

The QBO, the annual cycle, and their interactions: Isolating periodic modes with Koopman analysis

CLAIRE VALVA,^a EDWIN P. GERBER^a

^a *Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, New York*

ABSTRACT: The Quasi-Biennial Oscillation (QBO) is the dominant mode of variability in the equatorial stratosphere. It is characterized by alternating descending easterly and westerly jets over a period of approximately 28 months. It has long been known that the QBO interactions with the annual cycle, e.g., through variation in tropical upwelling, leading to variations in the descent rate of the jets and, resultingly, the QBO period. Understanding these interactions, however, has been hindered by the fact that conventional measures of the QBO convolve these interactions. Koopman formalism, derived from dynamical systems, allows one to decompose spatio-temporal datasets (or nonlinear systems) into spatial modes that evolve coherently with distinct frequencies. We use a data-driven approximation of the Koopman operator on zonal-mean zonal-wind to find modes that correspond to the annual cycle, the QBO, and the nonlinear interactions between the two. From these modes, we establish a data-driven index for a “pure” QBO that is independent of the annual cycle and investigate how the annual cycle modulates the QBO. We begin with what is already known, quantifying the Holton-Tan effect, a nonlinear interaction between the QBO and the annual cycle of the polar stratospheric vortex. We then use the pure QBO to do something new, quantifying how the annual cycle changes the descent rate of the QBO, revealing annual variations with amplitudes comparable to the 30m day^{-1} mean descent rate. We compare these results to the annual variation in tropical upwelling and interpret then with a simple model.

1. Introduction

The Quasi-Biennial Oscillation (QBO) is the leading mode of variability of the equatorial stratosphere. It is characterized by downward propagating easterly and westerly wind regimes with a period of approximately 28 months (Baldwin et al. 2001). While a tropical phenomenon, the QBO is known to affect other regions of the atmosphere such as extratropical surface variability (Garfinkel and Hartmann 2010; Anstey and Shepherd 2014). The most well known teleconnection is the Holton-Tan effect: a warming of the boreal polar stratosphere during easterly QBO (Holton and Tan 1982).

While the oscillations of the QBO are perhaps the most regular mode in the climate system that are not directly linked to the diurnal or seasonal cycles, the period of the QBO ranges from 24 to 34 months. A key mechanism for the range of periods is the interaction between the QBO and the annual cycle (Hampson and Haynes 2004; Krismer et al. 2013; Rajendran et al. 2016). As the mean QBO period is not an integer multiple of the annual cycle, any interactions between the two would occur irregularly. (The same could be said for interactions between the QBO and other regular oscillations in the climate system, e.g., the El Niño Southern Oscillation or the 11-year solar cycle.) A consequence of this irregular interaction means that it is hard to quantify the effects of the annual cycle on the QBO and what the QBO would look like during a given period in the absence of the annual cycle. For example, commonly used indices of the QBO — including the value

of zonal-mean zonal-wind at one (or more) pressure levels or representations in EOF space, e.g., Wallace et al. (1993) — will necessarily include nonlinear modulations of the QBO phase by the annual cycle. While the use of deseasonalized data to study the QBO can remove mean annual effects, the resulting index will still retain nonlinear interactions between the annual cycle and the QBO without a more complex data analysis technique to account for them.

In this study, we propose the use of Koopman methods to help solve this problem. We use a Koopman decomposition to categorize and separate nonlinear interactions of the QBO and the annual cycle, which allows for improved quantitative understanding of the QBO in the context of its irregular interaction with the annual cycle. Koopman operator theory translates between finite-dimensional nonlinear dynamical systems and linear (albeit infinite-dimensional) systems. As such, data-driven approximations of Koopman operators can provide decompositions of the climate system without assuming linearity. These decompositions identify a collection of quasi-periodic modes and corresponding frequencies from which one analyze interactions between two modes of known frequencies.

The separation between the known frequencies of the annual cycle and the QBO allow Koopman methods to isolate the phenomena and analyze the interaction between these separately. An immediate result of this property is that we can create an objective QBO index (and corresponding resulting “pure” QBO Koopman mode) that is independent of the annual cycle.

To motivate the necessary mathematics, we entice readers with the potential of this method. Figure 1 compares

Corresponding author: Claire Valva, clairev@nyu.edu

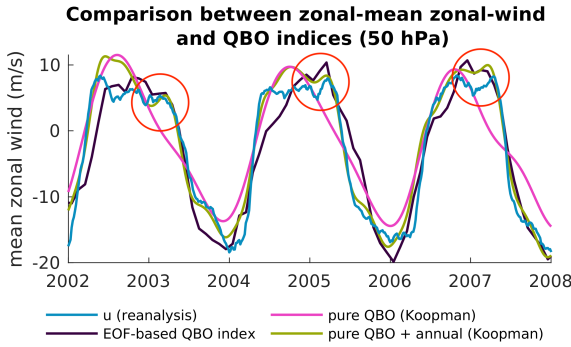


FIG. 1. Comparison between raw ERA5 reanalysis zonal-mean zonal-wind (blue), an EOF based QBO index (dark purple), a pure QBO mode (Koopman derived, pink), and a combination QBO and annual cycle mode (Koopman derived, green) at 50 hPa. Reanalysis and the Koopman derived modes are mean zonal-mean zonal-wind between -10° and 10° N. The Koopman derived modes have been normalized to have the same mean as the reanalysis winds. Similarly, the EOF index has been rescaled to have the same standard deviation as the Koopman derived mode (pure QBO with annual interactions) and the mean of the reanalysis for comparative purposes. A “shoulder” in the zonal-wind when the westerly QBO phase occurs during early boreal winter is indicated by red circles.

indices from two Koopman derived QBO modes to raw ERA5 reanalysis zonal-wind (Hersbach et al. 2020) and an empirical orthogonal function (EOF) QBO index derived from Singapore radiosonde monthly mean winds (from Newman and NASA/GSFC (2023), computed as in Wallace et al. (1993)), all at 50 hPa. Each index tracks the oscillation of the zonal-mean zonal-wind at 50 hPa fairly similarly, switching from easterly and westerly jets around the same time. The pure QBO Koopman mode more smoothly tracks this evolution without impact from the annual cycle (or higher frequency variability). The EOF based index, as well as our “combination” Koopman mode (which is explicitly designed to include nonlinear interactions between the QBO and annual cycle) retain interannual variability. In particular one observes a shifting “shoulder” (highlighted with red circles) when the westerly phase of the QBO falls early in the calendar year. This feature is a result of the interaction between the QBO and the annual cycle, which will be more more illuminatingly discussed in Section 3b. The conventional EOF based QBO index was computed from deseasonalized data, but clearly retains seasonal effects, more resembling the Koopman mode built to include them.

The remainder of the manuscript is organized as follows. In Section 2 we give a brief overview of Koopman theory and our data-driven algorithm, followed by a description of how to create and interpret aggregate Koopman modes in Section 3, where the term *aggregate* refers to the inclusion of harmonic frequencies which gives a more complete representation. We use these aggregate Koopman modes to

analyze the Holton-Tan effect and variation of QBO descent rates with the annual cycle in Section 4. Finally, in Section 5 we discuss other possibilities for use of this method, including application to other phenomena and model assessment.

2. Koopman methods: Theory and application

We begin with a brief explanation of Koopman operator theory and our data-driven Koopman approximation. For a complete overview of the computational method, as well as precise mathematical statements of Koopman operators, we refer the reader to the appendix of Froyland et al. (2021) for a detailed description of the algorithm and to Das et al. (2021) for the development of the algorithm and relevant spectral convergence results.

a. Koopman formalism

Consider a dynamical system with states denoted $\{x_t\}_{t \in \mathbb{R}}$ which is associated with a continuous map φ^s that takes the state at time t and pushes it forward to time $s+t$, i.e., $\varphi^s(x_t) = x_{s+t}$. As a parallel, consider a numerical model. We may use a numerical model to approximate φ^s , integrating the dynamical system $\frac{d}{dt}x = f(x)$ to determine x_{s+t} based on the initial condition x_t .

Associated with dynamical systems are scalar functions g called *observables* that take the state space as their domain. For example, a possible g for the state space of the climate system would be the function that computes or “observes” the mean temperature of the atmosphere. Then $g(x_t)$ would be the mean temperature at time t .

The *Koopman operator* K^s describes the evolution of a dynamical system $\{x_t\}_{t \in \mathbb{R}}$ by the action on observables of the system:

$$(K^s g)(x_t) = g \circ \varphi^s(x_t) = g(x_{s+t}) \quad (1)$$

In our example, $K^s g$ maps the observation of mean atmosphere temperature forward by time s . By inspection, one can see that the operator K^s is linear when acting on observables (or functions) even when the dynamics φ are nonlinear. For a wide class of systems of interest — i.e., those with an invariant probability measure (see Das et al. (2021) for a precise statement) — there is an alternative formulation of K^s by a skew-adjoint linear operator V which is called the *Koopman generator* that is analogous to a derivative.

$$Vg = \lim_{t \rightarrow 0} \frac{K^t g - g}{t} \quad (2a)$$

$$K^t g = e^{tV} g \quad (2b)$$

Here, eq. (2a) defines the action of V onto the function g , and eq. (2b) shows how V generates the Koopman operator K^t .

This formulation allows a spectral decomposition of V to serve as coherent feature extraction for dynamical systems through the eigenproblem:

$$e^{tV} \zeta_k = e^{it\omega_k} \zeta_k, \quad (3)$$

where the eigenfunctions ζ_k evolve predictably according to the eigenfrequency ω_k . As such, the eigenfunctions ζ_k are quasi-periodic modes that capture the evolution operator. If the evolution of the full system (or a particular feature of interest) is well described by a reasonably small set of eigenfunctions ζ_k , then there is an efficient representation of the dynamics that is inherently predictable.

The Koopman formalism also allows for eigenfrequency-eigenfunction generation given already known eigenfunctions. Suppose that you have eigenfrequencies α and β which correspond to the eigenfunctions a and b . Then there will also be an eigenpair with eigenfrequency $\alpha + \beta$ and eigenfunction ab .

$$e^{Vt} ab = e^{i\alpha t} a e^{i\beta t} b = e^{i(\alpha+\beta)t} ab$$

We discuss the relevant consequences of this construction property in Section 3a, but one can think of this generation property as accounting for harmonics (of the form 2α and α^2) and nonlinear interactions between components of a dynamical system with different frequencies (here, α and β).

b. Data-driven Koopman approximation

We can find decompositions of the climate system without assuming linearity by implementing this Koopman formalism. Given data, we want to find the Koopman eigenfunctions ζ_k and associated frequencies ω_k to build a decomposition of Koopman modes. This method can be thought of as semi-analogous to decompositions provided by principal component analysis or Fourier analysis, but with an additional allowance for nonlinear effects. Previous uses of this method have included analysis of the El Niño Southern Oscillation (ENSO, Froyland et al. (2021)) and identification of the Madden-Julien Oscillation (MJO, Lintner et al. (2023)).

From given input data \mathcal{D} of size $N_t \times N_d$ (time by space), we approximate the Koopman generator V with N_k eigenfrequency-eigenfunction pairs (ω_k, ζ_k) of V . As V is a skew-adjoint operator, eigenvalues should be purely imaginary and come in complex conjugate pairs, and we index our eigendecomposition such that $\text{im}(\omega_k) = -\text{im}(\omega_{-k})$ and $\zeta_k = \bar{\zeta}_{-k}$. Each eigenfunction ζ_k will have mean zero. A perfect numerical approximation would have $\text{re}(\omega_k) = 0$.

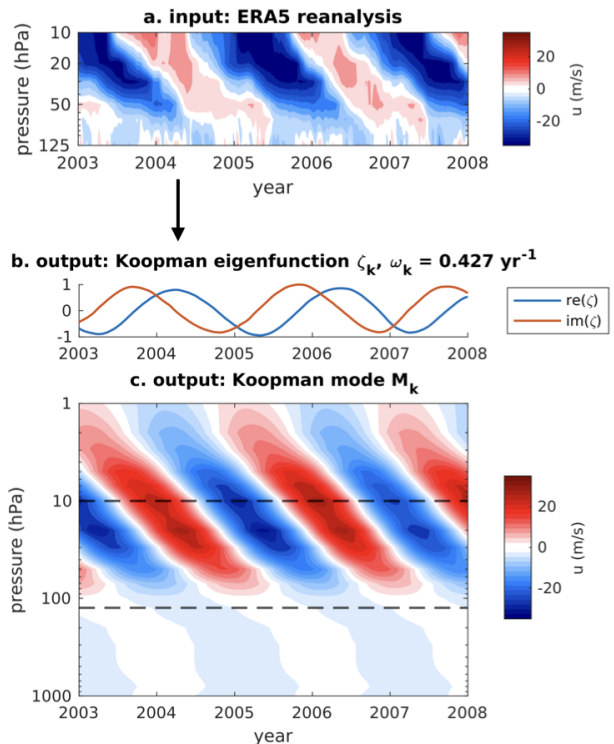


FIG. 2. Visual overview of Koopman decomposition method that shows the input data (a) mean of ERA5 zonal-mean zonal-wind between -10° and 10° N, (b) one of the N_k scalar Koopman eigenfunctions, here with frequency $\omega = 0.42737 \text{ yr}^{-1}$ and (c) zonal-mean, zonal-wind between -10° and 10° N of the associated mode M_k which is computed by projecting reanalysis data onto the eigenfunction. Note that the y-axis changes between top and bottom of the figures. The dashed lines in (c) correspond to the range of the data in (a).

Thus, we order the pairs (ω_k, ζ_k) such that $|\text{re}(\omega_k)|$ increase with k . This ordering serves as a proxy for numerical approximation error and has been used with empirical success (Das et al. 2021).

We can alternatively interpret the magnitude of $\text{re}(\omega_k)$ to inversely correspond to how periodic a given Koopman mode is: the smaller $|\text{re}(\omega_k)|$, the more periodic. The Koopman eigenpair corresponding to an extremely regular cycle (such as the annual or diurnal cycles) will be easier to resolve numerically due to its regularity. As such, $|\text{re}(\omega_k)| \ll 1$, as opposed to a cycle that has a varying period or other irregularity, where $|\text{re}(\omega_k)|$ is large. Additionally, there will always be an eigenfrequency $\omega_0 = 0$, where $\zeta_0 \equiv 1$ is constant to which corresponds to the mean state of a given system. We then have that $N_k = 2L + 1$, for L complex conjugate pairs of frequencies ω_k and the a single mean state corresponding to ω_0 .

From these eigenpairs, we compute *Koopman modes* M_k from linear projections of wanted target data onto the eigenfunction ζ_k . As we obtain M_k via projection, M_k is not limited to the size or variables of the initial input

data \mathcal{D} and can instead be extended to a larger physical area, different observed variable, or extended time period (similarly as EOFs). An important difference in output between the Koopman analysis and EOFs, however, is that the modes M_k are not orthogonal to each other; if they were, then the Koopman decomposition would assume linearity. If we were to compute M_k via a projection on the initial input data for all k we can recover the complete data, i.e., $\sum_k^{N_k} M_k = \mathcal{D}$ when the number of modes $N_k = N_t$. However, we choose to compute only a limited number of modes $N_k \ll N_t$ for better computational and analysis efficiency: for very large k , we expect the modes M_k to be less well-resolved, less approximately periodic, and to be of generally small magnitude. The goal is to find a small subset of well-resolved modes that capture the variability of interest.

In our study, we project the Koopman eigenfunctions onto data on all pressure levels while the input data for the eigenproblem is only levels between the upper troposphere to the middle stratosphere. In figure 2, we give a visual overview of the inputs and outputs of this method. We show a sample interval of the input data (zonal-mean zonal-wind from 10 to 125 hPa) and one of the output Koopman eigenpairs. The Koopman eigenfunction ζ_k is associated with eigenfrequency $0.42737 \text{ yr}^{-1} \approx 28$ month period, and the Koopman mode M_k is given from the projection zonal-mean zonal-wind from 1 to 1000 hPa onto ζ_k . This Koopman mode has many aspects of the QBO, as to be expected from its period, including the downward propagating easterly and westerly jets. As discussed in section 3, a more complete representation of the QBO requires inclusion of additional harmonics.

One limitation of the method worth reiterating is the assumption that the dynamical system is stationary in the sense that it has an invariant probability measure. Clearly, this assumption is unrealistic for post-industrial climate data, but a minor violation does not appear to greatly affect results. If the data were too non-stationary, the decomposition can result in nonconstant eigenfunctions with nonzero mean, essentially trends in the record. No trends appear in our analysis of the historical record, but it could be an issue with longer climate change projection simulations.

Finally, it has been found that approximations of this kind can be improved with a technique called *delay embeddings*, e.g., Giannakis (2019). Here, replace the input data \mathcal{D} with a larger $\hat{\mathcal{D}}$ of size $(N_t - (N_e - 1)) \times (N_d \cdot N_e)$. Now each row of $\hat{\mathcal{D}}$ includes observations from the $N_e - 1$ previous time points, i.e.,

$$\hat{\mathcal{D}}_t = (\mathcal{D}_t, \mathcal{D}_{t-1}, \dots, \mathcal{D}_{t-(N_e-1)}).$$

This acts to insert memory of previous time steps into the data. The computation of Koopman eigenvalues and eigenfunctions proceeds as previously described, where we substitute $\hat{\mathcal{D}}$ for \mathcal{D} . A rule-of-thumb often used to choose

the number of delays N_e to be on the order of the timescale of the phenomena of interest. For example, in this work studying the QBO, we use an embedding corresponding to approximately 23 months.

3. Construction of a pure QBO and related modes via Koopman decomposition

We perform and interpret a Koopman decomposition of zonal-mean zonal-wind of the stratosphere. We use these results to create the QBO index (seen in figure 1), as well as aggregate Koopman modes that correspond to the construction of a pure QBO mode, an annual cycle mode, and a QBO-annual cycle interaction mode.

We use five-day averages of zonal-mean zonal-wind ERA5 reanalysis (Hersbach et al. 2020) from 1979 to 2020 at seven pressure levels between 125 and 10 hPa. We choose a five-day average to decrease the high frequency variability and the computational expense. The data is density and area-weighted and then delay embedded with $N_e = 140$ samples. Results were similar varied embedding lengths. We compute 50 Koopman eigenfunctions (51 including the eigenpair corresponding to Koopman mode that is the time mean). The Koopman modes M_k are projected onto from ERA5 data at 21 pressure levels between 1000 and 1 hPa.

a. Interpretation of eigenfrequencies

Each Koopman mode has an associated frequency. For the well-resolved modes we want to be able to interpret these frequencies in terms of known phenomena. We emphasize that this method does not allow for targeting frequencies *a priori*, i.e., there is no point in which we specify that we would like to find an eigenfrequency of once every 28 months or of the annual cycle. The input data can, of course, affect results: the delay embedding was chosen to emphasize interannual variability. Additionally, if the data is too coarsely sampled, we can miss oscillations. In this case, the diurnal cycle cannot be observed given that our input data has a sampling frequency of 5^{-1} days. The pressure levels were chosen to prioritize variability at the levels of the QBO. Using input data that extended to 1 hPa decreased resolution of the results, given the prevalence of the semiannual oscillation.

In figure 3, we plot the all eigenvalues with positive frequencies (the eigenvalues of negative frequencies are simply the complex conjugates) in the complex plane, shaded by the log of the (density and area weighted) variance of each mode. As the eigenfunctions are ordered by their effective periodicity, one is not guaranteed that the modes capture a substantive fraction of the variance. The Koopman eigendecomposition is constructive, meaning that we expect harmonics of significant modes (modes with integer

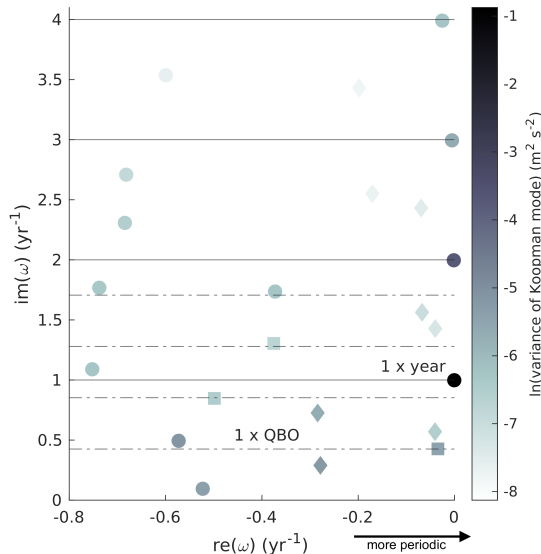


FIG. 3. Numerically computed Koopman eigenvalues plotted on the complex plane. The y-axis value is the frequency of each eigenmode, while the x-axis can be used as a proxy for the quality or the periodicity of the numerical mode (as a perfect computation would have $\text{re}(\omega)$ be 0). Color denotes the log of the (density and area weighted) variance of each Koopman mode, while dotted and solid horizontal lines denote frequencies corresponding to integer multiples of the QBO ($\sim 0.4286 \text{ yr}^{-1}$, denoted with square markers) and the seasonal cycle respectively. Diamond shaped markers denote QBO-annual cycle interaction frequencies, i.e., those of the form $n \pm 0.4286 m$ within a 10-day tolerance.

multiples of the base frequency) as well as interactions between modes (modes that are integer linear combinations of frequencies).

The leading eigenvalues — in both the sense of smallest real part (rightmost) and highest variance (darkest) — correspond to the annual cycle or harmonics thereof with integer valued frequencies. We also see frequencies that correspond to harmonics of the QBO, $0.4286 m$ for $m = 1, 2, \dots$, which correspond to a period of 28 months. Finally, we have identified modes that we call “QBO-annual cycle interaction” modes with frequencies that are approximately equal to $n + 0.4286 m$ for integers m and n within a selected tolerance, chosen here to have a period that is within 10 days (twice the sampling frequency) of the desired period.

Some eigenfrequencies remain unidentified. Of these, we conjecture that some could be sorted into the three previously discussed categories — particularly higher harmonics of the QBO or the interaction terms — but they do not fall into the desired frequency tolerance. Neglecting these modes minimally affects the results in later sections, as the unidentified modes are relatively small when compared to zonal-mean zonal-wind target data: the unidentified mode with the highest variance ($0.05 \text{ m}^2 \text{ s}^{-2}$) has a root-mean-square amplitude (over all levels) of 0.17 m s^{-1}

and a maximum value (of any level) of 2.6 m s^{-1} . For reference, the zonal-mean zonal-wind target data has a variance of $327.0 \text{ m}^2 \text{ s}^{-2}$ and root-mean-square amplitude of 12.9 m s^{-1} .

b. Aggregate Koopman modes

From the eigenfrequency groupings discussed in Section 3a, we can create corresponding aggregate Koopman modes which we will refer to as an annual cycle mode (made up of integer frequencies), a pure QBO mode (frequencies of the form $0.4286 m$ for integers m), and an annual cycle-QBO interaction mode (frequencies of the form $n \pm 0.4286 m$). The annual cycle aggregate mode is the sum the projection of 5 eigenfunctions (and their conjugates), the pure QBO aggregate mode is made up of the projections of 3 eigenfunctions, and the annual cycle-QBO interaction aggregate mode is made up of 8. Each of these aggregate modes are the sum of the Koopman modes M_k that correspond to the desired ω_k . For example, the additive mode for the annual cycle can be written as $M_{\text{annual}} = \sum_{k, \omega_k \approx n} M_k$.

The aggregate Koopman modes isolate particular phenomena: the pure QBO mode represents the evolution of the QBO in the absence of any interaction or interference by the seasonal cycle. The mean of the pure QBO mode in any month of the year, say December (not shown), is zero up to sampling uncertainty. Any impact of the annual cycle on the QBO will appear in the annual cycle-QBO interaction aggregate mode.

Figure 4 illustrates these aggregate modes and compares these to a slice of raw zonal-mean zonal-wind reanalysis (panel a). The sum of three aggregate modes (panel f) nearly captures the full reanalysis zonal-wind, but for high frequency variability and shows how the flow is dominated by the QBO and annual cycle. At first glance, the pure QBO mode (panel c) seems some to lack common characteristics of the QBO including the varying descent of the zero-line. The sum of the QBO and its interactions with the annual cycle (Panel e) captures the variability of its descent rate. When the annual cycle is added (panel f), the full evolution of the flow is recovered.

c. Koopman eigenfunction as QBO index

In addition to constructing aggregate Koopman modes, we can also use the Koopman eigenfunctions to create indices. Namely, we use this to create an “objective QBO index” from the single Koopman eigenfunction with frequency $\omega = 0.4274$ that is independent of annual effects. As we have a scalar eigenfunction, we can create a continuous index of the QBO defined by its phase angle. We plot this in figure 5 and use this index to divide the QBO into 4 phases: a pure westerly phase (denoted as 1 : W), a descending easterly phase (2 : $E \rightarrow W$), a pure easterly phase (3 : E), and a descending westerly phase (4 : $W \rightarrow E$). The

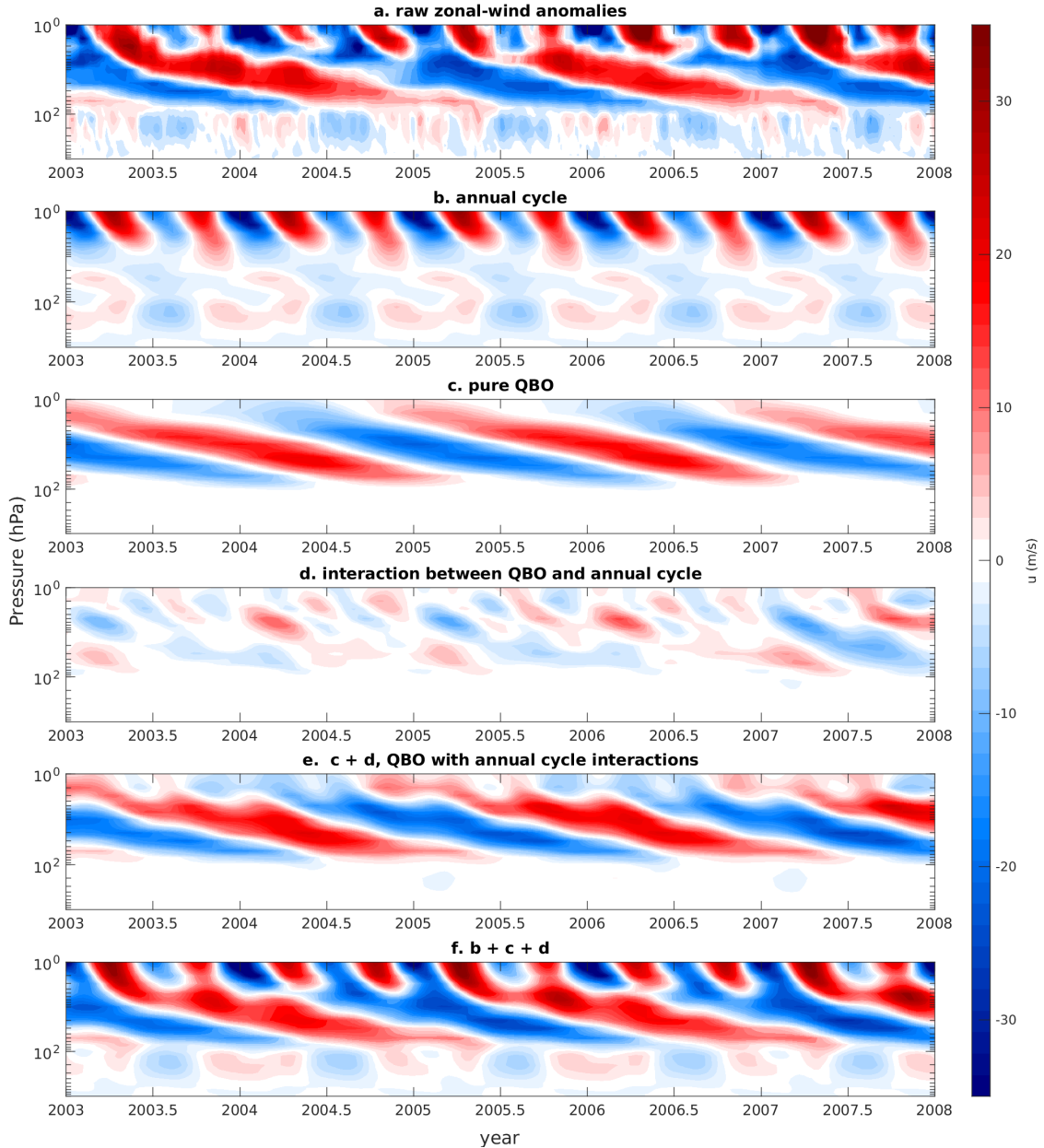


FIG. 4. Comparison between the tropical mean zonal-wind and the aggregate modes representing the annual, pure QBO, and interaction between the QBO and the annual cycle. We plot means of zonal-mean zonal-wind between -10° and 10° N of each component between 2003 and 2008. The time mean winds are removed for clarity. a: the raw wind anomalies. b: seasonal modes, i.e., approximately integer frequencies. c: pure QBO modes, frequencies that are multiples of 0.4286. d: seasonal-QBO interaction modes, frequencies of the form $n \pm 0.4286m$, e: total QBO, sum of panels (c) and (d), f: total seasonal and QBO effects, sum of panel (b) and (e). The final panel (f) reconstructs the total wind (a), except for short time variability.

number of phases is arbitrary, but four is chosen to balance a clarity of the QBO state and the need for sufficient data for statistics.

The timespan plotted includes the QBO disruption that took place in 2016 that interrupted the descent of the westerly phase of the QBO (Barton and McCormack 2017). It

materialized in this QBO index as the diversion and loop from the usual oscillation that returned to its usual path after the conclusion of the disruption.

Figure 5 examines whether our pure QBO mode captures known characteristics of the QBO. While the QBO index is defined from a single Koopman eigenfunction, we plot

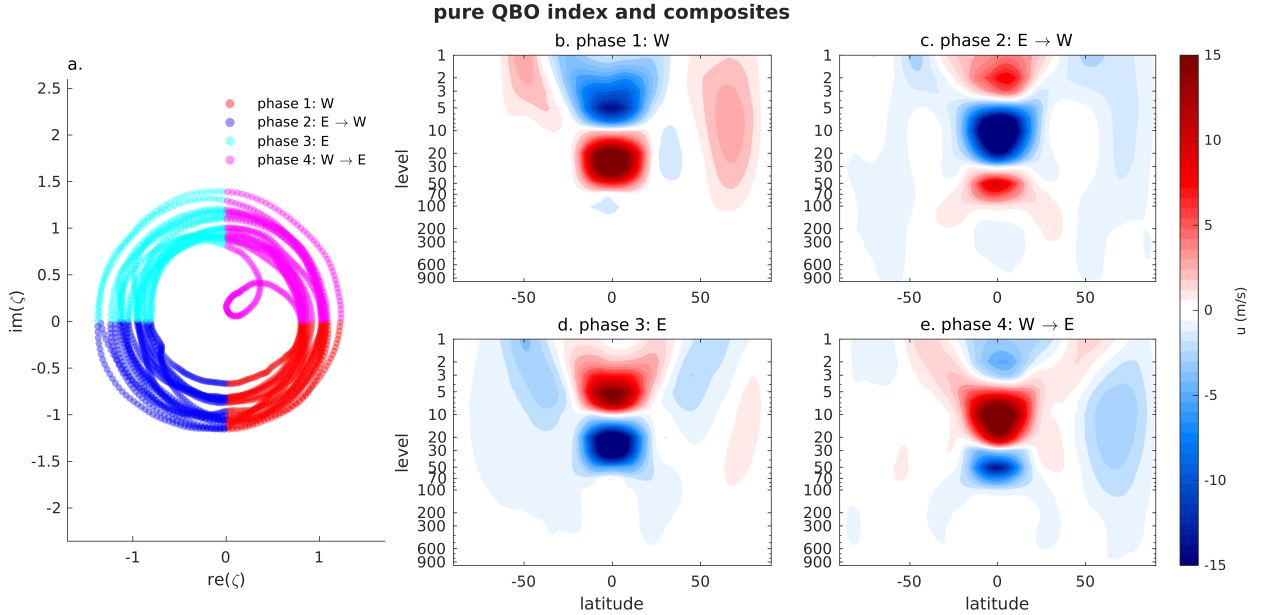


Fig. 5. We compare the scalar (complex) Koopman eigenfunction corresponding to eigenfrequency $\omega = 0.427$ to composites of the Koopman modes corresponding to splitting the QBO into four phases based on the quadrant of the Koopman eigenfunction. Colors of the scatter plot (panel a) of the Koopman eigenfunction correspond to phases one through four: a pure westerly phase (1 : W), a descending easterly phase (2 : $E \rightarrow W$), a pure easterly phase (3 : E), and a descending westerly phase (4 : $W \rightarrow E$). The “loop” in the pink scatter points (phase 4, descending westerly) corresponds to the disruption of the QBO that took place in 2016.

composites of the pure QBO aggregate mode in the figure to include variability associated with harmonics. The composites capture the arch, or horseshoe-shaped, wind anomalies that extend downward from the QBO region in the Subtropics, conjectured to be associated with the QBO mean meridional circulation by Garfinkel and Hartmann (2011). These can be identified in the phase 2, 3, and 4 composites as a downward extension of wind anomalies at beginning at approximately 50 hPa. Interestingly, the pure QBO mode captures modest strengthening (weakening) of the polar vortex during westerly QBO (easterly QBO), a shadow of the Holton-Tan effect. We investigate this particular aspect more in Section 4a.

4. Understanding QBO-annual cycle interaction with Koopman modes

We can now use the aggregate Koopman modes to investigate the nonlinear interaction between the QBO and the annual cycle. We first replicate known results regarding the Holton-Tan effect (Section 4a) as further proof of concept for the use of these modes. We then use them to quantify the effect of seasonality on the QBO descent rate via comparison to the pure QBO mode (Section 4b).

a. The Holton-Tan effect

The Holton-Tan effect, first noted in Holton and Tan (1982), refers to the QBO influence on the strength of

the wintertime stratospheric polar vortex in the Northern Hemisphere. Specifically, the polar vortex is weaker and warmer during easterly QBO (EQBO) and stronger and colder during westerly QBO (WQBO), as westerlies in the subtropics favor more equatorward propagation of planetary Rossby waves, thereby shielding the polar vortex (Lu et al. 2020).

In figure 6, we show composites of EQBO subtracted from WQBO in boreal winter (DJF) using the full reanalysis record of winds, versus the winds captured by our aggregate Koopman modes. Panel (a) is a composite formed from ERA5 reanalysis, where climatology EQBO (WQBO) is defined as mean reanalysis winds at 50 hPa over the season being less than -3 m s^{-1} (greater than 3 m s^{-1}) and the two are then subtracted. The remaining three panels are formed similarly, where we compute climatologies for a given Koopman mode over the EQBO and WQBO defined from reanalysis. We do not include the 2016 QBO disruption in these composite.

Use of the Koopman modes allows us to separate the Holton-Tan effect into a pure QBO component that is independent of the time of year (figure 6b is identical for any other season; not shown) and a nonlinear interaction component that is only present during DJF (figure 6d). Figure 6b and the QBO index composites (figure 5b, e) indicate a mean strengthening (weakening) of the polar vortex during WQBO (EQBO).

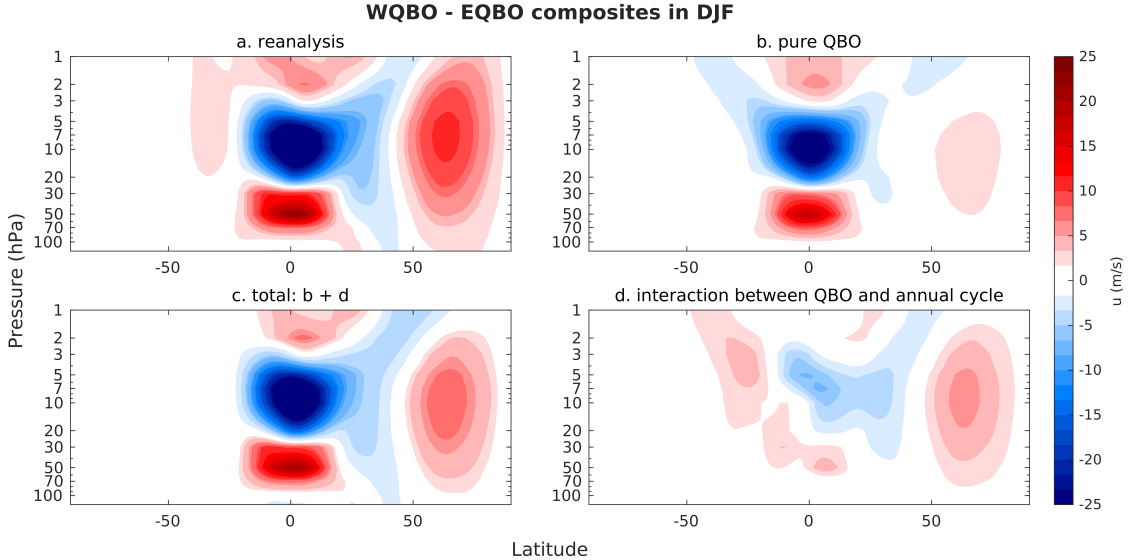


FIG. 6. Composites of zonal-mean zonal-wind during boreal winter (DJF) of WQBO-EQBO as measured by mean winds over the season at 50 hPa. Panel (a) is reanalysis — the typical composite shown for the Holton-Tan effect. Panel (b) and panel (d) are compositions of the pure QBO and QBO-annual cycle interaction modes. Panel (c) is the sum of (b) and (d). In the case of a perfect reconstruction from the Koopman modes, we would expect (c) to be the same as (a); Differences between these panels are on the order of 3 ms^{-1} .

The annual cycle-interaction aggregate mode captures the strengthening of the polar vortex in DJF. This strengthening is reversed in other seasons, where the aggregate mode is weakly negative. This cancels out the mean strengthening of the polar vortex in the pure QBO composites in other seasons.

One could create an equivalent figure where the determination of whether a given winter belongs to either EQBO or WQBO categories is done by using either the phase of the scalar QBO index or value of the zonal-wind from the pure QBO aggregate mode. While either of these would be a more “objective” way to define the phases of the QBO, we kept with the standard convention for consistency. The other approaches produce similar results; see Appendix b. Additionally, all the composites closely match similar composites in the literature (Baldwin et al. 2001; Lu et al. 2020).

The biggest differences between the Koopman reconstruction of the composite from reanalysis are present in the upper stratosphere, which indicates that we are excluding some stratospheric processes with our Koopman modes. We conjecture that this discrepancy may be partially due to the 11-year solar cycle, which has been suggested to influence the polar vortex and has a lower frequency than any of our resolved Koopman modes. In addition, this analysis does not account for ENSO interactions by design, which has also been noted to affect the polar vortex (Anstey and Shepherd 2014).

b. The annual cycle in QBO descent rate

Studies of the influence of the annual cycle on QBO descent rates (specifically the effects of tropical upwelling) have been done, notably using descent rate models, such as in Rajendran et al. (2018). Previous work was somewhat limited in that conventional methods cannot access a version of the QBO that is not influenced by the annual cycle. As such, an investigation into how the descent rate of the QBO is affected by the seasonal cycle must either reform the problem as seasonal preference of phase onsets, as done in (Dunkerton 1990), or look at trends in the data and compare to theoretical models.

The onset months of the QBO westerly or easterly phases have been observed to have a seasonal preference and the period of the QBO can vary significantly from the 28-month mean (Dunkerton 1990; Hampson and Haynes 2004). The perturbation of the QBO by the annual cycle is thought to be the primary reason for the variation in period between cycles, although ENSO and the solar cycle may also contribute.

The Koopman formulation of the QBO allow us to see the seasonal preference of onset months as a result of non-linear interactions of the annual cycle. Figure 7 contains histograms of zero-line crossings at 50 hPa in the pure QBO aggregate Koopman mode and the aggregate mode with additional annual cycle interactions. In the histogram of the aggregate mode that includes both the QBO and annual cycle interactions, we see preferential seasonal transitions from easterly to westerly winds (or vice versa) with a significant peak in spring (May) and secondary peak in fall

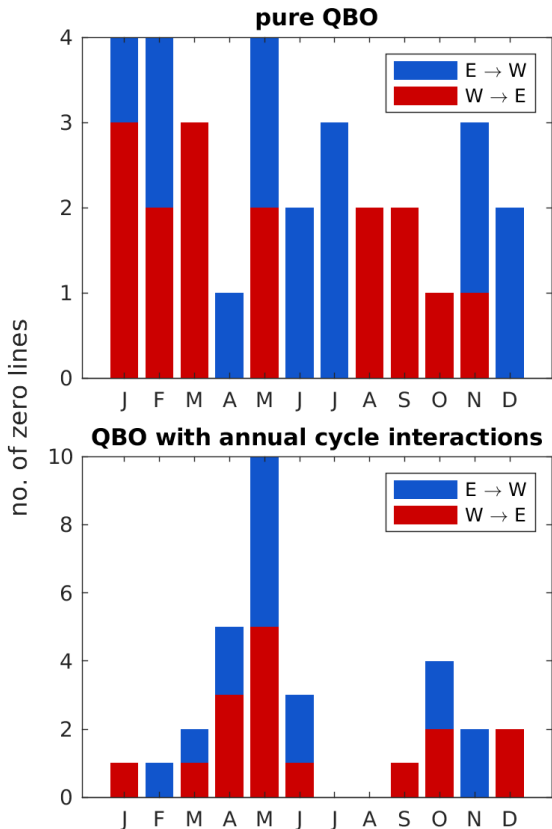


FIG. 7. Histograms of zero-line crossings at 50 hPa in the pure QBO mode (top) and the mode with nonlinear annual cycle interactions (bottom). Colors denote either an easterly to westerly transition (red) or a westerly to easterly transition (blue). While the histogram counting the pure QBO mode exhibits a uniform pattern zero-crossings, the bottom plot shows preferential transitions in boreal spring and no crossings in July or August.

(October) — similar to those analyzed in previous literature (Hampson and Haynes 2004). A transition in July or August has never been observed. In contrast, the transition months of our pure QBO mode show little preference in the timing of the onset. We hypothesize that these transition months are drawn from a uniform distribution, i.e., the pure QBO mode has no preferential onset.

We can test this with a χ^2 test, choosing the null hypothesis to be that the QBO transition month is drawn from a uniform distribution; the expectation for the number of QBO transitions in each month is 31/12, given the 31 transitions in the analyzed record, not counting the QBO disruption in 2016. For the pure QBO, we cannot reject the null hypothesis, with a $\chi^2 = 5$ and a p -value of 0.93. On the other hand, for the aggregate mode that contains the QBO and annual cycle interactions, we find that $\chi^2 = 32.8$

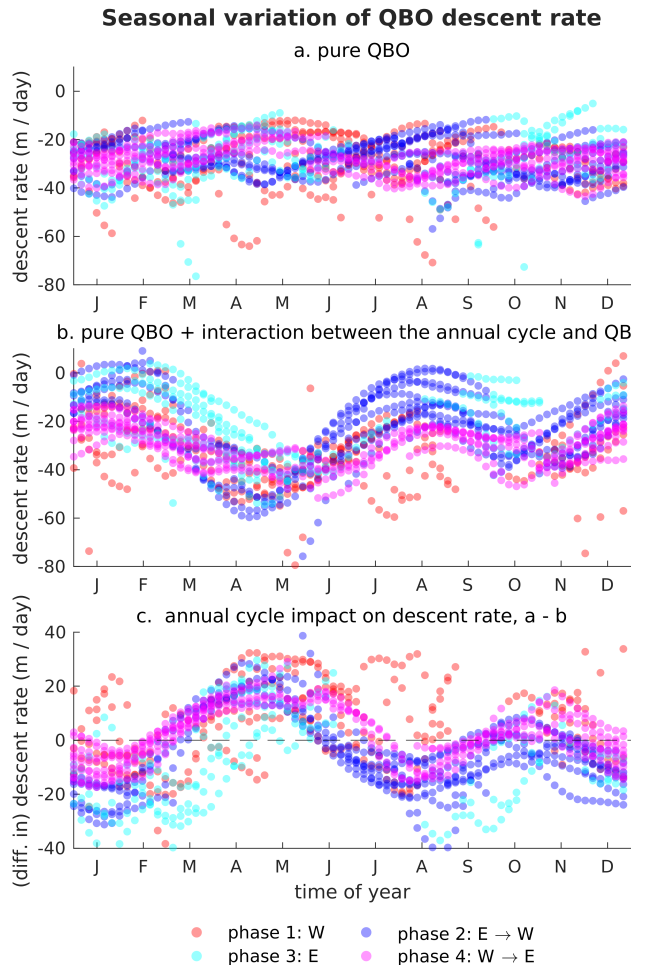


FIG. 8. Comparison of QBO zero-line descent rate by time of year between (a) the pure QBO mode and (b) the QBO with the addition of non-linear QBO-annual cycle interactions. Each scatter point is colored by the phase (as defined by the Koopman eigenfunction metric in figure 5) that the QBO was in when the rate was measured. Panel c is the difference between the two above panels, and represents the change in descent rate of the QBO due to the nonlinear interaction between the annual cycle and the QBO. A negative value means that the annual cycle-QBO interactions during this time of year *slow* the descent, rate while a positive difference means that the QBO will descend *faster*.

and $p = 5.5 \times 10^{-4}$, meaning that it would be quite unlikely for there to be a uniform preference of QBO transitions. This statistical result is not perfect, as there are a limited number of QBO transitions in the data and each transition is not truly independent of any other. However, the differences in these two aggregate modes strongly indicate that the distribution of QBO transitions is controlled by nonlinear interactions with the annual cycle.

To probe the annual variation in QBO further, we compute the descent rate of the QBO by tracking the zero-wind line, that is the descent of the $E \rightarrow W$ or $W \rightarrow E$ phase transition. One could equivalently track the westerly or

easterly maxima, but we have found that this introduces more noise (not shown) as the height of the maximum is more sensitive to small perturbations. For a given wind profile, we calculate the location of the mean zero-wind line between -10° and 10° N, estimating the exact pressure by interpolating in log-pressure (as recommended in Hersbach et al. (2020)). We restrict the zero-line to be at or between pressures 10 to 125 hPa, to limit issues of defining a descent rate on when there are multiple zero lines in a single snapshot (as is often the case when a new phase begins to descend from the upper stratosphere while the other is ending). We compute the descent rate as the derivative computed by a centered difference and then smoothed by a 30-day rolling window. We exclude data from periods where the zero-line “jumps up,” i.e., there is a change between descending easterly to westerly winds (or vice versa).

In figure 8, we plot the zero-crossing descent rates as a function of the time of year. Panel (a) shows the descent rate computed from the pure QBO Koopman mode, where color differentiates the phases of the QBO. The phases correspond in color and definition with those shown in figure 5. Consistent with the Koopman definition of the pure QBO mode, there is no dependence of the pure QBO descent rate on the time of year.

Panel b is the same as panel a, but computed from the aggregate Koopman mode that includes the nonlinear QBO-annual cycle interactions. The third panel (c) is the *difference* in descent rates between the two above panels and quantifies how much the amount annual cycle-QBO interactions either slow or speed up the descent of the QBO. We observe a marked semiannual modulation of descent rate in the QBO mode with additional annual effects, both in the raw descent rate (panel b) and more clearly in the difference from the pure QBO descent rate (panel c). This semiannual pattern is noisier for phases 1 and 3 (pure westerly or easterly). We expect phases 1 and 3 to be more variable than 2 and 4. Phases 2 and 4 are descending westerly and easterly respectively, where the zero-wind line cleanly descends between 10 and 125 hPa. Phases 1 and 3 contain “jumps” of the zero-line between the bottom of the QBO and the top, leading to much of the spread.

Koopman analysis extracts a semiannual variation in descent rates of order 30m day^{-1} , which is comparable in magnitude to the mean descent rate. There is pronounced annual variability, with a stronger peak in May than October. The peaks in May and October are consistent with the enhanced number of QBO transitions during these months. We can similarly explain the few QBO transitions in January and August, where descent of the zero-line essentially stalls completely.

Three potential mechanisms for the perturbation of QBO descent rates have been identified by Hampson and Haynes (2004) (1) the annual cycle in tropical upwelling caused by

the Brewer-Dobson circulation; (2) seasonal variations in the wave forcing that drive the QBO; and (3) seeding from above by the semiannual oscillation (SAO) (Hampson and Haynes 2004).

We first investigate to what extent seasonal variation in tropical upwelling is associated with variation in QBO descent rates. Figure 9 shows the relationship between variation in descent rates and tropical upwelling. We compute the linear relationship between the annual cycle of the transformed Eulerian mean of tropical upwelling (w^*) as a function of pressure to the mean annual cycle in descent rates. We consider the residual vertical velocity w^* between 3 and 100 hPa, as computed by Serva et al. (2024). The black markers correspond to annual mean descent rates regardless of QBO phase, while other colors denote subdivision into one of the QBO phases.

The left panel shows the regression coefficient and its 95% confidence interval at each pressure level. It indicates how much a 1m day^{-1} change in w^* impacts the descent rate of the QBO. Naively, if changes in the QBO descent rate were the result of a simple perturbation of the advection, we would expect a slope of -1 . The right panel shows the squared correlation coefficient or R^2 of the corresponding linear fits on the left, which suggest what fraction of the variability in mean QBO descent rates can be explained by variation in w^* .

The slope of the linear fits is negligible at 100 hPa, but quickly transitions to nearly around -1 from 50 to 15 hPa. This suggests that the amplitude of annual variations in w^* are consistent with the amplitude of annual variations in descent rate, agreeing with the naive expectation that one can qualitatively explain the descent rate changes simply by changes in advection.

The change in the correlation coefficient with height is similar to that of the slope. At 100 hPa the correlation coefficient is essentially zero, suggesting that the annual variation in w^* at the base of the tropical stratosphere has nothing to do with the annual variation of descent rate. As you go higher in the stratosphere, R^2 steadily increases with a maximum between 20 and 5 hPa, where 80% of the variation in descent rate are consistent with the variation in w^* . However, the value of the slope peaks lower than R^2 , due to the fact that the magnitude of variability in w^* increases with height in the stratosphere while the general shape of the seasonal variation remains the same (see the leftmost panel of figure 12). Another explanation for this vertical structure is that w^* exhibits a more annual cycle in the lower stratosphere (and hence unrelated to the dominant semiannual cycle that is observed) but becomes more decidedly semiannual in the upper stratosphere. As such, R^2 increases with height, where the variability of the QBO more strongly matches that of w^* .

In both the combined data and in each QBO phase, w^* potentially accounts for a significant amount of the seasonal

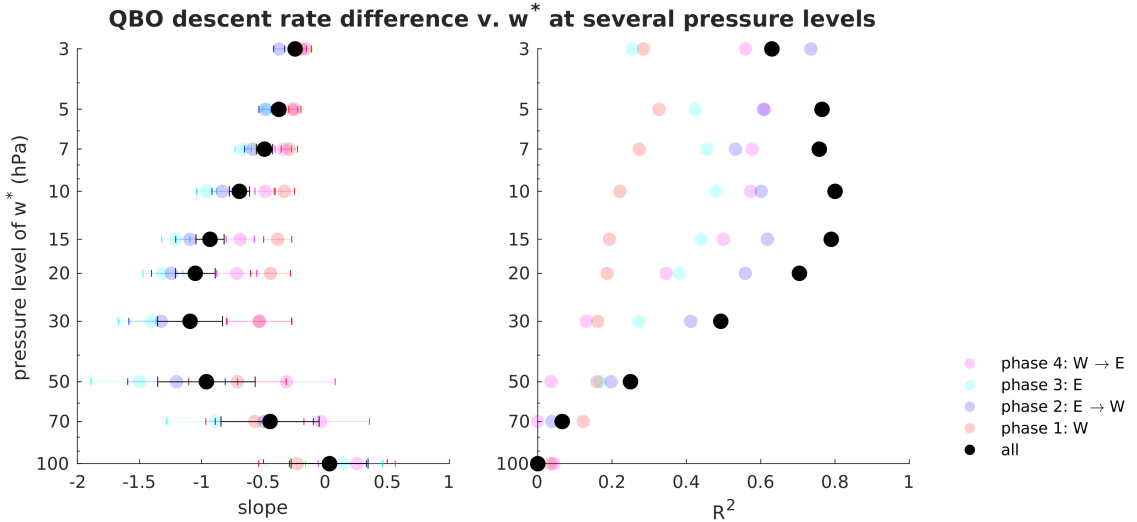


Fig. 9. Left: The slope at different pressure levels of the linear fit between descent rate (mean of time of year) vs w^* (mean between -25° and 25° N). Error bars denote the 95% confidence intervals of the slope. Black scatter points correspond to the time of year mean in descent rate over all phases of the QBO, while colored points correspond to each phase. Slope magnitudes peak at 30 hPa. Right: Corresponding R^2 (variability explained) values for each of the linear fits in the left figure. The R^2 values peak at 15 hPa for most phases.

variability in QBO descent rates. For the easterly and descending easterly phases (phases 2 and 3), the simple linear relationship appears to work well, though this is diminished for the westerly phases (1 and 4). As different phases of the QBO correspond to different average heights of the zero-crossings and to different dynamics, these changes are not necessarily surprising. We should be cautious not to over interpret these differences between QBO phases, as the subdivision of data decreases the signal-to-noise ratio.

The analysis in figure 9 also allows us to make predictions about how future changes in tropical upwelling may affect the speed of the QBO descent. For example, consider an increase of w^* by $0.1 \text{ mm s}^{-1} \approx 9 \text{ m day}^{-1}$, consistent with estimates from quadrupling of CO_2 experiments for several coupled chemistry models (Chiodo et al. 2018). Based on this linear relationship, an increase in upwelling of corresponds to an approximately equal slow down of the QBO descent, translating to an average of 9 m day^{-1} decrease in descent rates of the QBO and an increase of its period to 40 months. This prediction assumes there are no changes in other factors that could drive of the QBO descent, such as the wave forcing. Some models do predict an increase in QBO period, but others a decrease (Richter et al. 2022).

Correlation analysis cannot establish that w^* causes variations in QBO descent rates, as there could be a third factor that drives both the changes in the QBO and w^* . Seasonal variations in the wave forcing could influence both w^* and the QBO directly. Similarly, the SAO affects w^* in the upper stratosphere. The wave forcings of the QBO are difficult to diagnose; in particular, gravity waves are not well

resolved in the reanalysis. We therefore turn to a simple 1D model to untangle w^* from the wave forcing and the SAO.

c. Interpreting the results with a 1D model of the QBO

We use a model that is a hybrid of Holton and Lindzen (1972) and Plumb (1977). We solve the following equation for the zonal-wind (u):

$$\partial_t u + w^* \partial_z u - K \partial_z^2 u = G(z, t) + S(u, z) \quad (4)$$

where t and z are the time and vertical coordinate, w^* is the (prescribed) vertical advection, K is a constant diffusivity, $G(z, t)$ drives the SAO, and $S(u, z)$ is a monochromatic gravity wave forcing.

Following Holton and Lindzen (1972), the SAO is prescribed as

$$G(z, t) = 2G \frac{(z - 28 \text{ km})}{1000 \text{ m}} \omega_{\text{SAO}} \sin(2\pi \omega_{\text{SAO}} t) \text{ for } z > 28 \text{ km} \quad (5)$$

where ω_{SAO} is the frequency of the SAO.

The gravity wave forcing $S(u, z)$ assumes two identical gravity waves of opposite phase, as in Plumb (1977) and is detailed in Appendix a. The gravity wave forcing was tuned to give the QBO a 28 month period.

Each experiment described below reflects a 96 year run, where statistics are excluded for the first 12 years to avoid the influence of initiation. We investigate how prescribing different values of the model parameters affect annual descent rates of the QBO. We define the descent rate of the

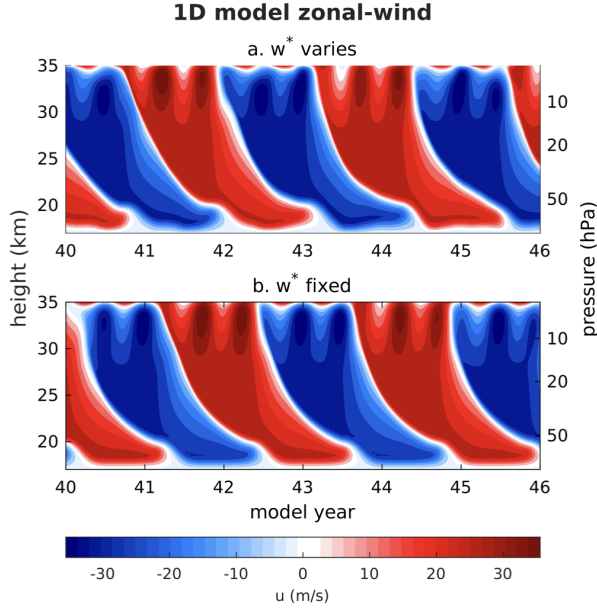


FIG. 10. Zonal-wind from the 1D model for model years 40 to 46, both with G (the parameter that sets the SAO strength) equal to 14. Panel a shows zonal-wind from the model where w^* periodically varies annually. Panel b shows zonal-wind from the model where w^* is fixed as the climatology.

QBO in the simple model the same way it is defined in the analysis of Koopman modes, by the change in height of the zero-wind line. Portions of the two integrations, with and without annually varying w^* are shown in figure 10.

We first ask if seeding of the SAO perturbs the descent rate of the QBO by augmenting the magnitude of the function G . We tested several values for the magnitude of $G(z, t)$, trying the values $G = 0, 3.28, 14,$ and 24 ms^{-1} . The value $G = 3.28 \text{ ms}^{-1}$ was estimated from reanalysis to match the SAO strength, while 14 ms^{-1} is the original magnitude from Holton and Lindzen (1972). We found that the descent rate of the QBO was completely insensitive to the value of G , even when other parameters (such as w^* were varied), indicating that the descent rate only depends on the wave driving and the upwelling in this model.

We next investigate how much variation the tropical upwelling — which is prescribed independently of the wave forcing — can explain the annual variation of the QBO in the model. We compare model simulations without an annual cycle, where w^* is prescribed to be the climatological mean at each level (analogous to the pure QBO Koopman mode) to simulations with an annual cycle, where prescribed w^* varies with the annual cycle, repeating annually as illustrated in figure 10. We quantify the effect of seasonal upwelling variation on the annual descent rate as the difference between the model run where w^* varies annually and where the model is the climatological mean.

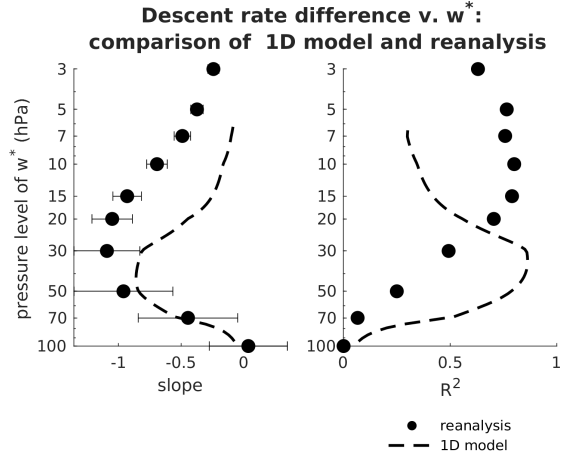


FIG. 11. The same as in figure 9 but for the 1D QBO model (dashed lines) and reanalysis QBO (black circles; same as in figure 9, average over all phases). As the 1D model has a top of $35 \text{ km} \approx 5.8 \text{ hPa}$, As such, w^* above 5 hPa was excluded from the analysis.

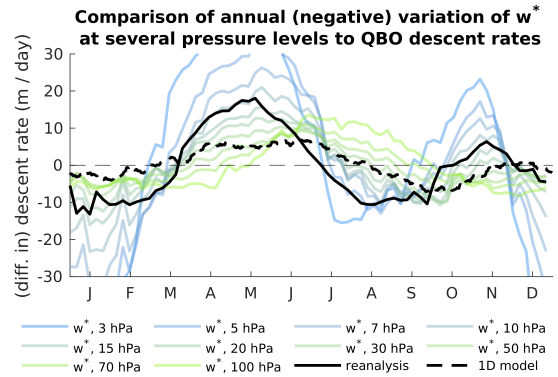


FIG. 12. The seasonal cycle in descent rates for reanalysis (the difference between the pure QBO mode and the mode with added annual-cycle interaction) and the 1D model of the QBO (the difference of model runs with a seasonally varying winds and without) with the *negative* deviation from climatology of w^* at several levels. The w^* anomalies are inverted because of their inverse correlation to the descent rate.

Figure 11 shows the linear relationship between w^* and annual variation of the 1D model descent rates as a function of pressure. For comparative purposes, we replot the data from the analysis of the Koopman modes in figure 9. The slope and correlation coefficient both peak at similar heights (between 60 and 30 hPa). Unlike the observed QBO, the slope that relates the 1D model descent rate to w^* never surpasses -1 , but the R^2 in the lower stratosphere is higher than in reanalysis. Part of this difference is explained by the extreme regularity of the 1D model compared to the climate system as well as the longer time period (84 years) in which we compute statistics. Both the slope

and correlation coefficient peak lower in the atmosphere for the model than in reanalysis.

To interpret these differences, we compare the 1D model descent rate with that of reanalysis more directly. Figure 12 compares the annual variation in descent rate from the 1D model (black, dotted line) and reanalysis (black, solid line) to the value of w^* at several pressure levels (various colors). The variation in descent rate in reanalysis is about twice that of the 1D model. We also see that while both the reanalysis and the 1D model have a semiannual pattern in annual descent rate, the timing of the peaks are different.

The fact that the descent rate of the 1D model is more highly correlated with upwelling at lower levels (and less correlated at upper levels) when compared to reanalysis can be understood from the strength of semiannual variation. Higher in the stratosphere, upwelling has a stronger semiannual variation, presumably in part due to the SAO, while lower down (where the 1D model is more sensitive to w^*), the annual cycle is more prevalent, causing different maxima in w^* . This appears to indicate that the true QBO is more affected by upwelling higher in the stratosphere than in the simple model. This could be due to the fact that stratospheric upwelling itself is more important in our atmosphere, or that the winds higher in the stratosphere are better correlated to another driver of annual changes in the QBO (such as the semiannual oscillation or variations in the wave forcing).

The results from this simple model are consistent with the hypothesis that much of the variability in QBO descent rates are controlled by mean deviations in tropical upwelling, even if the importance of winds at different height varies from the true system. However, upwelling alone cannot reproduce the full seasonal cycle. Annual variation in the wave driving must play an important role in the variation of the descent rate.

5. Conclusions and future directions

It has been long accepted interactions between the QBO and the annual cycle leads to variation in the downward propagation of the easterly and westerly phases of the oscillation, including their descent rates. Quantitative investigation of these interactions was difficult as conventional data-analysis methods are unable to easily untangle interactions between the QBO and the annual cycle. Koopman formalism is well-suited to this problem, as it is able to separate a periodic element of a system from its nonlinear interactions with other periodic modes of the system.

We have identified an objective, data-driven index of the QBO that is independent of the annual cycle and performed a decomposition of reanalysis data that separates the affects of the QBO, the annual cycle, and the nonlinear interactions between the two phenomena. This allowed an investigation of the strength of interactions between the QBO and the annual cycle, quantifying to what extent the Holton-Tan

effect is driven by nonlinear interactions between the QBO and the annual cycle. We quantified how the annual cycle perturbs the QBO descent rate, comparing it with annual variations in tropical upwelling, with the aid of a classic 1-D model of the QBO.

The annual variation of the descent rate is large, of amplitude of 30 m day^{-1} , equivalent to the mean descent rate. As a result, the QBO essentially stalls out from January to February and in July to August (figure 8), explaining the dearth of phase transitions in these months (figure 7). Variation in tropical upwelling can account for a large fraction of this variation, particularly in boreal winter. Changes in wave forcing, however, must also play a role, particularly in boreal summer.

Koopman formalism could have further desirable uses for QBO investigation. Applying this method to data that is not zonally averaged would allow for further insight (and possible discovery) of teleconnections between the QBO and higher latitudes that are present on smaller scales than phenomena like the polar vortex. Similar analysis could also be done for other variables of interest, such as temperature or outgoing longwave radiation to probe possible connections between the MJO and QBO (Yoo and Son 2016).

Alternatively, we also anticipate that the Koopman decomposition could allow for another avenue of QBO comparison in large climate model evaluations. While advances in gravity wave parameterizations have improved the QBO in comprehensive climate modes, replicating the Holton-Tan effect with observed strength remains a challenge (Garfinkel et al., 2018; Zhang et al., 2019). Using the decomposition as we did for the Holton-Tan effect (4a) could help to understand where climate models get the Holton-Tan effect wrong: is the QBO itself not well enough resolved, or is it the nonlinear interactions between the QBO and the annual cycle that are under accounted for.

Finally, the modes found from the Koopman decomposition could also be used for investigation and prediction of other aspects of the climate. Already, these methods have been used successfully to find indices with improved predictability in both the both ENSO and MJO when compared to classical methods (Wang et al. 2020; Lintner et al. 2023).

Acknowledgments. We acknowledge support from the US National Science Foundation through award OAC-2004572 and the US National Science Foundation Graduate Research Fellowship under Grant DGE-1839302.

Data availability statement. The code used to compute the Koopman decomposition can be found on github at: <https://github.com/claivalva/koopmanQBO>. Zonal-wind data from ERA5 can be found at <https://doi.org/10.24381/cds.bd0915c6> and the w^* from Serva et al. (2024) can be found at <https://doi.org/10.5281/zenodo.7081436>. The 1D QBO model is implemented in *PyTorch* and can be found at <https://github.com/DataWaveProject/qbo1d>.

APPENDIX

a. 1D model of the QBO

For our experiments in Section 4b comparing reanalysis and a 1D model, we use a simple model of the QBO that is a hybrid of Holton and Lindzen (1972) and Plumb (1977), as implemented in *PyTorch* by Connelly and Shamir (2022). We solve the following equation for the zonal-wind (u):

$$\partial_t u + w^* \partial_z u - K \partial_z^2 u = G(z, t) + S(u, z) \quad (\text{A1})$$

where t and z are the time and vertical coordinate, w^* is the (prescribed) vertical advection, K is the diffusivity ($0.3 \text{ m}^2 \text{ s}^{-1}$), $G(z, t)$ is the semiannual oscillation, and $S(u, z)$ is the wave forcing.

The wave forcing term parametrizes the momentum deposition at critical levels as $S(u, z) = \frac{-1}{\rho} \partial_z F(u, z)$ where $\rho(z)$ is the density profile and the wave flux $F(u, z)$ is written as:

$$F(u, z) = \sum_i A_i \exp\left(-\int_{z_L}^z \frac{\alpha(z)/N}{k_i(u - c_i)^2}\right) \quad (\text{A2})$$

where A_i and c_i are the wave amplitudes and wave speeds, and $\alpha(z)$ is the wave dissipation due to infrared cooling (see Holton and Lindzen (1972)). Here, we choose a two-wave set up and choose $A_i = \pm 1.18e - 3$ and $c_i = \pm 29$ in order to obtain a 28-month QBO period.

The semiannual oscillation is prescribed as:

$$G(z, t) = 2G \frac{(z - 28 \text{ km})}{1000 \text{ m}} \omega_{\text{SAO}} \sin(2\pi \omega_{\text{SAO}} t) \text{ for } z > 28 \text{ km} \quad (\text{A3})$$

where ω_{SAO} is the frequency of the SAO. We found that results of descent rates of the models were insensitive to the amplitude (G) of the SAO term.

We choose the model to have a domain top of $z_T = 35 \text{ km}$ and bottom of $z_L = 17 \text{ km}$ with a grid spacing of 250 m. The model is run for 96 years including a 12-year spin-up period with a time step of 12 hours. The upwelling

w^* is prescribed to either be the mean at each level or to be periodic over the year at each level. Log pressure interpolation was used to match grid spacing.

b. Alternative Holton-Tan figure

The figure A1 is a variant of figure 6, where we determine if each DJF is a WQBO or EQBO season based on the zonal-winds of the pure QBO Koopman aggregate mode. This is opposed to defining based on the unfiltered zonal-winds of reanalysis as in the figure in the main text. The two figures are similar, but not identical, as the pure index flags subtly different winters into the composite.

References

- Anstey, J. A., and T. G. Shepherd, 2014: High-latitude influence of the quasi-biennial oscillation. *Quarterly Journal of the Royal Meteorological Society*, **140** (678), 1–21, <https://doi.org/10.1002/qj.2132>.
- Baldwin, M. P., and Coauthors, 2001: The quasi-biennial oscillation. *Reviews of Geophysics*, **39** (2), 179–229, <https://doi.org/10.1029/1999RG000073>.
- Barton, C. A., and J. P. McCormack, 2017: Origin of the 2016 QBO Disruption and Its Relationship to Extreme El Niño Events. *Geophysical Research Letters*, **44** (21), 11,150–11,157, <https://doi.org/10.1002/2017GL075576>.
- Chiodo, G., L. M. Polvani, D. R. Marsh, A. Stenke, W. Ball, E. Rozanov, S. Muthers, and K. Tsigaridis, 2018: The Response of the Ozone Layer to Quadrupled CO₂ Concentrations. <https://doi.org/10.1175/JCLI-D-17-0492.1>.
- Connelly, D. S., and O. Shamir, 2022: A Pytorch 1D QBO model. URL <https://github.com/DataWaveProject/qbo1d>.
- Das, S., D. Giannakis, and J. Slawinska, 2021: Reproducing kernel Hilbert space compactification of unitary evolution groups. *Applied and Computational Harmonic Analysis*, **54**, 75–136, <https://doi.org/10.1016/j.acha.2021.02.004>.
- Dunkerton, T. J., 1990: Annual variation of deseasonalized mean flow acceleration in the equatorial lower stratosphere. *Meteorological Society of Japan, Journal*, **68**.
- Froyland, G., D. Giannakis, B. R. Lintner, M. Pike, and J. Slawinska, 2021: Spectral analysis of climate dynamics with operator-theoretic approaches. *Nature Communications*, **12** (1), 1–21, <https://doi.org/10.1038/s41467-021-26357-x>.
- Garfinkel, C. I., and D. L. Hartmann, 2010: Influence of the quasi-biennial oscillation on the North Pacific and El Niño teleconnections. *Journal of Geophysical Research: Atmospheres*, **115** (D20), <https://doi.org/10.1029/2010JD014181>.
- Garfinkel, C. I., and D. L. Hartmann, 2011: The Influence of the Quasi-Biennial Oscillation on the Troposphere in Winter in a Hierarchy of Models. Part I: Simplified Dry GCMs. *Journal of the Atmospheric Sciences*, **68** (6), 1273–1289, <https://doi.org/10.1175/2011JAS3665.1>.
- Giannakis, D., 2019: Data-driven spectral decomposition and forecasting of ergodic dynamical systems. *Applied and Computational Harmonic Analysis*, **47** (2), 338–396, <https://doi.org/10.1016/j.acha.2017.09.001>.

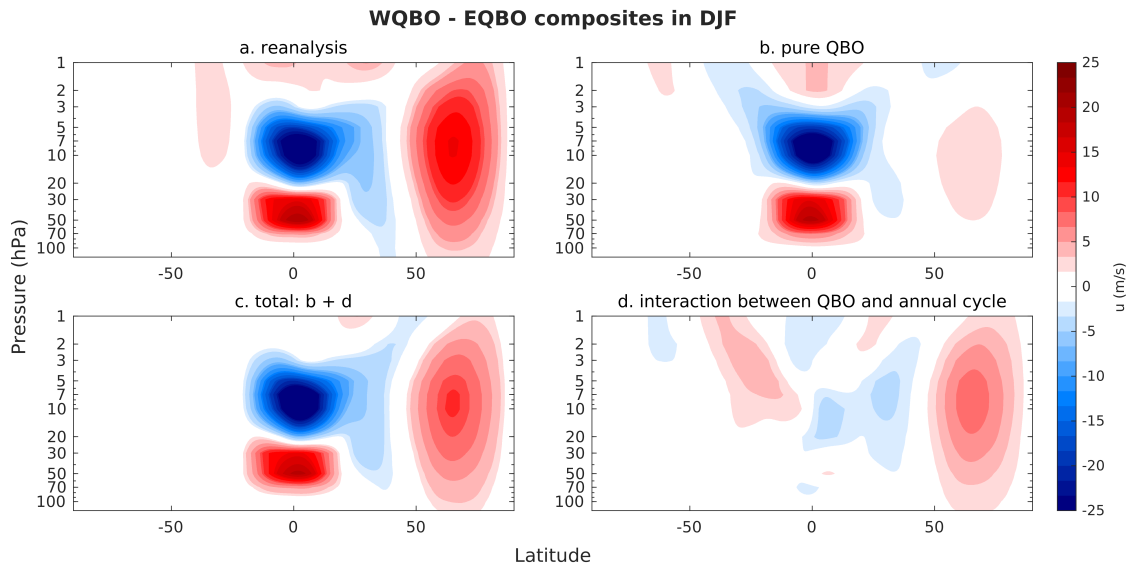


FIG. A1. Composites of zonal-mean zonal-wind during boreal winter (DJF) of WQBO-EQBO as measured by the winds of Koopman index composite mode 50 hPa. Panel (a) is reanalysis — the typical composite shown for the Holton-Tan effect. Panels (b) and panel (d) are compositions of the pure QBO and QBO-annual cycle interaction modes. Panel (c) is the sum of (b) and (d). In the case of a perfect reconstruction from the Koopman modes, we would expect (c) to be the same as (a); differences between these panels are on the order of 3 m s^{-1} .

- Hampson, J., and P. Haynes, 2004: Phase Alignment of the Tropical Stratospheric QBO in the Annual Cycle. *Journal of the Atmospheric Sciences*, **61** (21), 2627–2637, <https://doi.org/10.1175/JAS3276.1>.
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146** (730), 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Holton, J. R., and R. S. Lindzen, 1972: An Updated Theory for the Quasi-Biennial Cycle of the Tropical Stratosphere. *Journal of the Atmospheric Sciences*, **29** (6), 1076–1080, [https://doi.org/10.1175/1520-0469\(1972\)029<1076:AUTFTQ>2.0.CO;2](https://doi.org/10.1175/1520-0469(1972)029<1076:AUTFTQ>2.0.CO;2).
- Holton, J. R., and H.-C. Tan, 1982: The Quasi-Biennial Oscillation in the Northern Hemisphere Lower Stratosphere. *Journal of the Meteorological Society of Japan. Ser. II*, **60** (1), 140–148, <https://doi.org/10.2151/jmsj1965.60.1.140>.
- Krismer, T. R., M. A. Giorgetta, and M. Esch, 2013: Seasonal aspects of the quasi-biennial oscillation in the Max Planck Institute Earth System Model and ERA-40. *Journal of Advances in Modeling Earth Systems*, **5** (2), 406–421, <https://doi.org/10.1002/jame.20024>.
- Lintner, B. R., D. Giannakis, M. Pike, and J. Slawinska, 2023: Identification of the Madden-Julian Oscillation With Data-Driven Koopman Spectral Analysis. *Geophysical Research Letters*, **50** (10), e2023GL102743, <https://doi.org/10.1029/2023GL102743>.
- Lu, H., M. H. Hitchman, L. J. Gray, J. A. Anstey, and S. M. Osprey, 2020: On the role of Rossby wave breaking in the quasi-biennial modulation of the stratospheric polar vortex during boreal winter. *Quarterly Journal of the Royal Meteorological Society*, **146** (729), 1939–1959, <https://doi.org/10.1002/qj.3775>.
- Newman, P. A., and NASA/GSFC, 2023: The EOFs of the 1980-present QBO from the Singapore sonde monthly mean winds. URL https://acd-ext.gsfc.nasa.gov/Data_services/met/qbo/.
- Plumb, R. A., 1977: The Interaction of Two Internal Waves with the Mean Flow: Implications for the Theory of the Quasi-Biennial Oscillation. *Journal of the Atmospheric Sciences*, **34** (12), 1847–1858, [https://doi.org/10.1175/1520-0469\(1977\)034<1847:TLOTIW>2.0.CO;2](https://doi.org/10.1175/1520-0469(1977)034<1847:TLOTIW>2.0.CO;2).
- Rajendran, K., I. M. Moroz, S. M. Osprey, and P. L. Read, 2018: Descent Rate Models of the Synchronization of the Quasi-Biennial Oscillation by the Annual Cycle in Tropical Upwelling. *Journal of the Atmospheric Sciences*, **75** (7), 2281–2297, <https://doi.org/10.1175/JAS-D-17-0267.1>.
- Rajendran, K., I. M. Moroz, P. L. Read, and S. M. Osprey, 2016: Synchronisation of the equatorial QBO by the annual cycle in tropical upwelling in a warming climate. *Quarterly Journal of the Royal Meteorological Society*, **142** (695), 1111–1120, <https://doi.org/10.1002/qj.2714>.
- Richter, J. H., and Coauthors, 2022: Response of the Quasi-Biennial Oscillation to a warming climate in global climate models. *Quarterly Journal of the Royal Meteorological Society*, **148** (744), 1490–1518, <https://doi.org/10.1002/qj.3749>.
- Serva, F., B. Christiansen, P. Davini, J. von Hardenberg, G. van den Oord, T. J. Reerink, K. Wyser, and S. Yang, 2024: Changes in Stratospheric Dynamics Simulated by the EC-Earth Model From CMIP5 to CMIP6. *Journal of Advances in Modeling Earth Systems*, **16** (4), e2023MS003756, <https://doi.org/10.1029/2023MS003756>.
- Wallace, J. M., R. L. Panetta, and J. Estberg, 1993: Representation of the Equatorial Stratospheric Quasi-Biennial Oscillation in EOF Phase Space. *Journal of the Atmospheric Sciences*, **50** (12), 1751–1762, [https://doi.org/10.1175/1520-0469\(1993\)050<1751:ROTESQ>2.0.CO;2](https://doi.org/10.1175/1520-0469(1993)050<1751:ROTESQ>2.0.CO;2).
- Wang, X., J. Slawinska, and D. Giannakis, 2020: Extended-range statistical ENSO prediction through operator-theoretic techniques for

nonlinear dynamics. *Scientific Reports*, **10** (1), 2636, <https://doi.org/10.1038/s41598-020-59128-7>.

Yoo, C., and S.-W. Son, 2016: Modulation of the boreal winter-time Madden-Julian oscillation by the stratospheric quasi-biennial oscillation. *Geophysical Research Letters*, **43** (3), 1392–1398, <https://doi.org/10.1002/2016GL067762>.