October 16, 2025

Martin Singh,
Executive Editor,
Weather and Climate Dynamics

Dear Dr. Singh,

We are pleased to submit the revised version of our manuscript entitled "The role of synoptic circulations in lower-tropospheric dry static energy variability over a South Asian heatwave hotspot" for your consideration.

We have revised our manuscript based on the thoughtful comments provided by the two external reviewers, and additionally made changes to improve readability and clarify details as much as possible. To facilitate their evaluation, we have prepared a merged PDF consisting of:

1. A manually annotated revised manuscript – Additions are marked in red to highlight changes made in response to reviewer feedback.
2. A latexdiff-generated track-changes version – Included for completeness, though it omits certain citations and references due to tool limitations.

We hope this format supports a smooth review process. Please let us know if any further clarification is needed.

Sincerely,
Hardik M. Shah
On behalf of all co-authors

Hardik

The role of synoptic circulations in lower-tropospheric dry static energy variability over a South Asian heatwave hotspot

Hardik M. Shah[1] and Joy M. Monteiro[1,2]

[1]Department of Earth and Climate Science, Indian Institute of Science Education and Research Pune, Pune, Maharashtra, 411008, India

[2]Department of Data Science, Indian Institute of Science Education and Research Pune, Pune, Maharashtra, 411008, India

Correspondence: Hardik M. Shah (hardik.shah.reach@gmail.com)

Abstract. We examine the role of the synoptic-scale circulation in the distribution of daily changes of 600-900 hPa dry static energy (DSE) in a heatwave hotspot in northwest South Asia. Using a combination of linear regression and decision trees, we identify how the quasilinear (mean-eddy) and nonlinear (eddy-eddy) components of the flow contribute to different parts of this distribution. We show that the presence of synoptic eddies leads to strong correlations in the quasilinear components due to quasigeostrophy, allowing us to identify periods of upper tropospheric eddy activity. We show that the synoptic eddies induce a zonal quasilinear component which plays an important role in governing the magnitude and sign of daily DSE changes. Nonlinear components are observed to play an important role in the tails of this distribution, and we show that the specific nonlinear components that are involved depend on the phase of growth or decay of DSE and of the pre-existing DSE anomaly. We identify energetically distinct configurations involved in the tails of this distribution, and identify eddy configurations corresponding to each of these energetic configurations. Our analysis thus provides a discrete set of "regimes" which can be used to classify extreme daily DSE changes, and provides a more nuanced approach to compositing extreme events which is sensitive to the dynamics underpinning each event.

1 Introduction

Recent work has expanded our understanding of the likely processes involved in shaping regional distributions of near-surface temperature (Loikith and Neelin, 2015; Ruff and Neelin, 2012; Petoukhov et al., 2008). Due to the known association of large scale eddies with midlatitude weather extremes (Francis and Vavrus, 2012; Petoukhov et al., 2013), one of the questions of interest has been the role of large-scale dynamics in driving different moments characterizing these distributions (Schneider et al., 2014; Horton et al., 2015; Garfinkel and Harnik, 2016). Given that our ability to understand the impacts of global warming is currently limited by our ability to understand changes in atmospheric circulation (Shepherd, 2014, 2015), understanding the links between circulation and temperature variability in the current climate is important to diagnose the fidelity of climate models in representing these links, leading to improved confidence in climate projections.

Earlier attempts to understand the link between large-scale dynamics and temperature variability used quasilinear models of atmospheric dynamics to address free tropospheric temperature variability, where large-scale dynamics dominate. The

choice of quasilinear models was presumably governed by their success in reproducing atmospheric flow statistics (Schneider and Walker, 2006; O'Gorman and Schneider, 2007). However, quasilinear models without nonlinear processes predict Gaussian temperature distributions (Schneider et al., 2014), and are therefore unlikely to provide a reasonable explanation for the observed temperature distributions. While quasilinear models can produce skew by the incorporation of multiplicative noise processes (Sardeshmukh and Sura, 2009), Garfinkel and Harnik (2016) showed that the inclusion of eddy-eddy interactions can qualitatively reproduce the observed patterns of temperature distributions. There has also been a strong indication that skewness is related to distance from the jetstream via the meridional movement of synoptic systems. We note that the importance of eddy-eddy interactions is not limited to understanding temperature variability, but has been studied in a variety of geophysical contexts (Svirsky et al., 2023; Bouchet et al., 2019; Delsole and Farrell, 1996; Marston and Tobias, 2023). More recent work has tried to build on these initial insights, by using climate models for verifying the importance of nonlinear processes in generating non-Gaussian tails (Linz et al., 2018), and formulating nonlinear approximation schemes (Tamarin-Brodsky et al., 2019).

While such literature has helped expand our understanding of free-tropospheric temperature variability, the widely employed quasilinear models ignore the role played by the zonal mean flow. This choice may have been made due to their interest in the deformation flow field, but this choice makes it problematic to apply their results in an Eulerian framework, which is more relevant for regional temperature variability. Understanding near-surface temperature variability comes with the additional challenge of disentangling the effects of atmospheric macroturbulence and local processes such as land-atmosphere interactions and boundary layer feedback (Miralles et al., 2014; Dirmeyer et al., 2018; Chen and Dirmeyer, 2019). Lately, there have been frameworks proposed for quantifying the contribution of different proximate processes towards driving temperature tendency in different deciles of near-surface temperature (Linz et al., 2020).

Studies which focus on the tails of near-surface temperature distribution have established the association between heatwaves and dynamical processes such as blocking events and quasi-resonant amplification (Petoukhov et al., 2016; Kornhuber et al., 2019; Rao et al., 2021). The synoptic patterns associated with such dynamical processes are usually studied using composites (Ratnam et al., 2016; Rohini et al., 2016). While this strategy is useful for understanding "average" conditions, it can obscure the diversity of pathways that could lead to the outcomes of interest. For example, the composite picture presented in Monteiro and Caballero (2019) while studying extreme wet-bulb temperature events was found to average over at least two mechanisms, which depended on the phase of the wave packet over the region of interest (Pandey et al., 2020).

Since it is unlikely that we will be able to predict the occurrence of extreme events on seasonal timescales, it has been suggested that subseasonal-to-seasonal predictions might benefit from the prediction of waveguides rather than wave driven extremes (White et al., 2021). Furthermore, the relationship between upper tropospheric forcing and near-surface response is not always clear (White et al., 2021), and it might be a useful exercise to understand the impacts of Rossby wave packets on lower tropospheric quantities that are directly related to near-surface temperature. Such an exercise might allow evaluating seasonal "propensity of extremes" (Prodhomme et al., 2022) directly using such quantities instead. Thus, understanding the variability of such quantities is not only interesting from an atmospheric turbulence viewpoint but also from a seasonal predic-

tion viewpoint. One possible approach (which we take in this study) to addressing the question of understanding near-surface temperature variability and extremes could comprise the following steps:

1. Identifying a lower tropospheric quantity that is highly associated with near-surface temperature and is strongly influenced by the atmospheric circulation

2. Identifying quasilinear and nonlinear contributions to different parts of the distribution of this quantity

3. Characterizing the different pathways to extreme values of this quantity and relating them to characteristics of the circulation

In this study, we identify the lower tropospheric (600-900 hPa) dry static energy $\mathcal{S}_{Tot}$ (where the subscript "Tot" stands for total), as a suitable quantity whose daily changes are both highly correlated with near-surface temperature and have a significant contribution due to large-scale advection in our region of interest, a heatwave hotspot in South Asia. We study the variability of advection-driven daily changes in $\mathcal{S}_{Tot}$, during March-April, which are a part of the heatwave season in this region. We study the contribution of quasilinear and nonlinear advective fluxes to the distribution of advection driven daily changes in $\mathcal{S}_{Tot}$. We use an interpretable machine learning algorithm, the decision tree, to develop a nuanced picture of combinations of advective contributions that lead to extreme daily changes in $\mathcal{S}_{Tot}$ and show how a combination of these results and compositing leads to a richer description of the different ways in which the upper tropospheric circulation affects daily changes in $\mathcal{S}_{Tot}$. Finally, we summarize our work with an analysis of the lifecycle of daily changes in $\mathcal{S}_{Tot}$, highlighting how circulation drives the accumulation, saturation, and ventilation of $\mathcal{S}_{Tot}$ over the region.

2 Data and Methods

2.1 Data

We have used the European Center for Medium-Range Weather Forecasts Reanalysis version 5 (ERA5; Hersbach et al. (2020)) reanalysis dataset (0.25° x 0.25° resolution) for analysing daily mean quantities. The period of analysis spans the months of March and April, from 1980 to 2022, with the exclusion of 82 dates due to data quality issues. We use unevenly spaced pressure levels, with higher resolution near the surface and the upper levels. The specific pressure levels (in hPa) are 100, 125, 150, 175, 200, 225, 250, 300, 400, 500, 600, 700, 750, 775, 800, 825, 850, 875, and 900.

2.2 Methodology

Our region of interest is the South Asian heatwave hotspot (Ratnam et al., 2016; Rohini et al., 2016), defined as the area enclosed between 25^oN and 31^oN, and 68^oE and 78^oE. We use dry static energy (DSE) for tracking the energy content of the atmospheric parcels. Since DSE is conserved for dry, large scale adiabatic flows, it is particularly suited for diagnosing energy advection driven by the resolved states of the large scale atmosphere. Using Reynolds decomposition, we decomposed

the DSE and velocity fields into daily climatology and anomaly components ($X = \bar{X} + X'$) where the climatology represents the background state and anomalies represent transient variations. Daily climatology was computed as a 10-day rolling mean of daily mean computed over the period of analysis.

Our analysis focuses on the volume of the lower troposphere bounded vertically between 600 hPa and 900 hPa, and horizontally by the spatial extent of the region of interest. As previously mentioned, the dry static energy content of air parcels contained in this volume is given by $\mathcal{S}_{Tot}$, and daily changes in $\mathcal{S}_{Tot}$ are given by $\delta\mathcal{S}_{Tot} := [\mathcal{S}_{\mathrm{Tot}}(t_i) - \mathcal{S}_{\mathrm{Tot}}(t_{i-1})]_{i=1}^{N}$ where N is the total number of days in the analysis. The daily DSE convergence into this volume is calculated as the mass-weighted advection of DSE into it, given by $\delta\mathcal{S}$. The local conservation law for DSE is expressed in equation 1, with ϵ representing the terms involving unresolved or parameterized energy fluxes. Starting with the point conservation law for DSE,

$$\frac{\partial(\mathrm{DSE})}{\partial t} = -\nabla \cdot (\mathbf{v}\,\mathrm{DSE}) + \epsilon \tag{1}$$

We analyzed both the divergent and non-divergent components of the RHS; Even though the divergent term was not small, advective contributions to the change in DSE are larger than the divergent contributions in the presence of synoptic eddies. The existing literature cited in the introduction (Schneider et al., 2014; Garfinkel and Harnik, 2016; Tamarin-Brodsky et al., 2019; Linz et al., 2020) has also used the advective component for studying the role of eddies in driving midlatitude temperature variability. Further, the ERA5 divergence term may be unreliable due to contamination by numerical noise, as recently acknowledged by in the Copernicus Knowledge Base 2025 (for Medium-Range Weather Forecasts , ECMWF) and recent evaluations (Mayer et al., 2021). Because our objective is not to close the DSE budget but to characterize the role of synoptic eddies in shaping quasilinear and nonlinear DSE flux relationships, we ignore the contribution due to divergence and focus solely on the advective component in this study. Upon neglecting the mass conservation terms and discretizing the quantities, equation 1 reduces to the advection form in equation 2.

$$\delta(\mathrm{DSE}) = \delta t\,(-\mathbf{v} \cdot \nabla \mathrm{DSE} + \epsilon) \tag{2}$$

Taking the mass-weighted integral of equation 2, and performing Reynolds decomposition on both sides,

$$\delta\mathcal{S}_{Tot} = \delta t\,(\oint_D (-\mathbf{v} \cdot \nabla \mathrm{DSE} + \epsilon)\,dM)\,/M \tag{3a}$$

$$\delta\bar{\mathcal{S}}_{Tot} + \delta\mathcal{S}'_{Tot} = \delta t\,(\oint_D -(v'\frac{\partial\overline{\mathrm{DSE}}}{\partial y} + v'\frac{\partial \mathrm{DSE}'}{\partial y} + \bar{v}\frac{\partial\overline{\mathrm{DSE}}}{\partial y} + \bar{v}\frac{\partial \mathrm{DSE}'}{\partial y} + \ldots)\,dM)\,/M \tag{3b}$$

Using shorthand notation for the convergence terms on the RHS,

$$\delta\bar{\mathcal{S}}_{Tot} + \delta\mathcal{S}'_{Tot} = v'\bar{\mathcal{S}}_y + v'\mathcal{S}'_y + \bar{v}\bar{\mathcal{S}}_y + \bar{v}\mathcal{S}'_y + \ldots + \varepsilon \tag{4}$$

Where the order of terms on the RHS is preserved between 3b and 4. Ignoring the small daily changes in $\bar{\mathcal{S}}_{Tot}$, the mass weighted daily climatology of lower tropospheric DSE, we arrive at the relationship,

$$\delta\mathcal{S}'_{Tot} = \delta\mathcal{S} + \varepsilon \tag{5}$$

The terms and operations used in equations 1-5 are defined as follows:

$$\delta\mathcal{S} = v'\bar{\mathcal{S}}_y + v'\mathcal{S}'_y + \bar{v}\bar{\mathcal{S}}_y + \bar{v}\mathcal{S}'_y + \ldots \tag{6a}$$

$$\mathbf{v} = (u, v, w) \tag{6b}$$

$$\oint_D = \int_{Z=Z_{900\,hPa}}^{Z=Z_{600\,hPa}} \int_{\lambda=68^o}^{\lambda=78^o} \int_{\phi=25^o}^{\phi=31^o} \tag{6c}$$

$$dM = \rho * (R_e \cos\phi\, d\lambda) * (R_e\, d\phi) * (dZ) \tag{6d}$$

$$M = \oint_D dM \tag{6e}$$

$$\delta\mathcal{S}'_{Tot} = (\oint_D \delta DSE'\, dM)/M \tag{6f}$$

$$\bar{\mathcal{S}}_{Tot} = (\oint_D \delta\overline{DSE}\, dM)/M \tag{6g}$$

where δt = 86400, $\phi =$ Degrees of Latitude, $\lambda =$ Degrees of Longitude, $Z =$ Geopotential, $R_e =$ Radius of the Earth. The integrals above are computed by using discrete spatial steps of length 0.25^o in the latitudinal and longitudinal directions. We note that the correlation between $\delta\mathcal{S}_{Tot}$ and $\delta\mathcal{S}$ was highest when we used 2-day average values of $\delta\mathcal{S}$ given by $\left[(\delta\mathcal{S}_{t_i} + \delta\mathcal{S}_{t_{i-1}})/2\right]_{i=1}^{N}$, where N is the total number of days in the analysis. Consequently, all variables related to advection ($\delta\mathcal{S}$ and its Reynolds components $v'\bar{\mathcal{S}}_y$, $v'\mathcal{S}'_y$, $\ldots$) used throughout this study, and the mass-weighted values of velocity and $\mathcal{S}'$ in the advection regimes section (5.1), represent 2-day averages.

We use the decision tree classification model, which has been extensively applied in climate science (Gagne et al., 2009; Burrows et al., 1995; Xu et al., 2020; Zhang et al., 2012; Wei et al., 2020). The model recursively partitions the input feature space, generating disjoint branches each corresponding to a set of decision rules that group the samples belonging to the same class together. Each branch consists of internal nodes representing the decision variables or predictors and their optimal split values, and a terminating leaf node representing the dependent variable filtered by the conditions corresponding to its branch. The split values are determined by the criterion of maximum reduction in information entropy (Shannon, 1948). The performance of a decision tree model is usually evaluated using the F1-score (harmonic mean of precision and recall for a given class), averaged over all classes of the dependent (class) variable (Hastie et al., 2001). The model output is visualized as an inverted tree. A detailed description of the decision tree classification model is given in Wei et al. (2020). We analyze the explainability of the identified mechanisms across the distribution of daily advection using linear regression and residual diagnostics.

The entropy associated with a categorical variable Y with m classes is given by

$$\mathrm{H}(Y) = -\sum_{i=1}^{m} p_i \log_2(p_i) : \mathrm{P}(Y = i) = p_i \tag{7}$$

We first use the decision tree model to identify the conditions primarily associated with different tercile based classes of $\delta\mathcal{S}$. We then analyze the explainability provided by the identified conditions across the distribution of $\delta\mathcal{S}$ using linear regression and residual analysis. Finally, we use the decision tree model for identifying conditions additionally associated with the extreme deciles of $\delta\mathcal{S}$. All calculations are performed using the Python packages xarray (Hoyer and Hamman, 2017), numpy (Harris et al., 2020), pandas (McKinney et al., 2010), statsmodels (Seabold and Perktold, 2010), scikit-learn (Pedregosa et al., 2011), and plotting was done with the help of Python packages matplotlib (Hunter, 2007) and cartopy (Met Office, 2010 - 2015).

3 Lower tropospheric advection links circulations with daily changes in T_{2m} anomaly

Positive T_{2m} anomaly (T'_{2m}) is associated with both increased fluxes of shortwave radiation and decreased sensible heat fluxes from the ground (Supplementary Fig. S1). In contrast to this, "mega heatwaves" (Domeisen et al., 2023; Miralles et al., 2014), were associated with increased sensible heat fluxes from dessicated surfaces which are known to contribute to feedback cycles sustaining surface warming over multiple days. On the other hand, we find that daily changes in T'_{2m} (which we denote by $\delta T'_{2m}$) are linked to $\delta\mathcal{S}_{Tot}$, which in turn are highly correlated with $\delta\mathcal{S}$ (Figs. 1a, 1b). We find that $\delta T'_{2m}$ is also strongly linked to $\delta\mathcal{S}$, especially at large magnitudes (Fig. 1c). This suggests that the surface fluxes are likely a slower timescale forcing, and that daily variability is primarily governed by the atmospheric flow. Further, we find that the lower tropospheric daily temperature changes are strongly explained by daily T_{2m} changes (Supplementary Fig. S2). Thus, the daily changes in lower tropospheric DSE capture a large proportion of the variation in the daily changes in T_{2m}. Further, March and April are characterized by climatologically low values of precipitation (Supplementary Fig. S3a) and cloud cover (Supplementary Fig. S3b). Therefore, we proceed with the hypothesis that daily mean advective fluxes of DSE explain daily mean changes of DSE sufficiently well, and we don't consider contributions from radiative heating and latent heat release in this work. While $\delta\mathcal{S}_{Tot}$ is unbiasedly approximated by $\delta\mathcal{S}$ in the body of its distribution, parameterized energy fluxes become important for extreme values of $\delta\mathcal{S}$, and cause the systematic overprediction of $\delta\mathcal{S}_{Tot}$ (Fig. 1b). We note the barotropic nature of $\mathcal{S}'$ and the eddy wind fields, that strengthen progressively with height (Supplementary Fig. S4), and that the vertical velocity field was strongly related to quasigeostrophic omega (Supplementary Fig. S5), highlighting the influence of upper tropospheric dynamics in driving surface anomalies. Motivated by the association between daily changes in near-surface temperature anomaly and advection driven daily changes in lower tropospheric DSE, and the influence of atmospheric circulation, we identify $\delta\mathcal{S}$ as our quantity of interest.

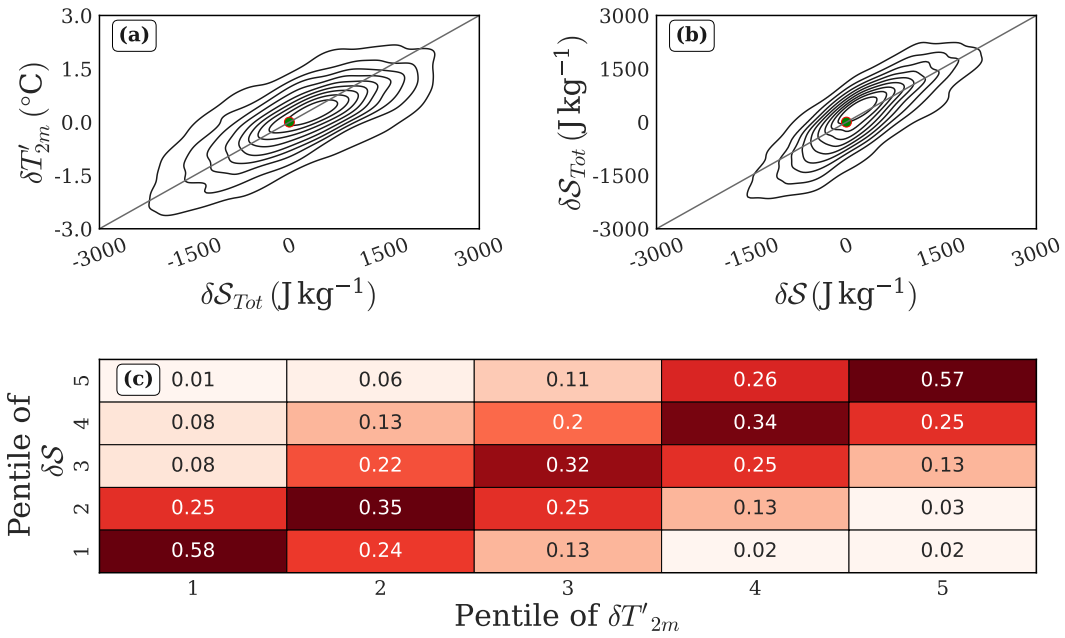

Figure 1. This figure shows the pairwise relationships between daily changes in T_{2m} anomaly ($\delta T'_{2m}$), daily changes in 600-900 hPa mass-weighted DSE ($\delta \mathcal{S}_{Tot}$), and daily 600-900 hPa mass-weighted advective convergence of DSE ($\delta \mathcal{S}$) during the combined periods of March and April. (a) The joint probability density plot between $\delta T'_{2m}$ and $\delta \mathcal{S}_{Tot}$. (b) The joint probability density plot between $\delta \mathcal{S}_{Tot}$ and $\delta \mathcal{S}$. For both these plots, the density contours are estimated using a non-parametric kernel density estimation method; the green dot with the red border is placed at the origin for reference. (c) Conditional probability distribution of pentiles of $\delta T'_{2m}$ given the pentile of $\delta \mathcal{S}$. We compute the conditional probability table by cross-tabulating the pentiles of $\delta T'_{2m}$ with the pentiles of $\delta \mathcal{S}$, and then normalizing each row by corresponding row totals to yield $P\left(\delta T'_{2m} \in P_i \,|\, \delta S \in P_j\right)$ where P_i is the i^{th} pentile of δT_{2m} and P_j is the j^{th} pentile of $\delta \mathcal{S}$. Pentiles refer to five equally populated bins based on ranked values.

4 Primary advective contributions to the $\delta \mathcal{S}$ budget

Our strategy to unravel the relationships governing the variability in advection of lower tropospheric DSE is to first find the underlying mechanisms that govern its sign, and subsequently study the drivers of its magnitude. A description of the magnitudes of the Reynolds components of $\delta \mathcal{S}$ is provided in the Supplementary Fig. S6. The magnitudes of these components are further justified by analyzing the magnitudes of the climatology and anomaly of $\mathbf{v}$ and $\nabla \mathcal{S}$ in the zonal, meridional, and vertical directions (Supplementary Table S1 and Supplementary Figs. S7). A zonally oriented climatology of $\mathcal{S}_{Tot}$ and horizontal winds is seen in Supplementary Fig. S8.

This section focuses on the first part of this strategy. $\delta \mathcal{S}$ closely resembles $\mathbb{N}(0, \sigma)$ in its body, with deviation in the tails (Supplementary Fig. S9). Due to the observed symmetry in the body of $\delta \mathcal{S}$, it is possible to model the separation of its signs by using tercile-based classification of $\delta \mathcal{S}$ with the classes labelled as Negative, Neutral, and Positive, as the dependent variable

in a decision tree classification model. The 33^{rd} and 66^{th} percentile values demarcated in Supplementary Figs. S6a, S6b. Even though this partitioning seems coarse-grained, we will subsequently show that most large $\delta\mathcal{S}$ events are contained in the Positive and Negative classes defined here.

A multivariate decision tree model for tercile-based classes of daily advection retrieves combinations of two of the largest magnitude Reynolds components, namely $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$ (Fig. 2a). The model is able to identify at least one decision rule for each of the three classes. The decision rules represented by the branches of the tree are described by the conditional probability statements presented in Table 1. The Positive and Negative classes are identified with a larger probability of occurrence than the Neutral class. The decision rules associated with the Positive and Negative classes are based on the in-phase combination of $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$, while the Neutral class is associated with out-of-phase combinations of $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$.

Advection Class (k)	**Node #**	**Conditions (C)**	$\mathbf{P(Y = k \mid C)}$
Negative	1	$\bar{u}\mathcal{S}'_x \leq 0, w'\bar{\mathcal{S}}_z \leq 0$	0.75
Neutral	2, 3	$(\bar{u}\mathcal{S}'_x \leq 0,\ w'\bar{\mathcal{S}}_z > 0) \mid (\bar{u}\mathcal{S}'_x > 0,\ w'\bar{\mathcal{S}}_z \leq 0)$	0.5
Positive	4	$\bar{u}\mathcal{S}'_x > 0, w'\bar{\mathcal{S}}_z > 0$	0.75

Table 1. This table describes the decision tree model presented in Fig. 2a. The decision rules and associated class probabilities are approximately the same for both months; therefore, we present the combined results. The means of the involved variables, and their scalar multipliers of σ, are both ~ 0, so expressions such as 0.1σ are approximated as 0 for ease of readability. For example, $\bar{u}\mathcal{S}'_x \leq -0.1\sigma(\bar{u}\mathcal{S}'_x)$ is simplified to $\bar{u}\mathcal{S}'_x \leq 0$. Table headers are described as follows. **Advection Class (K)**: The tercile based classes of $\delta\mathcal{S}$ modeled by the decision tree classification model. **Node #**: The leaf node numbers identifying each advection class. **Conditions (C)**: The set of decision rules that characterize the branches associated with the corresponding leaf node(s) identifying a given advection class. $\mathbf{P(Y = K \mid C)}$: The probability of occurrence of the identified advection class, given the corresponding decision rules.

The model generated decision rules given in Table 1 are based on a selection of the variables and split points that maximize entropy reduction (section 2.2). However, the underlying correlation structure of the dataset is not accounted for by the model, specifically, multicollinearity. For example, consider a variable that independently explains a large fraction of the variance in the dependent variable, but a much smaller fraction when included with one or more of the remaining model predictors. The model algorithm would not identify this variable as important, and its confounded influence would remain unidentified. We find that $v'\bar{\mathcal{S}}_y$ is the confounder in our case, due to a strong inverse association of $v'\bar{\mathcal{S}}_y$ with both other quasilinear components, $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$ (Fig. 2b). $v'\bar{\mathcal{S}}_y$ is important to decipher the interaction pathways involving $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$. Moreover, its magnitude is also comparable to these variables (Supplementary Fig. S6c), making its quantitative effect significant. Thus, $v'\bar{\mathcal{S}}_y$ is additionally identified as a key component in the analysis of $\delta\mathcal{S}$. It will be shown later (Fig. 4a) that the sum of all three quasilinear terms is a strong primary indicator of advection.

The inverse relationship between $v'\bar{\mathcal{S}}_y$ and $w'\bar{\mathcal{S}}_z$ is traced back to an in-phase relationship between mass-weighted integrals of v' and w' over the volume of the box. Due to the vertical coherence of the meridional and vertical velocity anomaly fields (Supplementary Fig. S4) and corresponding spatial derivatives of $\bar{\mathcal{S}}$ (Supplementary Fig. S10), the quasilinear components

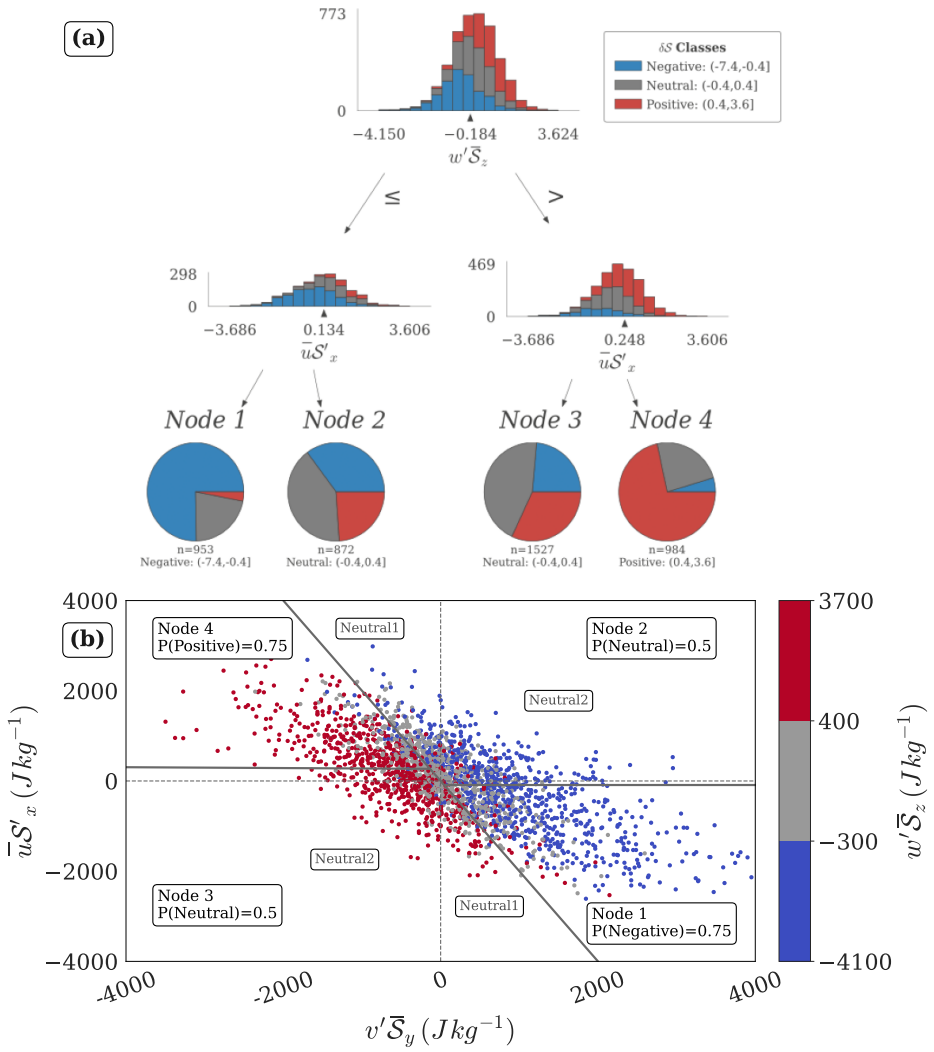

Figure 2. This figure examines the relationships governing the differences between terciles of daily advection for the combined period of March and April. (a) Standardized decision tree model for tercile based classes of $\delta\mathcal{S}$, across March-April. The root node and the internal (decision) nodes are represented by stacked bar charts of the class frequencies, and the distribution of the classes in the leaf nodes is represented by the pie charts. The variables are standardized, and the split thresholds at each node reflect the number of standard deviations from the mean. The model achieves an F1-score of ~ 0.6, with early-stop criteria as follows: entropy decrease ≥ 0.03 at each level, $\geq 40\%$ sample size for level splitting, $\geq 20\%$ sample size for leaf nodes. (b) Scatterplot showing the relationship between the unstandardized values of $v'\bar{\mathcal{S}}_y$ and $\bar{u}\mathcal{S}'_x$, colored by terciles of $w'\bar{\mathcal{S}}_z$ across March-April. The solid grey lines approximate decision boundaries, with each region labeled by the corresponding Node # and dominant class probability. The labels "Neutral1" and "Neutral2" correspond to the two kinds of deviations from the primary relation described in section 4.1. Pairwise correlations: $r(v'\bar{\mathcal{S}}_y, w'\bar{\mathcal{S}}_z) = -0.75$, $r(v'\bar{\mathcal{S}}_y, \bar{u}\mathcal{S}'_x) = -0.75$, $r(\bar{u}\mathcal{S}'_x, w'\bar{\mathcal{S}}_z) = 0.25$.

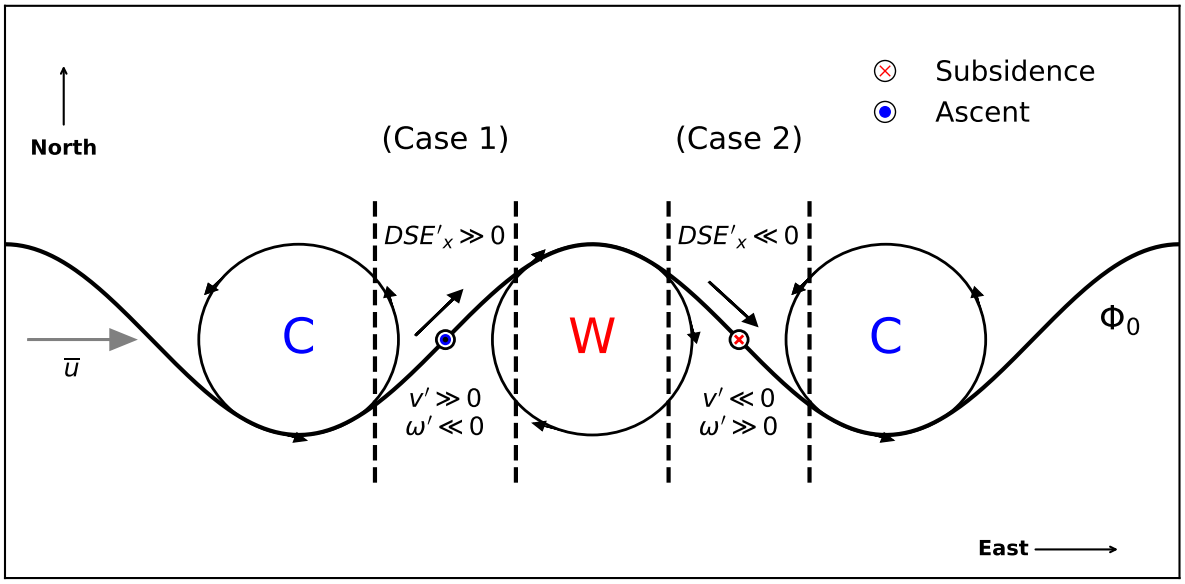

Figure 3. This schematic clarifies the relationship between the large scale quantities v', w', and $\mathcal{S}'_x$. Consider a climatological zonal flow ($\bar{u}$) and a geopotential contour of value Φ_0. Let the geopotential contour be displaced meridionally in a sinusoidal manner, with cyclonic and anticyclonic circulations induced along the minima and maxima of the wave, respectively. The corresponding cold and warm anomalies are indicated by letters 'C' and 'W' respectively. We demarcate the region of interest by a set of dotted vertical lines, and place it first at the head of the cold anomaly (case 1) and then at the head of the warm anomaly (case 2). In case 1, the region experiences a local maximum $v'(\gg 0)$ along with a local minimum $w'(\ll 0)$ both of which advect DSE under a strong climatological DSE gradient in y and z directions; further, since the region here is placed at the edge between the cold and warm anomalies, $DSE'_x > 0$ is induced and leads to advection of DSE by $\bar{u}$. Such a configuration sees advection components maximize in magnitude with signs as follows: $v'\bar{\mathcal{S}}_y > 0, w'\bar{\mathcal{S}}_z < 0, \bar{u}\mathcal{S}'_x < 0$. For case 2, all the above signs are reversed and the observed advection components are as follows: $v'\bar{\mathcal{S}}_y < 0, w'\bar{\mathcal{S}}_z > 0, \bar{u}\mathcal{S}'_x > 0$. Such configurations yield the relationships making up the primary mechanism.

behave like linear transformations of the mass-weighted quantities v', w', which are related through the quasigeostrophic omega equation. At the same time, $w'\bar{\mathcal{S}}_z$ and $\bar{u}\mathcal{S}'_x$ are not strongly correlated, suggesting the role of other mechanisms, either external to the system or related to different latent relationships with respect to $v'\bar{\mathcal{S}}_y$, governing at least one of the two variables. For example, mass-weighted $\mathcal{S}'_x$, which is scaled by a vertically coherent $\bar{u}$, is likely a function of the location and sign of one or multiple eddies interacting with the region or the result of advection by flows not associated with coherent structures.

4.1 Phenomenology

We plot the decision spaces generated by the tree model on the scatterplot involving all three quasilinear variables identified above (Fig. 2b). However, it is unclear what physical conditions they represent. Thus, we analyze individual instances and obtain a phenomenology of the large scale spatial patterns of winds and $\mathcal{S}'_{Tot}$ corresponding to different regions on the scatterplot.

Both periods (March and April) showed similar results in the following sections; therefore, we focus on presenting an in-depth analysis of April for the rest of this paper with the corresponding figures for March presented in the Appendix.

The decision spaces associated with the Negative and Positive classes have been annotated as Node 1 and Node 4, respectively, along with corresponding probability of occurrence in Fig. 2b. The large values of $|v'\bar{\mathcal{S}}_y|$ ($> \sigma$) are associated with a highly predictable sign and magnitude of the other quasilinear quantities, $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$. Smaller values ($|v'\bar{\mathcal{S}}_y| < \sigma$) were associated with variability in the signs of both other quantities. The extreme deciles (1 and 10) of advection driven daily changes in $\mathcal{S}$ almost exclusively occupied the decision spaces associated with the Negative and Positive Nodes 1 and 4 (Fig. 2b; Supplementary Fig. S11), and represented various configurations of anticyclones and cyclones located in proximity to the region of interest.

The decision spaces associated with the Neutral class have been annotated as Node 2 and Node 3 along with probabilities of the Neutral class in Fig. 2b. These decision spaces only weakly prefer the Neutral class with lesser but comparable probabilities of occurrence of the other classes. We find that these are associated with two kinds of deviations from the expected relationships established above. The first kind is where the meridional and vertical quasilinear components are in phase. We label the regions on the scatterplots representing this kind of deviation as "Neutral1". These points are observed to lie between the lines $\bar{u}\mathcal{S}'_x = -2 * v'\bar{\mathcal{S}}_y$ and $v'\bar{\mathcal{S}}_y = 0$, where $\bar{u}\mathcal{S}'_x * v'\bar{\mathcal{S}}_y < 0$. We find that they represent conditions where mass-weighted integrals of v' and w' acted out of phase, i.e., ascending northerly winds and subsiding southerly winds, as expected by the quasigeostrophic relationship. These conditions are characterized by small magnitude ($< \sigma$) of $v'\bar{\mathcal{S}}$, and the horizontal dominance of $\bar{u}\mathcal{S}'_x$. Despite the weak meridional winds, instances of large vertical winds are observed in association with mainly the zonal arm of a cyclone or anticyclone located directly to the north or south of the region, when both the meridional arms of an eddy overlap with the region. Here, v' is rendered small due to averaging over opposite phases, while large w' persists due to homogeneity of its phase across the region, which is in agreement with the phase of quasigeostrophic omega (Supplementary Fig. S5). $\mathcal{S}'_x$ is purely a function of the differently signed eddies interacting with different zonal halves of the region, and is not sensitive to the v', w' relationship. In cases without eddies, w' is small as v', and only the zonal mean flow advects a large $\mathcal{S}'$ into the region, under a $\mathcal{S}'_x$ governed by no particular structures. Typical configurations associated with the "Neutral1" regions are presented in Supplementary Figs. S12a, S12b.

The second kind of deviation is represented by points in the first and third quadrants, where the meridional and zonal quasilinear advection terms are in phase. We label these regions as "Neutral2" in the scatterplot Fig. 2b. The inverse phase relationship between the meridional and vertical quasilinear terms is preserved even at small magnitudes. It represents coherent configurations given by zonally oriented arms of southern/ northern large scale eddies and ridge-like conditions over the region. $\bar{u}\mathcal{S}'_x$ is reversed in sign either because of intensification of $\mathcal{S}'$ along the southern edge of High Mountain Asia, or due to weak larger scale structures of $\mathcal{S}'_x$ not always directly linked with circulation structures (the quantities are 600-900 hPa averaged, and generally, a given location could be influenced by diabatic forcing in the lower vertical levels, or be affected by other proximate eddies). Typical configurations associated with the "Neutral2" regions are presented in Supplementary Figs. S12c, S12d.

4.2 Explainability and residual bias

We assess the performance of the quasilinear advective contributions by regressing the daily advection of dry static energy into the lower troposphere, $\delta\mathcal{S}$, against the sum of its quasilinear components, $\delta\mathcal{S}_{QL} = v'\bar{\mathcal{S}}_y + w'\bar{\mathcal{S}}_z + \bar{u}\mathcal{S}'_x$.

Using the coefficient of determination (R^2), we find that that $\delta\mathcal{S}_{QL}$ explains $\sim 65\%$ of the variability in $\delta\mathcal{S}$. To evaluate its performance across the distribution of daily advection, we examine residuals stratified by deciles of $\delta\mathcal{S}$. In the central deciles, the residuals (observation - prediction) are more or less symmetrically distributed about zero (Fig. 4a), indicating an unbiased model fit. Towards the extremes, however, the residuals exhibit a systematic directional bias (as implied by the increasing magnitude of the mean residuals towards extreme deciles in Fig. 4a), indicating the insufficiency of driving mechanisms in the extremes of the model.

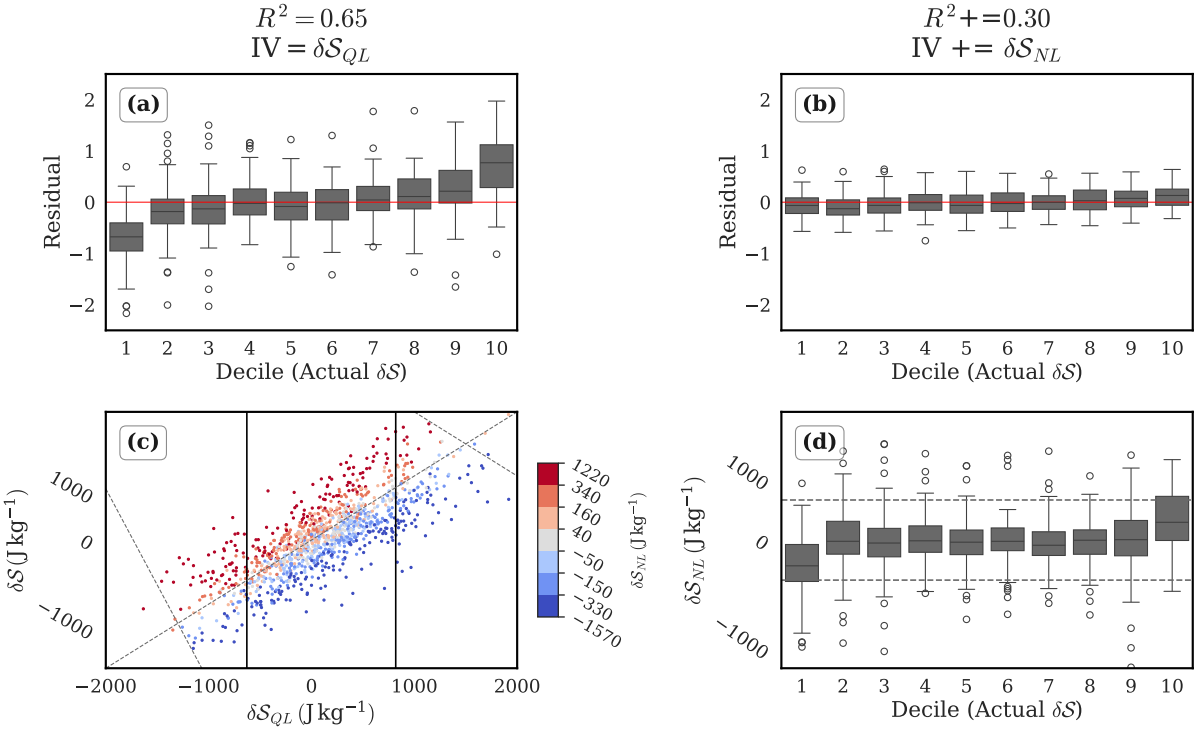

Figure 4. This figure presents two regression models for $\delta\mathcal{S}$, and complementary plots based on unstandardized data to aid interpretability, for the period of April. (a) Boxplots of residuals grouped by deciles of $\delta\mathcal{S}$, for the fitted standardized regression model $\delta\mathcal{S} \sim 0.8 * \delta\mathcal{S}_{QL} + 0.004$. (b) Same as (a), but for the fitted standardized regression model $\delta\mathcal{S} \sim 0.95 * \delta\mathcal{S}_{QL} + 0.6 * \delta\mathcal{S}_{NL} + 0.002$. Model performance summaries are provided in Supplementary Tables S2 and S3. (c) Scatterplot between the unstandardized variables, $\delta\mathcal{S}$ and $\delta\mathcal{S}_{QL}$, colored by quantile based categories of $\delta\mathcal{S}_{NL}$. The two negatively sloped dotted lines connect the peak values of $\delta\mathcal{S}$ and $\delta\mathcal{S}_{QL}$ in the right and left tails, highlighting the asymmetric saturation effect of $\delta\mathcal{S}_{NL}$. (d) Boxplot of unstandardized values of $\delta\mathcal{S}_{NL}$ grouped by deciles of $\delta\mathcal{S}$.

The bias in the primary contributions model is corrected by including the sum of nonlinear variables ($\delta\mathcal{S}_{NL} = v'\mathcal{S}'_y + w'\mathcal{S}'_z + u'\mathcal{S}'_x$). The addition of $\delta\mathcal{S}_{NL}$ eliminates the non-zero mean of residuals in the tails, and reduces the variance of residuals in all the deciles of $\delta\mathcal{S}$, with the largest reduction in the tails (Fig. 4b). The importance of $\delta\mathcal{S}_{NL}$ in the tails of $\delta\mathcal{S}$ is verified by the nonzero mean of $\delta\mathcal{S}_{NL}$ in the extreme deciles of $\delta\mathcal{S}_{QL}$ (Fig. 4d). We infer that the quasilinear contributions are active throughout the distribution of $\delta\mathcal{S}$, and the nonlinear contributions primarily modulate the extremes.

Together, the quasilinear and nonlinear contributions account for $\sim 95\%$ of the variability in $\delta\mathcal{S}$. The rest of the component terms on the RHS of equation 6a account for the remaining $\sim 5\%$ of the variability in $\delta\mathcal{S}$. We shall refer to the quasilinear contributions as the primary contribution, and the nonlinear contributions as the secondary contribution for the remainder of the paper.

Figure 4c shows that the magnitude of $\delta\mathcal{S}$ peaks earlier than $\delta\mathcal{S}_{QL}$ in both tails, suggesting that the main function of the nonlinear terms in the tails is to act against the quasilinear contribution, effectively saturating the growth of $\delta\mathcal{S}$. While this in itself may not seem surprising and has been noted elsewhere (Garfinkel and Harnik, 2016), we also observe a larger difference between the peak values of $\delta\mathcal{S}$ and $\delta\mathcal{S}_{QL}$ in the right tail as compared to the left tail. This asymmetry suggests that the saturative effect of the nonlinear terms is stronger there.

5 Secondary advective contributions to the $\delta\mathcal{S}$ budget

The fact that the linear regression model based on $\delta\mathcal{S}_{QL}$ alone causes a skewed distribution of residuals in the extreme deciles of $\delta\mathcal{S}$, and that the residuals are not symmetric about the tails, suggests a crucial role for nonlinearities in determining the asymmetry of the distribution of $\delta\mathcal{S}$. Specifically, the distribution of residuals in the left extreme decile (decile 1) and the right extreme decile (decile 10) is not exactly equal and opposite, highlighting the different influences of $\delta\mathcal{S}_{NL}$ in these regions. While it has been previously suggested that the meridional nonlinear term alone can provide a satisfactory explanation for the observed asymmetry (skewness in their case) (Tamarin-Brodsky et al., 2019), it is unclear whether this is true for our case as well. Thus, the next natural question is to identify the nonlinear components driving the tails of $\delta\mathcal{S}$.

As in the previous section, we use the decision tree model to provide the leading explanatory variables for the extreme deciles 1 and 10 of $\delta\mathcal{S}$. Alongside nonlinear terms, we incorporate the pre-existing anomaly defined as the previous day's total lower tropospheric DSE anomaly, $\mathcal{S}'_{Tot,\,Lag1}$, as an explanatory variable. This accounts for the possibility that the drivers of large positive $\delta\mathcal{S}$ may differ depending on whether advection acts upon a pre-existing positive or negative $\mathcal{S}'_{Tot,\,Lag1}$ anomaly. Our results show that not only are distinct nonlinear components active in the left and right tails, but that the influence of these components also varies with the sign of $\mathcal{S}'_{Tot,\,Lag1}$. The decision tree model identifies the combinations of variables separating the extreme deciles of $\delta\mathcal{S}$ (Fig. 5) and leads to the following picture:

- When the **pre-existing anomaly is negative**, ($\mathcal{S}'_{Tot,\,Lag1} < 0.51\sigma$; we round the level split thresholds to second decimal) large values of advection – both positive and negative – are associated with the vertical term $w'\mathcal{S}'_z$, as seen in the leftmost branches terminating in Node 1 and Node 2. Positive values of $w'\mathcal{S}'_z > 0.16\sigma$ (extending to 5.38σ) correlate strongly

with the positive extremes of $\delta\mathcal{S}$ (Node 2), indicating a strong role of $w'\mathcal{S}'_z$ in dissipating pre-existing negative anomalies. Conversely, small negative values of $w'\mathcal{S}'_z < 0.16\sigma$ (with a very thin tail extending to -8σ) are associated with Node 1, indicating the potential role of confounding nonlinear variables.

- When the **pre-existing anomaly is positive** ($\mathcal{S}'_{Tot,\,Lag1} > 0.51\sigma$), large negative values of $\delta\mathcal{S}$ (Node 3 and Node 4) correspond to strong dissipation of a pre-existing positive anomalies. The model suggests two pathways, associated with *the horizontal terms* $v'\mathcal{S}'_y$ and $u'\mathcal{S}'_x$ respectively. $v'\mathcal{S}'_y < -0.24\sigma$ identifies the Negative extreme decile of $\delta\mathcal{S}$ with a 0.85 probability (Node 3). When its values exceed -0.24σ, but when $u'\mathcal{S}'_x > -0.93\sigma$, there is a a 0.75 probability of occurrence of the Negative extreme decile of $\delta\mathcal{S}$. Large positive values of $\delta\mathcal{S}$ (Node 5) correspond to strong amplification of a pre-existing positive anomaly, and are preferred when $v'\mathcal{S}'_y$ is approximately in its positive half, in combination with small magnitude values of either sign of $u'\mathcal{S}'_x$ (as seen from the histogram of δS at the node where $u'\mathcal{S}'_x$ is split into Nodes 4 and 5).

To contextualize the model results with raw data, we inspect the nonlinear-quasilinear relationships in the context of $\mathcal{S}'_{Tot,\,Lag1}$ (Figs. 5b, 5c, 5d). First, we see that the nonlinear terms achieve large magnitudes in the extremes of their quasi-linear counterparts, corroborating observations from Fig. 4d. Strong, nearly orthogonal relationships emerge in the meridional and vertical directions. Recollecting the strong association between different quasilinear terms and $\delta\mathcal{S}_{QL}$ (Supplementary Fig. S13), it is seen that $w'\mathcal{S}'_z > 0$ (first quadrant in Fig. 5d) is strongly associated with dissipation of a pre-existing negative anomaly (since $w'\bar{\mathcal{S}}_z > 0$ corresponds to $\delta\mathcal{S} > 0$). $w'\mathcal{S}'_z < 0$ has fewer instances associated with $\delta\mathcal{S}_{QL} < 0$, where it causes amplification of a pre-existing negative anomaly. On the other hand, $v'\mathcal{S}'_y$ does not observe as strong a relationship with $\delta\mathcal{S}_{QL}$ in spite of an apparently weak inverse relationship, when $\mathcal{S}'_{Tot,\,Lag1} < 0$. These observations justify the model's choice of $w'\mathcal{S}'_z$ in separating the extreme deciles of advection when $\mathcal{S}'_{Tot,\,Lag1} < 0$. When $\mathcal{S}'_{Tot,\,Lag1} > 0$, the horizontal nonlinear-quasilinear relationships seem more active than the vertical relationship, which is reflected in the model output. $v'\mathcal{S}'_y$ seems to weakly prefer being in phase with $\delta\mathcal{S}_{QL}$, although we see a large frequency of instances with $v'\mathcal{S}'_y > 0$ corresponding to $\delta\mathcal{S}_{QL} < 0$ conditions. $u'\mathcal{S}'_x$ shows no clear preference except for a strongly negative skew for both signs of pre-existing anomaly. These findings underscore the strong dependency of the sign and magnitudes of the nonlinear terms on their quasilinear counterparts, particularly at the extremes. The schematic in Fig. 6 contextualizes some of the phase relationships observed in Figs. 4a and 4b.

This analysis highlights the variety of eddy-eddy interactions involved, and the dependence of nonlinear contributions on the quasilinear part of the flow. However, further analysis is needed to establish a direct link between circulation patterns and the relationship between the quasilinear and nonlinear advection components.

5.1 Advection regimes

We propose a diagnostic framework to isolate advection regimes by analyzing the phase relationships between primary $\delta\mathcal{S}_{QL}$ and secondary $\delta\mathcal{S}_{NL}$ components alongside total advection, conditioned on the sign of pre-existing lower tropospehric DSE anomaly, $\mathcal{S}'_{Tot,\,Lag1}$. To this end, we construct a scatterplot of $\delta\mathcal{S}_{QL}$ versus $\delta\mathcal{S}_{NL}$, with data points color-coded by deciles of

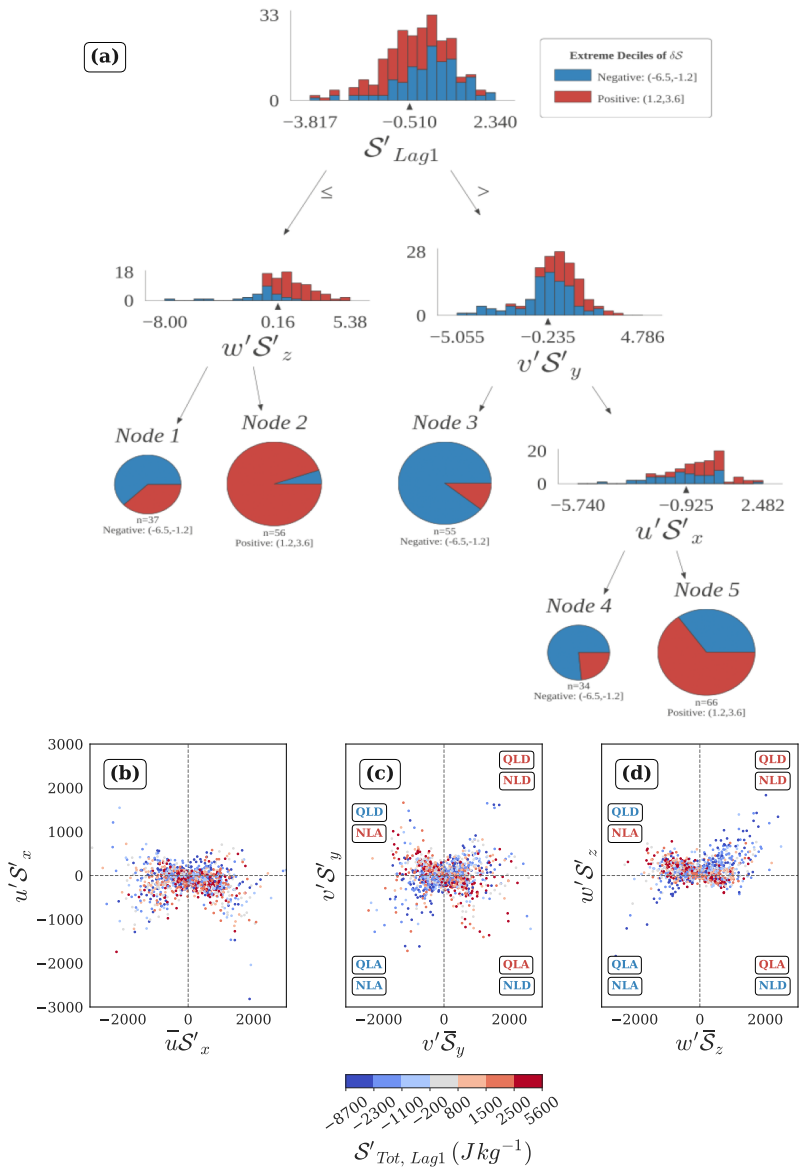

Figure 5. (a) Standardized decision tree model identifying the drivers of decile-based extremes of $\delta\mathcal{S}$ for April. The model achieves an F1-score of ~ 0.78, with early-stop criteria as follows: entropy decrease ≥ 0.03 at each level, $\geq 30\%$ sample size for level splitting, $\geq 15\%$ sample size for leaf nodes. (b-d) Nonlinear-quasilinear relationship in the zonal, meridional, and vertical directions, colored by quantile based categories of $\mathcal{S}'_{Tot,\,Lag1}$. All the daily mean advection terms are expressed in $J\,kg^{-1}$. Quadrants in (c) and (d) are labeled by the role of the advection components in amplifying or decaying the predominant sign of the pre-existing lower tropospheric DSE anomaly ($\mathcal{S}'_{Tot,\,Lag1}$) in that quadrant, with red for positive and blue for negative. For example, in (c), the first quadrant represents cases corresponding to quasilinear decay (QLD) and nonlinear decay (NLD) of a pre-existing negative anomaly by $v'\bar{\mathcal{S}}_y > 0$ and $v'\mathcal{S}'_y > 0$, respectively. Quadrants where advection aligns with the anomaly sign are labeled QLA (Quasilinear Amplification) and NLA (Nonlinear Amplification).

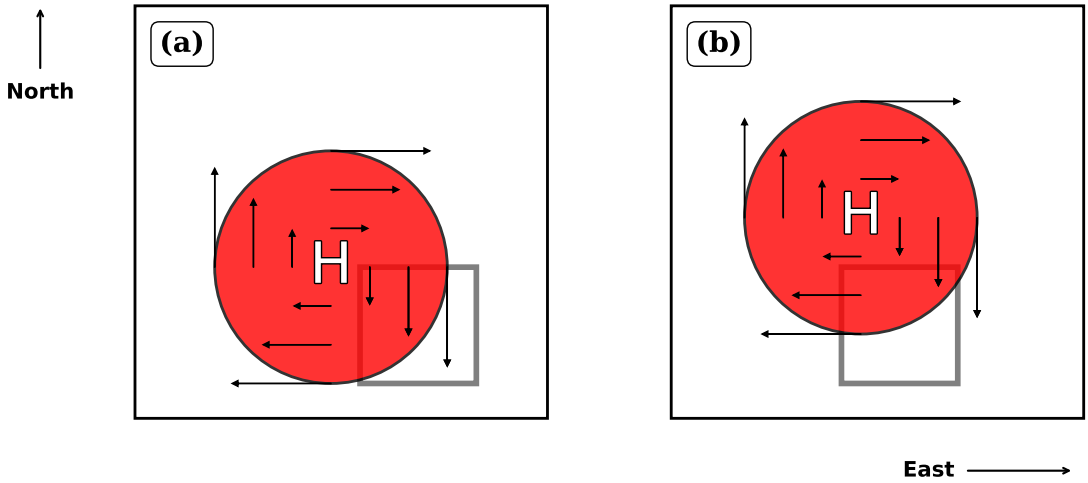

Figure 6. This schematic explains the eddy configurations responsible for some of the NL-QL phase interactions observed in Figs. 5a and 5b. The region of interest is shown by the transparent rectangle with thick grey borders. An anticyclonic warm core eddy associated with a high pressure anomaly is shown by the red circle, and the black solid arrows represent the circulation direction and strength which is proportional to the size of the arrows. The background conditions are given by: $\overline{DSE}_y < 0$, $\overline{DSE}_x \sim 0$. a) When an anticyclonic eddy is placed beyond the northwest corner of the region, the associated northerly flow ($v' < 0$) advects DSE' into the region due to $DSE'_y > 0$. Such a configuration causes nonlinear amplification of the DSE' of the region. At the same time, the meridional quasilinear advection term acts to dissipate the DSE' with upgradient transport of $\overline{DSE}$. Thus, the second quadrant in Fig. 5b is associated with quasilinear decay and nonlinear amplification by the meridional advection components. b) When the anticyclonic eddy is placed close to the north of the region, the associated easterly flow ($u' < 0$) advects DSE' out of the region due to $DSE'_x < 0$. Such a configuration causes nonlinear decay of the DSE' of the region, while the zonal quasilinear term acts to amplify the DSE' of the region due to strongly positive ($\bar{u} \gg 0$) mean zonal flow. Thus, the $\mathcal{S}'_{Tot,\,Lag1} > 0$ data points in the fourth quadrant of Fig. 5a are associated with quasilinear amplification and nonlinear decay of a pre-existing anomaly by the zonal advection components.

$\delta\mathcal{S}$, allowing us to visualize how combinations of quasilinear and nonlinear contributions map onto the phase and magnitude of net advection. We identify dominant nonlinear drivers for specific regions of the joint distribution, revealing systematic patterns in the interaction between the $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ components.

We find that for a given sign of pre-existing anomaly, the extreme deciles of $\delta\mathcal{S}_{QL}$, when combined with different phases of $\delta\mathcal{S}_{NL}$, usually correspond to distinct eddy configurations with respect to the region of interest. For instance, pairings of extreme deciles of $\delta\mathcal{S}_{QL}$ with the central deciles of $\delta\mathcal{S}_{NL}$ yield different eddy configurations than pairings with its extreme deciles. Pairings with extreme deciles of $\delta\mathcal{S}_{NL}$ is usually associated with a distinct nonlinear advection term modulating the effect of $\delta\mathcal{S}_{QL}$, while pairing with central deciles of $\delta\mathcal{S}_{NL}$ is associated with multiple nonlinear terms that are simultaneously of large magnitude and out of phase, justifying the small magnitude of $\delta\mathcal{S}_{NL}$. When $\delta\mathcal{S}_{QL}$ is in its central deciles, fewer pairings with the extreme deciles of $\delta\mathcal{S}_{NL}$ correspond to coherent eddy configurations. Thus, we identify distinct nonlinear components associated with most eddy configurations where high or low magnitudes of $\delta\mathcal{S}_{NL}$ interact to amplify or decay the contributions from the extreme deciles of $\delta\mathcal{S}_{QL}$. We refer to such combinations identified in the phase space as advection regimes.

5.1.1 Pre-existing positive anomaly

For daily cases with a pre-existing negative anomaly ($\mathcal{S}'_{Tot,\,Lag1} < 0$), we define Growth by positive daily advective tendency ($\delta\mathcal{S} > 0$) and Decay by negative daily advective tendency ($\delta\mathcal{S} < 0$), reflecting the phase combination of $\delta\mathcal{S}$ relative to the pre-existing positive phase of $\mathcal{S}'_{Tot,\,Lag1}$. As established earlier, $\delta\mathcal{S}$ is explained by $\delta\mathcal{S}_{QL}$ to the first order (Fig. 4a), but in the tails of $\delta\mathcal{S}$, both $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ can be comparable in magnitude (Fig. 4c). We study the joint distribution of $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$, and define Growth regimes for $\mathcal{S}'_{Tot,\,Lag1} > 0$ conditioned on $\delta\mathcal{S} > 0$ as follows:

- QL + NL Growth: Nonlinear amplification of Growth, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \gg 0$, i.e., when both $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$ are simultaneously in their positive extreme deciles
- NL Growth: Nonlinear Growth, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \sim 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 10th decile but $\delta\mathcal{S}_{QL}$ is in its central deciles (deciles 2 to 9)
- QL Growth: Quasilinear Growth, defined by $\delta\mathcal{S}_{QL} \gg 0$ and $\delta\mathcal{S}_{NL} \sim 0$, i.e., when $\delta\mathcal{S}_{QL}$ is in its 10th decile but $\delta\mathcal{S}_{NL}$ is in its central deciles (deciles 2 to 9)
- NL Saturated Growth: Nonlinear saturation of Growth, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \gg 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 1st decile and $\delta\mathcal{S}_{QL}$ is in its 10th decile

Similarly, we define Decay regimes for $\mathcal{S}'_{Tot,\,Lag1} > 0$ conditioned on $\delta\mathcal{S} < 0$ as follows:

- QL + NL Decay: Nonlinear amplification of Decay, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \ll 0$, i.e., when both $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$ are simultaneously in their negative extreme deciles

- NL Decay: Nonlinear Decay, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \sim 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 1st decile but $\delta\mathcal{S}_{QL}$ is in its central deciles (deciles 2 to 9)
- QL Decay: Quasilinear Decay, defined by $\delta\mathcal{S}_{QL} \ll 0$ and $\delta\mathcal{S}_{NL} \sim 0$, i.e., when $\delta\mathcal{S}_{QL}$ is in its 1st decile but $\delta\mathcal{S}_{NL}$ is in its central deciles (deciles 2 to 9)
- NL Saturated Decay: Nonlinear saturation of Decay, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \ll 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 10th decile and $\delta\mathcal{S}_{QL}$ is in its 1st decile

The above regimes are demarcated in the $\delta\mathcal{S}_{QL} - \delta\mathcal{S}_{NL}$ phase space for a pre-existing positive anomaly in Fig. 7, and the identified nonlinear combinations representing these regimes have been annotated in the figure. We observe that $v'\mathcal{S}'_y$ contributes to the nonlinear growth of a pre-existing positive anomaly, while $u'\mathcal{S}'_x$ plays an important role in the nonlinear saturation of its growth. These observations are in agreement with the description in Fig. 6. We also find that both the horizontal nonlinear terms drive nonlinear decay. During QL growth and decay regimes, there can be multiple nonlinear combinations representing large but out of phase nonlinear terms: For example, the QL Decay regime is characterized by two nonlinear combinations of comparable sample size, $u'\mathcal{S}'_x < 0, v'\mathcal{S}'_y > 0$ and $u'\mathcal{S}'_x < 0, v'\mathcal{S}'_y < 0, w'\mathcal{S}'_z > 0$. These findings underscore the variability in nonlinear-quasilinear dynamics and provide a framework for interpreting the evolution of advection structures modifying a pre-existing positive anomaly. The roles of these nonlinear drivers in driving growth and decay of $\mathcal{S}'_{Tot,\,Lag1} > 0$ are better clarified by inspecting the nonlinear-quasilinear scatterplots in the context of the phase and magnitude of $\delta\mathcal{S}$, $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ (Appendix Figs. A1, A2).

We find that a given advection regime consists of multiple nonlinear combinations. We have annotated the nonlinear combinations with the highest frequencies of occurrence in Fig. 7 – multiple regimes have been annotated if their frequencies are comparably large – and plotted the highest frequency combination in Fig. 8. The dominant nonlinear combinations associated with the advection regimes for $\mathcal{S}'_{Lag1} > 0$ conditions are clarified via composite maps (Fig. 8), and described in relation to the spatial configuration of eddy fields below.

1. NL Growth ($v'\mathcal{S}'_y \gg 0$): When the northerly winds associated with either a western/ northwestern anticyclonic disturbance engulf the region, the positive meridional gradient of DSE' intensifies – reaching about half the magnitude of the negative meridional gradient of $\overline{DSE}$. Simultaneously, the area-averaged northerly v' gains strength and enters its positive extreme decile. Such a configuration enhances the meridional anomalous advection of DSE' into the region, thereby amplifying the effect of positive $\delta\mathcal{S}_{QL}$ associated with such a configuration.

2. QL Growth ($u'\mathcal{S}'_x < 0, v'\mathcal{S}'_y > 0$): When comparable magnitudes of anomalous northerly and easterly winds associated with an anticyclonic disturbance operate in the eastern and southern halves of the region, under comparable magnitudes of $\mathcal{S}'_y > 0$ and $\mathcal{S}'_x < 0$, the easterly winds remove a significant part of the $\mathcal{S}'$ deposited by the northerly winds, rendering $\delta\mathcal{S}_{NL} \sim 0$. As a result, the net advection is a function of $\delta\mathcal{S}_{QL}$ alone in this configuration.

3. NL Saturated Growth ($u'\mathcal{S}'_x \ll 0$): When an anticyclone centered to the north of the region has advanced sufficiently into the region, the associated easterly winds pass through the southern half of the region, and the strength of $u' < 0$

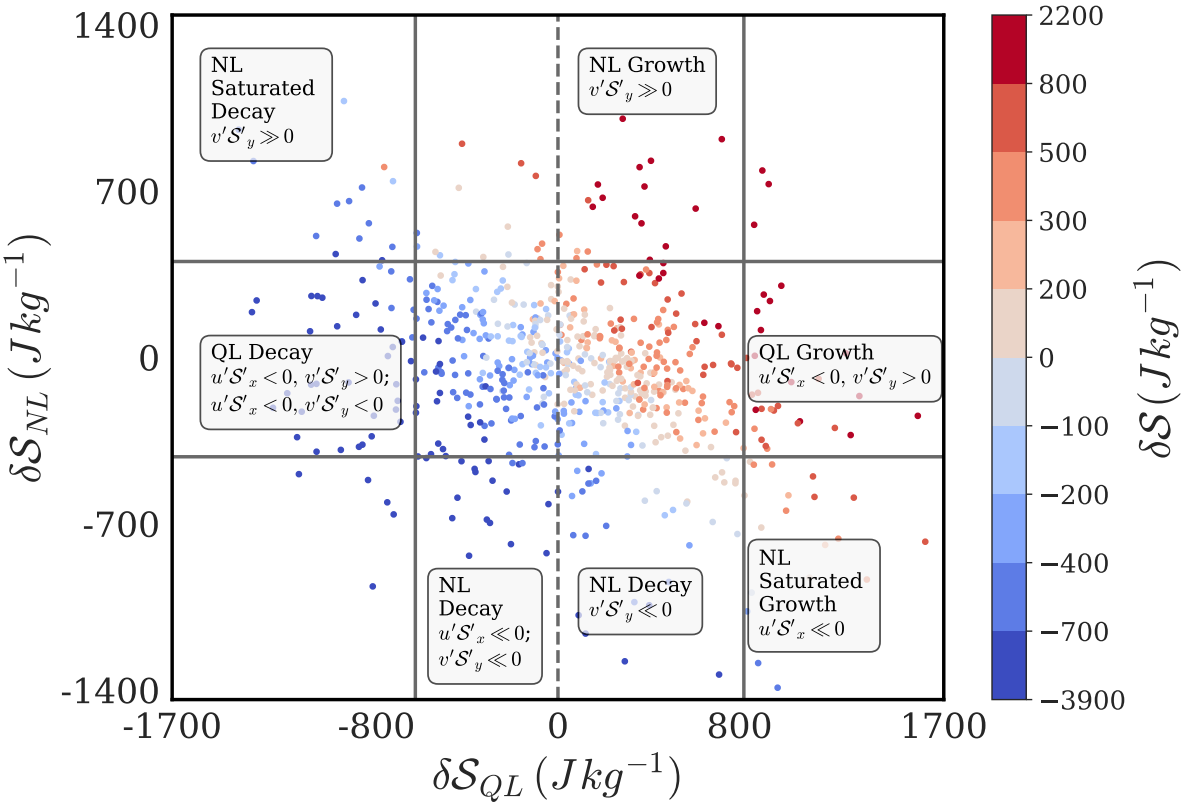

Figure 7. This figure shows the phase relationship between $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ colored by deciles of $\delta\mathcal{S}$, for pre-existing positive lower tropospheric DSE anomaly ($\mathcal{S}'_{Tot,\,Lag1} > 0$) conditions during April. Solid grey lines demarcate the 10$^{\text{th}}$ and 90$^{\text{th}}$ percentiles of $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$. Regime definitions are provided in section 5.1.1.

increases to more than half of the zonal mean flow. A strongly negative zonal gradient of DSE' is maintained due to the anomalously warm conditions associated with the anticyclone primarily acting on the western boundary of the region. Consequently, u' acts to remove a large part of the DSE' deposited by $\bar{u}$. The magnitude of v' is small in this configuration and does not yield a considerable magnitude of $v'\mathcal{S}'_y$ despite a large magnitude of positive $\mathcal{S}'_y$.

4. NL Decay ($u'\mathcal{S}'_x \ll 0$): When a cyclonic disturbance to the north/ west interacts primarily with the western boundary of the (anomalously warm) region of interest, the zonal eddy winds become aligned with the climatological zonal flow. Simultaneously, the zonal gradient of DSE' becomes positive and has a much larger magnitude than its meridional counterpart, because the northern boundary of the region is not affected similarly by the cyclonic disturbance. Coupled with anticyclonic winds to the south of the region, the zonal eddy winds gain strength, and drive the nonlinear decay of the pre-existing positive anomaly of DSE even without a large role of the quasilinear advection component. We note that the NL Decay regime has another component in Fig. 7, characterized solely by $v'\mathcal{S}'_y \gg 0$ within the positive half

of $\delta\mathcal{S}_{QL}$. This subset corresponds to 6 samples exhibiting extreme negative $\delta\mathcal{S}$. Upon inspection of its composite map, we find that it does not represent a coherent eddy configuration. We exclude this component from Fig. 8 and focus our interpretation on the $u'\mathcal{S}'_x \ll 0$ dominated branch of NL Decay described above.

5. NL Saturated Decay ($v'\mathcal{S}'_y \gg 0$): When a cyclone centered to the northwest of the anomalously warm region induces a negative meridional gradient of DSE', strong southerly winds act to deposit DSE' into the region. At the same time, the net quasilinear advection of DSE associated with the cyclone acts to deposit DSE' out of the region. This out of phase relationship is associated with nonlinear saturation of quasilinear decay.

6. QL Decay ($u'\mathcal{S}'_x < 0$, $v'\mathcal{S}'_y > 0$): The QL regime is revisited in the decay phase when a cyclone has advanced sufficiently into the region with anomalous conditions corresponding to strong and opposing zonal and meridional gradients of DSE', and strong westerly and southerly anomalous winds. In such a configuration, the nonlinear advection terms cancel out, rendering $\delta\mathcal{S}_{NL}$ of low magnitude.

We study the conditional probability distribution of deciles of $\mathcal{S}'_{Tot,\,Lag1}$ given the occurrence of the advection regimes active during $\mathcal{S}'_{Tot,\,Lag1} > 0$ conditions (Table 2). NL Growth and QL Growth regimes are predominantly concentrated in the central deciles of $\mathcal{S}'_{Tot,\,Lag1} > 0$, suggesting that linear and nonlinear growth mechanisms are most active during moderate positive anomalies. In contrast, the NL Saturated Growth regime preferentially occurs during more extreme positive values of $\mathcal{S}'_{Tot,\,Lag1}$. NL Decay and QL Decay regimes are most active during moderately positive $\mathcal{S}'_{Tot,\,Lag1}$ conditions, and NL Saturated Decay during neutral $\mathcal{S}'_{Tot,\,Lag1}$ conditions. Overall, most of the advection regimes are distributed across neutral, moderate and extreme positive deciles of $\mathcal{S}'_{Tot,\,Lag1}$, indicating that large Growth and Decay processes remain active throughout the lifecycle of $\mathcal{S}'_{Tot,\,Lag1} > 0$.

Advection Regime	**5**	**6**	**7**	**8**	**9**	**10**
NL Growth	0.19	0.25	0.19	0.12	0.19	0.06
NL Saturated Growth	0.13	0.20	0.13	0.00	0.07	0.47
QL Growth	0.11	0.22	0.11	0.33	0.11	0.11
NL Decay	0.14	0.10	0.19	0.19	0.24	0.14
NL Saturated Decay	0.00	0.54	0.15	0.15	0.15	0.00
QL Decay	0.06	0.25	0.28	0.16	0.09	0.16

Table 2. This table shows the conditional probability distribution of ranked deciles of $\mathcal{S}'_{Tot,\,Lag1}$ given the occurrence of an advection regime defined for $\mathcal{S}'_{Lag1} > 0$ conditions during April. We compute the conditional probability table by cross-tabulating the deciles of $\mathcal{S}'_{Tot,\,Lag1}$ with the advection regimes identified for $\mathcal{S}'_{Lag1} > 0$ conditions in Fig. 8, and then normalizing each row by corresponding row totals. Each cell value represents the conditional probability $P\left(\mathcal{S}'_{Lag1} \in D_i \,\middle|\, \text{regime} \in A_j\right)$ where D_i is the i^{th} decile of $\mathcal{S}'_{Tot,\,Lag1}$ and A_j is the advection regime on the j^{th} row.

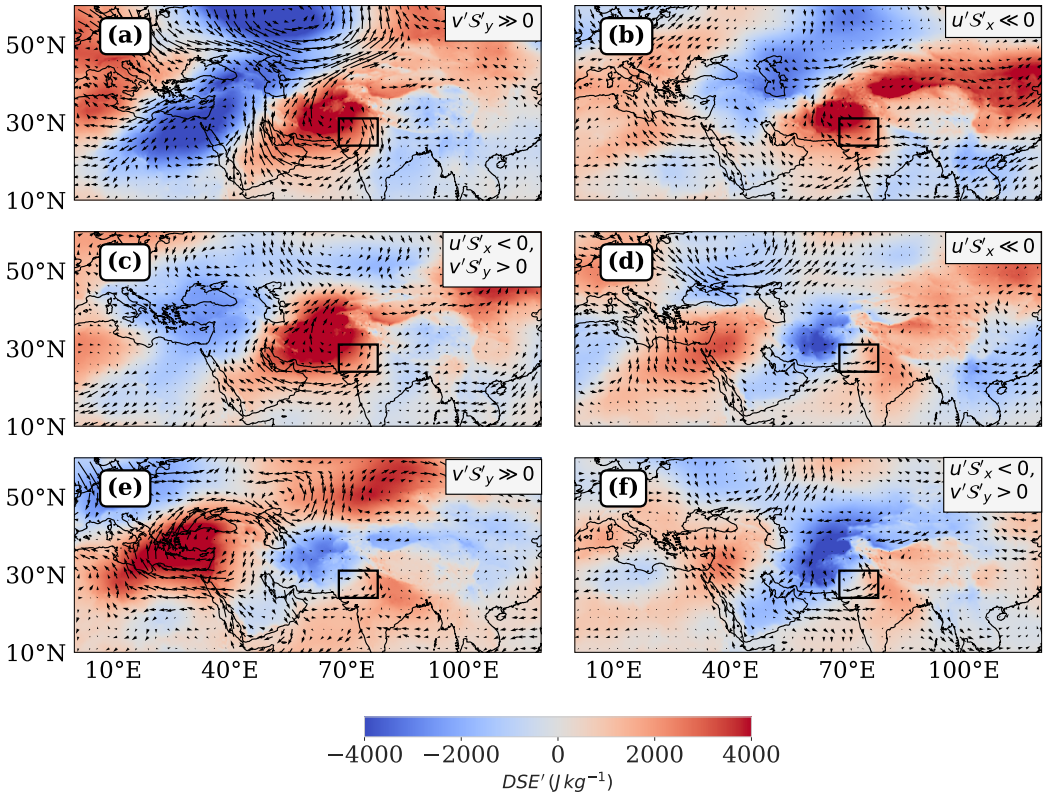

Figure 8. This figure shows the composite representations of horizontal eddy wind vectors and $\mathcal{S}'_{Tot}$ corresponding to the dominant nonlinear combination per advection regime amplifying and dissipating a pre-existing positive lower tropospheric DSE anomaly ($\mathcal{S}'_{Tot,\,Lag1} > 0$) during April, as defined in Fig. 7. The order of regimes is as follows: (a) NL Growth, (b) NL Saturated Growth, (c) QL Growth, (d) NL Decay, (e) NL Saturated Decay, (f) QL Decay. The Supplementary Table S4 clarifies the representativeness of an advection regime by the plotted nonlinear combination in the column "% Contribution". The color coding represents $\mathcal{S}'_{Tot}$ values ranging from -4000 to 4000 $J\,kg^{-1}$, with stronger hues indicating larger magnitudes.

5.1.2 Pre-existing negative anomaly

For daily cases with a pre-existing negative anomaly ($\mathcal{S}'_{Tot,\,Lag1} < 0$), we define Growth by negative daily advective tendency ($\delta\mathcal{S} < 0$) and Decay by positive daily advective tendency ($\delta\mathcal{S} > 0$), reflecting the phase combination of $\delta\mathcal{S}$ relative to the pre-existing negative phase of $\mathcal{S}'_{Tot,\,Lag1}$. We study the joint distribution of $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$, and define Growth regimes for $\mathcal{S}'_{Tot,\,Lag1} < 0$ conditioned on $\delta\mathcal{S} < 0$ as follows:

- QL + NL Growth: Nonlinear amplification of Growth, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \ll 0$, i.e., when both $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$ are simultaneously in their negative extreme deciles
- NL Growth: Nonlinear Growth, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \sim 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 1st decile but $\delta\mathcal{S}_{QL}$ is in its central deciles (deciles 2 to 9)
- QL Growth: Quasilinear Growth, defined by $\delta\mathcal{S}_{QL} \ll 0$ and $\delta\mathcal{S}_{NL} \sim 0$, i.e., when $\delta\mathcal{S}_{QL}$ is in its 1st decile but $\delta\mathcal{S}_{NL}$ is in its central deciles (deciles 2 to 9)
- NL Saturated Growth: Nonlinear saturation of Growth, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \ll 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 10th decile and $\delta\mathcal{S}_{QL}$ is in its 1st decile

Similarly, we define Decay regimes for $\mathcal{S}'_{Tot,\,Lag1} < 0$ conditioned on $\delta\mathcal{S} > 0$ as follows:

- QL + NL Decay: Nonlinear amplification of Decay, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \gg 0$, i.e., when both $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$ are simultaneously in their positive extreme deciles
- NL Decay: Nonlinear Decay, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \sim 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 10th decile but $\delta\mathcal{S}_{QL}$ is in its central deciles (deciles 2 to 9)
- QL Decay: Quasilinear Decay, defined by $\delta\mathcal{S}_{QL} \gg 0$ and $\delta\mathcal{S}_{NL} \sim 0$, i.e., when $\delta\mathcal{S}_{QL}$ is in its 10th decile but $\delta\mathcal{S}_{NL}$ is in its central deciles (deciles 2 to 9)
- NL Saturated Decay: Nonlinear saturation of Decay, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \gg 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 1st decile and $\delta\mathcal{S}_{QL}$ is in its 10th decile

The above regimes are demarcated in the $\delta\mathcal{S}_{QL} - \delta\mathcal{S}_{NL}$ phase space for a pre-existing negative anomaly in Fig. 9, and the identified nonlinear combinations representing these regimes have been annotated in the figure. When $\mathcal{S}'_{Tot,\,Lag1} < 0$, $u'\mathcal{S}'_x$ contributes to nonlinear amplification of growth, and $v'\mathcal{S}'_y$ aids nonlinear saturation of the growth of the negative anomaly, as observed in Fig. 9. In the decay phase, the roles of the nonlinear terms are reversed, and $u'\mathcal{S}'_x$ drives nonlinear saturation of decay while $v'\mathcal{S}'_y$ drives nonlinear amplification of decay. We see different nonlinear combinations active in the QL Decay regime; e.g. $u'\mathcal{S}'_x < 0,\ v'\mathcal{S}'_y < 0,\ w'\mathcal{S}'_z > 0$ and $u'\mathcal{S}'_x < 0, v'\mathcal{S}'_y > 0$. As before, the roles of these nonlinear terms in driving growth and decay of $\mathcal{S}'_{Tot,\,Lag1} < 0$ conditions are better understood by inspecting nonlinear-quasilinear scatterplots in the context of the phase and magnitude of $\delta\mathcal{S}$, $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ (Appendix Figs. A3, A4).

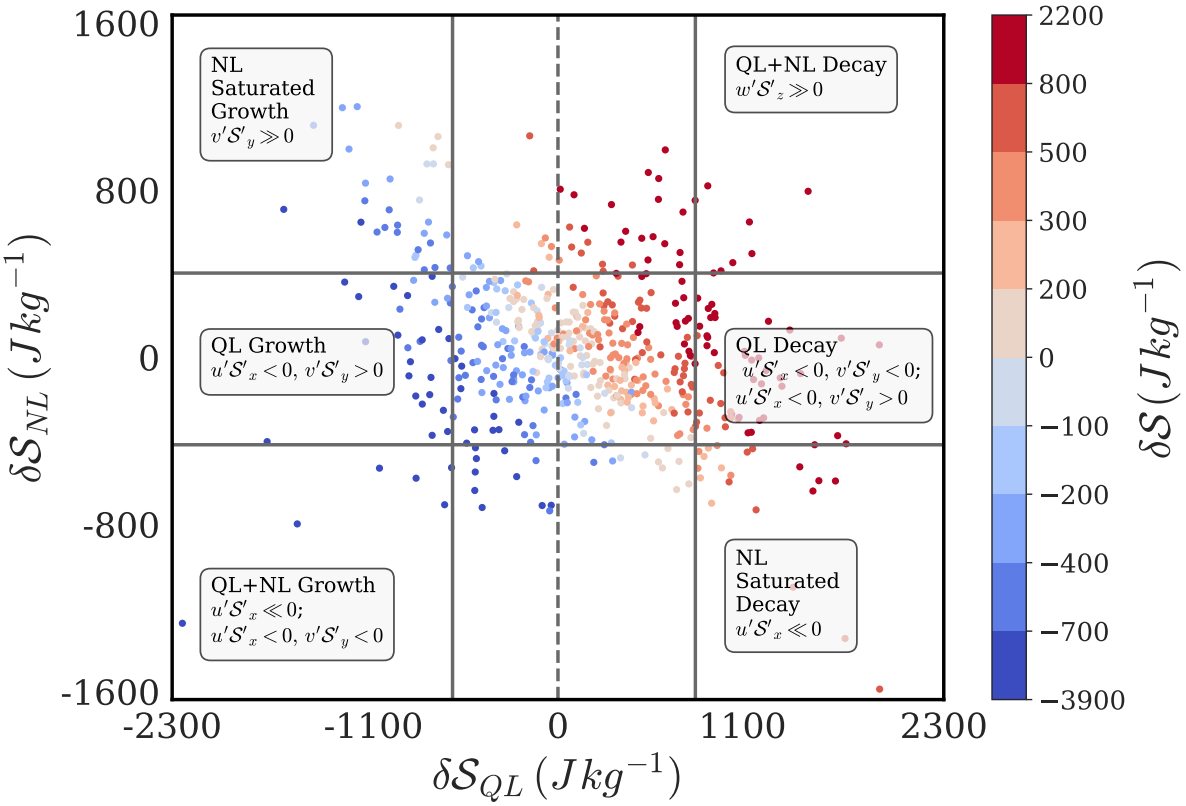

Figure 9. This figure shows the phase relationship between $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ colored by deciles of $\delta\mathcal{S}$, for pre-existing negative lower tropospheric DSE anomaly ($\mathcal{S}'_{Tot,\,Lag1} < 0$) conditions during April. Solid grey lines demarcate the 10th and 90th percentiles of $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$. Regime definitions are provided in section 5.1.2.

The circulations associated with the dominant nonlinear combinations representing the advection regimes for $\mathcal{S}'_{Lag1} < 0$ conditions (Fig. 9) are clarified via composite maps of eddy wind and DSE fields (Fig. 10), and described in relation to the spatial configuration of eddy fields below.

1. QL + NL Growth ($u'\mathcal{S}'_x \ll 0$): When a cyclonic disturbance is located to the northwest of the region, predominantly its easterly winds advect cold DSE into the region. This effect is intensified when the easterly flow is reinforced by easterly jet streaks, resulting in nonlinear amplification of the growth of a pre-existing negative DSE anomaly.

2. NL Saturated Growth ($v'\mathcal{S}'_y \ll 0$): When a cyclonic disturbance centered to the west induces a negative meridional gradient of DSE', the eddy winds over the region are dominated by southerly flow that advects DSE' into the region. In this configuration, positive nonlinear advection in the meridional direction acts in opposite phase to the negative quasilinear advection, leading to nonlinear saturation of the growth of the negative DSE anomaly.

3. QL Growth ($u'\mathcal{S}'_x < 0,\ v'\mathcal{S}'_y > 0$): When comparable magnitudes of eddy southerlies and westerlies associated with a cyclonic disturbance prevail in the eastern and southern halves of the region respectively, and coincide with similarly strong but opposite signed gradients of DSE' in respective directions, the southerly winds remove a substantial portion of the colder DSE' deposited by the westerly winds. This mutual cancellation of opposing contributions renders $\delta\mathcal{S}_{NL} \sim 0$. As a result, the net advection is governed solely by $\delta\mathcal{S}_{QL}$ in this configuration.

4. QL + NL Decay ($w'\mathcal{S}'_z \gg 0$): When northerly anticyclonic eddy winds are strongly coupled with subsiding vertical eddy winds, acting on a negative meridional and positive vertical gradient of DSE', the meridional nonlinear advection term anomalously cools the region, while the vertical nonlinear advection term anomalously warms it—with the latter exerting a dominant effect. The vertical nonlinear advection term acts in phase with net quasilinear advection, resulting in nonlinear amplification of the decay of a pre-existing negative DSE anomaly.

5. NL Saturated Decay ($u'\mathcal{S}'_x \ll 0$): We find a configuration representative of a trough-ridge couplet, with the southward dipping arm of cyclonically curved easterly winds of the ridge sweeping across the region. Coupled with a warm trough to the northwest of the region, the easterly arm advects cold DSE' into the region, while the northerly winds are associated with positive quasilinear advection of DSE'. Thus, the zonal nonlinear advection acts in the opposite phase of the quasilinear advection associated with such a configuration.

6. QL Decay ($u'\mathcal{S}'_x < 0,\ v'\mathcal{S}'_y < 0,\ w'\mathcal{S}'_z \gg 0$): The QL regime is revisited in the decay phase when northeasterly winds advect a southeastward horizontal gradient of DSE' to amplify the pre-existing negative DSE', but are opposed by large and positive nonlinear advection in the vertical, rendering $\delta\mathcal{S}_{NL}$ of low magnitude.

Thus, for a negative pre-existing anomaly, the roles of the meridional and zonal nonlinear terms were strongly reversed in contributions to amplification and dissipation. $v'\mathcal{S}'_y$ was out of phase with $\delta\mathcal{S}$ during both growth and decay since $\mathcal{S}'_y$ was of the same sign as $\bar{\mathcal{S}}_y$, and was out of phase with $\delta\mathcal{S}_{QL}$. $u'\mathcal{S}'_x$ strongly aided the amplification of the negative pre-existing anomaly since the easterly winds associated with cyclonic winds lined up with the climatological zonal winds, which were nearly always easterly. $w'\mathcal{S}'_z$ made significant contributions to the amplification phase.

We examine the conditional probability distribution of the deciles of $\mathcal{S}'_{Tot,\,Lag1}$ given the occurrence of the advection regimes active during $\mathcal{S}'_{Tot,\,Lag1} < 0$ conditions (Table 3). QL + NL Growth exhibits a strong preference for the left tail of $\mathcal{S}'_{Tot,\,Lag1}$, nonlinearly amplifying the growth of pre-existing extreme negative values of $\mathcal{S}'_{Tot,\,Lag1}$. NL Saturated Growth spans both extreme and moderate negative values of $\mathcal{S}'_{Tot,\,Lag1} < 0$, while QL Growth is most active during moderate values. QL + NL Decay and QL Decay are strongly concentrated in the left tail of $\mathcal{S}'_{Tot,\,Lag1}$, and NL Saturated Decay is more evenly distributed across the negative range of $\mathcal{S}'_{Tot,\,Lag1}$. Notably, all the Growth and Decay regimes here preferentially occupy moderate and extreme deciles of $\mathcal{S}'_{Tot,\,Lag1} < 0$ – a contrast to the regime behavior observed for $\mathcal{S}'_{Tot,\,Lag1} > 0$ conditions in section 5.1.1.

We repeat the entire analysis for the period of March, and arrive at distinct advection regimes associated with it (Appendix Figs. B1, B2). Overall, there are indications that such mapping between circulation patterns and the NL-QL phase space can effectively map different phases of the transition of a large scale disturbance over the region of interest. The conditional

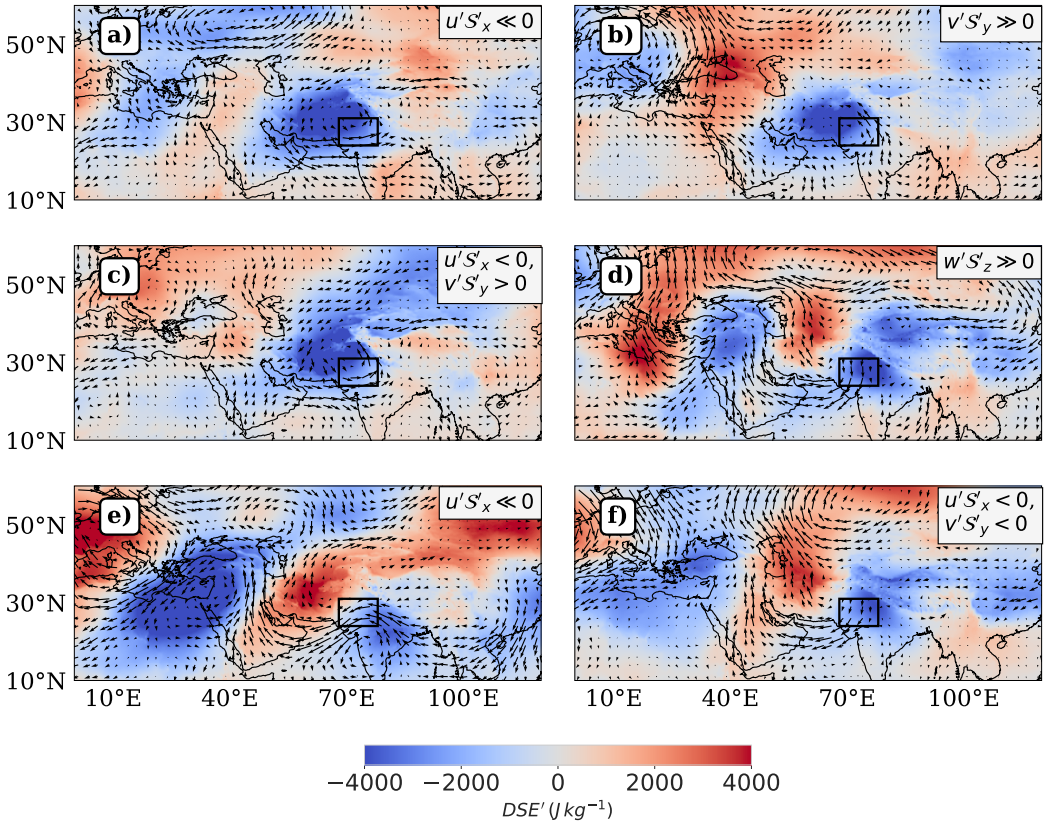

Figure 10. This figure shows the composite representations of horizontal wind anomaly vectors and $\mathcal{S}'_{Tot,\,Lag1}$ corresponding to the dominant nonlinear combination per advection regime amplifying and dissipating a pre-existing negative lower tropospheric DSE anomaly ($\mathcal{S}'_{Tot,\,Lag1} < 0$) during April, as defined in Fig. 7. The order of regimes is as follows: (a) QL + NL Growth, (b) NL Saturated Growth, (c) QL Growth, (d) QL + NL Decay, (e) NL Saturated Decay, f) QL Decay. The Supplementary Table S5 clarifies the representativeness of an advection regime by the plotted nonlinear combination in the column "% Contribution". The color coding represents $\mathcal{S}'_{Tot}$ values ranging from -4000 to $4000\ J\,kg^{-1}$, with stronger hues indicating larger magnitudes.

Advection Regime	1	2	3	4	5
QL + NL Growth	0.33	0.27	0.07	0.20	0.13
NL Saturated Growth	0.15	0.37	0.26	0.30	0.09
QL Growth	0.17	0.17	0.26	0.30	0.09
QL + NL Decay	0.71	0.00	0.29	0.00	0.00
NL Saturated Decay	0.14	0.29	0.24	0.10	0.24
QL Decay	0.34	0.28	0.19	0.16	0.03

Table 3. This table shows the same conditional probability distribution defined in Table 2, but filtered for cases with a negative pre-existing anomaly ($\mathcal{S}'_{Lag1} < 0$) during April. We compute the conditional probability table by cross-tabulating the deciles of $\mathcal{S}'_{Tot,\,Lag1}$ with the advection regimes identified for $\mathcal{S}'_{Lag1} < 0$ conditions in Fig. 10, and then normalizing each row by corresponding row totals.

probability distribution of deciles of $\mathcal{S}'_{Tot,\,Lag1}$ in Table 3f shows similar indications as to nonlinear and quasilinear regime preferences of the pre-existing negative anomaly states as Table 2f did for growth and decay of the pre-existing positive anomaly states.

6 Summary and Discussion

In this study, we have aimed to identify the contribution of atmospheric circulation to daily change in daily mean lower tropospheric DSE during March-April in a subtropical heatwave hotspot in South Asia. We choose to study DSE changes because we saw that while the daily temperature field is best explained by surface and radiative processes, *daily changes* in temperature are better explained by advection of DSE in the lower troposphere. Thus, understanding daily lower tropospheric DSE changes helps understand the variability in daily temperature changes over the region, particularly when the change is large.

We find that the velocity and DSE anomaly fields are vertically coherent, so the choice of this quantity allows us to focus on the impacts of upper tropospheric forcing on the lower tropospheric DSE budget, regardless of the form of the forcing itself (wave propagation, breaking, etc.,). Forcing due to balanced dynamics manifests itself as correlations between the different quasilinear advective terms, whereas such correlations may be absent for other kinds of circulations (such as monsoon intraseasonal oscillations) which may be more prevalent during or after the onset of the monsoon in this region.

We use decision trees to identify primary and secondary drivers of DSE variability because decision trees are both explainable and provide a relative quantification of the contribution of individual physical processes to DSE variability. Each model output is supported with an analysis of the raw data, ensuring the decision rules are both interpretable and grounded in robust underlying relationships.

We find that the "primary" quasilinear relationship capturing the inverse relationships between the meridional and vertical quasilinear components, and the meridional and zonal quasilinear components of advection effectively separates the presence

of eddies from incoherent flows, especially at large magnitudes. Smaller magnitudes of advection can represent eddy structures even when these relationships are not obeyed, but most instances with small magnitudes represent the absence of eddies.

Further, we find that the events that constitute the tail of the distribution of $\delta\mathcal{S}$ represent a rich variety of eddy configurations or "advection regimes" that are missing in the literature of temperature variability. We also find a strong presence of quasilinear regimes in the tails of $\delta\mathcal{S}$, and, counter to expectation, the presence of nonlinear regimes in the body of $\delta\mathcal{S}$. We obtain a circulation map for some of these regimes that are satisfactorily explained in terms of the horizontal configuration of the eddies.

The following list summarises the important observations and takeaways from our analysis:

1. We find that the quasilinear advective terms are important across all terciles of daily advective DSE changes. In contrast, the nonlinear terms are important primarily in the tails of the distribution. This observation is in line with previous studies, which propose that atmospheric macroturbulence is predominantly quasilinear – characterised by weak eddy-eddy interactions – and that eddy-eddy interactions are important for explaining the tail behavior of such macroturbulence. Further, the particular eddy-eddy interactions active in the tails of the distribution differ between extreme positive and negative advective DSE changes, depending on the sign of the pre-existing DSE anomaly.

2. We see that the combined vertical and zonal advection of DSE overwhelms the meridional advection of DSE, leading to an interesting situation where *southward* advection leads to an increase in lower tropospheric DSE, opposite to what one would expect if temperature variability was considered as a problem of one-dimensional mixing of the climatological equator-pole temperature gradient by eddies.

3. In the presence of eddies, the quasigeostrophically driven meridional-vertical wind coupling becomes associated with a zonal gradient of DSE' across the region of interest that persists during the transition of an eddy due to a preferred zonal direction of propagation (Fig. 3). The importance of vertical advection may be linked to the large static stability in our region of interest. The advection of DSE' by zonal mean flow is comparable in magnitude to the primary meridional and vertical terms due to large zonal flows prevalent during the considered times of the year. Thus, the zonal term has decisive control over the net effect of the quasilinear coupling involving all these three terms. This contribution appears to have been ignored in previous analyzes due to the focus on advection of isotherms by eddy flow fields, ignoring the zonal mean flow inhomogeneities between regions of a given latitude.

4. In addition to the large eddy winds in the meridional and vertical directions and a large zonal gradient of DSE' in the presence of an eddy, different eddy configurations render one or more of the following large: meridional gradient of DSE', vertical gradient of DSE' and zonal eddy winds (u'). Since the nonlinear advection terms result from the product of these terms in respective directions, different nonlinear terms become large under different eddy configurations and become important in the tails of the distribution of $\delta\mathcal{S}$. We note that eddies can also exist in the body of the distribution of daily advection of DSE, but we do not analyze them separately since they represent small advective daily changes in DSE.

5. $u'\mathcal{S}'_x$ is also neglected in studies considering nonlinear advection processes (Tamarin-Brodsky et al., 2019), presumably due to the Lagrangian approach which is not affected by the secondary gradients induced across a region. We find that zonal nonlinear advection term is mostly negative, irrespective of the sign of the pre-existing anomaly. This is because the signs of the anomalous zonal velocity and DSE'_x are opposite in the case of a northwestern anticyclone as well as a northwestern cyclone. The role of $u'\mathcal{S}'_x$ is important in driving nonlinear saturation of growth and nonlinear amplification of decay of a pre-existing positive anomaly. We find that $v'\mathcal{S}'_y$ aligns in phase with $u'\mathcal{S}'_x$ during instances of large negative advection that have nonlinear contributions, irrespective of the sign of the pre-existing anomaly; and it aligns with the phase of $\delta\mathcal{S}_{QL}$ during instances of large positive advection, nonlinearly amplifying both the growth of a pre-existing positive anomaly and the decay of a pre-existing negative anomaly. It is also found that $w'\mathcal{S}'_z$ is a strong driver of amplification and decay of negative anomalies, with a larger influence in causing decay as seen by its positively skewed distribution.

6. We identify different regimes of interaction between quasilinear and nonlinear advection terms, and show that each corresponds to distinct states of the pre-existing anomaly and flow patterns related to the phasing and location of upper-tropospheric eddies. We demonstrate that large $\delta\mathcal{S}$ can not only be driven by large quasilinear advection (as has been observed before), but also by large nonlinear advection when quasilinear advection is small. On the other hand, large values of $\delta\mathcal{S}$ can be driven purely by quasilinear advection when the individual nonlinear terms are large in magnitude but opposite in phase, yielding a negligible net effect; such cases are also consistently mapped to eddy configurations. Thus, we go beyond suggesting the qualitative importance of nonlinear advection to temperature variability, and identify the particular nonlinear terms that are important for different flow configurations.

7. We obtain the contribution of identified regimes to different stages of the anomaly lifecycle, and find that different configurations corresponding to the same sign of extreme advection occupy different stages. Such an observation across both signs of pre-existing anomaly and both periods of analysis, March and April, strongly suggests the existence of different pathways to extreme states of low tropospheric DSE.

Thus, we clarify the relationship between the energetics and flow patterns in our region of interest. We demonstrate that our approach allows for a more nuanced picture of DSE variability (and extreme variability in particular) than can be achieved by traditional composites which are insensitive to the contributions of individual quasilinear and nonlinear terms.

The decision tree based methodology for identifying terms governing the differences in sign of advection may be used as a reference to identify the primary dynamical relationships interacting with a region, and the regression based analysis for identifying regions in the distribution that obey different or supplementary sets of dynamics. While we have used this methodology to understand the distribution of $\delta\mathcal{S}$, we expect this approach should be useful in any such analysis of Reynolds-decomposed conservation laws, and could potentially also include the contributions due to the diabatic terms, if available.

Understanding the flow configurations that constitute the tail of $\delta\mathcal{S}$ might be used to identify contextual variables governing the interrelationships between terms of advection and provide a concise summary of the relationship between the statistics and

dynamics related to an observable. In the case of balanced large scale dynamics, our framework for identifying circulation patterns can help examine the structures of upper tropospheric flows interacting with a region.

Combined with the analysis mapping these circulations to stages of the anomaly lifecycle, our framework has the potential to identify persistent structures sustaining anomalies, and if there are distinct configurations associated with growth and decay. A similar approach was taken by Moron et al. (2010), who characterized distinct "weather types" associated with rainfall anomalies. However, there is still some work needed for using such a framework for relating circulations to T_{2m} variability, which is affected by other processes as previously noted.

In our case, we have also observed large durations of heatwaves associated with balanced dynamics during May, Jun, which will be studied using this framework in future work. Another useful practical application might be to explain periods of large observed skewness using such a framework. Further, the phenomenology of circulation structures obtained by studying the PDF of advection in this manner can be used towards a circulation-oriented evaluation of climate models and seasonal forecasts.

Appendix A: Nonlinear drivers of growth and decay of a pre-existing anomaly: April

A1 Growth of positive anomaly

We analyze the nonlinear terms driving the amplification and saturation phases of growth of a pre-existing positive anomaly (as defined in section 5.1.1) during April by inspecting the nonlinear-quasilinear relationship in each of the three directions (Fig. A1) in the context of the phase relationships between $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$, as follows:

- Growth of a pre-existing positive anomaly is strongly associated with positive quasilinear advective convergence (Figs. A1a, A1b and A1c). So, the problem reduces to inspecting the nonlinear drivers of the magnitude of $\delta\mathcal{S}_{NL}$ for each sign of its value.

- Figs. A1a, A1b and A1c suggest that for large magnitude of $\delta\mathcal{S}_{QL} > 0$, $v'\mathcal{S}'_y$ is consistently positive, while $u'\mathcal{S}'_x$ is consistently negative, and $w'\mathcal{S}'_z$ is positive but of low magnitude. Figs. A1d and A1e clarify that the opposite phases of $v'\mathcal{S}'_y$ and $u'\mathcal{S}'_x$ do not overlap, since $\delta\mathcal{S}_{NL} > 0$ when $v'\mathcal{S}'_y \gg 0$, and $\delta\mathcal{S}_{NL} < 0$ when $u'\mathcal{S}'_x \ll 0$.

- **Nonlinear Amplification**: When $\delta\mathcal{S}_{NL} > 0$, all three nonlinear terms act in phase with it. However, $v'\mathcal{S}'_y$ has the largest magnitude during most instances and likely drives $\delta\mathcal{S}_{NL} > 0$ when $\delta\mathcal{S} > 0$ and $\delta\mathcal{S}_{QL} > 0$.

- **Nonlinear Saturation**: When $\delta\mathcal{S}_{NL} < 0$, the largest nonlinear contributions come from $u'\mathcal{S}'_x$, and coincide with $\delta\mathcal{S}_{QL} > 0$. This identifies $u'\mathcal{S}'_x$ as the driver of saturation of growth of a pre-existing positive anomaly.

A2 Decay of positive anomaly

Using a similar approach, we analyze conditions associated with the decay of a pre-existing positive anomaly (Fig. A2).

- Decay of a pre-existing positive anomaly can be associated with positive as well as negative quasilinear advective convergence (Figs. A2a, A2b and A2c).
- **Nonlinear Amplification**: $u'\mathcal{S}'_x$ amplifies decay by acting in phase with $\delta\mathcal{S}_{QL} < 0$ (Fig. A2e).
- Focussing on the days corresponding to $\delta\mathcal{S}_{QL} > 0$, large magnitude of $v'\mathcal{S}'_y < 0$ (Fig. A2d) independently drives nonlinear decay.
- **Nonlinear Saturation**: Here, we focus on the days corresponding to $\delta\mathcal{S}_{QL} < 0$. $u'\mathcal{S}'_x$, $v'\mathcal{S}'_y$, and $w'\mathcal{S}'_z$ act in phase (> 0) to drive saturation of decay by opposing the phase of $\delta\mathcal{S}_{QL} < 0$ with $\delta\mathcal{S}_{NL} > 0$ (Figs. A2d, A2e, A2f). The magnitude of $v'\mathcal{S}'_y$ is largest, and is the likely driver of saturation of decay of a pre-existing positive anomaly.

A3 Growth of negative anomaly

We analyze the potential nonlinear advection terms driving the amplification and saturation phases of growth of a pre-existing negative anomaly, as follows.

- Growth of the pre-existing negative anomaly is strongly associated with negative quasilinear advective convergence barring a few low magnitude positive instances (Figs. A3a, A3b and A3c). So, the problem reduces to inspecting the nonlinear drivers of magnitude of $\delta\mathcal{S}_{NL}$ for each sign of its value.
- The meridional nonlinear term has large excursions on the positive side (A3d), while the zonal nonlinear term has large excursions on the negative side (A3e), suggesting that $v'\mathcal{S}'_y$ may act to saturate, while $u'\mathcal{S}'_x$ may act to amplify the growth of the anomaly.
- **Nonlinear Amplification**: For $\delta\mathcal{S}_{NL} < 0$, both the horizontal terms are in phase, but $u'\mathcal{S}'_x$ is consistently larger in magnitude (Figs. A3e, A3f), affirming its role in amplifying growth.
- **Nonlinear Saturation**: For $\delta\mathcal{S}_{NL} > 0$, both the horizontal nonlinear terms are predominantly positive. However, $v'\mathcal{S}'_y$ shows a consistently larger magnitude than $u'\mathcal{S}'_x$ (joint distribution not shown), affirming the role of $v'\mathcal{S}'_y$ in saturating growth of the negative anomaly.
- $w'\mathcal{S}'_z$ exhibits lower activity as compared to the horizontal nonlinear terms, indicating its limited influence in the growth phase.

A4 Decay of negative anomaly

Next, we analyze the nonlinear advection terms associated with the decay of a negative pre-existing anomaly (Fig. A4).

- Decay of a pre-existing negative anomaly is strongly associated with positive quasilinear advective convergence barring a few low magnitude negative instances (Figs. A4a, A4b and A4c). So, the problem reduces to inspecting the nonlinear drivers of magnitude of $\delta\mathcal{S}_{NL}$ for each sign of its value.

– $v'\mathcal{S}'_y$ makes large contributions to both sides of zero, but its negative excursions align with large positive values of $u'\mathcal{S}'x$ or $w'\mathcal{S}'_z$. This reduces its independent role in saturation.

– **Nonlinear Amplification**: $w'\mathcal{S}'_z$ is strongly associated with large instances of $\delta\mathcal{S}_{NL} > 0$ (Fig. A4f), corroborating the observations in Fig. 4d.

– **Nonlinear Saturation**: $u'\mathcal{S}'_x$ is consistently of large magnitude when $\delta\mathcal{S}_{NL} < 0$ and $\delta\mathcal{S}_{QL} > 0$ (Fig. A4e), emerging as a strong driver of saturation of decay of a pre-existing negative anomaly.

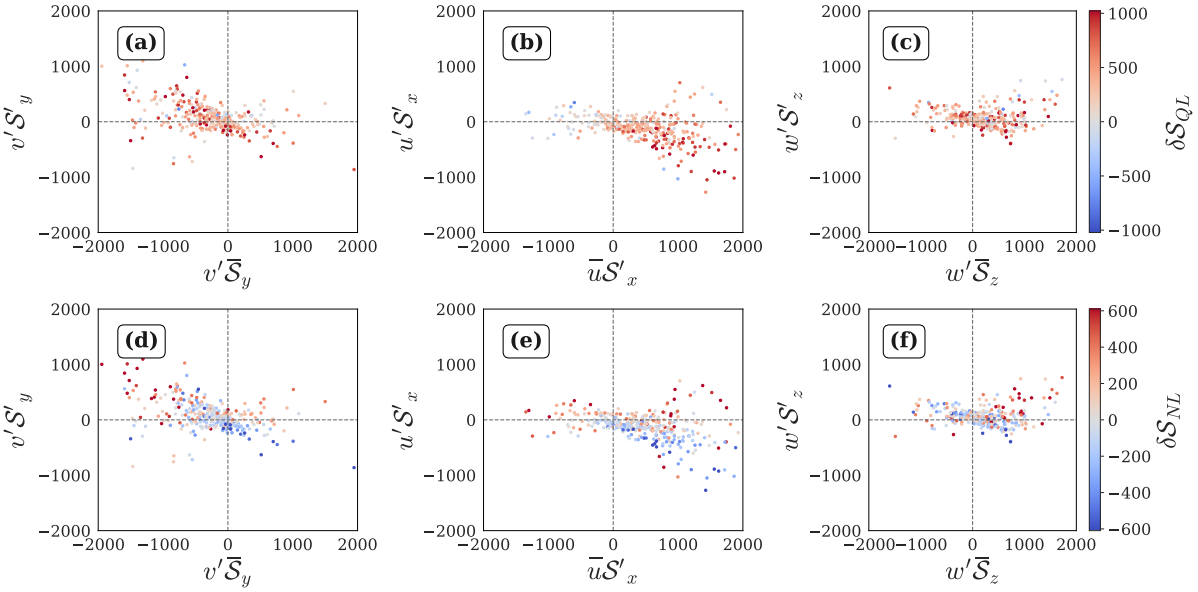

Figure A1. Nonlinear-quasilinear relationships for $\mathcal{S}'_{Tot,Lag1} > 0$ conditions, filtered for $\delta\mathcal{S} > 0$ days during April. (a-c) Quantiles of $\delta\mathcal{S}_{QL}$, (d-f) quantiles of $\delta\mathcal{S}_{NL}$, with blue for negative values and red for positive values and stronger colors indicating larger magnitudes. All quantities are expressed in $J\,kg^{-1}$.

Appendix B: Advection regimes: March

We repeated the analysis for March and identified the advection regimes for Growth and Decay of positive lower tropospheric DSE anomalies in Fig. B1, and for negative lower tropospheric DSE anomalies in B2, respectively.

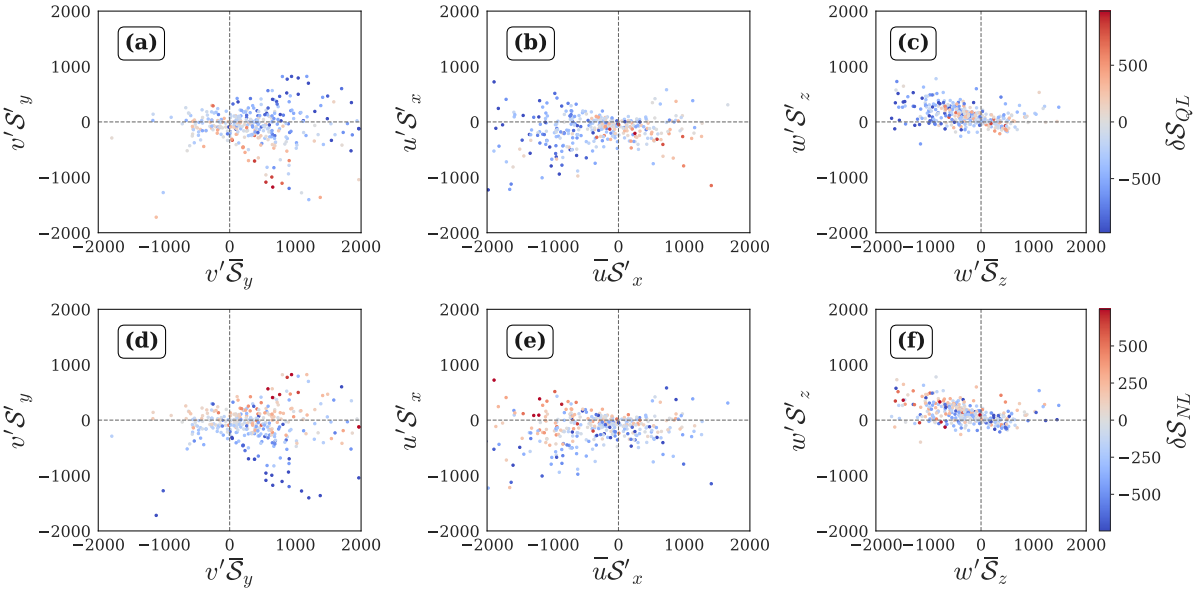

Figure A2. Nonlinear-quasilinear relationships for $\mathcal{S}'_{Tot,Lag1} > 0$ conditions, filtered for $\delta\mathcal{S} < 0$ days during April. (a-c) Quantiles of $\delta\mathcal{S}_{QL}$, (d-f) quantiles of $\delta\mathcal{S}_{NL}$, with blue for negative values and red for positive values, and stronger colors indicating larger magnitudes. All quantities are expressed in $J\,kg^{-1}$.

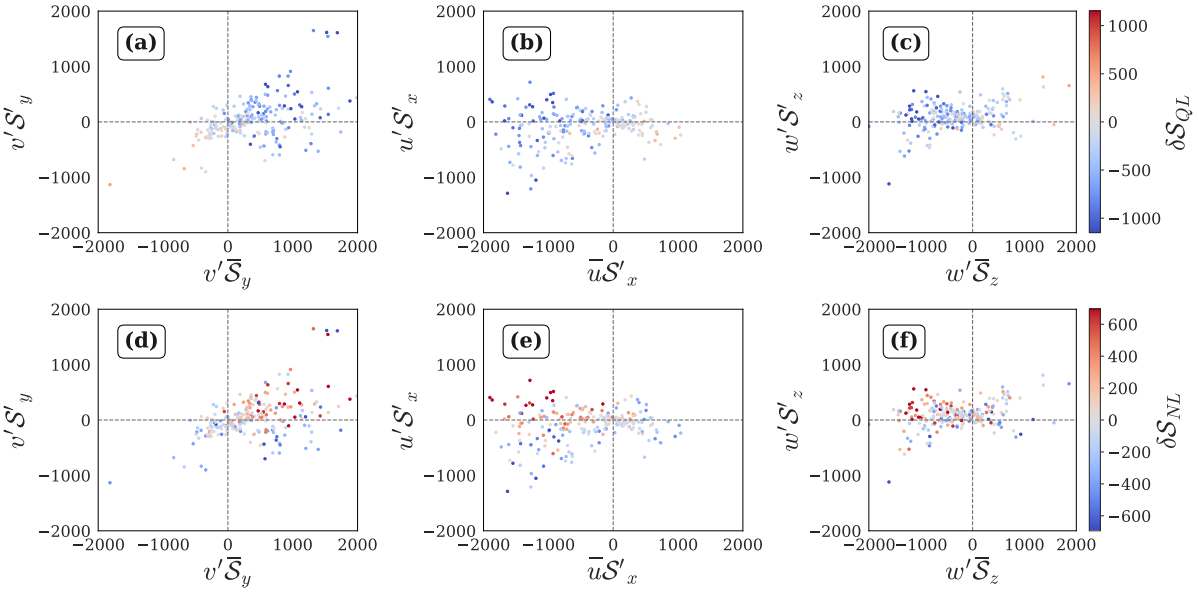

Figure A3. Nonlinear-quasilinear relationships for $\mathcal{S}'_{Tot,Lag1} < 0$ conditions, filtered for $\delta\mathcal{S} < 0$ days during April. (a-c) Quantiles of $\delta\mathcal{S}_{QL}$, (d-f) quantiles of $\delta\mathcal{S}_{NL}$, with blue for negative values and red for positive values, and stronger colors indicating larger magnitudes. All quantities are expressed in $J\,kg^{-1}$.

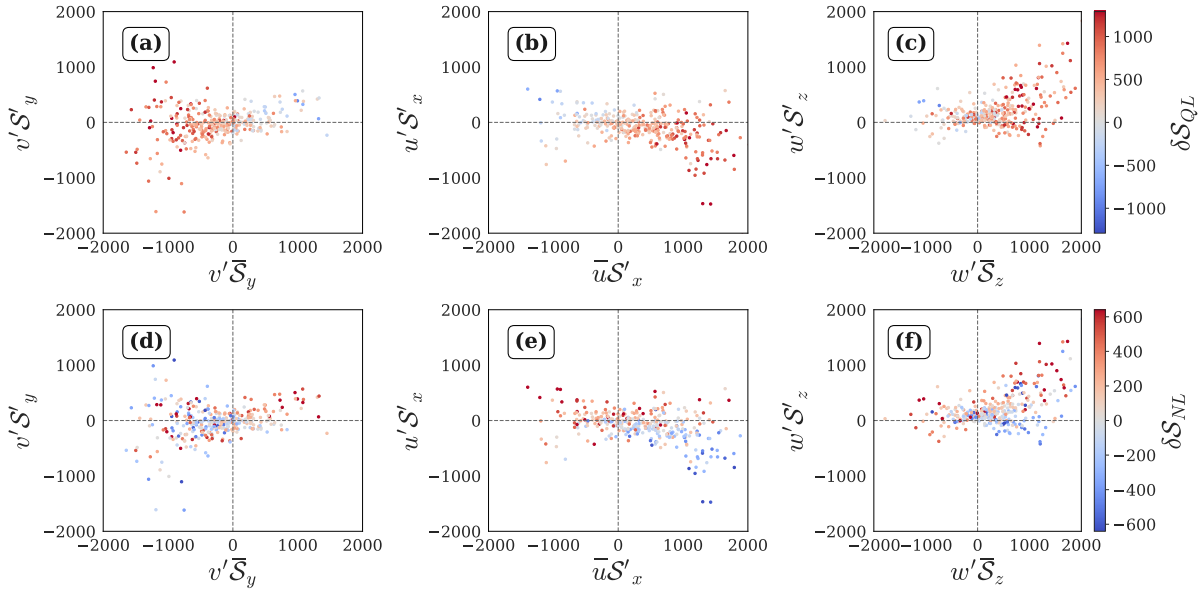

Figure A4. Nonlinear-quasilinear relationships for $\mathcal{S}'_{Tot,\,Lag1} < 0$ conditions, filtered for $\delta\mathcal{S} > 0$ days during April. (a-c) Quantiles of $\delta\mathcal{S}_{QL}$, (d-f) quantiles of $\delta\mathcal{S}_{NL}$, with blue for negative values and red for positive values, and stronger colors indicating larger magnitudes. All quantities are expressed in $J\,kg^{-1}$.

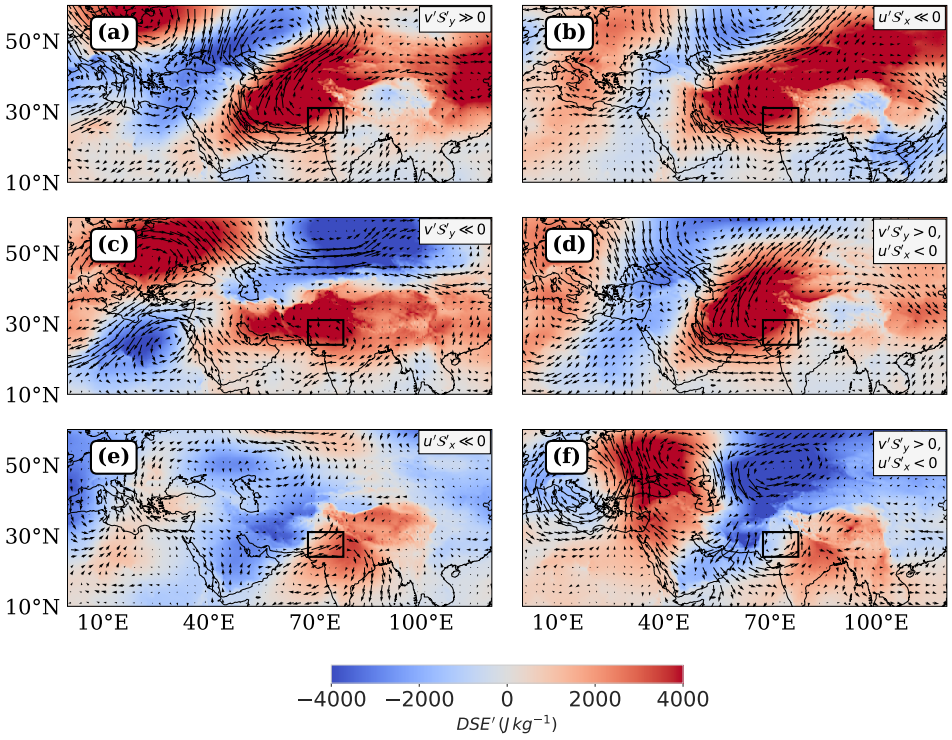

Figure B1. Composite representations of horizontal wind anomaly vectors and $\mathcal{S}'_{Tot}$ corresponding to advection regimes associated with Growth and Decay of $\mathcal{S}'_{Tot,\,Lag1} > 0$ conditions during March. The color coding represents $\mathcal{S}'_{Tot}$ values ranging from -4000 to 4000 $J\,kg^{-1}$.

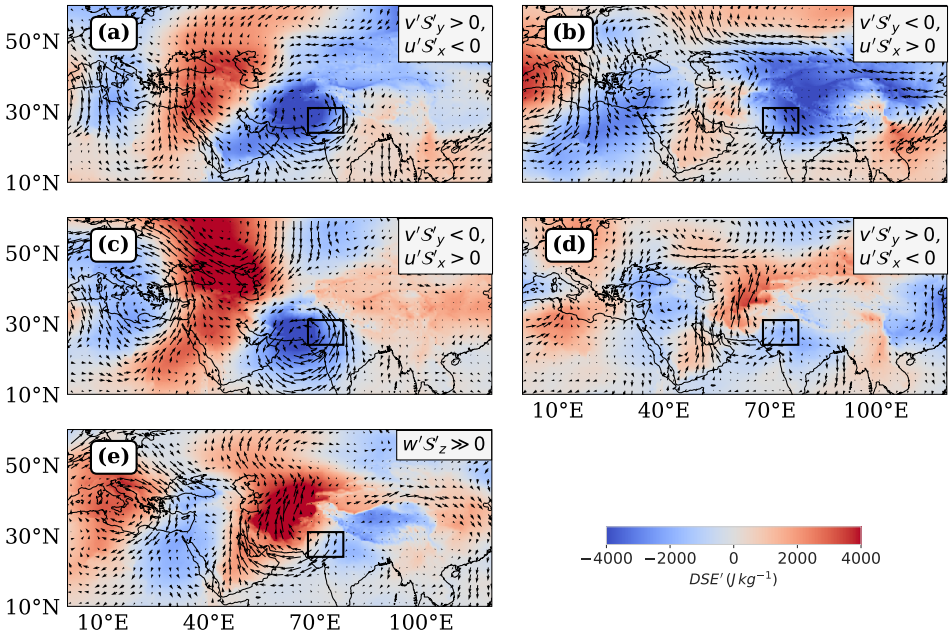

Figure B2. Composite representations of horizontal wind anomaly vectors and mass-weighted lower tropospheric DSE' corresponding to advection regimes amplifying and dissipating $\mathcal{S}'_{Tot,\,Lag1} < 0$ conditions during March. The color coding represents $\mathcal{S}'_{Tot}$ values ranging from -4000 to $4000\ J\,kg^{-1}$.

Author contributions. Both authors jointly conceptualized the project and contributed to the development of the methodology. Hardik M. Shah developed the computer code used for data preparation and analysis, applied statistical techniques for analysis and modeling, prepared data visualization and presentation ideas, and the original draft. Joy M. Monteiro acquired the funding for this project, acted as the project administrator and supervisor, arranged for the computing resources, provided mentorship, and helped with review and revision of the manuscript.

Competing interests. No competing interests are present.

Acknowledgements. This work was supported by the Department of Science and Technology grant DST/INT/ISR/P-40/2023 and Ministry of Earth Sciences Monsoon Mission 3 grant IITM/MM-III/ 2023/IND-4. We thank Vishal Dixit and Nili Harnik for helpful discussions shaping our methodology and interpretation.

References

Bouchet, F., Nardini, C., and Tangarife, T.: Kinetic Theory and Quasilinear Theories of Jet Dynamics, in: Zonal Jets: Phenomenology, Genesis, and Physics, edited by Galperin, B. and Read, P. L., Cambridge University Press, ISBN 978-1-107-04388-6, 2019.

Burrows, W. R., Benjamin, M., Beauchamp, S., Lord, E. R., McCollor, D., and Thomson, B.: CART Decision-Tree Statistical Analysis and Prediction of Summer Season Maximum Surface Ozone for the Vancouver, Montreal, and Atlantic Regions of Canada, Journal of Applied Meteorology and Climatology, 34, 1848 – 1862, https://doi.org/10.1175/1520-0450(1995)034<1848:CDTSAA>2.0.CO;2, 1995.

Chen, L. and Dirmeyer, P. A.: The relative importance among anthropogenic forcings of land use/land cover change in affecting temperature extremes, Climate Dynamics, 52, 2269–2285, https://doi.org/10.1007/s00382-018-4250-z, 2019.

Delsole, T. and Farrell, B. F.: The Quasi-Linear Equilibration of a Thermally Maintained, Stochastically Excited Jet in a Quasigeostrophic Model, Journal of the Atmospheric Sciences, 53, 1781–1797, https://doi.org/10.1175/1520-0469(1996)053<1781:TQLEOA>2.0.CO;2, 1996.

Dirmeyer, P. A., Chen, L., Wu, J., Shin, C. S., Huang, B., Cash, B. A., and others: Verification of land-atmosphere coupling in forecast models, reanalyses, and land surface models using flux site observations, Journal of Hydrometeorology, 19, 375–392, https://doi.org/10.1175/JHM-D-17-0152.1, 2018.

Domeisen, D. I. V., Eltahir, E. A. B., Fischer, E. M., Knutti, R., Perkins-Kirkpatrick, S. E., Schär, C., Seneviratne, S. I., Weisheimer, A., and Wernli, H.: Prediction and projection of heatwaves, Nature Reviews Earth & Environment, 4, 36–50, https://doi.org/10.1038/s43017-022-00371-z, 2023.

for Medium-Range Weather Forecasts (ECMWF), E. C.: ERA5: Data Documentation – Known Issues, https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation#ERA5:datadocumentation-Knownissues, accessed October 2025, 2025.

Francis, J. A. and Vavrus, S. J.: Evidence linking Arctic amplification to extreme weather in mid-latitudes, Geophysical Research Letters, 39, https://doi.org/https://doi.org/10.1029/2012GL051000, 2012.

Gagne, D. J., McGovern, A., and Brotzge, J.: Classification of Convective Areas Using Decision Trees, Journal of Atmospheric and Oceanic Technology, 26, 1341 – 1353, https://doi.org/10.1175/2008JTECHA1205.1, 2009.

Garfinkel, C. I. and Harnik, N.: The Non-Gaussianity and Spatial Asymmetry of Temperature Extremes Relative to the Storm Track: The Role of Horizontal Advection, J. Climate, 30, 445–464, https://doi.org/10.1175/JCLI-D-15-0806.1, 2016.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Fernández del Río, J., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357–362, https://doi.org/10.1038/s41586-020-2649-2, 2020.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2001.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Horton, D. E., Johnson, N. C., Singh, D., Swain, D. L., Rajaratnam, B., and Diffenbaugh, N. S.: Contribution of changes in atmospheric circulation patterns to extreme temperature trends, Nature, 522, 425–427, https://doi.org/10.1038/nature14550, 2015.

Hoyer, S. and Hamman, J.: xarray: N-D labeled arrays and datasets in Python, Journal of Open Research Software, 5, https://doi.org/10.5334/jors.148, 2017.

Hunter, J. D.: Matplotlib: A 2D graphics environment, Computing in Science & Engineering, 9, 90–95, https://doi.org/10.1109/MCSE.2007.55, 2007.

Kornhuber, K., Osprey, S., Coumou, D., Petri, S., Petoukhov, V., Rahmstorf, S., and Gray, L.: Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern, Environmental Research Letters, 14, 054 002, https://doi.org/10.1088/1748-9326/ab13bf, publisher: IOP Publishing, 2019.

Linz, M., Chen, G., and Hu, Z.: Large-Scale Atmospheric Control on Non-Gaussian Tails of Midlatitude Temperature Distributions, Geophysical Research Letters, 45, 9141–9149, https://doi.org/https://doi.org/10.1029/2018GL079324, 2018.

Linz, M., Chen, G., Zhang, B., and Zhang, P.: A Framework for Understanding How Dynamics Shape Temperature Distributions, Geophysical Research Letters, 47, e2019GL085 684, https://doi.org/https://doi.org/10.1029/2019GL085684, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL085684, 2020.

Loikith, P. C. and Neelin, J. D.: Short-Tailed Temperature Distributions over North America and Implications for Future Changes in Extremes, Geophysical Research Letters, 42, 8577–8585, https://doi.org/10.1002/2015GL065602, 2015.

Marston, J. B. and Tobias, S. M.: Recent Developments in Theories of Inhomogeneous and Anisotropic Turbulence, Annual Review of Fluid Mechanics, 55, 351–375, https://doi.org/10.1146/annurev-fluid-120720-031006, 2023.

Mayer, J., Mayer, M., and Haimberger, L.: Consistency and Homogeneity of Atmospheric Energy, Moisture, and Mass Budgets in ERA5, Journal of Climate, 34, 3955 – 3974, https://doi.org/10.1175/JCLI-D-20-0676.1, 2021.

McKinney, W. et al.: Data structures for statistical computing in python, in: Proceedings of the 9th Python in Science Conference, vol. 445, pp. 51–56, Austin, TX, 2010.

Met Office: Cartopy: a cartographic python library with a Matplotlib interface, Exeter, Devon, https://scitools.org.uk/cartopy, 2010 - 2015.

Miralles, D. G., Teuling, A. J., Heerwaarden, C. C. v., and Arellano, J. V.-G. d.: Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation, Nature Geoscience, 7, 345–349, https://doi.org/10.1038/ngeo2141, 2014.

Monteiro, J. M. and Caballero, R.: Characterization of Extreme Wet-Bulb Temperature Events in Southern Pakistan, Geophysical Research Letters, 46, 10 659–10 668, https://doi.org/https://doi.org/10.1029/2019GL084711, 2019.

Moron, V., Robertson, A. W., and Qian, J.-H.: Local versus regional-scale characteristics of monsoon onset and post-onset rainfall over Indonesia, Climate Dynamics, 34, 281–299, https://doi.org/10.1007/s00382-009-0547-2, 2010.

O'Gorman, P. A. and Schneider, T.: Recovery of Atmospheric Flow Statistics in a General Circulation Model without Nonlinear Eddy-Eddy Interactions, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL031779, 2007.

Pandey, K., Monteiro, J. M., and Natarajan, V.: An Integrated Geometric and Topological Approach for the Identification and Visual Analysis of Rossby Wave Packets, Monthly Weather Review, 148, 3139–3155, https://doi.org/10.1175/MWR-D-20-0014.1, publisher: American Meteorological Society, 2020.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Petoukhov, V., Eliseev, A. V., Klein, R., and Oesterle, H.: On statistics of the free-troposphere synoptic component: an evaluation of skewnesses and mixed third-order moments contribution to the synoptic-scale dynamics and fluxes of heat and humidity, Tellus A: Dynamic Meteorology and Oceanography, https://doi.org/10.1111/j.1600-0870.2007.00276.x, 2008.

Petoukhov, V., Rahmstorf, S., Petri, S., and Schellnhuber, H. J.: Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes, Proc Natl Acad Sci U S A, 110, 5336–5341, https://doi.org/10.1073/pnas.1222000110, 2013.

Petoukhov, V., Petri, S., Rahmstorf, S., Coumou, D., Kornhuber, K., and Schellnhuber, H. J.: Role of quasiresonant planetary wave dynamics in recent boreal spring-to-autumn extreme events, Proceedings of the National Academy of Sciences, 113, 6862–6867, https://doi.org/10.1073/pnas.1606300113, _eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.1606300113, 2016.

Prodhomme, C., Materia, S., Ardilouze, C., White, R. H., Batté, L., Guemas, V., Fragkoulidis, G., and García-Serrano, J.: Seasonal Prediction of European Summer Heatwaves, Clim Dyn, 58, 2149–2166, https://doi.org/10.1007/s00382-021-05828-3, 2022.

Rao, V. B., Rao, K. K., Mahendranath, B., Lakshmi Kumar, T. V., and Govardhan, D.: Large-scale connection to deadly Indian heatwaves, Quarterly Journal of the Royal Meteorological Society, 147, 1419–1430, https://doi.org/https://doi.org/10.1002/qj.3985, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3985, 2021.

Ratnam, J. V., Behera, S. K., Ratna, S. B., Rajeevan, M., and Yamagata, T.: Anatomy of Indian heatwaves, Scientific Reports, 6, 24 395, https://doi.org/10.1038/srep24395, 2016.

Rohini, P., Rajeevan, M., and Srivastava, A. K.: On the Variability and Increasing Trends of Heat Waves over India, Scientific Reports, 6, 26 153, https://doi.org/10.1038/srep26153, 2016.

Ruff, T. W. and Neelin, J. D.: Long tails in regional surface temperature probability distributions with implications for extremes under global warming, Geophysical Research Letters, 39, https://doi.org/https://doi.org/10.1029/2011GL050610, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011GL050610, 2012.

Sardeshmukh, P. D. and Sura, P.: Reconciling Non-Gaussian Climate Statistics with Linear Dynamics, Journal of Climate, 22, 1193–1207, https://doi.org/10.1175/2008JCLI2358.1, publisher: American Meteorological Society, 2009.

Schneider, T. and Walker, C. C.: Self-Organization of Atmospheric Macroturbulence into Critical States of Weak Nonlinear Eddy–Eddy Interactions, J. Atmos. Sci., 63, 1569–1586, https://doi.org/10.1175/JAS3699.1, 2006.

Schneider, T., Bischoff, T., and Plotka, H.: Physics of Changes in Synoptic Midlatitude Temperature Variability, J. Climate, 28, 2312–2331, https://doi.org/10.1175/JCLI-D-14-00632.1, 2014.

Seabold, S. and Perktold, J.: statsmodels: Econometric and statistical modeling with python, in: 9th Python in Science Conference, 2010.

Shannon, C. E.: A Mathematical Theory of Communication, Bell System Technical Journal, 27, 379–423, https://doi.org/https://doi.org/10.1002/j.1538-7305.1948.tb01338.x, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1948.tb01338.x, 1948.

Shepherd, T. G.: Atmospheric Circulation as a Source of Uncertainty in Climate Change Projections, Nature Geoscience, 7, 703–708, https://doi.org/10.1038/ngeo2253, 2014.

Shepherd, T. G.: Climate science: the dynamics of temperature extremes, Nature, 522, 425–427, https://doi.org/10.1038/nature14540, 2015.

Svirsky, A., Herbert, C., and Frishman, A.: Statistics of Inhomogeneous Turbulence in Large-Scale Quasigeostrophic Dynamics, Physical Review E, 108, 065 102, https://doi.org/10.1103/PhysRevE.108.065102, 2023.

Tamarin-Brodsky, T., Hodges, K., Hoskins, B. J., and Shepherd, T. G.: A Dynamical Perspective on Atmospheric Temperature Variability and Its Response to Climate Change, Journal of Climate, 32, 1707–1724, https://doi.org/10.1175/JCLI-D-18-0462.1, 2019.

Wei, W., Yan, Z., and Jones, P. D.: A decision-tree approach to seasonal prediction of extreme precipitation in eastern China, International Journal of Climatology, 40, 255–272, https://doi.org/https://doi.org/10.1002/joc.6207, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.6207, 2020.

White, R., Kornhuber, K., Martius, O., and Wirth, V.: From Atmospheric Waves to Heatwaves: A Waveguide Perspective for Understanding and Predicting Concurrent, Persistent and Extreme Extratropical Weather, Bulletin of the American Meteorological Society, 103, 1–35, https://doi.org/10.1175/BAMS-D-21-0170.1, 2021.

Xu, W., Ning, L., and Luo, Y.: Wind Speed Forecast Based on Post-Processing of Numerical Weather Predictions Using a Gradient Boosting Decision Tree Algorithm, Atmosphere, 11, https://doi.org/10.3390/atmos11070738, 2020.

Zhang, X., Jiang, H., Jin, J., Xu, X., and Zhang, Q.: Analysis of acid rain patterns in northeastern China using a decision tree method, Atmospheric Environment, 46, 590–596, https://doi.org/https://doi.org/10.1016/j.atmosenv.2011.03.004, 2012.

The role of synoptic circulations in lower-tropospheric ~~DSE~~ dry static energy variability over a South Asian heatwave hotspot

Hardik M. Shah[1] and Joy M. Monteiro[1,2]

[1]Department of Earth and Climate Science, Indian Institute of Science Education and Research Pune, Pune, Maharashtra, 411008, India

[2]Department of Data Science, Indian Institute of Science Education and Research Pune, Pune, Maharashtra, 411008, India

Correspondence: Hardik M. Shah (hardik.shah.reach@gmail.com)

Abstract. We examine the role of the synoptic-scale circulation in the distribution of ~~day-day~~ daily changes of 600-900 hPa dry static energy (~~DSE~~DSE) in a heatwave hotspot in northwest South Asia. Using a combination of linear regression and decision trees, we identify how the quasilinear (mean-eddy) and nonlinear (eddy-eddy) components of the flow contribute to different parts of this distribution. We show that the presence of synoptic eddies leads to strong correlations in the quasilinear components due to quasigeostrophy, allowing us to identify periods of upper tropospheric eddy activity. We show that the synoptic eddies induce a zonal quasilinear component which plays an important role in governing the magnitude and sign of ~~DSE~~ daily DSE changes. Nonlinear components are observed to play an important role in the tails of this distribution, and we show that the specific nonlinear components that are involved ~~depends~~ depend on the phase of growth or decay of ~~DSE and the background DSE~~ DSE and of the pre-existing DSE anomaly. We identify energetically distinct configurations involved in the tails of this distribution, and identify eddy configurations corresponding to each of these energetic configurations. Our analysis thus provides a discrete set of "regimes" which can be used to classify extreme ~~DSE~~ daily DSE changes, and provides a more nuanced approach to compositing extreme events which is sensitive to the dynamics underpinning each event.

1 Introduction

~~The observed regional distributions of near-surface temperature exhibit significant departures from being a Gaussian in the tails, and recent work has expanded our understanding of the likely processes involved in shaping these distributions~~ Recent work has expanded o (**???**). Due to the known association of large scale eddies with midlatitude weather extremes (**??**), one of the questions of interest has been the role of large-scale dynamics in driving different moments characterizing these distributions (**???**). Given that our ability to understand the impacts of global warming ~~are~~ is currently limited by our ability to understand changes in atmospheric circulation (**??**), understanding the links between circulation and temperature variability in the current climate is important to diagnose the fidelity of climate models in representing these links, leading to improved confidence in climate projections.

Earlier attempts to understand the link between large-scale dynamics and temperature variability used quasilinear models of atmospheric dynamics to address free tropospheric temperature variability, where large-scale dynamics dominate. The choice

of quasilinear models ~~is~~ was presumably governed by ~~the~~ their success in reproducing atmospheric flow statistics (**??**). However, quasilinear models without nonlinear processes predict Gaussian temperature distributions (**?**), and are therefore unlikely to provide a reasonable explanation for the observed temperature distributions. While ~~it is possible for quasilinear models to~~ quasilinear models can produce skew by the incorporation of multiplicative noise processes (**?**), **?** showed that the inclusion of eddy-eddy interactions can qualitatively reproduce the observed patterns of temperature distributions. There has also been a strong indication that skewness is related to distance from the jetstream via the meridional movement of synoptic systems. We note that the importance of eddy-eddy interactions is not limited to understanding temperature variability, but has been studied in a variety of geophysical contexts (**????**). More recent work has tried to build on these initial insights, by using climate models for verifying the importance of nonlinear processes in generating non-Gaussian tails (**?**), and formulating nonlinear approximation schemes (**?**).

While such literature has helped expand our understanding of free-tropospheric temperature variability, the widely employed quasilinear models ignore the role played by the zonal mean flow. This choice may have been made due to their interest in the deformation flow field, but this choice makes it problematic to apply their results in an Eulerian framework, which is more relevant for regional temperature variability. Understanding near-surface temperature variability comes with the additional challenge of disentangling the effects of atmospheric macroturbulence and local processes such as land-atmosphere interactions and boundary layer feedback (**???**). Lately, there have been frameworks proposed for quantifying the contribution of different proximate processes towards driving temperature tendency in different deciles of near-surface temperature (**?**).

Studies which focus on the tails of near-surface temperature distribution have established the association between heatwaves and dynamical processes such as blocking events and quasi-resonant amplification (**???**). The synoptic patterns associated with such dynamical processes are usually studied using composites (**??**). While this strategy is useful for understanding "average" conditions, it can obscure the diversity of pathways that could lead to the outcomes of interest. For example, the composite picture presented in **?** while studying extreme wet-bulb temperature events was found to average over at least two mechanisms, which depended on the phase of the wave packet over the region of interest (**?**).

Since it is unlikely that we will be able to predict the occurrence of extreme events on seasonal timescales, it has been suggested that subseasonal-to-seasonal predictions might benefit from the prediction of waveguides rather than wave driven extremes (**?**). Furthermore, the relationship between upper tropospheric forcing and near-surface response is not always clear (**?**), and it might be a useful exercise to understand the impacts of Rossby wave packets on lower tropospheric quantities that are directly related to near-surface temperature. Such an exercise might allow evaluating seasonal "propensity of extremes" (**?**) directly using such quantities instead. Thus, understanding the variability of such quantities is not only interesting from an atmospheric turbulence viewpoint but also from a seasonal prediction viewpoint. One possible approach (which we take in this study) to addressing the question of understanding near-surface temperature variability and extremes could comprise the following steps:

1. Identifying a lower tropospheric quantity that is highly associated with near-surface temperature and is strongly influenced by the atmospheric circulation

2. Identifying quasilinear and nonlinear contributions to different parts of the distribution of this quantity

3. Characterizing the different pathways to extreme values of this quantity and relating them to characteristics of the circulation

In this study, we identify the lower tropospheric (600-900 hPa) dry static energy ~~$\mathcal{S}$~~ $\mathcal{S}_{Tot}$ (where the subscript "Tot" stands for total), as a suitable quantity whose daily changes are both highly correlated with near-surface temperature and have a significant contribution due to large-scale advection in our region of interest, a heatwave hotspot in South Asia. We study the variability of advection-driven daily changes in ~~$\mathcal{S}$~~ $\mathcal{S}_{Tot}$, during March-April, which are a part of the heatwave season in this region. We study the contribution of quasilinear and nonlinear advective fluxes to the distribution of ~~changes in $\mathcal{S}$. We then~~ advection driven daily changes in $\mathcal{S}_{Tot}$. We use an interpretable machine learning algorithm, the decision tree, to develop a nuanced picture of combinations of advective contributions that lead to extreme ~~changes in $\mathcal{S}$~~ daily changes in $\mathcal{S}_{Tot}$ and show how a combination of these results and compositing leads to a richer description of the different ways in which the upper tropospheric circulation affects ~~changes in $\mathcal{S}$~~daily changes in $\mathcal{S}_{Tot}$. Finally, we summarize our work with ~~a comprehensive~~ an analysis of the lifecycle of ~~changes in $\mathcal{S}$~~daily changes in $\mathcal{S}_{Tot}$, highlighting how ~~the circulation leads to~~ circulation drives the accumulation, saturation, and ventilation of ~~$\mathcal{S}$~~ $\mathcal{S}_{Tot}$ over the region.

2 Data and Methods

2.1 Data

We ~~utilized~~ have used the European Center for Medium-Range Weather Forecasts Reanalysis version 5 (ERA5; **?**) reanalysis dataset (0.25° x 0.25° resolution) for analysing daily mean quantities. The ~~dataset is representative of the synoptic scale atmosphere, and is well-suited for understanding the advection processes driving large scale variations over our region of interest. For our analysis, we used the 2m-temperature (T2m) to represent near-surface temperature, and incorporated the 4D variables of temperature, geopotential, and the three components of velocity to study dry static energy. The~~ period of analysis spans the months of March and April, from 1980 to ~~2022.~~ 2022, with the exclusion of 82 dates due to data quality issues. We use unevenly s

2.2 Methodology

Our region of interest is the ~~northwest Indian~~ South Asian heatwave hotspot (**??**), defined as the area enclosed between ~~25^{o}N and 31^o~~25^{o}N and 31^{o}N, and ~~68^{o}E and 78^o~~68^{o}E and 78^{o}E. We use dry static energy (DSE) for tracking the energy content of the atmospheric parcels. Since DSE is conserved for dry, large scale adiabatic flows, it is particularly suited for diagnosing energy advection driven by the resolved states of the large scale atmosphere. Using Reynolds decomposition, ~~the DSE~~ we decomposed the DSE and velocity fields ~~were decomposed~~ into daily climatology and anomaly components ($X = \bar{X} + X'$) ~~to distinguish~~

~~between background states and transient variations~~where the climatology represents the background state and anomalies represent transien

Daily climatology was computed as a 10-day rolling mean of daily mean computed over the period of analysis.

Our analysis ~~is focussed~~ focuses on the volume ~~enclosed between the pressure levels 600 hPa and 900 hPa over the defined~~ ~~area. The~~ of the lower troposphere bounded vertically between 600 hPa and 900 hPa, and horizontally by the spatial extent of the region of the dry static energy content of air parcels contained in this volume is ~~represented~~ given by $\mathcal{S}_{Tot}$, ~~and the~~ and daily changes in $\mathcal{S}_{Tot}$ are given The daily DSE convergence into this volume is calculated as the mass-weighted advection of DSE into ~~this layer.~~

it, given by $\delta\mathcal{S}$. The local conservation law for DSE is expressed in equation 1, with ϵ representing the terms involving unresolved or parameterized energy fluxes. Starting with the point conservation law for DSE,

$$\frac{\partial(\mathrm{DSE})}{\partial t} = -\nabla \cdot (\mathbf{v}\,\mathrm{DSE}) + \epsilon \tag{1}$$

We analyzed both the divergent and non-divergent components of the RHS; Even though the divergent term was not small, advective co

Upon neglecting the mass conservation terms ~~, the equation reduces to an advection~~ and discretizing the quantities, equation ~~(??). Corresponding quantities representing~~ 1 reduces to the advection form in equation 2.

$$\delta(\mathrm{DSE}) = \delta t\,(-\mathbf{v} \cdot \nabla \mathrm{DSE} + \epsilon) \tag{2}$$

Taking the mass-weighted ~~daily integrals are given in equations 3 and ??, with equation 6a providing the~~ integral of equation 2, and performing Reynolds decomposition on both sides,

$$\delta\mathcal{S}_{Tot} = \delta t\,\Big(\oint_D (-\mathbf{v} \cdot \nabla \mathrm{DSE} + \epsilon)\,dM\Big)/M \tag{3a}$$

$$\delta\bar{\mathcal{S}}_{Tot} + \delta\mathcal{S}'_{Tot} = \delta t\,\Big(\oint_D -(v'\frac{\partial\overline{\mathrm{DSE}}}{\partial y} + v'\frac{\partial \mathrm{DSE}'}{\partial y} + \bar{v}\frac{\partial\overline{\mathrm{DSE}}}{\partial y} + \bar{v}\frac{\partial \mathrm{DSE}'}{\partial y} + \ldots)\,dM\Big)/M \tag{3b}$$

Using shorthand notation for ~~advective convergence terms .~~ the convergence terms on the RHS,

$$\delta\bar{\mathcal{S}}_{Tot} + \delta\mathcal{S}'_{Tot} = v'\bar{\mathcal{S}}_y + v'\mathcal{S}'_y + \bar{v}\bar{\mathcal{S}}_y + \bar{v}\mathcal{S}'_y + \ldots + \varepsilon \tag{4}$$

Where the order of terms on the RHS is preserved between 3b and 4. Ignoring the small daily changes in $\bar{\mathcal{S}}_{Tot}$, the mass weighted daily climatology of lower tropospheric DSE, we arrive at the relationship,

$$\delta\mathcal{S}'_{Tot} = \delta\mathcal{S} + \varepsilon \tag{5}$$

The terms and operations used in equations 1-5 are defined as follows:

$$\delta\mathcal{S} = v'\bar{\mathcal{S}}_y + v'\mathcal{S}'_y + \bar{v}\bar{\mathcal{S}}_y + \bar{v}\mathcal{S}'_y + \ldots \tag{6a}$$

$$\mathbf{v} = (u, v, w) \tag{6b}$$

$$\oint_D = \int_{Z=Z_{900\,hPa}}^{Z=Z_{600\,hPa}} \int_{\lambda=68^o}^{\lambda=78^o} \int_{\phi=25^o}^{\phi=31^o} \tag{6c}$$

$$dM = \rho * (R_e \cos\phi\, d\lambda) * (R_e\, d\phi) * (dZ) \tag{6d}$$

$$M = \oint_D dM \tag{6e}$$

$$\delta\mathcal{S}'_{Tot} = (\oint_D \delta DSE'\, dM)/M \tag{6f}$$

$$\bar{\mathcal{S}}_{Tot} = (\oint_D \delta\overline{DSE}\, dM)/M \tag{6g}$$

where δt = 86400, $\phi =$ Degrees of Latitude, $\lambda =$ Degrees of Longitude, $Z =$ Geopotential, $R_e =$ Radius of the Earth. The integrals above are computed by using discrete spatial steps of length 0.25^o in the latitudinal and longitudinal directions. We note that the correlation between $\delta\mathcal{S}_{Tot}$ and $\delta\mathcal{S}$ was ~~maximized~~ highest when we used 2-day ~~(day0 and lag1)~~ average values of ~~advection~~$\delta\mathcal{S}$ given by $\left[(\delta\mathcal{S}_{t_i} + \delta\mathcal{S}_{t_{i-1}})/2\right]_{i=1}^{N}$, where N is the total number of days in the analysis. Consequently, all ~~values~~ variables related to advection ($\delta\mathcal{S}$ ~~,~~ and its Reynolds components $v'\bar{\mathcal{S}}_y$, $v'\mathcal{S}'_y$, ~~etc.~~ ...) used throughout this study, and the mass-weighted values of velocity and $\mathcal{S}'$ in the ~~advective~~ advection regimes section (**??**5.1), represent ~~such~~ 2-day averages.

$$\frac{\partial(\mathrm{DSE})}{\partial t} = -\nabla\cdot(\mathbf{v}\,\mathrm{DSE}) + \epsilon$$

$$\delta(\mathrm{DSE}) = \delta t\,(-\mathbf{v}\cdot\nabla\mathrm{DSE} + \epsilon)$$

~~Taking the mass-weighted integral of equation 1, and using Reynolds decomposition to expand both sides,~~

$$\delta\mathcal{S}_{Tot} = \delta t\,(\oint_D (-\mathbf{v}\cdot\nabla\mathrm{DSE} + \epsilon)\, dM)\,/M$$

$$\delta\mathcal{S}_{Tot} = \delta\mathcal{S} + \varepsilon$$

where,

$$\delta\mathcal{S}_{Tot} = \delta\mathcal{S}'_{Tot} + \delta\bar{\mathcal{S}}_{Tot}$$

$$\delta\mathcal{S}'_{Tot} = (\oint_D \delta DSE'\, dM)/M$$

$$\bar{\mathcal{S}}_{Tot} = (\oint_D \delta\overline{DSE}\, dM)/M$$

$$\delta\mathcal{S} = \delta t\,(\oint_D -(v'\frac{\partial\overline{\mathrm{DSE}}}{\partial y} + \overline{v}\frac{\partial\mathrm{DSE}'}{\partial y} + v'\frac{\partial\mathrm{DSE}'}{\partial y} + ...)\, dM)\,/M$$

$$= v'\bar{\mathcal{S}}_y + \bar{v}\mathcal{S}'_y + v'\mathcal{S}'_y + \ldots$$

such that,

$$\mathbf{v} = (u, v, w)$$

$$\oint_D = \int_{Z=Z_{900\,hPa}}^{Z=Z_{600\,hPa}} \int_{\lambda=68^o}^{\lambda=78^o} \int_{\phi=25^o}^{\phi=31^o}$$

$$dM = \rho * (R_e \cos\phi\, d\lambda) * (R_e\, d\phi) * (dZ)$$

$$M = \oint_D dM$$

where δt = 86400, ϕ = Degrees of Latitude, λ = Degrees of Longitude, Z = Geopotential, R_e = Radius of the Earth.

We use the decision tree classification model, which has been extensively applied in various domains for many years (**??**). The model output is visualized as an inverted tree with a root node (e.g. Node 0 in Fig. **??**a), split nodes or decision nodes (e.g. Node 1, Node 4 in Fig. **??**a), and the "terminating" leaf nodes (e.g. Nodes 2, 3, 5, and 6 in Fig. **??**a). The model

We use the decision tree classification model, which has been extensively applied in climate science (**?????**). The model recursively partitions the input feature space, generating disjoint branches each corresponding to a set of decision rules that group the samples belonging to the same class together. For example, in Fig. **??**a, the branch connecting Node 0, Node 1 and Node 2 identifies the Negative class with a $\sim$ 0.75 probability. A decision node includes an independent variable and its corresponding

value for splitting the data on. This splitting value is determined by the model to maximize the reduction in information entropy (equation 7) (?). Each branch consists of internal nodes representing the decision variables or predictors and their optimal split values, and

The performance of a decision tree model is usually evaluated using the mean F1-score (harmonic mean of precision and recall for a given class), which is the arithmetic mean of the F1-scores computed for all classes (?). averaged over all classes of the dependent (clas (?). The model output is visualized as an inverted tree. A detailed description of the decision tree classification model is given in ?. We analyze the explainability of the identified mechanisms across the distribution of daily advection using linear regression and residual d

The entropy associated with a categorical variable Y with m classes is given by:

The entropy associated with a categorical variable Y with m classes is given by

$$\mathrm{H}(Y) = -\sum_{i=1}^{m} p_i \log_2(p_i) \,|\, \mathrm{P}(Y = i) = p_i \tag{7}$$

We identify the primary contributions to advection using the decision tree algorithm to model tercile based equiprobable classes of δS, namely Negative, Neutral, and Positive. We tune the model parameters to achieve a performance score (F1-score) of at least 0.6 using a tree depth of 2 (excluding the root node). The predictors are standardized with respect to mean and standard deviation for each season (March and Apr), for facilitating interpretation and readability of the decision tree model. We analyze the effectiveness of the identified mechanisms We first use the decision tree model to identify the conditions primarily associated w We then analyze the explainability provided by the identified conditions across the distribution of daily advection $\delta\mathcal{S}$ using linear regression and a diagnostic, residual analysis. Using Finally, we use the decision tree model , we identify secondary contributions differentiating the 1st and 10th for identifying conditions additionally associated with the extreme deciles of $\delta\mathcal{S}$, using a tree depth of 3 for allowing nonlinear interactions.

All calculations were . All calculations are performed using the Python packages xarray (?), numpy (?), pandas (?), statsmodels (?), scikit-learn (?), and plotting was done with the help of Python packages matplotlib (?) and cartopy (?).

3 Lower Tropospheric Advection tropospheric advection links Circulations circulations with Changes daily changes in T2mT_{2m} anomaly

Positive T2m anomaly T_{2m} anomaly (T'_{2m}) is associated with anomalies of fluxes of both both increased fluxes of shortwave radiation and sensible heat into the ground, indicating decreased sensible heat fluxes from the ground (Supplementary Fig. S5over the region of interest (Supplementary Fig. S1). In contrast to this, during "mega heatwaves" studied by (??) , (??), were associated with increased sensible heat fluxes from dessicated surfaces have been observed which are known to contribute to feedback cycles sustaining significant surface warming over multiple days. On the other hand, we find

that ~~changes in T2m anomaly~~ daily changes in T'_{2m} (which we denote by $\delta T'_{2m}$) are linked to ~~the changes in $\mathcal{S}_{Tot}$~~$\delta\mathcal{S}_{Tot}$, which in turn are highly correlated with ~~the 2-day averaged advective contributions to the total daily change (Fig~~$\delta\mathcal{S}$ (Figs. 1a, 1b). We find that $\delta T'_{2m}$ is also strongly linked to $\delta\mathcal{S}$, especially at large magnitudes (Fig. 1c). This suggests that the surface fluxes are likely a slower timescale forcing, and that daily variability is primarily governed by the atmospheric flow. Further, we find that the lower tropospheric daily temperature changes are strongly explained by daily T_{2m} changes (Supplementary Fig. S

While $\delta\mathcal{S}_{Tot}$ is unbiasedly approximated by $\delta\mathcal{S}$ in the body of its distribution, parameterized energy fluxes become important for extreme values of $\delta\mathcal{S}$, and cause the systematic overprediction of $\delta\mathcal{S}_{Tot}$ (Fig. 1b). ~~Further, we~~ We note the barotropic nature of ~~the wind fields and $\mathcal{S}$ (Supplementary Fig. S6), which are driven by upper tropospheric forcing (Supplementary Fig. S7). Motivated by the association between T2m anomaly changes and advection driven changes in the lower troposphere, and the influence of atmospheric circulation, we identify $\delta\mathcal{S}$ as our quantity of interest.~~ $\mathcal{S}'$ and the eddy wind fields, that strengthen progressively wi

Motivated by the association between daily changes in near-surface temperature anomaly and advection driven daily changes in lower tro

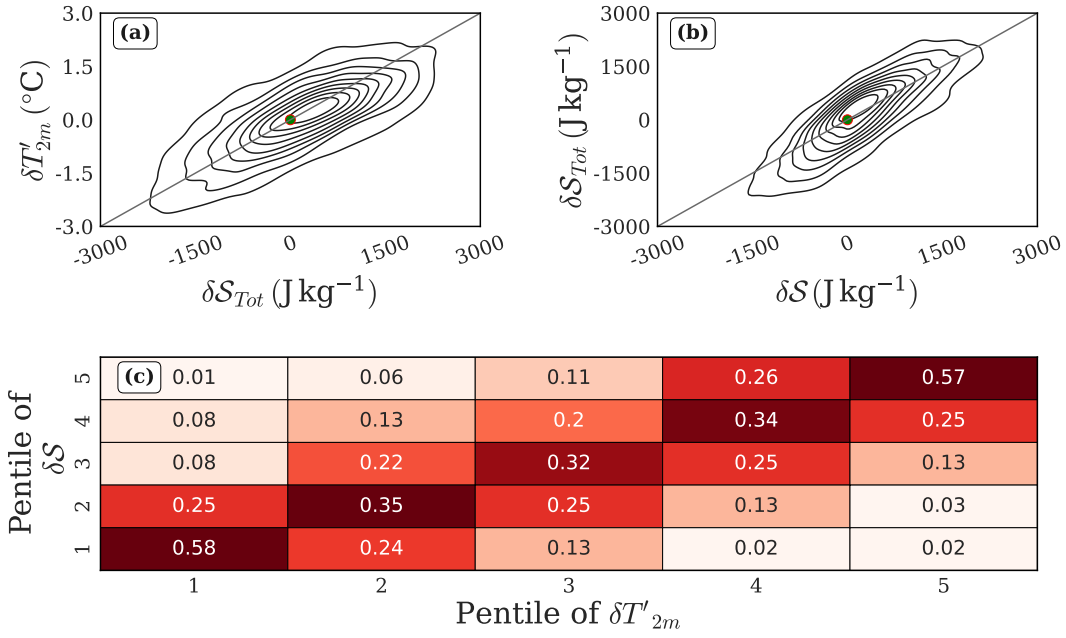

Figure 1. ~~For March and April combined, a)~~ This figure shows the ~~relationship~~ pairwise relationships between daily ~~change~~ changes in ~~T2m~~ T_{2m} anomaly ($\delta T'_{2m}$)~~and~~, daily ~~change~~ changes in ~~total~~ 600-900 hPa mass-weighted ~~DSE~~ DSE ($\delta\mathcal{S}_{Tot}$)~~b) the relationship between $\delta\mathcal{S}_{Tot}$~~, and ~~the advective contributions to $\delta\mathcal{S}_{Tot}$~~ daily 600-900 hPa mass-weighted advective convergence of DSE ($\delta\mathcal{S}$) during the combined periods of March and April. (a) The joint probability density plot between $\delta T'_{2m}$ and $\delta\mathcal{S}_{Tot}$. (b) The joint probability density plot between $\delta\mathcal{S}_{Tot}$ and $\delta\mathcal{S}$. For both these plots, the de

4 Primary ~~Advective Contributions~~ advective contributions to the $\delta\mathcal{S}$ ~~Budget~~budget

Our strategy to unravel the relationships governing the variability in ~~$\delta\mathcal{S}$~~ advection of lower tropospheric DSE is to first find the underlying mechanisms that govern its sign, and subsequently study ~~drivers of~~ the drivers of its magnitude. A description of the ~~magnitude~~ magnitudes of the Reynolds components of $\delta\mathcal{S}$ is provided in the Supplementary Fig. ~~S1 and its supporting text. Further, the~~ S6. The magnitudes of these components are further justified by analyzing the ~~the magnitude of anomalies and climatologies of each the~~ magnitudes of the climatology and anomaly of $\mathbf{v}$ and $\nabla\mathcal{S}$ ~~component~~ in the zonal, meridional, and vertical directions (Supplementary Table S1 and ~~Figs. S2, S3).~~ Supplementary Figs. S7). A zonally oriented climatology of $\mathcal{S}_{Tot}$ and horizo

This section focuses on the first part of this strategy. $\delta\mathcal{S}$ closely resembles $\mathbb{N}(0,\sigma)$ in its body, with deviation in the tails (Supplementary Fig. ~~S8~~S9). Due to the observed symmetry in the body of $\delta\mathcal{S}$, it is possible to model the separation of its signs by using tercile-based classification of $\delta\mathcal{S}$ with the classes labelled as Negative, Neutral, and Positive, as the dependent variable in a decision tree classification model~~(e.g., using the 33$^{\text{rd}}$ and 66$^{\text{th}}$~~. The 33$^{\text{rd}}$ and 66$^{\text{th}}$ percentile values demarcated in Supplementary Figs. ~~S1 a, b)~~S6a, S6b. Even though this partitioning seems coarse-grained, we will subsequently show that most large $\delta\mathcal{S}$ events are contained in the Positive and Negative classes defined here.

~~Multivariate models of depth 2 for each of March and April retrieve~~ A multivariate decision tree model for tercile-based classes of daily advection retrieves combinations of two of the largest magnitude Reynolds components, namely $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$ (~~Figs. ??a~~and ~~??a~~Fig. 2a). The ~~models are~~ model is able to identify at least one ~~set of decision rules for nodes associated with~~ decision rule for each of the three classes. ~~Some decision rules identified as~~ The decision rules represented by the branches of the tree are described ~~in Table ??~~by the conditional probability statements presented in Table 1. The Positive and Negative classes are ~~better separated, or~~ identified with a larger probability of occurrence ~~,~~ than the Neutral class~~,~~. The decision rules associated with the Positive and Negative classes are ~~associated with~~ based on the in-phase combination of $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$~~. The~~ , while the Neutral class is associated with out-of-phase combinations of ~~these components~~$\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$.

The model generated decision rules given in Table ~~??~~ 1 are based on a selection of the variables and split points that maximize entropy reduction (section 2.2). However, the underlying correlation structure of the dataset is not accounted for by the model, specifically, multicollinearity. For example, consider a variable that independently explains a large fraction of the variance in the dependent variable, but a much smaller fraction when included with one or more of the remaining model predictors. The model algorithm would not identify this variable as important, and its confounded influence would remain unidentified. We find that $v'\bar{\mathcal{S}}_y$ is the confounder in our case, due to a strong inverse association of $v'\bar{\mathcal{S}}_y$ with both other quasilinear components, $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$ (~~Figs. ??b, ??b~~Fig. 2b). $v'\bar{\mathcal{S}}_y$ is important to decipher the interaction pathways involving $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$. Moreover, its magnitude is also comparable to these variables (Supplementary Fig. ~~S1 b~~S6c), making its quantitative effect significant. Thus, $v'\bar{\mathcal{S}}_y$ is additionally identified as a key component in the analysis of $\delta\mathcal{S}$. It will be shown later (Fig. 4~~c~~a) that the sum of all ~~the~~ three quasilinear terms is a strong primary indicator of advection.

The inverse relationship between $v'\bar{\mathcal{S}}_y$ and $w'\bar{\mathcal{S}}_z$ is traced back to an in-phase relationship between mass-weighted integrals of v' and w' over the volume of the box. Due to the vertical coherence of ~~both the~~ the meridional and vertical velocity anomaly

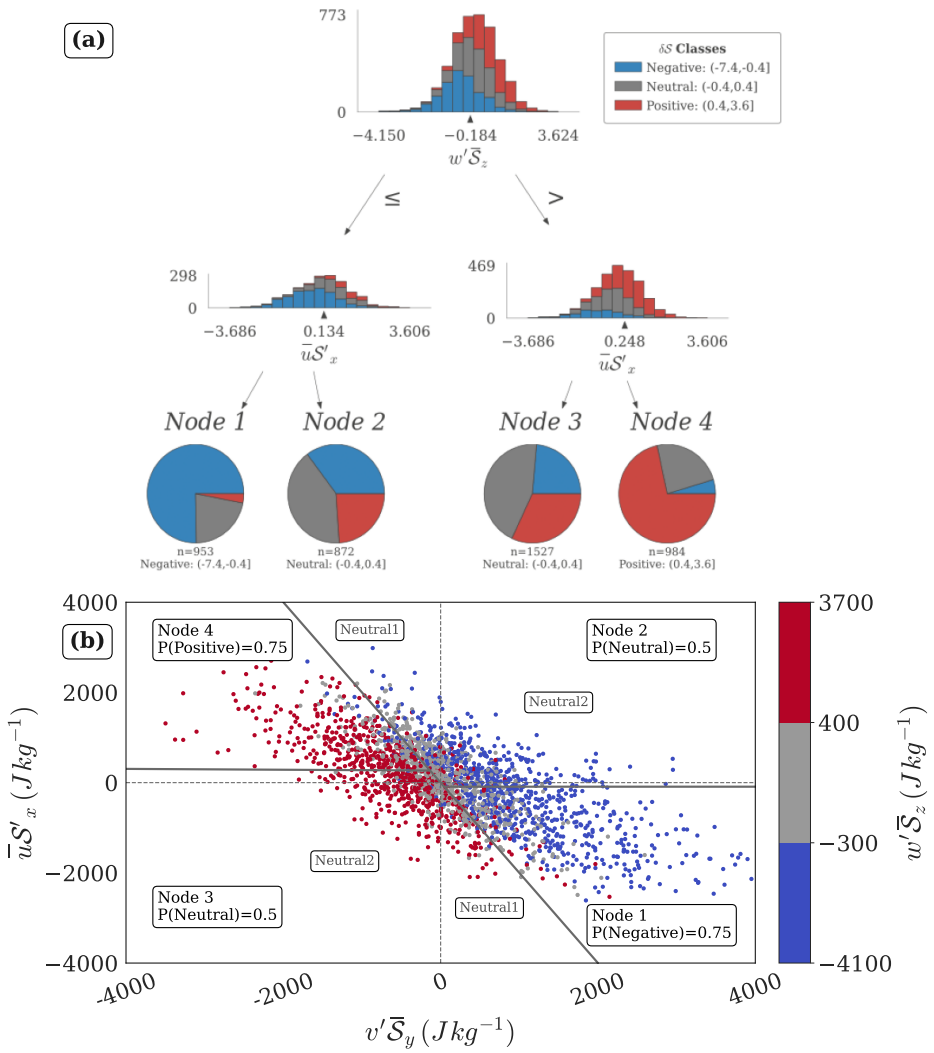

Figure 2. ~~a)~~ ~~Representations~~ This figure examines the relationships governing the differences between terciles of daily advection for the ~~standardized~~ combined period of March and April. (a) Standardized decision tree model for tercile based classes of $\delta\mathcal{S}$, ~~for March~~across March-April. The root node and the internal (decision) nodes are represented by stacked bar charts ~~for the relative frequency~~ of the ~~classes modeled by~~ class frequencies, and the ~~algorithm. The~~ distribution of the classes in the leaf nodes is represented by the pie charts. ~~We use standardized variables with values given by z-scores.~~ The ~~splitting~~ variables are standardized, and the split thresholds at each node ~~are expressed as~~ reflect the number of standard deviations from the mean~~of the splitting variable chosen by the algorithm~~. The ~~apriori probability of each class (Node 0) is 0.33. The performance is given by~~ model achieves an F1-score of ~ 0.6~~. The Positive and Negative classes were better identified than the Neutral class~~, with ~~mean F1-scores of approximately 0.65 and 0.55~~early-stop criteria as follows: entropy decrease ≥ 0.03 at each level, ~~respectively~~$\geq 40\%$ sample size for level splitting, $\geq 20\%$ sample size for leaf nodes. ~~b)~~ ~~Scatterplots show~~ (b) Scatterplot showing the ~~untreated~~ relationship between the unstandardized values of $v'\bar{S}_y$ and $\bar{u}S'_x$, colored by terciles of $w'\bar{S}_z$ ~~for March~~across March-April. The solid grey lines ~~demarcate~~ approximate decision boundaries~~delineating the decision spaces of different nodes. Each space is annotated~~, with each region labeled by the corresponding Node # ~~,~~ and ~~the probability associated with its~~ dominant class probability. The ~~pairwise~~ labels "Neutral1" and "Neutral2" correspond to the two kinds of deviations from the primary relation described in section 4.1. Pairwise correlations~~between these variables are given by~~: ~~$r(v'\bar{S}_y, w'\bar{S}_z) = -0.79$~~$r(v'\bar{S}_y, w'\bar{S}_z) = -0.75$, $r(v'\bar{S}_y, \bar{u}S'_x) = -0.75$, ~~$r(\bar{u}S'_x, w'\bar{S}_z) = 0.35$~~$r(\bar{u}S'_x, w'\bar{S}_z) = 0.25$

~~Leaf Node Label~~ **Advection Class (k)**	~~Advection Class (K)~~ **Node #**	~~Conditions (C)~~ **Conditions (C)**	~~P(Y~~
~~Node 2~~ Negative	1	$\bar{u}\mathcal{S}'_x \leq 0, w'\bar{\mathcal{S}}_z \leq 0$	
~~Node 3, Node 5~~ Neutral	~~$(\bar{u}\mathcal{S}'_x \leq 0, w'\bar{\mathcal{S}}_z > 0) \mid (\bar{u}\mathcal{S}'_x > 0, w'\bar{\mathcal{S}}_z \leq 0)$~~ 2, 3	$(\bar{u}\mathcal{S}'_x \leq 0, w'\bar{\mathcal{S}}_z > 0) \mid (\bar{u}\mathcal{S}'_x > 0, w'\bar{\mathcal{S}}_z \leq 0)$	
~~Node 6~~ Positive	4	$\bar{u}\mathcal{S}'_x > 0, w'\bar{\mathcal{S}}_z > 0$	

Table 1. ~~Description of~~ This table describes the decision tree ~~models~~ model presented in ~~Figs~~Fig. ~~??~~2a, ~~??a~~. The decision rules and associated ~~probabilities for a given~~ class probabilities are approximately the same for both months; therefore, ~~so~~ we present the combined results. The ~~mean for both~~ means of the involved variables, and ~~the~~ their scalar multipliers of σ, are both ~ 0~~; hence~~, so expressions such as 0.1σ ~~is~~ are approximated as 0 for ease of readability. For example, $\bar{u}\mathcal{S}'_x \leq -0.1\sigma(\bar{u}\mathcal{S}'_x)$ is simplified ~~as~~ to $\bar{u}\mathcal{S}'_x \leq 0$. Table headers are described as follows. ~~**Leaf Node Label**~~**Advection Class (K)**: ~~Identifies a specific leaf node in~~ The tercile based classes of $\delta\mathcal{S}$ modeled by the decision tree classification model. ~~**Advection Class (K)**~~**Node #**: ~~Specifies the~~ The leaf node numbers identifying each advection class~~(Negative, Neutral, Positive) associated with the node~~. **Conditions (C)**: ~~Lists the~~ The set of decision rules that ~~define~~ characterize the ~~branch terminating in~~ branches associated with the ~~given~~ corresponding leaf node(s) identifying a given advection class. $\mathbf{P(Y = K \mid C)}$: ~~Represents the~~ The probability of occurrence of the identified advection class~~conditional to~~, given the corresponding decision rules.

fields (Supplementary Fig. ~~S6) and spatial detivatives~~ S4) and corresponding spatial derivatives of $\bar{\mathcal{S}}$ (Supplementary Fig. ~~S4~~S10), the quasilinear components behave like linear transformations of the mass-weighted quantities v', w', which ~~were~~ are related through the ~~quasi-geostrophic~~ quasigeostrophic omega equation. ~~We also found that the vertical velocity was almost entirely explained by the quasi-geostrophic omega equation (Supplementary Fig. S7).~~

At the same time, $w'\bar{\mathcal{S}}_z$ and ~~$\bar{u}\mathcal{S}'_x$~~ $\bar{u}\mathcal{S}'_x$ are not strongly correlated, suggesting the role of other mechanisms, either external to the system or related to different latent relationships with respect to $v'\bar{\mathcal{S}}_y$, governing at least one of the two variables. For ~~e.g.~~example, mass-weighted $\mathcal{S}'_x$, which is scaled by a vertically coherent $\bar{u}$, is likely a function of the location and sign of one or multiple eddies interacting with the region or the result of advection by flows not associated with coherent structures.

4.1 Phenomenology

We plot the decision spaces generated by the tree model on the scatterplot involving all ~~the~~ three quasilinear variables identified above (~~Figs. ??b, ??b~~Fig. 2b). However, it is unclear what physical conditions they represent. Thus, we analyze individual instances ~~,~~and obtain a phenomenology of the large scale spatial patterns of winds and ~~$\mathcal{S}'$~~ $\mathcal{S}'_{Tot}$ corresponding to different regions on the scatterplot. Both periods (March and April) showed similar results in the following sections; therefore, we focus on ~~preenting~~ presenting an in-depth analysis of April for the rest of this paper with the corresponding figures for March presented in the Appen

The decision spaces associated with the Negative and Positive classes have been annotated as Node 1 and Node 4, respectively, along w

The large values of $|v'\bar{\mathcal{S}}_y|$ ($> \sigma$) are associated with a highly predictable sign and magnitude of the other quasilinear quantities, $\bar{u}\mathcal{S}'_x$ and $w'\bar{\mathcal{S}}_z$. Smaller values ($|v'\bar{\mathcal{S}}_y| < \sigma$) were associated with variability in the signs of both other quantities. The extreme deciles (1 and 10) of advection driven daily changes in $\mathcal{S}$ almost exclusively occupied the decision spaces associated with the

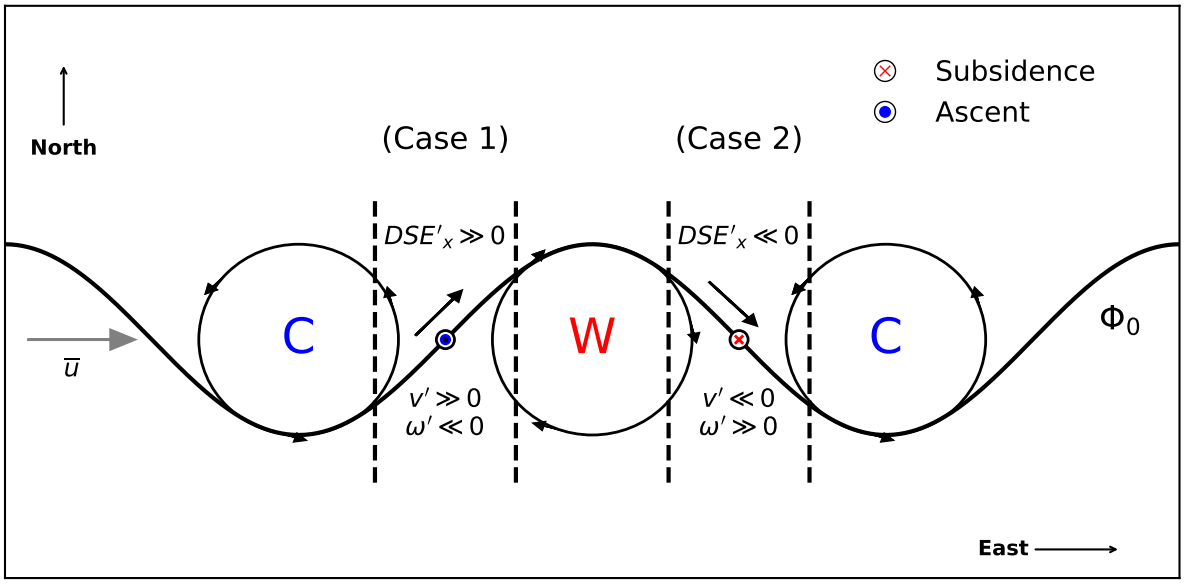

Figure 3. ~~a)~~ ~~Representations of the standardized decision tree model for tercile based classes of $\delta\mathcal{S}$, for April. Model description as in Fig. ??. The performance is given by F1-score of ~ 0.6. The Positive and Negative classes are better identified than the Neutral class, with mean F1-scores of approximately 0.65 and 0.55, respectively.~~ ~~**b)**~~ ~~Scatterplots show the untreated values of $v'\bar{\mathcal{S}}_y$ and $\bar{u}\mathcal{S}'_x$, colored by tercils of $w'\bar{\mathcal{S}}_z$ for April. The solid grey lines demarcate approximate decision boundaries delineating the decision spaces of different nodes. Each space is annotated with the Node #, and the probability associated with its dominant class. The pairwise correlations between these variables are given by: $r(v'\bar{\mathcal{S}}_y, w'\bar{\mathcal{S}}_z) = -0.7$, $r(v'\bar{\mathcal{S}}_y, \bar{u}\mathcal{S}'_x) = -0.7$, $r(\bar{u}\mathcal{S}'_x, w'\bar{\mathcal{S}}_z) = 0.1$.~~ This schematic clarifies the relationship between the large scale quantities v', w', and $\mathcal{S}'_x$. Consider a climatological zonal flow ($\bar{u}$) and a geopotential cor

~~Positive and Negative Nodes 2 and 6 (Figs. ??b, ??b~~Negative and Positive Nodes 1 and 4 (Fig. 2b; Supplementary Fig. ~~S9~~S11),
235 and represented various configurations of anticyclones and cyclones located in proximity to the region of interest.

The decision spaces associated with the Neutral class have been annotated as Node ~~3 and Node 5~~ 2 and Node 3 along with probabilities of the Neutral class in ~~Figs. ??b, ??b~~Fig. 2b. These decision spaces only weakly prefer the Neutral class with lesser but comparable probabilities of occurrence of the other classes. We find that these are associated with two kinds of deviations from the expected relationships established above. ~~First,~~ The first kind is where the meridional and vertical quasilinear
240 components are in phase. We label the regions on the scatterplots representing this kind of deviation as "Neutral1". These points are observed to lie between the lines $\bar{u}\mathcal{S}'_x = -2 * v'\bar{\mathcal{S}}_y$ and $v'\bar{\mathcal{S}}_y = 0$, where $\bar{u}\mathcal{S}'_x * v'\bar{\mathcal{S}}_y < 0$. We find that they represent conditions where mass-weighted integrals of v' and w' acted out of phase, i.e., ascending northerly winds ~~or~~ and subsiding southerly winds~~instead of the other way around~~, as expected by the ~~quasi-geostrophic~~ quasigeostrophic relationship. These conditions are characterized by small magnitude ($< \sigma$) of $v'\bar{\mathcal{S}}$, and the horizontal dominance of $\bar{u}\mathcal{S}'_x$. ~~In spite of the small~~
245 Despite the weak meridional winds, instances of large vertical winds are observed in association with mainly the zonal arm of a cyclone or anticyclone located directly to the north or south of the region, when both the meridional arms of ~~a curved structure~~

an eddy overlap with the region~~to varying extents albeit of weak strength. Here, v' is a result of the area averaging, and w', of the quasi-geostrophic omega equation.~~ . Here, v' is rendered small due to averaging over opposite phases, while large w' persists due to homog
$\mathcal{S}'_x$ is purely a function of the differently signed eddies interacting with different zonal halves of the region, and is not sensitive to the v', w' relationship. In cases without eddies, w' is small as v', and only the zonal mean flow advects a large $\mathcal{S}'$ into the region, under a $\mathcal{S}'_x$ governed by no particular structures. Typical configurations associated with the "Neutral1" regions are presented in Supplement

The second kind of deviation is represented by points in the first and third quadrants, where the meridional and zonal quasilinear advection terms are ~~observed to be~~ in phase. We label these regions as "Neutral2" in ~~both the scatterplots (Figs. ??b, ??b).~~ the scatterplot Fig. 2b. The inverse phase relationship between the meridional and vertical quasilinear terms is preserved even ~~under small magnitude, and~~ at small magnitudes. It represents coherent configurations given by zonally oriented arms of southern/ northern large scale eddies and ridge-like conditions over the region. $\bar{u}\mathcal{S}'_x$ is reversed in sign either because of intensification of $\mathcal{S}'$ along the southern edge of High Mountain Asia, or due to weak larger scale structures of $\mathcal{S}'_x$ not always directly linked with circulation structures (the quantities are 600-900 hPa averaged, and generally, a given location could be influenced by diabatic forcing in the lower vertical levels, or be affected by other proximate eddies). Typical configurations associated with the "Neutral2" regions are presented in Supplementary Figs. S12c, S12d.

4.2 ~~Efficacy~~Explainability and residual bias

We ~~evaluate the explainability and efficacy~~ assess the performance of the quasilinear advective contributions by regressing the ~~advective contributions to daily change in $\mathcal{S}$~~ daily advection of dry static energy into the lower troposphere, $\delta\mathcal{S}$, against the sum of ~~the quasilinear terms, denoted by $\delta\mathcal{S}_{QL}$. We assess explainability using the~~ its quasilinear components, $\delta\mathcal{S}_{QL} = v'\bar{\mathcal{S}}_y + w'\bar{\mathcal{S}}_z + \bar{u}\mathcal{S}'_x$.

Using the coefficient of determination (R^2~~of~~), we find that that $\delta\mathcal{S}_{QL}$ explains $\sim 65\%$ of the variability in $\delta\mathcal{S}$. To evaluate its performance across the ~~model, and evaluate efficacy by examining the~~ distribution of daily advection, we examine residuals stratified by deciles of $\delta\mathcal{S}$. In the central deciles, the residuals ~~by decile of the dependent variable. It is seen that $\delta\mathcal{S}_{QL}$ explains $\sim 65\%$ of the variability present in the dependent variable. The residuals~~ (observation - prediction) are more or less symmetrically distributed about zero ~~in the body of the distribution~~ (Fig. 4a), ~~and show an increasingly directional bias as we approach the extremes~~(indicating an unbiased model fit. Towards the extremes, however, the residuals exhibit a systematic directional bias (as implied by the increasing magnitude of the mean residuals towards extreme deciles in Fig. 4a)~~. That is, while $\delta\mathcal{S}_{QL}$ predicts the sign of $\delta\mathcal{S}$ correctly, and estimates the magnitude accurately in the central half of its distribution, it systematically underpredicts the magnitude with increase in distance from the body of the distribution of $\delta\mathcal{S}$~~, indicating the insufficiency of driving mechanisms in the extremes of the model.

~~This bias is corrected only by the inclusion of the~~ The bias in the primary contributions model is corrected by including the sum of nonlinear variables ~~($\delta\mathcal{S}_{NL} = v'\mathcal{S}'_y + w'\mathcal{S}'_z + u'\mathcal{S}'_x$). The addition of $\delta\mathcal{S}_{NL}$), which corrects the skewness~~ eliminates the non-zero mean of residuals in the tails~~and explains a smaller fraction of the unexplained variance uniformly throughout~~

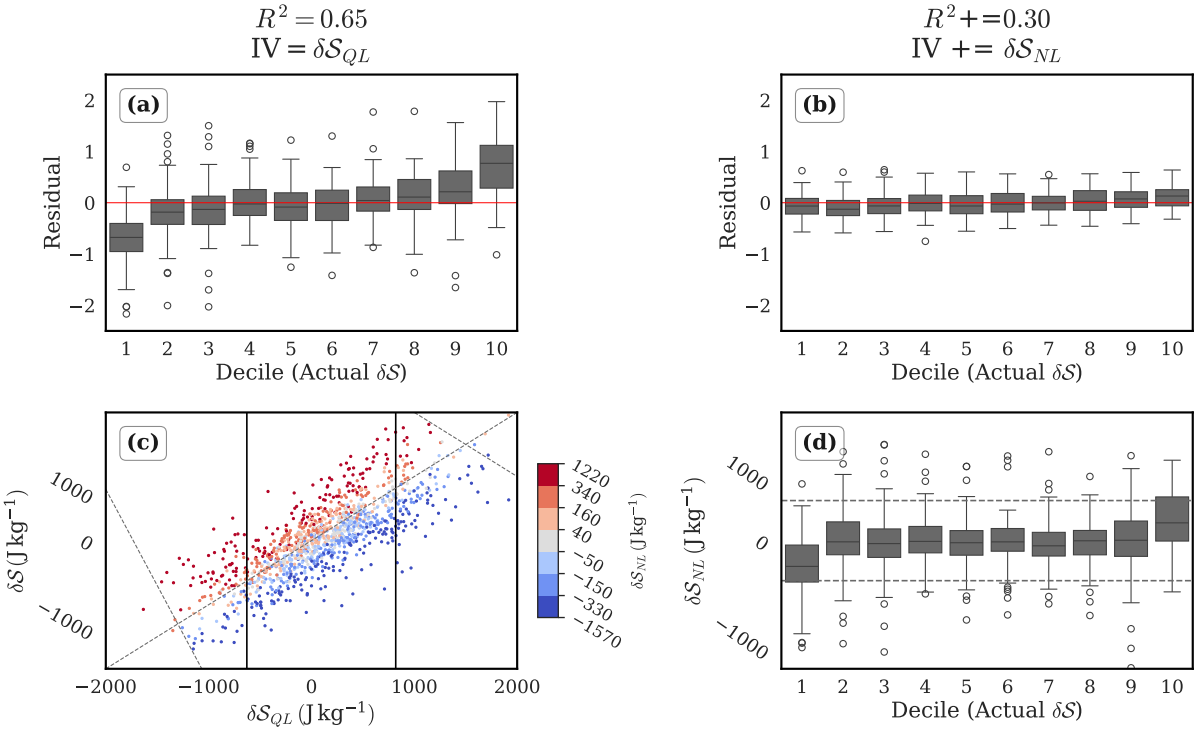

Figure 4. ~~For~~ This figure presents two regression models for $\delta\mathcal{S}$, and complementary plots based on unstandardized data to aid interpretability, for the pe
(c) Scatterplot between the ~~period of April, performing linear regression with all~~ unstandardized variables~~standardized: a)~~ , $\delta\mathcal{S}$ ~~against~~ and $\delta\mathcal{S}_{QL}$ ~~as the predictor~~, ~~we plot the boxplot of residuals~~ colored by ~~groups~~ quantile based ~~on deciles~~ categories of ~~$\delta\mathcal{S}$; b) $\delta\mathcal{S}$ against $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ as~~ . The two negatively sloped dotted lines connect the ~~predictors, we plot the boxplot of residuals by groups based on deciles of $\delta\mathcal{S}$; c) the scatterplot of model predictions against actual~~ peak values of $\delta\mathcal{S}$ ~~for the model~~ and $\delta\mathcal{S}_{QL}$ in ~~a; d)~~ the ~~scatterplot of model predictions against actual values of $\delta\mathcal{S}$ for the model in b; e) Scatterplot between the untreated variables, $\delta\mathcal{S}$~~ right and ~~$\delta\mathcal{S}_{QL}$~~ left tails, ~~colored by~~ highlighting the asymmetric saturation effect of $\delta\mathcal{S}_{NL}$~~; f.~~ (d) Boxplot of ~~untreated~~ unstandardized values ~~:~~ of $\delta\mathcal{S}_{NL}$ grouped by ~~groups based on~~ deciles of $\delta\mathcal{S}$. ~~(OLS regression output summaries uploaded as PDF files in Supplementary Information.)~~

the body , and reduces the variance of residuals in all the deciles of $\delta\mathcal{S}$, with the largest reduction in the tails (Fig. 4b). The combination $\delta\mathcal{S}_{QL}$ and importance of $\delta\mathcal{S}_{NL}$ explain $\sim 95\%$ of the variability in daily advection of $\mathcal{S}$. in the tails of $\delta\mathcal{S}$ is verified by the nonzero mean of $\delta\mathcal{S}_{NL}$ in the extreme deciles of $\delta\mathcal{S}_{QL}$ (Fig. 4d). We infer that the quasilinear contributions are active throughout the distribution of $\delta\mathcal{S}$, and the nonlinear contributions are mainly active in primarily modulate the extremes. The other quasilinear and constant terms do not add significant explanatory power to the model due to the small magnitude of $\bar{v}$, $\bar{w}$, and $\bar{\mathcal{S}}_x$ (Supplementary Table S1).

Together, the quasilinear and nonlinear contributions account for $\sim 95\%$ of the variability in $\delta\mathcal{S}$. The rest of the component terms on the

We shall refer to the quasilinear contributions as the primary contribution, and the nonlinear contributions as the secondary contribution , secondary contribution for the remainder of the paper.

Figure 4e c shows that the magnitude of $\delta\mathcal{S}$ peaks earlier than $\delta\mathcal{S}_{QL}$ in both tails, suggesting that the main function of the nonlinear terms in the tails is to act against the quasilinear contribution, effectively saturating the growth of $\delta\mathcal{S}$. While this in itself may not seem surprising and has been noted elsewhere (**?**), we also observe a larger magnitude of slope difference between the peak values of $\delta\mathcal{S}$ and $\delta\mathcal{S}_{QL}$ in the right tail which as compared to the left tail. This asymmetry suggests that the saturative effect of the nonlinear terms is stronger there.

5 Secondary Advective Contributions advective contributions to the $\delta\mathcal{S}$ Budgetbudget

The fact that the linear regression model based on $\delta\mathcal{S}_{QL}$ alone causes a skewed distribution of residuals in the extreme deciles of $\delta\mathcal{S}$, and that the residuals are not symmetric about the tails, suggests a crucial role for nonlinearities in determining the asymmetry of the distribution of $\delta\mathcal{S}$. Specifically, the distribution of residuals in the left extreme decile (decile 1) and the right extreme decile (decile 10) is not exactly equal and opposite, highlighting the different influences of $\delta\mathcal{S}_{NL}$ in these regions. While it has been previously suggested that the meridional nonlinear term alone can provide a satisfactory explanation for the observed asymmetry (skewness in their case) (**?**), it is unclear whether this is true for our case as well. Thus, the next natural question is to identify the nonlinear components driving the tails of $\delta\mathcal{S}$.

Like As in the previous section, we use the decision tree model to provide the leading explanatory variables for the extreme deciles 1 and 10 , of $\delta\mathcal{S}$. Alongside nonlinear terms, we incorporate "background conditions" the pre-existing anomaly defined as the previous day's $\mathcal{S}'_{Tot}$ anomaly(denoted by total lower tropospheric DSE anomaly, $\mathcal{S}'_{Tot,Lag1}$), as explanatory variables. We expect $\mathcal{S}'_{Tot,Lag1}$ to be relevant since a large positive $\delta\mathcal{S}$ occurring during a positive $\mathcal{S}'_{Tot,Lag1}$ may have different drivers than when it occurs during a negative $\mathcal{S}'_{Tot,Lag1}$ (and vice versa). We find , as an explanatory variable. This accounts for the possibility that the drivers of large positive $\delta\mathcal{S}$ may differ depending on whether advection acts upon a pre-existing p Our results show that not only are different distinct nonlinear components active in the left and right tails, but that these components also differ for different $\mathcal{S}'_{Tot,Lag1}$. but that the influence of these components also varies with the sign of $\mathcal{S}'_{Tot,Lag1}$.

The decision tree model identifies the combinations of variables separating the extreme deciles of $\delta\mathcal{S}$ (Fig. 5) and leads to the following picture:

– ~~For negative background conditions,~~ When the pre-existing anomaly is negative, ($\mathcal{S}'_{Tot,Lag1} < 0.51\sigma$; we round the level split thresh large values of advection ~~(~~ both positive and negative ~~)~~ are associated with the vertical term $w'\mathcal{S}'_z$~~(~~, as seen in the left-most branches ~~corresponding to $\mathcal{S}'_{Tot,Lag1} < 0.5\sigma$,~~ terminating in Node ~~2 and Node 3). Large positive~~ 1 and Node 2. Positive values of $w'\mathcal{S}'_z$ ~~(0.16σ,5σ)~~ $> 0.16\sigma$ (extending to 5.38σ) correlate strongly with the positive extremes of $\delta\mathcal{S}$ (Node ~~3~~2), indicating a strong role of $w'\mathcal{S}'_z$ in dissipating ~~the negative background~~ pre-existing negative anomalies. Conversely, small negative values of $w'\mathcal{S}'_z < 0.16\sigma$ (with a very thin tail extending ~~the range~~ to -8σ) are associated with ~~the Node 2~~Node 1, indicating the potential role of confounding nonlinear variables.

– ~~When $\mathcal{S}'_{Tot,Lag1} > 0$,~~ When the pre-existing anomaly is positive ($\mathcal{S}'_{Tot,Lag1} > 0.51\sigma$), large negative values of $\delta\mathcal{S}$ (Node ~~5 and Node 7~~3 and Node 4) correspond to strong dissipation of ~~positive background~~ a pre-existing positive anomalies. The model suggests two pathways, associated with *the horizontal terms* $v'\mathcal{S}'_y$ and $u'\mathcal{S}'_x$ respectively. ~~$v'\mathcal{S}'_y < -0.2\sigma$~~ $v'\mathcal{S}'_y < -0.24\sigma$ identifies the Negative extreme decile of $\delta\mathcal{S}$ with a 0.85 probability (Node ~~5~~3). When its values exceed ~~-0.2σ, $u'\mathcal{S}'_x$ is associated with~~ -0.24σ, but when $u'\mathcal{S}'_x > -0.93\sigma$, there is a a 0.75 probability of occurrence of the Negative extreme decile of $\delta\mathcal{S}$. ~~However, based on magnitude based arguments as the previous point, both the horizontal nonlinear terms are weak drivers.~~

– ~~When $\mathcal{S}'_{Tot,Lag1} > 0$, large~~ Large positive values of $\delta\mathcal{S}$ (Node ~~8~~5) correspond to strong amplification of ~~positive background anomaly.~~ a pre-existing positive anomaly, and are preferred when $v'\mathcal{S}'_y$ is approximately in its positive half, in combination with small magnitude values of either sign of $u'\mathcal{S}'_x$ (as seen from the histogram of δS at the node where $u'\mathcal{S}'_x$ is split into Nodes

~~In order to~~ To contextualize the model results with ~~the unfiltered, untreated~~ raw data, we inspect the nonlinear-quasilinear relationships in the context of $\mathcal{S}'_{Tot,Lag1}$ (~~Fig~~Figs. 5b, ~~c,~~ 5c, 5d). First, we see that the nonlinear terms achieve large magnitudes in the extremes of their quasilinear counterparts, corroborating observations from Fig. 4~~f~~d. Strong, nearly orthogonal ~~,~~ relationships emerge in the meridional and vertical directions. Recollecting the strong association between different quasilinear terms and $\delta\mathcal{S}_{QL}$ (Supplementary Fig. ~~S10~~S13), it is seen that ~~$w'\mathcal{S}'_z$~~ $w'\mathcal{S}'_z > 0$ (first quadrant in Fig. 5d) is strongly associated with dissipation of ~~negative background conditions~~ a pre-existing negative anomaly (since $w'\bar{\mathcal{S}}_z > 0$ corresponds to $\delta\mathcal{S} > 0$). $w'\mathcal{S}'_z < 0$ has fewer instances associated with $\delta\mathcal{S}_{QL} < 0$, where it causes amplification of ~~negative background conditions~~a pre-existing negative anomaly. On the other hand, $v'\mathcal{S}'_y$ does not observe as strong a relationship with $\delta\mathcal{S}_{QL}$ in spite of an apparently weak inverse relationship, when $\mathcal{S}'_{Tot,Lag1} < 0$. These observations justify the model's choice of $w'\mathcal{S}'_z$ in separating the extreme deciles of advection when $\mathcal{S}'_{Tot,Lag1} < 0$. When $\mathcal{S}'_{Tot,Lag1} > 0$, the horizontal nonlinear-quasilinear relationships seem more active than the vertical relationship, which is reflected in the model output. $v'\mathcal{S}'_y$ seems to weakly prefer being in phase with $\delta\mathcal{S}_{QL}$, although we see a large frequency of ~~instanes~~ instances with $v'\mathcal{S}'_y > 0$ corresponding to $\delta\mathcal{S}_{QL} < 0$ conditions. $u'\mathcal{S}'_x$ shows no clear preference except for a strongly negative skew for both ~~kinds of background conditions~~signs of pre-existing anomaly. These findings ~~finding~~ underscore the strong dependency of the sign and magnitudes of the nonlinear terms on their quasilinear counterparts, particularly at the extremes. The schematic in Fig. 6 contextualizes some of the pha

This analysis highlights the variety of eddy-eddy interactions involved, and the dependence of nonlinear contributions on the quasilinear part of the flow. However, further ~~exploration~~ analysis is needed to ~~address potential confounding factors, and uncover latent advection structures corresponding to flow patterns~~establish a direct link between circulation patterns and the relationship between the quasilinear and nonlinear advection components.

5.1 ~~Nonlinear-Quasilinear Interactions~~Advection regimes

We propose a diagnostic framework to isolate ~~advective regimes by evaluating the phase interactions between primary and secondary components,~~ advection regimes by analyzing the phase relationships between primary $\delta\mathcal{S}_{QL}$ and secondary $\delta\mathcal{S}_{NL}$ components alongside total advection~~. Specifically, we employ a scatterplot of $\delta\mathcal{S}_{QL}$ versus~~ , conditioned on the sign of pre-existing lower tropospehric DSE anomaly, $\mathcal{S}'_{Tot,Lag1}$. To this end, we construct a scatterplot of $\delta\mathcal{S}_{QL}$ versus $\delta\mathcal{S}_{NL}$, with data points

We find that for a given sign of pre-existing anomaly, the extreme deciles of $\delta\mathcal{S}_{QL}$, when combined with different phases of $\delta\mathcal{S}_{NL}$, usu When $\delta\mathcal{S}_{QL}$ is in its central deciles, fewer pairings with the extreme deciles of $\delta\mathcal{S}_{NL}$ correspond to coherent eddy configurations. Thus, we identify distinct nonlinear components associated with most eddy configurations where high or low magnitudes of $\delta\mathcal{S}_{NL}$ ~~, with points color-coded by the deciles of $\delta\mathcal{S}$. On such a plot, we demarcate~~ interact to amplify or decay the contributions from the extreme deciles of ~~all three variables and identify the nonlinear drivers for various regions thus obtained.~~ $\delta\mathcal{S}_{QL}$. We refer to such combin

5.1.1 ~~Positive background conditions~~

~~For positive background conditions ($\mathcal{S}'_{Tot,Lag1} > 0$), we implement the proposed strategy in Fig. 7. Amplification of growth is defined as $\delta\mathcal{S} > 0$ when $\delta\mathcal{S}_{QL} > 0$ and $\delta\mathcal{S}_{NL} > 0$, while saturation of growth is defined as $\delta\mathcal{S} > 0$ when $\delta\mathcal{S}_{QL} > 0$ and $\delta\mathcal{S}_{NL} < 0$. Amplification of decay is~~

5.1.2 Pre-existing positive anomaly

For daily cases with a pre-existing negative anomaly ($\mathcal{S}'_{Tot,Lag1} < 0$), we define Growth by positive daily advective tendency ($\delta\mathcal{S} > 0$) an As established earlier, $\delta\mathcal{S}$ is explained by $\delta\mathcal{S}_{QL}$ to the first order (Fig. 4a), but in the tails of $\delta\mathcal{S}$, both $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ can be comparable We study the joint distribution of $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$, and define Growth regimes for $\mathcal{S}'_{Tot,Lag1} > 0$ conditioned on $\delta\mathcal{S} > 0$ as follows:

- QL + NL Growth: Nonlinear amplification of Growth, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \gg 0$, i.e., when both $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$ are simultaneously in their positive extreme deciles
- NL Growth: Nonlinear Growth, defined ~~as~~ by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \sim 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 10th decile but $\delta\mathcal{S}_{QL}$ is in its central deciles (deciles 2 to 9)

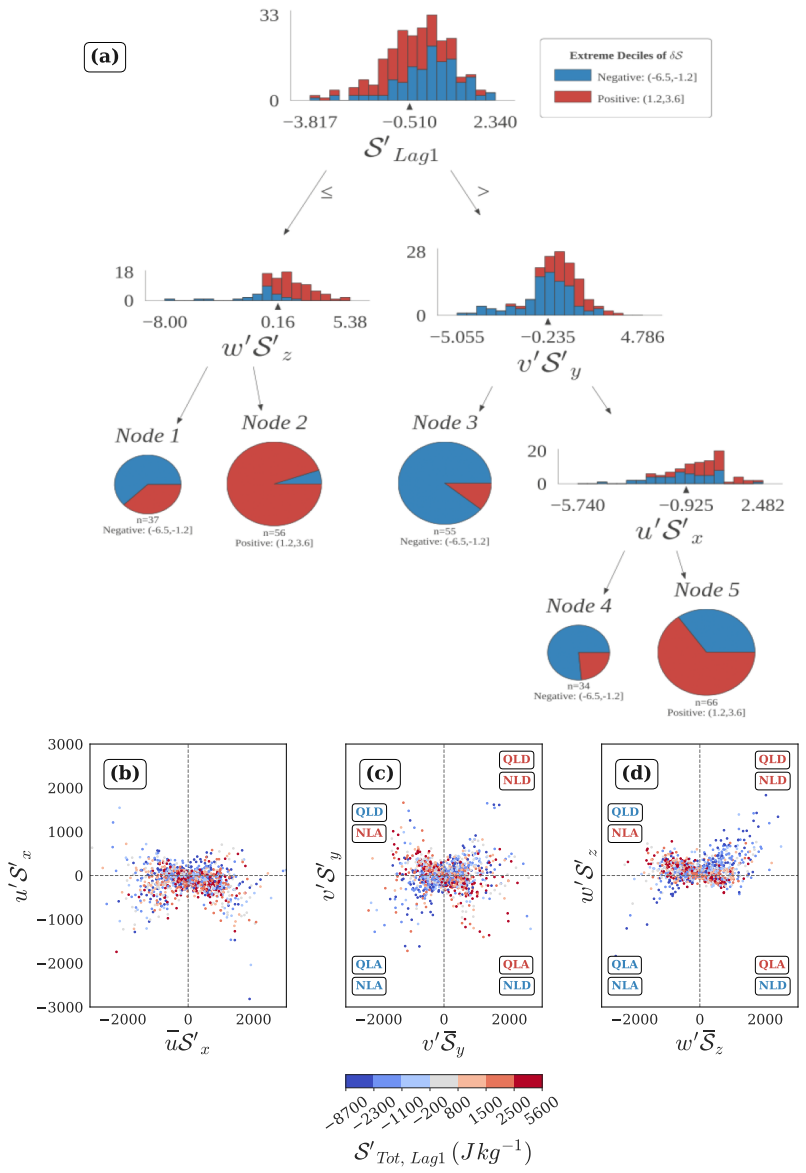

Figure 5. (a) ~~The standardized~~ Standardized decision tree model ~~for~~ identifying the drivers of ~~decile based~~ decile-based extremes of $\delta\mathcal{S}$ ~~filtered~~ for April~~; F1-score = 0.78, with the highest importance scores for $w'\mathcal{S}'_z$ and $\mathcal{S}'_{Tot,Lag1}$, followed by $v'\mathcal{S}'_y$ and $\bar{u}\mathcal{S}'_x$, respectively~~. ~~The nonlinear-quasilinear~~ The model achieves an F1-score of ~ 0.78, with early-stop criteria as follows: entropy decrease ≥ 0.03 at each level, $\geq 30\%$ sample size for level splitti (b-d) Nonlinear-quasilinear relationship in the ~~unfiltered, untreated dataset, for the b)~~ zonal, ~~c)~~ meridional, and ~~d)~~ vertical directions, colored by quantile based categories of $\mathcal{S}'_{Tot,Lag1}$. All the daily mean advection ~~quantities in the scatterplots~~ terms are expressed in Jkg^{-1}. Quadrants in (c) and (d) are labeled by the role of the advection components in amplifying or decaying the predominant sign of the pre-existing lower tropospheric DSE anomaly ($\mathcal{S}'_{Tot,Lag1}$) in that quadrant, with red for positive and blue for negative. For example, in (c), the first quadrant represents cases corresponding to quasilinear decay (QLD) and nonlinear decay (NLD) of a pre-existing negative anomaly by $v'\bar{\mathcal{S}}_y > 0$ and $v'\mathcal{S}'_y > 0$, respectively. Quadrants where advection aligns with the anomaly sign are labeled QLA (Quasilinear Amplification) and NLA (Nonlinear Amplification).

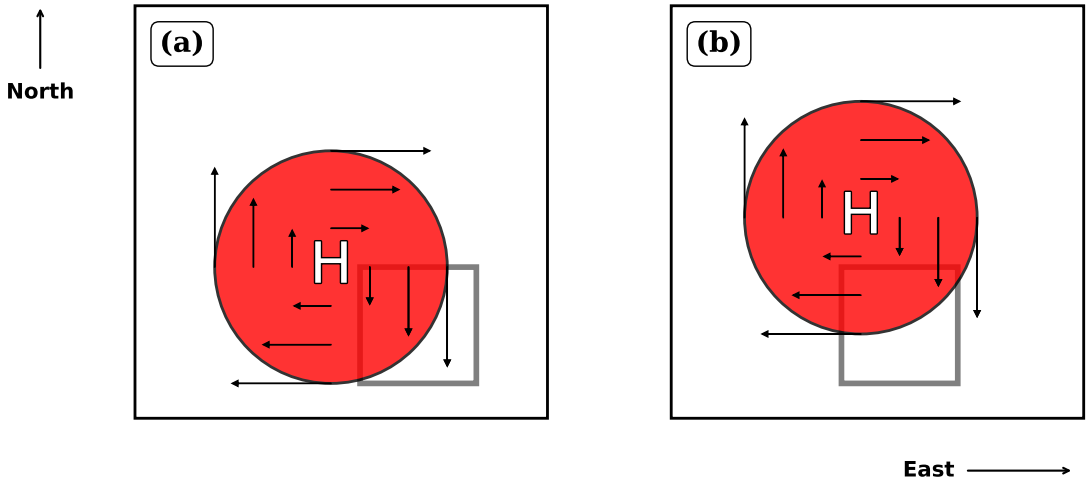

Figure 6. This schematic explains the eddy configurations responsible for some of the NL-QL phase interactions observed in Figs. 5a and 5b. The region of interest is shown by the transparent rectangle with thick grey borders. An anticyclonic warm core eddy associated with a high pressure anomaly is shown by the red circle, and the black solid arrows represent the circulation direction and strength which is proportional to the size of the arrows. The background conditions are given by: $\overline{DSE}_y < 0$, $\overline{DSE}_x \sim 0$. a) When an anticyclonic eddy is placed beyond the northwest corner of the region, the associated northerly flow ($v' < 0$) advects DSE' into the region due to $DSE'_y > 0$. Such a configuration causes nonlinear amplification of the DSE' of the region. At the same time, the meridional quasilinear advection term acts to dissipate the DSE' with upgradient transport of $\overline{DSE}$. Thus, the second quadrant in Fig. 5b is associated with quasilinear decay and nonlinear amplification by the meridional advection components. b) When the anticyclonic eddy is placed close to the north of the region, the associated easterly flow ($u' < 0$) advects DSE' out of the region due to $DSE'_x < 0$. Such a configuration causes nonlinear decay of the DSE' of the region, while the zonal quasilinear term acts to amplify the DSE' of the region due to strongly positive ($\bar{u} \gg 0$) mean zonal flow. Thus, the $\mathcal{S}'_{Tot,Leg1} > 0$ data points in the fourth quadrant of Fig. 5a are associated with quasilinear amplification and nonlinear decay of a pre-existing anomaly by the zonal advection components.

- QL Growth: Quasilinear Growth, defined by $\delta\mathcal{S}_{QL} \gg 0$ and $\delta\mathcal{S}_{NL} \sim 0$, i.e., when $\delta\mathcal{S}_{QL}$ is in its 10th decile but $\delta\mathcal{S}_{NL}$ is in its central deciles (deciles 2 to 9)

- NL Saturated Growth: Nonlinear saturation of Growth, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \gg 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 1st decile and $\delta\mathcal{S}_{QL}$ is in its 10th decile

Similarly, we define Decay regimes for $\mathcal{S}'_{Tot,Lag1} > 0$ conditioned on $\delta\mathcal{S} < 0$ as follows:

- QL + NL Decay: Nonlinear amplification of Decay, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \ll 0$, i.e., when both $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$ are simultaneously in their negative extreme deciles

- NL Decay: Nonlinear Decay, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \sim 0$, i.e., when ~~$\delta\mathcal{S}_{QL} < 0$ and $\delta\mathcal{S}_{NL} < 0$, while saturation of decay is defined as $\delta\mathcal{S} < 0$ when $\delta\mathcal{S}_{QL} < 0$ and $\delta\mathcal{S}_{NL} > 0$.~~ $\delta\mathcal{S}_{NL}$ is in its 1st decile but $\delta\mathcal{S}_{QL}$ is in its central deciles (deciles 2 to 9)

- QL Decay: Quasilinear Decay, defined by $\delta\mathcal{S}_{QL} \ll 0$ and $\delta\mathcal{S}_{NL} \sim 0$, i.e., when $\delta\mathcal{S}_{QL}$ is in its 1st decile but $\delta\mathcal{S}_{NL}$ is in its central deciles (deciles 2 to 9)

- NL Saturated Decay: Nonlinear saturation of Decay, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \ll 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 10th decile and $\delta\mathcal{S}_{QL}$ is in its 1st decile

~~The identified regimes highlight the role of~~ The above regimes are demarcated in the $\delta\mathcal{S}_{QL} - \delta\mathcal{S}_{NL}$ phase space for a pre-existing positi

We observe that $v'\mathcal{S}'_y$ ~~in anomaly growth amplification , and the role of~~ contributes to the nonlinear growth of a pre-existing positive anomaly, while $u'\mathcal{S}'_x$ ~~in saturation of growth. Conversely, in the decay phase, $u'\mathcal{S}'_x$ appears to contribute to amplifying decay, while $v'\mathcal{S}'_y$ plays a more prominent role in saturation of decay, as seen in the QL Decayregime where $v'\mathcal{S}'_y > 0$. Notably, in the NL Decay regime, $v'\mathcal{S}'_y$ makes a minor contribution to amplification of decay.~~ plays an important role in the nonlinear saturation of its These observations are in agreement with the description in Fig. 6. We also find that both the horizontal nonlinear terms drive nonlinear de

These findings underscore the variability in nonlinear-quasilinear dynamics ~~across the anomaly's lifecycle~~ and provide a framework for interpreting the evolution of advection structures ~~under positive background conditions~~modifying a pre-existing positive anomaly. The roles of these nonlinear drivers in driving growth and decay of $\mathcal{S}'_{Tot,Lag1} > 0$ are better ~~understood by inspecting unfiltered~~ clarified by inspecting the nonlinear-quasilinear scatterplots in the context of the phase and magnitude of $\delta\mathcal{S}$, $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ (~~Supplementary Figs. S11, S12~~Appendix Figs. A1, A2).

5.1.3 ~~Negative background conditions~~

~~For negative background conditions ($\mathcal{S}'_{Tot,Lag1} < 0$), the regimes of growth and decay ($\delta\mathcal{S} < 0$ and $\delta\mathcal{S} > 0$ respectively) are shown in Fig. 9. Amplification of growth is defined as $\delta\mathcal{S} < 0$ when $\delta\mathcal{S}_{QL} < 0$ and $\delta\mathcal{S}_{NL} < 0$, while saturation of growth is defined as $\delta\mathcal{S} < 0$ when $\delta\mathcal{S}_{QL} < 0$ and $\delta\mathcal{S}_{NL} > 0$. Amplification of decay is defined as $\delta\mathcal{S} > 0$ when $\delta\mathcal{S}_{QL} > 0$ and $\delta\mathcal{S}_{NL} > 0$, while saturation of decay is defined as $\delta\mathcal{S} > 0$ when $\delta\mathcal{S}_{QL} > 0$ and $\delta\mathcal{S}_{NL} < 0$~~We find that a given advection regime consists of multiple no

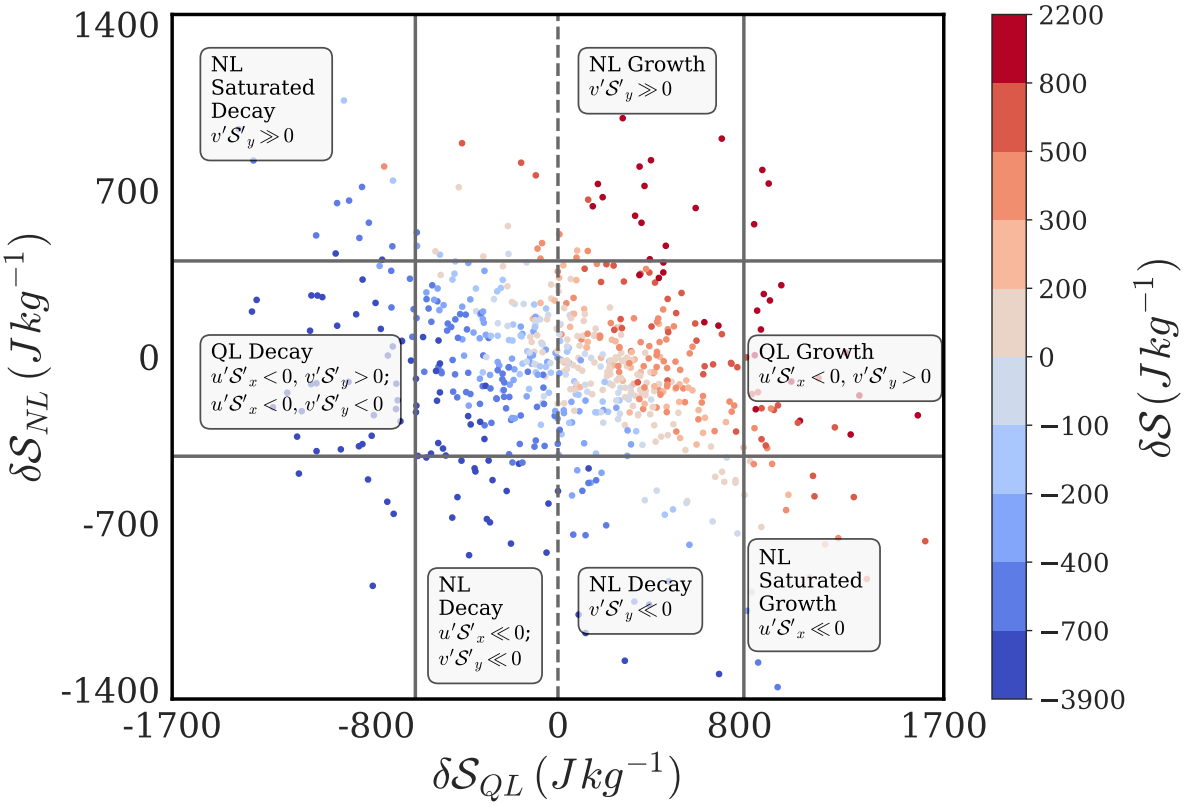

Figure 7. ~~Under $\mathcal{S}'_{Tot,Lag1} > 0$ conditions during April, the phase relationship between $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$, colored by deciles of unfiltered $\delta\mathcal{S}$. The solid grey lines demarcate the 10th and 90th percentiles of $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$. Growth of the positive background anomaly is defined as $\delta\mathcal{S} > 0$ conditions. Decay of the positive background anomaly is defined as $\delta\mathcal{S} < 0$. Different regions in the phase space are annotated as "QL" or "NL", representing the characteristic roles of quasilinear and nonlinear advection, respectively, along with the drivers of NL advection where applicable. QL Growth (Decay) is defined as $\delta\mathcal{S}_{QL}$ in its 10th (1st) decile and $\delta\mathcal{S}_{NL}$ between its 2nd and 9th deciles. NL Growth (Decay) is defined as $\delta\mathcal{S}_{NL}$ in its 10th (1st) decile and $\delta\mathcal{S}_{QL}$ between its 2st and 9th deciles. NL saturated Growth (Decay) is defined as $\delta\mathcal{S}_{QL}$ in its 10th (1st) decile, and $\delta\mathcal{S}_{NL}$ in its 1st (10th) decile.~~ This figure shows the phase relationship between $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ colored by deciles of $\delta\mathcal{S}$, for pre-existing positive lower tropospheric DSE anomaly (

The dominant nonlinear combinations associated with the advection regimes for $\mathcal{S}'_{Lag1} > 0$ conditions are clarified via composite maps (Fig.

When $\mathcal{S}'_{Tot,Lag1} < 0$8), $u'\mathcal{S}'_x$ contributes to amplification, and $v'\mathcal{S}'_y$ aids saturation of the growth of the negative anomaly, as observed in Fig. 9. As before, the roles of these nonlinear drivers in driving growth and decay of $\mathcal{S}'_{Tot,Lag1} < 0$ conditions are better understood by inspecting unfiltered nonlinear-quasilinear scatterplots in the context of $\delta\mathcal{S}$, $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ (Supplementary Figs. S13, S14). described in relation to the spatial configuration of eddy fields below.

5.2 Advective Regimes

We find that extreme deciles of $\delta\mathcal{S}_{QL}$, in combination with each of the extreme deciles of $\delta\mathcal{S}$, and the body of $\delta\mathcal{S}_{NL}$, usually correspond to distinct eddy configurations with respect to the region of interest. Extreme deciles of $\delta\mathcal{S}_{NL}$ combined with extreme deciles of $\delta\mathcal{S}_{QL}$ also sometimes represent distinct eddy configurations. In the body of $\delta\mathcal{S}_{QL}$, extreme deciles of $\delta\mathcal{S}_{NL}$ represent distinct eddy configurations only in a few instances. Thus, we have identified nonlinear drivers for most of the ways in which $\delta\mathcal{S}_{NL}$ interacts with $\delta\mathcal{S}_{QL}$ in the extreme deciles of $\delta\mathcal{S}_{QL}$. The skew of $\delta\mathcal{S}_{NL}$ in the extreme deciles of $\delta\mathcal{S}_{QL}$ (Fig. 4f) can be attributed to eddy configurations driven by nonlinear drivers, that represent eddy-eddy covariances.

The circulations associated with regions of the phase space in Fig. 7 are clarified in Fig. 8, by plotting the composite representations of mass-weighted quantities averaged over the box, wherever they corresponded to eddy configurations. Descriptions as follows.

1. NL Growth ($v'\mathcal{S}'_y v'\mathcal{S}'_y \gg 0$): When the northerly winds associated with either a western/ northwestern anticyclonic disturbance engulf the region, the magnitude of $\mathcal{S}'_y > 0$ strengthens to positive meridional gradient of DSE' intensifies – reaching about half the magnitude of the negative climatological gradient meridional gradient of $\overline{DSE}$. Simultaneously, the area averaged area-averaged northerly v' gains significant strength to almost completely lie in strength and enters its positive extreme decile. Such a configuration enhances the meridional anomalous advection of $\mathcal{S}'$ DSE' into the region, thereby amplifying the effect of positive $\delta\mathcal{S}_{QL}$ associated with such a configuration.

2. QL Growth ($u'\mathcal{S}'_x < 0, v'\mathcal{S}'_y > 0$): When comparable magnitudes of anomalous northerly and easterly winds associated with an anticyclonic disturbance operate in the eastern and southern halves of the region, under comparable magnitudes of $\mathcal{S}'_y > 0$ and $\mathcal{S}'_x < 0$, the easterly winds remove a significant part of the $\mathcal{S}'$ deposited by the northerly winds, rendering $\delta\mathcal{S}_{NL} \sim 0$. As a result, the net advection is a function of $\delta\mathcal{S}_{QL}$ alone in this configuration.

3. NL Saturated Growth ($u'\mathcal{S}'_x u'\mathcal{S}'_x \ll 0$): When an anticyclone centered to the north of the region has advanced sufficiently into the region, the the associated easterly winds pass through the southern half of the region., and the strength of $u' < 0$ increases to A strongly negative $\mathcal{S}'_x$ zonal gradient of DSE' is maintained due to the anticyclone, and the strength of $u' < 0$ increases to more than half of the zonal mean flow. Thusthe anomalously warm conditions associated with the anticyclone primarily acting on the Consequently, u' acts to dissipate a comparable magnitude of $\mathcal{S}'$ as remove a large part of the DSE' deposited by $\bar{u}$.

The magnitude of v' is small in this configuration and does not yield a considerable magnitude of $v'\mathcal{S}'_y$ ~~in spite of~~ despite a large magnitude of positive $\mathcal{S}'_y$.

4. NL Decay (~~$u'\mathcal{S}'_x$~~$u'\mathcal{S}'_x \ll 0$): When a cyclonic disturbance to the north/ west interacts primarily with the western boundary of the (anomalously warm) region of interest, ~~u' becomes aligned with and comparable in magnitude to $\bar{u}$. Simultaneously, the sign of $\mathcal{S}'_x$ becomes positive , and its magnitude is much larger than $\mathcal{S}'_y$ because the northern boundary of the region is not affected similarly. Thus, the zonal nonlinear component becomes comparable to its quasilinear counterpart, and such a configuration involves significant contributions by $u'\mathcal{S}'_x$ in causing decay of the positive background anomaly~~the zonal eddy winds b Coupled with anticyclonic winds to the south of the region, the zonal eddy winds gain strength, and drive the nonlinear decay of the We note that the NL Decay regime has another component in Fig. 7, characterized solely by $v'\mathcal{S}'_y \gg 0$ within the positive half of $\delta\mathcal{S}_{QL}$. This subset corresponds to 6 samples exhibiting extreme negative $\delta\mathcal{S}$. Upon inspection of its composite map, we find that it does not represent a coherent eddy configuration. We exclude this component from Fig. 8 and focus our interpretation on the $u'\mathcal{S}'_x \ll 0$ dominated branch of NL Decay described above.

5. ~~QL Decay~~NL Saturated Decay ($v'\mathcal{S}'_y \gg 0$): When a cyclone centered to the northwest of the anomalously warm region induces a neg

6. QL Decay ($u'\mathcal{S}'_x < 0, v'\mathcal{S}'_y > 0$): The QL regime is revisited in the decay phase when a cyclone has advanced sufficiently into the region with anomalous conditions corresponding to strong and opposing zonal and meridional gradients of ~~$\mathcal{S}'$~~DSE', and strong westerly and southerly anomalous winds. In such a configuration, the nonlinear advection terms cancel out, rendering $\delta\mathcal{S}_{NL}$ of low magnitude.

~~(Figure 8f)Further, we see that~~ We study the conditional probability distribution of deciles of $\mathcal{S}'_{Tot,Lag1}$ given the occurrence of the advection regimes active during $\mathcal{S}'_{Tot,Lag1} > 0$ conditions (Table 2). NL Growth and QL Growth regimes are predominantly concentrated in the central deciles of $\mathcal{S}'_{Tot,Lag1} > 0$, suggesting that linear and nonlinear growth mechanisms are most active during moderate positive anomalies. In contrast, the NL Saturated Growth regime ~~predominantly occurs during the largest positive background anomaly conditions. This indicates that saturation of growth happens in the middle stages of an anticyclone's transition over a region, as observed earlier in Fig. 4c. Both quasilinear regimes (growth and decay) favour moderately positive background conditions. Both kinds of circulations correspond to cases when the corresponding disturbance has advanced sufficiently into the region, growing or decaying the positive background anomaly. It is interesting to note that~~ preferentially occurs during more extreme positive values of $\mathcal{S}'_{Tot,Lag1}$. NL Decay and QL Decay regimes are most active during moderately positive $\mathcal{S}'_{Tot,Lag1}$ conditions, and NL Saturated Decay during neutral $\mathcal{S}'_{Tot,Lag1}$ conditions. Overall, most of the ~~quasilinear regime does not correspond to insignificant magnitudes of the nonlinear terms. Rather nonlinear advection happens to be large in magnitude in at least two directions, but opposite in sign so that their net effect is negligible. On the other hand, the NL decay regime is active during more extreme~~ advection regimes are distributed across neutral, moderate and extreme positive deciles of $\mathcal{S}'_{Tot,Lag1}$ ~~days. This drives the idea that such a configuration is encountered earlier in a cyclone's transition over the region.~~ , indicating that large Growth and Decay processes remain active throughout the lifecycle of $\mathcal{S}'_{Tot,Lag1} > 0$.

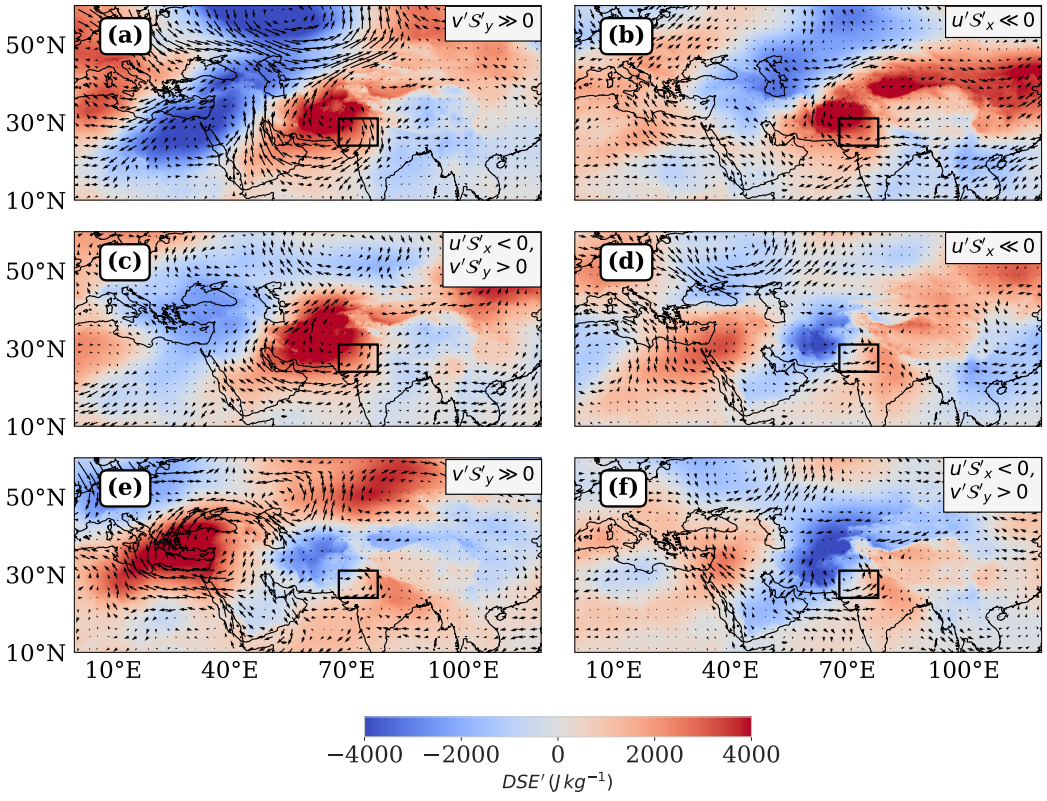

Figure 8. ~~During April,~~ This figure shows the composite representations of horizontal eddy wind ~~anomaly~~ vectors and ~~S'~~ S'_{Tot} corresponding to the dominant nonlinear combination per advection ~~regimes~~ regime amplifying and dissipating ~~$S' > 0$ conditions~~a pre-existing positive lower tropospheric DSE anomaly ($S'_{Tot,Log1} > 0$) during April, as defined in Fig. 7. The order of regimes is as follows: (a) NL Growth, (b) NL Saturated Growth, (c) QL Growth, (d) NL Decay, (e) ~~QL~~ NL Saturated Decay~~. The colors range from blue to red in relation with values of S' from −4000 to 4000 Jkg^{-1}.~~ , (f) ~~The percentage occurrence of advective regimes identified in Fig~~QL Decay. ~~8 by decile of background anomaly.~~ The Supplementary Table S4 clarifies the representativeness of an advection regime by the plotted nonlinear combination in the column "% Contribution" The color coding represents S'_{Tot} values ranging from -4000 to 4000 $J\,kg^{-1}$, with stronger hues indicating larger magnitudes.

The circulations associated with regions of the phase space in Fig. 9 are clarified in Fig. 10, by plotting the composite representations of mass-weighted quantities averaged over the box, wherever they corresponded to eddy configurations. Descriptions as follows.

Advection Regime	5	6	7	8	9	10
NL Growth	0.19	0.25	0.19	0.12	0.19	0.06
NL Saturated Growth	0.13	0.20	0.13	0.00	0.07	0.47
QL Growth	0.11	0.22	0.11	0.33	0.11	0.11
NL Decay	0.14	0.10	0.19	0.19	0.24	0.14
NL Saturated Decay	0.00	0.54	0.15	0.15	0.15	0.00
QL Decay	0.06	0.25	0.28	0.16	0.09	0.16

Table 2. This table shows the conditional probability distribution of ranked deciles of $\mathcal{S}'_{Tot,Lag1}$ given the occurrence of an advection regime defined for $\mathcal{S}'_{Lag1} > 0$ conditions during April. We compute the conditional probability table by cross-tabulating the deciles of $\mathcal{S}'_{Tot,Lag1}$ with the advection regimes identified for $\mathcal{S}'_{Lag1} > 0$ conditions in Fig. 8, and then normalizing each row by corresponding row totals. Each cell value represents the conditional probability $P\left(\mathcal{S}'_{Lag1} \in D_i \middle| \text{regime} \in A_j\right)$ where D_i is the i^{th} decile of $\mathcal{S}'_{Tot,Lag1}$ and A_j is the advection regime on the j^{th} row.

5.1.1 Pre-existing negative anomaly

For daily cases with a pre-existing negative anomaly ($\mathcal{S}'_{Tot,Lag1} < 0$), we define Growth by negative daily advective tendency ($\delta\mathcal{S} < 0$) an
We study the joint distribution of $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$, and define Growth regimes for $\mathcal{S}'_{Tot,Lag1} < 0$ conditioned on $\delta\mathcal{S} < 0$ as follows:

- QL + NL Growth: Nonlinear amplification of Growth, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \ll 0$, i.e., when both $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$ are simultaneously in their negative extreme deciles
- NL Growth: Nonlinear Growth, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \sim 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 1st decile but $\delta\mathcal{S}_{QL}$ is in its central deciles (deciles 2 to 9)
- QL Growth: Quasilinear Growth, defined by $\delta\mathcal{S}_{QL} \ll 0$ and $\delta\mathcal{S}_{NL} \sim 0$, i.e., when $\delta\mathcal{S}_{QL}$ is in its 1st decile but $\delta\mathcal{S}_{NL}$ is in its central deciles (deciles 2 to 9)
- NL Saturated Growth: Nonlinear saturation of Growth, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \ll 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 10th decile and $\delta\mathcal{S}_{QL}$ is in its 1st decile

Similarly, we define Decay regimes for $\mathcal{S}'_{Tot,Lag1} < 0$ conditioned on $\delta\mathcal{S} > 0$ as follows:

- QL + NL Growth: When the southerly winds associated with a western cyclonic disturbance slightly offset to the south engulf the region, the magnitude of $\mathcal{S}'_y > 0$ strengthens to about a fourth of the magnitude Decay: Nonlinear amplification

of Decay, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \gg 0$, i.e., when both $\delta\mathcal{S}_{NL}$ and $\delta\mathcal{S}_{QL}$ are simultaneously in their positive extreme deciles

- NL Decay: Nonlinear Decay, defined by $\delta\mathcal{S}_{NL} \gg 0$ and $\delta\mathcal{S}_{QL} \sim 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 10th decile but $\delta\mathcal{S}_{QL}$ is in its central deciles (deciles 2 to 9)
- QL Decay: Quasilinear Decay, defined by $\delta\mathcal{S}_{QL} \gg 0$ and $\delta\mathcal{S}_{NL} \sim 0$, i.e., when $\delta\mathcal{S}_{QL}$ is in its 10th decile but $\delta\mathcal{S}_{NL}$ is in its central deciles (deciles 2 to 9)
- NL Saturated Decay: Nonlinear saturation of Decay, defined by $\delta\mathcal{S}_{NL} \ll 0$ and $\delta\mathcal{S}_{QL} \gg 0$, i.e., when $\delta\mathcal{S}_{NL}$ is in its 1st decile and $\delta\mathcal{S}_{QL}$ is in its 10th decile

The above regimes are demarcated in the $\delta\mathcal{S}_{QL} - \delta\mathcal{S}_{NL}$ phase space for a pre-existing negative anomaly in Fig. 9, and the identified no

When $\mathcal{S}'_{Tot,Lag1} < 0$, $u'\mathcal{S}'_x$ contributes to nonlinear amplification of growth, and $v'\mathcal{S}'_y$ aids nonlinear saturation of the growth of the negative ~~climatological gradient. Simultaneously, the area averaged southerly v' gains significant strength to almost completely lie in its positive extreme decile. Such a configuration enhances the negative advection of $\mathcal{S}'$ by anomalous winds into the region, thereby amplifying the effect of negative~~ anomaly, as observed in Fig. 9. In the decay phase, the roles of the nonlinear terms

As before, the roles of these nonlinear terms in driving growth and decay of $\mathcal{S}'_{Tot,Lag1} < 0$ conditions are better understood by inspecting nonlinear-quasilinear scatterplots in the context of the phase and magnitude of $\delta\mathcal{S}$, $\delta\mathcal{S}_{QL}$ ~~associated with such a configuration.~~ and $\delta\mathcal{S}_{NL}$ (Appendix Figs. A3, A4).

The circulations associated with the dominant nonlinear combinations representing the advection regimes for $\mathcal{S}'_{Lag1} < 0$ conditions (Fig

1. ~~NL saturated Growth : When the southerly winds associated with a cyclone advect positive S' from the south without much contribution from the zonal arm of the cyclone, the meridional nonlinear advection acts in opposite phase of the quasilinear advection ($\delta\mathcal{S}_{QL} < 0$)associated with such a configuration.~~ QL + NL Growth ($u'\mathcal{S}'_x \ll 0$): When a cyclonic disturbance is

2. NL Saturated Growth ($v'\mathcal{S}'_y \ll 0$): When a cyclonic disturbance centered to the west induces a negative meridional gradient of DSE

3. QL Growth ($u'\mathcal{S}'_x < 0$, $v'\mathcal{S}'_y > 0$): When comparable magnitudes of ~~anomalous southerly and easterly winds~~ eddy southerlies and westerlies associated with a cyclonic disturbance ~~operate~~ prevail in the eastern and southern halves of the region respectively, and coincide with similarly strong but opposite signed gradients of DSE' in respective directions, ~~under comparable magnitudes of $\mathcal{S}'_y < 0$ and $\mathcal{S}'_x > 0$,~~ the southerly winds remove a ~~significant part of the $\mathcal{S}'$~~ substantial portion of the colder DSE' deposited by the ~~northerly winds, rendering~~ westerly winds. This mutual cancellation of opposing contributions r $\delta\mathcal{S}_{NL} \sim 0$. As a result, the net advection is ~~a function of~~ governed solely by $\delta\mathcal{S}_{QL}$ ~~alone~~ in this configuration.

4. QL + NL Decay ~~: When an anticyclonic disturbance to the west interacts primarily with the western boundary of the (anomalously cold)region of interest, northerly v' is the major component of anomalous horizontal winds over the region.~~

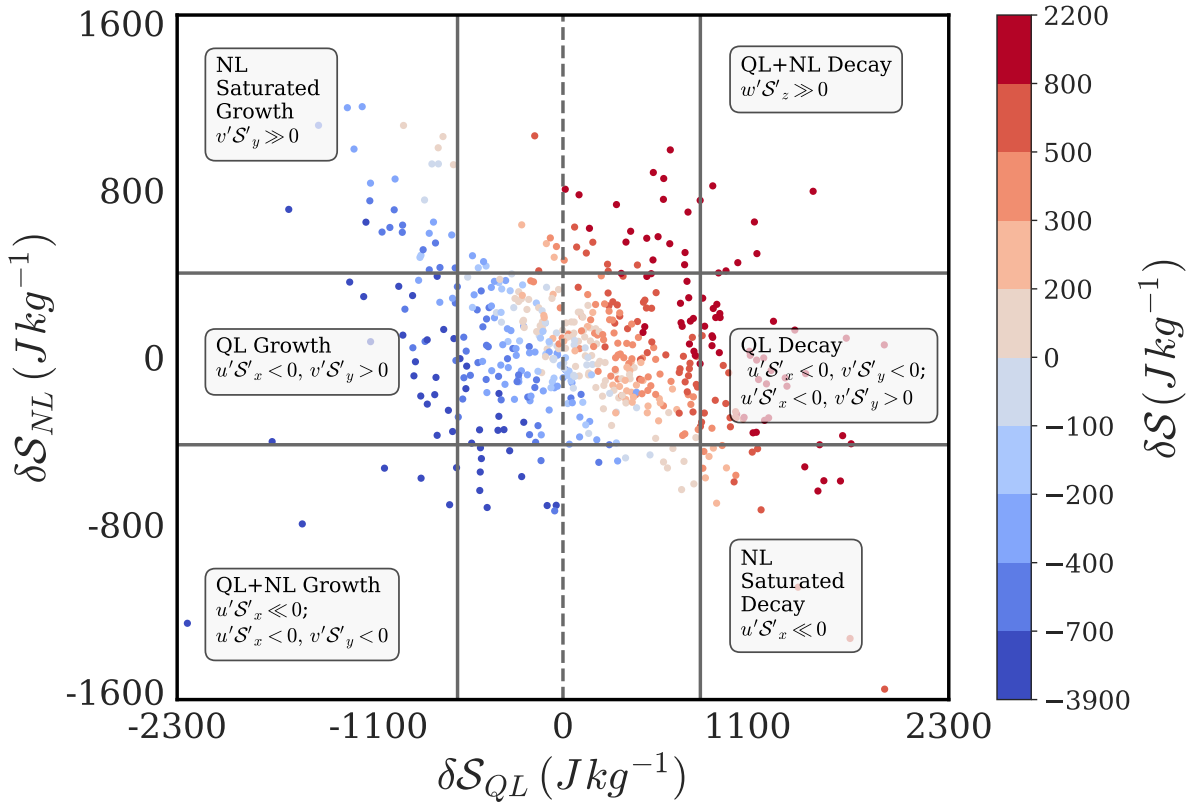

Figure 9. This figure shows the phase relationship between $\delta\mathcal{S}_{QL}$ and $\delta\mathcal{S}_{NL}$ colored by deciles of $\delta\mathcal{S}$, for pre-existing negative lower tropospheric DSE

~~When these winds advect over $\mathcal{S}'_y < 0$, the meridional winds advect colder $\mathcal{S}'$ into the region, amplifying the advection of colder $\overline{\mathcal{S}}$ due to $\overline{\mathcal{S}}_y < 0$.~~ ($w'\mathcal{S}'_z \gg 0$): When northerly anticyclonic eddy winds are strongly coupled with subsiding vertical eddy win

5. NL Saturated Decay ($u'\mathcal{S}'_x \ll 0$): We find a configuration representative of a trough-ridge couplet, with the southward dipping arm of cyclonically curved easterly winds of the ridge sweeping across the region. Coupled with a warm trough to the northwest of the region, ~~such a configuration involves~~ the easterly arm ~~advecting cold $\mathcal{S}$~~ advects cold DSE' into the region, while the northerly winds are associated with ~~($\delta\mathcal{S}_{QL} > 0$)~~positive quasilinear advection of DSE'. Thus, the zonal nonlinear advection acts in the opposite phase of the quasilinear advection associated with such a configuration.

6. QL Decay ~~: The QL regime is revisited in the decay phase when a cyclone has advanced sufficiently into the region with anomalous conditions corresponding to strong but opposing zonal and meridional gradients of $\mathcal{S}'$, and strong easterly and northerly anomalous winds. In such a configuration, the nonlinear advection terms cancel out, rendering $\delta\mathcal{S}_{NL}$ of low magnitude. During April, composite representations of horizontal wind anomaly vectors and $\mathcal{S}'$ corresponding to~~

~~advection regimes amplifying and dissipating $\mathcal{S}' < 0$ conditions, as defined in Fig. 7. The order of regimes is as follows: a) QL + NL Growth, b) NL Saturated Growth, c) QL Growth, d) QL + NL Decay, e) NL Saturated Decay, f) QL Decay. The colors range from blue to red in relation with values of $\mathcal{S}'$ from -4000 to $4000\ Jkg^{-1}$. g) The percentage occurrence of advective regimes identified in Fig. 10 by deciles of background anomaly.~~ ($u'\mathcal{S}'_x < 0$, $v'\mathcal{S}'_y < 0$, $w'\mathcal{S}'_z \gg 0$): The QL regime is rev

Thus, ~~there are indications that such mapping between circulation patterns and the NL-QL phase space can effectively map different phases of the transition of a large scale disturbance over the region of interest. Figure 10f shows similar indications as to nonlinear and quasilinear regime preferences of the negative background states as Fig. 8f did for growth and decay of the positive background anomaly.~~

~~For negative "background" conditions $\mathcal{S}'_{Tot,Lag1} < 0$~~for a negative pre-existing anomaly, the roles of the meridional and zonal nonlinear terms were strongly reversed in contributions to amplification and dissipation. $v'\mathcal{S}'_y$ was out of phase with $\delta\mathcal{S}$ during both ~~phases~~ growth and decay since $\mathcal{S}'_y$ was of the same sign as $\bar{\mathcal{S}}_y$, and was out of phase with $\delta\mathcal{S}_{QL}$. $u'\mathcal{S}'_x$ strongly aided the amplification of the ~~"background" state~~ negative pre-existing anomaly since the easterly winds associated with cyclonic winds lined up with the climatological zonal winds, which were nearly always easterly. $w'\mathcal{S}'_z$ made significant contributions to the amplification phase.

We examine the conditional probability distribution of the deciles of $\mathcal{S}'_{Tot,Lag1}$ given the occurrence of the advection regimes active during $\mathcal{S}'_{Tot,Lag1} < 0$ conditions (Table 3). QL + NL Growth exhibits a strong preference for the left tail of $\mathcal{S}'_{Tot,Lag1}$, nonlinearly amplifying the growth of pre-existing extreme negative values of $\mathcal{S}'_{Tot,Lag1}$. NL Saturated Growth spans both extreme and moderate negative values of $\mathcal{S}'_{Tot,Lag1} < 0$, while QL Growth is most active during moderate values. QL + NL Decay and QL Decay are strongly concentrated in the left tail of $\mathcal{S}'_{Tot,Lag1}$, and NL Saturated Decay is more evenly distributed across the negative range of $\mathcal{S}'_{Tot,Lag1}$. Notably, all the Growth and Decay regimes here preferentially occupy moderate and extreme deciles of $\mathcal{S}'_{Tot,Lag1} < 0$ – a contrast to the regime behavior observed for $\mathcal{S}'_{Tot,Lag1} > 0$ conditions in section 5.1.2.

Advection Regime	**1**	**2**	**3**	**4**	**5**
QL + NL Growth	0.33	0.27	0.07	0.20	0.13
NL Saturated Growth	0.15	0.37	0.26	0.30	0.09
QL Growth	0.17	0.17	0.26	0.30	0.09
QL + NL Decay	0.71	0.00	0.29	0.00	0.00
NL Saturated Decay	0.14	0.29	0.24	0.10	0.24
QL Decay	0.34	0.28	0.19	0.16	0.03

Table 3. This table shows the same conditional probability distribution defined in Table 2, but filtered for cases with a negative pre-existing anomaly ($\mathcal{S}'_{Lag1} < 0$) during April. We compute the conditional probability table by cross-tabulating the deciles of $\mathcal{S}'_{Tot,Lag1}$ with the advection regimes identified for $\mathcal{S}'_{Lag1} < 0$ conditions in Fig. 10, and then normalizing each row by corresponding row totals.

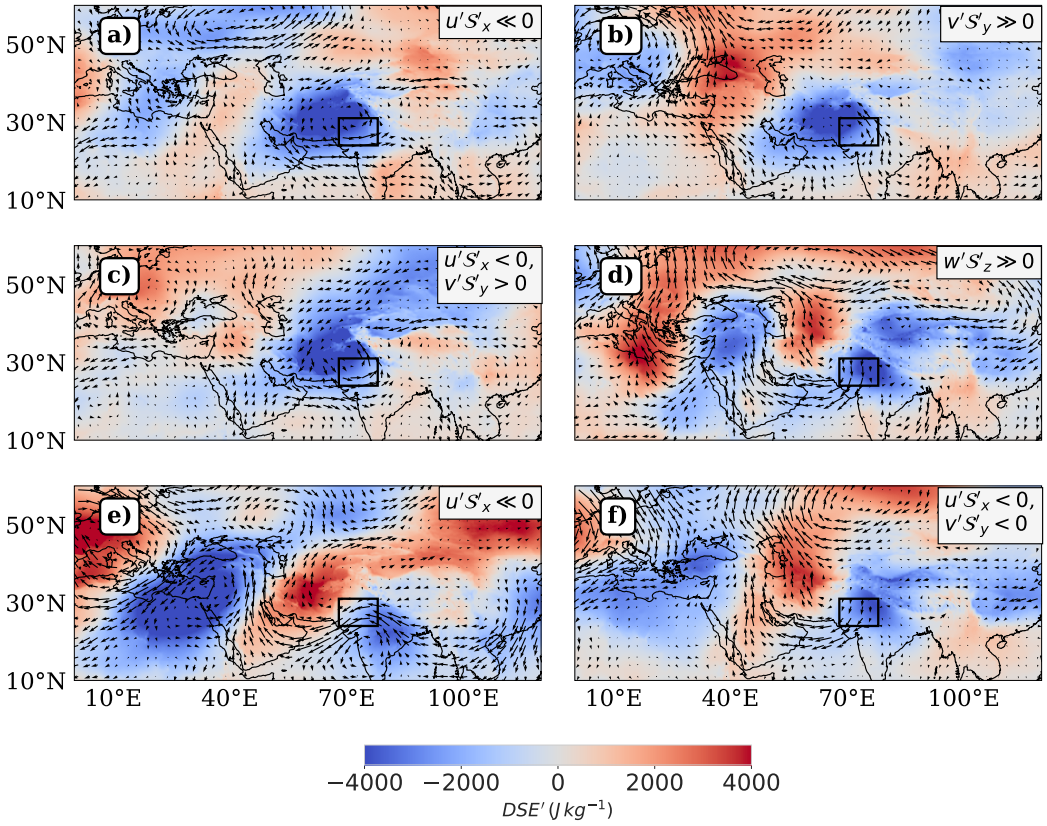

Figure 10. This figure shows the composite representations of horizontal wind anomaly vectors and $\mathcal{S}'_{Tot,Lag1}$ corresponding to the dominant nonlinear combination per advection regime amplifying and dissipating a pre-existing negative lower tropospheric DSE anomaly ($\mathcal{S}'_{Tot,Lag1} < 0$) during April, as defined in Fig. 7. The order of regimes is as follows: (a) QL + NL Growth, (b) NL Saturated Growth, (c) QL Growth, (d) QL + NL Decay, (e) NL Saturated Decay, f) QL Decay. The Supplementary Table S5 clarifies the representativeness of an advection regime by the plotted nonlinear combination in the column "% Contribution" The color coding represents $\mathcal{S}'_{Tot}$ values ranging from -4000 to 4000 $J\,kg^{-1}$, with stronger hues indicating larger magnitudes.

We repeat the entire analysis for the period of March, and arrive at distinct ~~advective~~ advection regimes associated with it (~~Supplementary Figs. S15, S16).~~ Appendix Figs. B1, B2). Overall, there are indications that such mapping between circulation patterns and the NL-QL phase space can effectively map different phases of the transition of a large scale disturbance over the region of interest. The conditional probability distribution of deciles of $\mathcal{S}'_{Tot,Lag1}$ in Table 3f shows similar indications as to nonlinear and quasilinear regime preferences of the pre-existing negative anomaly states as Table 2f did for growth and decay of the pre-existing positive anomaly states.

6 Summary and Discussion

In ~~the current~~ this study, we have aimed to identify the contribution of atmospheric circulation to daily change in daily mean lower tropospheric ~~DSE~~ DSE during March-April in ~~a~~ a subtropical heatwave hotspot in South Asia. We choose to study ~~DSE~~ DSE changes because we saw that while the daily temperature field is best explained by surface and radiative processes, *daily changes* in temperature are better explained by advection of ~~DSE~~ DSE in the lower troposphere. Thus, understanding ~~DSE~~ daily lower tropospheric DSE changes helps understand ~~day-day temperature variability~~ the variability in daily temperature changes over the region, particularly when the change is large.

We find that the velocity and ~~DSE~~ DSE anomaly fields are vertically coherent, so the choice of this quantity allows us to focus on the impacts of upper tropospheric forcing on the lower tropospheric ~~DSE~~ DSE budget, regardless of the form of the forcing itself (wave propagation, breaking, etc.,). Forcing due to balanced dynamics manifests itself as correlations between the different quasilinear advective terms, whereas such correlations may be absent for other kinds of circulations (such as monsoon intraseasonal oscillations) which may be more prevalent during or after the onset of the monsoon in this region.

We ~~choose to~~ use decision trees to identify primary and secondary drivers of ~~DSE~~ DSE variability because decision trees are both explainable and provide a relative quantification of the contribution of individual physical processes to ~~DSE~~ DSE variability. Each model output is supported with an analysis of the ~~unfiltered, untreated~~ raw data, ensuring the decision rules are both interpretable and grounded in robust underlying relationships.

We find that the "primary" quasilinear relationship capturing the inverse relationships between the meridional and vertical quasilinear components, and the meridional and zonal quasilinear components of advection effectively separates the presence of eddies from incoherent flows, especially at large magnitudes. Smaller magnitudes of advection can represent eddy structures even when these relationships are not obeyed, but most instances with small magnitudes represent the absence of eddies.

Further, we find that the events that constitute the tail of the distribution of ~~changes in $\mathcal{S}$~~ $\delta\mathcal{S}$ represent a rich variety of eddy configurations or "~~advective~~ advection regimes" that are missing in the literature of ~~heatwave studies~~temperature variability. We also find a ~~significant~~ strong presence of quasilinear regimes in the tails of ~~this distribution of advection, and~~$\delta\mathcal{S}$, and, counter to expectation, the presence of nonlinear regimes in the ~~quasilinearly governed body of this distribution.~~ body of $\delta\mathcal{S}$. We obtain a circulation map for some of these regimes that are satisfactorily explained in terms of the horizontal configuration of the eddi

The following list summarises the important observations and takeaways from our analysis:

1. We ~~observe~~ find that the quasilinear advective terms are important across all terciles of ~~the distribution of daily changes in DSE, whereas~~ daily advective DSE changes. In contrast, the nonlinear terms are important primarily in the tails of the distribution. This observation is ~~in-line~~ in line with previous studies, which propose that atmospheric macroturbulence is predominantly quasilinear – characterised by weak eddy-eddy interactions – ~~,~~ and that eddy-eddy interactions are important for explaining the tail ~~behaviour~~ behavior of such macroturbulence. Further, ~~different advection structures involving distinct nonlinear contributions are important for~~ the particular eddy-eddy interactions active in the tails of the distribution differ between extreme positive and negative ~~advection~~advective DSE changes, depending on the sign of the ~~background DSE~~ pre-existing DSE anomaly.

2. We see that ~~vertical (QG omega) advection of DSE~~ the combined vertical and zonal advection of DSE overwhelms the meridional advection of DSE, leading to an interesting situation where *southward* advection leads to an increase in lower tropospheric ~~DSE~~DSE, opposite to what one would expect if temperature variability was considered as ~~horizontal,~~ a problem of one-dimensional mixing of the climatological equator-pole temperature gradient by eddies. ~~The importance of vertical advection may be linked to the large static stability in our region of interest.~~

3. In the presence of eddies, ~~this~~ the quasigeostrophically driven meridional-vertical wind coupling becomes associated with ~~$\mathcal{S}'_x$~~ a zonal gradient of DSE' across the region of interest that persists ~~over~~ during the transition of ~~the~~ an eddy due to a preferred zonal direction of propagation ~~. Due to large zonal flows prevalent during the considered times of the year, the advection of $\mathcal{S}'$~~ (Fig. 3). The importance of vertical advection may be linked to the large static stability in our region of interest. The advection of DSE' by zonal mean flow is comparable in magnitude to the primary meridional and vertical terms due to large zonal flows prevalent during the considered times of the year. Thus, the zonal term has decisive control over the net effect of the ~~quasigeostrophic coupling~~ quasilinear coupling involving all these three terms. This contribution ~~seems~~ appears to have been ignored in previous ~~analyses~~ analyzes due to the focus on advection of isotherms by eddy flow fields, ignoring the zonal mean flow inhomogeneities between regions of a given latitude.

4. ~~The previous point implies a direct relationship between the presence of eddies and large values of v', w' and $\mathcal{S}'_x$ obeying the above-mentioned relationship. Further, different eddy configurations render one or more of $\mathcal{S}'_y$, u' and $\mathcal{S}'_z$ large, leading to nonlinear terms of large magnitude~~ In addition to the large eddy winds in the meridional and vertical directions and a large Since the nonlinear advection terms result from the product of these terms in respective directions, different nonlinear terms become large under different eddy configurations and become important in the tails of the distribution ~~. Thus, the tails of the~~ of $\delta\mathcal{S}$~~distribution is composed of different eddy configurations, and is unlikely to be explained by a single nonlinear term~~. We note that ~~eddy structures~~ eddies can also exist in the body of the distribution of daily advection of DSE, but we do not ~~analyse~~ analyze them separately since they represent small ~~day-day changes in $\mathcal{S}$~~advective daily changes in DSE.

5. $u'\mathcal{S}'_x$ is also neglected in studies considering ~~multiplicative noise~~ nonlinear advection processes (**?**), presumably due to the Lagrangian approach which is not affected by the secondary gradients induced across a region. ~~We find that $u'\mathcal{S}'_x$ plays an important role in restricting the growth of positive anomalies over our region of interest. $v'\mathcal{S}'_y$ amplifies the growth of positive anomalies, and the roles of both the horizontal nonlinear terms are largely reversed in the decay phase of positive anomalies. It is also found that $w'\mathcal{S}'_z$ is a strong driver of decay of negative anomalies, while $v'\mathcal{S}'_y$ acts to restrict the decay. Negative anomalies are amplified by $u'\mathcal{S}'_x$, and $w'\mathcal{S}'_z$ does not have a strong role to play in this phase.~~ We find that zonal nonlinear advection term is mostly negative, irrespective of the sign of the pre-existing anomaly. This is because It is also found that $w'\mathcal{S}'_z$ is a strong driver of amplification and decay of negative anomalies, with a larger influence in causing deca

6. We identify different regimes of interaction between quasilinear and nonlinear advection terms, and show that each corresponds to distinct ~~background states~~ states of the pre-existing anomaly and flow patterns related to the phasing and location of upper-tropospheric eddies. We demonstrate that large $\delta\mathcal{S}$ can ~~occur not only when quasilinear advection is large~~ not only be driven by large quasilinear advection (as has been observed before), but ~~can also exist~~ also by large nonlinear advection when quasilinear advection is small~~but nonlinear advection is large~~. On the other hand, large values of $\delta\mathcal{S}$ can be driven purely by quasilinear advection ~~represents nonlinear regimes where the individual nonlinear terms are large in magnitude but their net effect is negligible, and such cases are also consistently mapped to certain eddy configurations.~~

   when the individual nonlinear terms are large in magnitude but opposite in phase, yielding a negligible net effect; such cases are als
   Thus, we go beyond ~~just~~ suggesting the qualitative importance of nonlinear advection to temperature variability, ~~but~~ and identify the particular nonlinear terms that are important for different flow configurations.

7. We obtain the contribution of identified regimes to different stages of the anomaly lifecycle, and find that different configurations corresponding to the same sign of extreme advection occupy different ~~background statesstages~~. Such an observation across both signs of ~~the background anomaly ,~~ pre-existing anomaly and both periods ~~under analysis,~~ of analysis, March and April, strongly suggests the existence of different pathways to extreme ~~low tropospheric states of energy~~states of low tropospheric DSE.

Thus, we clarify the relationship between the energetics and flow patterns in our region of interest. We demonstrate that our approach allows for a more nuanced picture of ~~DSE~~ DSE variability (and extreme variability in particular) than can be achieved by traditional composites which are insensitive to the contributions of individual quasilinear and nonlinear terms.

The decision tree based methodology for identifying terms governing the differences in sign of advection may be used as a reference to identify the primary dynamical relationships interacting with a region, and the regression based analysis for identifying regions in the distribution that obey different or supplementary sets of dynamics. While we have used this methodology to understand the distribution of $\delta\mathcal{S}$, we expect this approach should be useful in any such analysis of Reynolds-decomposed conservation laws, and could potentially also include the contributions due to the diabatic terms, if available.

Understanding the flow configurations that constitute the tail of $\delta\mathcal{S}$ might be used to identify contextual variables governing the interrelationships between terms of advection and provide a concise summary of the relationship between the statistics and dynamics related to an observable. In the case of balanced large scale dynamics, our framework for identifying circulation patterns can help examine the structures of upper tropospheric flows interacting with a region.

Combined with the analysis mapping these circulations to stages of the anomaly lifecycle, our framework has the potential to identify persistent structures sustaining anomalies, and if there are distinct configurations associated with growth and decay. A similar approach was taken by **?**, who characterized distinct "weather types" associated with rainfall anomalies. However, there is still some work needed for using such a framework for relating circulations to ~~T2m~~ T_{2m} variability, which is affected by other processes as previously noted.

In our case, we have also observed large durations of heatwaves associated with balanced dynamics during May, Jun, which will be studied using this framework in future work. Another useful practical application might be to explain periods of large observed skewness using such a framework. Further, the phenomenology of circulation structures obtained by studying the PDF of advection in this manner can be used towards a circulation-oriented evaluation of climate models and seasonal forecasts.

Appendix A: Nonlinear drivers of growth and decay of a pre-existing anomaly: April

A1 Growth of positive anomaly

We analyze the nonlinear terms driving the amplification and saturation phases of growth of a pre-existing positive anomaly (as defined in

- Growth of a pre-existing positive anomaly is strongly associated with positive quasilinear advective convergence (Figs. A1a, A1b ar

- Figs. A1a, A1b and A1c suggest that for large magnitude of $\delta\mathcal{S}_{QL} > 0$, $v'\mathcal{S}'_y$ is consistently positive, while $u'\mathcal{S}'_x$ is consistently negative, and $w'\mathcal{S}'_z$ is positive but of low magnitude. Figs. A1d and A1e clarify that the opposite phases of $v'\mathcal{S}'_y$ and $u'\mathcal{S}'_x$ since $\delta\mathcal{S}_{NL} > 0$ when $v'\mathcal{S}'_y \gg 0$, and $\delta\mathcal{S}_{NL} < 0$ when $u'\mathcal{S}'_x \ll 0$.

- **Nonlinear Amplification**: When $\delta\mathcal{S}_{NL} > 0$, all three nonlinear terms act in phase with it. However, $v'\mathcal{S}'_y$ has the largest magnitude during most instances and likely drives $\delta\mathcal{S}_{NL} > 0$ when $\delta\mathcal{S} > 0$ and $\delta\mathcal{S}_{QL} > 0$.

- **Nonlinear Saturation**: When $\delta\mathcal{S}_{NL} < 0$, the largest nonlinear contributions come from $u'\mathcal{S}'_x$, and coincide with $\delta\mathcal{S}_{QL} > 0$. This identifies $u'\mathcal{S}'_x$ as the driver of saturation of growth of a pre-existing positive anomaly.

A2 Decay of positive anomaly

Using a similar approach, we analyze conditions associated with the decay of a pre-existing positive anomaly (Fig. A2).

- Decay of a pre-existing positive anomaly can be associated with positive as well as negative quasilinear advective convergence (Figs. A2a, A2b and A2c).

- **Nonlinear Amplification**: $u'\mathcal{S}'_x$ amplifies decay by acting in phase with $\delta\mathcal{S}_{QL} < 0$ (Fig. A2e).

- Focussing on the days corresponding to $\delta\mathcal{S}_{QL} > 0$, large magnitude of $v'\mathcal{S}'_y < 0$ (Fig. A2d) independently drives nonlinear decay.

- **Nonlinear Saturation**: Here, we focus on the days corresponding to $\delta\mathcal{S}_{QL} < 0$. $u'\mathcal{S}'_x$, $v'\mathcal{S}'_y$, and $w'\mathcal{S}'_z$ act in phase (> 0) to drive saturation of decay by opposing the phase of $\delta\mathcal{S}_{QL} < 0$ with $\delta\mathcal{S}_{NL} > 0$ (Figs. A2d, A2e, A2f). The magnitude of $v'\mathcal{S}'_y$ is largest, and is the likely driver of saturation of decay of a pre-existing positive anomaly.

A3 Growth of negative anomaly

We analyze the potential nonlinear advection terms driving the amplification and saturation phases of growth of a pre-existing negative anomaly, as follows.

- Growth of the pre-existing negative anomaly is strongly associated with negative quasilinear advective convergence barring a few lo

- The meridional nonlinear term has large excursions on the positive side (A3d), while the zonal nonlinear term has large excursions on the negative side (A3e), suggesting that $v'\mathcal{S}'_y$ may act to saturate, while $u'\mathcal{S}'_x$ may act to amplify the growth of the anomaly.

- **Nonlinear Amplification**: For $\delta\mathcal{S}_{NL} < 0$, both the horizontal terms are in phase, but $u'\mathcal{S}'_x$ is consistently larger in magnitude (Figs. A3e, A3f), affirming its role in amplifying growth.

- **Nonlinear Saturation**: For $\delta\mathcal{S}_{NL} > 0$, both the horizontal nonlinear terms are predominantly positive. However, $v'\mathcal{S}'_y$ shows a consistently larger magnitude than $u'\mathcal{S}'_x$ (joint distribution not shown), affirming the role of $v'\mathcal{S}'_y$ in saturating growth of the negative anomaly.

- $w'\mathcal{S}'_z$ exhibits lower activity as compared to the horizontal nonlinear terms, indicating its limited influence in the growth phase.

A4 Decay of negative anomaly

Next, we analyze the nonlinear advection terms associated with the decay of a negative pre-existing anomaly (Fig. A4).

- Decay of a pre-existing negative anomaly is strongly associated with positive quasilinear advective convergence barring a few low r

– $v'\mathcal{S}'_y$ makes large contributions to both sides of zero, but its negative excursions align with large positive values of $u'\mathcal{S}'x$ or $w'\mathcal{S}'_z$. This reduces its independent role in saturation.

– **Nonlinear Amplification**: $w'\mathcal{S}'_z$ is strongly associated with large instances of $\delta\mathcal{S}_{NL} > 0$ (Fig. A4f), corroborating the observations in Fig. 4d.

– **Nonlinear Saturation**: $u'\mathcal{S}'_x$ is consistently of large magnitude when $\delta\mathcal{S}_{NL} < 0$ and $\delta\mathcal{S}_{QL} > 0$ (Fig. A4e), emerging as a strong driver of saturation of decay of a pre-existing negative anomaly.

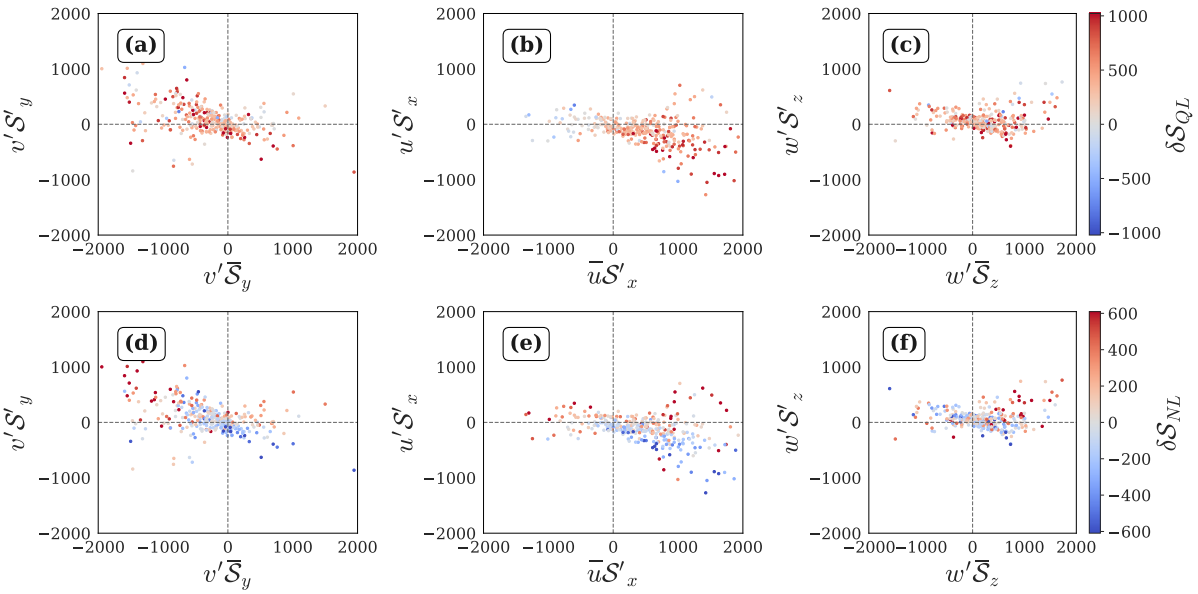

Figure A1. Nonlinear-quasilinear relationships for $\mathcal{S}'_{Tot,Lag1} > 0$ conditions, filtered for $\delta\mathcal{S} > 0$ days during April. (a-c) Quantiles of $\delta\mathcal{S}_{QL}$, (d-f) quantiles of $\delta\mathcal{S}_{NL}$, with blue for negative values and red for positive values and stronger colors indicating larger magnitudes. All quantities are expressed in Jkg^{-1}.

Appendix B: Advection regimes: March

We repeated the analysis for March and identified the regimes given in Figs. B1 and B2.

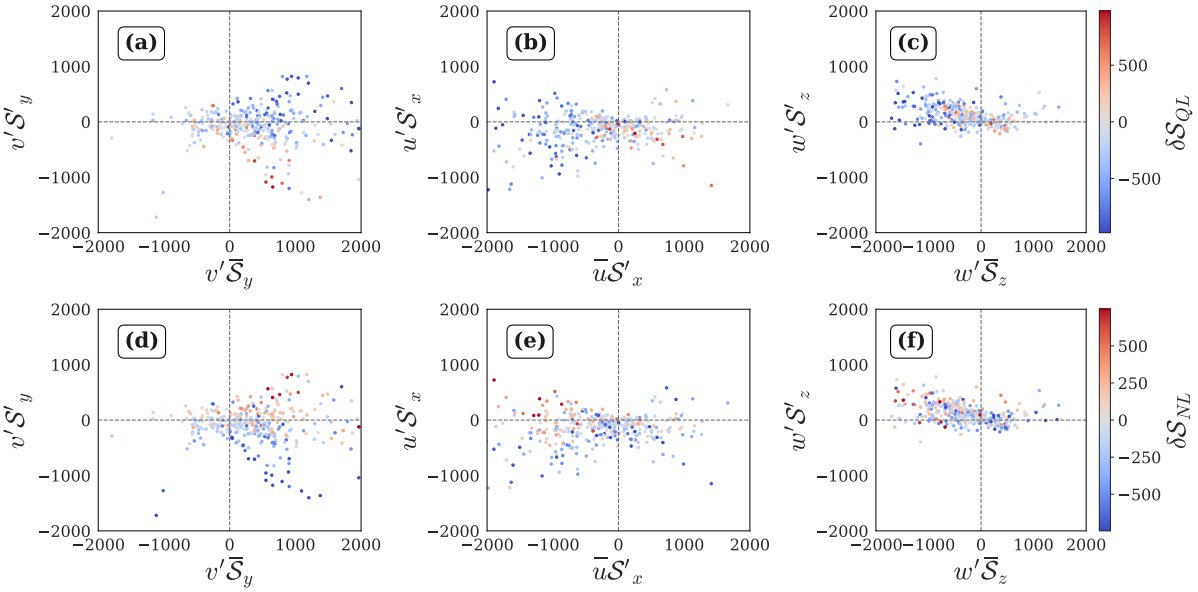

Figure A2. Nonlinear-quasilinear relationships for $\mathcal{S}'_{Tot,Log1} > 0$ conditions, filtered for $\delta\mathcal{S} < 0$ days during April. (a-c) Quantiles of $\delta\mathcal{S}_{QL}$, (d-f) quantiles of $\delta\mathcal{S}_{NL}$, with blue for negative values and red for positive values, and stronger colors indicating larger magnitudes. All quantities are expressed in $J\,kg^{-1}$.

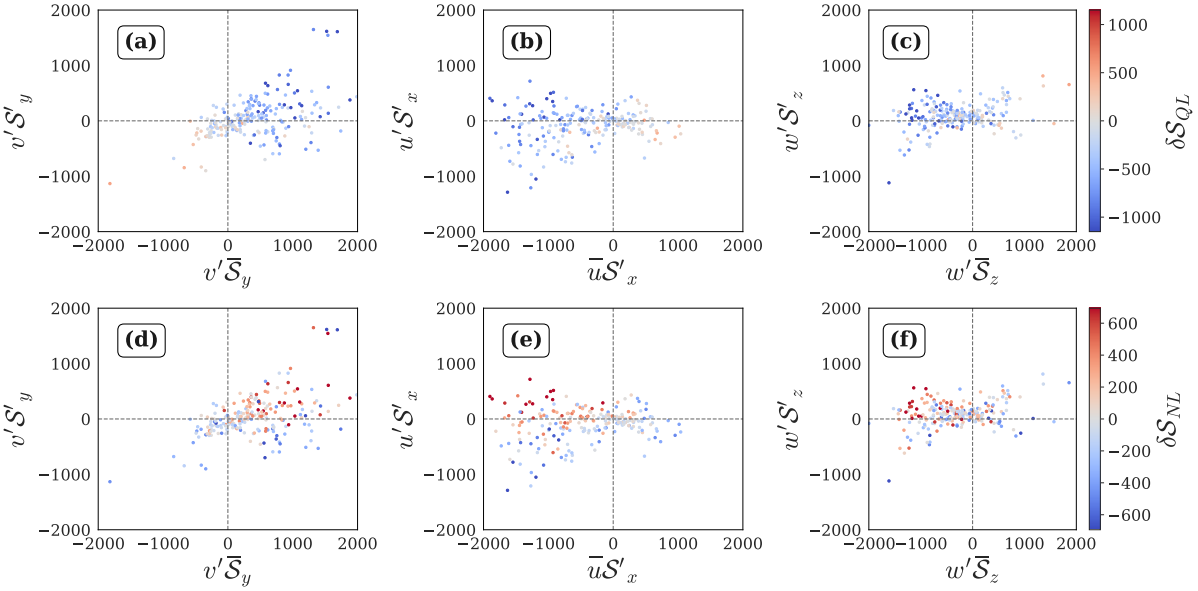

Figure A3. Nonlinear-quasilinear relationships for $\mathcal{S}'_{Tot,Log1} < 0$ conditions, filtered for $\delta\mathcal{S} < 0$ days during April. (a-c) Quantiles of $\delta\mathcal{S}_{QL}$, (d-f) quantiles of $\delta\mathcal{S}_{NL}$, with blue for negative values and red for positive values, and stronger colors indicating larger magnitudes. All quantities are expressed in $J\,kg^{-1}$.

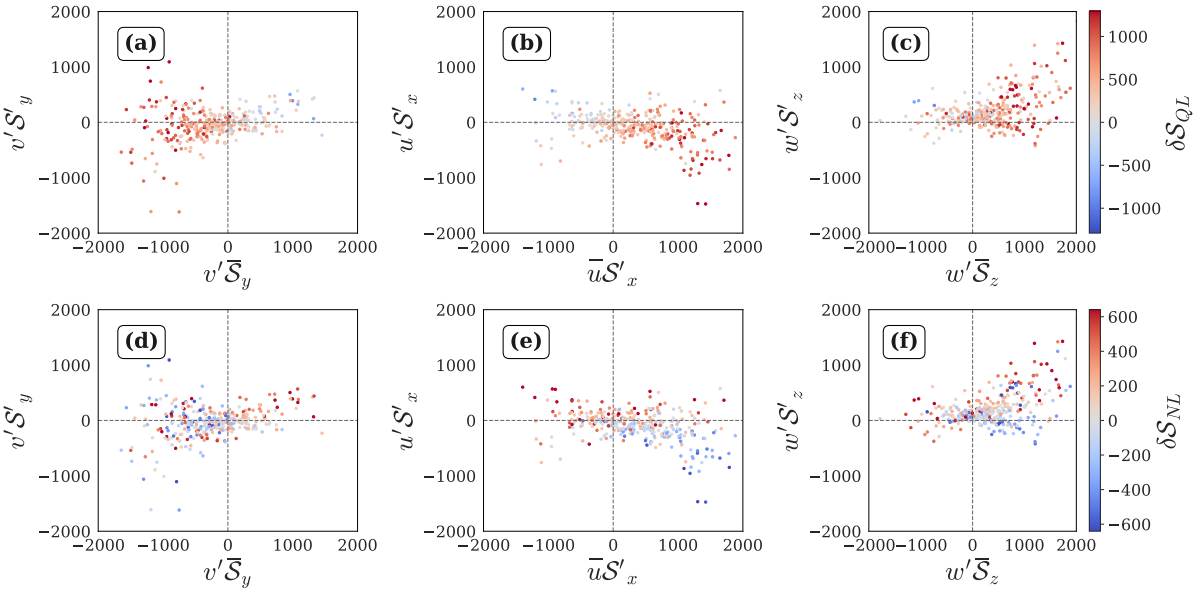

Figure A4. Nonlinear-quasilinear relationships for $\mathcal{S}'_{Tot,Lag1} < 0$ conditions, filtered for $\delta\mathcal{S} > 0$ days during April. (a-c) Quantiles of $\delta\mathcal{S}_{QL}$, (d-f) quantiles of $\delta\mathcal{S}_{NL}$, with blue for negative values and red for positive values, and stronger colors indicating larger magnitudes. All quantities are expressed in $J\,kg^{-1}$.

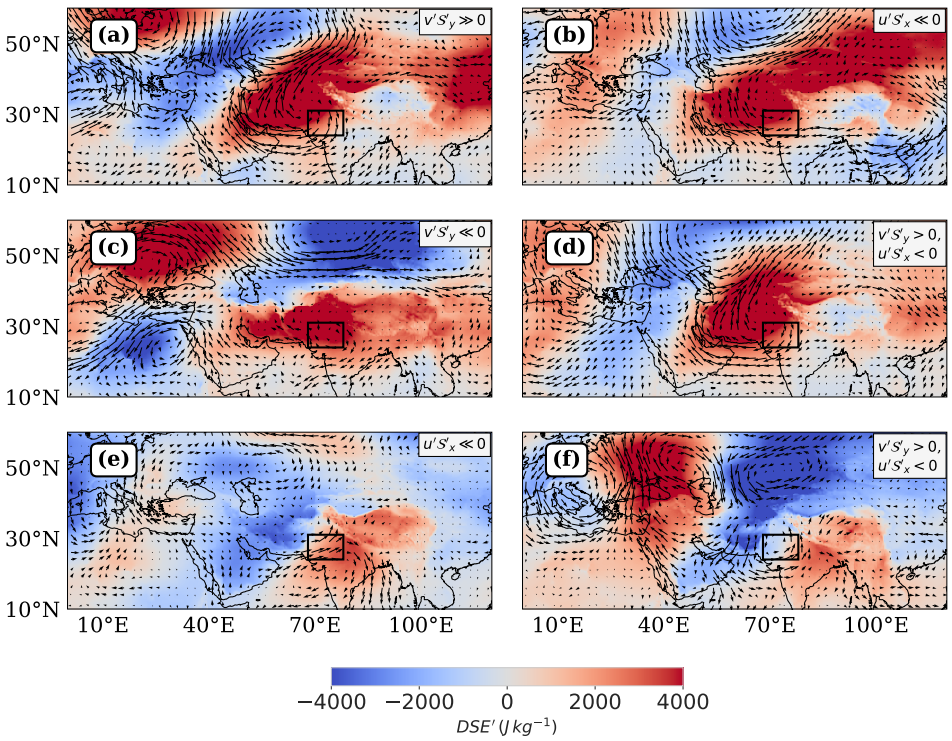

Figure B1. Composite representations of horizontal wind anomaly vectors and $\mathcal{S}'_{Tot}$ corresponding to advection regimes associated with Growth and Decay of $\mathcal{S}'_{Tot,Lag1} > 0$ conditions during March. The color coding represents $\mathcal{S}'_{Tot}$ values ranging from -4000 to 4000 $J\,kg^{-1}$.

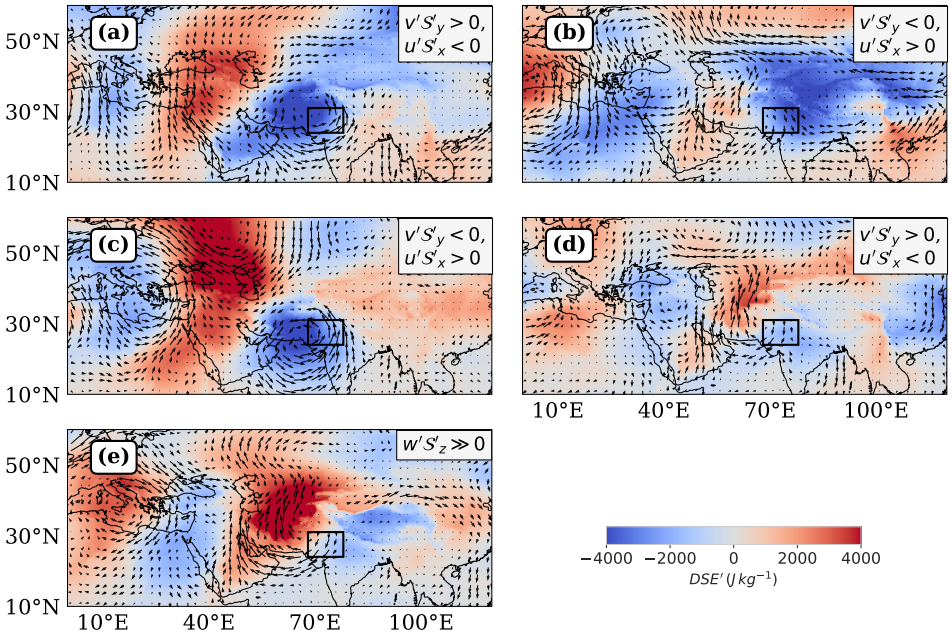

Figure B2. Composite representations of horizontal wind anomaly vectors and mass-weighted lower tropospheric DSE' corresponding to advection regimes amplifying and dissipating $\mathcal{S}'_{Tot,Lag1} < 0$ conditions during March. The color coding represents $\mathcal{S}'_{Tot}$ values ranging from -4000 to $4000\ J\,kg^{-1}$.

Author contributions. Both authors jointly conceptualized the project and contributed to the development of the methodology. Hardik M. Shah developed the computer code used for data preparation and analysis, applied statistical techniques for analysis and modeling, prepared data visualization and presentation ideas, and the original draft. Joy M. Monteiro acquired the funding for this project, acted as the project administrator and supervisor, arranged for the computing resources, provided mentorship, and helped with review and revision of the manuscript.

Competing interests. No competing interests are present.

Acknowledgements. This work was supported by the Department of Science and Technology grant DST/INT/ISR/P-40/~~2023.~~ 2023 and Ministry of Earth Sciences Monsoon Mission 3 grant IITM/MM-III/ 2023/IND-4. We thank Vishal Dixit and Nili Harnik for helpful discussions shaping our methodology and interpretation.

References

Bouchet, F., Nardini, C., and Tangarife, T.: Kinetic Theory and Quasilinear Theories of Jet Dynamics, in: Zonal Jets: Phenomenology, Genesis, and Physics, edited by Galperin, B. and Read, P. L., Cambridge University Press, ISBN 978-1-107-04388-6, 2019.

Burrows, W. R., Benjamin, M., Beauchamp, S., Lord, E. R., McCollor, D., and Thomson, B.: CART Decision-Tree Statistical Analysis and Prediction of Summer Season Maximum Surface Ozone for the Vancouver, Montreal, and Atlantic Regions of Canada, Journal of Applied Meteorology and Climatology, 34, 1848 – 1862, https://doi.org/10.1175/1520-0450(1995)034<1848:CDTSAA>2.0.CO;2, 1995.

Chen, L. and Dirmeyer, P. A.: The relative importance among anthropogenic forcings of land use/land cover change in affecting temperature extremes, Climate Dynamics, 52, 2269–2285, https://doi.org/10.1007/s00382-018-4250-z, 2019.

Delsole, T. and Farrell, B. F.: The Quasi-Linear Equilibration of a Thermally Maintained, Stochastically Excited Jet in a Quasigeostrophic Model, Journal of the Atmospheric Sciences, 53, 1781–1797, https://doi.org/10.1175/1520-0469(1996)053<1781:TQLEOA>2.0.CO;2, 1996.

Dirmeyer, P. A., Chen, L., Wu, J., Shin, C. S., Huang, B., Cash, B. A., and others: Verification of land-atmosphere coupling in forecast models, reanalyses, and land surface models using flux site observations, Journal of Hydrometeorology, 19, 375–392, https://doi.org/10.1175/JHM-D-17-0152.1, 2018.

Domeisen, D. I. V., Eltahir, E. A. B., Fischer, E. M., Knutti, R., Perkins-Kirkpatrick, S. E., Schär, C., Seneviratne, S. I., Weisheimer, A., and Wernli, H.: Prediction and projection of heatwaves, Nature Reviews Earth & Environment, 4, 36–50, https://doi.org/10.1038/s43017-022-00371-z, 2023.

for Medium-Range Weather Forecasts (ECMWF), E. C.: ERA5: Data Documentation – Known Issues, https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation#ERA5:datadocumentation-Knownissues, accessed October 2025, 2025.

Francis, J. A. and Vavrus, S. J.: Evidence linking Arctic amplification to extreme weather in mid-latitudes, Geophysical Research Letters, 39, https://doi.org/https://doi.org/10.1029/2012GL051000, 2012.

Gagne, D. J., McGovern, A., and Brotzge, J.: Classification of Convective Areas Using Decision Trees, Journal of Atmospheric and Oceanic Technology, 26, 1341 – 1353, https://doi.org/10.1175/2008JTECHA1205.1, 2009.

Garfinkel, C. I. and Harnik, N.: The Non-Gaussianity and Spatial Asymmetry of Temperature Extremes Relative to the Storm Track: The Role of Horizontal Advection, J. Climate, 30, 445–464, https://doi.org/10.1175/JCLI-D-15-0806.1, 2016.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Fernández del Río, J., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357–362, https://doi.org/10.1038/s41586-020-2649-2, 2020.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2001.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Horton, D. E., Johnson, N. C., Singh, D., Swain, D. L., Rajaratnam, B., and Diffenbaugh, N. S.: Contribution of changes in atmospheric circulation patterns to extreme temperature trends, Nature, 522, 425–427, https://doi.org/10.1038/nature14550, 2015.

Hoyer, S. and Hamman, J.: xarray: N-D labeled arrays and datasets in Python, Journal of Open Research Software, 5, https://doi.org/10.5334/jors.148, 2017.

Hunter, J. D.: Matplotlib: A 2D graphics environment, Computing in Science & Engineering, 9, 90–95, https://doi.org/10.1109/MCSE.2007.55, 2007.

Kornhuber, K., Osprey, S., Coumou, D., Petri, S., Petoukhov, V., Rahmstorf, S., and Gray, L.: Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern, Environmental Research Letters, 14, 054 002, https://doi.org/10.1088/1748-9326/ab13bf, publisher: IOP Publishing, 2019.

Linz, M., Chen, G., and Hu, Z.: Large-Scale Atmospheric Control on Non-Gaussian Tails of Midlatitude Temperature Distributions, Geophysical Research Letters, 45, 9141–9149, https://doi.org/https://doi.org/10.1029/2018GL079324, 2018.

Linz, M., Chen, G., Zhang, B., and Zhang, P.: A Framework for Understanding How Dynamics Shape Temperature Distributions, Geophysical Research Letters, 47, e2019GL085 684, https://doi.org/https://doi.org/10.1029/2019GL085684, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL085684, 2020.

Loikith, P. C. and Neelin, J. D.: Short-Tailed Temperature Distributions over North America and Implications for Future Changes in Extremes, Geophysical Research Letters, 42, 8577–8585, https://doi.org/10.1002/2015GL065602, 2015.

Marston, J. B. and Tobias, S. M.: Recent Developments in Theories of Inhomogeneous and Anisotropic Turbulence, Annual Review of Fluid Mechanics, 55, 351–375, https://doi.org/10.1146/annurev-fluid-120720-031006, 2023.

Mayer, J., Mayer, M., and Haimberger, L.: Consistency and Homogeneity of Atmospheric Energy, Moisture, and Mass Budgets in ERA5, Journal of Climate, 34, 3955 – 3974, https://doi.org/10.1175/JCLI-D-20-0676.1, 2021.

McKinney, W. et al.: Data structures for statistical computing in python, in: Proceedings of the 9th Python in Science Conference, vol. 445, pp. 51–56, Austin, TX, 2010.

Met Office: Cartopy: a cartographic python library with a Matplotlib interface, Exeter, Devon, https://scitools.org.uk/cartopy, 2010 - 2015.

Miralles, D. G., Teuling, A. J., Heerwaarden, C. C. v., and Arellano, J. V.-G. d.: Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation, Nature Geoscience, 7, 345–349, https://doi.org/10.1038/ngeo2141, 2014.

Monteiro, J. M. and Caballero, R.: Characterization of Extreme Wet-Bulb Temperature Events in Southern Pakistan, Geophysical Research Letters, 46, 10 659–10 668, https://doi.org/https://doi.org/10.1029/2019GL084711, 2019.

Moron, V., Robertson, A. W., and Qian, J.-H.: Local versus regional-scale characteristics of monsoon onset and post-onset rainfall over Indonesia, Climate Dynamics, 34, 281–299, https://doi.org/10.1007/s00382-009-0547-2, 2010.

O'Gorman, P. A. and Schneider, T.: Recovery of Atmospheric Flow Statistics in a General Circulation Model without Nonlinear Eddy-Eddy Interactions, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL031779, 2007.

Pandey, K., Monteiro, J. M., and Natarajan, V.: An Integrated Geometric and Topological Approach for the Identification and Visual Analysis of Rossby Wave Packets, Monthly Weather Review, 148, 3139–3155, https://doi.org/10.1175/MWR-D-20-0014.1, publisher: American Meteorological Society, 2020.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Petoukhov, V., Eliseev, A. V., Klein, R., and Oesterle, H.: On statistics of the free-troposphere synoptic component: an evaluation of skewnesses and mixed third-order moments contribution to the synoptic-scale dynamics and fluxes of heat and humidity, Tellus A: Dynamic Meteorology and Oceanography, https://doi.org/10.1111/j.1600-0870.2007.00276.x, 2008.

Petoukhov, V., Rahmstorf, S., Petri, S., and Schellnhuber, H. J.: Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes, Proc Natl Acad Sci U S A, 110, 5336–5341, https://doi.org/10.1073/pnas.1222000110, 2013.

Petoukhov, V., Petri, S., Rahmstorf, S., Coumou, D., Kornhuber, K., and Schellnhuber, H. J.: Role of quasiresonant planetary wave dynamics in recent boreal spring-to-autumn extreme events, Proceedings of the National Academy of Sciences, 113, 6862–6867, https://doi.org/10.1073/pnas.1606300113, _eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.1606300113, 2016.

Prodhomme, C., Materia, S., Ardilouze, C., White, R. H., Batté, L., Guemas, V., Fragkoulidis, G., and García-Serrano, J.: Seasonal Prediction of European Summer Heatwaves, Clim Dyn, 58, 2149–2166, https://doi.org/10.1007/s00382-021-05828-3, 2022.

Rao, V. B., Rao, K. K., Mahendranath, B., Lakshmi Kumar, T. V., and Govardhan, D.: Large-scale connection to deadly Indian heatwaves, Quarterly Journal of the Royal Meteorological Society, 147, 1419–1430, https://doi.org/https://doi.org/10.1002/qj.3985, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3985, 2021.

Ratnam, J. V., Behera, S. K., Ratna, S. B., Rajeevan, M., and Yamagata, T.: Anatomy of Indian heatwaves, Scientific Reports, 6, 24 395, https://doi.org/10.1038/srep24395, 2016.

Rohini, P., Rajeevan, M., and Srivastava, A. K.: On the Variability and Increasing Trends of Heat Waves over India, Scientific Reports, 6, 26 153, https://doi.org/10.1038/srep26153, 2016.

Ruff, T. W. and Neelin, J. D.: Long tails in regional surface temperature probability distributions with implications for extremes under global warming, Geophysical Research Letters, 39, https://doi.org/https://doi.org/10.1029/2011GL050610, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011GL050610, 2012.

Sardeshmukh, P. D. and Sura, P.: Reconciling Non-Gaussian Climate Statistics with Linear Dynamics, Journal of Climate, 22, 1193–1207, https://doi.org/10.1175/2008JCLI2358.1, publisher: American Meteorological Society, 2009.

Schneider, T. and Walker, C. C.: Self-Organization of Atmospheric Macroturbulence into Critical States of Weak Nonlinear Eddy–Eddy Interactions, J. Atmos. Sci., 63, 1569–1586, https://doi.org/10.1175/JAS3699.1, 2006.

Schneider, T., Bischoff, T., and Plotka, H.: Physics of Changes in Synoptic Midlatitude Temperature Variability, J. Climate, 28, 2312–2331, https://doi.org/10.1175/JCLI-D-14-00632.1, 2014.

Seabold, S. and Perktold, J.: statsmodels: Econometric and statistical modeling with python, in: 9th Python in Science Conference, 2010.

Shannon, C. E.: A Mathematical Theory of Communication, Bell System Technical Journal, 27, 379–423, https://doi.org/https://doi.org/10.1002/j.1538-7305.1948.tb01338.x, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1948.tb01338.x, 1948.

Shepherd, T. G.: Atmospheric Circulation as a Source of Uncertainty in Climate Change Projections, Nature Geoscience, 7, 703–708, https://doi.org/10.1038/ngeo2253, 2014.

Shepherd, T. G.: Climate science: the dynamics of temperature extremes, Nature, 522, 425–427, https://doi.org/10.1038/nature14540, 2015.

Stiglic, G. and al, e.: Comprehensive Decision Tree Models in Bioinformatics, PLOS ONE, https://doi.org/10.1371/journal.pone.0033812, 2012.

Svirsky, A., Herbert, C., and Frishman, A.: Statistics of Inhomogeneous Turbulence in Large-Scale Quasigeostrophic Dynamics, Physical Review E, 108, 065 102, https://doi.org/10.1103/PhysRevE.108.065102, 2023.

Tamarin-Brodsky, T., Hodges, K., Hoskins, B. J., and Shepherd, T. G.: A Dynamical Perspective on Atmospheric Temperature Variability and Its Response to Climate Change, Journal of Climate, 32, 1707–1724, https://doi.org/10.1175/JCLI-D-18-0462.1, 2019.

Wei, W., Yan, Z., and Jones, P. D.: A decision-tree approach to seasonal prediction of extreme precipitation in eastern China, International Journal of Climatology, 40, 255–272, https://doi.org/https://doi.org/10.1002/joc.6207, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.6207, 2020.

White, R., Kornhuber, K., Martius, O., and Wirth, V.: From Atmospheric Waves to Heatwaves: A Waveguide Perspective for Understanding and Predicting Concurrent, Persistent and Extreme Extratropical Weather, Bulletin of the American Meteorological Society, 103, 1–35, https://doi.org/10.1175/BAMS-D-21-0170.1, 2021.

Xu, W., Ning, L., and Luo, Y.: Wind Speed Forecast Based on Post-Processing of Numerical Weather Predictions Using a Gradient Boosting Decision Tree Algorithm, Atmosphere, 11, https://doi.org/10.3390/atmos11070738, 2020.

Zhang, X., Jiang, H., Jin, J., Xu, X., and Zhang, Q.: Analysis of acid rain patterns in northeastern China using a decision tree method, Atmospheric Environment, 46, 590–596, https://doi.org/https://doi.org/10.1016/j.atmosenv.2011.03.004, 2012.