Diagnosing the Horizontal Propagation of Rossby Wave Packets along the Midlatitude Waveguide

GABRIEL WOLF and VOLKMAR WIRTH

University of Mainz, Mainz, Germany

(Manuscript received 16 September 2016, in final form 10 April 2017)

ABSTRACT

It has been suggested that upper-tropospheric Rossby wave packets propagating along the midlatitude waveguide may play a role for triggering severe weather. This motivates the search for robust methods to detect and track Rossby wave packets and to diagnose their properties. In the framework of several observed cases, this paper compares different methods that have been proposed for these tasks, with an emphasis on horizontal propagation and on a particular formulation of a wave activity flux previously suggested by Takaya and Nakamura. The utility of this flux is compromised by the semigeostrophic nature of upper-tropospheric Rossby waves, but this problem can partly be overcome by a semigeostrophic coordinate transformation. The wave activity flux allows one to obtain information from a single snapshot about the meridional propagation, in particular propagation from or into polar and subtropical latitudes, as well as about the onset of wave breaking. This helps to clarify the dynamics of individual wave packets in cases where other, more conventional methods provide ambiguous or even misleading information. In some cases, the “true dynamics” of the Rossby wave packet turns out to be more complex than apparent from the more conventional diagnostics, and this may have important implications for the predictability of the wave packet.

1. Introduction

Midlatitude weather is dominated by a succession of synoptic-scale cyclones and anticyclones. These, in turn, are often accompanied by longer-lived and larger-scale Rossby wave packets (RWPs) in the upper troposphere (Chang 2005; Wirth and Eichhorn 2014). This suggests that the dynamics of such RWPs play an important role for the prediction of midlatitude cyclones and the associated weather, which arguably explains the recent interest in upper-tropospheric wave packets.

An example for the downscale connection between an upper-tropospheric RWP and severe weather is the heavy rain event over central Europe in August 2002 (Shapiro and Thorpe 2004), which led to devastating flooding of the Elbe River (Ulbrich et al. 2003a,b). In this case, the associated RWP formed some 10 days before the rain event far upstream over the western Pacific Ocean. Hovmöller diagrams indicate a steady eastward progression of this RWP during its lifetime (Glatt and Wirth 2014). Rossby wave breaking toward the end of the life cycle resulted in the formation of a cutoff cyclone over Europe, which led to the heavy precipitation event. The rain was poorly forecast by the operational centers even on the relatively short time scale of a few days (Grazzini and van der Grijn 2002). Assuming that large-scale and long-lived dynamical features should generally be predictable on a time scale longer than just a few days, this suggest that there may be room for improvements concerning the forecast of the severe weather event, and that the dynamics of the RWP may play an important role in this context. On the other hand, the above picture of the downscale connection and the inheritance of predictability of smaller scales from larger scales is likely to be too simplistic, because the smaller scales may also have a reverse impact on the larger scales. Such an upscale impact was, for instance, diagnosed for the abovementioned RWP in August 2002 (Enomoto et al. 2007), where a tropical
storm interacted with the wave packet thus triggering Rossby wave breaking and cutoff low formation during the final stage.

The foregoing motivates the present study, in which we aim to improve our understanding for the propagation of upper-tropospheric RWPs. An established method to diagnose the evolution of an RWP uses so-called Hovmöller diagrams, which display a meridional average of a suitable variable in a longitude–time diagram (Hovmöller 1949; Glatt et al. 2011). Hovmöller diagrams are useful, because they are easy to compute and provide a rough overview of the spatiotemporal evolution of an RWP. Obviously, through meridional averaging one essentially loses information regarding the meridional direction such as meridional propagation or the existence of multiple RWPs in the meridional direction. Another technique diagnoses RWPs on longitude–latitude maps and tracks them in time, thus providing information about both the zonal and the meridional direction (Souders et al. 2014b, a; Grazzini and Vitart 2015). Despite some obvious advantages in comparison with Hovmöller diagrams, diagnosing and tracking of RWPs is far from straightforward and may occasionally yield misleading results. In particular, diagnosing RWP objects on a longitude–latitude map requires a number of choices, like for instance choosing an algorithm to compute the envelope of a wave packet and picking a threshold.

Neither Hovmöller diagrams nor RWP tracking inherently provides information about the propagation of RWPs at any given time; rather, it can only be derived by reference to a sequence of consecutive points in time. This is in contrast to diagnostics involving some form of a flux, which by design is a vector and may be designed such as to indicate the direction of propagation. In this study we make use of wave activity and its associated flux. The concept of wave activity and wave activity flux is attractive, because it involves a conservation relation for conservative flows, in distinct contrast with wave energy (Eliassen and Palm 1961). Suitable formulations possess the so-called group-velocity property, which means that for almost-plane wave packets the wave activity flux is equal to $c_A A$, where $A$ is the wave activity and $c_A$ denotes the group velocity of the wave packet. In addition, the wave activity flux does not only provide information about the propagation of wave packets, but also about their interaction with the mean flow (Edmon et al. 1980).

A formulation that is appropriate in the current context is the wave activity flux of Takaya and Nakamura (2001). One particular feature of this formulation is its phase independence; this means that it discounts individual troughs and ridges and focuses on the dynamics of the entire wave packet instead (Danielson et al. 2006). Focus on the entire wave packet is desirable, for instance, when studying model errors as opposed to initial condition errors (Gray et al. 2014), and it would be interesting to find out whether the wave packet as a whole is more predictable than the location of individual troughs and ridges within a wave packet (Lee and Held 1993). Further to its phase independence, the wave activity flux of Takaya and Nakamura is able to diagnose both the zonal and the meridional propagation of wave packets on a background flow, which may be weakly varying in the zonal direction as well as in time. For stationary waves on a zonal background flow, the flux of Takaya and Nakamura (2001) reduces to the flux of Plumb (1985), and a further zonal average recovers the original formulation of Eliassen and Palm (1961).

It is the goal of the present paper to compare various methods for diagnosing the propagation of upper-tropospheric RWPs, with the aim to work out their strengths and weaknesses. This will be done in the framework of case studies of past episodes. In particular, we will show the added value of the wave activity flux of Takaya and Nakamura (2001), which allows one to obtain information that is not readily available from the other methods. At the same time, we will emphasize some aspects regarding RWP dynamics and their predictability, which we believe are underappreciated to date. Throughout this paper we are only concerned about the horizontal propagation of RWPs along the upper-tropospheric midlatitude waveguide.

The remainder of this paper is organized as follows. In section 2 we present the data used in our case studies as well as the methods applied in our analysis. The results of our analysis are then presented in section 3. Finally, section 4 provides a short discussion and summarizes our conclusions.

2. Data and methods

a. Data

As data source for our case studies we used ERA-Interim reanalyses (Dee et al. 2011). We retrieved both components of the horizontal wind as well as geopotential on 300hPa on a latitude–longitude grid with $1.5° \times 1.5°$ resolution.

b. Wave packets and the semigeostrophic transformation

A zonally aligned Rossby wave packet is said to be present if the zonal dependence of the meridional wind $v$ can be written as
where $\lambda$ is longitude, $C$ is the so-called carrier wave, and $A$ is the slowly varying amplitude. The amplitude function $A(\lambda)$ is nonnegative everywhere and is spatially localized. The latter means that it has a maximum in the center of the wave packet and smoothly decays to smaller values toward the leading and trailing edge of the wave packet. The carrier wave $C(\lambda)$ oscillates between positive and negative values and varies on a much shorter spatial scale than $A$. The amplitude $A$ will also be referred to as envelope in the following. The task of envelope reconstruction is tantamount as to compute $A(\lambda)$ when $\nu(\lambda)$ is given.

For real Rossby wave packets, the carrier wave $C(\lambda)$ is not a plane wave $\propto \sin(s \lambda)$ with some constant wave-number $s$. Rather, the semigeostrophic nature of Rossby waves makes troughs generally thinner than ridges, such that the distance between consecutive minima/maxima in $\nu$ is smaller than the distance between consecutive maxima/minima (Hoskins 1975). For illustration we show a snapshot of the upper-tropospheric flow in Fig. 1a, where this effect can clearly be seen. This fundamental property of Rossby waves implies that envelope reconstruction with the Hilbert transform technique (Zimin et al. 2003) produces a tendency to split a single wave packet into several fragments (Wolf and Wirth 2015). Therefore, unless stated otherwise, we perform a semigeostrophic coordinate transformation in this paper before the respective diagnostic is applied to the data. Details like the equations for the transformation and their implementation are given in Wolf and Wirth (2015). The geostrophic wind, which is part of the formula, is computed from geopotential height anomalies. The effect of the transformation is illustrated in Fig. 1b; apparently, consecutive extrema of $\nu$ are now approximately equidistant in longitude, in distinct contrast to Fig. 1a. In other words, the transformation reduces the asymmetry between troughs and ridges and, therefore, improves the envelope reconstruction using the Hilbert transform (Wolf and Wirth 2015). As will be shown in this paper, it is also beneficial for other diagnostics of wave packet detection and quantification.

c. Definition of a background flow

Some techniques require the definition of a background flow (denoted by an overbar) as well as a perturbation (denoted by a prime) also referred to as “anomaly.” In the present study this is done by applying to the time series a low-pass and bandpass filter with a Bloomfield convergence window (Doblas-Reyes and Deque 1998). The anomaly is obtained by restriction to periods of 3–30 days, while the background state is
obtained by restriction to periods longer than 30 days. Periods shorter than 3 days are considered noise and will be discarded. The low-frequency limit for consecutive positive and negative perturbations (1/30 day⁻¹) is chosen similar to the 20-day moving average in Zimin et al. (2006).

d. Methods to extract the wave packet envelope

We perform envelope reconstruction from the meridional wind \( v(\lambda, \varphi) \) using the methods of Zimin et al. (2003) and Zimin et al. (2006). Both methods extract the envelope with the help of a Hilbert transform. In Zimin et al. (2003) the Hilbert transform is carried out in the zonal direction separately for each latitude, while in Zimin et al. (2006) it is carried out along streamlines of the background flow. For both methods, we restrict the meridional wind to a certain range of zonal wavenumbers. We first use a fixed range of zonal wavenumbers, but as a second step we suggest as a possible improvement a latitude-dependent range. Details regarding the wavenumber ranges will be given below when presenting the results of the first case study.

e. Wave activity flux according to Takaya and Nakamura

The horizontal components of the wave activity flux of Takaya and Nakamura (2001) are given by

\[
F(\lambda, \varphi) = \frac{p \cos(\varphi)}{2|\nabla|} \left( \frac{\partial^2}{\partial l^2} - \frac{\Phi'}{f a \cos(\varphi)} \frac{\partial u'}{\partial l} \right) + \theta \left[ -u'v' + \frac{\Phi'}{af} \frac{\partial u'}{\partial \varphi} \right] + \nabla \left[ u'^2 + \frac{\Phi'}{af} \frac{\partial u'}{\partial \varphi} \right] + C_u M, \tag{2}
\]

where \( \mathbf{v} = (u, v) \) is the horizontal wind, \( \Phi \) is geopotential, \( f = 2\Omega \sin(\varphi) \) is the Coriolis parameter, \( \Omega \) is the angular velocity of Earth’s rotation, \( a \) is the radius of Earth, and \( \varphi \) denotes latitude. The associated wave activity

\[
M = \frac{p}{2} \left( \frac{q^2}{2|\nabla| q} + \frac{e}{|\nabla| - C_p} \right) \tag{3}
\]

is calculated through a combination of enstrophy and wave action. The quantity \( M \) is designed such as to render the wave activity phase independent for small-amplitude waves, with \( e \) and \( q \) being the eddy kinetic energy and enstrophy, respectively, exactly calculated as in Takaya and Nakamura (2001). The vector \( \mathbf{C}_u \) in Eq. (2) represents the phase speed \( C_p \) in the direction of the background flow. The quantity \( C_p \) is calculated similarly as in Takaya and Nakamura (2001) by computing a 24-h time-lag gridpoint-based correlation of the geopotential height anomaly. The resulting field of \( \mathbf{C}_u \) turns out to vary on a very small spatial scale, which is somewhat inconsistent with the underlying concept of phase propagation of the wave as a whole. We therefore apply a cosine-square-weighted smoothing of 30° in longitude and 9° in latitude direction, resulting in a gradual change of the phase speed inside an RWP (H. Nakamura 2013, personal communication). Finally, the entire flux is smoothed by a cosine-square-weighted average including the grid points inside the adjacent 4.5° in the meridional direction and 7.5° in the zonal direction.

f. Wave packet tracking

To highlight and visualize the spatiotemporal evolution of the investigated RWPs, a newly developed tracking method for RWP objects will be applied. Basically, the RWP objects are defined by the areas of the envelope field that exceed a specific threshold value. The RWP trajectories are calculated by temporal tracking the center of mass of the RWP objects in longitude–latitude maps. The aim of the developed tracking method was to keep it as simple as possible. To get useful results and to deal with the difficulties related to the tracking of RWPs, the main effort was put into the definition of RWP objects and the choice of an appropriate threshold value. The technical details and the exact definitions for the tracking method are described in the appendix.

Originally we created and tested our tracking algorithm mainly for the methods of Zimin et al. (2003) and Zimin et al. (2006), although in principle it can be applied to any phase-independent measure of the RWPs. Later we also applied our tracking algorithm to the wave activity flux, which works well for the cases investigated in this paper. As the wave activity flux is somewhat noisier than the methods of Zimin et al. (2003) and Zimin et al. (2006) (high wavenumbers are neglected), the tracking of RWPs defined by the noisier wave activity flux could lead to difficulties in some cases. This may be a concern when applying the algorithm to large datasets to create a climatology. Future research should investigate how the wave activity flux can optimally be exploited for tracking.
For instance, this could be in the form of directly applying the tracking to the flux (with possibly further modifications) or to combine the tracking of RWPs calculated by the method of Zimin et al. (2006) with the wave activity flux as a hybrid method, to make use of the advantages/additional information of the wave activity flux. Furthermore, it appears more logical to directly track wave activity $M$ instead of the modulus of the flux $F$. However, for practical reasons we decided to track the wave activity flux. While the calculation of the stationary contribution of $F$ is unproblematic, some difficulties can occur when computing $M$. First, the expression for $M$ contains the phase propagation speed $C_p$, which is quite noisy and requires stronger additional smoothing. Second, both denominators of $M$ can become close to zero; following Takaya and Nakamura (2001), $M$ is set to zero in those regions. In all cases investigated in the current paper, tracking of the wave activity flux works without problems.

3. Results

a. Benefit of the semigeostrophic coordinate transformation for the wave activity flux diagnostic

The derivations leading to the formulation $F$ are based on a number of assumptions, one of which posits that the carrier wave is a sine function [see beginning of section 2 in Takaya and Nakamura (2001)]. As was explicitly argued in Wolf and Wirth (2015), the latter assumption is not satisfied for observed wave packets owing to the semigeostrophic nature of Rossby waves. The latter implies that generally troughs are narrower than ridges (see Fig. 1). This generates artifacts in methods which reconstruct the envelope using the Hilbert transform like Zimin et al. (2003) or Zimin et al. (2006). To alleviate the problem, Wolf and Wirth (2015) suggested to perform a semigeostrophic coordinate transformation before reconstructing the envelope from the meridional wind field. Indeed, similar problems arise when computing the wave activity flux $F$, and we are going to show in this section that the use of a semigeostrophic coordinate transformation is also beneficial in this case.

The effect of the transformation on $F$ for a specific case is demonstrated in Fig. 2. Without the transformation (Fig. 2a) the RWP shows two maxima in $|F|$ at the location of the troughs (at 120° and 70°W) and a clear minimum at the location of the ridge (at 100°W). In a similar case, Danielson et al. (2006) argued that the two relative maxima indicate lower-tropospheric sources of wave activity resulting from the surface cyclones associated with the troughs. However, here we argue that this effect is at least partly due to the semigeostrophic nature of the Rossby wave packet and, in this sense, spurious. Indeed, after the application of the semigeostrophic coordinate transformation the diagnostic $|F|$ indicates a more coherent wave packet with less pronounced relative maxima and minima (Fig. 2b). We conclude that the two maxima in the original analysis (Fig. 2a) are at
least partly spurious and cannot entirely be attributed to related sources of wave activity. But even after the transformation, the strongest flux values are still located inside the two troughs. This can be expected, as the applied transformation of two-dimensional semigeostrophic theory does account only for the spatial asymmetry between troughs and ridges. For the full three-dimensional semigeostrophic theory, troughs also tend to be stronger than ridges, which further leads to an asymmetry in strength (Snyder et al. 1991). Hence, the applied transformation does not fully get rid of the trough–ridge asymmetry, but it does reduce the asymmetry, which therefore leads to a more coherent wave signal. In the following, we will always apply the semigeostrophic coordinate transformation before further analysis.

b. Added value from the wave activity flux diagnostic

In the framework of individual cases we will now show that the wave activity flux diagnostic has some added value that cannot be obtained from more conventional diagnostics based on the meridional wind or its envelope. As we will see, a key asset in all cases is the information about the meridional (in addition to the zonal) propagation direction.

1) ONSET OF A WAVE PACKET

To obtain a better overview over the spatiotemporal evolution of the RWP from Figs. 1 and 2, we present a corresponding Hovmöller diagram with the calculated RWP trajectory in Fig. 3. Note that this diagram includes the date of the snapshots from the two previous figures. The quantity plotted in this Hovmöller diagram is the magnitude of the zonal component of $F$. The rational for us to average the magnitude of the zonal component of $F$ instead of $|F|$ is not obvious at first sight, but it can be understood by comparing the following two hypothetical RWPs. Let the first one be purely zonally propagating and, therefore, the meridional average of $|F|$ and the magnitude of the zonal component of $F$ are identical. Let the second RWP be identical to the first one, but rotated by an angle $\alpha$ in the horizontal plane. By doing so the second wave packet has a nonvanishing meridional propagation direction. The rotation leads to a larger meridional width of the wave packet [increase by $1/\cos(\alpha)$], which results in a larger value on the Hovmöller diagram from averaging $|F|$ in the meridional direction. If we average, instead, the magnitude of the zonal component of $F$, the two wave packets are represented by the same strength in the Hovmöller diagram, since for the rotated wave packet the use of the zonal component of $F$ [which is smaller than $|F|$ by a factor $\cos(\alpha)$] exactly compensates for the effect of the increased meridional width of the wave packet. Therefore, averaging the zonal component of $F$ gives equal weight to RWPs, which have the same strength $|F|$ but that differ in their meridional component. Figure 3 indicates a coherent RWP starting west of $135^\circ$W at the beginning of August and propagating all the way across the Atlantic Ocean toward Asia. As in Fig. 2, the troughs still appear as relative maxima in Fig. 3 despite the use of the semigeostrophic coordinate transformation.

It is instructive to compare the wave activity flux diagnostic with more conventional diagnostics based on envelope reconstruction of the meridional wind. For instance, Glatt and Wirth (2014) investigated the same RWP as the one studied here. The authors followed Zimin et al. (2003) and applied a Hilbert transform to the meridional wind latitude by latitude while keeping only zonal wavenumbers 5–14. For a fair comparison, we repeated their analysis on the 300-hPa level (which we use in the present analysis) instead of the 250-hPa level (which they used in their analysis). In addition to the slightly different pressure level, we also applied the semigeostrophic coordinate transformation before the envelope reconstruction. The result can be seen in Fig. 4a. A comparison of this figure with the original analysis of
Glatt and Wirth (2014, see their Fig. 3) indicates that the semigeostrophic coordinate transformation leads to a much more coherent structure in the Hovmöller representation of the RWP. Details are discussed in Wolf and Wirth (2015).

After 5 August, the two Hovmöller diagrams in Figs. 3 and 4a show an overall similar evolution of the RWP, although the behavior is somewhat smoother in Fig. 4a. However, during the early stage of the RWP life cycle there is a distinct difference between the two Hovmöller diagrams. The evolution of $F$ in Fig. 3 seems to indicate a change of group speed between 3 and 4 August, with nearly vanishing group velocity before that date. The onset of the related RWP trajectory on this diagram can be traced back to the end of July at 135°W (blue line on the Hovmöller diagram). On the other hand, the onset of the RWP in Fig. 4a seems to be on 3 August at around 170°W.

We aim to understand the different evolution of the RWP in the above two diagnostics by considering its trajectory on a longitude–latitude map. This is shown atop of the Hovmöller diagram in both figures. The RWP identified by $F$ in Fig. 3 originates at high latitudes around 75°N, 135°W propagating southward. Owing to the meridional averaging, a Hovmöller diagram can only represent zonal propagation; an RWP with a purely meridional group velocity has zero group velocity in the zonal direction. Consequently, the small zonal group velocity of the RWP in the Hovmöller diagram of Fig. 3 is a result of the predominantly meridional direction of propagation. On the other hand, the diagnostic in Fig. 4a completely misses the generation of the wave packet over the Beaufort Sea. Instead, it connects the North American RWP with an anomaly over the mid-Pacific. This is consistent with the path of the RWP (see the trajectory on the latitude–longitude map atop the Hovmöller diagram in Fig. 4a), which starts at about 45°N and 170°W. The poor representation of meridionally propagating RWPs is a known deficiency of the method of Zimin et al. (2003). The underlying reason is that the envelope is reconstructed along circles of constant latitude, such that anomalies at different latitudes cannot readily be interpreted as part of the same wave packet.

For a deeper understanding of the dynamics of this RWP, we show latitude–longitude snapshots of $F$ at different times in Fig. 5. The chosen dates are separated by 3 days each, starting at 0000 UTC 3 August 2002. Obviously, unlike a Hovmöller diagram, this sequence of snapshots includes information about the meridional
propagation. The first snapshot (Fig. 5a) shows a time briefly after the formation of the RWP. The arrows clearly indicate that the RWP (between 180° and 135°W) originates at high latitudes, consistent with the composite at the top of Fig. 3. The formation of the RWP is associated with an overturning ridge, which can be seen from the contours of geopotential north of 60°N. This ridge started to develop its northeastward tilt during the previous days (not shown).

As mentioned earlier, the method of Zimin et al. (2003) is prone to misinterpretation in cases of meridionally propagating RWPs. To eliminate this weakness, a few years later an improved method was designed by Zimin et al. (2006). Instead of calculating the envelope field along latitude circles, it is now calculated along streamlines of the background flow. The authors showed that generally this leads to an improvement, although the modification renders the diagnostic computationally more expensive and sensitively dependent on the exact definition of the background flow. In our implementation we used the background flow as described in section 2c. As a modification to Zimin et al. (2006), we use in our analysis that component of the horizontal wind, which is perpendicular to the streamlines; this adaption proved to be beneficial in the present case as well as in other cases not mentioned here. In addition, we adapt the selected wavenumbers to the length of the streamlines in order to obtain always RWPs with the given wavenumbers in the zonal direction. Instead of using a fixed wavenumber range we further choose wavenumbers depending on the latitude from which the envelope is calculated. We choose a cosine decay of the wavenumber range that mainly captures the dominant part of the latitude-dependent power spectra of the meridional wind. The latitude-dependent power spectra of the meridional wind and the chosen wavenumber range can be seen in Fig. 6.
The chosen wavenumber range follow a cosine decay to higher latitudes. By this definition, the chosen wavenumbers will not be discrete, but continuous. For the application to the Fourier transform in the method of Zimin et al. (2006) we allow a pro rata contribution of nondiscrete wavenumbers. This means, for a wavenumber range of 4.4 to 13.2, we take the full contribution of the wavenumbers 5–13, 60% of wavenumber 4, and 20% of wavenumber 14. By doing so, we get a smooth transition of the envelope field at different latitudes.

With the improvements in Zimin et al. (2006) and our additional refinements, this method is also able to diagnose the high-latitude origin of the current RWP (Fig. 4b). However, the trajectory calculation in this special case is quite sensitive to the choice of the threshold. With a slightly higher threshold we were not able to calculate a trajectory originating from high latitudes, but just starting a few days later around 50°N. Further, the use of the wind perpendicular to the background flow seems crucial to calculate trajectories with a meridional propagation direction.

The main critical point of the method of Zimin et al. (2006) may be, however, the necessary choice of the background flow. The method uses the Hilbert transform of a wave quantity (e.g., meridional wind) along the calculated background streamlines. This means that the wave quantity at remote locations is used to calculate the resulting RWPs. If the calculated background streamlines incorrectly do not connect all the anomalies representing a wave signal, the resulting envelope field will have lower values, possibly even fall below the chosen threshold. This is also the critical point for the onset of the RWP connected with the Elbe flooding. The onset occurs at the edge of a large-scale ridge that existed only for a limited time period. Therefore, the calculated background streamlines depend quite sensitive on the exact choice for the time filtering of the background flow. The wave activity flux on the other hand is calculated locally and, therefore, does not depend on the successful connection of related anomalies.

This example of the RWP associated with the Elbe flooding suggests that the method of Zimin et al. (2006) is in general able to represent wave packets moving along the midlatitude waveguide. But in some cases, where the waveguide is displaced for a shorter time period, the method sometimes does not represent the RWPs correctly. This problem results from the calculation along streamlines where remote anomalies are used to calculate the RWP strength. Because of the local calculation of the wave activity flux, this diagnostic is more adequate to represent such events.

2) VARIATIONS IN THE STRENGTH OF A WAVE PACKET

We now return to Fig. 5 and discuss the evolution of the RWP after its onset. During its early stage (Fig. 5b), the RWP remains quite compact and propagates eastward. Three days later (Fig. 5c) the RWP is strongly elongated in the zonal direction. Part of the wave activity seems to be transported to the southern part of the North Atlantic Ocean, while another part of wave activity continues to be transported eastward toward Europe. Both at this time and three days later (Fig. 5d), the vector $F$ is slightly diffusive between Europe and Iceland. The RWP seems to interact with the larger-scale ridge over central and eastern Europe, which implies that the RWP is losing wave activity to the background flow or by dissipation. At the same time, a significant fraction of the wave activity is being transferred to the subtropical jet over the Mediterranean with farther eastward propagation, but the overall values of $|F|$ are somewhat reduced. The latter is consistent with the diffusiveness of $F$ during the previous days.

In summary, the sequence of snapshots of the wave activity flux in Fig. 5 provides us with some detailed insight into the dynamics of this RWP, which can be obtained neither by the diagnostic of Zimin et al. (2003) nor the diagnostic of Zimin et al. (2006). In particular, the confluence or diffluence of $F$ helped us to understand changes in strength and compactness of the RWP. In regions of convergence the RWP tends to increase in strength. On the other hand, in regions of diffluence,
wave activity is spread over a broad region, which is associated with a decrease in the strength of the RWP. Consequential, the downstream development is then less effective and wave activity can be more easily lost due to dissipation. Without an additional source of wave activity, such an RWP will become weaker farther downstream. As we did not make use of the vertical flux component, such an additional source of wave activity could be given by the convergence of the vertical component of the wave activity flux.

3) PROPAGATION INTO AND FROM THE SUBTROPICS

The second case stems from an episode in December 1990. Figure 7 shows a conventional Hovmöller diagram of the upper-tropospheric meridional wind. Overall, the spatiotemporal evolution on this diagram is rather complex, but there is a well-defined RWP between 6 and 9 December propagating eastward across North America and the North Atlantic Ocean. This picture seems to be corroborated by individual snapshots of the meridional wind on longitude–latitude maps in Figs. 8a and 8b, indicating a single large wave packet extending from the Pacific Ocean (around 40°N, 180°) across the northern part of the North Atlantic all the way toward Europe (around 20°E).

However, snapshots of F tell a different story (Fig. 8c). The arrows in this plot suggest that what appeared to be a single large wave packet in the above diagnostics, is in reality a sequence of two wave packets. The first of the two wave packets extends across North America (160°E–90°W) propagating southward into the subtropics, while the second one extends across the Atlantic Ocean (north of 40°N, 70°W–45°E) originating from farther south. This new interpretation in terms of two separate wave packets becomes even more consistent when considering the situation 1.5 days later (Fig. 8d), where |F| clearly indicates two separate wave packets. In hindsight, this behavior might possibly have been guessed from the meridional wind field in Fig. 8b, but one would be hard pressed to make this statement without the additional information from the wave activity flux.

How does the improved method of Zimin et al. (2006) fare in this case? We implemented this method with the additional adaptions described earlier. Instead of two distinctive RWPs it rather shows one RWP (not shown here). To be fair, it should be mentioned that if one would apply a 3-day high-pass filter to cut off the very fast transients, as it is done for the wave activity flux, then one would also end up with two separate RWPs. On the other hand, including the very high-frequency transients for the wave activity flux still shows a separation between the two RWP. The wave behavior for this case is quite complex, as apparently quite different time scales are important for the RWP evolution. But even so, the wave activity flux could give us some further insight by the additional information of propagation direction.

4) DIAGNOSING THE ONSET OF WAVE BREAKING

We finally consider the same cases as in the previous subsections in order to show the potential of F for diagnosing the onset of wave breaking. According to linear wave theory, a strong jet acts as a waveguide (Martius et al. 2010), which means that RWPs tend to propagate along the jet. This idea is based on the fundamental concept of Rossby wave dynamics: the core of a jet coincides with strong meridional curvature of the zonal wind corresponding to a strong meridional gradient of potential vorticity (e.g., Dritschel and McIntyre 2008). A strong gradient of potential vorticity, in turn, represents a strong restoring force for parcels away from its unperturbed trajectory, which guarantees close to linear wave behavior. The jet in this context is considered to be part of the background flow, and to the extent that the jet makes some gentle excursions in the meridional direction, the RWP follows these excursions on its otherwise zonal and eastward trajectory. Anomalies due to the RWPs are measured with respect to the jet direction (i.e., following streamlines of the background flow).
If the direction of propagation of a wave packet has a large component away from the waveguide, this means that wave activity is transported into a region of relatively smaller gradients of potential vorticity. This implies a smaller restoring force and, therefore, a tendency away from linear wave behavior toward more nonlinear dynamics. We, therefore, suggest that the onset of wave breaking should coincide with the vector $\mathbf{F}$ having a significant component away from the direction of the waveguide. More specifically, we take the angle $\alpha$ between the first summand of $\mathbf{F}$ in Eq. (2) (without the term of $\mathbf{C}_\mu M$) and the streamlines of the background flow as an indicator for the propensity of the RWP to transit to a nonlinear stage and, possibly, to wave breaking. We do not include the second summand of $\mathbf{F}$ in Eq. (2), because this term just represents a projection of the phase propagation speed times the wave activity onto the background flow, which reduces the angle $\alpha$ everywhere inside of the RWPs. Real wave breaking is characterized by strong convergence of the wave activity flux. Therefore, as a further criterion for the wave breaking indicator we consider only those regions that show high values of WAF convergence or divergence. By including this condition we can exclude to some extent the RWPs that are just dissipating or having a diffuent WAF signal spread over a wide area. Taken this into account, we calculate the following simple RWP breaking index:

\[ \text{Bindex} = |\theta_{\text{WAF}}| - \frac{1}{2} \theta_{\text{WAF}}^2 \]
where WPB_{cond} is the condition of strong enough WAF convergence or divergence. This condition is given by smoothing the WAF divergence over 9° in latitude and longitude directions and considering only those regions exceeding a certain threshold (then WPB_{cond} = 1, otherwise WPB_{cond} = 0). The subjectively chosen threshold is 7 × 10⁻⁵ m² s⁻², which one would get by an RWP with only zonal components of \( F \) represented by a divergence of \( \Delta F \) per \( \Delta \varphi = 75 \text{ m}^2 \text{ s}^{-2} \) per 15° longitude at 50°N.

Both the example of the Elbe flood RWP as well as one of the RWP contained in the December 1990 case include nonlinear wave behavior with cutoff formation. The December 1990 case includes an RWP over the east Atlantic at 1200 UTC 6 December 1990 (Fig. 8c), leading to a strong trough over western Europe (Fig. 8d) and, finally, a cutoff cyclone during the ensuing evolution (Fig. 9a). The situation at 1200 UTC 8 December 1990 right before cutoff formation is analyzed in Fig. 9a, where the color shading quantifies the RWP breaking index defined in Eq. (4). The formation of the cutoff itself is indicated by the contour lines showing the geopotential height at consecutive times. Apparently, there are large values of |WPB| close to the location of the forming cutoff low around 0°. Thus, the onset of wave breaking is associated with large values of |WPB| at a time when this wave breaking has not yet materialized and is, therefore, not yet obvious from the meridional wind field itself. Further, the high |WPB| values at 120°W are also connected to apparently nonlinear wave behavior, indicated by the dashed geopotential contour lines.

As mentioned in the introduction, the RWP associated with the Elbe flooding also contained a nonlinear stage with cutoff formation. The cutoff formed around 11 August 2002. Therefore, we applied the same calculation of WPB also to this case. From Fig. 9b it can be seen that also this RWP breaking was preceded by strong values of WPB.

More complete information from the WPB index can be seen in the Hovmöller diagram in Fig. 10. This figure shows that the values of WPB during the wave breaking events in Fig. 9 are quite strong, at least for the given two RWPs. The additional strong values of WPB at 0° around 30 November 1990 is connected with a converging and breaking RWP into a large-scale European ridge connected with a strong ridge amplification. Therefore, for these cases the value of WPB could be used to successfully indicate the stage of nonlinear wave behavior and wave breaking onset. Whether or not the WPB is indeed a good indicator of wave breaking in more general cases will be considered in future investigations. In any case we believe that it can be used to indicate a more nonlinear behavior of the wave and, therefore, care must be taken by interpreting the wave activity flux, as it is valid only for weakly nonlinear waves.
4. Discussion and conclusions

Several methods have been proposed in the past in order to identify and characterize the dynamics of upper-tropospheric Rossby wave packets. In this paper we highlighted some benefits of the wave activity flux formulation of Takaya and Nakamura (2001), providing added value over more conventional methods based on the meridional wind alone. Our focus was on the horizontal propagation of RWPs, thus considering only the horizontal components of the wave activity flux.

First, we pointed out that the wave activity flux of Takaya and Nakamura (2001) is based on the assumption of an almost plane wave, which is usually violated when dealing with observed cases of RWPs. Instead, owing to the semigeostrophic nature of real RWPs, troughs are thinner than ridges. As a consequence, there are relative maxima in the strength of the wave activity flux associated with the troughs. These must be considered as at least partly spurious and should not be interpreted as indicating necessarily sources of wave activity. Following the ideas of Wolf and Wirth (2015) we showed that the problem can at least partly be overcome by applying a semigeostrophic coordinate transformation to reestablish the assumption of an almost-plane wave. It follows that the interpretation of relative maxima in the wave activity flux requires care, because they may be partly an artifact of the semigeostrophic nature of RWPs.

In the framework of individual cases we showed that the wave activity flux diagnostic provides a representation of Rossby wave packets that is “sharper” and easier to interpret than more conventional diagnostics based on the meridional wind and its envelope. Often, the key asset was the additional information regarding the meridional propagation of the RWP. This allowed us, for instance, to trace the origin of a particular RWP to polar latitudes, which helped to interpret the somewhat curious behavior of this RWP on a conventional Hovmöller diagram. In another case, RWP propagation into or out of the subtropics indicated that what appeared to be a single extended RWP at first sight, was more likely a succession of two independent RWPs. In other words, neighboring but independent RWPs could more easily be separated by drawing on the additional information of their propagation direction. Such a new perspective on the episode may have important implications for an evaluation of predictability. A single large-scale RWP should be predictable on a rather long time scale owing to its quasi-linear nature. On the other hand, propagation into the subtropics and reemergence of a secondary RWP from the subtropics is likely to involve highly
nonlinear dynamics, which may drastically limit the predictability during the later stage of the episode.

We also argued that the wave activity flux can be used to indicate the onset of wave breaking (i.e., the transition of the wave into its nonlinear stage). As a related measure we suggested to consider the angle between wave activity flux and the background flow in regions of stronger wave activity flux divergence or convergence. Large angles indicate the transport of wave activity away from the waveguide, which is generally connected with the deformation of anomalies followed by wave breaking. The convergence–divergence criterion is included to distinguish dissipation processes from wave breaking events. Dealing with wave breaking, it should be mentioned that the wave activity flux is actually only applicable for weakly nonlinear waves. Therefore, the suggested measure of wave breaking could at the same time be seen as “sanity check” for the interpretation of the flux. In regions where this wave breaking measure shows high values and therefore possibly nonlinear wave behavior, the flux must be considered with caution. Future research should elucidate how this measure for the onset of wave breaking compares with other wave breaking diagnostics.

Of course there are some caveats, and consideration of the wave activity flux cannot solve all questions in connection with RWPs. For instance, we did not make any use of the vertical component of the wave activity flux, nor did we consider the conservation relation connecting the wave activity flux with the associated wave activity. The latter seems to be difficult in light of the fact that the formulation of Takaya and Nakamura (2001) is based on linear wave theory thus neglecting higher-order terms in the wave amplitude. Real wave packets, on the other hand, typically have quite large amplitudes, resulting in an additional (but unknown) term in the conservation relation and thus preventing its quantitative application. A way out of this dilemma might be a formulation for finite-amplitude wave activity based on developments of Nakamura and Solomon (2010) or Methven (2013). However, until recently these formulations only considered zonally averaged budgets. Huang and Nakamura (2016) generalized this finite-amplitude wave activity to local wave activity, which looks like a promising tool to analyze zonally propagating wave packets.

The wave activity flux is, of course, not the only diagnostic that is capable of investigating RWPs in detail. Similar results can also be obtained by using energy fluxes (Orlanski and Chang 1993; Danielson et al. 2006). Here we focused on the wave activity flux and highlighted the useful aspects that can be derived from this diagnostic. The major advantage of the wave activity flux compared to the wave energy flux is the nearly phase independence of this diagnostic. Therefore, the wave activity flux seems to be quite a useful diagnostic if one is interested in the evolution of the wave packet as a whole. If one is rather interested in the evolution of the individual troughs and ridges inside the wave packet, then the energy fluxes would be more useful.

Overall we conclude that the use of a suitable form of a wave activity flux can give multiple useful insights into the dynamics of RWPs, insights that could not be obtained from more conventional methods like Hovmöller diagrams of the meridional wind, envelope reconstructions based on the meridional wind, or wave packet tracking. It remains a challenge to utilize this additional information for answering questions about the climatological behavior of Rossby wave packets or their predictability.

Acknowledgments. We thank the reviewers for their constructive comments, leading to improvements of the text. We acknowledge funding from the German Research Foundation (DFG) in the framework of the Research Group PANDOWAE and the Transregional Collaborative Research Center SFB/TRR 165 “Waves to Weather.”

APPENDIX

Wave Packet Tracking

RWPs are tracked by analyzing their envelopes on consecutive longitude–latitude maps as follows. First, we reconstruct the envelope using one of the methods described in the sections 2d and 2e; the envelope is additionally smoothed by a Hann filter that includes the grid points inside the adjacent 7.5° in the meridional direction. Then we apply a double threshold in order to identify the RWPs. Coherent regions with values above the upper threshold define RWP objects, while coherent regions with values above the lower threshold define RWP areas. A specific RWP is defined as the sum of all RWP objects that are contained in one single RWP area. For illustration see Fig. A1, where the RWP area is given by the area encircled by the black dashed line and the RWP objects are given by the areas encircled by the black solid lines. The three RWP objects are all located in the same RWP area and therefore are part of the same RWP. The center of mass and the properties of the RWP are calculated by using the RWP objects and not the RWP areas.

The tracking is done by simple overlap of RWP areas and RWP objects of the previous time step. We use a 12-h difference between successive time steps. If an RWP area at the time of investigation (t) has at least one
common grid point with an RWP object of the previous time step \((t - \Delta t)\), then the two related RWPs are identified as part of the same RWP trajectory. If an RWP area at time \(t\) has an overlap with two or more RWP objects of separated RWPs at time \(t - \Delta t\), the related trajectories are merged at time \(t\) to one trajectory. In case of two or more RWP areas at time \(t\) having an overlap with the same RWP object at time \(t - \Delta t\), the previous trajectory of the RWP at time \(t\) splits into two or more trajectories at time \(t\). For being able to track RWPs also at an early stage of their onset phase or up to the end of their decay phase (i.e., when the RWPs are quite weak), the tracking is done also by simple overlap of RWP areas at time \(t\) if no RWP object exists inside the RWP areas at time \(t - \Delta t\). This tracking of RWP areas during the onset and decay phase of the RWPs allows on the one hand a tracking over longer time spans, but on the other hand this frequently can lead to undesired merging of low threshold areas. Therefore, the tracking is only pursued until the first merging of RWP areas without included RWP objects occur or until RWP areas vanish because of decreased wave strength (below the low threshold). If an RWP splits into two or more RWPs and these RWPs merge again in a time span of two days or less, the RWPs and all their properties are represented by one single trajectory and the temporarily splitting is effectively ignored. This single trajectory of such temporarily split RWPs is realized by recalculating a single center of mass by a weighted average of the different RWPs at the short time span in which these RWP were separated. This recalculation is also done in case of simple merging events without previous splitting, if the merging trajectory existed only for no more than two days.

The double threshold with the separation of the RWP into RWP area and RWP objects is done in order to minimize the usual difficulties occurring for a simple threshold-based tracking method. These difficulties include frequent RWP splitting by the application of a single high threshold and frequent RWP merging of not necessarily connected RWPs by the use of a single low threshold. The use of a double threshold can mitigate both disadvantages while broadly keeping the advantages of a single low or single high threshold. By the use of the low threshold (RWP area) the frequency of unwanted RWP splitting events can be strongly reduced. This can be easily seen in Fig. A1. At this date, the meridional wind (blue and red contours) suggests a coherent RWP extending from about 160°W to 0°, which is captured by the low threshold (dashed black line in Fig. A1). By using only the high threshold (solid black lines in Fig. A1) for the tracking, the RWP would split into three separate RWPs. A further advantage of a single low threshold is that RWPs can be tracked over longer time spans, while a single high threshold would put the focus more on the time spans when the RWPs are fully developed with high envelope values. This can be clearly seen in the Hovmöller diagram in Fig. A2, where RWP trajectories are calculated by the use of a single low threshold (red lines), a single high threshold (white dashed lines), and a double threshold (blue lines), which is realized by a combination of the two single thresholds. One example for this tracking over longer time spans by the use of a lower threshold can be given by
the use of the double threshold (blue lines) shows that the trajectories calculated by using only the high value of the double threshold can be realized by an arctangent function:

$$\tau_h(\tau^* E) = \tau_{0,h} + \frac{2\Delta}{\pi} \arctan \left( \frac{\tau^* E - \tau_{0,h}}{\Delta/10} \right). \quad (A1)$$

where $\tau_{0,h}$ defines the basic threshold value, $\Delta$ gives the range of possible values of the arctangent function accounting for a seasonal dependence of the threshold, and $E$ is the mean envelope field, which is defined as the average range of possible values of the arctangent function accounting for a seasonal dependence of the threshold, and $E$ is the mean envelope field, which is defined as the average value of the RWP measure over the Northern Hemisphere between $15^\circ$ and $85.5^\circ$N. The product $\tau^* E$ represents a specific relative threshold, which means that the threshold $\tau_h$ is a function of this relative threshold.

For the use of a double threshold we include a second lower threshold, based on the same function, but with a slightly weaker slope at $\tau_{0,h}$. The function is given by

$$\tau_s(\tau^* E) = \tau_{0,s} + \frac{2\Delta}{\pi} \arctan \left( \frac{\tau^* E - \tau_{0,s}}{\Delta/10} \right). \quad (A2)$$

The double threshold used in this paper is given in Fig. A3. The chosen values for the calculation of the trajectories by the use of the method of Zimin et al. (2006) with our additional refinements are $\tau_{0,h} = 25 \text{ m s}^{-1}$, $\tau_{0,s} = 20 \text{ m s}^{-1}$, $\tau^* = 3$, and $\Delta = 10 \text{ m s}^{-1}$. These values are chosen to result in meaningful trajectories for all seasons. As this object-based tracking of a continuous field depends to
some extent on the subjectively chosen threshold values, the results show a certain sensitivity to these threshold values, but we would say it is about the same as for other methods. The sensitivity for the tracking based on a single threshold for one specific time period can be seen in Fig. A2. For this time period we calculated the trajectories not only with the use of the above presented double threshold (blue lines), but also by using only the low threshold (red lines) and the high threshold (white dashed lines) contribution of the double threshold. A comparison of these trajectories shows already an overall good agreement. The differences and with it the sensitivity can be summarized by two main points. First, the high threshold trajectories (white dashed lines) can be identified only over a shorter time range. This leads to a splitting of actual coherent RWPs and the missing of the early onset and the final decay phase of the RWPs. Second, low threshold trajectories (red lines) are able to identify an earlier onset and later decay phase of the RWPs, but sometimes tend to identify a merging of actual separated RWPs. The double threshold, which is a combination of the low and high threshold, can mainly get rid of both disadvantages and therefore decreases the sensitivity of the results on the choice of the threshold values. It is also true that one can also find time periods in which the sensitivity on the actual threshold is even higher than in the given example of Fig. A2, but especially because of this sensitivity we think that the use of a double threshold renders this issue less problematic. The advantage of the double threshold lies in the circumstances that it does not depend on one threshold, which must be suitable to prevent undesired splitting and allow longer RWP identification (which can be done by lowering the threshold), as well as prevent undesired merging of weak RWP signals (which can be done by increasing the threshold). By using a double threshold, the high and low threshold can be chosen to deal with these issues rather separately, allowing the individual high and low threshold values to be higher or lower than it would be advisable for a single threshold.

For the calculation of the trajectories by the use of the method of Zimin et al. (2003) we did use \( \tau_{0,h} = 30 \, \text{m s}^{-1} \), \( \tau_{0,s} = 25 \, \text{m s}^{-1} \), \( \tau^* = 2.8 \), and \( \Delta = 8 \, \text{m s}^{-1} \); and for the trajectories based on the wave activity flux of Takaya and Nakamura (2001) we did use \( \tau_{0,h} = 30 \, \text{m s}^{-1} \), \( \tau_{0,s} = 60 \, \text{m s}^{-1} \), \( \tau^* = 4 \), and \( \Delta = 20 \, \text{m s}^{-1} \). Those values were suitable for all chosen cases in this paper.

REFERENCES


Grazzini, F., and G. van der Grijn, 2002: Central European floods during summer 2002. ECMWF Newsletter, No. 96, ECMWF, Reading, United Kingdom, 18–28.


