deformation times wind speed	$m \cdot s^{-2}$	Reap (1996)
	16:	Flow deformation times vertical temperature gradient	$K \cdot m^{-1} \cdot s^{-1}$	Reap (1996)
$*$	17:	Magnitude of residual of nonlinear balance equation	s^{-2}	Knox (1997) McCann (2001) Koch and Caracena (2002)
	18:	Magnitude of horizontal divergence	s^{-1}	Ellrod and Knapp (1992)
$*\vee$	19:	Version 1 of North Carolina State University Index	s^{-3}	Kaplan et al. (2004) Kaplan et al. (2006)
	20:	Negative absolute vorticity advection	s^{-2}	Bluestein (1992)
	21:	Magnitude of relative vorticity advection	s^{-2}	Knox et al. (2008)

3.2.4 2050-2080 CAT projections using diagnostics

In order to refine climate projections of CAT it is important to know the skill of each diagnostic under severe turbulence conditions, for which observation model pairs are required. The acquisition of such a data set is described in the next section.

3.3 Data extraction methodology

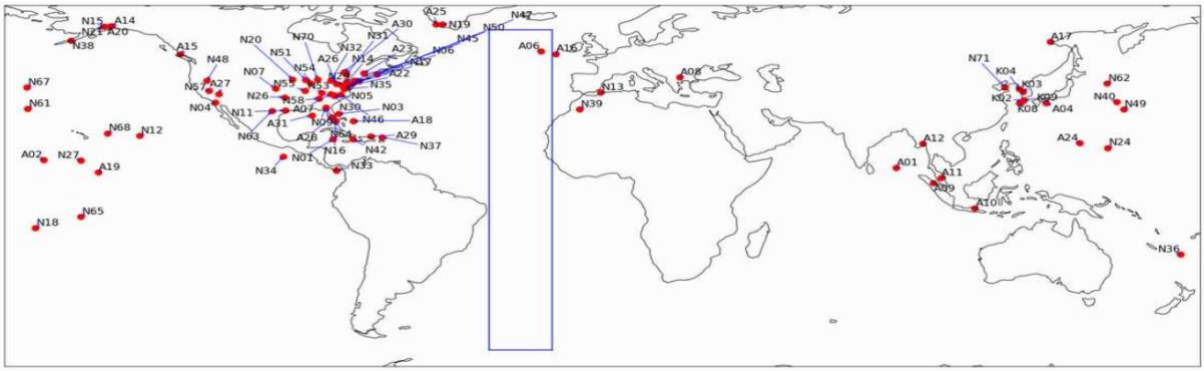
In order to build the observation model pairs, firstly a set of events where aircraft had encountered severe CAT were required. Secondly a gridded data set containing the standard meteorological variables to allow the calculation of the 21 turbulence diagnostics.

3.3.1 CAT events

In this study, an observation set containing only severe turbulence encounters is needed. The Aviation Herald (AvHerald, n.d.) and the United States' National Transportation Safety Board (NTSB, n.d.) archive reports of a wide range of accidents and incidents involving turbulence. These archives were searched for the term “CAT” in the article title or main text, to ensure other forms of turbulence described in section 3.2 were omitted. Further quality checks were undertaken so that any events under FL100 were excluded, as were events that took place directly over or downwind of the Andes, Rockies, Alps, Himalayas or Greenland, in case they were instances of BLT or MWT were misinterpreted as CAT. Any events declared as CAT occurring in regions of obvious convective activity (potentially NCT) according to the article or in the vicinity of another aircraft (potentially WVT) were also excluded. However despite this due diligence it is difficult to rule out the possibility that some of these events were caused by (or had a contribution from) other forms of turbulence such as CIT.

Finally, in order for the reports to be usable, three pieces of information for each event were required: the date and time; the latitude and longitude; and the pressure altitude (flight level).

In total 80 CAT events (occurring between 1994-2019) severe enough to be reported were identified across the globe as shown in Figure 3.1. The coordinates in time and space for each of the 80 CAT events can be found in Appendix B. Of the 80 CAT events identified, 36 were found over the continental US and are



shown in Figure 3.2. 26 events were extracted from (AvHerald, n.d.); 53 events from (NTSB, n.d.); and finally one well documented event from Lee and Chun (2018).

Figure 3.1: The geographical location of the 80 CAT events used in this study. The blue rectangle over the Atlantic ocean represents a four-dimensional hyper-cube of data referred to as the "Atlantic climatology", a sample from which global CAT diagnostic distributions were inferred. Events with A, N & K as the first letter are events garnered from AvHerald.com, NTSB and Lee and Chun (2018) respectively.

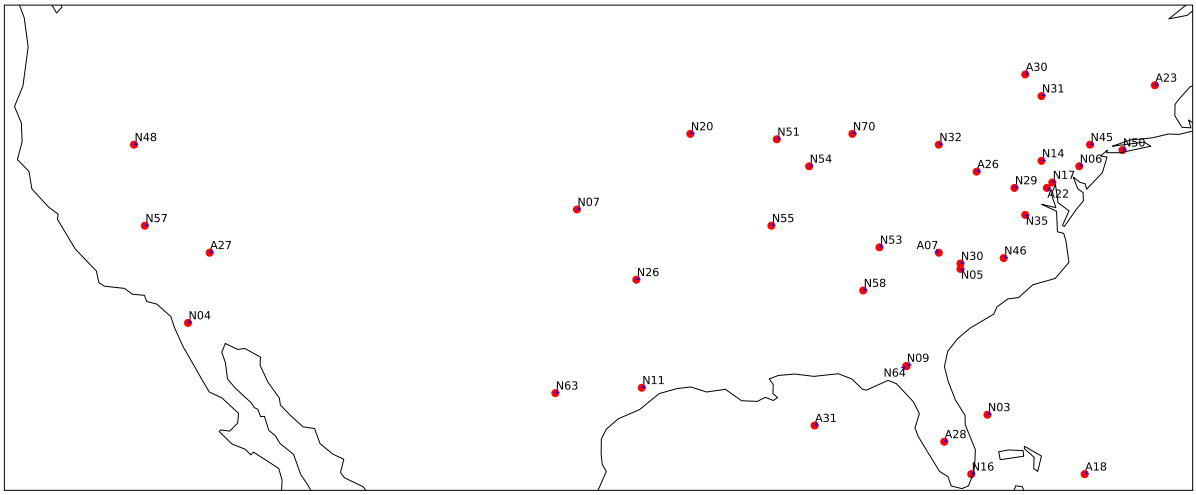


Figure 3.2: The North American events zoomed in.

3.3.2 ERA5

An atmospheric reanalysis dataset is one that produces a representation of the state of the atmosphere atmospheric data a historical period by assimilating observations over that period to produce a gridded dataset constrained both by the physics of the model and the assimilated observations. They are commonly used to study past weather and climate.

ERA5 (Hersbach et al., 2020) recently replaced ERA-Interim as the ECMWF's current climate reanalysis dataset. When completed, it will cover the time period from 1950 to the present. The model contains 137 pressure levels in the vertical and has a TL639/N320 horizontal resolution, which corresponds to about 0.28 degrees at the equator. Events occurring between the years of 2000-2006 inclusive were downloaded from ERA5.1 (only available for the years 2000-2006 inclusive) as the original ERA5 dataset had a cold bias in the lower stratosphere during this period (Simmons et al., 2020).

3.4 Evaluating diagnostic skill

3.4.1 Method

Using the coordinates of the 80 CAT events, their individual latitudes, longitudes, times, and altitudes were rounded to the nearest ERA5 grid point (the nearest quarter of a degree), time step (hourly) and pressure level. As aircraft altitudes are often reported in meters or feet for each event, these were first converted into atmospheric pressure in mb using (NOAA, n.d.):

$$\text{mb} = 1013.25 \times (1 - \frac{\text{meters}}{44307.69396})^{.190284} \tag{3.2}$$

A 3D cube centred on these rounded coordinates ±2.5 degrees in latitude and longitude across the 298-197~hPa range for the nearest hour to the event in question was then downloaded from ERA5. A 3D cube of gridpoints rather than a single point was downloaded as many diagnostics have first and second order derivatives which require neighbouring grid points. Some events had geographical uncertainty associated with them and consequently, diagnostic values at multiple neighbouring grid points were required (See Appendix B). The 21 diagnostics (See Table 3.2) were then calculated for the grid point(s) closest to the event. This process was repeated for each of the 80 CAT events.

Calculating the 21 diagnostics in ERA5 produces absolute values. However, rather than absolute values, we are interested in how extreme these values are for the ERA5 200-300hPa atmosphere specifically (diagnostic absolute values are model-dependent). To get a sense of this, and to facilitate inter-diagnostic comparison, the raw diagnostic values are first converted into cosine of latitude weighted percentiles. To convert a given diagnostic's absolute value into a weighted percentile relative to the ERA5 atmosphere, that diagnostic's climatological distribution in ERA5 needs to be known. To estimate the climatological distribution a sub sample of data was taken from 6 hourly timesteps for the year 2016, at ten pressure levels between 298-197~hPa, and over the geographical area -60-60N and 34-15W depicted by the blue box in Figure 3.1. This sample, hereafter referred to as the "Atlantic climatology" was chosen as the subsample to compute the climatological distributions from as it permits a wide range of latitudes whilst remaining over the ocean, guaranteeing the exclusion of MWT. For a discussion on the implications of choosing the Atlantic climatology sub sample, please refer to Appendix C.

To check whether 2016 was a representative year or not, the 21 CAT diagnostic percentiles for a single example event (A16) that occurred in 2016 (near the west coast of Ireland) were calculated relative to both a 2016 time series and a 10 year (2007-2016) time series at the nearest ERA5 gridpoint to A16's location in time and space (See Figure 3.4). The correlation between 2016 and the 2007-2016 diagnostic percentiles was high (r=99.8%) and led to the conclusion that 2016 was sufficiently representative.

The pressure levels were chosen for the Atlantic climatology as these are the pressure levels between which aircraft typically cruise and also where CAT peaks. This final assumption can be asserted by computing the 21 CAT diagnostics across the Atlantic Climatology domain and normalising them and producing an averaged profile shown in Figure 3.3.

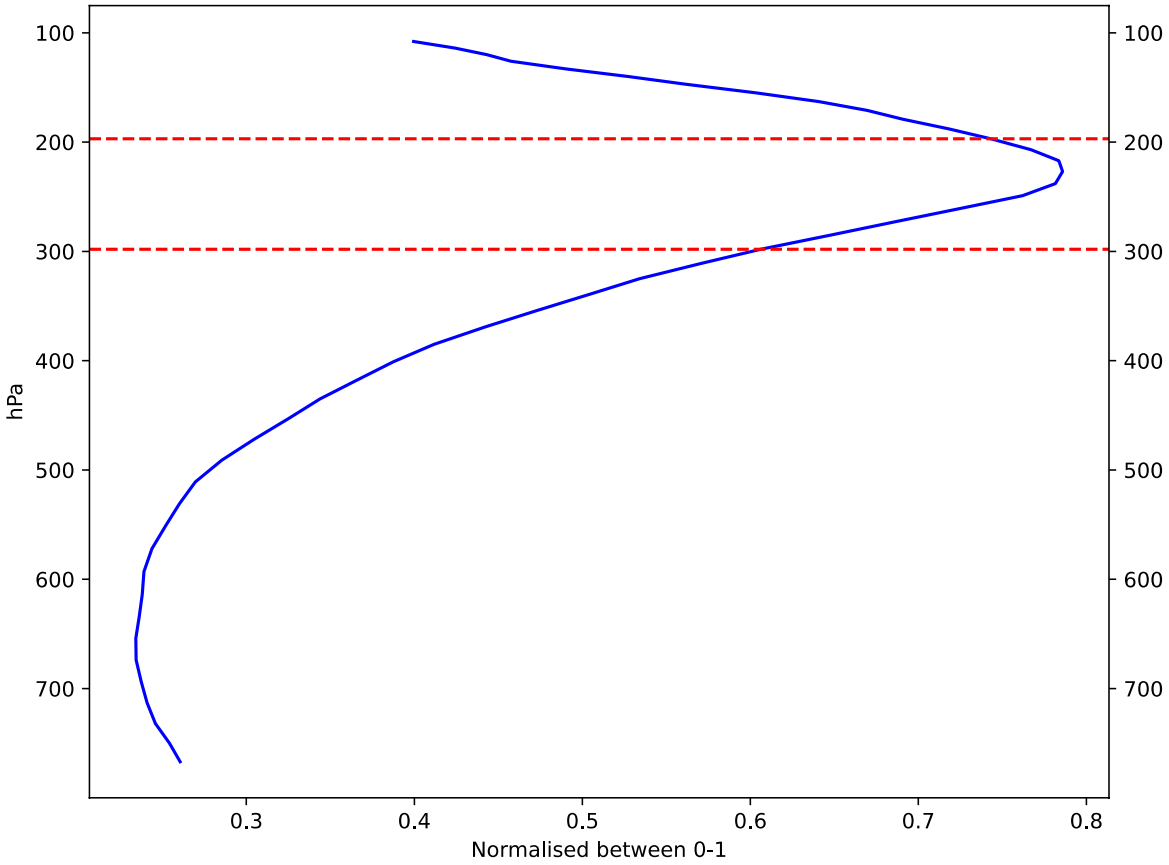


Figure 3.3: The 21-diagnostic mean with altitude. The figure was computed by taking the time and area mean of the Atlantic climatology for each diagnostic in turn. Then in order to calculate a 21 diagnostic-mean altitude profile, each diagnostic was first normalised between 0-1 (lowest-highest value for that diagnostic's profile). The red dotted lines indicate aircraft typical cruising ranges. Altitudes above and below this range were excluded from the Atlantic climatology sample.

The values of the 21 diagnostics were calculated at each grid point of the Atlantic climatology. A weighted percentile function was then applied to each of the Atlantic climatologies' 21 diagnostic distributions. This gave a correspondence between raw diagnostic values and percentiles with 0.1 percentile resolution for each diagnostic. A weighted percentile was chosen over an unweighted one to take into account the decreasing grid box area at higher latitudes.

Finally, to evaluate the skill of the 21 diagnostics at diagnosing injury-causing CAT, the diagnostics need to be ranked by their 80-event means, expressed as a percentile of the climatological distribution.

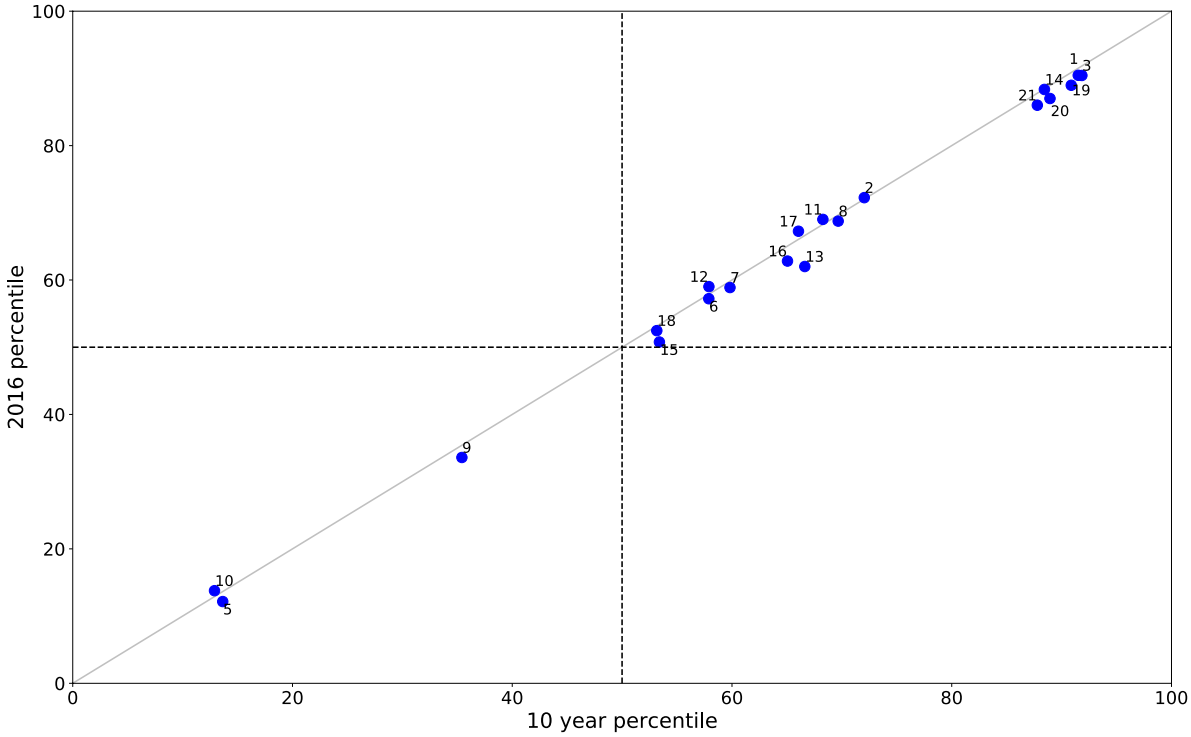


Figure 3.4: A comparison between diagnostic percentiles calculated from 2016 only distributions versus a 2007-2016 distribution for the event A16. Both the 1 year (8784 hourly values) and 10 year distributions (87672 hourly values) were drawn from the nearest ERA5 gridpoint to the A16 event. The numbers refer to the diagnostics numbers which are listed in Table 3.2.

In the case of events with low location accuracy for example the event which was "approximately 500 nautical miles north west of Tokyo" a mean of the percentiles was taken across the grid points with a longitude - latitude bounding that had a high probability of encompassing the aircraft's true geographical location.

3.4.2 Results

For each diagnostic, the distribution of percentiles across the 80 events are plotted as histograms shown in Figure 3.5. The vertical dashed lines show the 80-event mean (green if ≥66.7%, red if <60%, grey if in between these values) and its percentile value is shown in the top left of its respective subplot. Diagnostic percentile distributions that are statistically significantly different to a uniform distribution (which is typically the case if random coordinates are chosen) according to a chi-squared test (p=0.05/24) are marked with an asterisk to their top right.

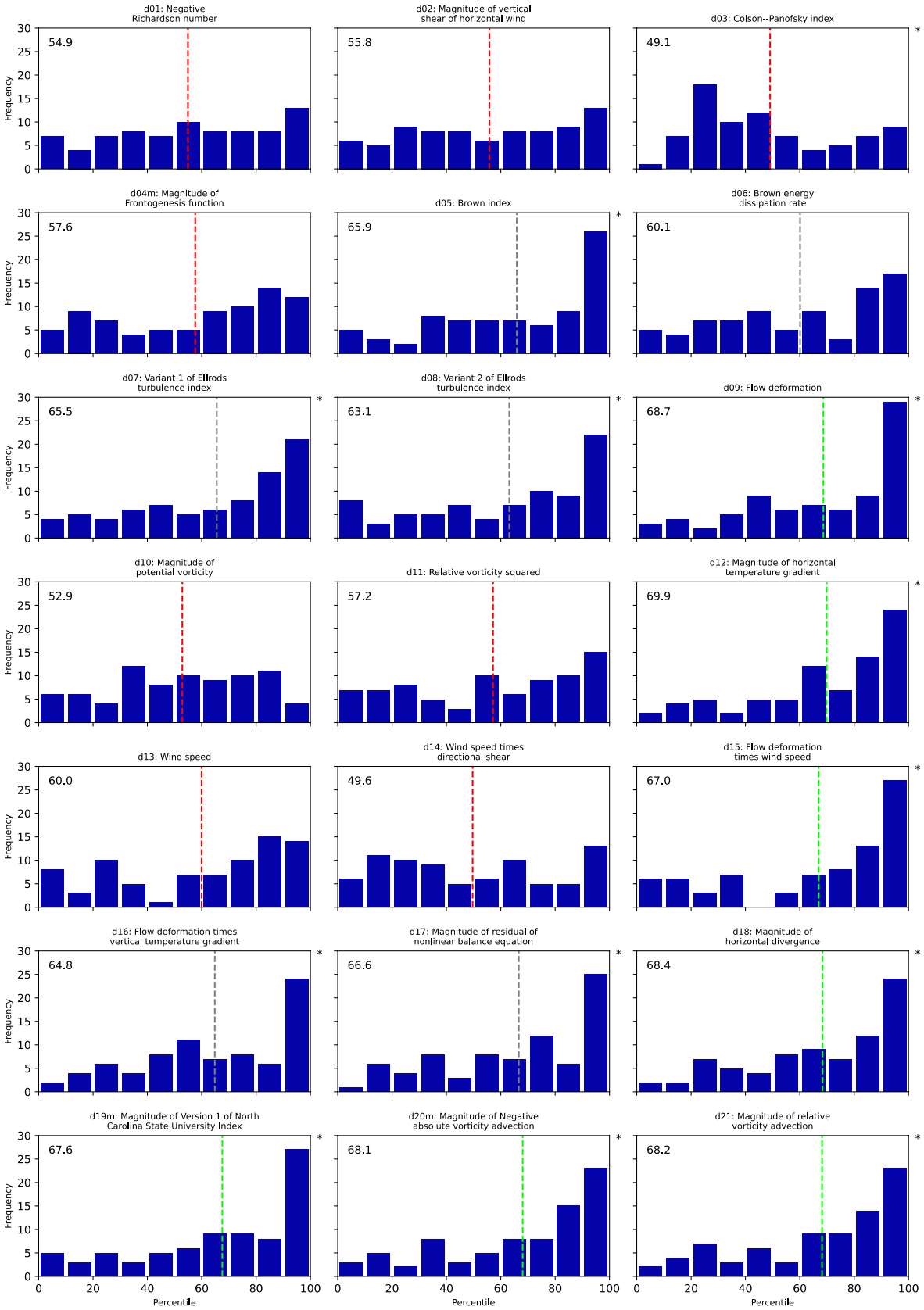


Figure 3.5: Event percentile distributions (n=80 events) for each of the 21 CAT diagnostics examined in this study. The vertical dashed lines show the 80-event mean (green if ≥66.7%, red if <60%, grey if in between these values) and its percentile value is shown in the top left of its respective subplot.

Diagnostic percentile distributions that are statistically significantly different to a uniform distribution according to a chi-squared test (p=0.05/24) are marked with an asterisk to their top right.

Diagnostics 4,19 & 20 conventionally exclude negative values, but it was found that a large fraction of grid points (in some cases up to 50%) were negative, so the decision to not exclude them was taken. This results in a U-shaped distribution, as shown in the top panels of Figure 3.6. Modified versions of these three diagnostics were created in which their absolute values were taken instead of negative values being excluded. The modified diagnostics 4, 19, and 20 are given in bottom panel of Figure 3.6.

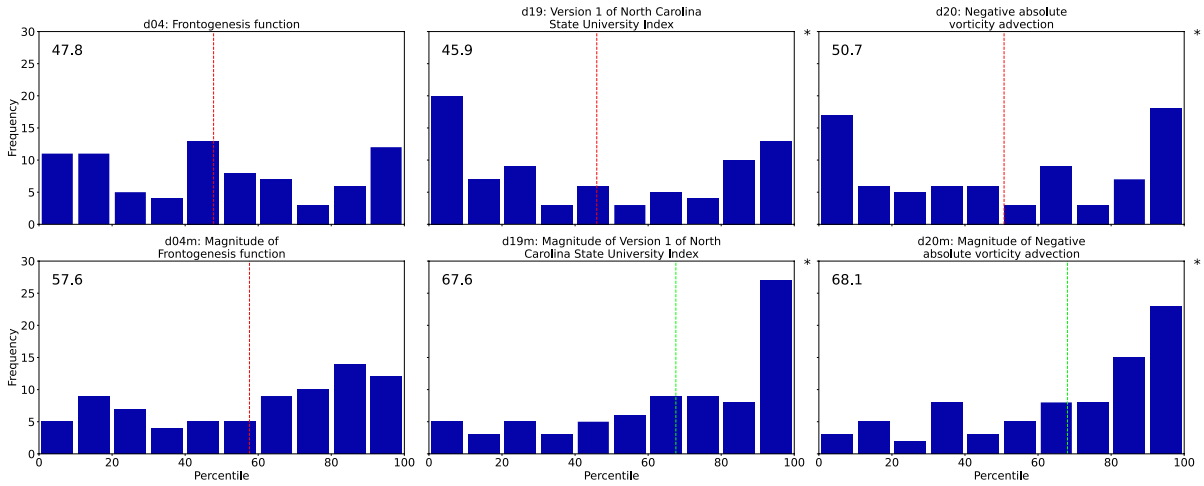


Figure 3.6: Diagnostics 4, 19, and 20 compared with modified versions in which the absolute values are taken. Diagnostic percentile distributions that are statistically significantly different from a uniform distribution according to a chi-squared test (p=0.05/24) are marked with an asterisk to their top right.

Figures 3.5 and Table 3.3 indicate that the diagnostics fall into three broad skill categories when it comes to diagnosing injury causing CAT: high skill with mean percentiles greater than 66.7%. There are seven diagnostics in this category.

There are seven in the medium skill range of 60-66.7%

Finally there are ten diagnostics in the low skill range with a percentile mean of less than 60.0%.

If a turbulence diagnostic exhibited no skill, a uniform distribution would be detected. A chi-squared test was applied to each diagnostic to find those that are significantly different from a uniform distribution. The significance level was set at p=0.05 / n, where n is the total significant tests carried out equalling 24. n is effectively the Bonferroni correction and is applied as the number of tests implemented is >20 and therefore a possibility that one of the tests returned will give a false negative if the significance level is not reduced. Diagnostics that have distributions which are statistically not uniform are indicated with an asterisk to their top right in Figures 3.5 and 3.6. A high-skill diagnostic such as d18, magnitude of horizontal divergence, exhibits a strong negative skew with a large proportion of the 80 events in the 90-100th percentile bins. This in contrast to a low-skill diagnostic such as d10, where the distribution of percentiles is reasonably uniform and not statistically distinguishable from a diagnostic with zero skill.

Table 3.3: The 80-event percentile means for each of the 21+3 diagnostics.

SIG TEST?	Diagnostic number, name and units		80-event mean
∗∨	d12:	Magnitude of horizontal temperature gradient ($K \cdot m^{-1}$)	69.4
	d18:	Magnitude of horizontal divergence (s^{-1})	69.0
	d09:	Flow deformation (s^{-1})	68.9
	d21:	Magnitude of relative vorticity advection (s^{-2})	68.4
	d20m:	Magnitude of Negative absolute vorticity advection (s^{-2})	68.3
∗∨	d19m:	Magnitude of Version 1 of North Carolina State University Index (s^{-3})	68.1
∨	d15:	Flow deformation times wind speed ($m \cdot s^{-2}$)	67.0
∗	d17:	Magnitude of residual of nonlinear balance equation (s^{-2})	66.5
	d05:	Brown index (s^{-1})	66.4
∗∨	d07:	Variant 1 of Ellrods turbulence index (s^{-2})	66.1
	d16:	Flow deformation times vertical temperature gradient ($K \cdot m^{-1} \cdot s^{-1}$)	64.9
	d08:	Variant 2 of Ellrods turbulence index (s^{-2})	63.8
	d06:	Brown energy dissipation rate ($J \cdot kg^{-1} \cdot s^{-1}$)	60.9
∨	d13:	Wind speed ($m \cdot s^{-1}$)	60.0
∗∨	d04m:	Magnitude of Frontogenesis function ($m^2 \cdot s^{-3} \cdot K^{-2}$)	58.3
	d11:	Relative vorticity squared (s^{-2})	57.4
	d02:	Magnitude of vertical shear of horizontal wind (s^{-1})	56.6
∗	d01:	Negative Richardson number (—)	55.5
	d10:	Magnitude of potential vorticity (PVU)	53.2
	d20:	Negative absolute vorticity advection (s^{-2})	50.7
	d14:	Wind speed times directional shear ($rad \cdot s^{-1}$)	50.4
∗	d03:	Colson–Panofsky index (kt^2)	49.0
∗∨	d04:	Frontogenesis function ($m^2 \cdot s^{-3} \cdot K^{-2}$)	47.5
∗∨	d19:	Version 1 of North Carolina State University Index (s^{-3})	46.3

Diagnostic d12, the Magnitude of horizontal temperature gradient, performed the best. This is due to jet streams, which are driven by latitudinal temperature gradients via the thermal wind balance being the main sources of CAT (Reiter, 1963a; Reiter, 1963b). It therefore stands to reason that a large horizontal temperature gradient, especially in the latitudinal direction, might be associated with a stronger jet stream and, consequently, stronger CAT. Flow deformation, DEF is given by equation 3.3.

$$\mathrm{DEF} = \sqrt{\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right)^2} \tag{3.3}$$

where *u* and *v* are the zonal and meridional wind components respectively and *x* and *y* are distances in the longitudinal and latitudinal directions. Diagnostics d5, d6, d7, d8, d9, d15 & d16 containing deformation (*DEF)* sit within the medium or high skill categories and all except d16, *DEF* × vertical temperature gradient is significantly different from a zero-skill diagnostic. This agrees with the findings of (Ellrod and Knapp, 1992) that flow deformation is an important source of CAT. Both diagnostics containing the horizontal divergence (d18 and d17) also appear in the top eight most skilled diagnostics.

Diagnostic d01, the negative Richardson number, which is at the core of theoretical basis for being a key indicator of the CAT potential in a given region of atmosphere performs poorly with a mean percentile score of 55.5%. The poor performance of the Richardson number was noted by (Kaplan et al., 2005) and is perhaps due inadequate vertical resolution. Two of the three diagnostics explicitly containing the

Richardson number, d03: Colson-Panofsky index, and d19: Version 1 of North Carolina State University Index, performed very poorly, with 80-event means of 49.0% and 46.3% respectively.

Seven diagnostics used on the upper levels in GTG2 (Sharman et al., 2006), were studied here. However, only two of them d12, magnitude of horizontal temperature gradient, and d15, $DEF \times U$ were among the high skilled (>66.7%) diagnostics. Also, three diagnostics that feature in GTG2 had 80-event percentile means within 5% of 50%. These were, d03: Colson-Panofsky index, d19: Version 1 of North Carolina State University Index; d04: frontogenesis function. There are similar skill results for the GTG2 mid level diagnostics analysed in this study. However, only 20-25% of the 80 events examined in this study occurred at mid-levels (FL100-FL200).

When assessing GTG2 diagnostic performance, consideration should be given that US flights use GTG2 and will have planned a flight route to avoid regions of forecast turbulence. Since 75% of the events used here were likely to have been US commercial aircraft, this may introduce bias. However, the highest scoring diagnostic (d12, magnitude of horizontal temperature gradient) is used in GTG2, and the other GTG2 diagnostics are fairly evenly distributed in terms of mean percentile (See Table 3.3) meaning a bias is unlikely significant.

Kim et al. (2011) noted differential performance in GTG diagnostic skill dependent on the season. The 80 events used here are evenly distributed over the year (DJF=23 events, MAM=22 events, JJA=18 events, and SON=17 events). However there are not enough events in any one season alone to get reliable results at this time.

This study's methodology implicitly makes several assumptions. First, we assume that ERA5 is an accurate representation of the atmosphere at typical aviation altitudes. Figure 16 & 17 of Hersbach et al. (2020) show that ERA5's temperature and wind bias are at the 200~hPa bias is 0.1~K and ± 0.1 ms^{-1} respectively meaning ERA5 is a very good representation of the atmosphere. Simmons et al. (2020) showed ERA5 has a cold bias between 2000-2006 in the stratosphere. Although the strongest cold bias is in the upper stratosphere there was still an additional 0.1~K bias at 200~hPa. To minimise effects to the study the ERA5.1 rerun for the 11 events that fell within this time period was used.

Second, it is assumed that all events are CAT and not other forms of turbulence. As discussed in section 3.3 only select events with CAT mentioned were chosen minimising the chance of incorrect identification. In effect this probably meant that many legitimate CAT events were screened out due to being referred to as generic turbulence. This conservative approach gives confidence that the vast majority of the 80 events are indeed CAT.

It was also assumed that the reported temporal and spatial coordinates of the events are accurate. This assumption was shown to be largely correct when nine duplicate events garnered from AvHerald with similar coordinates to those reported by the NTSB were examined. In most cases, the reported coordinates

were very close (e.g., within an hour, a single ERA5 pressure level, or degree of latitude and longitude). However, one or two larger discrepancies were also noted (e.g., 6 hours or 4000 m of altitude difference). Despite this, the similarities in the coordinates between the aforementioned crossover events was close enough to ensure a high degree of confidence in their accuracy.

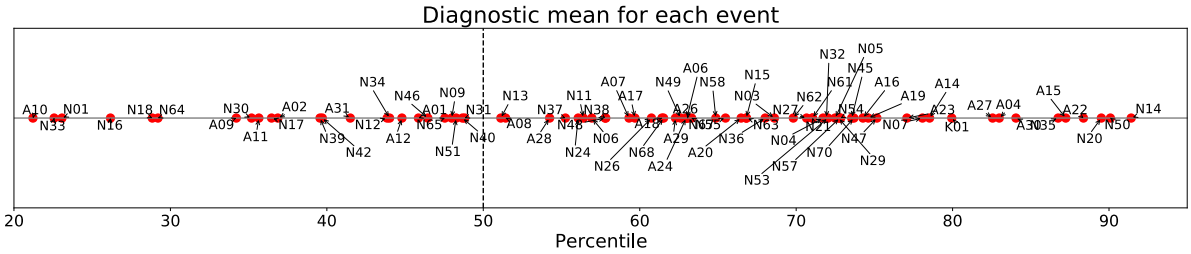


Figure 3.7: The 21-diagnostic percentile means for each of the 80 CAT events. Events with A, N, and K as the first letter are events garnered from AvHerald.com, NTSB, and Lee and Chun (2018), respectively.

Figure 3.7 shows the distribution of average diagnostic percentile mean for each of the 80 events which range from 21.6%-91.4%, mean of the 21-diagnostic means is 61.5%. Events with extremely low 21-diagnostic percentile means could be due to one or more sources of error combining meaning there could be a geolocation error in regards to the position of the aircraft and that of the predicted turbulence.

In addition to the assumptions stated above, a further assumption is that the 2016 Atlantic climatology subset is representative in both time and space. The geographical position over the ocean was chosen to ensure the absence of MWT. However, many of the events are over land, and Dutton and Panofsky (1970) suggest that CAT may be more prevalent here than over the ocean. Whilst the CAT diagnostics are specifically designed to diagnose CAT, they may be affected by orography within the IFS, the model behind ERA5. This may artificially shift the climatological distributions used to calculate the percentiles for each event higher or lower introducing uncertainty.

3.5 Revising CAT projections

One of the key sources of uncertainty in future CAT projections is the spread between predictions made using different diagnostics. In Storer et al. (2017), a uniformly weighted mean is taken across the diagnostics to provide an estimate of where changes in CAT are most likely to occur. However, if some diagnostics are more skilled than others, then a uniform weighting perhaps isn't the most appropriate choice when taking the mean across the diagnostics. In this section, two different weighting schemes that take into account the different diagnostic skills derived in section 3.4 are applied to (1) the North Atlantic CAT projection in Williams (2017) and (2) the global CAT projections of Storer et al. (2017) to further refine these projections.

3.5.1 Revisiting North Atlantic CAT projections

Williams (2017) calculated the median increases in CAT over the North Atlantic in a doubled-CO2 simulation relative to a pre-industrial control simulation by applying the same 21 CAT diagnostics used

in this study. The 21-diagnostic-mean increase recalculated from Williams (2017) original data between the two simulations in all five CAT severity thresholds is shown in the top row (Uniform) of Table 3.4. Implicit in Williams' (2017) diagnostic-mean increases is the assumption that all 21 CAT diagnostics are equally skilled. However, Figures 3.5 and 3.6 suggest that some diagnostics are better than others when it comes to diagnosing injury-causing CAT, hence a uniform weighting can now be improved upon using the results of this paper.

Table 3.4: Projected increases in CAT in five different strength categories for three different weighting methods (plus a mean of the linear and non-linear methods) for a doubled-CO_2 experiment versus a pre-industrial control simulation. Following Williams (2017).

Weighting Method	Light	Light–Moderate	Moderate	Moderate–Severe	Severe
Uniform (U)	+58.7%	+72.0%	+83.9%	+106.2%	+141.1%
Linear (L)	+57.9%	+70.8%	+83.0%	+105.0%	+138.9%
Non-linear (NL)	+54.7%	+66.5%	+80.4%	+101.2%	+129.0.7%
Mean(L, NL)	+56.3%	+68.7%	+81.7%	+103.1%	+134.4%
Mean(L, NL) - (U)	−2.4%	−3.3%	−2.3%	−3.1%	−6.7%

In Table 3.4, in addition to Williams' (2017) full uniform weighting approach (Uniform), two other weighting methods are used to recalculate the 21-diagnostic-mean increases in CAT over the North Atlantic for the five severity thresholds in Williams (2017). In the linear weighting method (L), diagnostics are weighted according to their 80-event means (see X in equation 3.4; Table 3.3). In the non-linear method (NL), a different approach is taken. Hypothetically, had one of the diagnostics achieved a perfect skill (an 80-event mean of 100%), the temptation would be to give it an infinite weight and disregard all the other diagnostics. Conversely diagnostics with zero skill (80-event means of 50%) should receive zero weighting. Therefore a weighting regime of X/(1-X) was applied to each diagnostic, where:

$$X = \left|\frac{2*(eventpercentilemean-50)}{100}\right| \qquad (3.4)$$

Table 3.4 shows that the mean increases in CAT prevalence per threshold for a doubled-CO2 versus a pre-industrial control simulation are typically slightly less when a non-uniform weighting scheme based on the skill of the 21 diagnostics is applied. The projected change in CAT varies depending on the weighting

method used, but the average across the two methods indicates a future projected CAT increase of 2.3-6.7% less than that projected by Williams (2017) using a uniform weighting. The threshold for which the estimate was revised downwards most 6.7% was the severe CAT threshold.

3.5.2 Revisiting Global CAT projections

Storer et al. (2017), showed a global increase in DJF Moderate or Greater (MOG) CAT at 200 hPa by the period 2050-2080 calculated from the RCP8.5 climate change scenario relative to a preindustrial control run of a CMIP5 climate model (HadGEM2-ES). Figure 3.8(A) is generated by taking an equally weighted mean across 20 of the 21 diagnostics used in this study. The magnitude of potential vorticity was excluded because it was found to produce unphysical results. Storer et al. (2017) found large increases in MOG CAT of the order of several hundred percent in the mid-latitudes, offset by a small but statistically insignificant decrease in the tropics.

The results from Storer et al. (2017) shown in Figure 3.8(A) are recalculated using the linear (Fig. 3.8 panel B) and non-linear (Fig. 3.8, panel C) weighting methods described in the previous subsection. Difference plots of linear minus uniform and non-linear minus uniform are given in Fig. 3.9.

Not a lot appears to change between the linear and uniform weighting schemes.

A large change is evident between the non-linear and uniform weighting schemes, with an extra increase in projected DJF MOG CAT of around 100-150% over the Atlantic corridor and eastern North America, eastern Pacific, Europe and along the Southern hemisphere jet stream region. In addition to this increase, around 100% less DJF MOG CAT is evident across a much smaller patch just south of the North Atlantic Corridor.

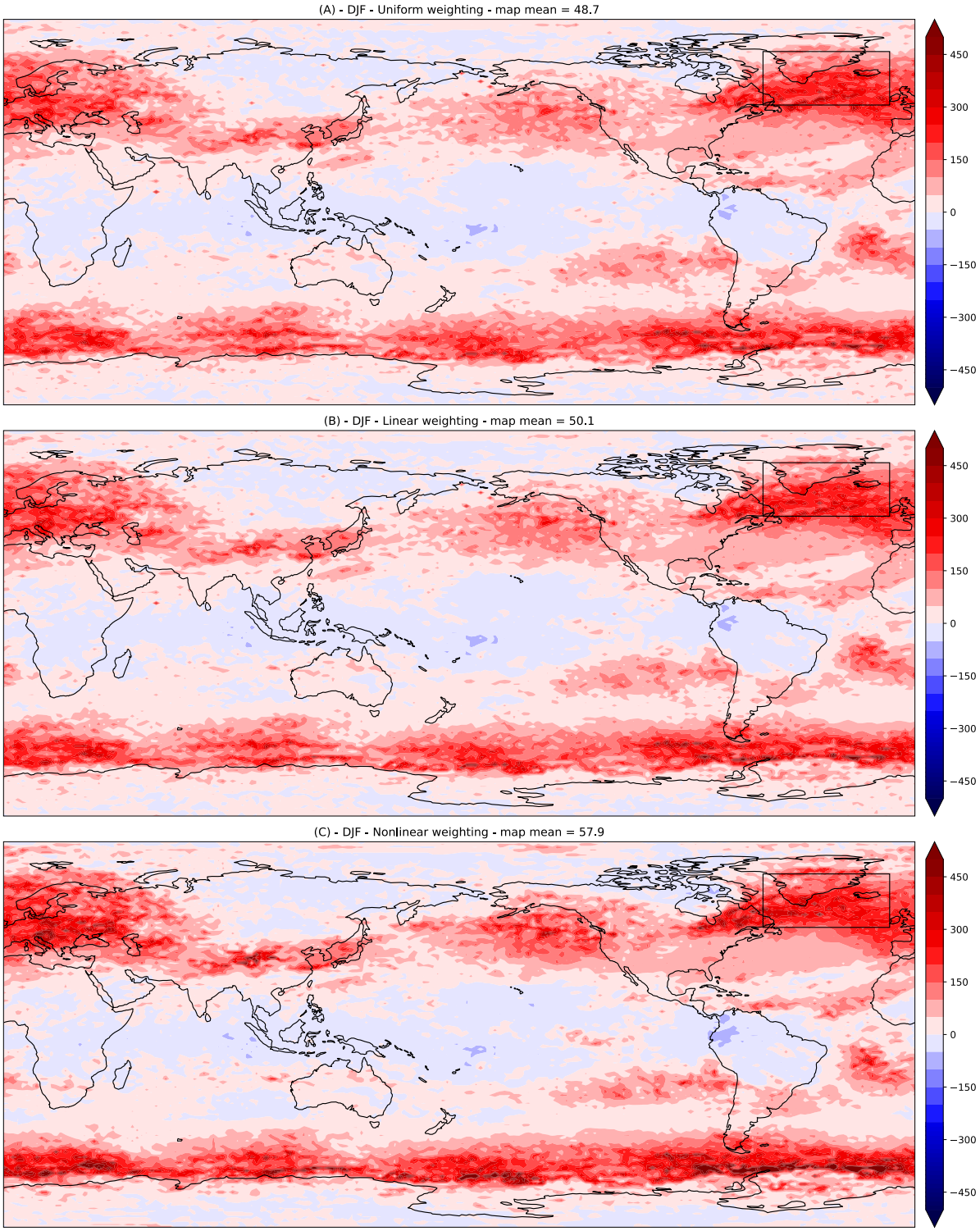


Figure 3.8: DJF 20-diagnostic average of the percentage change between a HadGEM2-ES RCP8.5 2050-2080 climate change simulation and a pre-industrial control at 200 hPa. The 20 diagnostics are weighted using (A) the uniform weighting method as in Storer et al. (2017), (B) the linear weighting method and (C) the non-linear weighting method. The rectangle is the geographical area considered in Williams (2017).

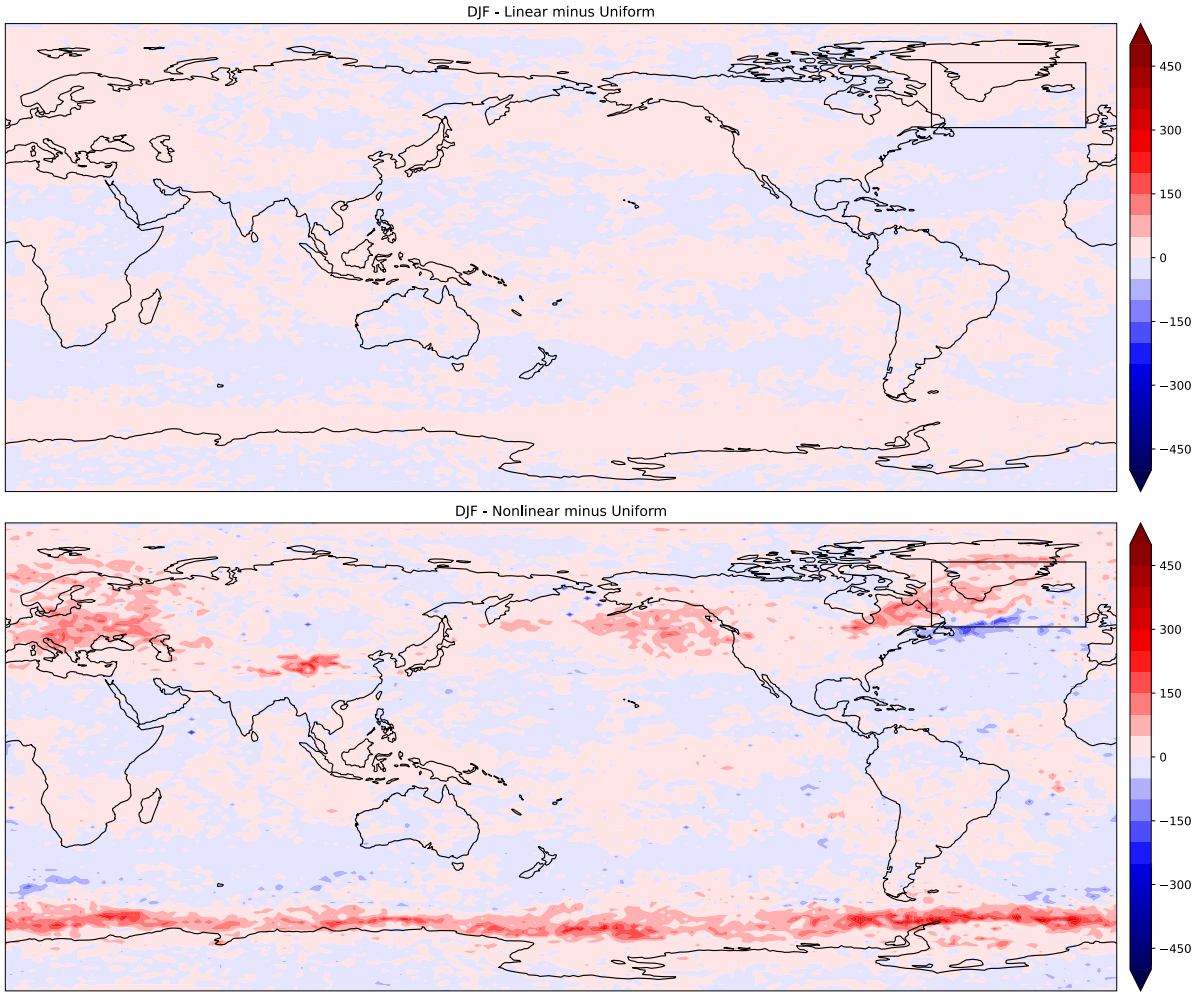


Figure 3.9: Difference plots between the (A) linear weighting and (B) non-linear method and the uniform weighting methods shown in Fig. 3.8.

In summary, the new weighting schemes seem to imply a smaller projected increase in CAT in the case of Williams (2017), but a larger projected increase in the case of Storer et al. (2017). Although the latter is a global study and the former a regional study, this result remains if the work of Storer et al. (2017) is restricted to just the northern North Atlantic. Both studies agree that CAT will increase with climate change and this conclusion remains true after applying our weighting. The reason for this result is clearly related to diagnostic 9 (Flow deformation), diagnostic 12 (Magnitude of horizontal temperature gradient), diagnostic 15 (Flow deformation times wind speed), diagnostic 18 (Magnitude of horizontal divergence) and diagnostic 21 (Magnitude of relative vorticity advection) which together make up about half of the weight. It is also worth remembering that an RCP8.5 climate change scenario run until the period 2050-2080 using the CMIP5 HadGEM2-ES climate model, as used in Storer et al. (2017), is a different climate change scenario than the doubling of CO2 used in Williams (2017) and run on a CMIP3 GFDL-CM2.1 climate model.

3.6 Conclusions

In this study, the skill of 24 CAT diagnostics (21 unmodified and 3 modified) at diagnosing turbulence was evaluated using 80 injury-causing CAT events encountered by commercial aircraft. A large variation in skill was observed, which can be roughly divided into three classes: high skill (80-event percentile means of ≥66.7%; 7 diagnostics), middle skill (80-event percentile means of 60-66.7%; 7 diagnostics), and low skill (80-event percentile means of <60%; 10 diagnostics).

The three highest scoring diagnostics were, in descending order: the magnitude of horizontal temperature gradient (80-event percentile means of 69.4%), magnitude of horizontal divergence (69.0%) and flow deformation (68.9%), while ten diagnostics had 80-event percentile means under 60.0% or had percentile distributions that were statistically indistinguishable from zero skill (or both). It was found that all seven diagnostics containing the flow deformation term in their equations occurred in either the medium or high skill categories (80-event means over 60.0%) and all but one of them had percentile distributions which were significantly different from a zero-skill diagnostic according to a chi-squared test. It was also found that diagnostics containing the Richardson number performed less well than might be expected from CAT theory and previous work.

One major finding of this research was that modified versions of the negative absolute vorticity advection, North Carolina State University Index, and frontogenesis function diagnostics were found to outperform the original versions, significantly in the case of the first two (68.3% versus 50.7%, 68.1% versus 46.3%, and 58.3% versus 47.5%, respectively). These modified diagnostics, as well as the high/middle skill set of diagnostics observed here, warrant further testing in CAT forecasting systems such as GTG to see if they generate improvements in injury-causing forecasting skill.

The fact that not all diagnostics are equally skilled when it comes to injury-producing turbulence is novel in the context of climate projections of CAT, because previous works e.g. Storer et al. (2017) and Williams (2017) implicitly assumed equal skill through their choice of a uniform weighted diagnostic mean. Given that injury-causing turbulence provides most of the damage to passengers, crew, and airframes of all turbulence encounters, knowledge of its diagnostic skill weighting will be useful for CAT forecasters going forward.

Using this evidence of differential diagnostic skill, the mean increases in injury-causing CAT across two different sets of climate projects were refined. While the refinement of CAT projections in the North Atlantic corridor of Williams (2017) suggested a slightly smaller increase in injury causing CAT with climate change across all five thresholds, the refinement of the projections of Storer et al. (2017) indicated a larger increase in moderate or greater CAT, both in the North Atlantic corridor region and globally. These results echo previous work demonstrating a challenging future in relation to CAT for the aviation industry. Not only is a large increase in CAT likely, but unfortunately, the areas with the largest increases are likely to coincide with the busiest air spaces over the North Atlantic corridor and Europe.

There are of course considerable uncertainties with such CAT climate projections. Humanity's future response to climate change and the degree to which trends in globalisation continue or not, would be two such examples. However, the potential for large changes in CAT along with the corresponding impacts on aviation safety suggests at the very least that contingency planning is advisable.

This study has specifically looked at CAT and future work might extend the same methodology to other types of turbulence, such as CIT and MWT, in order to give industry decision-makers a more holistic picture of future turbulence, on which to base future planning. Commercial aircraft will continue to encounter severe CAT going into the future and while unfortunate, the analysis done here could be extended to cover these events too.

Chapter 4

Evidence for Large Increases in Clear-Air Turbulence over the Past Four Decades

4.1 Introduction

Turbulence is estimated to cost the aviation industry around US$200 million annually in the USA alone (Eichenbaum, 2003). These costs arise partly from additional airframe fatigue, requiring maintenance and subsequent loss of productivity, as well as occasional airframe damage. Additionally, passengers and crew suffer injuries, some requiring costly hospital treatment. Some aircraft turbulence occurs in well-defined locations, such as over mountain ranges or within the vicinity of convective storms, and is largely avoidable. However, clear-air turbulence (CAT) is difficult to observe in advance of an aircraft's track using remote sensing methods. Furthermore, it is still challenging for aviation meteorologists to forecast CAT, partly because current Numerical Weather Prediction (NWP) models have grid sizes that are many times larger than the turbulent eddies that affect aircraft. Hence, operational forecasters use empirical turbulence diagnostics (e.g. Ellrod and Knapp, 1992; Brown, 1973; Dutton, 1980; Knox, 1997; Knox et al., 2008; McCann et al., 2012) calculated from the temperature and wind fields of NWP output. In recent years, these diagnostics have been combined into multi-diagnostic forecasts (Sharman et al., 2006).

Williams and Joshi (2013) applied 21 CAT diagnostics to the 200 hPa pressure surface (corresponding to a flight level of approximately 39,000 ft) of a climate model using a doubled-CO2 scenario. They found that the future occurrence of moderate-or-greater (MOG) CAT increased substantially during winter in the North Atlantic. MOG CAT is defined by the Federal Aviation Administration (FAA, 2006) as the point at which unsecured objects begin to move, and at which people find it difficult to move around inside the cabin. Williams (2017) expanded the analysis to examine turbulence increases at different severity levels (light, moderate, and severe) and found an increase in the frequency of diagnosed CAT for nearly all threshold–diagnostic pairs. Storer et al. (2017) analysed a CMIP5 simulation using the RCP8.5 scenario in 2050–2080 and compared it with a pre-industrial control simulation, covering the whole globe and each

season at different flight levels. Within the jet stream regions of both hemispheres, the RCP8.5 scenario relative to the pre-industrial control showed around a 300% increase in CAT.

Lee et al. (2019) examined three reanalysis datasets over 1979–2017 and found evidence of a 15% increase in vertical wind shear strength at 250 hPa over the North Atlantic (30–70°N, 10–80°W). As is well known, when the ratio of thermal stability to vertical wind shear (Richardson number Ri) is less than some critical value, typically 1/4, Kelvin–Helmholtz instabilities can form and ultimately result in CAT. Therefore, stronger vertical wind shear is expected to increase the amount of turbulence. However, studies examining whether the amount of CAT measured by aircraft has increased due to recent warming trends are limited by the availability of suitable data. The record of automated, quantitative eddy dissipation rate (EDR) turbulence measurements is too short. Pilot reports (PIREPs) have a longer record, but are not quantitative, and the geographical distribution of CAT based on PIREPS is limited in spatial and temporal extent (Wolff & Sharman, 2008). Furthermore, long-term improvements in operational CAT forecast skill should be acting to reduce the probability of encountering turbulence, even if the amount of turbulence in the atmosphere is increasing.

Outside aviation, CAT is also of interest as a mechanism allowing the mixing of air between the troposphere and stratosphere. Jaeger and Sprenger (2007; hereafter JS07) used ERA40 reanalysis data (Uppala et al., 2005) to compute a winter and summer northern hemisphere CAT climatology (1958–2001) near the dynamic tropopause, using four CAT diagnostics: Richardson number, negative squared Brunt–Väisälä frequency, negative potential vorticity (Gidel and Shapiro, 1979), and Ellrod's turbulence index (Ellrod and Knapp, 1992). An increase in all four of these indices over the north Atlantic, European, and US regions was found. However, for aviation purposes, aircraft fly along constant flight levels as opposed to the dynamic tropopause. Furthermore, the first 22 years of the climatology of JS07's uses reanalysis data from before the start of meteorological satellite era in 1979, leading to data quality concerns. Simmons et al. (2020) and Marlton et al. (2021) showed that the commissioning and decommissioning of meteorological satellites can introduce biases. Lee et al. (2023) have recently updated the work of JS07 (see Section 4.4 for a discussion).

Reanalysis packages now contain four decades of data entirely in the satellite era, during which the world has continued to warm. Therefore, the present study aims to analyse CAT trends during 1979–2020 in the ERA5 reanalysis dataset, which has more advanced model physics and higher vertical and horizontal resolution than ERA40. The 21 diagnostics used in Williams and Joshi (2013) and Williams (2017) will be computed, as opposed to the four in JS07, to yield an improved quantification of inter-diagnostic uncertainties. These 21 diagnostics have previously been validated using aircraft measurements of CAT and have generally been found to be skilful (e.g. Sharman et al., 2006). To make the results more applicable to global aviation, the diagnostics will be calculated on the 197 hPa pressure level, corresponding approximately to a flight level of 39,000 feet (FL390), rather than the tropopause, and for the entire globe as opposed to just the northern hemisphere considered by JS07 and Lee et al. (2023).

The layout of this paper is as follows. Section 4.2 describes a methodology to compute the number of times a year a given turbulence severity occurs. Section 4.3 presents the results. Section 4.4 concludes with a discussion and summary.

4.2 Methodology

Global ERA5 reanalysis data (Hersbach et al., 2020) containing zonal and meridional wind speed, dry bulb temperature, and geopotential height were extracted on the 197 hPa pressure level with 0.25° horizontal resolution at three hourly intervals from 1 January 1979 to 31 December 2020. To allow the computation of CAT diagnostics that require vertical derivatives, fields on the 188 and 206 hPa levels were also extracted. The 21 turbulence diagnostics were then calculated from the extracted reanalysis fields every three hours.

To allow an inter-diagnostic comparison, the uncalibrated CAT diagnostic values, each with different physical units, are compared with threshold values derived from a climatological probability distribution for each diagnostic, following Williams (2017). The reanalysis data were extracted on a fixed Gaussian grid, meaning more data points for a given surface area near the pole than the equator, and so the climatological probability distributions were latitudinally weighted. The latitudinally weighted distributions were calculated for the year 2000, a reference year chosen as being the 1979–2020 midpoint.

Following Williams (2017), the diagnostic values corresponding to the 97th, 99.1st, 99.6th, 99.8th, and 99.9th percentiles were then derived globally for the reference year 2000, corresponding, respectively, to the thresholds for light-or-greater (LOG), light-to-moderate-or-greater (LMOG), moderate-or-greater (MOG), moderate-to-severe-or-greater (MSOG), and severe-or-greater (SOG) turbulence. For each diagnostic and threshold, the number of exceedances at a given coordinate for each month, season, and year in the study period were computed. For each year, an average exceedance field was calculated by taking the mean of the 21 exceedance fields for each diagnostic. The diagnostic average exceedance was chosen over the unequal weightings given in chapter 3 because the latter apply specifically to injury causing turbulence, a specific rather than a more general category of aviation affecting turbulence.

To calculate temporal trends, linear regression was applied at each grid point using the annual-mean exceedance values for the 1979–2020 period for the five turbulence severities. Using the regression model at each grid point, fitted 1979 and 2020 exceedances were computed as a guide to the underlying turbulence statistics in the absence of interannual variability at the start and end of the period. In the rare cases that these fitted exceedances were negative, they were set to zero. Exceedances were converted into percentage probabilities of exceedance, by normalising by the number of three-hour periods in each year.

Absolute changes in the probability of encountering turbulence over the 42 years were computed by subtracting the fitted 1979 values from the fitted 2020 values. For example, a 1% increase over one year would be equivalent to around 29 extra exceedances, given the 2,920 three-hour periods in a single year. To calculate the relative changes, the absolute changes were divided by the fitted 1979 values and multiplied by 100, to yield a percentage relative change

4.3 Results

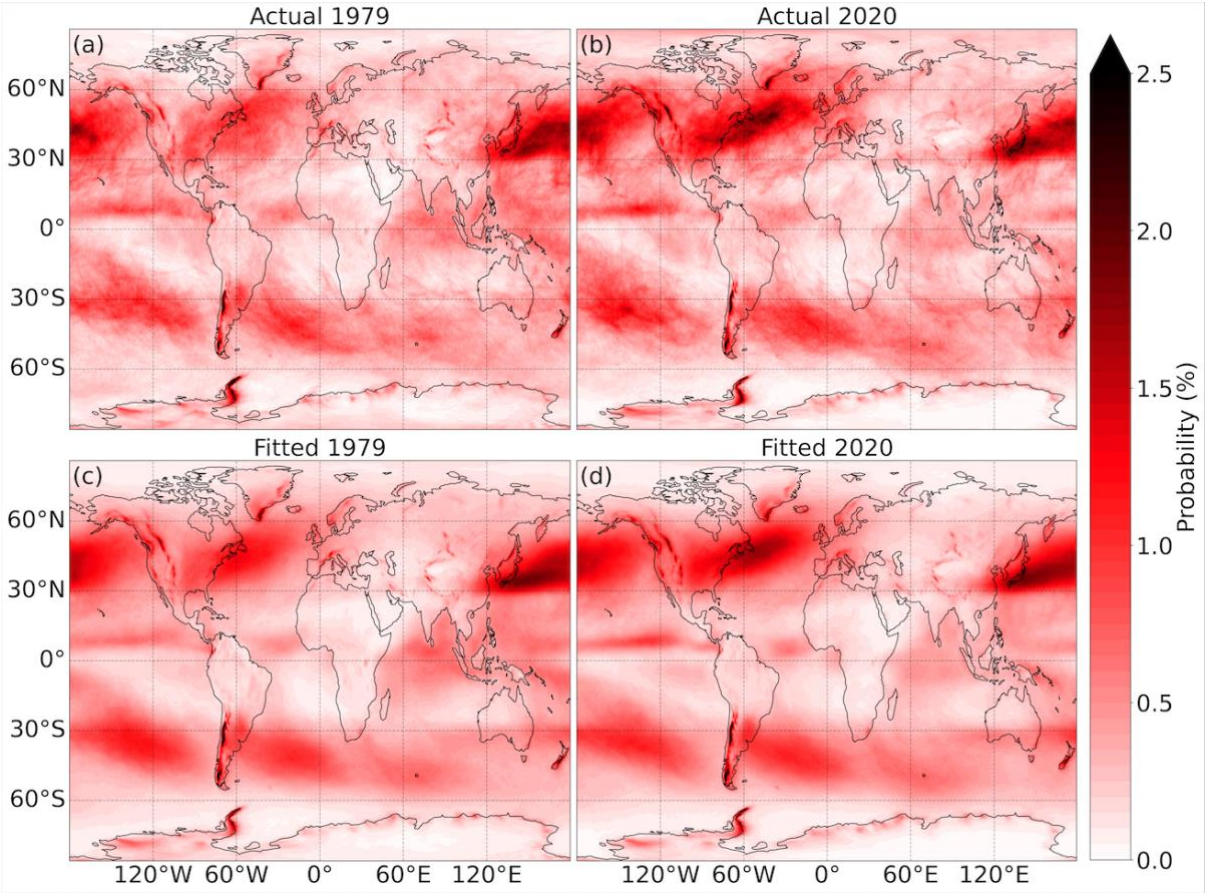


Figure 4.1: Annual-mean probabilities of encountering MOG CAT in (a) the year 1979, (b) the year 2020, (c) the year 1979 inferred from the linear regression model, and (d) the year 2020 inferred from the linear regression model. The probabilities are calculated from ERA5 at 197 hPa and are averaged over 21 CAT diagnostics.

Figure 4.1 shows global maps of the annual-mean diagnostic-mean MOG CAT probability in 1979 and 2020. The spatial structures of the probabilities inferred from the linear regression model are smoother than the raw annual fields, because interannual variability has been filtered out. The probability of diagnosed MOG CAT is generally larger over the oceans than the continents and is larger in the midlatitudes where

the atmospheric jet streams are located. A band of diagnosed MOG CAT is also evident along the equatorial oceans, as discussed in Williams and Storer (2022). The probability of MOG CAT in the Northern Hemisphere midlatitudes (30–60°N) is roughly double the corresponding probability in the Southern Hemisphere (30–60°S). Mountainous regions such as the Andes and Rockies have higher probabilities of MOG CAT, possibly due to mountain wave breaking. The same figure but for LOG and SOG CAT intensities are given in Figures 4.2 and 4.3 respectively. The spatial pattern is similar, save for a smaller/bigger amplitude for SOG/LOG and SOG is noisier than MOG/LOG being based on fewer data points.

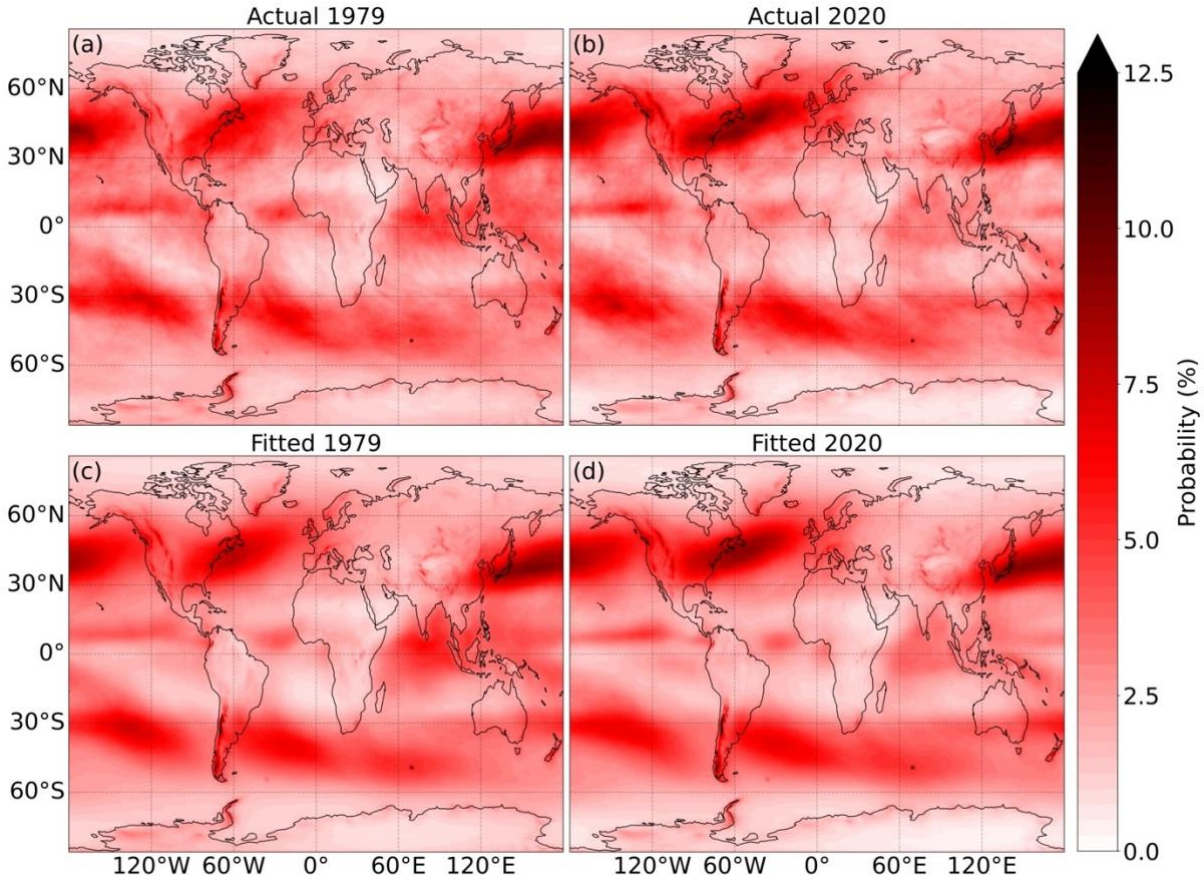


Figure 4.2. The same as for Figure 4.1, except for Light-Or-Greater Clear-air turbulence (LOG CAT) instead of Moderate-Or-Greater (MOG CAT)

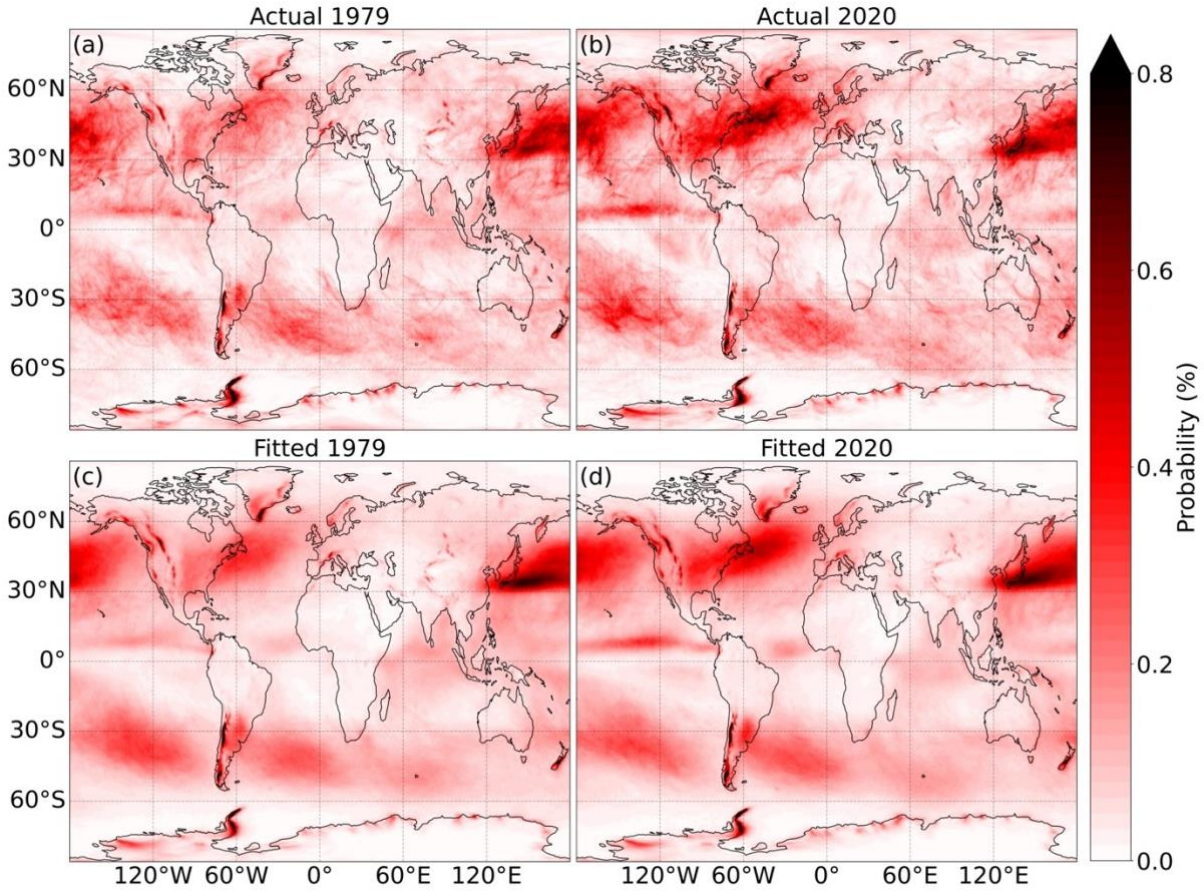


Figure 4.3: The same as for Figure 4.1, except for Severe-Or-Greater Clear-air turbulence (SOG CAT) instead of Moderate-Or-Greater (MOG CAT).

Figure 4.4: The change in ERA5's 197 hPa annual-mean diagnostic-mean MOG CAT probability over 1979–2020, showing (a) the absolute change and (b) the relative change. The changes are diagnosed from the linear trend. Stippling indicates statistical significance at the p=0.05 level, according to a two-sided Wald test (Fahrmeir et al., 2022) applied to the absolute change. The two boxes represent the North Atlantic (36–60°N and 55–10°W) and USA (30–55°N and 124–60°W) areas used in Figures 4.3 and 4.4 and Table 4.1.

Figure 4.4(a) shows the absolute change in the probability of MOG CAT from 1979 to 2020. In the North Atlantic sector, the diagnosed MOG CAT probability has increased by 0.3% in absolute terms, implying an extra 26 hours per year of diagnosed MOG CAT over the entire 42-year period, equivalent to an annual increment of around 40 minutes. A smaller increase is evident over the northern Pacific, but it is less pronounced. Figure 4.2(b) shows the corresponding relative change in the probability of MOG CAT from 1979 to 2020. Regions over the USA and the North Atlantic exhibit relative increases of up to 100%, indicating that both the absolute and relative changes over the period have been large. Other areas such as northern Brazil and parts of the coast of Antarctica also show large relative increases, despite the absolute

increases being more modest. In addition to the large changes seen over the USA and the North Atlantic, there are also statistically significant changes over Europe and the Middle East, as well as the South Atlantic and Eastern Pacific. The same figure but for LOG and SOG CAT intensities are given in Figures 4.6 and 4.7 respectively and a breakdown of the MOG figure 4.4. is given in figure 4.5.

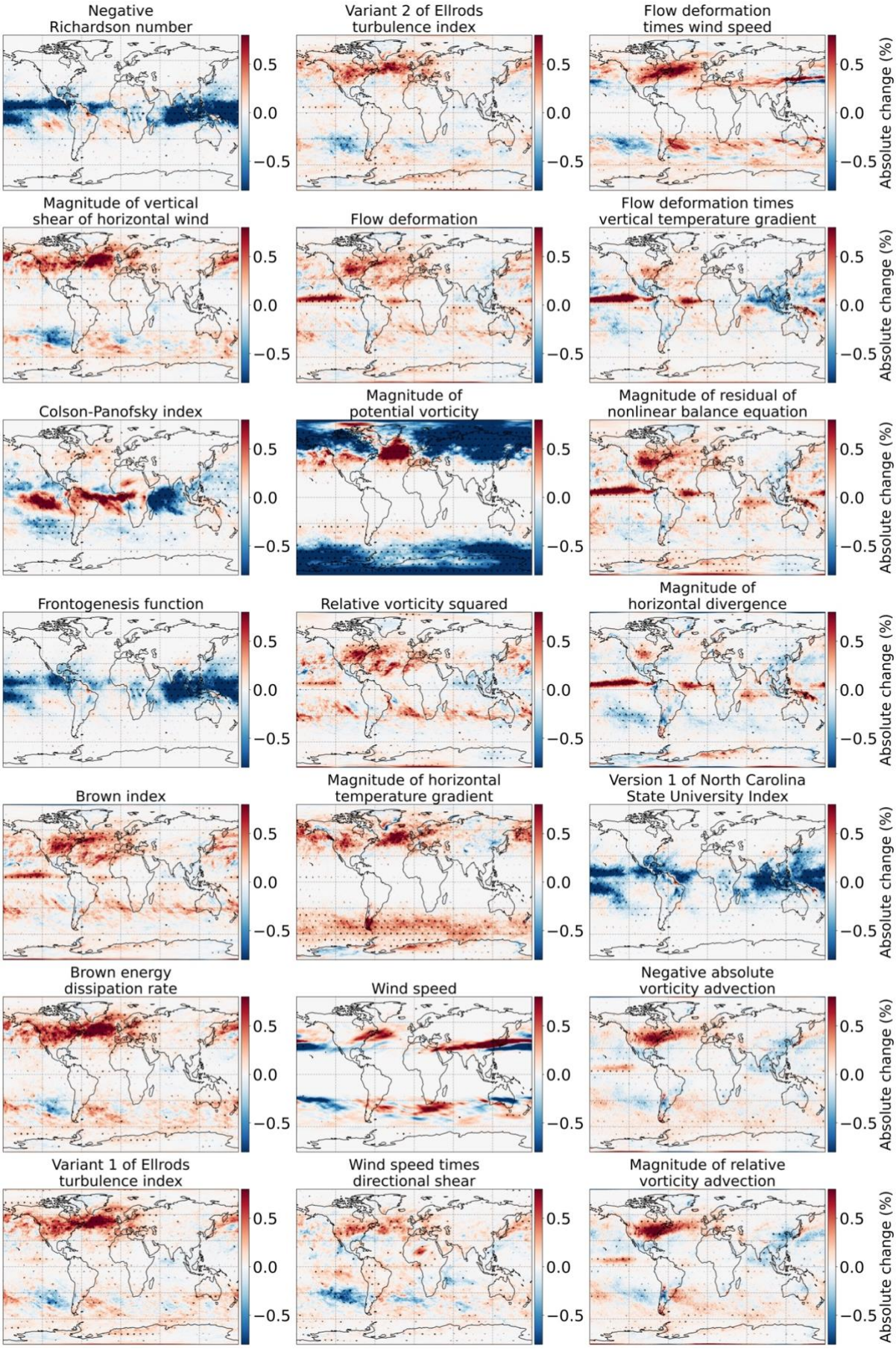


Figure 4.5: The MOG CAT absolute change (Panel a in figure 4.4), broken down by its 21 constituent CAT diagnostics.

Figure 4.5 shows the 21 diagnostic breakdown (all equally weighted) of the change in MOG CAT (Figure 4.4.). At least 9 of the diagnostics (CAT-d02: Magnitude of vertical shear of horizontal wind, CAT-d05: Brown index, CAT-d06: Brown energy dissipation rate, CAT-d07: Variant 1 of Ellrod's turbulence index, CAT-d08: Variant 2 of Ellrod's turbulence index, CAT-d12: Magnitude of horizontal temperature gradient, CAT-d15: Flow deformation times wind speed, CAT-d20: Negative absolute vorticity advection and CAT-d21: Magnitude of relative vorticity advection) contribute strongly to the large increase observed over USA and the North Atlantic and a similar set contribute to the smaller increase over the north Pacific. The overall increase over the equatorial eastern Pacific and Atlantic appear to be a balance between 5 diagnostics (CAT-d03: Colson--Panofsky index, CAT-d09: Flow deformation, CAT-d16: Flow deformation times vertical temperature gradient, CAT-d17: Magnitude of residual of nonlinear balance equation, CAT-d18: Magnitude of horizontal divergence) driving it higher and 3 diagnostics (CAT-d01: Negative Richardson number, CAT-d04: Frontogenesis function and CAT-d19: Version 1 of North Carolina State University Index) driving it lower. 2 of these 3 (d01 and d19) contain the Richardson Number within their formula, however, d03 (also containing the Richardson Number) suggests an increase.

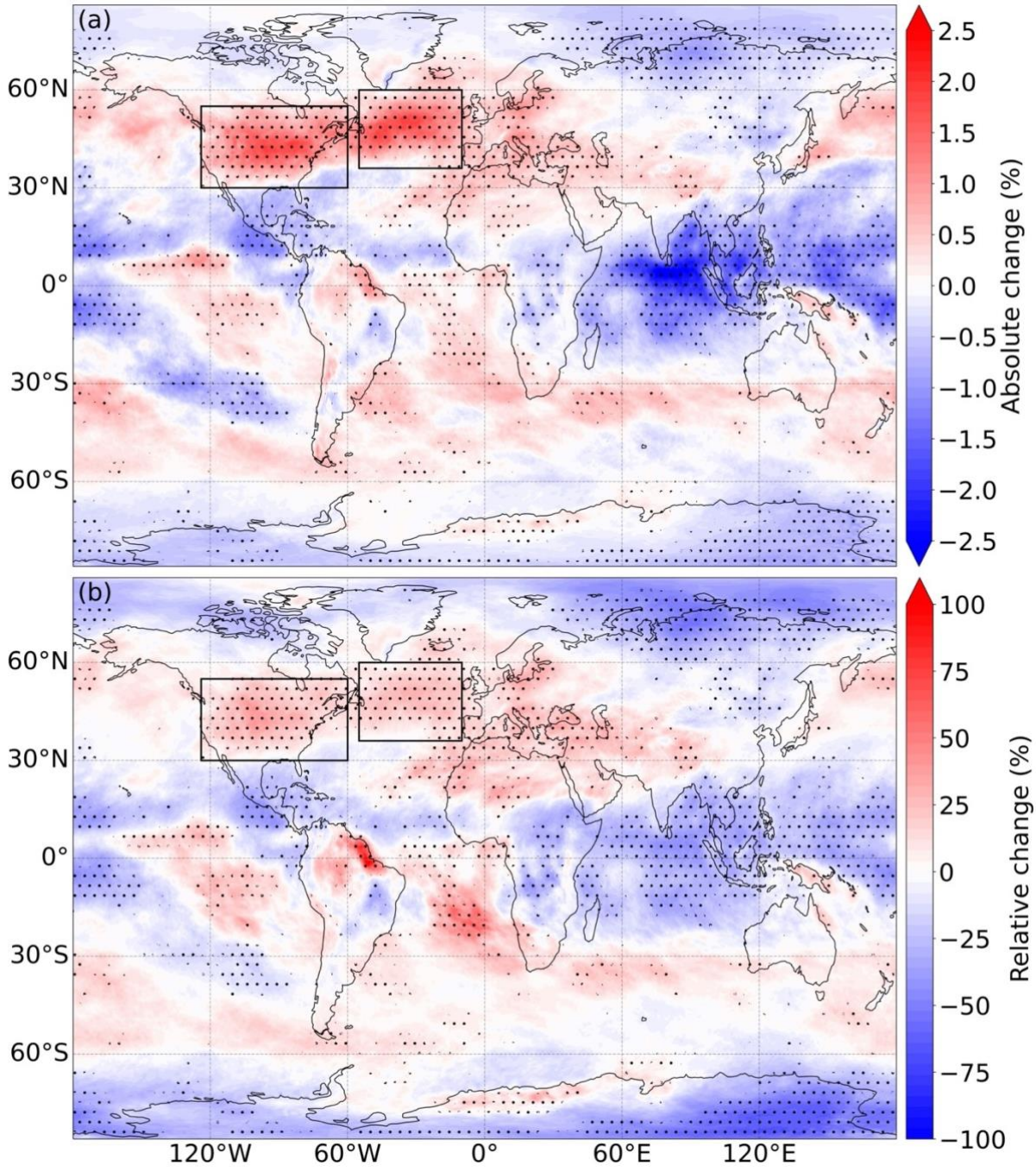


Figure 4.6: The same as for Figure 4.4, except for Light-Or-Greater Clear-air turbulence (LOG CAT) instead of Moderate-Or-Greater (MOG CAT).

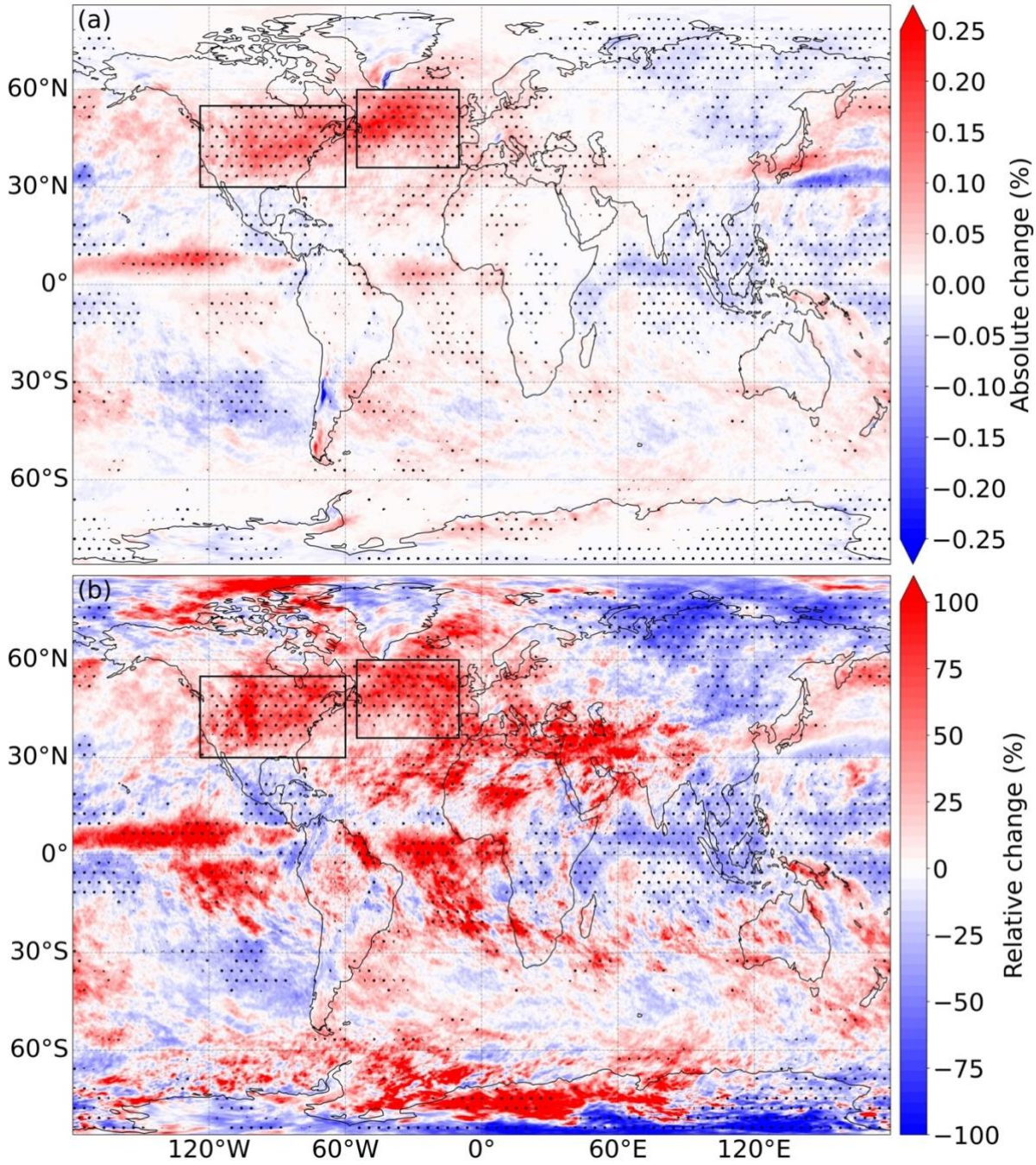


Figure 4.7: The same as for Figure 4.4, except for Severe-Or-Greater Clear-air turbulence (SOG CAT) instead of Moderate-Or-Greater (MOG CAT).

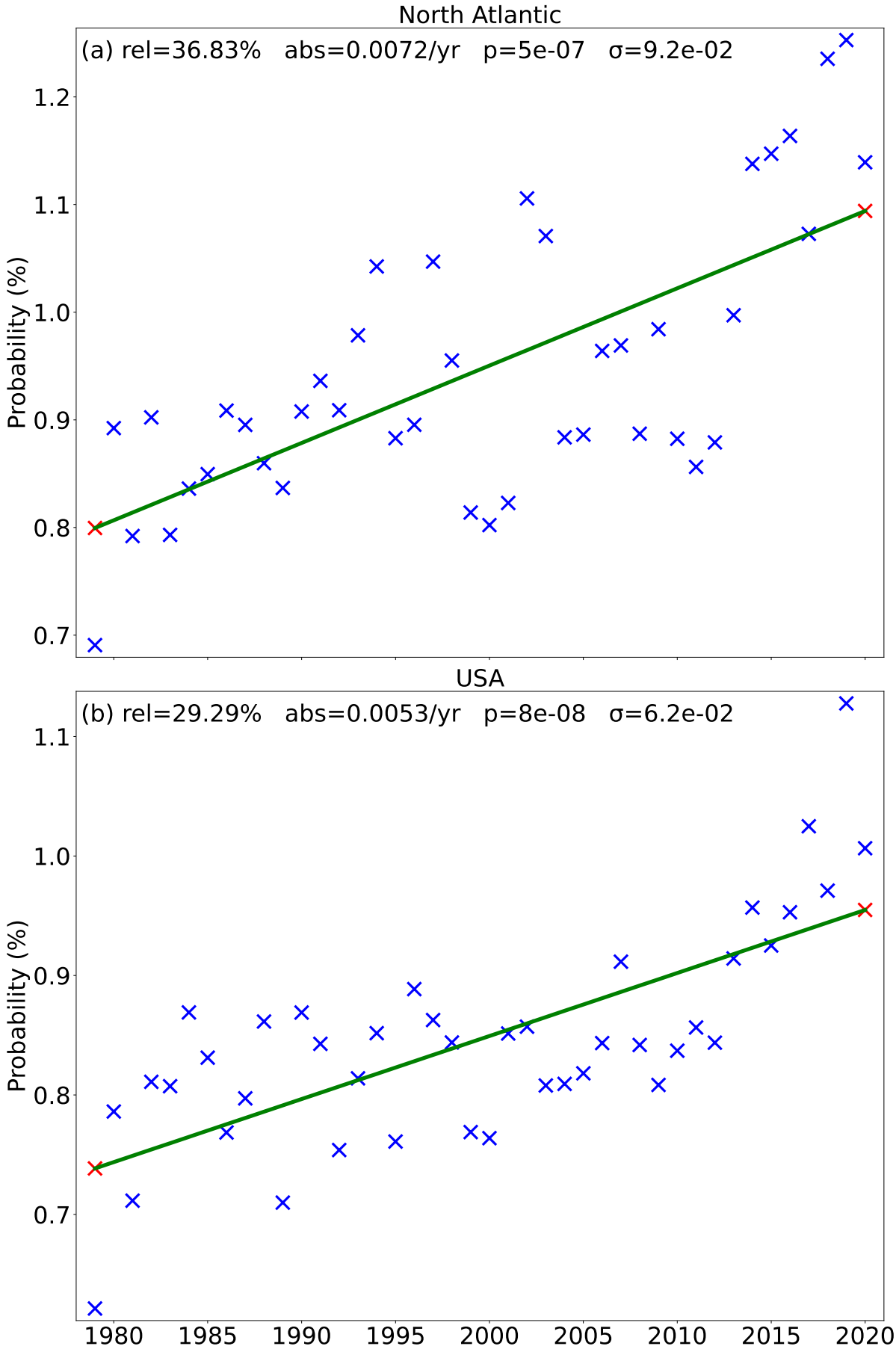


Figure 4.8: A linear regression analysis conducted on the ERA5 197 hPa annual-mean diagnostic-mean MOG CAT probability for the (a) North Atlantic and (b) USA boxes indicated in Figure 4.2. The 42 blue crosses in each panel indicate data from the 42 years, whereas the two red crosses show the fitted 1979 and 2020 values. Stated at the top of each panel are the relative change in the fit from 1979 to 2020 (%), the

absolute change per year calculated as the slope of the regression line (%/year), the p value for the slope, and the standard deviation of the residual (σ; %).

Figure 4.8 shows the annual probabilities of diagnosed MOG CAT for each year in the period 1979–2020 over the North Atlantic and USA, averaged over all diagnostics. There is a large increasing trend in both regions. For example, the North Atlantic starts at an absolute probability of 0.8% in 1979 and increases by 37% (in relative terms) to an absolute probability of nearly 1.1% in 2020. This increase equates to more than a whole day's worth of additional diagnosed MOG CAT exceedances per year in 2020 relative to 1979. The interannual variations in turbulence are noticeably greater in the North Atlantic than the USA, possibly because of the influence of the North Atlantic Oscillation (NAO); see Kim et al. (2016, 2020). Note that, over the 42-year period, the average latitude of the subpolar jet may have shifted, but this shift is negligible compared to the latitudinal extent of these boxes (Archer and Caldeira, 2008; Simmons, 2022). The same figure but for LOG and SOG CAT intensities are given in Figures 4.9 and 4.10 respectively (increases of 17% and 54% respectively)

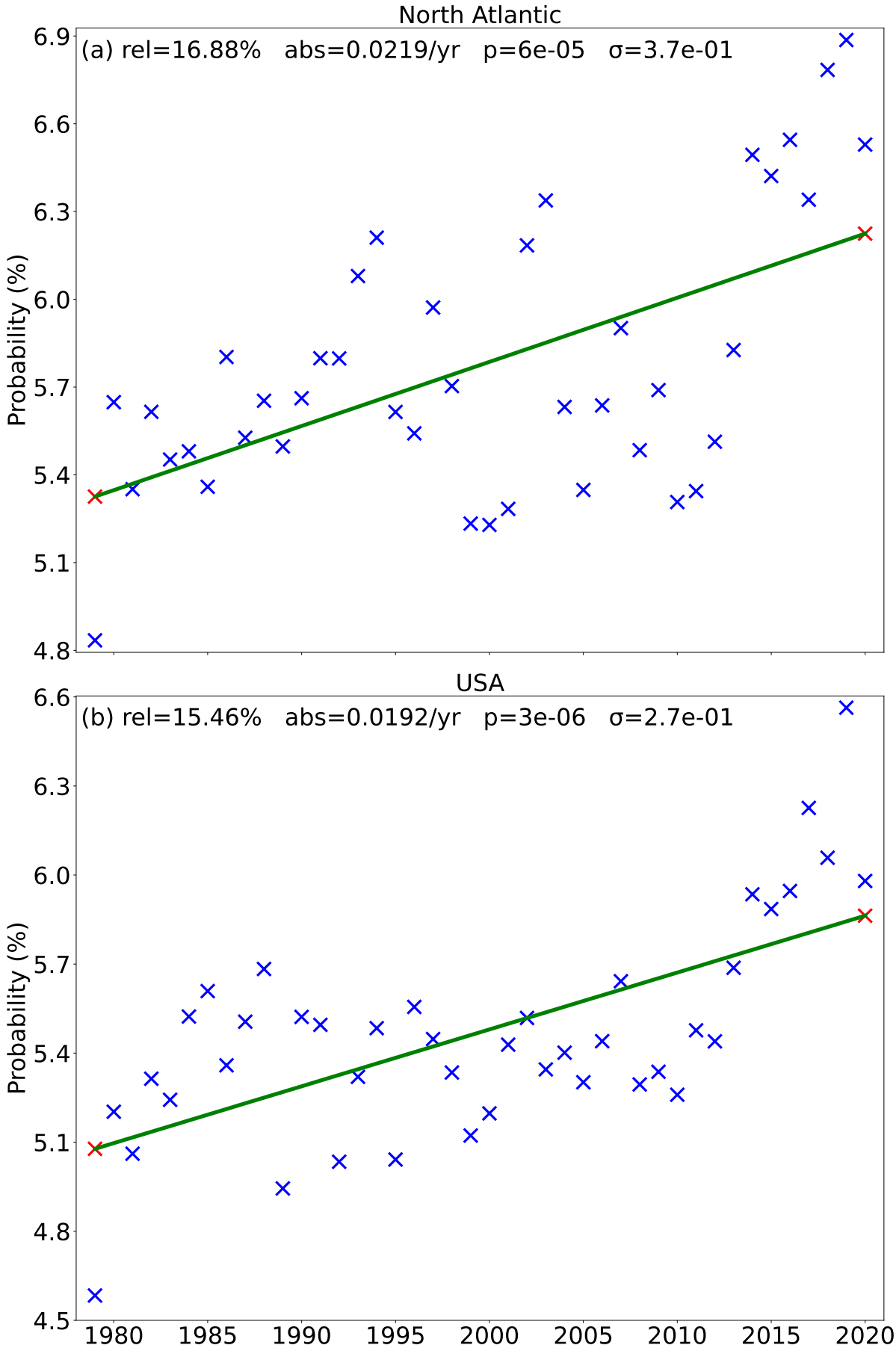


Figure 4.9: The same as for Figure 4.8, except for Light-Or-Greater Clear-air turbulence (LOG CAT) instead of Moderate-Or-Greater (MOG CAT).

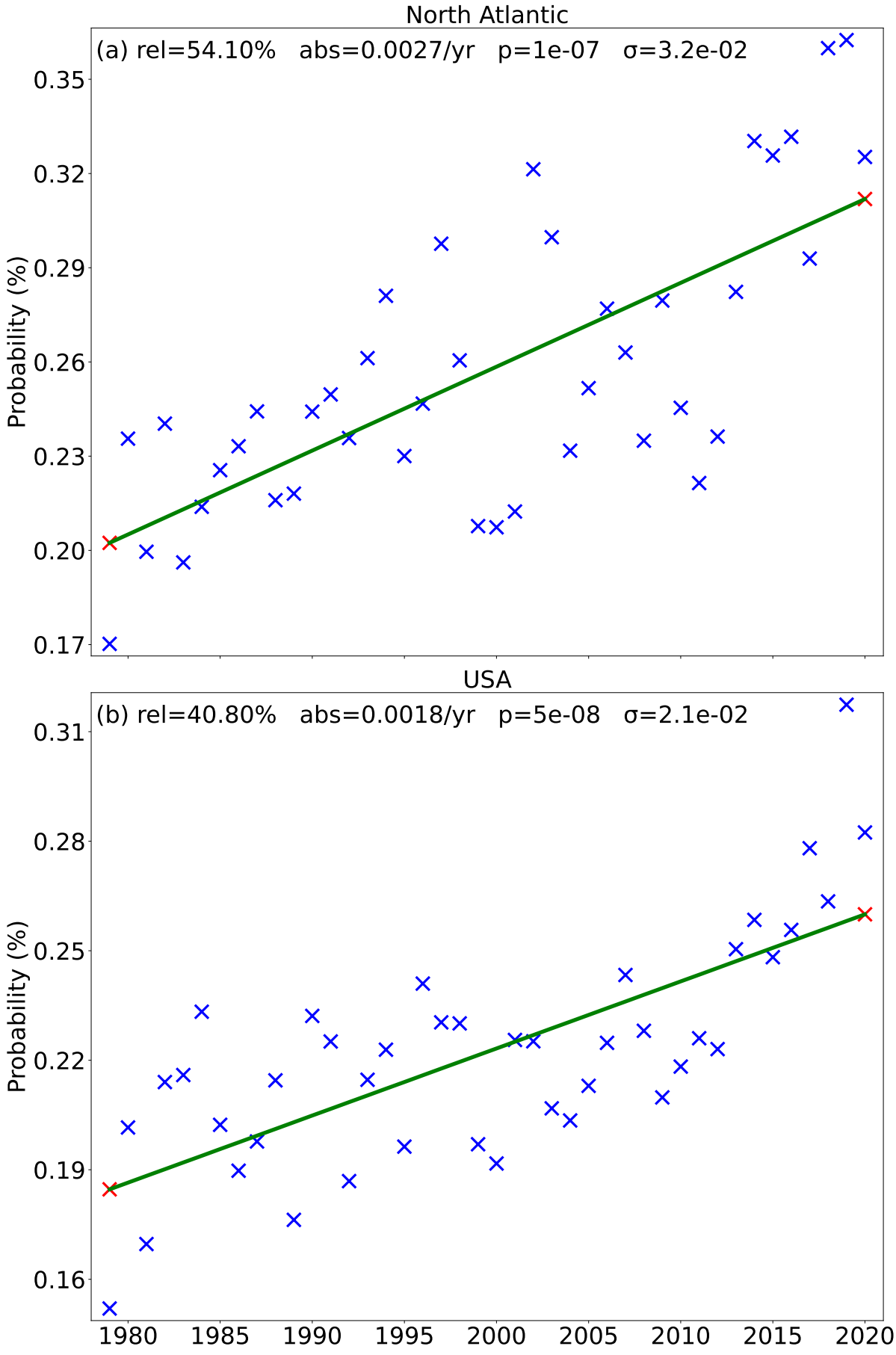


Figure 4.10: The same as for Figure 4.8, except for Severe-Or-Greater Clear-air turbulence (SOG CAT) instead of Moderate-Or-Greater (MOG CAT).

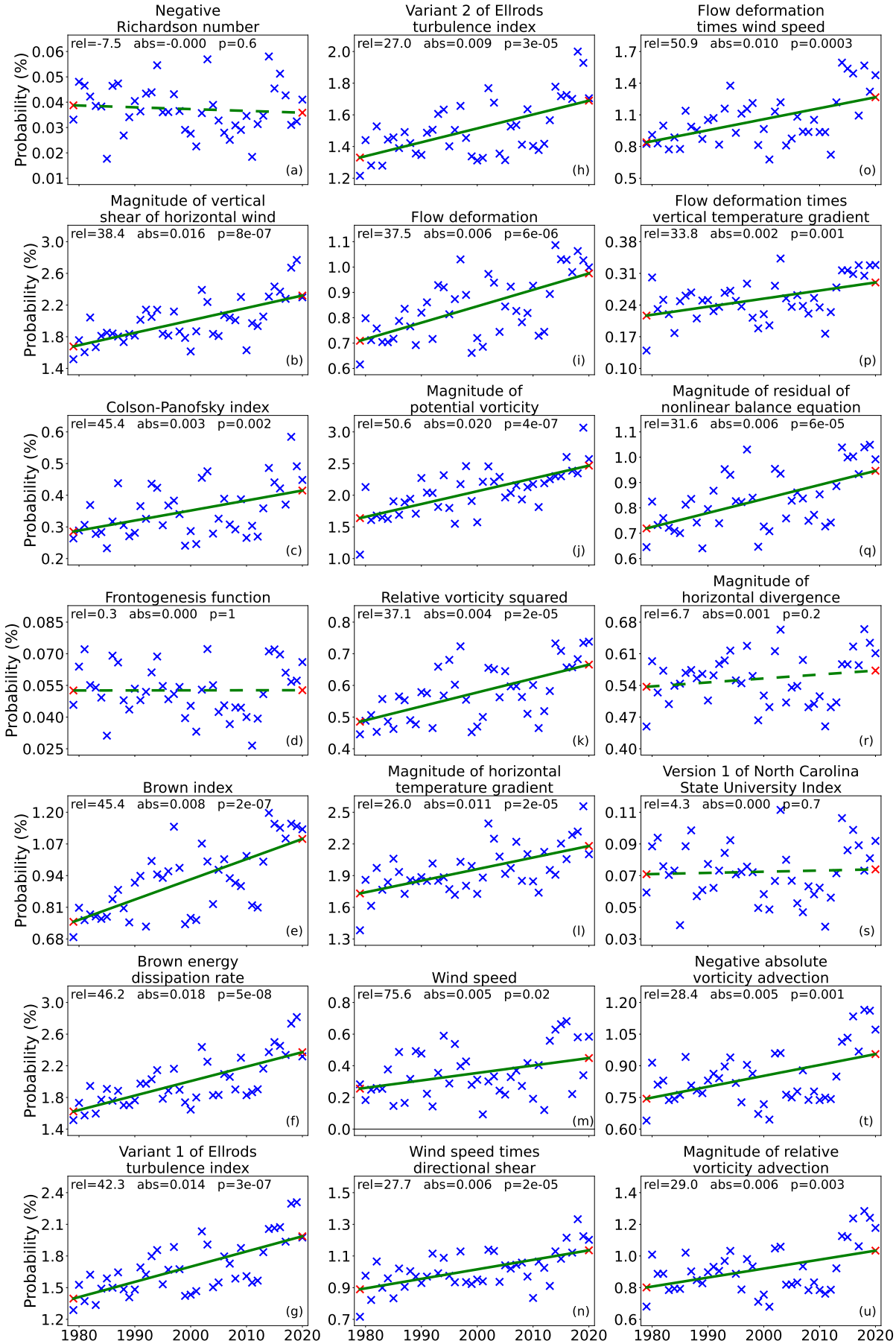


Figure 4.11: A linear regression analysis conducted on the ERA5 197 hPa annual-mean MOG CAT probability for the North Atlantic box indicated in Figure 4.8(a), for each of the 21 CAT diagnostics. The 42 blue crosses in each panel indicate data from the 42 years, whereas the two red crosses show the fitted

1979 and 2020 values. Solid (dashed) green lines indicate trends that are (are not) statistically significant at the p=0.05 level, according to a two-sided Wald test. Stated at the top of each panel are the relative change in the fit from 1979 to 2020 (%), the absolute change calculated as the slope of the regression line (%/year), and the p value for the slope.

Figure 4.11 expands on the North Atlantic analysis in Figure 4.8(a) by decomposing it into the 21 constituent CAT diagnostics. It is seen that 17 of the 21 diagnostics show significant (p<0.05) upward trends, with relative changes of up to 75.6%. The remaining 4 diagnostics show no significant trend, and none of the diagnostics shows a significant downward trend. These results indicate a high level of inter-diagnostic agreement that MOG CAT increased over the study period. The same figure but for LOG and SOG CAT intensities are given in Figures 4.12 and 4.13 respectively.

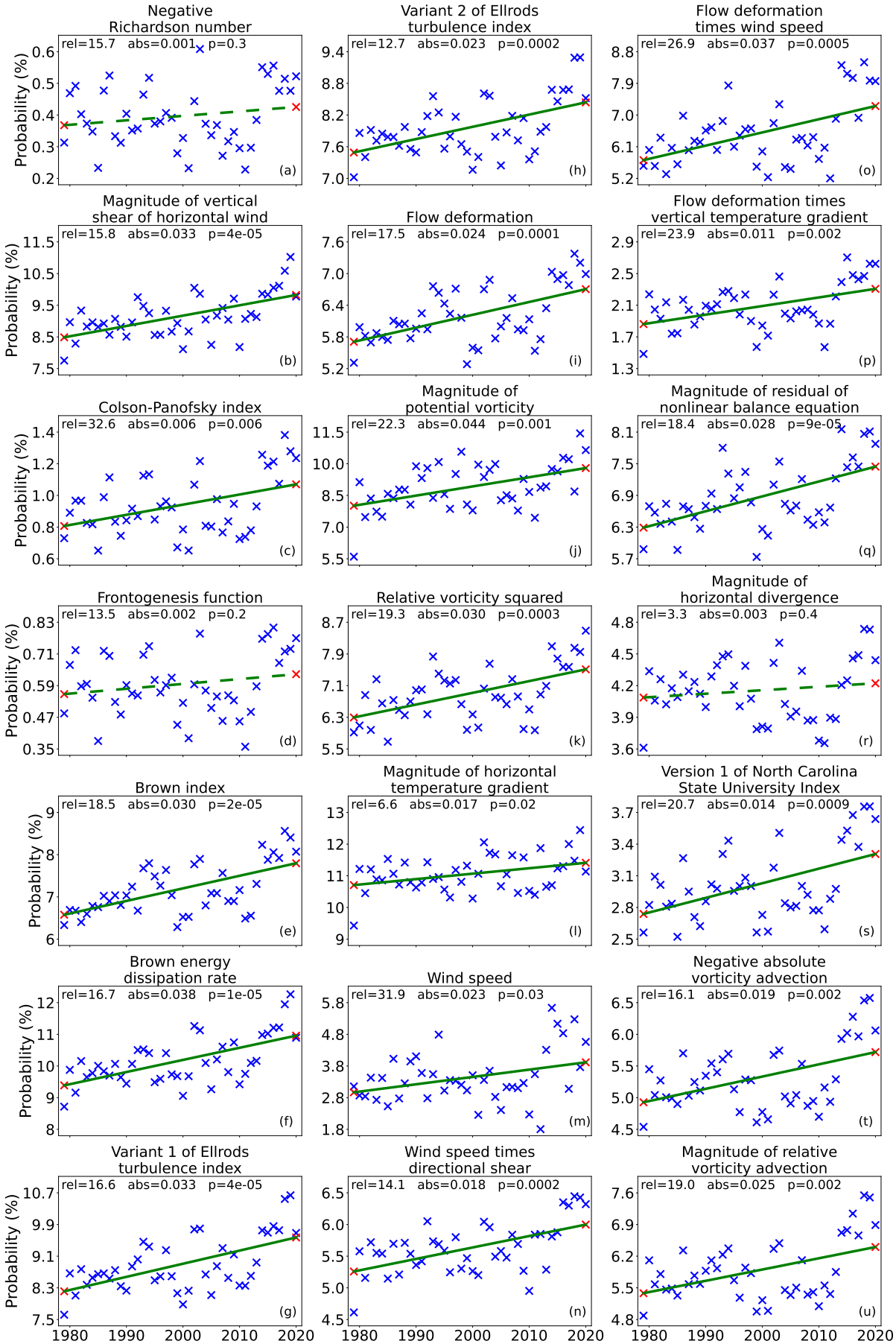


Figure 4.12: The same as for Figure 4.11, except for Light-Or-Greater Clear-air turbulence (LOG CAT) instead of Moderate-Or-Greater (MOG CAT).

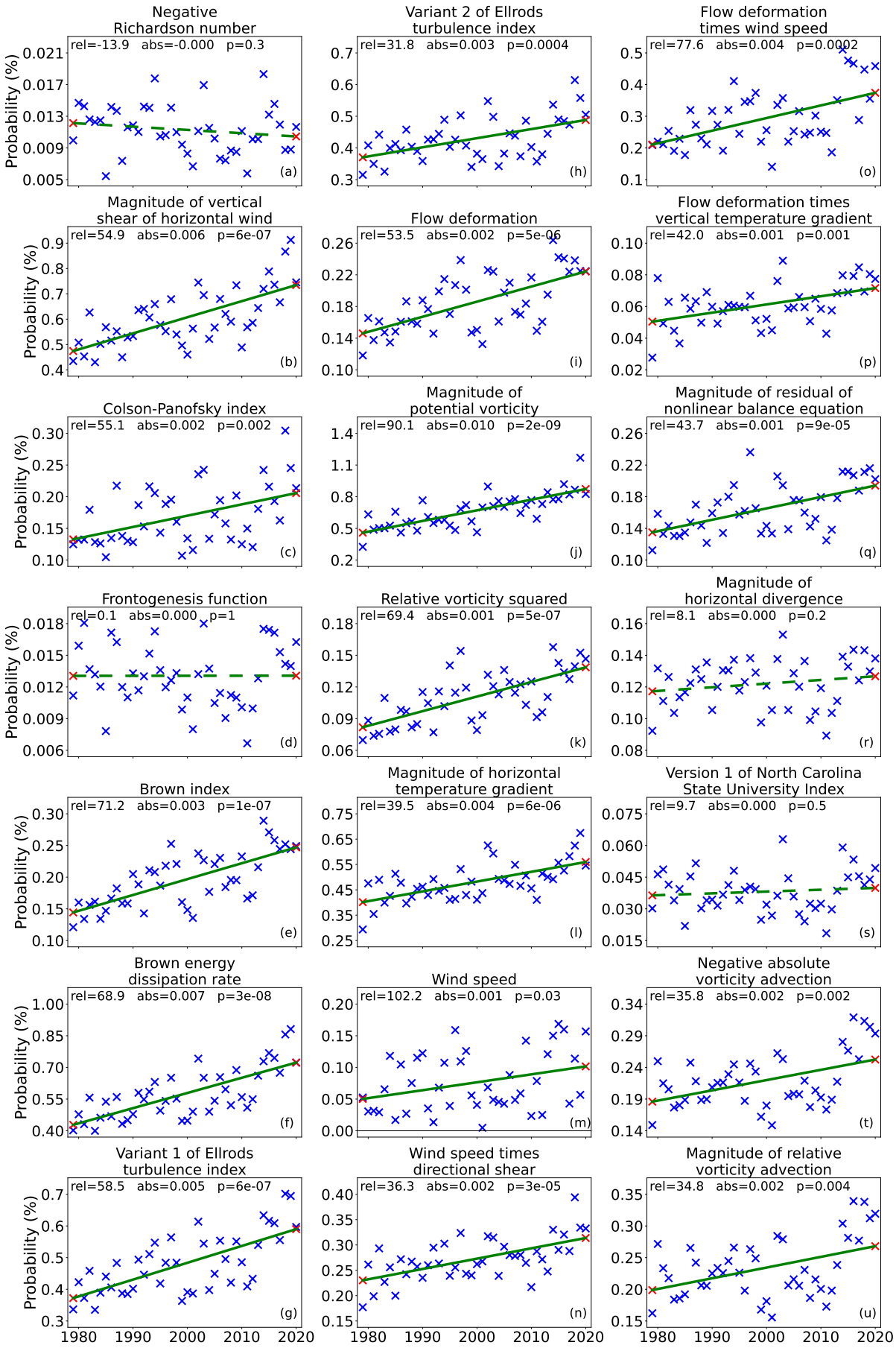

Negative Richardson number
rel=-13.9 abs=-0.000 p=0.3
(a)
Variant 2 of Ellrods turbulence index
rel=31.8 abs=0.003 p=0.0004
(h)
Flow deformation times wind speed
rel=77.6 abs=0.004 p=0.0002
(o)
Magnitude of vertical shear of horizontal wind
rel=54.9 abs=0.006 p=6e-07
(b)
Flow deformation
rel=53.5 abs=0.002 p=5e-06
(i)
Flow deformation times vertical temperature gradient
rel=42.0 abs=0.001 p=0.001
(p)
Colson-Panofsky index
rel=55.1 abs=0.002 p=0.002
(c)
Magnitude of potential vorticity
rel=90.1 abs=0.010 p=2e-09
(j)
Magnitude of residual of nonlinear balance equation
rel=43.7 abs=0.001 p=9e-05
(q)
Frontogenesis function
rel=0.1 abs=0.000 p=1
(d)
Relative vorticity squared
rel=69.4 abs=0.001 p=5e-07
(k)
Magnitude of horizontal divergence
rel=8.1 abs=0.000 p=0.2
(r)
Brown index
rel=71.2 abs=0.003 p=1e-07
(e)
Magnitude of horizontal temperature gradient
rel=39.5 abs=0.004 p=6e-06
(l)
Version 1 of North Carolina State University Index
rel=9.7 abs=0.000 p=0.5
(s)
Brown energy dissipation rate
rel=68.9 abs=0.007 p=3e-08
(f)
Wind speed
rel=102.2 abs=0.001 p=0.03
(m)
Negative absolute vorticity advection
rel=35.8 abs=0.002 p=0.002
(t)
Variant 1 of Ellrods turbulence index
rel=58.5 abs=0.005 p=6e-07
(g)
Wind speed times directional shear
rel=36.3 abs=0.002 p=3e-05
(n)
Magnitude of relative vorticity advection
rel=34.8 abs=0.002 p=0.004
(u)
Probability (%)
1980 1990 2000 2010 2020

Figure 4.13: The same as for Figure 4.11, except for Severe-Or-Greater Clear-air turbulence (SOG CAT) instead of Moderate-Or-Greater (MOG CAT).

threshold	LOG	LMOG	MOG	MSOG	SOG

season						
DJF	1979 turbulence (h)	128.9	45.6	22.3	12.1	6.4
	2020 turbulence (h)	155.6	59.3	30.6	17.2	9.6
	Absolute increase (h)	26.7	13.7	8.3	5.2	3.1
	Relative increase (%)	21	30	37	43	49
MAM	1979 turbulence (h)	90.4	27.2	11.8	5.7	2.7
	2020 turbulence (h)	113.4	38.9	18.6	9.7	5.0
	Absolute increase (h)	23.1	11.7	6.8	4.0	2.3
	Relative increase (%)	26	43	57	71	85
JJA	1979 turbulence (h)	114.1	36.5	16.1	7.7	3.6
	2020 turbulence (h)	124.5	43.8	21.1	10.9	5.5
	Absolute increase (h)	10.4	7.3	5.0	3.2	1.9
	Relative increase (%)	9	20	31	41	52
SON	1979 turbulence (h)	133.1	43.4	19.8	10.0	5.0
	2020 turbulence (h)	153.2	53.4	25.8	13.9	7.4
	Absolute increase (h)	20.1	10.0	6.0	3.9	2.4
	Relative increase (%)	15	23	31	39	47
Annual	1979 turbulence (h)	466.5	152.7	70.0	35.5	17.7
	2020 turbulence (h)	546.8	195.4	96.1	51.8	27.4
	Absolute increase (h)	80.2	42.7	26.1	16.3	9.7
	Relative increase (%)	17	28	37	46	55

Table 4.1: Fitted changes from 1979 to 2020 in the number of hours (per season and annually) spent in CAT, derived from ERA5 at 197 hPa using a diagnostic-mean calculation, for an average point in the North Atlantic box indicated in Figure 2. The changes are broken down by season for five turbulence strength thresholds: Light-Or-Greater (LOG), Light-to-Moderate-Or-Greater (LMOG), Moderate-Or-Greater (MOG), Moderate-to-Severe-Or-Greater (MSOG), and severe-or-greater (SOG). The changes are statistically significant for each combination of season and threshold ($p < 3\text{x}10^{-2}$ in all cases).

Table 4.1 shows the absolute and relative changes in hours of diagnosed CAT averaged in the North Atlantic, broken down by season and turbulence severity. There is a seasonal cycle, with autumn and winter having more CAT (of all strengths) than spring and summer, because of the seasonal cycle in jet stream

strength. The number of hours spent in turbulence generally decreases with increasing turbulence strength, because climatologically stronger turbulence is rarer. For LOG CAT, annually there were 466.5 hours of turbulence in 1979, increasing by 80.2 hours (17%) to 546.8 hours in 2020. In contrast, the relative increases are generally larger with increasing turbulence strength, consistent with the future projections of Williams (2017). For SOG CAT annually, there were 17.7 hours of turbulence in 1979, increasing by 9.7 hours (55%) to 27.4 hours in 2020.

4.4 Summary and Discussion

In this paper, trends in ERA5's 21-diagnostic-averaged CAT probability at 197 hPa over the period 1979–2020 were examined. The largest increases in both absolute and relative MOG CAT were found over the North Atlantic and continental United States, with statistically significant absolute increases of 0.3% (26 hours) and 0.22% (19 hours), respectively, over the total reanalysis period. Absolute changes are important in regard to aircraft damage, as every additional minute spent traversing turbulence causes fatigue and increases wear-and-tear on the airframe and increases maintenance costs and the potential for injuries, irrespective of whether the increase is on top of a low or high base rate.

The above two hotspots for increased CAT contrast with the East Asian and East Pacific hotspots identified by Lee et al. (2023). There is an important methodological difference between our studies, in addition to the different seasons and diagnostics used, that may account for these different results. This difference likely arises because many CAT diagnostics require the computation of vertical derivatives. To compute these derivatives, we used input fields at 206 hPa and 188 hPa to calculate the diagnostic values at 197 hPa. In contrast, Lee et al. (2023) appear to have used input fields at 200 hPa and 300 hPa to calculate the diagnostic values at 250 hPa. This means we have used much finer (and therefore more accurate) vertical finite differences to compute the gradients. Interestingly, of all the CAT diagnostics that the studies have in common, the two that do not require the computation of a vertical derivative (namely the deformation and divergence) are very similar between the two studies, lending support to this explanation for the differences. In this paper we have also examined 21 diagnostics, many more than Lee et al. (2023), and we have examined the whole globe rather than just the northern extra-tropics (20–60°N).

The corresponding relative increases for the end of the reanalysis period compared to the start are 37% for the North Atlantic and 29% for the USA. These relative changes are useful for diagnosing which regions are expected to become significantly more turbulent. For example, Figure 4.4(b) shows that over the northern coast of Brazil there is a 100% relative increase in MOG CAT, which informs us that even though the baseline occurrence is relatively low compared to other regions, the frequency has now doubled compared to the start of the period. This is of great importance, as these regions cover some of the globe's busiest flight corridors. Our study represents the best evidence yet that CAT has increased over the past four decades.

Regions over western ocean basins are hot spots for diagnosed CAT. This is partly because jet streams tend to be fastest over the ocean, due to the low surface roughness compared to the land. It is also partly because there is a large zonal temperature contrast between the ocean and continent at the western boundary, especially in winter, due to their different specific heat capacities. These horizontal temperature gradients contribute to vertical wind shear, which in turn contribute to CAT.

Future CAT projections by Williams and Joshi (2013) and Williams (2017) using climate models showed increases over the North Atlantic region in DJF for a doubling of CO2 relative to a pre-industrial control run (560 ppm and 280 ppm CO2, respectively). For example, an average across the 21 diagnostics used by Williams and Joshi (2013) gives an 83.9% increase in MOG CAT occurrence for doubled CO2. During the 1979–2020 period, CO2 concentrations rose by around 30% (from 335 to 410 ppm), yet our study finds a CAT increase of 37% in this region and season over this period, which (after accounting for the different CO2 rises) is more than we would expect from the climate model results. Williams and Storer (2022) also observed greater CAT increases in reanalysis data than a climate model. Taken together, these findings suggest that climate model simulations may underestimate future CAT increases.

Our analysis has used ERA5.1, which corrects for the known cold bias in the lower stratosphere during 2000–2006 in the previous version of ERA5 (Simmons et al., 2020). As with all reanalysis datasets, the quality and quantity of assimilated observational data generally improve over time, although ERA5 has good multidecadal consistency with the plentiful wind observations from aircraft and satellites near the tropopause (Simmons, 2022).

Future work should address the limitations of this study. The sensitivity of the results to using an equally weighted ensemble mean of CAT diagnostics should be explored. Trends in other forms of aviation-affecting turbulence apart from CAT, including convectively induced turbulence (CIT) and mountain wave turbulence (MWT), should be diagnosed from forthcoming reanalysis datasets, such as the planned ERA6 that will contain various convection diagnostics. The northern hemisphere's greater positive trend than the southern hemisphere also warrants further investigation. Turbulence data from aircraft could also be analysed, but the time period for which quantitative, automated measurements are available is far shorter than the 42 years covered here, making trend detection problematic. In the absence of a long-term record of quantitative aircraft turbulence measurements, reanalysis diagnostics provide the best available global picture of historic variations in CAT.

Chapter 5

Towards seasonal predictions of global clear-air turbulence

5.1 Introduction.

Turbulence is already a problem for the aviation industry, costing around US$200 million annually in the USA alone (Eichenbaum, 2003). The costs of Clear-air Turbulence (CAT) are partly due to the fact that CAT can be damaging to airframe, passengers and crew. CAT is difficult to observe meaning avoidance can be difficult. Compounding the problem further aircraft flying west to east tend to make frequent use of the jet streams' westerly flow to decrease journey times and increase fuel economy. However, to enter the flow, the aircraft may have to pass through regions of shear which produce CAT.

Work by Williams and Joshi (2013), Williams (2017) used an ensemble of 21 CAT diagnostics to investigate trends in north Atlantic CAT given a 2 x CO_2 climate change scenario. This was then followed by Storer et al. (2017) which used the same ensemble to look at global CAT under an RCP8.5 climate change scenario. Both studies found that the frequency of Moderate or Greater (MOG) turbulence could double in both the northern and southern hemisphere midlatitudes relative to the preindustrial period given a climate change scenario. Work by Prosser et al. (2023), used reanalysis data to show that some of this predicted increase has already occurred over the midlatitudes in both hemispheres over the past 4 decades (1979-2020) with the continental US and the North Atlantic showing particularly noticeable increases over that period.

In addition to the demonstrated long term trends, CAT also exhibits year to year variability as shown for example, in Figure 3 of Prosser et al. (2023) ENSO is the largest inter-annual perturbation in the Earth's climate system and the NAO has a profound impact on the weather regionally in the North Atlantic (Wallace and Hobbs, 2006). Kim et al. (2016) have undertaken work on the connection between the NAO and flight routing across the North Atlantic. They also looked at the probability of encountering turbulence using the TI1 turbulence index (Ellrod and Knapp, 1992) and found that east-bound flights in a -ve NAO scenario contained the greatest percentage (1.31%) of MOG turbulence along a wind optimised route. Subsequently only 1.02% of the east bound (EB) flight optimised routed contained MOG turbulence during a +ve NAO scenario. For west bound flights during +ve NAO and –ve NAO the percentage of MOG was found to be similar at 0.92 and 0.81% respectively. East-bound flights encounter more clear air turbulence given that

airlines enter jet-stream regions to reduce flight times and fuel usage. The proposed explanation for the much larger east-bound turbulence during a -NAO is that the jet stream tends to shift southwards putting aviation in the jet's cyclonic shear side which is prone to turbulence (Knox, 1997; Knox et al., 2008).

The impact of ENSO on Global CAT remains unclear. ENSO is predictable with a lead time of 6 months to a year (Tang et al., 2018). Wolff & Sharman (2008) showed that CAT frequency peaked over the continental US during a strong El Niño in 1998. Therefore, if a relationship between ENSO and CAT exists along significant flight corridors, then it could provide advantageous in forecasting CAT at seasonal timescales.

Cheyne (2020) examine the composite global vertical wind shear at 250hPa using NCEP-DOE Reanalysis 2 data in a selection of 5 particularly strong El Niño years (Oceanic Niño Index 3.4 anomaly > 1.25 K) and 5 particularly strong La Niña years (ONI3.4 anomaly < -1.25 K) They find that El Niño significantly affects vertical wind shear for 21% of the globe at 250 hPa and 23% for La Niña. While vertical wind shear is a diagnostic of global CAT, previous work (Sharman et al. 2006) has shown that a comprehensive ensemble of diagnostics is required to adequately represent CAT generation mechanisms. Here, the work of Cheyne (2020) is expanded using the ensemble of 21 diagnostics first used by Williams and Joshi (2013) applied to the ERA5 reanalysis dataset. Cheyne (2020) focused on the average value of vertical windshear whereas this work focuses on the upper the 97th, 99.1st, 99.6th, 99.8th and 99.9th percentiles of CAT diagnostic values at each given grid point in ERA5. These percentiles correspond to Light-or-greater (LOG), Light-to-moderate-or-greater (LMOG), Moderate-or-greater (MOG), Moderate-to-Severe-or-greater (MSOG) and Severe-or-greater (SOG) thresholds for turbulence intensity. This paper will be divided into three sections. Section 5.2 describes is the methodology and datasets, section 5.3 the results and section 5.4, a discussion of the results and ideas for further work.

5.2 Method

Temperature, *T*, and the zonal and meridional wind component fields, *u* and *v*, respectively at pressure levels 206, 197 and 188hPa at 3-hourly timesteps were extracted from ERA5 for the 42 years between 1979-2020. The *T*, *u*, and *v* fields for the three pressure levels were then interpolated from a N320 reduced Gaussian grid on to a regular 0.25 x 0.25 degree grid.

In order to calculate each diagnostic's threshold for the 5 thresholds used here: Light-or-greater (LOG), Light-to-moderate-or-greater (LMOG), Moderate-or-greater (MOG), Moderate-to-Severe-or-greater (MSOG) and Severe-or-greater (SOG), the 97th, 99.1st, 99.6th, 99.8th and the 99.9th percentile respectively was calculated for each of 21 diagnostics for the year 2000. The year 2000 was chosen as a reference because the whole 42 years would be too large to compute and because it was the midpoint of the study

period. Furthermore, when calculating these percentiles, the exceedance values were latitudinally weighted such that a quarter of a degree near the poles was not considered equal to one just off the equator.

Then for each grid point the frequency of exceedance for each of the CAT severity thresholds was then computed for each diagnostic and month. This was then converted to a frequency probability (i.e. 2% would equate to 5 exceedances in a 30 month containing 240 3 hourly timesteps.

In order to produce spatial maps of CAT, for each of the 21 diagnostics and 5 thresholds, the average exceedance percentage across the DJF of all years (The 'D' is the December in the preceding year) where an El Niño occurred was computed and plotted. Here El Niño years are those years with a score of >= 0.5 on the Oceanic Niño Index (ONI) 3.4, index.

El Niño	1980, 1983, 1987, 1988, 1992, 1995, 1998, 2003, 2005, 2007, 2010, 2015, 2016, 2019, 2020,
Neutral	1979, 1981, 1982, 1990, 1991, 1993, 1994, 2002, 2004, 2013, 2014, 2017,
La Niña	1984, 1985, 1986, 1989, 1996, 1997, 1999, 2000, 2001, 2006, 2008, 2009, 2011, 2012, 2018,
+NAO	1983, 1984, 1989, 1993, 1994, 1995, 1999, 2000, 2012, 2014, 2015, 2016, 2018, 2020,
Neutral	1980, 1990, 1991, 1992, 1998, 2005, 2006, 2007, 2008, 2009, 2013, 2016, 2017, 2019,
-NAO	1979, 1982, 1985, 1986, 1987, 1988, 1996, 1997, 2001, 2002, 2003, 2004, 2010, 2011,

Table 5.1: Table of El Niño, neutral and La Niña years as well as positive, neutral and negative NAO years.

A diagnostic average is then taken to give the El Niño CAT spatial map as shown in figure 2a. The procedure is the same for La Niña and Neutral where years with an ONI 3.4 index score of >=-0.5 and between -0.5 & 0.5 respectively are used. El Niño minus Neutral, El Niño minus La Niña and La-Nino minus Neutral plots are then created.

A diagnostic average exceedance was chosen over the unequal weightings given in chapter 3 because the latter apply specifically to injury causing turbulence, a specific rather than a more general class of aviation affecting turbulence.

The regression of CAT probability onto the ONI 3.4 index scores undertaken for several study areas for DJFs between 1979-2020. These study areas are defined in table 5.2 and shown in figure 5.2d and similar to those used by Cheyne (2020). For each study area or box and for each DJF the diagnostic and 3 month average probability were then calculated. These were then scatter plotted and linear regressions along with significance tests were performed.

Table 5.2: The latitudinal and longitudinal extent of various regions used in this study.

Region	Latitude Limits (°)		Longitude Limits (°)		ENSO/NAO
	Southern	Northern	Western	Eastern	
World	-90	90	-180	180	ENSO
NA Ocean	30	60	-70	-2.5	ENSO
US/Mexico	15	50	-140	-70	BOTH
US/NAO	30	60	-124	-10	ENSO
Europe	35	60	0	45	BOTH
East Asia	30	60	90	150	ENSO
NA Ocean-N	47.5	60	-70	-2.5	NAO
NA Ocean-S2	30	47.5	-43.5	-18.5	NAO
NA Ocean-S	30	47.5	-70	-2.5	NAO

Table 5.3: The 21 CAT diagnostics used in this study. The stars in the left hand column indicate which diagnostics are included in GTG (Sharman et al. 2006) See Appendix A for diagnostic equations.

No.	Diagnostic name	Units
01:*	Negative Richardson number	—
02:	Magnitude of vertical shear of horizontal wind	s^{-1}
03:*	Colson-Panofsky index	kt^2
04:*	Frontogenesis function	$m^2 \cdot s^{-3} \cdot K^{-2}$
05:	Brown index	s^{-1}
06:	Brown energy dissipation rate	$J \cdot kg^{-1} \cdot s^{-1}$
07:*	Variant 1 of Ellrod's turbulence index	s^{-2}
08:	Variant 2 of Ellrod's turbulence index	s^{-2}
09:	Flow deformation	s^{-1}
10:	Magnitude of potential vorticity	PVU
11:	Relative vorticity squared	s^{-2}
12:*	Magnitude of horizontal temperature gradient	$K \cdot m^{-1}$
13:	Wind speed	$m \cdot s^{-1}$
14:	Wind speed times directional shear	$rad \cdot s^{-1}$
15:	Flow deformation times wind speed	$m \cdot s^{-2}$
16:	Flow deformation times vertical temperature gradient	$K \cdot m^{-1} \cdot s^{-1}$
17:*	Magnitude of residual of nonlinear balance equation	s^{-2}
18:	Magnitude of horizontal divergence	s^{-1}
19:*	Version 1 of North Carolina State University Index	s^{-3}
20:	Negative absolute vorticity advection	s^{-2}
21:	Magnitude of relative vorticity advection	s^{-2}

5.3 Results

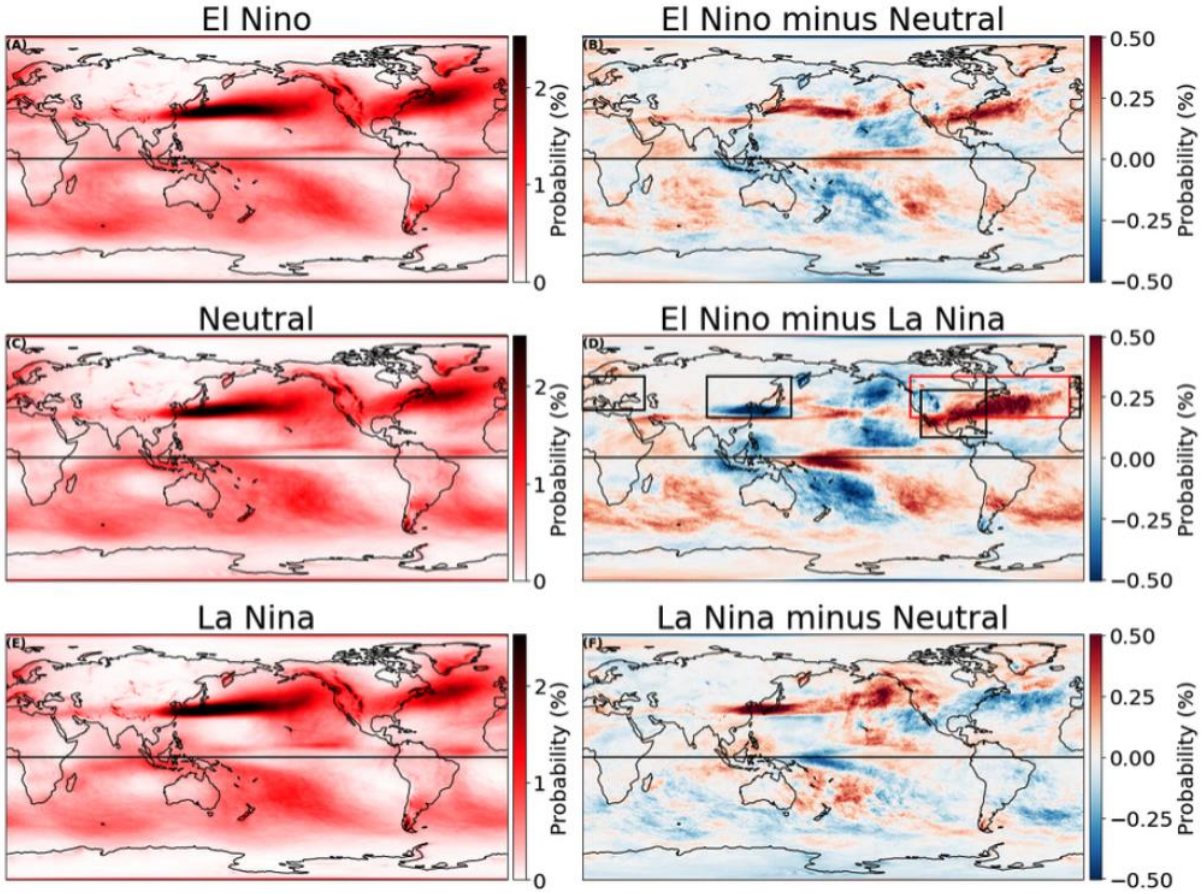


Figure 5.2: Panels A, C and E shows the probability of moderate strength CAT in percent in 15x El Niño years, 12x Neutral years and 15x La Niña years between 1979-2020. The right hand-side of the figure shows: the difference between El Niño and Neutral years (B), El Niño and La Niña years and (D) and La Niña and neutral years (F)

Figure 5.2 shows the 197 hPa diagnostic aggregated CAT distribution for (A) 15x El Nino years, (C) 12x Neutral years and (E) 15x La Niña years. In general, from the 3 permutations of difference plots, we see that the strongest difference are between the El Niño and La Niña composites with a majority of 'hot spots' (differences in CAT) occurring in and around the Pacific.

For the regions where CAT is stronger in the El Niño phase, there is a strong region covering the north of Mexico/south of the US continuing northeast over the gulf stream north Atlantic towards the UK. This pattern is mirrored (albeit not as strongly) in the Southern Hemisphere midlatitudes and there is also a Western equatorial region as well. The El Niño minus Neutral and La Niña minus Neutral plots indicate that the relationship is such that neutral conditions exist midway between El Niño and La Niña as might be expected.

For areas that present more CAT during a La Niña, there are 3 separate regions in the Pacific Ocean, one centred on equatorial Australasia and a band across Japan and into China. The comparison of the El Niño minus Neutral and La Niña minus Neutral plots indicates that while La Niña has the most CAT, El Niño and Neutral are pretty equally matched in the volume of CAT that they contain over Japan and East Asia.

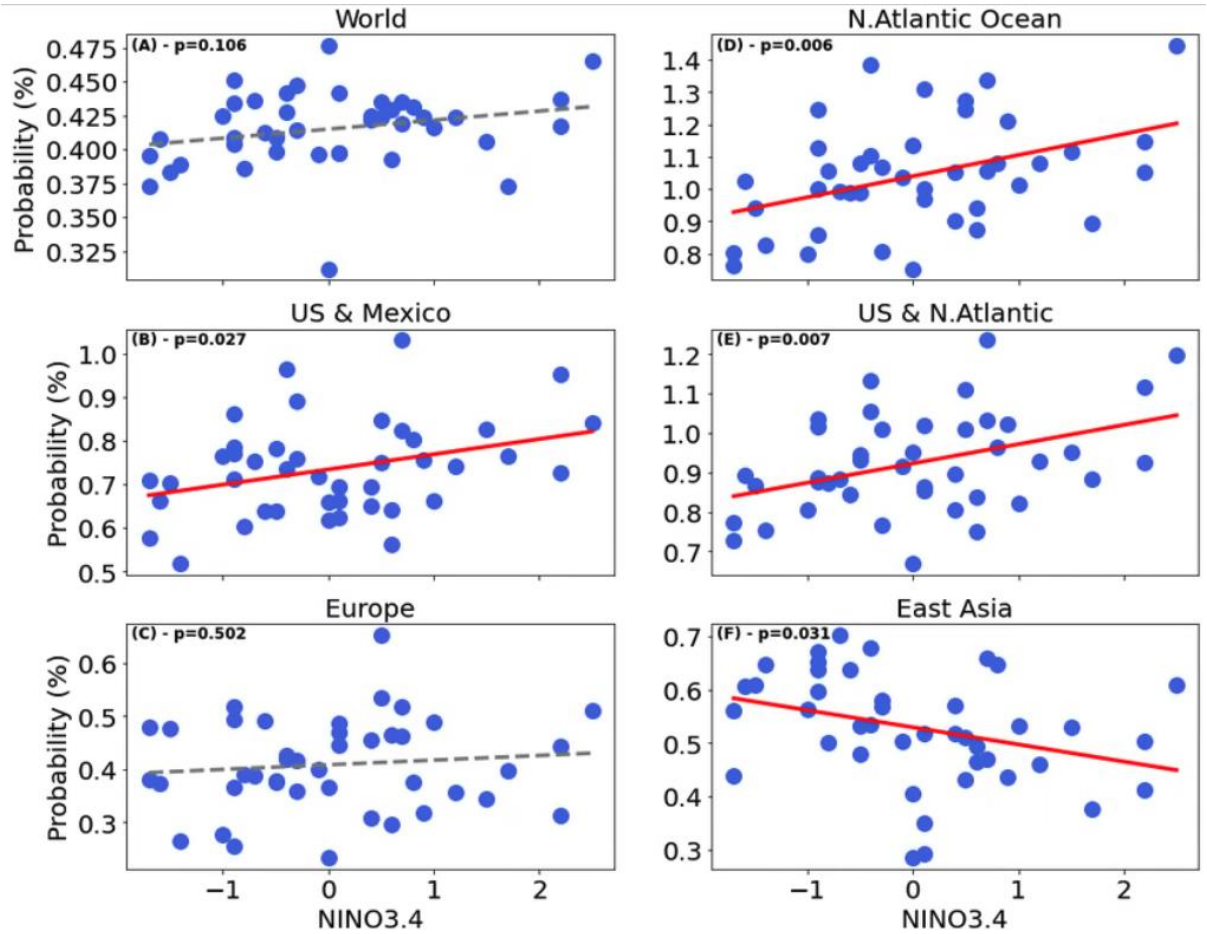


Figure 5.3: Shows the relationship between ONI 3.4, and the average CAT probability in various geographic box (see figure 5.2d and Table 5.2) between 1979-2020. The geographic boxes are (A) The World, (B) US and Mexico, (C) Europe, (D) The North Atlantic Ocean, (E) US and North Atlantic Ocean and (F) east Asia. Linear regression lines that are significant at the 5% level are shown with a solid red line.

Worldwide there is no significant trend between the state of ENSO and the amount of CAT, however over both the northern North Atlantic Ocean and the US and Mexico there is a significant positive relationship (p=0.006 and p=0.027 respectively). These regions combined give a strong positive relationship (significance of p=0.007). Not all regions show a positive relationship however with East Asia being the exception with a negative relationship between the ENSO phase and CAT prevalence (p=0.031), a pattern shown clearly in Figure 5.3. Europe seems to show no relationship one way or the other.

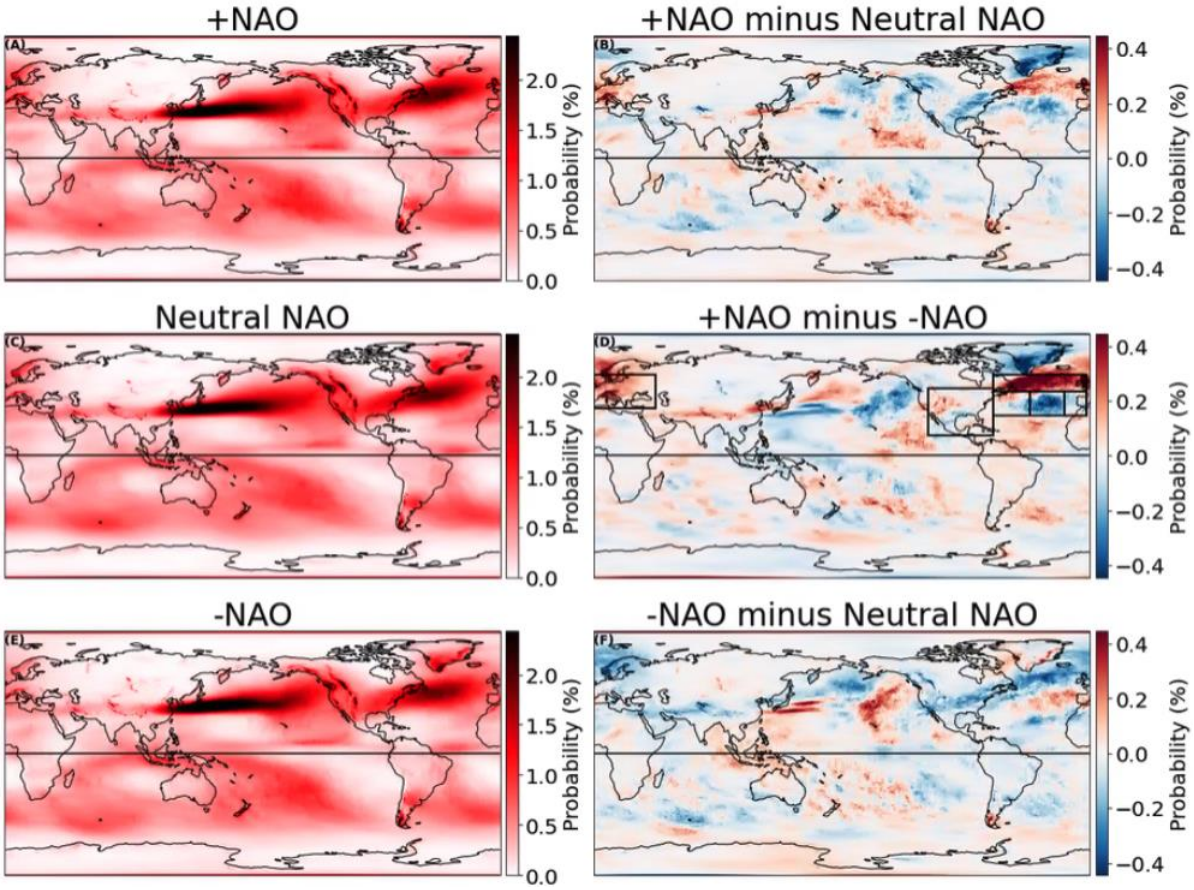


Figure 5.4: Panels A, C and E shows the probability of moderate strength CAT in percent in the 14 most positive NAO years 1979-2020 (A), 14x neutral years (C) and 14x most negative NAO years (E). The right hand-side of the figure shows: the difference between: the most positive and neutral NAO years (B), the most positive and negative NAO years (D) and the most negative and neutral NAO years (F).

Figure 5.4 shows the 197 hPa diagnostic aggregated CAT distribution for (A) 14x +NAO years, (C) 14x Neutral NAO years and (E) 14x -NAO years. The NAO data was obtained from Hurrell (2024) and +NAO/-NAO was defined as the 14 most/least positive DJF averages between 1979-2020. In comparison to ENSO, the NAO's impact on CAT seems to be more regional, more localised to the northern North Atlantic and Europe. The pattern over the north Atlantic is most clearly seen in the +NAO minus -NAO difference plot in panel D. There, a tripole structure exists where during +NAO conditions, greater CAT probability occurs on average over the northern half of the northern north Atlantic and Europe coinciding with the route most trans-Atlantic commercial aviation take. This is sandwiched between two regions of reduced average CAT. The opposite would be true from the perspective of an -NAO year.

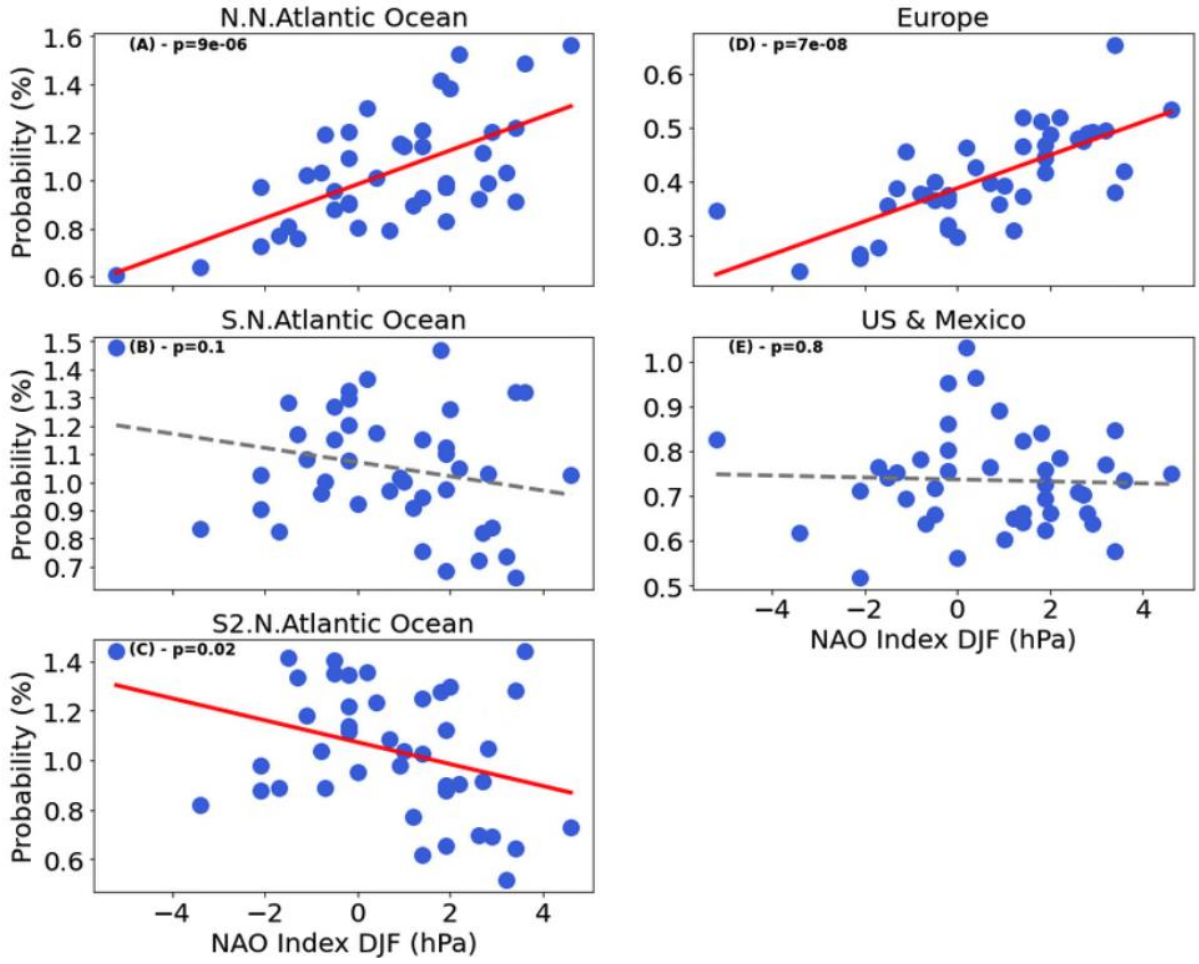


Figure 5.5: Shows the relationship between the Hurrell Station-Based Seasonal NAO Index and the average CAT probability in various geographic box (see figure 5.4d and Table 5.2) between 1979-2020. The geographic boxes are (A) Europe, (B) 47.5-60N North Atlantic Ocean, (C) US & Mexico, (D) 30-47.5N North Atlantic Ocean (Special longitude) and (E) 30-60N North Atlantic Ocean. Linear regression lines that are significant at the 5% level are shown with a solid red line.

As expected, Europe and the northern half of the north Atlantic (47.5-60 degrees latitude) show a very strong positive trend ($p<0.001$) between the NAO and 197hPa CAT. The southern half (30-47.5 degrees latitude) of the north Atlantic however does not show a significant relationship (Fig. 5.5, Panel E). If the longitude of the north Atlantic box is shrunk down from 67.5 degrees (-70W—2.5E) to 24.5 degrees (-43W—18.5E) a significant negative relationship can be found ($p=0.022$, Fig. 5.5, Panel D). In contrast to ENSO, the NAO seems to have no discernible relationship to the US and Mexico box with most of its impact being over the northern Atlantic or eastwards.

In figures 5.2-5.5 it has been shown that ENSO and the NAO definitely have an impact on the prevalence of 197hPa CAT. This leads to the question of whether there are other sources of variability in the Earth system that also contribute to CAT variability? Are ENSO and CAT even the top two most important sources of variability? To ascertain this, an Empirical orthogonal functions (EOF) analysis was conducted, the results of which are shown in figure 5.6. EOF analysis (also known as Principal Component Analysis)

is a technique often used in climate and environmental science to study potential spatial patterns of variability and how they change with time (see Hannachi, 2004 for a primer).

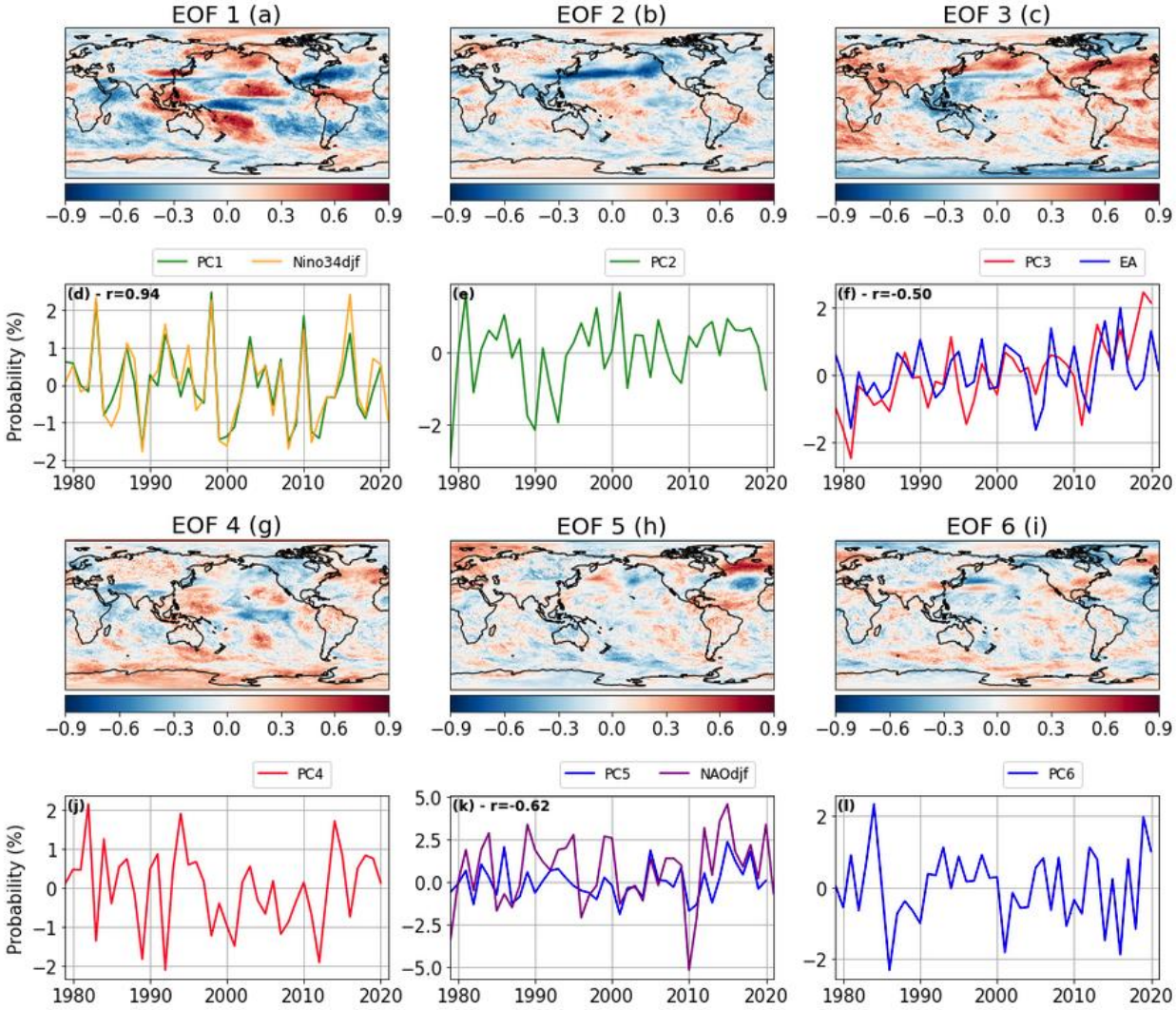


Figure 5.6: Panel a shows the 1st EOF and panel d shows the associated time series, PC1 (along with the ONI 3.4, timeseries for comparison) applied to 1979-2020 DJF CAT probabilities. Panel c and f show the 3rd EOF and the associated time series, PC3 (along with the East Atlantic Oscillation [EAO] timeseries for comparison). Panel h and k show the 5th EOF and the associated time series, PC5 (along with the NAO timeseries for comparison) respectively.

To check for other sources of variability, an EOF analysis was performed on the 42xDJF CAT dataset, the results of which can be seen in figure 5.6. The first EOF is clearly ENSO and matches up both spatially and temporally (r=0.94). The 2nd EOF (panels b and e) does indicate a source of variability in the midlatitude north Pacific, however this does not match up either spatially or physically with any known source of climate variability (e.g. the Pacific Decadal Oscillation [PDO], North Pacific Index [NPI], North Pacific Gyre Oscillation [NPGO], Atlantic Multidecadal Oscillation [AMO], Quasi-Biennial Oscillation [QBO]).

PC3 has a fairly high correlation with the East Atlantic Oscillation (r=0.5), but the spatial pattern is somewhat dissimilar. It has been included here as work by Irvine et al. (2013) has shown that the East

Atlantic Index (along with the NAO) influence jet stream latitude and are therefore likely to be relevant to the locations of CAT in the north Atlantic since 2/3rds of CAT occurs in the vicinity of the jet stream (Ellrod et al., 2003).

For EOF5, on top of having the correct spatial pattern, the PC5 timeseries has a high enough correlation with EOF5 (r=0.62) that it is highly likely to be the NAO.

EOF analysis requires that each EOF is strictly orthogonal with every other EOF/source of variability and this may not be strictly true between sources of climate variability such as ENSO and the NAO, therefore a correlation of 0.62 is likely sufficient to diagnose the NAO.

This analysis leads to the conclusion that ENSO is relevant for predicting values of 197 hPa CAT globally and the NAO (and possibly the East Atlantic index) are relevant for predicting 197hPa CAT over the north Atlantic region. But in this region, the question naturally arises as to the order of importance when it comes to predicting 197hPa CAT? To test this, multiple linear regression models were set up and used to predict values over the north Atlantic and Europe.

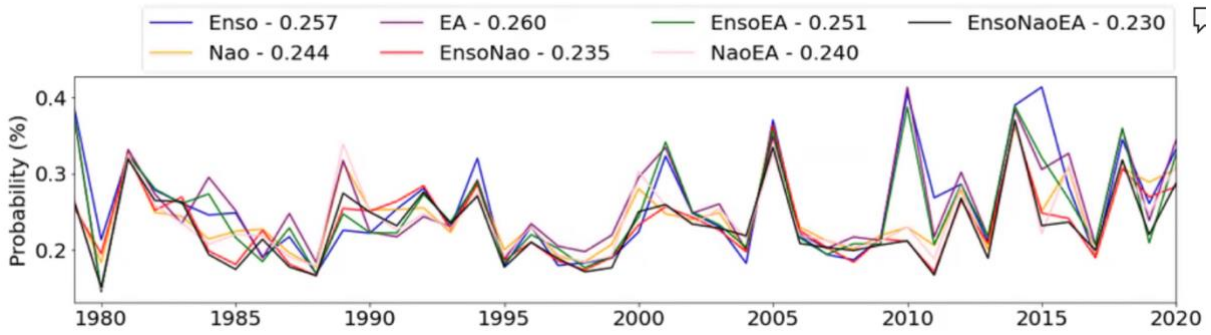


Figure 5.7: The sum of the difference between observed CAT probabilities and various linear regression models over a north Atlantic and European box (latitude 34N-60N and longitude –56W-22E). The legend indicates which of the 3 (ENSO, NAO and East Atlantic index) sources of variability have been included in the multiple linear regression model and the number to the right of the legend indicates the sum of the values over the 42 years.

The departure between these predictions and the observed values is shown in Figure 5.7. ENSO and the EA index by themselves as single linear regression did relatively poorly for predicting CAT values over the north Atlantic and Europe. The NAO by itself did best but was improved when combined with ENSO and this combination was superior to both the ENSO+EA and NAO+EA combinations over this region. The MLR model with all 3 combined together did best, but only marginally. The coefficients for the ENSO+NAO MLR model is shown in Figure 5.8.

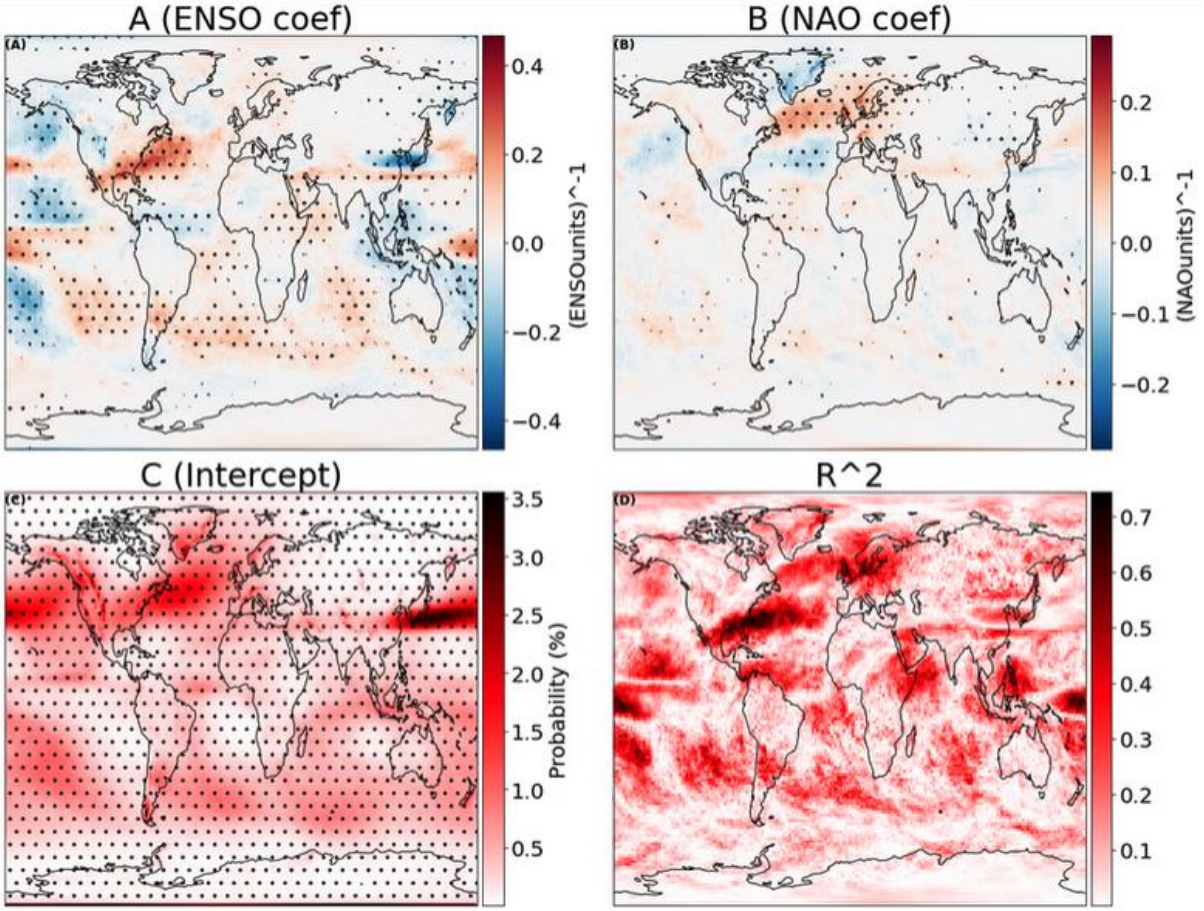


Figure 5.8: shows the results of a multiple linear regression of DJF ENSO (ONI 3.4, - 5N-5S 170W-120W), the DJF Hurrell Station-Based Seasonal NAO Index and diagnostic average CAT. As can be seen from panel A, the ENSO effect on CAT is fairly global in nature, while the NAO effect (panel B) is far more local to the north Atlantic. Using the ENSO and NAO coefficients for any given year, the 197 hPa CAT can be predicted as in Figure 5.9 row 3.

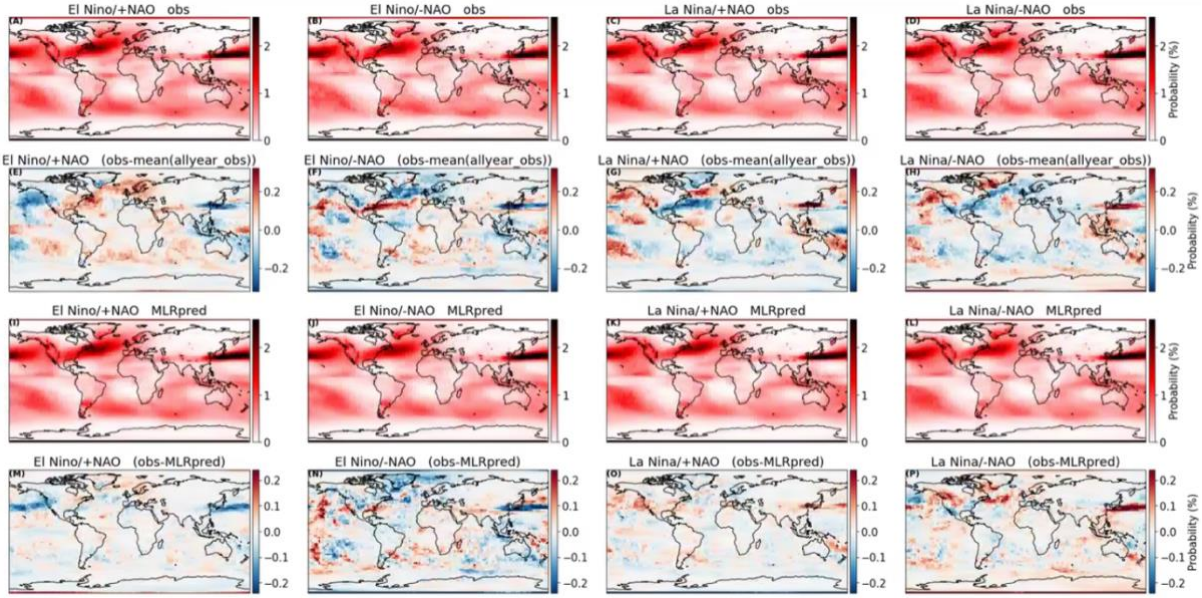


Figure 5.9: Row 1 shows the composite of CAT probability for all DJFs for years where both ENSO and the NAO were positive (panel A) ENSO was positive and the NAO negative (panel B), ENSO was negative and the NAO positive (panel C) and panel D for years where both ENSO and the NAO were negative. Panels e, f, g and h, while preserving the phase order of the row above, show the observed values of row 1 minus the average of all 42 observed DJFs of CAT probability (effectively showing how each of the 4 phases is different from average)
Panels i, j, k and l show the ENSO/NAO multi linear regression predicted DJF CAT probabilities. Panels m, n, o, and p show the observed values of row 1 minus the MLR predicted values for each phase combination (i.e. where the ENSO/NAO combined MLR fails to predict well).

The difference between the MLR predicted CAT and ERA5 observed CAT are perhaps hard to see by eye, hence row 4 shows the difference. The largest differences occur in the midlatitude western pacific with CAT being underpredicted by ~0.2% in the -ENSO/-NAO phase and overpredicted by ~0.2% in the +ENSO/-NAO phase. Similarly there is a region of underprediction/overprediction in the midlatitude north Atlantic in those same phases albeit with an error of closer to 0.1%. The 2nd row shows how each of the 4 phases is different from one another. Most strikingly the north Atlantic is completely different in an +ENSO/-NAO phase relative to a -ENSO/+NAO phase where the northern north Atlantic contains less CAT and the southern half more and vice-versa in the ENSO-/+NAO phase. This indicates that accurate prediction of CAT in the north Atlantic requires both knowledge of the NAO and ENSO as both contribute to the pattern.

5.4 Summary and Discussion

Analysing 42 years of DJF global cruise altitude (197hPa) ERA5 data, differing spatial patterns of CAT at 197hPa occurred during the El Niño, La Niña and Neutral phase of ENSO.

In particular, the region over the US and North Atlantic, a busy flight path, showed the biggest difference in CAT occurrence between an El Niño and a La Niña phase of ENSO (~0.4% higher) equivalent to 35 extra hours of MOG CAT per year in El Niño as opposed to La Niña).

The aforementioned Wolff and Sharman (2008) observation that incidences of CAT appeared to spike over the US during the strong 1998 El Niño appears to be consistent with this finding.

Japan and East Asia - also a busy flight area - showed the opposite pattern with up to a 0.45% difference equivalent to 39.5 hours more CAT in La Niña than El Niño or Neutral.

Europe, the third major area of aviation activity globally showed little effect of ENSO, which is perhaps unsurprising as Europe is the furthest from the source of ENSO in the Pacific Ocean.
As to a mechanistic understanding of why the US and East Asia are so affected, aside from the fact that they both border the Pacific, we can bring little light to bear.

We also aren't that clear on why the northern hemisphere is more affected than the southern, though this is likely due to reduced volumes of CAT in the southern hemisphere more generally (See Prosser et al., 2023).

The NAO also has a significant bearing on the volume of CAT at cruise altitude but only within the northern north Atlantic and Europe. In the northern north Atlantic, the probability of CAT is also ~0.4% higher in a +NAO than a -NAO (equivalent to 35 extra hours of MOG CAT per year) and this is particularly of consequence as it lines up with the main transatlantic flight routes between Europe and the US.

These finding are of consequence to the aviation industry as more fuel is required to either fly through, or around turbulent areas of the sky. If more/less CAT is encountered depending on the phase of ENSO, then differing amounts of fuel will be required and the industry may wish to adjust their purchasing as soon as the future state of ENSO is known to reasonable degree of certainty say 6 months in advance. The NAO is somewhat more difficult to forecast skilfully on such timescales, however limited sub seasonal 'forecasts of opportunity' may be identifiable in the coming years (Albers & Newman, 2021)

In order to ascertain the importance of ENSO and the NAO as key sources of variability impacting on CAT, an EOF analysis was performed. ENSO mapped onto EOF1 and the NAO mapped onto EOF5, EOF3 was potentially the East Atlantic Index, though this is by no means certain. EOF2 seemed to show a strong source of variability in the midlatitude norther Pacific, but this didn't match any known sources of variability wither spatially or temporally.

Because of the predictability of these sources of climate variability, a multiple linear regression was produced to predict the location of 197hPa CAT using values of ENSO and the NAO. The model appears

to work well in reconstructing the actual locations of CAT in most places (See Figure 5.9), although a remaining source of error occurs in the midlatitude northern Pacific. The aviation industry should be able to use this model to make short-term forecasts of CAT to fine-tune their requisioning of supplies (e.g. fuel) that are sensitive to CAT volume.

There are many ways this study could be expanded. In order to compute average climatologies we have employed a uniform weighting scheme across our 21 diagnostics that assumes no superior/inferior skill at diagnosing CAT. Other schemes could be employed and perhaps reveal different patterns.

CAT is not the only type of aviation affecting turbulence and an analysis that looks at how other forms (e.g.) CIT and MWT are affected by ENSO and the NAO would also be valuable to decision makers within the industry.

Chapter 6
Summary and conclusions

CAT is already a major concern for the aviation industry and with increasing levels of climate change, this concern is only likely to grow.

6.1 Main results of this thesis

1. The SOG CAT diagnostic skill of an ensemble of 21 CAT diagnostics have been evaluated. Some diagnostics perform better, others perform less well. Such an evaluation will be useful both for short term CAT forecasts and for longer term CAT-climate projections.

2. Global maps of 197hPa diagnosed-CAT have been created for the years 1979 and 2020 as well as maps of both relative and absolute change between these 2 years. The former maps are of interest in showing where the CAT exists globally/geographically and the latter maps indicate where the amount of diagnosed CAT has increased (or not) over the 40 years of the satellite era.

3. In the final section, the impact of ENSO and the NAO on 197 hPa diagnosed CAT was examined. ENSO has a global effect with the Pacific, East Asia, central/north America and the north Atlantic being particularly affected. The effect on diagnosed CAT of the NAO is more limited to the north Atlantic and Europe, changes which are very relevant to aviation due to the high volume of aviation traffic that fly through these regions.

6.2 Evaluating diagnostic skill

Every year passengers and crew get injured/hospitalised from encounters with CAT and being able to forecast this type accurately is of paramount importance.

Here based on 80 encounters with CAT severe enough that they made international news, the skill of an ensemble of 21 CAT diagnostics have been evaluated.

The 80-event average diagnostic skill ranged from around the 50% mark (no significant skill at all) for 6 diagnostics all the way to 69.4% for d12:The magnitude of horizontal temperature gradient). 6 other diagnostics were judged to have high skill. Prior to this evaluation, when making both regional and global projections of future CAT, the working assumption of Williams (2017) and Storer et al. (2017) was that

each of these diagnostics had equal skill. Since they do not, these projections have been modified to take into account two different skill based weighting schemes. For the North Atlantic projections of Williams 2017, both skill based weighting schemes slightly reduce the projected increase in CAT under a doubling of CO2 however for the global projections under an RCP8.5 scenario, the CAT projections are worse as a result of the skill weighting.

A key finding of this chapter was that modified versions of diagnostics 4, 19 and 20 (Frontogenesis function, the North Carolina State University Index and the negative absolute vorticity advection) where negative values were not excluded, performed significantly better at diagnosing SOG CAT than the original unmodified versions. This is useful information for CAT forecasters.

6.3 Changes in global CAT over the reanalysis period.

Although DJF and JJA northern hemisphere maps of 4 separate CAT diagnostics at tropopause altitude have already been produced by Jaegar and Sprenger in 2007, an annual diagnostic averaged CAT based on more recent reanalysis (ERA5) at cruise altitude (197hPa) have not. In the second results chapter, such maps for the years 1979 and 2020 have been produced. The maps show various interesting features about the global distribution of diagnosed CAT at cruise altitude. Most CAT is located around the jet streams of the midlatitudes, but is often heaviest over the oceanic western boundaries (e.g. over the Gulf stream and Kuroshio current). This is likely due to a combination of the land/sea temperature contrast contributing to vertical wind shear and the faster jet speeds over the ocean.

There is a certain amount of CAT diagnosed around the equator by some diagnostics, in particular the equatorial pacific.

Northern hemisphere midlatitude CAT appears to be noticeably stronger than in the Southern Hemisphere. In 1979, the region over the Kuroshio current was home to the greatest volume of CAT in the world and an amount far greater than that located over the Gulf Stream. By 2020 the CAT volume here had increased significantly to the point where it is broadly comparable.

If we look at MOG CAT, this region over the north Atlantic saw a statistically significant relative increase of 37% (equivalent to an extra 26 hours of CAT in the year 2020 relative to 1979) likely due to the 15% increase in vertical wind shear in the north Atlantic jet stream over the same period reported by Lee et al. (2019) For this region, 17 out of the 21 diagnostics saw a significant positive trend and only 4 did not. The 200 hPa atmosphere over continental USA are also a region of dramatic increase seeing 29% greater CAT over the 40 year satellite period (equivalent to an extra 19 hours of CAT in the year 2020 relative to 1979). For SOG these relative increases increase to 54% and 40% respectively. Both these regions are important areas of aviation traffic.
For the North Atlantic region, a seasonal breakdown was conducted and the most turbulent season was DJF, followed by SON, followed by JJA with MAM coming last with about a third less CAT than DJF.

Also within this region, 5 separate strengths of CAT were examined and while the total number of turbulent hours decreased as the strength threshold increased, the relative increase over the 40 years also increased with the strength threshold. For example MOG increased by 37% over the north Atlantic, but SOG increased by 54%. This is likely due to the rightmost extremes of a distribution increasing more in relative terms than a mid section would as the distribution as a whole moves further to the right (greater diagnostic values). This is bad news from the point of view of passengers and crew safety since SOG CAT is responsible for the majority of injuries and hospitalisations. Even if a double to tripling of midlatitude CAT occurs this century (as projected by Storer et al., 2017), passengers and crew may be spared the worst of it if the ability to forecast CAT improves substantially as time goes on. However, since this is not a given, the industry may wish to plan for a future in a more turbulent atmosphere.

Finally, the 37% increase in MOG CAT over the north Atlantic for a 30% increase in CO2 is slightly higher than the 89% increase in MOG CAT from a doubling of CO2 projected by Williams and Joshi (2013). This combined with the finding of Williams and Storer (2022) that climate models tend to underestimate CAT increases relative to reanalysis suggest that climate model simulations may underestimate future CAT increases.

Given the increases in CAT over the past 40 years and the projected increases, further research to improve the accuracy of forecasts will be critical to ensuring crew and passenger safety as the world continues to warm.

6.4 The impact of climate variability on diagnosed CAT

Work done by Cheyne (2020) indicated that ENSO had a large impact on global vertical wind shear, a key component (and diagnostic of CAT). To investigate whether other sources of variability impacted on CAT (and whether ENSO was the leading source of variability), an EOF was performed. ENSO was found to correspond to EOF1, the leading source of variability with the East Atlantic Oscillation possibly corresponding to EOF3. The NAO matched pretty robustly both spatially and temporally with EOF5, having more of a regional impact in the north Atlantic than elsewhere.

For ENSO, the US/Mexico and the North Atlantic regions all correlated positively with diagnosed CAT. East Asia correlated negatively. Other areas throughout the Pacific are also influenced by ENSO, though these are less relevant to aviation.

The NAO doesn't have such a global impact but does result in a tripole structure over the north Atlantic with the area from 30N - 47.5N as well as Greenland correlating negatively with the NAO and the region in-between (from 47.5N-60N) and Europe correlating positively.

Inputting both the phase of the NAO and ENSO into an MLR, CAT can be predicted to a fair degree of accuracy with the midlatitude Pacific being an exception. Given that the state of ENSO can be predicted some months ahead, this is likely to be of use to the aviation industry when planning their advanced requisition of fuel.

6.5 Limitations and further work

All three of these projects have limitations and possibilities for future work.

In an ideal world, we would want global maps of aviation affecting turbulence. The maps in this thesis on the other hand are of 21-diagnostic-averaged CAT. The CAT diagnostics used sometimes have the ability to diagnose other forms of aviation affecting turbulence (e.g. the Richardson number can diagnose CIT) however this ambiguity between turbulence categories (on which more research is probably warranted) not withstanding, other forms of turbulence e.g. CIT and MWT are not included in the analysis and were they included, the overall picture for both the distribution and trends in turbulence might change substantially. Aside from this, the work in chapter 3 demonstrates that not all turbulence diagnostics have equal skill at diagnosing injury causing turbulence, yet the global figures in chapter 4 implicitly assume – with their simple 21-diagnostic-average – that they do. As time goes on, more injury causing events will occur and these can be used to further refine the accuracy of diagnostic skill weightings, something which could well alter the spatial distribution and trends in injury causing CAT once applied to reanalysis and climate projections (e.g. Storer et al., 2017). In addition to injury causing events, more and more turbulence data (e.g. EDR and DEVG) from airlines (e.g. the IATA Turbulence Aware Platform; IATA, n.d.),is likely to become available and this data could also be used to evaluate the 21 diagnostics of Williams and Joshi (2013).

Only the ERA5 reanalysis was examined in this thesis but it would also be interesting to repeat the analysis done here in JRA-55, to see how sensitive the results are to the choice of reanalysis dataset, and ERA6 once it is released.

Finally, this research has raised many questions that ought to be looked into further. For example, understanding the mechanisms behind the questions of why is there substantially more diagnosed CAT in the NH midlatitudes than the SH, and why East Asia experiences less diagnosed CAT during an El-Nino when the US and Mexico experiences more? These questions and many more are just waiting to be explored.

Appendix A

Turbulence Equations

D01: Negative Richardson number

$$R_s T_m \left(\frac{\frac{T_l}{p_l^{.286}} - \frac{T_u}{p_u^{.286}}}{\frac{T_m}{p_m^{.286}}} \right) \left(\frac{\frac{dp}{p}}{(du)^2 + (dv)^2} \right) \tag{D1}$$

D02: Magnitude of vertical shear of horizontal wind

$$\frac{(du)^2 + (dv)^2)^{0.5}}{\frac{R_s}{g} T_m \left(\frac{dp}{p_m} \right)} \tag{D2}$$

D03: Colson-Panofsky index

$$\left(\frac{3600}{1852} \right)^2 \left((du)^2 + (dv)^2 \right) \left(\frac{1 - Ri}{0.5} \right) \tag{D3}$$

D04: Frontogenesis function

$$\frac{du}{d\theta} = \frac{du}{T_u \left(\frac{1E5}{p_u} \right)^{.286} - T_l \left(\frac{1E5}{p_l} \right)^{.286}}$$

$$\frac{dv}{d\theta} = \frac{dv}{T_u \left(\frac{1E5}{p_u} \right)^{.286} - T_l \left(\frac{1E5}{p_l} \right)^{.286}}$$

$$\frac{du}{d\theta} \left(\frac{du}{d\theta} \frac{du}{dx} + \frac{dv}{d\theta} \frac{du}{dy} \right) + \frac{dv}{d\theta} \left(\frac{du}{d\theta} \frac{dv}{dx} + \frac{dv}{d\theta} \frac{dv}{dy} \right) \tag{D4}$$

D05: Brown index

$$0.3\left(\frac{dv}{dx}-\frac{du}{dy}+f\right)^2 + \left(\left(\frac{dv}{dx}+\frac{du}{dy}\right)^2+\left(\frac{du}{dx}-\frac{dv}{dy}\right)^2\right)^{0.5} \quad \text{(D5)}$$

D06: Brown energy dissipation rate

$$\frac{1}{24}\left(0.3\left(\frac{dv}{dx}-\frac{du}{dy}+f\right)^2\right) + \frac{1}{24}\left(\left(\left(\frac{dv}{dx}+\frac{du}{dy}\right)^2+\left(\frac{du}{dx}-\frac{dv}{dy}\right)^2\right)^{0.5}\left(du^2+dv^2\right)\right) \quad \text{(D6)}$$

D07: Variant 1 of Ellrod's turbulence index

$$\left(\frac{\left(du^2+dv^2\right)^{0.5}}{\frac{R_s}{g}T\frac{dp}{p}}\right)\left(\left(\frac{dv}{dx}+\frac{du}{dy}\right)^2+\left(\frac{du}{dx}-\frac{dv}{dy}\right)^2\right)^{0.5} \quad \text{(D7)}$$

D08: Variant 2 of Ellrod's turbulence index

$$\left(\frac{\left(du^2+dv^2\right)^{0.5}}{\frac{R_s}{g}T\frac{dp}{p}}\right)\left(\left(\left(\frac{dv}{dx}+\frac{du}{dy}\right)^2+\left(\frac{du}{dx}-\frac{dv}{dy}\right)^2\right)^{0.5}-\frac{du}{dx}-\frac{dv}{dy}\right) \quad \text{(D8)}$$

D09: Flow deformation

$$\left(\left(\frac{dv}{dx}+\frac{du}{dy}\right)^2+\left(\frac{du}{dx}-\frac{dv}{dy}\right)^2\right)^{0.5} \quad \text{(D9)}$$

D10: Potential Vorticity

$$- g\zeta_a \frac{d\theta}{dp} \tag{D10}$$

D11: Vertical Relative Vorticity squared

$$\left(\frac{dv}{dx} - \frac{du}{dy}\right)^2 \tag{D11}$$

D12: Horizontal temperature Gradient

$$\left(\frac{dT^2}{dx} + \frac{dT^2}{dy}\right)^{0.5} \tag{D12}$$

D13: Wind speed

$$\left(u^2 + v^2\right)^{0.5} \tag{D13}$$

D14: Wind speed times directional shear

$$\left(u^2 + v^2\right)^{0.5} \left| \frac{\arctan\left(\frac{v_u}{u_u}\right) - \arctan\left(\frac{v_l}{u_l}\right)}{\frac{R_s}{g} T \frac{dp}{p}} \right| \tag{D14}$$

D15: Flow deformation times wind speed

$$\left(\left(\frac{dv}{dx} + \frac{du}{dy}\right)^2 + \left(\frac{du}{dx} - \frac{dv}{dy}\right)^2\right)^{0.5} \left(u^2 + v^2\right)^{0.5} \tag{D15}$$

D16: Flow deformation times vertical temperature gradient

$$\frac{|dt|}{\frac{R_s}{g} T \frac{dp}{p}} \left(\left(\frac{dv}{dx} + \frac{du}{dy}\right)^2 + \left(\frac{du}{dx} - \frac{dv}{dy}\right)^2\right)^{0.5} \tag{D16}$$

D17: Magnitude of residual of nonlinear balance equation

$$2\left|\frac{du}{dx}\frac{dv}{dy}-\frac{du}{dy}\frac{dv}{dx}\right| \quad \text{(D17)}$$

D18: Magnitude of horizontal divergence

$$\left|\frac{du}{dx}+\frac{dv}{dy}\right| \quad \text{(D18)}$$

D19: Version 1 of North Carolina State University index

$$\beta=\frac{2\Omega\cos(\phi)}{R_E}$$

$$\xi=\frac{dv}{dy}-\frac{du}{dy}$$

$$\frac{dQ}{dx}=\frac{d\xi}{dx}$$

$$\frac{dQ}{dy}=\frac{d\xi}{dy}$$

$$\frac{u\frac{du}{dx}+v\frac{dv}{dy}}{Ri\left(\frac{dQ}{dx}^2+\frac{dQ}{dy}^2\right)^{0.5}} \quad \text{(D19)}$$

D20: Negative absolute vorticity advection

$$-u\frac{dQ}{dx}-v\left(\beta+\frac{dQ}{dy}\right) \quad \text{(D20)}$$

D21: Magnitude of relative vorticity advection

$$\left|u\frac{dQ}{dx}+v\frac{dQ}{dy}\right| \quad \text{(D21)}$$

University of Reading

Appendix B

Table 5: Event time and space coordinates in ERA5 and source of information used to triangulate location. If the websites FlightAware (n.d.) (FA) or FlightRadar24 (n.d.) (FR24) were used they are noted under 'Other resources'. For the NTSB events (Those with an ID beginning N) The URL begins with "https://app.ntsb.gov/pdfgenerator/ReportGeneratorFile.ashx?EventID=" and ends with the source given in the table.

OldID	NewID	Date/time	Latitude	Longitude	hPa	Source	Other resources ?
1	A01	20/08/2019 10:00	8.25±0.0	38.5±0.0	227	http://avherald.com/h?article=4cbd6d41	FR24
5	A02	11/07/2019 15:00	11.25±0.0	-168.25±0.0	249	http://avherald.com/h?article=4ca4118b	FA/FR24
18	A04	22/11/2017 14:00	32.5±0.0	133.75±0.0	285	http://avherald.com/h?article=4b17dc45	FR24
35	A06	27/09/2018 08:00	52±0.0	-18.5±0.0	249	http://avherald.com/h?article=4be6a797	FA
39	A07	24/12/2015 07:00	36±0.0	-81.75±0.0	511	http://avherald.com/h?article=4bc4e625	FA
42	A08	30/07/2018 13:00	42.25±0.0	23.25±0.0	238	http://avherald.com/h?article=4bbcac74	FR24
69	A09	24/10/2017 10:00	2.5±0.0	99.75±0.0	385	http://avherald.com/h?article=4b034a60	FR24

70	A10	02/01/2018 07:00	-7±0.0	112.25±0.0	401	http://avherald.com/h?article=4b31e4f1	FR24
82	A11	16/07/2014 00:00	4.5±0.0	102±0.0	217	http://avherald.com/h?article=47763054	—
93	A12	01/05/2017 00:00	17.25±0.0	96.5±0.0	238	http://avherald.com/h?article=4a861a24	—
114	A14	20/02/2010 00:00	61.5±0.0	-147.75±0.0	285	http://avherald.com/h?article=4278f9a1	FA
127	A15	22/11/2016 22:00	51±0.0	-127±0.0	453	http://avherald.com/h?article=4a18df78	FA
137	A16	31/08/2016 04:00	51±0.0	-14±0.0	249	http://avherald.com/h?article=49d57cf9	—
201	A17	29/12/2016 14:00	55.5±0.0	135±2.5	285	http://avherald.com/h?article=4a3b6873	FA
202	A18	05/05/2016 18:00	25.75±0.0	-75±0.0	227	http://avherald.com/h?article=497e1876	FA
203	A19	19/02/2013 14:00	6.5±0.0	-151.75±0.0	217	http://avherald.com/h?article=45e08ad5	FA
204	A20	03/02/2012 00:00	61.25±0.25	-148.5±0.25	614	http://avherald.com/h?article=44a7490d	FA
207	A22	19/03/2013 21:00	39±0.0	-76.75±0.0	531	http://avherald.com/h?article=4608f3ce	FA
208	A23	08/05/2009 23:00	43.75±0.25	-71.75±0.25	217	http://avherald.com/h?article=43e4e36b	FA
209	A24	15/07/2010 02:00	17.5±0.0	143.75±0.0	227	http://avherald.com/h?article=431030b3&opt=0	FA
213	A26	18/11/2009 02:00	39.75±0.0	-80±0.0	285	http://avherald.com/h?article=4282bee7&opt=0	FA
214	A27	24/02/2008 21:00	36±0.0	-115.5±0.0	634	http://avherald.com/h?article=402a4c28	FA
215	A28	10/07/2009 20:00	27.25±0.0	-81.5±0.0	634	http://avherald.com/h?article=41ca88cd	FA
216	A29	08/02/2009 16:00	20±0.0	-69.75±0.0	207	http://avherald.com/h?article=41557486	FA
217	A30	07/01/2009	44.25±0.0	-77.75±0.0	418	no longer online	FA

		18:00					
218	A31	30/06/2008 19:00	28±0.0	-87.5±0.0	285	no longer online	FA
1002	N01	04/08/2016 00:00	19±0.0	-81±0.0	511	20160805X51304&AKey=1&RType=Final&IType=CA	—
1004	N03	28/01/2016 23:00	28.5±0.0	-79.5±0.0	273	20160129X24835&AKey=1&RType=Final&IType=CA	—
1005	N04	26/12/2015 21:00	32.75±0.0	-116.5±0.0	593	20160119X15230&AKey=1&RType=Final&IType=CA	—
1006	N05	23/12/2015 06:00	35.25±0.0	-80.75±0.0	511	20160119X20258&AKey=1&RType=Final&IType=CA	—
1008	N06	23/11/2015 19:00	40±0.0	-75.25±0.0	325	20151124X65446&AKey=1&RType=Final&IType=CA	—
1009	N07	24/10/2015 15:00	38±0.0	-98.5±0.0	339	20151026X81321&AKey=1&RType=Final&IType=CA	—
1011	N09	24/11/2014 19:00	30.75±0.0	-83.25±0.0	249	20141211X62917&AKey=1&RType=Final&IType=CA	—
1013	N11	18/10/2013 10:00	29.75±0.0	-95.5±0.0	227	20131121X13345&AKey=1&RType=Final&IType=CA	—
1014	N12	19/02/2013 15:00	20.25±0.0	-139.25±0.0	197	20131121X13345&AKey=1&RType=Final&IType=CA	—
1015	N13	20/03/2012 14:00	36.5±0.0	-0.5±0.0	227	20120322X25316&AKey=1&RType=Final&IType=LA	—
1016	N14	03/03/2012 22:00	40.25±0.0	-77±0.0	531	20120410X01241&AKey=1&RType=Final&IType=CA	—
1017	N15	02/02/2012 19:00	61.25±0.0	-150±0.0	593	20120203X61713&AKey=1&RType=Final&IType=CA	—
1018	N16	29/09/2011 20:00	25.75±0.0	-80.25±0.0	551	20110930X20857&AKey=1&RType=Final&IType=CA	—
1019	N17	16/05/2011 23:00	39.25±0.0	-76.5±0.0	593	20110615X10113&AKey=1&RType=Final&IType=CA	—
1020	N18	12/03/2010 11:00	-14.25±0.0	-170.75±0.0	511	20101203X90900&AKey=1&RType=Final&IType=LA	—
1023	N20	03/04/2010 12:00	41.5±0.0	-93.25±0.0	285	20100405X40816&AKey=1&RType=Final&IType=LA	—

1024	N21	20/02/2010 00:00	61.25±0.0	-150±0.0	298	20100222X84927&AKey=1&RType=Final&IType=CA	—
1027	N24	19/05/2008 13:00	15.75±0.0	152.25±0.0	227	20080723X01111&AKey=1&RType=Final&IType=FA	—
1029	N26	08/07/2007 20:00	34.75±0.0	-95.75±0.0	298	20071130X01878&AKey=1&RType=Final&IType=LA	—
1030	N27	27/06/2007 11:00	11±0.0	-157±0.0	261	20070720X00966&AKey=1&RType=Final&IType=LA	—
1032-5.1	N29	15/04/2006 20:00	39±0.0	-78.25±0.0	369	20060425X00486&AKey=1&RType=Final&IType=LA	—
1033-5.1	N30	11/07/2005 23:00	35.5±0.0	-80.75±0.0	217	20051028X01749&AKey=1&RType=Final&IType=MA	—
1034-5.1	N31	29/08/2004 18:00	43.25±0.0	-77±0.0	311	20050323X00350&AKey=1&RType=Final&IType=LA	—
1036-5.1	N32	22/07/2003 18:00	41±0.0	-81.75±0.0	311	20030902X01433&AKey=1&RType=Final&IType=CA	—
1037-5.1	N33	25/04/2003 03:00	7.25±0.0	-80±0.0	435	20030514X00654&AKey=1&RType=Final&IType=LA	—
1038-5.1	N34	06/04/2003 16:00	12.5±0.0	-96±0.0	325	20030425X00566&AKey=1&RType=Final&IType=LA	—
1039-5.1	N35	17/11/2002 23:00	37.75±0.0	-77.75±0.0	511	20021209X05575&AKey=1&RType=Final&IType=IA	—
1040-5.1	N36	01/05/2002 06:00	-24.25±0.0	174.25±0.0	285	20020506X00632&AKey=1&RType=Final&IType=LA	—
1041-5.1	N37	22/04/2002 07:00	19.75±0.0	-66.25±0.0	227	20020429X00592&AKey=1&RType=Final&IType=LA	—
1043-5.1	N38	12/12/2000 20:00	56±0.0	-160±0.0	285	20010110X00085&AKey=1&RType=Final&IType=IA	—
1048	N39	08/07/1999 18:00	30.25±0.0	-6.75±0.0	311	20001212X19286&AKey=1&RType=Final&IType=FA	—
1050	N40	20/01/1999 13:00	33±0.0	155±0.0	261	20001204X00114&AKey=1&RType=Final&IType=IA	—
1035-5.1	N42	25/07/2004 05:00	19±0.0	-75.25±0.0	217	20060111X00048&AKey=1&RType=Final&IType=MA	—
1047	N45	30/09/1999	41±0.0	-74.75±0.0	285	20001212X19871&AKey=1&RType=Final&IType=LA	—

		17:00					
1049	N46	09/02/1999 00:00	35.75±0.0	-78.75±0.0	511	20001205X00124&AKey=1&RType=Final&IType=LA	—
1055	N47	18/04/1998 19:00	43.25±0.5	-67.75±0.0	238	20001211X09935&AKey=1&RType=Final&IType=FA	—
1056	N48	04/03/1998 21:00	41±0.25	-119±0.5	238	20001211X09708&AKey=1&RType=Final&IType=LA	—
1057	N49	28/12/1997 14:00	30.25±2.75	157±1.25	285	20001208X09291&AKey=1&RType=Final&IType=MA	—
1058	N50	26/09/1997 20:00	40.75±0.0	-73.25±0.0	354	20001208X08916&AKey=1&RType=Final&IType=LA	—
1059	N51	07/09/1997 02:00	41.25±0.0	-89.25±0.0	694	20001208X08808&AKey=1&RType=Final&IType=LA	—
1062	N53	26/03/1997 19:00	36.25±0.0	-84.5±0.5	339	20001208X07493&AKey=1&RType=Final&IType=LA	—
1063	N54	02/03/1997 12:00	40±0.0	-87.75±0.5	418	20001208X07501&AKey=1&RType=Final&IType=LA	—
1065	N55	28/01/1997 15:00	37.25±0.0	-89.5±0.0	435	20001208X07272&AKey=1&RType=Final&IType=LA	—
1069	N57	20/11/1996 02:00	37.25±0.0	-118.5±0.0	311	20001208X07076&AKey=1&RType=Final&IType=LA	—
1070	N58	29/08/1996 20:00	34.25±0.0	-85.25±0.25	238	20001208X06588&AKey=1&RType=Final&IType=LA	—
1074	N61	01/11/1995 15:00	30.5±2.5	-173±0.0	238	20001207X04870&AKey=1&RType=Final&IType=LA	—
1076	N62	17/10/1995 10:00	40±0.0	152±0.0	261	20001207X04726&AKey=1&RType=Final&IType=LA	—
1078	N63	20/04/1995 03:00	29.5±0.0	-99.5±0.0	369	20001207X03300&AKey=1&RType=Final&IType=LA	—
1080	N64	05/07/1994 12:00	30.75±0.0	-83.25±0.0	511	20001206X01808&AKey=1&RType=Final&IType=LA	—
1082	N65	12/02/1994 14:00	-10±0.0	-157±0.0	238	20001206X00784&AKey=1&RType=Final&IType=LA	—
1093	N67	23/12/1988 11:00	38.5±0.0	-173.25±0.0	207	20001213X27446&AKey=1&RType=Final&IType=LA	—

1100	N68	04/03/1986 05:00	21.25±0.0	-149±0.0	261	20001213X32999&AKey=1&RType=Final&IType=FA	—
1104	N70	27/05/1982 15:00	41.5±0.0	-85.75±0.0	238	20020917X03237&AKey=1&RType=Final&IType=IA	—
1108	K01	13/02/2013 03:00	38.5±0.0	121.25±0.0	385	Lee and Chun (2018)	—

Appendix C: Addendum to Chapter 4 – Revisiting the results of chapter 3 in light of chapter 4.

Figure 4.1 shows that most ERA5 diagnosed CAT occurs over the western ocean basins. However figures 3.1 and 3.2 show that a large fraction of the 80 events occurred over land. In addition the Atlantic climatology (shown in figure 3.1), although over the ocean, tends to avoid the most turbulent locations at the North American western boundary.

So the question arises: Is the negative skew in figure 3.5 evidence of diagnostic and methodological skill or merely an artifact of the location of the 80 events biasing the results in a positive direction?

To investigate this, an initial set of sensitivity experiments were carried out. Here the temporal (both date and time) coordinates and altitudes of the 80 events were kept the same as in chapter 3, but the latitude/longitude geographical coordinates were chosen at random between 60N-60S (same latitudinal extent as the Atlantic climatology) in figure 4.14 and 90N-90S in figure 4.15. The results were as follows:

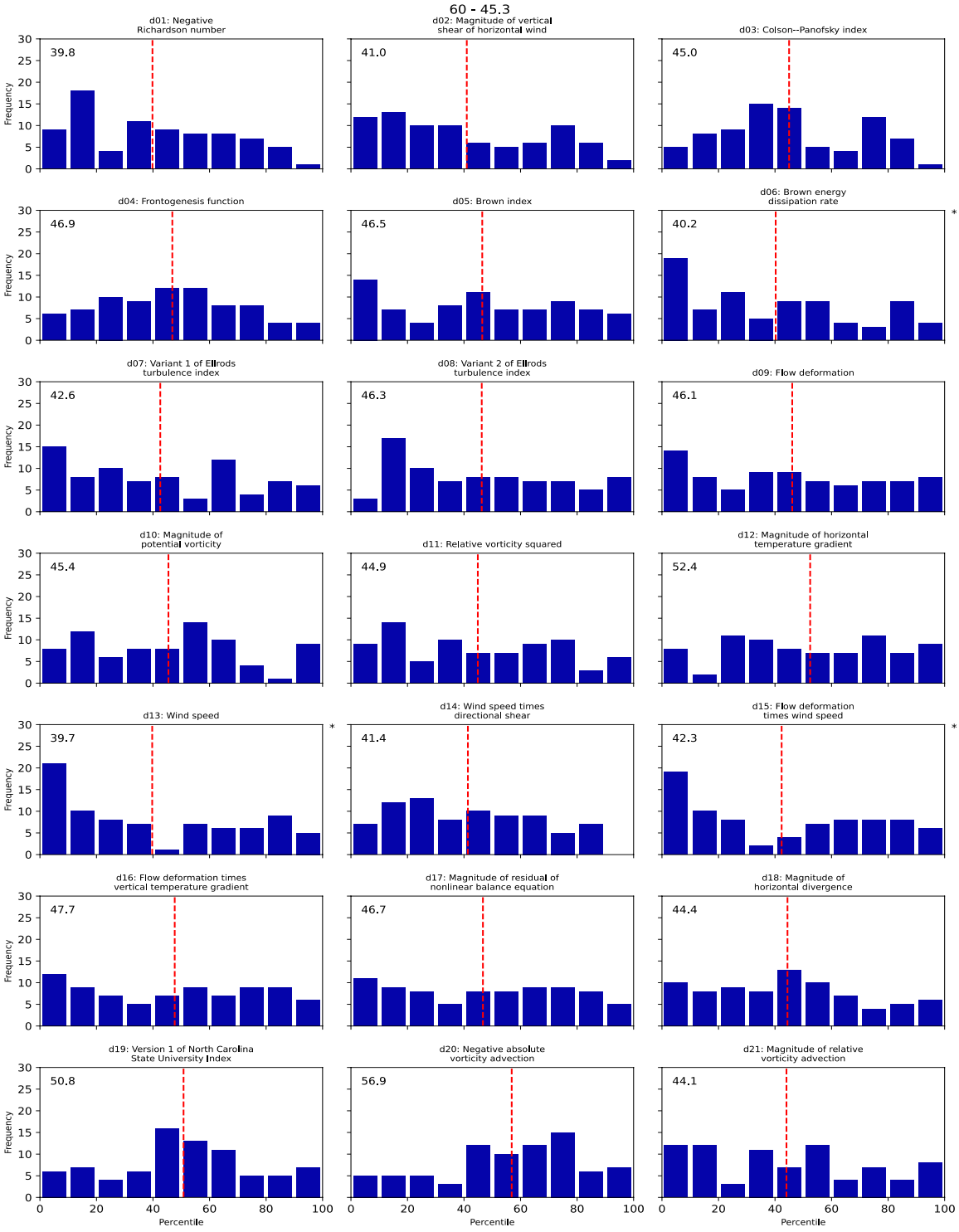


Figure 4.14: Event percentile distributions (n=80 events) for each of the 21 CAT diagnostics examined in this study. Temporal and altitude coordinates have been kept constant but geographical coordinates have been varied between 60N-60S. The vertical dashed lines show the 80-event mean (green if ≥66.7%, red if <60%, grey if in between these values) and its percentile value is shown in the top left of its respective subplot.

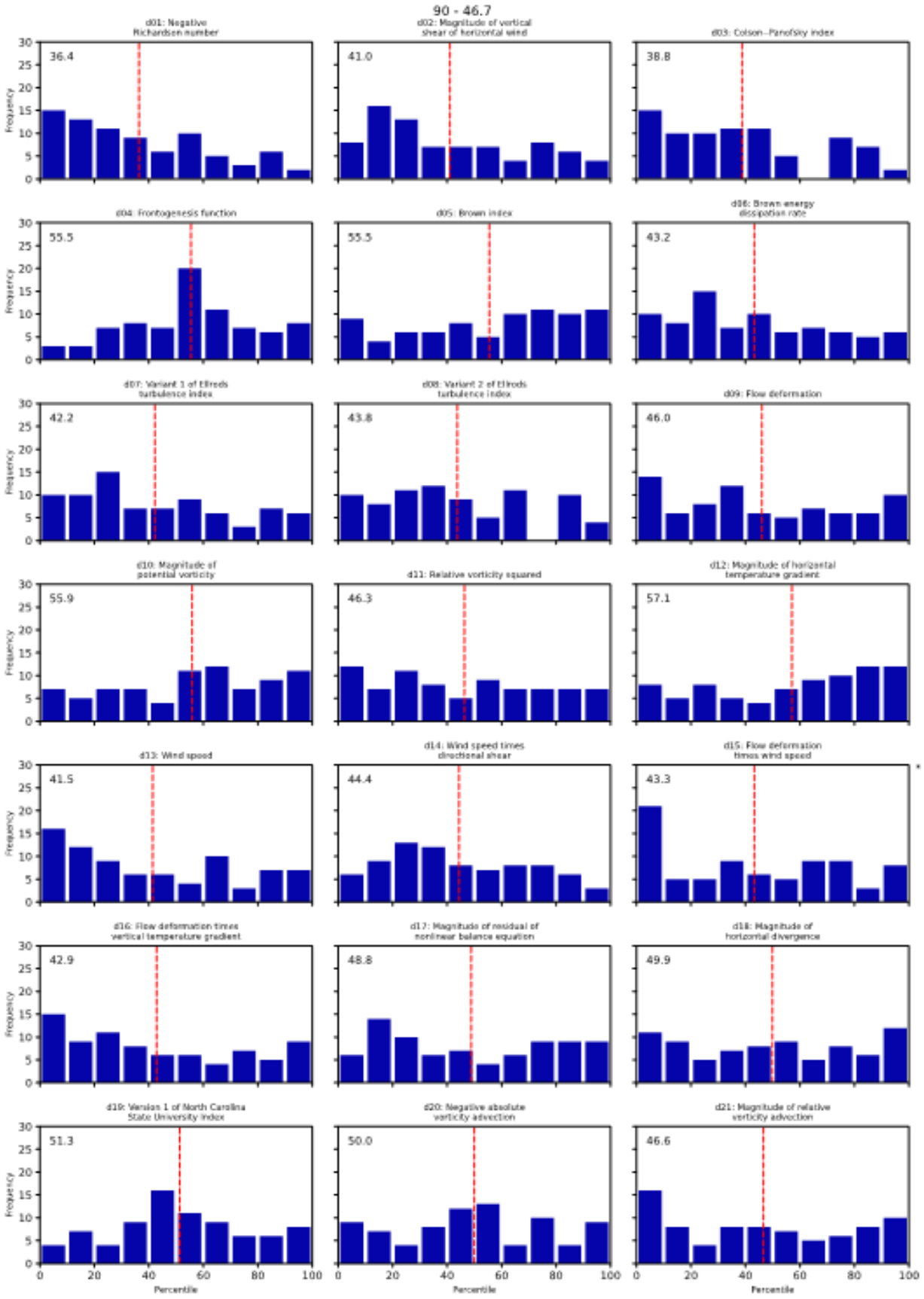

90 - 46.7
d01: Negative Richardson number
36.4
d02: Magnitude of vertical shear of horizontal wind
41.0
d03: Colson–Panofsky index
38.8
d04: Frontogenesis function
55.5
d05: Brown index
55.5
d06: Brown energy dissipation rate
43.2
d07: Variant 1 of Ellrods turbulence index
42.2
d08: Variant 2 of Ellrods turbulence index
43.8
d09: Flow deformation
46.0
d10: Magnitude of potential vorticity
55.9
d11: Relative vorticity squared
46.3
d12: Magnitude of horizontal temperature gradient
57.1
d13: Wind speed
41.5
d14: Wind speed times directional shear
44.4
d15: Flow deformation times wind speed
43.3
d16: Flow deformation times vertical temperature gradient
42.9
d17: Magnitude of residual of nonlinear balance equation
48.8
d18: Magnitude of horizontal divergence
49.9
d19: Version 1 of North Carolina State University index
51.3
d20: Negative absolute vorticity advection
50.0
d21: Magnitude of relative vorticity advection
46.6
Frequency
Percentile

Figure 4.15: Event percentile distributions (n=80 events) for each of the 21 CAT diagnostics examined in this study. Temporal and altitude coordinates have been kept constant but geographical coordinates have been varied between 90N-90S. The vertical dashed lines show the 80-event mean (green if ≥66.7%, red if <60%, grey if in between these values) and its percentile value is shown in the top left of its respective subplot.

The distributions in figures 4.14 and 4.15 look fairly flat with only 3 diagnostics diverging from uniform (according to a Chi-squared statistical test) in the former and one in the latter. The 21 diagnostic averages were 45.3 and 46.7 respectively which suggests that the choice of the Atlantic climatology is biasing the results low (more turbulence in the Atlantic than the global average), a result not unsurprising given a higher prevalence of turbulence over the ocean more generally in Figure 4.1.

But these results don't rule out the possibility that the negative skew in figure 3.5 occurred due to the 80 events just so happening to be concentrated in regions of climatologically high turbulence.

Therefore a further set of sensitivity experiments were carried out. Here, the latitude/longitude geographical coordinates and altitudes were kept the same as in chapter 3, however the temporal coordinates were varied. For figure 4.16, the times were chosen at random but for figure 4.17, a year was added to 40 of the events and a year subtracted from the remaining 40. This serves to change the times, but keep the seasonality constant.

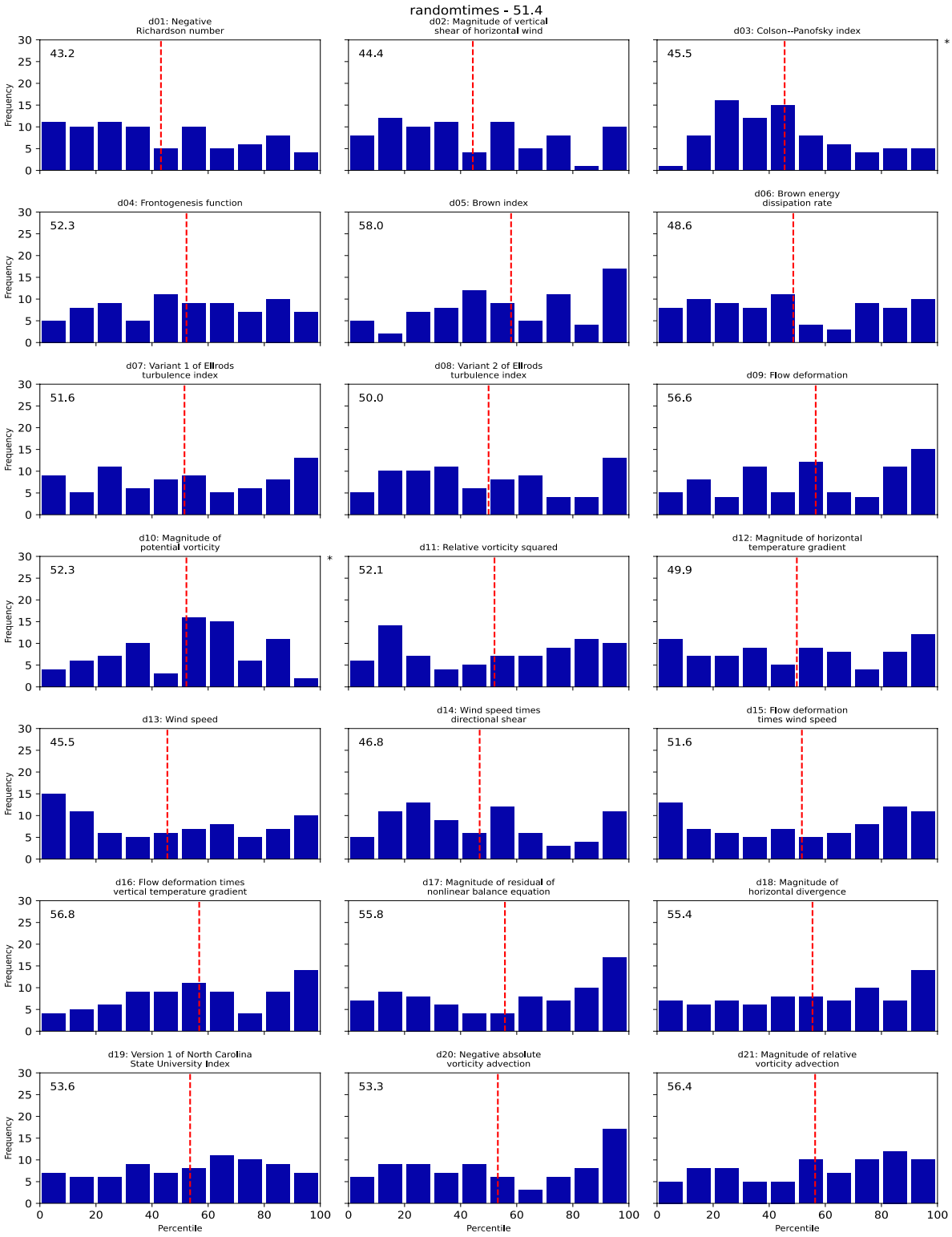


Figure 4.16: Event percentile distributions (n=80 events) for each of the 21 CAT diagnostics examined in this study. Geographical and altitude coordinates have been kept constant but temporal coordinates have been selected at random. The vertical dashed lines show the 80-event mean (green if ≥66.7%, red if <60%, grey if in between these values) and its percentile value is shown in the top left of its respective subplot.

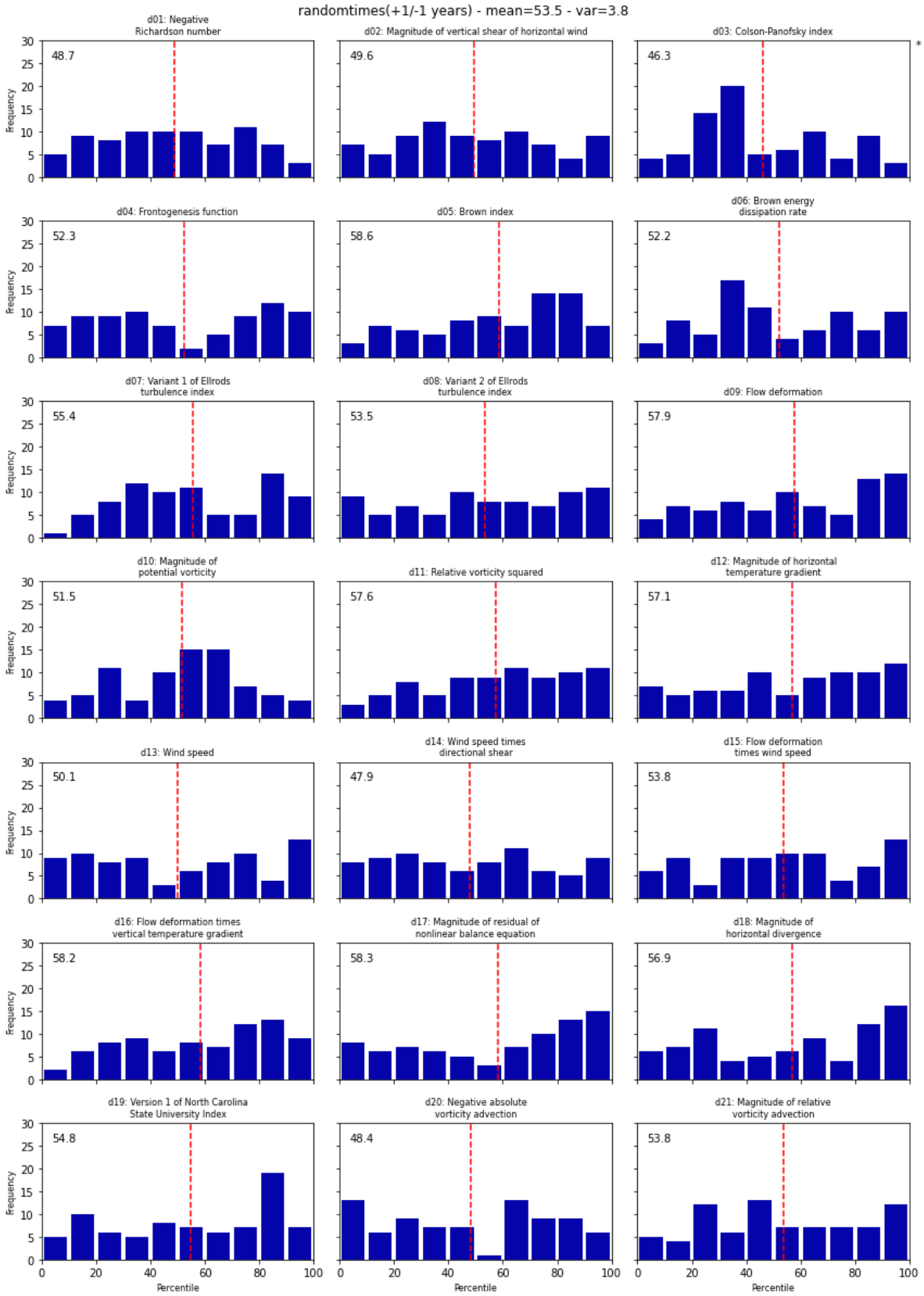

randomtimes(+1/-1 years) - mean=53.5 - var=3.8
d01: Negative Richardson number
48.7
d02: Magnitude of vertical shear of horizontal wind
49.6
d03: Colson-Panofsky index
46.3
d04: Frontogenesis function
52.3
d05: Brown index
58.6
d06: Brown energy dissipation rate
52.2
d07: Variant 1 of Ellrods turbulence index
55.4
d08: Variant 2 of Ellrods turbulence index
53.5
d09: Flow deformation
57.9
d10: Magnitude of potential vorticity
51.5
d11: Relative vorticity squared
57.6
d12: Magnitude of horizontal temperature gradient
57.1
d13: Wind speed
50.1
d14: Wind speed times directional shear
47.9
d15: Flow deformation times wind speed
53.8
d16: Flow deformation times vertical temperature gradient
58.2
d17: Magnitude of residual of nonlinear balance equation
58.3
d18: Magnitude of horizontal divergence
56.9
d19: Version 1 of North Carolina State University Index
54.8
d20: Negative absolute vorticity advection
48.4
d21: Magnitude of relative vorticity advection
53.8
Frequency
Percentile

Figure 4.17: Event percentile distributions (n=80 events) for each of the 21 CAT diagnostics examined in this study. Geographical and altitude coordinates have been kept constant but temporal coordinates have been altered by +/- one year. The vertical dashed lines show the 80-event mean (green if ≥66.7%, red if <60%, grey if in between these values) and its percentile value is shown in the top left of its respective subplot.

In summary, figure 3.5 showed a substantial negative skew where 13 of the diagnostics demonstrated a statistically significant departure from a uniform distribution and a diagnostic average of 59.6%.

The average of the diagnostic averages of the 2 random locations experiments (Figures 4.14 and 4.15) is 46% and the average of the diagnostic averages of the 2 random times experiments (Figures 4.16 and 4.17) is 52.45%.

Conclusion

52.45% is almost exactly half way between 46% and 59.6% which suggests that although the location of the 80 events may have contributed half to the negative skew in figure 3.5, the skill of the diagnostics and the method employed is likely to account for the other half.

While the random times experiments did present higher diagnostic averages than the random location ones, not even they exhibit negative skew and of the 3 diagnostics (between the 2 of them) which deviated statistically from a flat distribution, none of these deviated because of a negative skew.

References

Albers, J.R. and Newman, M., 2021. Subseasonal predictability of the north Atlantic oscillation. *Environmental Research Letters*, *16*(4), p.044024.

Archer, C.L. and Caldeira, K., 2008. Historical trends in the jet streams. *Geophysical Research Letters*, *35*(8).

AvHerald (n.d.), The aviation herald.
URL: https://avherald.com/ (Accessed 30 March 2024)

Barnes, E.A. and Polvani, L., 2013. Response of the midlatitude jets, and of their variability, to increased greenhouse gases in the CMIP5 models. *Journal of Climate*, *26*(18), pp.7117-7135.

Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D. and Soci, C., 2021. The ERA5 global reanalysis: Preliminary extension to 1950. *Quarterly Journal of the Royal Meteorological Society*, *147*(741), pp.4186-4227.

Bluestein, H.B., 1992. *Synoptic-dynamic Meteorology in Midlatitudes: Observations and theory of weather systems* (Vol. 2). Taylor & Francis.

Bluestein, H.B., 1993. Observations and theory of weather systems. *Synoptic-dynamic meteorology in midlatitudes*, p.594.

Brown, R., 1973. New indices to locate clear-air turbulence. *Meteor. Mag*, *102*, pp.347-361.

Buldovskii, G.S., Bortnikov, S.A. and RUBINSHTEIN, M., 1976. Forecasting of heavy turbulence zones in the upper troposphere. *Meteorologiia i Gidrologiia*, pp.9-18.

Cheyne, R., 2020. Does El Niño-Southern Oscillation affect global clear-air turbulence? (Unpublished bachelor dissertation). University of Reading, Reading, UK.

Coffel, E. and Horton, R., 2015. Climate change and the impact of extreme temperatures on aviation. *Weather, Climate, and Society*, *7*(1), pp.94-102.

Colson, D. and Panofsky, H.A., 1965. An index of clear air turbulence. *Quarterly Journal of the Royal Meteorological Society*, *91*(390), pp.507-513.

Coumou, D., Lehmann, J. and Beckmann, J., 2015. The weakening summer circulation in the Northern Hemisphere mid-latitudes. *Science*, *348*(6232), pp.324-327.

Delcambre, S.C., Lorenz, D.J., Vimont, D.J. and Martin, J.E., 2013. Diagnosing Northern Hemisphere jet portrayal in 17 CMIP3 global climate models: Twenty-first-century projections. *Journal of Climate*, *26*(14), pp.4930-4946.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, D.P. and Bechtold, P., 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, *137*(656), pp.553-597.

Demoz, B.B., Starr, D.O.C., Chan, K.R. and Bowen, S.W., 1998. Wavelet analysis of dynamical processes in cirrus. *Geophysical research letters*, *25*(9), pp.1347-1350.

Dong, B., Sutton, R.T., Shaffrey, L. and Harvey, B., 2022. Recent decadal weakening of the summer Eurasian westerly jet attributable to anthropogenic aerosol emissions. *Nature Communications*, *13*(1), p.1148.

Dutton, J.A. and Panofsky, H.A., 1970. Clear Air Turbulence: A Mystery May Be Unfolding: High altitude turbulence poses serious problems for aviation and atmospheric science. *Science*, *167*(3920), pp.937-944.

Dutton, M.J.O., 1980. Probability forecasts of clear-air turbulence based on numerical model output.

Eichenbaum, H., 2003. Historical overview of turbulence accidents and case study analysis. *MCR Federal Inc. Rep. Br-M021/080–1*, *82*.

Eick, D., 2014, September. Turbulence related accidents & incidents. In *Presentation at NCAR turbulence impact mitigation workshop* (Vol. 2, pp. 3-4).

Ellrod, G.P. and Knapp, D.I., 1992. An objective clear-air turbulence forecasting technique: Verification and operational use. *Weather and Forecasting*, *7*(1), pp.150-165.

Ellrod, G.P., Lester, P.F. and Ehernberger, L.J., 2003. Clear-air turbulence. Encyclopedia of Atmospheric Sciences, JR Holton et al., Eds.

Elvidge, A.D., Vosper, S.B., Wells, H., Cheung, J.C., Derbyshire, S.H. and Turp, D., 2017. Moving towards a wave-resolved approach to forecasting mountain wave induced clear air turbulence. *Meteorological Applications*, *24*(3), pp.540-550.

Endlich, R.M., 1964. The mesoscale structure of some regions of clear-air turbulence. *Journal of Applied Meteorology and Climatology*, *3*(3), pp.261-276.

ERA5 (n.d.), ECMWF Reanalysis v5 (ERA5).
URL: https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5 (Accessed, 30 March 2024)

FAA, 2006: Preventing injuries caused by turbulence. Advisory Circular 120-88A, Federal Aviation Administration, Washington, DC.

Fahrmeir, L., Kneib, T., Lang, S. and Marx, B.D., 2022. Regression models. In *Regression: Models, methods and applications* (pp. 23-84). Berlin, Heidelberg: Springer Berlin Heidelberg.

FlightAware (n.d.), 'Flight aware'.
URL: https://uk.flightaware.com/

FlightRadar24 (n.d.), 'Flight radar 24'.
URL: https://www.flightradar24.com/

Fovell, R., Durran, D. and Holton, J.R., 1992. Numerical simulations of convectively generated stratospheric gravity waves. *Journal of Atmospheric Sciences*, *49*(16), pp.1427-1442.

Gidel, L.T. and Shapiro, M.A., 1979. The role of clear air turbulence in the production of potential vorticity in the vicinity of upper tropospheric jet stream-frontal systems. *Journal of the Atmospheric Sciences*, *36*(11), pp.2125-2138.

Gill, P.G. and Stirling, A.J., 2013. Including convection in global turbulence forecasts. *Meteorological Applications*, *20*(1), pp.107-114.

Gill, P.G., 2016. Aviation turbulence forecast verification. *Aviation turbulence: Processes, detection, prediction*, pp.261-283.

Graham, R.M., Hudson, S.R. and Maturilli, M., 2019. Improved performance of ERA5 in Arctic gateway relative to four global atmospheric reanalyses. *Geophysical Research Letters*, *46*(11), pp.6138-6147.

Gratton, G., Padhra, A., Rapsomanikis, S. and Williams, P.D., 2020. The impacts of climate change on Greek airports. *Climatic Change*, *160*, pp.219-231.

Guest, F.M., Reeder, M.J., Marks, C.J. and Karoly, D.J., 2000. Inertia–gravity waves observed in the lower stratosphere over Macquarie Island. *Journal of the atmospheric sciences*, *57*(5), pp.737-752.

Gultepe, I. and Starr, D.O.C., 1995. Dynamical structure and turbulence in cirrus clouds: Aircraft observations during FIRE. *Journal of Atmospheric Sciences*, *52*(23), pp.4159-4182.

Gultepe, I., Sharman, R., Williams, P.D., Zhou, B., Ellrod, G., Minnis, P., Trier, S., Griffin, S., Yum, S.S., Gharabaghi, B. and Feltz, W., 2019. A review of high impact weather for aviation meteorology. *Pure and applied geophysics*, *176*, pp.1869-1921.

Haman, K.E., 2011, December. On self-maintenance of clear-air turbulence. In *Journal of Physics: Conference Series* (Vol. 318, No. 7, p. 072009). IOP Publishing.

Hannachi, A., 2004. A primer for EOF analysis of climate data. *Department of Meteorology, University of Reading*, *1*(29), pp.3-3.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D. and Simmons, A., 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), pp.1999-2049.

Hurrell, James &, Phillips, Adam & National Center for Atmospheric Research Staff (Eds). Last modified 2023-07-10 "The Climate Data Guide: Hurrell North Atlantic Oscillation (NAO) Index (station-based)." Retrieved from https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-station-based on 2024-05-04.

Irvine, E.A., Hoskins, B.J., Shine, K.P., Lunnon, R.W. and Froemming, C., 2013. Characterizing North Atlantic weather patterns for climate-optimal aircraft routing. *Meteorological Applications*, *20*(1), pp.80-93.

Irvine, E.A., Shine, K.P. and Stringer, M.A., 2016. What are the implications of climate change for trans-Atlantic aircraft routing and flight time?. *Transportation Research Part D: Transport and Environment*, *47*, pp.44-53.

Jaeger, E.B. and Sprenger, M., 2007. A Northern Hemispheric climatology of indices for clear air turbulence in the tropopause region derived from ERA40 reanalysis data. *Journal of Geophysical Research: Atmospheres*, *112*(D20).

Kaplan, M.L., Lux, K.M., Cetola, J.D., Huffman, A.W., Riordan, A.J., Slusser, S.W., Lin, Y.L., Charney, J.J. and Waight, K.T., 2004. *Characterizing the severe turbulence environments associated with commercial aviation accidents: A real-time turbulence model (RTTM) designed for the operational prediction of hazardous aviation turbulence environments* (No. NASA/CR-2004-213025).

Kaplan, M.L., Huffman, A.W., Lux, K.M., Charney, J.J., Riordan, A.J. and Lin, Y.L., 2005. Characterizing the severe turbulence environments associated with commercial aviation accidents. Part 1: A 44-case study synoptic observational analyses. *Meteorology and Atmospheric Physics*, *88*, pp.129-152.

Kaplan, M.L., Charney, J.J., Waight, K.T., Lux, K.M., Cetola, J.D., Huffman, A.W., Riordan, A.J., Slusser, S.D., Kiefer, M.T., Suffern, P.S. and Lin, Y.L., 2006. Characterizing the severe turbulence environments associated with commercial aviation accidents. A real-time turbulence model (RTTM) designed for the operational prediction of hazardous aviation turbulence environments. *Meteorology and Atmospheric Physics*, *94*, pp.235-270.

Karnauskas, K.B., Donnelly, J.P., Barkley, H.C. and Martin, J.E., 2015. Coupling between air travel and climate. *Nature Climate Change*, *5*(12), pp.1068-1073.

Kauffmann, P., 2002. The business case for turbulence sensing systems in the US air transport sector. *Journal of Air Transport Management*, *8*(2), pp.99-107.

Kim, J.H. and Chun, H.Y., 2010. A numerical study of clear-air turbulence (CAT) encounters over South Korea on 2 April 2007. *Journal of Applied Meteorology and Climatology*, *49*(12), pp.2381-2403.

Kim, J.H. and Chun, H.Y., 2011. Statistics and possible sources of aviation turbulence over South Korea. *Journal of applied meteorology and climatology*, *50*(2), pp.311-324.

Kim, J.H., Chun, H.Y., Sharman, R.D. and Keller, T.L., 2011. Evaluations of upper-level turbulence diagnostics performance using the Graphical Turbulence Guidance (GTG) system and pilot reports (PIREPs) over East Asia. *Journal of applied meteorology and climatology*, *50*(9), pp.1936-1951.

Kim, J.H., Chan, W.N., Sridhar, B., Sharman, R.D., Williams, P.D. and Strahan, M., 2016. Impact of the North Atlantic Oscillation on transatlantic flight routes and clear-air turbulence. *Journal of Applied Meteorology and Climatology*, *55*(3), pp.763-771.

Kim, J.H., Kim, D., Lee, D.B., Chun, H.Y., Sharman, R.D., Williams, P.D. and Kim, Y.J., 2020. Impact of climate variabilities on trans-oceanic flight times and emissions during strong NAO and ENSO phases. *Environmental Research Letters*, *15*(10), p.105017.

Knox, J.A., 1997. Possible mechanisms of clear-air turbulence in strongly anticyclonic flows. *Monthly weather review*, *125*(6), pp.1251-1259.

Knox, J.A., 2001, June. The breakdown of balance in low potential vorticity regions: Evidence from a clear air turbulence outbreak. In *Preprints, 13th Conf. on atmospheric and oceanic fluid dynamics, Breckenridge, CO, Amer. Meteor. Soc* (Vol. 64, p. 67).

Knox, J.A., McCann, D.W. and Williams, P.D., 2008. Application of the Lighthill–Ford theory of spontaneous imbalance to clear-air turbulence forecasting. *Journal of the Atmospheric Sciences*, *65*(10), pp.3292-3304.

Koch, S.E. and Caracena, F., 2002. Predicting clear-air turbulence from diagnosis of unbalanced flow. In *10 th Conference on Aviation, Range, and Aerospace Meteorology(ARAM), Portland, OR* (pp. 359-363).

Koch, S.E., Jamison, B.D., Lu, C., Smith, T.L., Tollerud, E.I., Girz, C., Wang, N., Lane, T.P., Shapiro, M.A., Parrish, D.D. and Cooper, O.R., 2005. Turbulence and gravity waves within an upper-level front. *Journal of the atmospheric sciences*, *62*(11), pp.3885-3908.

Kopeć, J.M., Kwiatkowski, K., de Haan, S. and Malinowski, S.P., 2016. Retrieving atmospheric turbulence information from regular commercial aircraft using Mode-S and ADS-B. *Atmospheric Measurement Techniques*, *9*(5), pp.2253-2265.

Lane, T.P., Doyle, J.D., Plougonven, R., Shapiro, M.A. and Sharman, R.D., 2004. Observations and numerical simulations of inertia–gravity waves and shearing instabilities in the vicinity of a jet stream. *Journal of the atmospheric sciences*, *61*(22), pp.2692-2706.

Lane, T.P., Sharman, R.D., Trier, S.B., Fovell, R.G. and Williams, J.K., 2012. Recent advances in the understanding of near-cloud turbulence. *Bulletin of the American Meteorological Society*, *93*(4), pp.499-515.

Lee, D.B. and Chun, H.Y., 2018. A numerical study of aviation turbulence encountered on 13 February 2013 over the Yellow Sea between China and the Korean Peninsula. *Journal of Applied Meteorology and Climatology*, *57*(4), pp.1043-1060.

Lee, S.H., Williams, P.D. and Frame, T.H., 2019. Increased shear in the North Atlantic upper-level jet stream over the past four decades. *Nature*, *572*(7771), pp.639-642.

Lee, D.S., Fahey, D.W., Skowron, A., Allen, M.R., Burkhardt, U., Chen, Q., Doherty, S.J., Freeman, S., Forster, P.M., Fuglestvedt, J. and Gettelman, A., 2021. The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmospheric environment*, *244*, p.117834.Lee, J. H., Kim, J. H., Sharman, R. D., Kim, J. and Son, S. W., 2023. Climatology of Clear-Air Turbulence in Upper Troposphere and Lower Stratosphere in the Northern Hemisphere Using ERA5 Reanalysis Data. *Journal of Geophysical Research: Atmospheres*, *128*(1), 2022JD037679.

Lester, P.F., 1993. Turbulence: A new perspective for pilots. Jeppesen Sanderson. *Inc., Englewood, CO*.

Lorenz, D.J. and DeWeaver, E.T., 2007. Tropopause height and zonal wind response to global warming in the IPCC scenario integrations. *Journal of Geophysical Research: Atmospheres*, *112*(D10).

Ma, S. and Zhou, T., 2016. Robust strengthening and westward shift of the tropical Pacific Walker circulation during 1979–2012: A comparison of 7 sets of reanalysis data and 26 CMIP5 models. *Journal of Climate*, *29*(9), pp.3097-3118.

Mancuso, R.L. and Endlich, R.M., 1966. Clear air turbulence frequency as a function of wind shear and deformation. *Monthly weather review*, *94*(9), pp.581-585.

Marlton, G.J., 2016. *On the development, characterisation and applications of a balloon-borne atmospheric turbulence sensor* (Doctoral dissertation, University of Reading).

Marlton, G., Charlton-Perez, A., Harrison, G., Polichtchouk, I., Hauchecorne, A., Keckhut, P., Wing, R., Leblanc, T. and Steinbrecht, W., 2021. Using a network of temperature lidars to identify temperature biases in the upper stratosphere in ECMWF reanalyses. *Atmospheric Chemistry and Physics*, *21*(8), pp.6079-6092.

McCann, D.W., 2001. Gravity waves, unbalanced flow, and aircraft clear air turbulence. *National Weather Digest*, *25*(1/2), pp.3-14.

McCann, D.W., Knox, J.A. and Williams, P.D., 2012. An improvement in clear-air turbulence forecasting based on spontaneous imbalance theory: the ULTURB algorithm. *Meteorological Applications*, *19*(1), pp.71-78.

Meneguz, E., Wells, H. and Turp, D., 2016. An automated system to quantify aircraft encounters with convectively induced turbulence over Europe and the Northeast Atlantic. *Journal of Applied Meteorology and Climatology*, *55*(5), pp.1077-1089.

Miles, J.W. and Howard, L.N., 1964. Note on a heterogeneous shear flow. *Journal of Fluid Mechanics*, *20*(2), pp.331-336.

NCAR-RAL (n.d.), Turbulence.
URL: https://ral.ucar.edu/aap/turbulence (Accessed 30 March 2024)

NOAA (n.d.), Pressure altitude.
URL: https://www.weather.gov/media/epz/wxcalc/pressureAltitude.pdf (Accessed 30 March 2024)

NTSB, (n.d.), The national transport safety board.
URL: https://www.ntsb.gov/safety/data/Pages/Data_Stats.aspx (Accessed 30 March 2024)

IATA, (n.d.), IATA Turbulence Aware Platform
URL: https://www.iata.org/en/services/data/safety/turbulence-platform/ (Accessed 24 June 2024)

Pinto, J.O., Grim, J.A. and Steiner, M., 2015. Assessment of the High-Resolution Rapid Refresh model's ability to predict mesoscale convective systems using object-based evaluation. *Weather and Forecasting*, *30*(4), pp.892-913.

Polvani, L.M. and Kushner, P.J., 2002. Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. *Geophysical Research Letters*, *29*(7), pp.18-1.

Polvani, L.M., Waugh, D.W., Correa, G.J. and Son, S.W., 2011. Stratospheric ozone depletion: The main driver of twentieth-century atmospheric circulation changes in the Southern Hemisphere. *Journal of Climate*, *24*(3), pp.795-812.

Prosser, M.C., Williams, P.D., Marlton, G.J. and Harrison, R.G., 2023. Evidence for large increases in clear-air turbulence over the past four decades. *Geophysical Research Letters*, *50*(11), e2023GL103814.

Reap, R.M., 1996. Probability forecasts of clear-air turbulence for the contiguous US Technical Procedures Bulletin No. 430. *National Weather Service Office of Meteorology, Department of Commerce, Washington DC*.

Reiter, E.R., 1963a. *Jet streams and turbulence* (Doctoral dissertation, Colorado State University. Libraries).

Reiter, E.R., 1963b. *Nature and observation of high-level turbulence especially in clear air*. NWRF.

Reiter, E.R., 1963c. Jet-stream meteorology. *(No Title)*.

Reiter, E.R. and Nania, A., 1964. Jet-stream structure and clear-air turbulence (CAT). *Journal of Applied Meteorology and Climatology*, *3*(3), pp.247-260.

Ren, D., Dickinson, R.E., Fu, R., Bornman, J.F., Guo, W., Yang, S. and Leslie, L.M., 2019. Impacts of climate warming on maximum aviation payloads. *Climate Dynamics*, *52*, pp.1711-1721.

Ren, D., Fu, R., Dickinson, R.E., Leslie, L.M. and Wang, X., 2020. Aviation impacts on fuel efficiency of a future more viscous atmosphere. *Bulletin of the American Meteorological Society*, *101*(10), pp.E1761-E1780.

Roach, W.T. and Bysouth, C.E., 2002. How often does severe clear air turbulence occur over tropical oceans?. *Weather*, *57*(1), pp.8-19.

Romps, D.M., Seeley, J.T., Vollaro, D. and Molinari, J., 2014. Projected increase in lightning strikes in the United States due to global warming. *Science*, *346*(6211), pp.851-854.

Salmon, R., 1998. *Lectures on geophysical fluid dynamics*. Oxford University Press.

Schwartz, B., 1996. The quantitative use of PIREPs in developing aviation weather guidance products. *Weather and Forecasting*, *11*(3), pp.372-384.

Shapiro, M.A., 1976. The role of turbulent heat flux in the generation of potential vorticity in the vicinity of upper-level jet stream systems. *Monthly Weather Review*, *104*(7), pp.892-906.

Shapiro, M.A., 1978. Further evidence of the mesoscale and turbulent structure of upper level jet stream–frontal zone systems. *Monthly Weather Review*, *106*(8), pp.1100-1111.

Shapiro, M.A., 1980. Turbulent mixing within tropopause folds as a mechanism for the exchange of chemical constituents between the stratosphere and troposphere. *Journal of Atmospheric Sciences*, *37*(5), pp.994-1004.

Sharman, R., Tebaldi, C., Wiener, G. and Wolff, J., 2006. An integrated approach to mid-and upper-level turbulence forecasting. *Weather and forecasting*, *21*(3), pp.268-287.

Sharman, R.D., Trier, S.B., Lane, T.P. and Doyle, J.D., 2012. Sources and dynamics of turbulence in the upper troposphere and lower stratosphere: A review. *Geophysical Research Letters*, *39*(12).

Sharman, R.D. and Pearson, J.M., 2017. Prediction of energy dissipation rates for aviation turbulence. Part I: Forecasting nonconvective turbulence. *Journal of Applied Meteorology and Climatology*, *56*(2), pp.317-337.

Sherwood, S.C. and Nishant, N., 2015. Atmospheric changes through 2012 as shown by iteratively homogenized radiosonde temperature and wind data (IUKv2). *Environmental Research Letters*, *10*(5), p.054007.

Simmons, A., Soci, C., Nicolas, J., Bell, B., Berrisford, P., Dragani, R., Flemming, J., Haimberger, L., Healy, S., Hersbach, H. and Horányi, A., 2020. Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5. 1.

Simmons, A.J., 2022. Trends in the tropospheric general circulation from 1979 to 2022. *Weather and Climate Dynamics*, *3*(3), pp.777-809.

Stang (2020) METR2023 - Lecture 7 - Segment 2: Gradient Wind Balance. https://www.youtube.com/watch?v=79s6RUE9184 (Accessed 26/03/2020).

Ștefan, S., Antonescu, B., Urlea, A.D., Buzdugan, L., Andrei, M.D., Necula, C. and Voinea, S., 2020. Study of clear air turbulence related to tropopause folding over the romanian airspace. *Atmosphere*, *11*(10), p.1099.

Stendel, M., Francis, J., White, R., Williams, P.D. and Woollings, T., 2021. The jet stream and climate change. In *Climate change* (pp. 327-357). Elsevier.

Storer, L.N., Williams, P.D. and Joshi, M.M., 2017. Global response of clear-air turbulence to climate change. *Geophysical Research Letters*, *44*(19), pp.9976-9984.

Storer, L.N., Williams, P.D. and Gill, P.G., 2019a. Aviation turbulence: dynamics, forecasting, and response to climate change. *Pure and Applied Geophysics*, *176*, pp.2081-2095.

Storer, L.N., Gill, P.G. and Williams, P.D., 2019b. Multi-model ensemble predictions of aviation turbulence. *Meteorological Applications*, *26*(3), pp.416-428.

Storer, L.N., Gill, P.G. and Williams, P.D., 2020. Multi-diagnostic multi-model ensemble forecasts of aviation turbulence. *Meteorological Applications*, *27*(1), p.e1885.

Strong, C. and Davis, R.E., 2007. Winter jet stream trends over the Northern Hemisphere. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *133*(629), pp.2109-2115.

Tang, Y., Zhang, R.H., Liu, T., Duan, W., Yang, D., Zheng, F., Ren, H., Lian, T., Gao, C., Chen, D. and Mu, M., 2018. Progress in ENSO prediction and predictability study. *National Science Review*, *5*(6), pp.826-839.

Tenenbaum, J., Williams, P.D., Turp, D., Buchanan, P., Coulson, R., Gill, P.G., Lunnon, R.W., Oztunali, M.G., Rankin, J. and Rukhovets, L., 2022. Aircraft observations and reanalysis depictions of trends in the North Atlantic winter jet stream wind speeds and turbulence. *Quarterly Journal of the Royal Meteorological Society*, *148*(747), pp.2927-2941.

Tetzner, D., Thomas, E. and Allen, C., 2019. A validation of ERA5 reanalysis data in the Southern Antarctic Peninsula—Ellsworth land region, and its implications for ice core studies. *Geosciences*, *9*(7), p.289.

Trier, S.B. and Sharman, R.D., 2009. Convection-permitting simulations of the environment supporting widespread turbulence within the upper-level outflow of a mesoscale convective system. *Monthly Weather Review*, *137*(6), pp.1972-1990.

Trier, S.B., Sharman, R.D. and Lane, T.P., 2012. Influences of moist convection on a cold-season outbreak of clear-air turbulence (CAT). *Monthly Weather Review*, *140*(8), pp.2477-2496.

Uccellini, L.W., Brill, K.F., Petersen, R.A., Keyser, D., Aune, R., Kocin, P.J. and des Jardins, M., 1986. A report on the upper-level wind conditions preceding and during the shuttle Challenger (STS 51L) explosion. *Bulletin of the American Meteorological Society*, *67*(10), pp.1248-1265.

Uccellini, L.W. and Koch, S.E., 1987. The synoptic setting and possible energy sources for mesoscale wave disturbances. *Monthly weather review*, *115*(3), pp.721-729.

Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D.C., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A. and Li, X., 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *131*(612), pp.2961-3012.

Vallis, G.K., Zurita-Gotor, P., Cairns, C. and Kidston, J., 2015. Response of the large-scale structure of the atmosphere to global warming. *Quarterly Journal of the Royal Meteorological Society*, *141*(690), pp.1479-1501.

Wallace, J.M. and Hobbs, P.V., 2006. *Atmospheric science: an introductory survey* (Vol. 92). Elsevier.

Williams, P.D. and Joshi, M.M., 2013. Intensification of winter transatlantic aviation turbulence in response to climate change. *Nature Climate Change*, *3*(7), pp.644-648.

Williams, P.D., 2016. Transatlantic flight times and climate change. *Environmental Research Letters*, *11*(2), p.024008.

Williams, P.D., 2017. Increased light, moderate, and severe clear-air turbulence in response to climate change. *Advances in Atmospheric Sciences*, *34*(5), pp.576-586.

Williams, P.D., Alexander, M.J., Barnes, E.A., Butler, A.H., Davies, H.C., Garfinkel, C.I., Kushnir, Y., Lane, T.P., Lundquist, J.K., Martius, O. and Maue, R.N., 2017. A census of atmospheric variability from seconds to decades. *Geophysical Research Letters*, *44*(21), pp.11-201.

Williams (2021), Improved turbulence forecasts for the aviation sector, https://results2021.ref.ac.uk/impact/2bbca9b9-cc5f-4ad7-b7ad-7e1b2393e8d3 (Accessed 31 March 2024)

Williams, P.D. and Storer, L.N., 2022. Can a climate model successfully diagnose clear-air turbulence and its response to climate change?. *Quarterly Journal of the Royal Meteorological Society*, *148*(744), pp.1424-1438.

Wills, R.C., White, R.H. and Levine, X.J., 2019. Northern Hemisphere stationary waves in a changing climate. *Current climate change reports*, *5*, pp.372-389.

Wolff, J.K. and Sharman, R.D., 2008. Climatology of upper-level turbulence over the contiguous United States.
Journal of Applied Meteorology and Climatology, *47*(8), pp.2198-2214.

Woollings, T., Hannachi, A. and Hoskins, B., 2010. Variability of the North Atlantic eddy-driven jet stream. *Quarterly Journal of the Royal Meteorological Society*, *136*(649), pp.856-868.