Rossby Wave Packets on the Midlatitude Waveguide—A Review

VOLKMAR WIRTH AND MICHAEL RIEMER
Institute for Atmospheric Physics, Johannes Gutenberg University Mainz, Mainz, Germany

EDMUND K. M. CHANG
School of Marine and Atmospheric Sciences, Stony Brook University, State University of New York, Stony Brook, New York

OLIVIA MARTIUS
Oeschger Centre for Climate Change Research, and Institute of Geography, University of Bern, Bern, Switzerland

(Manuscript received 6 January 2017, in final form 20 March 2018)

ABSTRACT

Rossby wave packets (RWPs) are Rossby waves for which the amplitude has a local maximum and decays to smaller values at larger distances. This review focuses on upper-tropospheric transient RWPs along the midlatitude jet stream. Their central characteristic is the propagation in the zonal direction as well as the transfer of wave energy from one individual trough or ridge to its downstream neighbor, a process called “downstream development.” These RWPs sometimes act as long-range precursors to extreme weather and presumably have an influence on the predictability of midlatitude weather systems. The paper reviews research progress in this area with an emphasis on developments during the last 15 years. The current state of knowledge is summarized including a discussion of the RWP life cycle as well as Rossby waveguides. Recent progress in the dynamical understanding of RWPs has been based, in part, on the development of diagnostic methods. These methods include algorithms to identify and track RWPs in an automated manner, which can be used to extract the climatological properties of RWPs. RWP dynamics have traditionally been investigated using the eddy kinetic energy framework; alternative approaches based on potential vorticity and wave activity fluxes are discussed and put into perspective with the more traditional approach. The different diagnostics are compared to each other and the strengths and weaknesses of individual methods are highlighted. A recurrent theme is the role of diabatic processes, which can be a source for forecast errors. Finally, the paper points to important open research questions and suggests avenues for future research.

1. Introduction

The upper-tropospheric flow in midlatitudes is predominantly directed from west to east. Embedded deviations from the zonal direction give rise to undulations that are referred to as Rossby waves (Rossby et al. 1939; Rossby 1940; Haurwitz 1940). Rossby waves owe their existence to the occurrence of gradients of potential vorticity (PV; Hoskins et al. 1985). In the atmosphere, a northward gradient of background PV is, to leading order, provided by the combination of the rotation and the sphericity of Earth [for an introductory-level text on Rossby waves see Rhines (2002)]. Rossby waves are in distinct contrast to other types of waves such as gravity waves or sound waves, which rely on gravity or the compressibility of air, respectively, for their basic restoring mechanism.

The atmospheric general circulation cannot be understood without reference to Rossby waves because they transfer energy, moisture, and momentum across
Deviations from the zonal flow are often diagnosed using the meridional wind $v$. This variable is particularly well suited for the identification of Rossby waves, because its zonal Fourier spectrum has a strong contribution at the spatial scales of interest. In an idealized “incarnation” of a Rossby wave, $v$ is purely sinusoidal with a constant amplitude. Along a latitude circle on an upper-tropospheric quasi-horizontal surface (such as a pressure surface or an isosurface of potential temperature $\theta$), this can be written in the following form:

$$v(\lambda, t) = A \cos(s\lambda - \omega t),$$  

(1)

where $A > 0$ is the amplitude, $\lambda$ denotes longitude (measured in radians), $t$ denotes time, $\omega$ is the angular frequency, and $s$ is the zonal wavenumber, which measures the number of full wavelengths as one goes around the globe. Planetary-scale Rossby waves are typically characterized by zonal wavenumbers $s = 1, 2,$ and $3$, whereas synoptic-scale Rossby waves are characterized by higher wavenumbers.

Another idealized incarnation of a Rossby wave would be a single trough or ridge being equivalent to a single dipole of $v$. However, except in very rare cases, a Rossby wave appears neither as a purely sinusoidal circumglobal wave nor as a single trough or ridge. Instead, the amplitude $A(\lambda, t)$ is a function of longitude and time, giving rise to so-called Rossby wave packets (RWPs) with a finite number of troughs and ridges and being zonally confined to a limited region (Fig. 1). The underlying sinusoidal factor $\cos s\lambda$ (dotted line) is referred to as the carrier wave. The two red lines depict $\pm A(\lambda)$; they enclose the actual RWP $v(\lambda)$ (blue line). The zonally varying amplitude $A(\lambda)$ is sometimes referred to as the “envelope.” The spatial variation of the envelope (red) is more gradual than that of the carrier wave (dotted) or the RWP signal (blue).

A real world example is presented in Fig. 2. Figure 2a shows the midlatitude jet with large meridional undulations over North America. Over the rest of hemisphere, the jet is more zonally oriented. An alternative representation is provided in Fig. 2b, showing contours of geopotential height $Z$. The large-scale flow is directed along $Z$ contours to a good approximation; correspondingly, the $Z$ contours have strong undulations over North America. This behavior is also reflected in the meridional wind $v$ in Fig. 2b, which shows the characteristic pattern of an RWP over North America with alternating areas of positive and negative values. By contrast, $v$ is small in regions where the jet is more zonally aligned (as over central Asia). The envelope of $v$ computed along latitude circles (Fig. 2c) clearly indicates the presence of an RWP over North America and the absence of Rossby wave activity over Eurasia between $40^\circ$ and $60^\circ$N. Finally, in terms of Ertel (1942) PV (Fig. 2d), the same RWP appears as a sequence of large-amplitude troughs (i.e., tongue-like equatorward protrusions of high values of PV) and somewhat broader low-PV ridges.

### b. The paradigm of downstream development

According to linear wave theory, key features of any wave can be derived from its dispersion relation. For barotropic Rossby waves on the beta plane with a constant and purely zonal basic flow $u_0$, this relation reads as follows:

$$\omega = u_0 k - \frac{k \beta}{k^2 + l^2},$$  

(2)

where $k$ and $l$ are the Cartesian wavenumbers in the zonal and meridional direction, respectively; and $\beta$ represents the northward gradient of planetary vorticity [Rossby 1945; for a derivation in modern terminology see, e.g., Holton (2004)]. Note that $k$ is related to the spherical integer zonal wavenumber $s$ by $s = ka \cos \phi$, where $a$ is Earth’s radius and $\phi$ denotes latitude (Andrews et al. 1987). From the dispersion relation one can compute the zonal phase speed as $c = \omega/k$ and the zonal group velocity as $c_g = \partial \omega / \partial k$. The phase speed describes the speed of propagation of individual troughs and ridges, whereas the group velocity describes the speed of propagation of the entire RWP. Generally $c_g$ differs from $c$, pointing to the dispersive nature of Rossby waves (Rossby 1945; Holmäller 1949). Because $\beta$ is positive, one obtains $c_g > c$, which means that the envelope of Rossby waves moves eastward faster than individual troughs and ridges. This behavior is shown more explicitly in Fig. 3, where the envelope propagates eastward by about $40^\circ$ longitude in 48 h, while individual
troughs and ridges propagate eastward only by about 10° in 48 h. As a consequence, new troughs (labeled D and E) have formed within the 48-h time span to the east of the original RWP. This phenomenon is called “downstream development,” because midlatitude flow is from west to east and the new trough forms to the east, which is downstream of the original RWP. A mechanistic explanation in terms of PV will be provided later in section 3f.

The dispersive nature of Rossby waves also implies that an initial wave packet with a limited zonal extent gradually extends over a larger region as time proceeds.

c. Different types of RWPs and their propagation

There are different types of RWPs associated with different temporal and spatial scales, different types of forcing, and specific properties of the background flow.

![Figure 2](image1.png)

**FIG. 2.** An RWP over North America at 0000 UTC 7 Aug 2002. (a) Magnitude of the horizontal wind at 300 hPa (color shading, in m s\(^{-1}\)). (b) Meridional wind \(v\) at 300 hPa (color shading, in m s\(^{-1}\)) and isolines of 300-hPa geopotential height \(Z\) (black contours, every 150 m). (c) Envelope (color shading, in m s\(^{-1}\)) of the meridional wind \(v\) at 300 hPa. (d) Ertel potential vorticity on the 330-K isentrope (color shading, in PVU; 1 PVU = 10\(^{-6}\) K kg\(^{-1}\) m\(^2\) s\(^{-1}\)). The data used for this figure are from the ERA-Interim project (Dee et al. 2011).

![Figure 3](image2.png)

**FIG. 3.** Northern Hemisphere RWP at (a) 0000 UTC 7 Aug and (b) 0000 UTC 9 Aug 2002. Both panels show the 300-hPa meridional wind \(v\) (blue line, in m s\(^{-1}\)) restricted to zonal wavenumbers 4–10 and averaged over 40°–60°N, as well as the corresponding envelope \(\pm A\) (red lines). The position of individual troughs is diagnosed by the change from negative to positive \(v\) and denoted by capital letters A, B, C,.... The upstream trough A has practically vanished from the RWP during the 48-h time span, while two downstream troughs D and E have been created during that time.
Sometimes an RWP is little more than an individual trough or ridge (Sanders 1988); sometimes an RWP may span a substantial part of the globe (Branstator 2002). Some RWPs propagate along great circles, while others are ducted in the zonal direction.

Although there is a seamless transition between the different types of RWPs (Branstator 2014), it has proven useful to distinguish a few prototypes. Teleconnection studies have traditionally based such a distinction on the associated time scale (Blackmon et al. 1984). Slow fluctuations (>30 days) are typically associated with zonally extended tropical or subtropical forcing producing train-like wave patterns along great circles (Wallace and Gutzler 1981). In the neighborhood of the forcing the vertical structure of the perturbations corresponds to a first baroclinic mode (which means that anomalies in the upper and lower troposphere are out of phase), while the far-field response has a nearly equivalent-barotropic structure [which means that upper- and lower-tropospheric fluctuations are in phase; Hoskins and Karoly (1981)]. In contrast, fluctuations on the weekly to submonthly time scale are associated with wave patterns with a preference for zonal propagation (Blackmon et al. 1984; Hsu and Lin 1992). Again, their vertical structure tends to be equivalent barotropic (Hoskins and Karoly 1981; Wirth and Eichhorn 2014). Finally, there are transient, higher-frequency wave patterns with zonal propagation along the midlatitude jet. These wave packets may arise from baroclinic instability of the zonal flow in midlatitudes (Simmons and Hoskins 1978) and are often characterized by an out-of-phase interaction between the upper troposphere and the surface. While the low-frequency wave trains appear in longer-term (e.g., monthly) averages, the higher-frequency wave packets usually disappear in longer-term averages and are most clearly visible on instantaneous maps.

The propagation of RWPs in a baroclinic atmosphere can be estimated by computing the group velocity from the corresponding dispersion relation, which is a generalization of (2). In principle, Rossby waves propagate in all directions. However, the typical wind speeds in the stratosphere imply that synoptic-scale RWPs can usually not propagate upward into the stratosphere and are, hence, confined to the troposphere (Charney and Drazin 1961). For this reason, the discussion of RWP propagation often focuses on horizontal propagation on the sphere. The direction of horizontal propagation, in turn, depends on the relative magnitude of the horizontal wavenumbers $k$ and $l$ as well as on the background flow. In a purely zonal background flow with constant angular velocity (i.e., $u_0 \propto \cos \phi$, with $\phi$ denoting latitude), RWPs propagate along great circles (Hoskins and Karoly 1981). The real flow is more complex, and the existence of a strong jet leads to preferential propagation in the along-jet direction (Hsu and Lin 1992; Hoskins and Ambrizzi 1993; Branstator 2002; Schwierz et al. 2004b).

Unfortunately, there is no consistent terminology regarding the different types of RWPs. The low-frequency variety has traditionally been referred to as “Rossby wave trains,” and the more transient variety along the midlatitude waveguide has interchangeably been referred to as Rossby wave trains or “Rossby wave packets.” Here, we suggest to distinguish these two types by reserving the term Rossby wave packet for the transient synoptic-scale variety along the midlatitude waveguide.

d. Why are midlatitude RWPs interesting?

Recently, there has been a renewed interest in transient RWPs along the midlatitude waveguide. One reason is that RWPs often occur as precursors to extreme weather. Extreme weather refers to surface weather that falls into the tail(s) of the respective local distribution (e.g., precipitation exceeding the 95th percentile). To the extent that weather events inherit predictability from larger-scale dynamical features such as RWPs (Anthes et al. 1985), a better understanding of the RWPs may help to improve the weather forecast, and this is particularly relevant in case of extreme weather. An example is the episode in August 2002, when a quasi-stationary low pressure system over central Europe was associated with a long-lived precursor RWP (Fig. 2c); that low pressure system brought heavy precipitation in parts of Europe resulting in catastrophic flooding of the river Elbe. The forecast of this event was rather poor as little as a few days ahead of time. Each stage of the RWP life cycle may be subject to forecast errors, and it is important to obtain a better understanding of what stages and which processes contribute most strongly to poor forecasts. In particular, the role of diabatic processes has been discussed in this context (e.g., Rodwell et al. 2013).

Another area of recent interest is the predictability of the RWPs themselves and their role in downstream forecast error propagation. In practice, forecasts are likely to be limited by errors occurring on the large scales (Durran and Gingrich 2014). For instance, the amplitude of Rossby waves may grow locally because of baroclinic (Eady 1949) or barotropic (Lorenz 1972) instability, thus limiting their predictability. In addition, RWPs can transmit errors downstream (e.g., Rodwell et al. 2013), and there may be upscale error growth from the convective scale all the way to the planetary scale (Stensrud and Anderson 2001; Zhang et al. 2007).

These issues and questions gave rise to a novel research focus, in which RWPs are considered as meteorologically meaningful entities. While these efforts drew on previous knowledge about RWPs, they transcended...
more conventional approaches. In particular, there has been work to define RWPs in an automated manner, to follow them in a quasi-Lagrangian way, and to determine the properties of RWP objects.

**e. Scope and outline of this review**

This review provides an overview of the current knowledge about transient RWPs and their propagation along the upper-tropospheric midlatitude Rossby waveguide. It includes, in particular, new diagnostic methods, the association of RWPs with extreme weather, and the role of RWPs for predictability. We restrict our attention to transient RWPs on the synoptic to subplanetary scale with a finite extent in the zonal direction, having a characteristic time scale from a few days to a week or two. This usually implies RWP propagation in the along-jet direction as well as zonal localization of related processes. We also consider climatological properties of such RWPs. We exclude from consideration long-lived, quasi-stationary Rossby wave trains. We also exclude from consideration issues involving stratosphere–troposphere coupling, which were discussed in a recent review by Kidston et al. (2015).

A recurrent theme will be the role of diabatic processes for RWP dynamics, because these are believed to be particularly important in the context of predictability.

Our review aims to include all relevant publications from the last 15 years. This roughly covers the time period of the World Meteorological Organization research program called The Observing System Research and Predictability Experiment (THORPEX), which started in 2003 (Shapiro and Thorpe 2004). One particular part of this research program was the investigation of dynamical processes such as midlatitude RWPs and their role for weather prediction (Parsons et al. 2017). The THORPEX period includes the activities of a research group on the Predictability and Dynamics of Weather Systems in the Atlantic–European Sector (PANDOWAE), which was the main German contribution to THORPEX during the years 2008–14 (also see the PANDOWAE Special Collection, http://journals.ametsoc.org/topic/pandowae). We do not aim to provide a comprehensive coverage of all work done during the pre-THORPEX era, but we will present the key ideas and developments from that earlier time.

The review is organized as follows. We begin with a discussion of the midlatitude Rossby waveguide in section 2. Section 3 reviews the methods that have been designed to diagnose RWPs (i.e., their associated features, properties, and processes). Section 4 then discusses the science regarding the dynamics of RWPs and their life cycle. Climatological aspects are covered in section 5. The role of RWPs for weather forecasting is discussed in section 6, and their role in the occurrence of extreme weather is covered in section 7. Finally, section 8 summarizes the main achievements during the last 15 years, discusses caveats and limitations, and indicates outstanding issues and open questions.

**2. The midlatitude Rossby waveguide**

Preferred pathways of Rossby wave propagation have been associated with the notion of a Rossby waveguide. Waveguides are important for both high-frequency transient perturbations (Chang and Yu 1999) and low-frequency quasi-stationary perturbations (e.g., Blackmon et al. 1984; Hoskins and Ambrizzi 1993; Branstator 2002; Ding and Wang 2005; Branstator 2014; O’Kane et al. 2016). In this section we will elucidate the concept of a waveguide, introduce the concept of jet streams as efficient Rossby waveguides, discuss diagnostic tools to capture the Rossby waveguide, and present recent results.

**a. The concept of a waveguide**

In the framework of linear theory, a wave is defined as a small deviation from the basic state, and the basic state has an important influence on the propagation of the wave. A Rossby waveguide is a structure or property of the basic state that creates a propensity for an RWP to propagate along a certain path. There are two limiting situations that allow a quasi-analytic solution of the linear wave equation and, hence, a more specific definition of the term “waveguide.” Although the real atmosphere rarely (if ever) comes close to either of these limits, they are useful in a heuristic sense, to establish terminology, to define diagnostic tools, and to acquire qualitative insight.

The first limit is an atmosphere in which the latitudinal distribution of PV is piecewise constant with step discontinuities between the homogeneous patches (Platzman 1949, 1968). Each step discontinuity constitutes an idealized representation of a jet stream, because the latter are characterized by strong horizontal PV gradients (e.g., Swanson et al. 1997; Schwierz et al. 2004b; Dritschel and McIntyre 2008; Martius et al. 2010). Observations indicate that strong instantaneous PV gradients are, indeed, concentrated in a rather small fraction of the hemisphere (Schwierz et al. 2004b). This can be seen in Fig. 4a, which shows instantaneous gradients of PV on a quasi-horizontal isentrope. Flows with step discontinuities in PV have been investigated, for instance, by Swanson et al. (1997), Heifetz et al. (2004), Martius et al. (2010), and Hoskins and James (2014). In the case of a single PV discontinuity, the wave can only propagate along this PV front both in terms of phase speed and group velocity, and its amplitude decays exponentially in the direction perpendicular to the PV front. This behavior is consistent with the PV perspective on balanced dynamics, because Rossby waves are...
fundamentally related to material displacements across the background PV gradient, allowing the flow to create local anomalies of PV (Hoskins et al. 1985).

PV gradients arise from the latitudinal variation of planetary vorticity as well as from gradients of relative vorticity and static stability. The former is often referred to as the \( \beta \)-attributable gradient (Holton 2004). Strong gradients of relative vorticity and static stability are typically collocated with the jet streams along the tropopause breaks. On instantaneous maps, they show up in the form of narrow and elongated bands of strong PV gradients on jet-crossing isentropes (e.g., Shapiro and Keyser 1990; Schwierz et al. 2004a; Hoskins and James 2014, see also our Fig. 4a). The associated quasigeostrophic PV gradients are an order of magnitude stronger than the \( \beta \)-attributable gradient (Schwierz et al. 2004a). These bands are sometimes referred to as jet waveguides or tropopause waveguides. The PV gradient bands are narrow in the sense that their “... width is much smaller than the typical wavelengths of Rossby waves” (Harvey et al. 2016, p. 775). On a vertical cross section through the dynamical tropopause, these waveguides are collocated with steplike changes in the height of the dynamical tropopause or in the PV on jet-crossing isentropes (e.g., Shapiro and Keyser 1990; Schwierz et al. 2004a; Hoskins and James 2014). To the extent that the strong PV gradients in the upper troposphere are concentrated around the jet stream, being surrounded by regions of weak PV gradients in both the

---

**Fig. 4.** (a) Instantaneous PV gradient [color, in PVU (1000 km)\(^{-1}\)] on the 325-K isentrope at 0000 UTC 25 Oct 2008, the black line indicates the 2-PVU isoline. (b) PV gradient of the 30-day mean centered on 25 Oct 2008 [color shading, in PVU (1000 km)\(^{-1}\)] and the 2-PVU contour (black line). (c) September–November mean (1979–2013) PV gradient on the 330-K isentrope [color shading, in PVU (1000 km)\(^{-1}\)] with the black contours indicating the 2-, 3-, 4-, and 5-PVU isolines. The vertical color bar refers to (a) and the horizontal color bar refers to (b) and (c).
vertical and meridional directions, RWP s are expected to primarily propagate in the zonal direction along the jet streams. This behavior is, indeed, confirmed in observations (e.g., Hoskins and Ambrizzi 1993; Branstator and Teng 2017).

The second limit is a situation in which the background flow varies only marginally on the scale characterizing the wave. The underlying mathematical theory is usually referred to as WKBJ theory (Lighthill 1967). Similarly as in geometric optics, one performs “ray tracing” (i.e., one determines paths along which the RWPs propagate). Important in this context is a scalar diagnostic called the “refractive index”. The refractive index can be computed from the basic state and diagnoses RWP propagation in the sense that ray paths are bent toward higher values (Hoskins and Karoly 1981; Held 1983; Hoskins and Ambrizzi 1993; Lee and Feldstein 1996). A basic state with a strong zonal jet yields a zonally oriented local maximum of the refractive index. The local maximum is flanked by so-called turning latitudes on either side, where RWP propagation changes from poleward to equatorward and vice versa. Ray paths oscillate between the two turning latitudes and are, thus, ducted in the zonal direction along the local maximum (Hoskins and Ambrizzi 1993). It follows that a basic state with a strong zonal jet represents a zonal waveguide.

For both theoretical concepts the definition of the basic state (i.e., the background flow) is critical because its properties determine the properties of the waves and their propagation characteristics (Hoskins and Karoly 1981; Branstator 1983; Hoskins and Ambrizzi 1993). Often the background flow is defined as a low-pass-filtered or time-averaged state of the total flow field. In this case, the instantaneous sharp PV gradients and, hence, the waveguide property get—to a certain extent—lost, because the extratropical jet stream exhibits pronounced variability on various time scales (e.g., Lorenz and Hartmann 1993). This phenomenon can also be observed in Fig. 4, where the PV gradients of the 30-day mean (Fig. 4b) and the climatological mean (Fig. 4c) are approximately one-fifth as strong as the instantaneous PV gradients (Fig. 4a). Note also that the 30-day mean still contains a distinct structure over Europe that is the imprint of long-lasting quasi-stationary weather systems during that period, while the maxima of the climatological PV gradient are quasi-zonally oriented. As an improvement, Methven and Berrisford (2015) proposed the use of the so-called modified Lagrangian mean (MLM) state as a slowly evolving background state; it is defined as the zonally symmetric state obtained through adiabatic rearrangement of the flow such that every PV contour lying within an isentropic layer encloses the same mass and circulation as in the full flow. The rationale behind the MLM is the conservation of mass and circulation within a PV isoline on an isentropic surface. The computation of the MLM results in a stronger background flow than the climatological time average and, hence, stronger PV gradients.

If one transcends linear theory and accounts for nonlinear effects, the waves do have an impact on the background state. In practice it may, therefore, be an advantage to use a nonstationary background flow, which implicitly accounts for the feedback of the waves on the waveguide. For strongly nonlinear flows, the concept of PV mixing provides a conceptual framework to describe the feedback mechanism (Dritschel and McIntyre 2008; Hoskins and James 2014; Methven and Berrisford 2015). PV mixing by nonlinear eddies, such as breaking Rossby waves, reduces the PV gradients within the mixing region, but at the same time it sharpens the gradients at the edges of the mixing region. The breaking waves, hence, reinforce an old waveguide or create new waveguides [see Dritschel and McIntyre (2008) and Hoskins and James (2014) for more details].

b. Jet waveguides, baroclinic eddies, and storm tracks

The theoretical arguments above indicate that jet streams serve as efficient Rossby waveguides, suggesting an explanation why the observed waveguides are collocated with the jet streams. On Earth, there are essentially two jet streams: the subtropical jet and the extratropical, eddy-driven jet [Lee and Kim (2003) discuss the relevant dynamics]. As the terminology suggests, baroclinic eddies are a key ingredient for the latter [see Lorenz and Hartmann (2003) or Robinson (2006) for more detailed discussions]. Regions where baroclinic eddies occur preferentially are often referred to as storm tracks, and Wallace et al. (1988) refer to these regions as baroclinic waveguides. It is, therefore, not a coincidence that the extratropical Rossby waveguide and the storm tracks are closely related to each other and partly collocated (Hoskins and Valdes 1990; Chang and Orlanski 1993; Chang et al. 2002; Swanson 2007; Lu et al. 2010; Novak et al. 2015).

c. Diagnostic approaches

In the past, the identification of waveguides has followed essentially two techniques. The first technique diagnoses the pathways of individual transient eddies and accumulates related statistical information. For instance, Hsu and Lin (1992) found preferred pathways for eddy propagation by computing the teleconnectivity of low- or band-passed-filtered eddy streamfunctions, while Chang and Yu (1999) used time-lagged one-point correlation maps based on the meridional wind in the upper troposphere. In essence, these approaches go back to the
earlier teleconnection studies of Wallace and Gutzler (1981) and Blackmon et al. (1984), which are based on the correlation between the perturbations at different locations on the globe.

The second technique tries to extract information about waveguides from the background flow—that is, without explicit reference to the eddies—based on either of the two theories sketched earlier, invoking either PV gradients or the index of refraction. For instance, waveguides have been diagnosed by computing isentropic gradients of PV (Fig. 4; Schwierz et al. 2004b) or ln(PV) (Martius et al. 2010). Also, gradients of absolute vorticity on an upper-tropospheric pressure level have been used for this purpose, as they can be considered as an approximation to isentropic gradients of PV for the large-scale flow (Newman and Sardeshmukh 1998). Other authors made use of the ideas from WKBJ theory by computing the refractive index from observed background flows (Hoskins and Ambrizzi 1993; Yang and Hoskins 1996). Importantly, both techniques yield qualitatively similar results in some cases (Hoskins and Ambrizzi 1993). This renders the results more robust and shows that heuristic concepts can be useful even when the underlying assumptions and approximations are only marginally satisfied.

d. Efficient waveguides and waveguide interactions

Although theory provides some guidance, the question remains what makes a jet a good waveguide in the real atmosphere. The issue can be addressed by experimentation with numerical models. It turns out that the ducting property of a jet waveguide in a climate model depends on the strength of the jet stream (Branstator 2002). This result was later corroborated by idealized simulations with a barotropic model, indicating that a jet must be strong and, in particular, narrow in order to make it a good waveguide (Manola et al. 2013). The notion that narrow jets are good waveguides for propagation in the zonal direction is also found in Holman et al. (2014), who argued that the seasonal evolution of the time mean jet over the eastern Pacific has an important impact on Rossby wave propagation.

Depending on the background flow, there is sometimes more than one waveguide present in the meridional direction. If the waveguides are well separated meridionally, they can be treated as independent from each other. However, if these multiple waveguides are close to each other, RWPs may transfer from one waveguide to the other (Hoskins and Ambrizzi 1993; Martius et al. 2010). The latter happens relatively frequently over the eastern Pacific and over Europe, where waves have been seen to transfer from the extratropical onto the subtropical waveguide (Martius et al. 2010; Röthlisberger et al. 2016).

3. Diagnostic frameworks

In this section we present diagnostic frameworks that have specifically been developed to study RWPs and their dynamics. The reader who wants to proceed immediately to the science of RWPs can skip this section and later refer to it as needed.

a. Hovmöller diagrams

The 1940s experienced an unprecedented increase in the availability of upper-air data (Stickler et al. 2010). Not surprisingly, it was in the late 1940s that Cressman (1948) and Hovmöller (1949) developed a diagram, in which the latitudinal average of a variable from an upper-tropospheric surface is plotted as a function of longitude and time. An example is given in Fig. 5. This diagram, which is now commonly referred to as a Hovmöller diagram, condenses the spatiotemporal information and greatly facilitates the task of following the evolution of RWPs. It effectively exploits the excellent pattern recognition skills of the human brain and makes the process of downstream development evident. For instance, pairs of blue and red patches in Fig. 5 represent individual troughs and ridges; however, these patches are organized into a small number of “superstructures” that form an envelope to the individual patches, such as the one extending from 180° on 3 August all the way to 90°E on 15 August. This superstructure represents an RWP, and the speed of propagation on the Hovmöller diagram corresponds to its group velocity.

Historically, the Hovmöller diagram was born in a forecasting environment, especially as a tool to detect downstream development (Persson 2017). The diagram has been applied to various variables such as geopotential (Hovmöller 1949), the meridional wind (Chang 1993), the meridional wind squared (Chang 1993; Lee and Held 1993), and eddy kinetic energy (Chang and Orlanski 1993). Despite its general utility, there is a major limitation: the latitudinal average implied in its construction prevents one from obtaining information regarding latitudinal propagation, and this gives rise to occasional misinterpretations (Wolf and Wirth 2017).

There have been a number of adaptations and refinements to the classical Hovmöller diagram. For instance, Glatt et al. (2011) proposed to apply latitudinal weighting instead of averaging over a fixed latitude band, with the weighting function being proportional to the zonal variance of the meridional wind. This algorithm self-adjusts to the optimum range of latitudes and avoids the need to predetermine a fixed latitude band. Another algorithm makes the latitudinal band depend even on longitude with the aim to follow the main waveguide (Martius et al. 2006). A systematic comparison between
different types of Hovmöller diagrams shows that the refinements are beneficial in situations where otherwise the wave signal would leave the averaging band of the conventional version (Glatt et al. 2011). Overall, however, the plots from the different versions look quite similar, such that a simpler version may often be the method of choice owing to its robustness and ease of implementation.

b. Envelope reconstruction

In some applications, one aims to focus on the RWP as an entity, which is tantamount to discounting the phase of individual troughs and ridges within the wave packet (see Fig. 2c). This aim can be achieved by extracting the envelope of a wave-like variable such as, preferably, the meridional wind $v$ or some form of anomaly $v'$ (e.g., a deviation from the zonal mean or from a time average). In the early days, the task of envelope reconstruction was often performed using complex demodulation, which requires the choice of a reference wavenumber and implies some spatial smoothing (Lee and Held 1993; Chang and Yu 1999; Chang 2000). The need to specify a single reference wavenumber does not seem to be a serious issue in practical applications (Chang and Yu 1999), although it may create problems in specific situations (Zimin et al. 2003).

Zimin et al. (2003) suggested an alternative method for envelope reconstruction involving the Hilbert transform along circles of constant latitude. This method, which is well known in digital signal processing (Gabor 1946), can be combined with a restriction of the zonal wavenumbers to a user-specified interval. In contrast to complex demodulation, the Hilbert transform method does not require one to prespecify a single reference wavenumber.

In many applications, envelope reconstruction implicitly assumes that RWPs are oriented purely in the zonal direction (e.g., Zimin et al. 2003). Obviously, this is not always true. As an improvement, Zimin et al. (2006) suggested to apply the envelope reconstruction along streamlines of the background flow rather than along latitude circles. The background flow can be obtained from the wind data through a low-pass filter in time; it represents, in some broad sense, the waveguide along which the wave packets propagate. Zimin et al. (2006) convincingly showed that their new method performs better than the original method of Zimin et al. (2003). On the downside, the new method is computationally more expensive and less straightforward to implement. For instance, the need to specify a background flow opens some ambiguities and poses challenges when applying the algorithm to real-time forecasts. Also, it is not clear to date to what extent the algorithm allows a truly automated implementation.

Both the methods of Zimin et al. (2003) and Zimin et al. (2006) implicitly assume that RWPs are almost plane wave packets, which unfortunately is not satisfied for real RWPs. Rather, owing to the semigeostrophic nature of Rossby waves (Hoskins 1975), troughs are usually narrower than ridges, something that is well known to any synoptician and that is apparent in Fig. 2. This phenomenon gives rise to spurious wiggles in the envelope field when applying either of these two methods (Wolf and Wirth 2015). Combined with a threshold, these wiggles lead to a tendency for an RWP object to split into several fragments (see section 3c). This problem could be reduced by spatial smoothing or filtering, but this creates a tendency toward a spurious merger in other cases, which is undesirable. As a way out, Wolf and Wirth (2015) proposed a method, which reduces the tendency for spurious fragmentation without the need to apply spatial filtering or smoothing. Their key idea is to use the semigeostrophic coordinate transformation in order to perform the envelope reconstruction in semigeostrophic space rather than in physical space.

c. RWP objects

The Hovmöller diagram has proven useful to follow the space–time evolution of RWPs, but at its core it remains a field-based approach. This means that it produces a scalar field that requires interpretation by a human as part of the analysis, thus precluding any straightforward automation. This motivated several recent initiatives to design computer-based algorithms, which allow one to automatically identify and track RWPs. The definition of such RWP objects finally
makes explicit the idea that RWPs are meteorologically relevant entities. It opens new opportunities, like systematically screening reanalysis data and producing statistics regarding specific RWP properties such as their size, duration, and location of generation and decay.

Early approaches for RWP object identification were based on latitudinally averaged fields. A rather straightforward algorithm identifies objects as maxima of the envelope of the meridional wind exceeding a specified threshold (Grazzini and Lucarini 2011). Additional criteria on the translation between adjacent time steps are applied to guarantee the smooth propagation of these objects. Another approach is to identify objects as coherent regions on a suitable Hovmöller diagram (Glatt and Wirth 2014). Figure 6 provides an example showing how the RWP preceding the Elbe flood in August 2002 appears as an object. It is instructive to compare this Hovmöller diagram with the more conventional Hovmöller diagram of the meridional wind for the same period (Fig. 5). The abovementioned superstructure that appears in the beholder’s eye when looking at Fig. 5 turns into a somewhat inhomogeneous, but well-defined patch in Fig. 6. Incidentally, the distinct relative minima of the purple-colored patch at 90° and 15°W in Fig. 6 are, in part, due to the semigeostrophic nature of Rossby waves.

Algorithms that are based on latitudinal averages are doomed to fail in the event of multiple wave packets at different latitudes on a split waveguide. This problem motivated Souders et al. (2014a) to design a method that allows one to track RWPs both in longitude and latitude in a fully automated fashion. Another such algorithm was suggested by Wolf and Wirth (2017), using a partly self-adapting double threshold to reduce the tendency of spurious fragmentation and spurious merger. There has been no systematic investigation yet clarifying how the algorithmic differences between Souders et al. (2014a) and Wolf and Wirth (2017) impact the respective results.

d. Eddy kinetic energy

An RWP represents deviations from a background state, and it is natural to consider eddy kinetic energy (EKE) as a measure to quantify aspects of the RWP. Orlanski and Katzfey (1991) developed a local eddy energy framework to examine the life cycle of baroclinic RWPs. Eddies are defined as transients (i.e., deviations from a time mean flow). The equation governing the evolution of the eddy kinetic energy $K_e$ can be written as

$$\frac{\partial K_e}{\partial t} + \nabla \cdot \mathbf{S} = -\omega \mathbf{a}' - \mathbf{v}_h' \cdot [(\mathbf{v}' \cdot \nabla) \mathbf{v}_h'] - \mathbf{v}_h' \cdot ([(\mathbf{v}' \cdot \nabla) \mathbf{v}_h'] + \text{dissipation} \, , \quad (3)$$

where

$$K_e = \frac{1}{2} \mathbf{v}_h' \cdot \mathbf{v}_h' \, . \quad (4)$$

In these equations, the overbar denotes a time mean, the prime denotes deviation from the time mean, and subscript $h$ denotes the horizontal components of a vector (because in the primitive equations it is only the horizontal part of the kinetic energy that participates in the conservation of total energy following the 3D flow); $\mathbf{v}$ is the velocity vector, $\omega$ is the pressure vertical velocity, and $\alpha$ is the specific volume. In (3), the first two terms on the right-hand side represent baroclinic and barotropic conversion, respectively. The third term on the right-hand side also represents an energy transfer between the mean flow and the eddies, but averages out to zero in the time mean. Following Orlanski and Sheldon (1993), the energy flux $\mathbf{S}$ can be written as follows:

$$\mathbf{S} = \mathbf{v} K_e + \mathbf{v}' \Phi' - k \times \left[ \mathbf{v} \frac{\Phi^2}{2f(y)} \right] \, , \quad (5)$$

where $\Phi$ denotes geopotential. The first two terms on the right correspond to the advective flux and the geopotential flux, respectively. For Rossby waves, only the ageostrophic geopotential flux $\mathbf{v} \Phi'$ participates in the energy propagation because the geostrophic part of the geopotential flux is nondivergent. Orlanski and Sheldon (1993) recast the ageostrophic geopotential flux
into the last two terms on the right of (5). The flux $S$, together with the advective flux of eddy available potential energy, is to leading order equal to the total eddy energy times the group velocity (Chang and Orlanski 1994), and reduces to the eddy energy flux of barotropic Rossby waves (Hayes 1977) under linear quasigeostrophic scaling. Thus, $S$ represents the propagation of the envelope of the wave packet.

An example from the Elbe flood episode is given in Fig. 7a, which can be compared with the more standard diagnostics provided in Fig. 2. The maxima of eddy kinetic energy correspond to regions of strong meridional flow. Overall, the energy flux $S$ is directed from the upstream end of the RWP toward its downstream end, but the field $S$ is characterized by a large amount of structure related to the individual troughs and ridges. Part of this structure is due to nonlinear self-advection, which is included as part of the first term on the right-hand side of (5). Of interest is the energy flux at the downstream (eastern) end of the RWP, transporting energy into a previously less disturbed region. This downstream energy transport leads to the growth of a new energy center within the next 12 h (not shown) and signifies the eastward propagation of the RWP into this region.

d. Wave activity flux

Although eddy kinetic energy and its associated flux provide important insight into RWP dynamics, there is a crucial drawback: eddy energy is not globally conserved even under purely conservative (i.e., adiabatic and frictionless) conditions. By contrast, wave activity $A$ and the associated wave activity flux $F$ do satisfy a conservation relation,

$$\frac{\partial A}{\partial t} + \nabla \cdot F = 0,$$

for conservative flow. It means that the flux $F$ completely accounts for the local rate of change of wave activity $A$, and there are no further adiabatic sources and sinks, in contrast to the EKE budget equation (3). As a consequence, wave activity is globally conserved during the conservative propagation of a wave packet. This facilitates following the evolution of a wave packet, because wave activity is merely transferred from one location to another. In addition, the wave activity flux can be defined such that it equals the wave activity times the group velocity for waves in a slowly-varying background flow (Vanneste and Shepherd 1998).

The earliest expression for wave activity and its associated flux was designed to apply to zonal averages only (Eliassen and Palm 1961). This diagnostic would not be useful in our present context, as it does not allow one to follow the zonal propagation of an RWP. The first formulation that can be applied to snapshots of transient RWP provided by Takaya and Nakamura (2001), based on earlier work by Takaya and Nakamura (1997). Their method is designed with the aim to render both the wave activity and the associated flux phase independent (see also Esler and Haynes 1999a); in other words, these fields are meant to characterize the wave packet proper, discounting the phase information associated with individual troughs and ridges. In practice, complete phase independence is often not achieved, and this is partly due to the abovementioned semigeostrophic nature of Rossby waves (Wolf and Wirth 2015). It is, therefore, not surprising that the use of the semigeostrophic coordinate transformation increases the phase independence of the retrieved flux, which makes it easier to focus on the wave packet proper.

An example of applying the wave activity flux $F_{TN}$ is shown in Fig. 7b, to be compared with the corresponding EKE analysis in Fig. 7a. The flux $F_{TN}$ is large in the region of the wave packet, and its spatial distribution is smoother than for the flux $S$. In other words, $F_{TN}$ is more related to the wave packet as a whole than to individual troughs and ridges. Unfortunately, the strength of $F_{TN}$ is not entirely independent of the wave’s phase, which compromises the utility of $F_{TN}$ to a certain extent. Also, the flux $F_{TN}$ turns singular where the PV gradient of the background state is zero and where the phase speed of the wave equals the speed of the background flow (i.e., at the so-called critical levels). In practical applications, the method has therefore been restricted to regions where these problems do not occur.

Despite these caveats, the flux $F_{TN}$ may provide useful information. For instance, Danielson et al. (2006) studied the difference between eddy kinetic energy diagnostics and wave activity flux diagnostics in the framework of a case study. They compared the results from both diagnostics and found good quantitative agreement at least during the early linear stage of their wave packet. In particular, the directional information of the energy and wave activity fluxes yielded consistent results, and the baroclinic conversion term corresponded to the vertical component of the wave activity flux. Similarly, Wolf and Wirth (2017) found the information about the direction of propagation to be valuable for diagnosing the origin and the final stage of an RWP in a specific case. Thus, the wave activity flux appears to be a powerful complement to an eddy energy diagnostic.

There is yet another class of fluxes that allow one to diagnose the accumulated effect of transient eddies on the time mean flow. Examples include the $E$ vector of Hoskins et al. (1983), the flux of Plumb (1986), and the Trenberth flux (Trenberth 1986), which is a local version of the Eliassen–Palm flux (Eliassen and Palm 1961). A
variant of the Trenberth flux proves useful to diagnose the tilt of the eddies (Drouard et al. 2015). To our knowledge there have been so far only very few studies applying these types of fluxes to investigate RWPs, presumably because these fluxes cannot be used to diagnose instantaneous flow situations.

f. PV-based diagnostics

Analyzing Rossby waves in terms of PV goes back to the early work of Rossby (1940). Ever since, PV and the associated twin concepts of PV advection and PV inversion have proven highly successful to obtain a conceptual understanding of balanced dynamics (Hoskins et al. 1985).

Here, we provide an illustration of downstream development of an RWP from the PV perspective in Fig. 8 [cf. Fig. 9.13 in Hoskins and James (2014)]. We assume that there is a strong northward PV gradient somewhere in the middle of a zonally aligned channel. A single trough (i.e., a southward excursion of PV contours) at initial time \( t_1 \) corresponds to an isolated positive PV anomaly, which is associated with a localized counterclockwise flow anomaly (Fig. 8a). This flow anomaly advects the initial PV contours southward on the western side of the PV anomaly and northward on the eastern side of the PV anomaly. The net effect is that the trough is shifted westward, while a new ridge is generated to the east of the initial trough. Therefore, at a somewhat later time \( t_2 \), there is a trough–ridge couplet. At the subsequent time interval \( t_2 \to t_3 \), both the trough and the ridge (i.e., the entire RWP) are shifted westward through the action of the induced wind anomalies, and on the eastern flank another new trough is being generated (Fig. 8b). Thus, at time \( t_3 \) there is a trough–ridge–trough triplet. Adding a westerly basic flow, which exactly opposes the westerly phase propagation, one obtains the picture as shown in Figs. 8c, 8d, and 8e, which feature the essential signatures of downstream development (see Hoskins 1990). In essence, there are two elements responsible for the effect. First, the velocity
field obtained from PV inversion is of somewhat larger scale than the PV anomaly itself, such that the wind and the associated advection induced by a PV trough (or ridge) extends beyond the PV trough (or ridge) itself. Second, the wave packet is finite in longitude at any time, being limited both at its western and eastern edge. The trailing and the leading PV anomalies affect the PV contour just outside of the wave packet in an opposite manner: while the trailing PV anomaly creates a same-signed new PV anomaly at the western edge, the leading PV anomaly creates an opposite-signed PV anomaly at the eastern edge.

In a baroclinic atmosphere, surface temperature is equivalent to a very shallow PV distribution that typically has an equatorward rather than a poleward gradient (Bretherton 1966; Schneider et al. 2003). Similar considerations as in Fig. 8 apply, but they should now give rise to “upstream development” close to the surface, implying the generation of new anomalies on the western edge. Hence, in a baroclinic atmosphere a localized perturbation will disperse downstream in the upper troposphere and upstream near the surface, leading to an overall spreading of originally localized perturbations (e.g., Simmons and Hoskins 1979; Shapiro et al. 1999).

These central ideas of “PV thinking” can be applied in a quantitative manner. Nielsen-Gammon and Lefevre (1996) introduced a quasigeostrophic framework that diagnoses the contributions of individual, physically meaningful PV anomalies to the tendency of geopotential height. Variables are first partitioned into a large-scale background state and perturbations thereof. Following simple models of baroclinic instability by vertically interacting Rossby waves (Eady 1949; Hoskins et al. 1985), the PV field is further partitioned into upper- and lower-level anomalies. For the individual PV anomalies, the associated perturbations in geopotential, and, thus, the geostrophic wind are derived by piecewise PV inversion (Davis 1992).

Nielsen-Gammon and Lefevre considered the evolution of RWPs in terms of the evolution of the geopotential associated with the upper-level PV anomalies. The governing tendency equation contains six individual terms, which arise from the different combinations of the partitioned wind and PV field in the PV advection term [see Eq. (3.7) and Fig. 4 in Nielsen-Gammon and Lefevre (1996)]. Most important are advection of the background PV by (i) the wind associated with upper-level PV anomalies, which represents downstream development; (ii) the wind associated with lower-level PV anomalies, which represents baroclinic amplification; and (iii) advection of the upper-level PV anomaly by the background wind, which represents deformation. In addition, there are three nonlinear interaction terms, which are usually small.

![Fig. 8](https://example.com/fig8.png)

FIG. 8. Qualitative explanation of barotropic downstream development in the framework of PV thinking. Each panel represents a latitude–longitude map with the lines depicting contours of PV at different times $t_i$ ($i = 1, 2, 3$). The PV gradient of the background atmosphere is northward (i.e., upward in the figure). (left) The circled plus and minus signs represent PV anomalies; the curved arrows represent the wind field associated with these PV anomalies; the solid PV contour depicts an early stage, while the dashed contour represents the situation a short time later (i.e., after the wind had some time to advect the initial contour). (a) The time interval $t_1 \rightarrow t_2$ and (b) time interval $t_2 \rightarrow t_3$. (right) The scenario in a frame of reference in which the phase of the troughs and ridges is stationary.
The approach of Nielsen-Gammon and Lefèvre can be generalized for Ertel's (1942) PV instead of quasigeostrophic PV. This was done by Teubler and Riemer (2016). PV anomalies are defined as deviations from an approximately steady background state and are partitioned into upper- and lower-level anomalies. Then, piecewise PV inversion is applied to deduce advective tendencies associated with the respective anomalies. Teubler and Riemer considered RWPs in terms of PV anomalies on isentropic surfaces intersecting the tropopause. They performed piecewise PV inversion under nonlinear balance (Charney 1955; Davis 1992) and evaluated the tendency of the spatially integrated PV anomaly of individual troughs and ridges [see Eqs. (4) and (5) in Teubler and Riemer (2016)].¹

As in Nielsen-Gammon and Lefèvre (1996), Teubler and Riemer's diagnostic yields tendency terms that represent downstream development and baroclinic amplification, respectively. In addition, their method is able to quantify the tendencies due to the upper-level divergent flow and diabatic PV modification.

There are some important differences between these PV frameworks and the EKE framework described in section 3d. One point to note is that the energy (and geopotential height perturbation) associated with a given PV anomaly does not only depend on the size and magnitude of the PV anomaly, but also on its shape. This behavior occurs because different parts of a given PV anomaly may interfere destructively or constructively in terms of their associated wind field and, hence, EKE (Farrell 1982; Badger and Hoskins 2001). The quantitative impact of the shape of a PV anomaly on EKE becomes apparent when PV inversion is performed. The barotropic conversion of the EKE framework and the deformation term of the height-tendency equation of Nielsen-Gammon and Lefèvre (1996) correspond to changes in the shape of the perturbation under mean-flow deformation and shear. A comparable deformation term is missing from the Ertel-PV framework of Teubler and Riemer (2016) because the PV tendency is spatially averaged and the PV anomalies are not inverted to yield the associated wind and geopotential height anomalies.

Another difference is in the treatment of diabatic processes. Diabatic heating does not directly enter the EKE equation, but instead appears indirectly through an enhancement of baroclinic conversion (e.g., Gutowski et al. 1992). In addition, the upper-level divergent flow caused by diabatic heating below may also contribute to the ageostrophic geopotential flux term and, thus, be mis-interpreted as wave propagation. The Ertel-PV framework diagnoses diabatic processes more explicitly. The diabatic term in the Ertel-PV tendency equation quantifies instantaneous diabatic PV modification. In addition, the diabatic secondary circulation contributes to the divergent term in the framework of Teubler and Riemer (2016). Because large upper-level divergence is usually associated with latent heat release below, the divergent term can be considered to be at least partly due to diabatic processes. In addition, the divergent term includes the effect of the secondary circulations associated with the adiabatic balanced dynamics, and further analyses are needed to demonstrate whether and to what extent the diabatic contribution is indeed dominant. Diabatically produced PV anomalies contribute, in the subsequent development, to other tendency terms, for instance those representing baroclinic growth or wave propagation. Such indirect diabatic effects are inherent in all diagnostics discussed herein and signify the intrinsic coupling of balanced flow and moist physics. Some more details on the differences and commonalities of these frameworks are given in Teubler and Riemer (2016, see their section 3f).

4. The dynamics of RWPs and their life cycle

This section reviews our knowledge about the dynamics of RWPs and the understanding of associated processes. There is a particular focus on their quasi-zonal propagation because this feature is unique to wave packets and cannot be studied in the framework of circumglobal Rossby modes. We also include here the work that has been done about the role of diabatic processes for RWP dynamics. Figure 9 illustrates key processes that govern the propagation and maintenance of RWPs, and we will repeatedly refer to this figure during our discussion.

a. Propagation and localization of RWPs

1) Early interpretations

In the 1940s, observational studies by Namias and Clapp (1944), Cressman (1948), and Hovmöller (1949) indicated that downstream development occurs in the midlatitudes, and this was immediately interpreted as Rossby wave dispersion following the barotropic theory presented by Rossby (1945). Somewhat later it was found that the downstream intensification mainly occurs in the upper troposphere, and it was put into question whether this can be regarded as a purely barotropic process (Krishnamurti et al. 1977).

Meanwhile, Merkine (1977) and Thacker (1975) adopted into geophysical fluid dynamics a linear asymptotic theory that was originally developed in plasma physics to examine the growth and spread of localized perturbations...
in an unstable flow (Briggs 1964). These results showed how a localized perturbation in an unstable flow can grow and spread as a wave packet. However, this theory also predicts upstream (in addition to downstream) spreading of baroclinic waves with respect to existing disturbances, which was in conflict with the observations that were available at that time. We will return to this issue below.

Although observational studies continued to show downstream development, it was still unclear whether the dispersion is related to equivalent barotropic or baroclinic Rossby waves, as the upstream spreading predicted by the baroclinic instability theory was not observed (Joung and Hitchman 1982). At the same time, composite studies and regression analyses conducted to examine the structure and evolution of synoptic-scale baroclinic waves (Blackmon et al. 1984; Lim and Wallace 1991) indicated that these waves appear as wave packets that are largely advected eastward by the 700-hPa flow without displaying any signs of downstream development. These failures to find downstream development in synoptic-scale waves were later explained by Chang (1993) as being due to heavy time filtering of the geopotential height data (which is dominated by low-frequency variability) leading to a distortion of the temporal evolution of these wave packets.

2) THE DYNAMICS OF DOWNSTREAM BAROCLINIC DEVELOPMENT

The energetics of observed wave packets has been studied extensively using the eddy kinetic energy framework described in section 3d (Orlanski and Katzfey 1991; Chang 2000). Wave growth at the leading edge is generally dominated by the convergence of the ageostrophic geopotential flux (Fig. 9). In the center of the wave packet where the waves are mature, downstream radiation of the ageostrophic geopotential flux is balanced by strong baroclinic growth and convergence of the ageostrophic geopotential flux from farther upstream. At the upstream end, energy decay is mainly due to downstream radiation of the ageostrophic geopotential flux not being sufficiently balanced by the weaker baroclinic conversion there. Overall, the structure and energetics of these baroclinic RWPs resemble those displayed at the leading edge of growing baroclinic RWPs found in the modeling study of Simmons and Hoskins (1979).

These results have been summarized into a downstream baroclinic development paradigm by Orlanski and Sheldon (1995). Under this paradigm, the ultimate source of energy for the growth and maintenance of the RWP is baroclinic conversion from mean flow available potential energy. However, the bulk of the conversion does not occur in the downstream growing part of the RWP, but in the mature part toward the center and upstream portion of the RWP (Fig. 9). On average, contributions from barotropic conversion are rather small. Nevertheless, Chang (2000) also showed that the importance of individual processes can vary from case to case, and barotropic conversion can be important in either the growing...
or the decaying phase of a disturbance. Results based on these individual case studies have been confirmed by the composite structure of observed long-lived RWPs (Chang 2001). These RWPs are shown to be fueled by baroclinic conversion (or, equivalently, by upward wave activity fluxes) at the upstream end of the wave packet, and disperse downstream because of downstream directing energy/wave activity fluxes.

Wave dispersion in these studies is predominantly in the zonal direction, and local growth and decay of individual troughs and ridges are dominated by the convergence and divergence of eddy energy or wave activity fluxes (Fig. 9). This behavior is in contrast to the well-known paradigm of an eddy life cycle with baroclinic growth and barotropic decay, which is based on circumglobal normal modes and emphasizes wave propagation in the meridional direction (Simmons and Hoskins 1978). In other words, this paradigm does not appear to describe the observed behavior of individual troughs or ridges very well. Yet, the traditional eddy life cycle may well be applicable to the evolution of the zonally averaged behavior, which represents the averaged effect over a single RWP or multiple RWPs (Randel and Stanford 1985; Chang 2005b).

The downstream baroclinic development paradigm may be relevant in more practical terms. For instance, it helps one to better understand the development of surface cyclones in specific situations. A cyclone that is located in the center of an RWP can keep intensifying for a longer period than a cyclone that is located at the upstream end of an RWP (Decker and Martin 2005). The reason is that the cyclone in the center of an RWP can benefit from upstream energy fluxes to a larger extent than the upstream cyclone. Yet, whether this is a general behavior has not been demonstrated.

Apart from affecting the evolution of individual trough–ridge systems and wave packets, downstream baroclinic development has implications for the dynamics of storm tracks as a whole (Hoskins et al. 1983; Chang et al. 2002). Over the storm-track entrance region, which is geographically anchored by a region of enhanced baroclinicity (Hoskins and Valdes 1990), upper-tropospheric disturbances can tap into this baroclinicity and strongly amplify to trigger the growth of RWPs [see also section 4b (1), which focuses on RWP generation]. These RWPs then propagate downstream, feeding wave development into the less baroclinic downstream portion of the storm track. This process continues until the RWPs propagate into a region of strong deformation, which gives rise to barotropic decay (i.e., Rossby wave breaking), and terminates the storm track (Lee 1995; Swanson et al. 1997; Kaspi and Schneider 2011, 2013). Thus, RWPs can act as a conduit to extend storm tracks away from zones of enhanced baroclinicity into regions that are less favorable for baroclinic growth (Chang and Orlanski 1993).

3) THE UPSTREAM EDGE OF AN RWP

Linear theory suggests that baroclinic RWPs grow and expand longitudinally, with downstream development occurring at upper levels and upstream development near the surface (Merkine 1977; Simmons and Hoskins 1979). Observed RWPs generally maintain more or less the same longitudinal extent over their lifetime (Figs. 5 and 6). Given that downstream development is being observed at their downstream end (Fig. 5), this means that upstream development must not only be suppressed, but eddies on the upstream end of the RWP must decay. Surface friction and the beta effect tend to slightly suppress upstream development (Simmons and Hoskins 1979; Orlanski and Chang 1993), but linear modeling studies with beta and reasonable surface damping suggest that initially localized disturbances still tend to spread both upstream and downstream (Orlanski and Chang 1993; Swanson and Pierrehumbert 1994). The decay on the upstream part of the RWP has been hypothesized to be due to strong barotropic shear generated by upstream radiation of barotropic Rossby waves (Swanson and Pierrehumbert 1994). Alternatively, weakly nonlinear theory in an idealized configuration indicates that localization of an RWP may be achieved by nonlinear self-focusing of wave activity (Esler 1997; Esler and Haynes 1999b). However, observational studies by Chang (2001) did not find evidence to support either mechanism.

Why is upstream development not systematically observed? Thornicroft and Hoskins (1990) interpreted upstream development as a possible mechanism for secondary cyclone formation on trailing cold fronts. While several subsequent studies have followed up on this hypothesis to explain secondary frontal cyclones (e.g., Shapiro et al. 1999; Wernli et al. 1999), systematic upstream development has still not been observed. Wernli et al. (1999) suggested that the structure of the disturbances formed because of upstream development may be very sensitive to the structure of the parent disturbance that gives rise to the upstream development, and this may make upstream development hard to identify.

Another argument to explain the apparent lack of upstream development suggests that the effect of upstream-directed ageostrophic fluxes, which is responsible for upstream development, may be partly canceled by nonlinear advection and upward ageostrophic fluxes (Rivière et al. 2015). Nevertheless, upstream growth of new disturbances is found in idealized nonlinear simulations of initially localized perturbations (Simmons and Hoskins 1979; Orlanski and Chang 1993; Swanson and Pierrehumbert 1994). Therefore, it remains unclear
how effective this mechanism can generally be in inhibiting upstream development.

b. RWP generation and modification

1) RWP GENERATION

The generation of an RWP can be detected as the first occurrence of a ridge or trough on a previously undisturbed waveguide. However, not all of these initial wave disturbances develop into an RWP. In the seminal work of Simmons and Hoskins (1979), the authors used an idealized model setup, where a perturbation was embedded into an initially unperturbed jet. More recent idealized work prescribed perturbations “external” to the jet (e.g., Schwierz et al. 2004b). In the real atmosphere, a variety of processes and dynamical features may adopt the role of the initial perturbation, including recurving tropical cyclones (e.g., Jones et al. 2003a, to be discussed below; see Fig. 10), mesoscale convective systems (e.g., Rodwell et al. 2013), and warm conveyor belt outflows (e.g., Madonna et al. 2014b). All of these systems are associated with strong diabatic heating due to latent heat release in clouds. Further perturbations are provided by vortex-like stratospheric systems (Kew et al. 2010), also known as “coherent tropopause disturbances” (Pyle et al. 2004) or “tropopause polar vortices” (Cavallo and Hakim 2009). In addition, breaking Rossby waves on the midlatitude waveguide can excite new RWPs on the subtropical waveguide (e.g., Martius et al. 2010). This process is particularly relevant for wave initiation on the subtropical jet over North Africa. For a detailed discussion, see Röthlisberger et al. (2016), who introduce a criterion for the automated identification of synoptic-scale Rossby wave initiation on the extratropical and the subtropical waveguides.

Baroclinic conversion (Fig. 9) is an important mechanism for the amplification of an initially small perturbation, consistent with the observation that RWPs form preferentially in regions of enhanced baroclinicity (e.g., Chang et al. 2002). Applying the method of Röthlisberger et al. (2016) to ERA-Interim data (Dee et al. 2011), the northwestern Pacific, North America, and the North Atlantic have been identified as the three main regions of RWP generation. Further notable generation occurs over North Africa and the Middle East. These results are consistent with those based on other methodologies that will be discussed in section 5.

Another scenario for RWP generation starts from a circumglobal Rossby wave with near-constant amplitude. Esler (2004) suggested that there may be a natural tendency for such a wave to disintegrate into a succession of several wave packets. In his idealized model, a homogeneous circumglobal wave spontaneously developed undulations in its envelope because of a weakly nonlinear instability similar to the one found by Benjamin and Feir (1967). This idea is consistent with the simulations of Lee and Held (1993) showing that zonally localized RWPs tend to spontaneously develop in baroclinically unstable flows. However, the analysis of Esler (2004) is based on a number of assumptions and simplifications, and it is not clear to date to what extent it applies to the real atmosphere.

2) THE ROLE OF DIABATIC PROCESSES

Much of the work on RWP dynamics, including the paradigm of downstream baroclinic development, is based on models of dry balanced flow. This leaves open the question as to the role of diabatic processes including the release of latent heat and the impact of radiation. Recently, interest in the influence of diabatic processes on RWPs has increased, partly motivated by the connection of these processes to large medium-range forecast errors (section 6).

The effect of moist processes and associated latent heat release on baroclinic development has been extensively investigated through case studies, numerical experiments, and theoretical arguments. It is well established that moist processes invigorate baroclinic development and amplify surface cyclones (e.g., Danard 1964; Tracton 1973; Sanders and Gyakum 1980; Bosart 1981; Gyakum 1983; Golding 1984; Thorpe and Emanuel 1985; Emanuel et al. 1987; Davis and Emanuel 1991; Gutowski et al. 1992; Reed et al. 1992; Davis et al. 1993; Lapeyre and Held 2004). It is also well established that latent heat release locally modifies the tropopause and, hence, the jet structure (e.g., Kleinschmidt 1950; Hoskins and Berrisford 1988; Davis et al. 1993; Stoecking 1996; Dickinson et al. 1997; Wernli and Davies 1997; Bosart 1999; Henderson et al. 1999; Pomroy and Thorpe 2000). More recent work addressed the following important questions: Do these local modifications exhibit a downstream impact (i.e., do they modify the propagation of RWPs)? And how important are modifications by latent heat release compared to dry dynamics?

Early studies demonstrated a clear impact of latent heat release on individual weather events in the downstream region [e.g., a cyclone as in Hoskins and Berrisford (1988), or heavy precipitation as in Massacand et al. (2001)]. More recent studies interpreted the downstream impact in terms of RWP modification and demonstrated the influence of latent heat release within warm conveyor belts (e.g., Grams et al. 2011; Madonna et al. 2014a), organized convective systems (e.g., Rodwell et al. 2013), deep monsoon convection (Stensrud 2013), and tropical cyclones that recurve into the midlatitudes [see section 4b(3)].
The PV perspective (Hoskins et al. 1985) provides a useful framework to study the impact of diabatic processes on balanced flow. Davis et al. (1993) distinguished two general diabatic effects on the PV distribution. First, there is the material nonconservation of PV in the presence of diabatic processes (Ertel 1942), which is the so-called direct diabatic effect. Second, there is the indirect diabatic effect that comprises (i) PV advection by the balanced flow associated with diabatically produced PV anomalies and (ii) PV advection by secondary circulations that are associated with diabatic processes. We adopt this terminology here.

The intensification of surface cyclones by latent heat release can be explained by the diabatic generation of positive, low-level PV anomalies (e.g., Davis and Emanuel 1991; Davis et al. 1993; Stoelinga 1996; Davis et al. 1996; Campa and Wernli 2012; Binder et al. 2016). These low-level PV anomalies enhance the baroclinic coupling (e.g., Gutowski et al. 1992), increase RWP amplitude, and thus constitute an important indirect diabatic effect. However, in some cases, the increase of wave amplitude may in turn promote Rossby wave breaking (e.g., Davis et al. 1993; Riemer and Jones 2014), and thus initiate the decay of an RWP.

The direct impact of latent heat release on RWPs is due to the generation of an upper-tropospheric negative PV anomaly, which amplifies ridges and may induce downstream effects (e.g., Ahmadi-Givi et al. 2004; Chagnon et al. 2013; Chagnon and Gray 2015). Arguably, the more important effect on RWP amplitude, however, is due to the indirect impact from upper-tropospheric divergent outflow associated with latent heat release below. This process, too, amplifies ridges (see our Fig. 9 and, e.g., Davis et al. 1996; Riemer and Jones 2010; Piaget et al. 2015; Teubler and Riemer 2016; Schneider et al. 2017). The indirect diabatic ridge amplification is of similar importance for RWP amplitude as baroclinic growth and exhibits a large case-to-case variability (Teubler and Riemer 2016). Furthermore, the upper-tropospheric outflow may locally reduce phase propagation, in particular the progression of troughs, and may induce trough deformation and a reduction of trough amplitude (e.g., Pantillon et al. 2013a; Riemer and Jones 2014). It has, however, not been examined explicitly if this process reduces the group velocity of the affected RWP.

Besides latent heat release, the influence of longwave radiation on RWPs is a current research focus. Radiation is thought to sharpen the tropopause, and hence the tropopause waveguide, as the large vertical gradient of water vapor in the tropopause region implies large gradients in radiative cooling (Zierl and Wirth 1997). Chagnon and Gray (2015) hypothesized that radiative processes thereby influence the propagation of Rossby waves even on the weather time scale. Based on idealized models, it was shown that a sharper waveguide is associated with a faster phase speed (Harvey et al. 2016). More generally, longwave radiative cooling has been demonstrated to substantially modify the PV distribution near the tropopause within RWPs (Chagnon et al. 2013; Chagnon and Gray 2015; Teubler and Riemer 2016). These modifications tend to amplify troughs and weaken ridges. It is not clear, however, to what extent these radiative tendencies should be interpreted as modification of RWP amplitude or to what extent they modify the background state. It is also not yet sufficiently clear if the radiative sharpening of the tropopause occurs fast enough to be relevant on the synoptic time scale. In summary, while recent studies illustrate the potential importance of longwave radiation for RWP dynamics, fundamental questions still remain.

**Fig. 10.** Illustration of the excitation of an RWP by the extratropical transition of a tropical cyclone. (a) 120, (b) 156, and (c) 192 h, respectively, into an idealized numerical experiment with an initially straight jet. The shading denotes potential temperature θ, and the bold solid contours depict wind speed ($\approx 45 \text{ m s}^{-1}$) on the dynamical tropopause, here defined as the 2-PVU surface. Thin contours show surface pressure, every 5 hPa, dashed for 990 hPa and lower. The horizontal scale is given in (a). [The figure is Figs. 2a–c from Riemer et al. (2008), with modifications.]
For the sake of completeness, we note that the PV distribution near the tropopause may also be modified by other nonconservative processes like, for instance, turbulent mixing (Gidel and Shapiro 1979; Keyser and Rotunno 1990; Baumgart et al. 2018). Therefore, such processes may potentially influence RWPs. Yet, to our knowledge no systematic research has been conducted to this effect.

3) **Extratropical Transition of Tropical Cyclones**

Tropical cyclones that interact with the midlatitude jet may exert a large impact on the midlatitude flow, in particular those storms that move farther poleward and undergo extratropical transition (ET; Jones et al. 2003b; Evans et al. 2017). This impact is not restricted to the vicinity of the storm, but may extend over a large geographical region and is often associated with large medium-range forecast uncertainty and errors (Keller et al. 2018, manuscript submitted to Mon. Wea. Rev.). The latter, as well as the climatological aspects, will be discussed in subsequent sections (section 5c and 6c, respectively). Here, we are concerned with the dynamics and the physical processes governing the impact.

The impact of ET on the midlatitudes is mediated by the generation of a new RWP or the modification of an existing one. The general characteristics of this interaction can be illustrated by the generation of an RWP during ET in a highly idealized scenario with an initially straight jet (Fig. 10). Early during the interaction, the formation of a jet streak and of a ridge–trough couplet is evident downstream of the recurving tropical cyclone (Fig. 10a). Subsequently, the jet streak and the ridge–trough couplet amplify and promote the development of a new downstream cyclone (Fig. 10b). The emerging RWP then disperses farther downstream and promotes the development of another surface cyclone (Fig. 10c), following the conceptual model of downstream baroclinic development. From the energetics perspective, the recurring tropical cyclone thus constitutes a source of eddy kinetic energy, which subsequently disperses downstream as an RWP (Harr and Dea 2009; Keller et al. 2014; Chen 2015; Quinting and Jones 2016).

Three key mechanisms are distinguished that govern the generation and amplification of the ridge–trough couplet:

1) ridge building downstream of ET by the upper-tropospheric divergent outflow from (i) the ET system itself (e.g., Davis et al. 2008; Riemer et al. 2008; Riemer and Jones 2010; Archambault et al. 2013; Grams et al. 2013a), (ii) the emerging warm conveyor belt along the developing warm front ahead of the storm (Torn 2010; Grams et al. 2013a), and (iii) predecessor rain events (Galarneau et al. 2010; Wang et al. 2009; Grams and Archambault 2016);

2) ridge building by warm advection caused by the cyclonic circulation of the ET system (Bosart and Lackmann 1995; Riemer et al. 2008; Riemer and Jones 2010, 2014; Quinting and Jones 2016); and

3) amplification of the trough downstream by the upper-tropospheric outflow anticyclone (Riemer et al. 2008; Riemer and Jones 2010, 2014).

The first mechanism describes how outflow associated with latent heat release below can perturb the midlatitude waveguide, thereby generating or amplifying an RWP (cf. “divergent amplification” in Fig. 9). The second mechanism signifies baroclinic growth (cf. “baroclinic conversion” in Fig. 9). In a barotropic framework, the second and the third mechanisms can be considered, in more general terms, as the interaction of a vortex with a jet, which has been investigated in idealized scenarios in the context of RWP generation by Ferreira and Schubert (1999) and Schwierz et al. (2004b).

The distinct jet streak developing during ET (Figs. 10a,b) is not only an intrinsic part of the amplifying baroclinic wave (Rotunno et al. 1994; Wandishin et al. 2000), but is also strengthened externally by the ET system (Riemer and Jones 2010; Grams et al. 2013a; Archambault et al. 2015). This external part of the jet streak formation is associated with the increase of the $\theta$ gradient on the dynamical tropopause (i) directly by the frontogenetical effect of the divergent outflow and (ii) by a further increase of $\theta$ on the equatorward side of the jet (where $\theta$ is already high) because of the presence of the outflow anticyclone.

Under which conditions does ET have the strongest impact on midlatitude RWPs? It turns out that in this context the phasing plays a crucial role (i.e., the relative position between the tropical cyclone and the midlatitude wave pattern) (Ritchie and Elsberry 2007; Riemer and Jones 2010, 2014; Grams et al. 2013b; Keller et al. 2014). For a strong impact, the tropical cyclone needs to move into a region conducive to midlatitude development (i.e., ahead of a trough) and approximately remain in this relative position (“phase locking”). The tropical cyclone’s upper-level outflow can reduce the wave’s phase speed and may thus promote phase locking (Riemer et al. 2008; Riemer and Jones 2010, 2014; Pantillon et al. 2013a,b, 2015). In a phase-locked configuration, the same ridge–trough couplet may be amplified over an extended period of time (Riemer et al. 2008; Archambault et al. 2015); the recurring tropical cyclone can be likened to a local wavemaker (Riemer et al. 2008) and ET can be interpreted in terms of a
resonant interaction (Hodyss and Hendricks 2010; Scheck et al. 2011a,b).

In addition to the phase-locking mechanism, the leading edge of an RWP has been hypothesized to be an optimal location where ET leads to the most pronounced and longest-lasting RWP amplification (Riemer and Jones 2010). By contrast, the impact on a preexisting, high-amplitude wave pattern with well-developed surface cyclones can be expected to be less pronounced (Riemer and Jones 2014). This difference in the impact, observed in idealized numerical experiments, can partly be explained by the key role of the downstream surface cyclone in communicating ET’s impact into the farther downstream region: consistent with the paradigm of downstream baroclinic development, amplified RWPs occur in these experiments only when the downstream cyclone has a faster and stronger development, too (Riemer et al. 2008; Riemer and Jones 2010). This faster and stronger development, however, does not occur when the downstream cyclone is already well developed (i.e., approximately vertically stacked) when influenced by ET (Riemer and Jones 2014).

Amplification of the downstream cyclone in the context of ET has only partly been investigated so far. On the one hand, there is clear evidence that the amplification of the downstream trough and the formation of a jet streak provide favorable upper-tropospheric conditions (Riemer and Jones 2010; Grams et al. 2013b; Riemer et al. 2014). On the other hand, at least for the jet streak, a small number of case studies did not reveal a consistently positive contribution to downstream cyclone development (Riemer et al. 2014). More generally, the relative location of upper- and lower-tropospheric features and not only their amplitude is of importance for baroclinic growth (e.g., Hoskins et al. 1985) and thus impacts the development of the downstream cyclone. However, this phasing aspect has not yet been considered explicitly in the context of RWP modification by ET.

Furthermore, the availability of moisture in the downstream region is important for the magnitude of the downstream impact (Riemer et al. 2008; Riemer and Jones 2010; Grams et al. 2011; Torn and Hakim 2015; Grams and Archambault 2016). Besides amplifying baroclinic conversion, upper-level divergence associated with warm conveyor belt outflow enhances ridge building downstream of the downstream cyclone. This observed importance of moist processes for the downstream impact of ET is consistent with the notion that upper-level divergence may make a first-order contribution to RWP amplitude.

Although most studies have focused on the generation or amplification of RWPs, in some cases ET may actually decrease the amplitude of an existing RWP. For instance, ET may induce wave breaking by deforming the upstream trough and, thus, initiate the decay of the RWP (e.g., Pantillon et al. 2013a; Riemer and Jones 2014; Pantillon et al. 2015). Consistently, in some cases the impact of ET on midlatitude development was, in fact, found to be detrimental for the evolution of the downstream surface cyclone (e.g., Agusti-Panareda 2008). The question that remains is: What is the climatological impact of ET on midlatitude waveliness? We will come back to this question later in section 5c.

c. Nonlinear effects: RWP decay and interaction with the mean flow

Earlier we have seen that nonlinearity may inhibit upstream development and maintain the zonal localization of RWPs (section 4a). In addition, nonlinearity is also important for the interaction between RWPs and the mean flow as well as for the decay of RWPs. The latter two effects will be discussed in this subsection.

Long-lived RWPs may have an impact on the zonal mean flow and hence on the background flow seen by the perturbations in linear theory. For instance, the poleward eddy heat flux associated with long-lived RWPs in the Southern Hemisphere acts to reduce the meridional temperature gradient, while the concomitant momentum flux convergence acts to sharpen the jet (Chang 2001, 2005b). Apparently, much of the mean flow modification occurs in the vicinity of the RWP. Most of the events with significant growth and decay of hemispheric total eddy kinetic energy can be traced back to a single RWP, especially when the RWP propagates across the highly baroclinic region in the south Indian Ocean.

The final stage of the RWP life cycle is usually associated with Rossby wave breaking (Fig. 9), that is, the highly nonlinear, irreversible deformation of PV contours (McIntyre and Palmer 1984). Four different types of wave breaking have been distinguished, depending on whether the wave breaks cyclonically or anticyclonically and whether the poleward or equatorward part of the breaking wave dominates (Davies et al. 1991; Thorncroft et al. 1993; Wernli and Sprenger 2007; Gabriel and Peters 2008; Masato et al. 2012). As a consequence of wave breaking, the zonal progression of the RWP is terminated and the wave energy is dispersed in the meridional direction. Concomitantly, Rossby wave breaking is associated with meridional eddy momentum fluxes, which result in changes of the position and strength of jet streams and hence the waveguides for subsequent RWPs.

d. Conclusions

The overarching theme of this section was the recognition that different physical processes act at different spatial locations within an RWP, and that the redistribution of eddy energy or wave activity in the downstream direction
is an important term in the relevant budget equations (Fig. 9). These processes, in combination with interactions with the RWP environment, give rise to a fairly complex RWP life cycle, which requires refined diagnostic methods for their investigation.

5. Climatological properties of RWPs
   a. Results based on regression and composite techniques

   Regression techniques have proven valuable to document the prevalence of RWPs in midlatitudes of both hemispheres, both in the warm and in the cold seasons (Chang 1993; Berbery and Vera 1996; Chang 1999; Chang and Yu 1999). These studies also identified preferred waveguides over which RWPs tend to be most coherent and that generally correspond to regions with strong PV gradients (see section 2).

   Cyclone-relative composites have been used to reveal the relationship between surface explosive cyclogenesis and upper-level troughs (e.g., Sanders 1986). A similar strategy was used by Chang (2005a) to identify the connection between deep surface cyclones and precursor RWPs. He found two pathways of RWP precursors to surface cyclones over the western Pacific: one over Siberia and another over the subtropical jet waveguide across southern Asia. The occurrence of an RWP on either of these tracks increases the probability of occurrence of a deep and rapidly developing surface cyclone over the western Pacific a few days later by roughly a factor of 2. A similar technique has been used to show that deep surface cyclones over Europe are preceded on average by long-lived (up to two weeks) precursor RWPs that extend over more than 360° in longitude during their lifetime (Wirth and Eichhorn 2014). The RWPs preceding European surface cyclones have a longer lifetime and a larger longitudinal extent than the RWPs preceding western Pacific cyclones. This difference may be related to the fact that long-lived RWPs tend to form over the Pacific and these propagate across North America and the North Atlantic to initiate cyclogenesis over Europe (see the discussion below in section 5b).

   Another composite study found that RWPs excited by cyclogenesis in the western North Pacific display an asymmetric shape in the zonal direction with a rather abrupt decay in the upstream direction and a more gradual exponential decay in the downstream direction (Hakim 2003). RWPs in the North Pacific often seed new wave packets over the North Atlantic, but not vice versa. Seeding of Pacific wave packets through Atlantic wave packets seems to be suppressed because the latter often deviate from the zonal direction and propagate equatorward into the sub-tropics. It follows that Northern Hemisphere RWPs rarely circumnavigate the entire hemisphere. By contrast, in the Southern Hemisphere, the RWPs may well circumnavigate the hemisphere more than once (Chang 2000).

   Composites on the level of a Hovmöller diagram are also useful to investigate how the properties of upper-tropospheric RWPs depend on the conditions in the lower stratosphere (Williams and Colucci 2010). The strength of the stratospheric polar vortex turns out to affect both the wavelength and the phase speed of the diagnosed RWPs, and these connections are mediated through differences in the vertical wind shear.

   b. Recent results based on RWP objects

   The first attempt to retrieve statistical information about RWP objects (see section 3c) was presumably the one by Grazzini and Lucarini (2011). They applied their algorithm to 50 years of daily data from a dataset that combined ERA-40 (Uppala et al. 2005) with operational ECMWF analyses; they found that the region of preferred occurrence of RWPs as well as their properties have a strong seasonal dependence. In addition, they noted preferred RWP genesis over the western and central Pacific and the western Atlantic as well as preferred RWP decay in the eastern parts of the Pacific and Atlantic Oceans.

   The same dataset was used in another study to compute the frequency distributions of the lifetime as well as the preferred location of generation and decay of RWP objects (Glatt and Wirth 2014). Preferred regions for generation and decay show substantial differences between the seasons. Particularly interesting is the situation for spring (Fig. 11). For the shorter-lived RWPs, there are two distinct and about equally important regions with generation of RWPs corresponding to the western side of the Pacific and the Atlantic Ocean, respectively. On the other hand, for the longer-lived RWPs, the peak over eastern North America is much smaller than the peak over the western Pacific, suggesting that the long-lived RWPs are those that are able to propagate across the North American continent and continue their trajectory across the Atlantic Ocean. The Asian continent, by contrast, seems to be a major obstacle for any RWP propagation, which is consistent with the finding that Europe is the dominant region for RWP decay.

   A climatology that is based on RWP tracking in both longitude and latitude was first provided by Souders et al. (2014b), using the algorithm of Souders et al. (2014a). In their analysis, Souders et al. (2014b) considered the frequency of occurrence, generation, and decay of RWPs, as well as their group velocity and seasonal behavior. Figure 12 provides an example,
showing a frequent occurrence of RWPs along the well-known storm tracks (Chang et al. 2002), as well as a preference for RWP onset over the western parts of the ocean basins in the Northern Hemisphere. The figure indicates that cyclogenesis and the onset of RWPs are closely connected and should be considered as the lower- and upper-tropospheric signatures, respectively, of three-dimensional flow structures that span the entire depth of the troposphere. The object-based RWP climatologies add value by focusing on the upper troposphere and quantifying where, on average, the RWPs come into existence, how long they live, how far they propagate, and where they decay (see Fig. 11). The interannual variability was shown to be strongly associated with large-scale flow regimes such as ENSO and the Arctic Oscillation (Souders et al. 2014b).

Although there is a fair amount of agreement between the RWP object climatologies derived from different methods, there are also discrepancies regarding specific aspects (Souders et al. 2014b). For instance, there is a large difference in the frequency of occurrence of long-lived RWPs, with the algorithm of Souders et al. (2014a) yielding much higher values than that of Glatt and Wirth (2014). In addition, there are differences in the formation regions, which are presumably due to different definitions of the location of RWP onset (Souders et al. 2014b). Such remaining differences appear unfortunate, but they cannot be avoided because of the lack of a unique definition of RWP objects.

c. Climatological impact of the extratropical transition of tropical cyclones

There is a statistically significant increase of RWP amplitude and occurrence frequency (cf. climatology) in the North Pacific and in the Indian Ocean after the recurvature of a tropical cyclone (Archambault et al. 2013, 2015; Torn and Hakim 2015; Quinting and Jones 2016). In the North Pacific, this amplification may last for 4–10 days (Archambault et al. 2013, 2015) with a relative increase in amplitude of about 30% (Quinting and Jones 2016), and RWPs occur 10%–30% more frequently (Quinting and Jones 2016) and exhibit larger wavelengths and a greater downstream extent (Torn and Hakim 2015). August–November is particularly favorable and December and June particularly unfavorable for RWP amplification by recurving tropical cyclones in the North Pacific (Archambault et al. 2013). By contrast, no such modifications are found in the North Atlantic (Torn and Hakim 2015; Quinting and Jones 2016). The reasons for this geographical difference are not yet understood and constitute an interesting current research question.

As a common characteristic, extratropical transition in the North Atlantic and the North Pacific is, on average, associated with the quasi-stationary amplification of the downstream ridge–trough couplet (Torn and Hakim 2015). This result is consistent with the reduction of the wave’s phase propagation by the tropical cyclone’s outflow (section 4b). In the North Pacific, however, it is not clear whether the ridge–trough amplification should be considered as the amplification of a preexisting upstream RWP (Archambault et al. 2015) or as the generation of a new RWP (Torn and Hakim 2015). This difference in the interpretation is likely due to differences in the details of how these studies have created the respective composites in which RWPs are identified. Unambiguously, the development of midlatitude cyclones during fall and winter involves preexisting RWPs that amplify during cyclone development.

From a climatological perspective, the amplification of downstream RWPs in the North Pacific is sensitive to the strength of the interaction of the recurving tropical cyclone with the midlatitude flow (Archambault et al. 2013, 2015), as well as to the existence of an upstream trough (Quinting and Jones 2016; Torn and Hakim 2015). Arguably, both sensitivities are two sides of the same coin, signifying a “synergistic interaction” between the tropical cyclone and the trough (Quinting and Jones 2016). Our interpretation of this synergistic interaction is that strong interaction implies phase locking between the recurving tropical cyclone and the midlatitude wave pattern, and that the existence of an upstream trough promotes a phase-locked configuration (section 4b). In contrast, the amplification of RWPs in a climatological sense is relatively insensitive to the intensity or size of the recurving
tropical cyclone, and whether or not it reintensifies as an extratropical cyclone after completing extratropical transition (Archambault et al. 2013).

6. RWPs and weather forecasting

As mentioned in the introduction, a major incentive to investigate RWPs stems from their reputed role for weather forecasting and, in particular, their association with extreme weather. These two aspects will be discussed in the current section (weather forecasting) and the subsequent section 7 (extreme weather), respectively.

a. How well are RWPs and the waveguide predicted?

The first and rather obvious question is to what extent numerical weather prediction (NWP) models are able to represent and predict the RWPs as well as the waveguide that they propagate on. A particular case of an RWP was examined by Glatt and Wirth (2014), indicating that the RWP was well captured only in those forecasts for which the wave packet was already in existence at the time of initialization; moreover, the forecast became rather poor after a lead time of as little as 5 days. The latter result seems to be at odds with the commonly held view that large-scale phenomena such as RWPs should be predictable on a rather long time scale. However, this evaluation was for a single case only involving a single forecast model; further systematic studies are required to possibly generalize these results.

Regarding the waveguide, several operational forecast models are fraught with a spurious decrease of the PV gradient next to the tropopause (which is a key property of the waveguide), as well as a northerly bias of the waveguide position (Gray et al. 2014; Giannakaki and Martius 2016). The misrepresentation of the sharpness of the waveguide can lead to errors in the jet speed and the phase speed of the waves along the waveguide. Although there is a partial cancellation of errors, they still accumulate and may become noticeable in forecasts exceeding a lead time of 4–5 days (Harvey et al. 2016).

Similar to the sharpness of the waveguide, the amplitude of upper-tropospheric ridges is systematically underrepresented in operational global forecast models (Gray et al. 2014). This underrepresentation is consistent with an underrepresentation of latent heat release and the associated transport of air from the lower to the upper troposphere, and thus upper-tropospheric divergence (Gray et al. 2014; Teubler and Riemer 2016). The misrepresentation of latent heat release in the warm conveyor belts of midlatitude cyclones has been argued to lead to the observed systematic error in RWP amplitude (Martínez-Alvarado et al. 2016). To date, however, it is not clear whether errors in warm conveyor
belts are indeed systematic (Madonna et al. 2015). Whether systematic or not, errors caused by latent heat release may propagate and grow within an RWP (discussed in more detail below), eventually leading to errors that affect the whole RWP structure even in medium-range forecasts, thus leading to exceptionally large forecast errors (“forecast busts,” Rodwell et al. 2013).

b. Zonal propagation of forecast errors and uncertainties

There are indications that analysis differences (Hollingsworth et al. 1985), the impact of observations (Barwell and Lorenc 1985), and growing errors (Langland et al. 2002) all propagate in the zonal direction at speeds faster than those of individual troughs and ridges. This was clearly illustrated by analyses during the National Oceanic and Atmospheric Administration (NOAA) Winter Storm Reconnaissance Programs (Szunyogh et al. 2000, 2002); the impact of targeted observations spreads in a manner resembling the zonal propagation of a growing RWP during its linear phase.

Consistent with these earlier results, the vertical structure and evolution of forecast errors resemble the propagation and dispersion of a linear RWP (Hakim 2005). Similarly, uncertainties tend to spread at a speed similar to the group velocity of RWPs, often accompanying the growth of an RWP (Anwender et al. 2008; Sellwood et al. 2008; Zheng et al. 2013). Results from these studies form the basis of the application of several tools that have originally been developed to diagnose short-range forecast uncertainties—namely the ensemble transform Kalman filter (Bishop et al. 2001) and ensemble sensitivity analysis (Torn and Hakim 2008)—to the medium range at a longer time scale than the development of individual synoptic trough–ridge systems (Sellwood et al. 2008; Majumdar et al. 2010; Chang et al. 2013; Zheng et al. 2013).

c. Upscale error growth affecting RWPs

The predictability of RWPs depends on the upscale growth of errors and uncertainties from the convective scale to the synoptic scale (e.g., Zhang et al. 2003). Figure 13 illustrates such upscale error growth from the PV perspective, which emphasizes the tropopause waveguide and RWPs as focal points. At forecast day 2, PV errors exhibit localized, mesoscale maxima near the midlatitude waveguide (Fig. 13a). At forecast day 6, there are large errors that affect the two RWPs present in this forecast (Fig. 13b).

PV forecast errors in general exhibit their largest values in the vicinity of the midlatitude waveguide (i.e., within RWPs) (Dirren et al. 2003; Davies and Didone 2013; Baumgart et al. 2018). This observation casts some further doubt on the general view that RWPs as large-scale flow features exhibit and provide enhanced predictability (see section 6d). There is general agreement that the fastest error growth is associated with latent heat release and occurs on the convective scale. Such errors then affect the mesoscale by adjustment to balance (Zhang et al. 2007; Bierdel et al. 2017) and by displacement of PV gradients by the divergent flow (Baumgart et al. 2018). Standard models for error growth (Hohenegger and Schär 2007; Zhang et al. 2007) assume that baroclinic instability plays the leading role for the amplification and upscale growth of forecast errors from the mesoscale up to the scale of RWPs. However, there may be other growth mechanisms that are unrelated to baroclinic instability (e.g., Snyder 1999; Davies and Didone 2013; Harvey et al. 2016). In particular, track bifurcation of tropical cyclones during extratropical transition is such a mechanism (e.g., Riemer and Jones 2014, see section 6e).

To quantify the different contributions to the amplification of forecast errors, PV-error tendency equations have been derived and evaluated (Davies and Didone 2013; Baumgart et al. 2018). Using the PV diagnostic of Teubler and Riemer (2016, here summarized in section 3f), the different contributions to error growth in the case illustrated in Fig. 13 were quantified by Baumgart et al. (2018). Interestingly, differences in the interaction between upper- and lower-tropospheric PV anomalies, which signifies baroclinic growth in the PV framework, contribute less than 15% to error growth in this case. Differences in the interaction of upper-tropospheric PV anomalies, which represent near-tropopause dynamics, make the most important contribution to the growth of those errors that are later associated with RWPs. Baumgart et al. (2018) demonstrated that in their case the error growth is dominated by nonlinearities in the Rossby wave dynamics, complementing essentially linear mechanisms previously proposed by Snyder (1999) and Harvey et al. (2016).

d. Potential for extended-range predictability from RWPs

Another important question is whether and to what extent the presence of an RWP has an impact on predictability. The relevance of downstream development for weather forecasting was presumably recognized as soon as it was detected in observations (Namias and Clapp 1944; Cressman 1948; Hovmöller 1949), as well as in model simulations (Simmons and Hoskins 1979). Indeed, as discussed earlier, Chang (2005a) and Wirth and Eichhorn (2014) suggested that the development of deep surface cyclones over the western Pacific and over Europe is frequently preceded by upstream RWPs. The presence of upstream wave packets significantly increases...
the probability of cyclone development over these regions, whereas the absence of upstream wave packets decreases the probability for cyclone development.

The potential benefit from the existence of an RWP for weather forecasts hinges on the idea that the larger spatial scale and longer lifetime of RWP envelopes compared with individual troughs and ridges imply a higher degree of predictability (Lee and Held 1993). Whether this is indeed the case is still uncertain, as only recently have there been efforts to evaluate how well RWPs are predicted by NWP models (Glatt and Wirth 2014). More systematic studies are needed to clarify the issue.

Nevertheless, several studies have provided indications that atmospheric predictability may be enhanced by the presence of RWPs. For instance, large-scale flow patterns that are conducive to heavy precipitation south of the Alps, which are characterized by upstream RWPs (Martius et al. 2008), are associated with a higher-than-average predictive skill of the European Centre for Medium-Range Weather Forecasts (ECMWF) global forecasting system (Grazzini 2007). More recently, ensemble analyses and sensitivity experiments have demonstrated that the increased forecast uncertainty originates from the interaction of the recurving tropical cyclone with the midlatitude jet. From there, a “cone” of increased uncertainty evolves in space and time, approximately defined by the phase and group speed of the associated midlatitude RWP (as depicted in Hovmöller diagrams; Harr et al. 2008; Anwender et al. 2008; Riemer and Jones 2010; Grams et al. 2013b; Pantillon et al. 2013a,b; Aiyyer 2015; Quinting and Jones 2016; Torn 2017). This notion is similar to the idea that forecast errors themselves propagate and grow like linear RWPs (Hakim 2005).

e. Forecast uncertainty associated with the extratropical transition of tropical cyclones

It has been forecasters’ wisdom for a long time that the ET of a tropical cyclone constitutes a distinct source of forecast uncertainty in the midlatitudes (Jones et al. 2003b). More recently, ensemble analyses and sensitivity experiments have demonstrated that the increased forecast uncertainty originates from the interaction of the recurving tropical cyclone with the midlatitude jet. From there, a “cone” of increased uncertainty evolves in space and time, approximately defined by the phase and group speed of the associated midlatitude RWP (as depicted in Hovmöller diagrams; Harr et al. 2008; Anwender et al. 2008; Riemer and Jones 2010; Grams et al. 2013b; Pantillon et al. 2013a,b; Aiyyer 2015; Quinting and Jones 2016; Torn 2017). This notion is similar to the idea that forecast errors themselves propagate and grow like linear RWPs (Hakim 2005).
Thereby, forecast uncertainty can increase over a large geographical region. The consensus emerging from the above studies is that the increased forecast uncertainty in midlatitudes is due to the projection of the uncertainty in the evolution of the ET system onto the subsequent evolution of the RWP.

The evolution of the ET system itself is very sensitive to the phasing with respect to the midlatitude wave pattern (Hanley et al. 2001; Ritchie and Elsberry 2007). Case studies and idealized modeling demonstrate that this large sensitivity translates into uncertainties of the evolution of the downstream flow (Riemer and Jones 2010, 2014; Grams et al. 2013b, 2015). Singular vector (Reynolds et al. 2009) and ensemble (Anwender et al. 2010; Keller et al. 2011) analyses corroborate this notion. Figure 14 provides an illustration of this sensitivity to phasing from a set of idealized numerical experiments (Riemer and Jones 2010). These experiments consider the development of baroclinic waves from a localized initial perturbation in a channel configuration. A reference experiment without a tropical cyclone is compared to three experiments, in which a tropical cyclone interacts with the developing wave. These three experiments differ only in the phasing of the tropical cyclone and the wave pattern. A large impact of the tropical cyclone on the developing RWP, as compared to the reference experiment, and a large sensitivity of this impact with respect to phasing, can be inferred in Fig. 14 from the differences in the undulation of the dynamical tropopause.

One prominent source of this large sensitivity is the existence of bifurcation points in the storm’s steering flow during the onset of ET, implying that small uncertainties in the track of the tropical cyclone may quickly amplify (Scheck et al. 2011b; Grams et al. 2013b; Riemer and Jones 2014). While the existence of bifurcation points is a generic feature in idealized Rossby wave scenarios, it is yet unclear what percentage of real-atmospheric ET cases is indeed affected by this bifurcation-like behavior. In addition, several studies have proposed that the complex, multiscale processes that govern the structure and intensity evolution of the ET system and its interaction with the upstream trough are a further prominent source of uncertainty (Jones et al. 2003b; Davis et al. 2008; Riemer et al. 2008; Pantillon et al. 2013a). To date, this source of uncertainty has not yet been explicitly investigated, and its relative role in generating large forecast uncertainty of the midlatitude downstream flow is only poorly understood.

The above-quoted studies on predictability downstream of ET indicate a considerable case-to-case variability. There are a few, partly speculative, attempts to explain this variability, but the underlying reasons have not been clarified satisfactorily. The extent to which downstream forecast uncertainty increases has been related to the extent of RWP amplification (Quinting and Jones 2016), to whether a preexisting RWP is amplified or a new one is generated (Torn and Hakim 2015), to the strength of the upper-tropospheric tropical cyclone–jet interaction (Grams et al. 2015), to uncertainties in the moisture distribution (Torn and Hakim 2015), to the general sensitivity of the midlatitude flow (Grams et al. 2015), and to the amplitude of the downstream ridge prior to ET (Torn 2017).

7. RWPs as precursors to extreme weather

Recently, the connection between RWPs and extreme weather has found increased interest in the scientific literature. These studies were motivated by the societal importance of extreme weather and the sometimes poor performance of forecasting these weather events on the medium range (e.g., Grazzini and van der Grijn 2002; Shapiro and Thorpe 2004). In many of the investigated cases, the extreme weather was associated with a meridionally elongated upper-tropospheric trough (i.e., a breaking Rossby wave), which has been known to be conducive to extreme weather for a long time (e.g., Bosart et al. 1996; Massacand et al. 1998; Roebber et al. 2002; Galanneau et al. 2012; Parker et al. 2014; Bosart et al. 2017). Interesting in the context of this paper is the fact that such an elongated trough is sometimes part of an RWP during its decay stage. This means that the corresponding RWP can be seen as a long-range precursor to the local extreme weather event.

A link between precursor RWPs and extreme weather has been documented for numerous cases with a variety of weather events: for strong surface cyclones (Chang 2005a; Wirth and Eichhorn 2014); for extreme temperatures (Marenco et al. 2002; Sprenger et al. 2013; Chen et al. 2015; Jacques-Coper et al. 2016; Bosart et al. 2017; Fragkoulidis et al. 2018); for floods and precipitation in England (Krishnamurti et al. 2003; Davies 2015), Switzerland (Martius et al. 2008; Piaget et al. 2015; Barton et al. 2016), Canada (Milrad et al. 2015), Antarctica (Welker et al. 2014), and Saudi Arabia (de Vries et al. 2016); for extreme winds in North Africa (Wiegand et al. 2011); and for a combination of extremes (Bosart et al. 2017). These studies contain examples of weather extremes from both hemispheres located in the subtropics, extratropics, and the higher latitudes. Long-lasting RWPs preceding extreme weather exist both on the extratropical and the subtropical waveguide (Feldstein and Dayan 2008; Li et al. 2015; de Vries et al. 2016).

However, not each extreme weather event is by necessity associated with a precursor RWP. Figure 15 provides an example, showing the meridional wind in the upper troposphere during an episode with a series of five precipitation events in southern Switzerland. Three of these five precipitation events coincided with the dissipation of
an RWP that previously had formed over the Pacific Ocean and propagated across the Atlantic Ocean toward Europe. In the figure, these precursor RWPs can be detected through the characteristic patterns on the Hovmöller diagram (cf. with the similar patterns in Fig. 5). The event on 27 September was associated with a cutoff low that had formed from the upper-level PV trough located over Europe on 23 September. The last event on 12 October was associated with an upper-level trough that had formed downstream of a strong blocking ridge over the Atlantic Ocean. Hence, while the presence of a deep trough over southwestern Europe was found for all five precipitation events, only some, but not all, of these troughs were associated with a long-lived precursor RWP.

A key question, of course, is whether or not the presence of a precursor RWP helps to improve the prediction of the extreme weather event. The answer to this question hinges on two presumptions: (i) that the occurrence of RWPs results in enhanced predictability in general, and (ii) that there is a systematic connection between RWPs and weather extremes. The first of these presumptions was discussed in section 6, where we showed that there is no clear answer yet. The second of these presumptions has been the topic of the current section, where we saw that there is a connection in numerous cases, but it is not yet clear to what extent this connection is a systematic one. It transpires that the question whether or not precursor RWPs help to predict extreme weather is still open.

8. Summary, caveats, and future directions

Rossby wave packets (RWPs) on the midlatitude Rossby waveguide have been studied ever since their first description in the 1940s by C. G. Rossby and collaborators. Some of the key properties (such as downstream development, section 1b) have been known and at least partly understood right from the beginning. More systematic investigations followed since the late 1970s, using baroclinic models in judicious configurations as well as detailed regression and composite analyses based on observed data, drawing heavily on the concepts of eddy kinetic energy and related fluxes (section 3d). The last 15 years have seen a renewed interest in RWPs, partly because of their putative role for forecast error propagation (section 6) and their connection with extreme weather (section 7). Currently, RWPs constitute a vibrant research topic addressing relevant questions. In the remainder of this section we aim to critically summarize the current state of affairs and identify avenues for future research.

a. Recent developments

An important line of recent research concerned the development of algorithms to define RWP objects (section 3c). This object-oriented approach gives explicit credit to the idea that RWPs are meteorologically meaningful entities with specific properties and different stages of a life cycle (sections 4b and 4c). For the first time, RWPs have been tracked in space and time in an automated fashion (section 3c), and these new algorithms were used to extract climatological information (section 5b) and perform object-oriented verification in terms of RWPs. The results are broadly consistent with previous knowledge from the analysis of storm tracks (Chang et al. 2002), but the object-oriented approach provided additional information, like for instance statistics about the onset and decay of RWPs. Clearly, the use of these methods is at its early stage, and we believe that its potential has not been exploited to its full extent yet.

Another novel line of research was the systematic application of a PV framework to analyze the dynamics of RWPs (section 3f), which is complementary to the more established eddy kinetic energy framework. The PV framework is currently being used to study the evolution of forecast error formation, growth, and downstream propagation (section 4b). In particular, it lends itself to studying the role of diabatic processes, which are of leading-order importance for RWP amplification (section 4b). Wave amplification is local in the sense that it is typically associated with latent heat release in a specific cyclone, but the impact may then propagate downstream thus affecting the entire RWP.

There has also been an increased interest in the waveguide concept during the past few years (section 2), with studies to elucidate the role of specific waveguide properties for the propagation of RWPs. A special topic
of continued interest is the extratropical transition of tropical cyclones [sections 4b(3) and 5c], where diabatic processes strongly affect the interaction of the transitioning tropical cyclone with midlatitude RWPs, and where these RWPs can play a key role in transmitting errors downstream (section 6e). In addition, new studies suggested that the presence of RWPs may modify the forecast quality (section 6) and that, in some cases, RWPs occur as precursors to extreme weather (section 7).

b. Remaining issues and future directions

There are many outstanding issues and open questions. In particular, the role of RWPs as precursors to extreme weather and the implications for forecast quality has been addressed only for specific cases or flow configurations. More systematic studies seem necessary to generalize these results. For instance, are there specific conditions that make an RWP a precursor to extreme weather? Other questions concern the formation of errors in predicting RWP propagation and RWP modification, for example, under what conditions are RWPs particularly sensitive to perturbations? There are numerous open questions regarding the relative contribution of individual processes to error growth associated with RWPs, in particular regarding the relative role of barotropic, baroclinic, and moist processes. Also, the large case-to-case variability apparent in this context needs to be better understood.

While the dynamics at the leading edge of an RWP (viz., downstream baroclinic development) are relatively well understood, more work is required to fully understand the spatial coherence of RWPs and the processes at their trailing edge. It is desirable to learn more about the dependence of RWP characteristics on their generation mechanism, such as the excitation by extratropical transition, by other forms of organized convection, and by different types of cyclogenesis. Also, diabatic processes still pose a challenge, like for instance their climatological role for RWP generation and modification in comparison with dry dynamics. In addition, more research is needed toward an unambiguous and dynamically meaningful identification of the background waveguide and wave–mean flow interaction.

Issues remain in connection with the object-oriented methods for RWP identification. All these algorithms involve some user-defined input such as thresholds, filter details, or the choice of a background flow, and the results depend to a certain extent on these choices. The key question is how to fix or calibrate related parameters. A possible technique is to resort to a comparison with a conventional Hovmöller diagram (Grazzini and Lucarini 2011; Glatt and Wirth 2014). This technique means that effectively the calibration is done by different individuals, which represents an element of subjectivity. On the level of a case study, this approach appears well justified. However, the impact of such calibration is less obvious when the algorithm is applied in an automated fashion (i.e., blindly) to long time series of data in order to produce RWP climatologies. Souders et al. (2014a) tried to back up their calibration through the simultaneous use of energy fluxes, but even this method retains elements of subjectivity and calibration is always performed on a finite number of cases. Not surprisingly, systematic comparisons between the different methods indicate some differences, and in the end the definition of an RWP object remains somewhat elusive (Glatt et al. 2011; Souders et al. 2014b).

These caveats concerning RWP objects do not mean that these diagnostics are useless. One way to deal with this situation is to use multiple such diagnostics for the same analysis (Glatt et al. 2011). The latter approach was taken by Ahmadi-Givi et al. (2014) who investigated RWPs at the end of the North Atlantic storm track; through a combination of diagnostics, they found key differences in RWP behavior in two specific cases and
were able to explain differences in the subsequent evolution over the Mediterranean.

A substantial part of the diagnostic armory to study RWPs is based on linear wave theory. However, observed RWPs have often large amplitudes such that linear theory is not a good approximation and should not be applied in a quantitative manner. Recently, there have been efforts to define measures of zonal-mean finite-amplitude wave activity that lend themselves to practical implementations (Nakamura and Solomon 2010, 2011; Methven 2013). Some work is needed to make these concepts applicable to zonally confined RWPs, and such developments are just about to emerge (Huang and Nakamura 2016, 2017; Nakamura and Huang 2017). Future work will reveal to what extent these finite-amplitude concepts are superior to more conventional concepts from linear theory and what additional insight they are able to produce. An ultimate limitation with the concept of an RWP lies in the fact that it requires some underlying “waviness” as a particular property of the flow, and that the definition of a “perturbation” requires the definition of a corresponding “background state,” which is not unique (section 2). The future is going to reveal how far one can go with the concept of RWPs along the midlatitude waveguide in order to better understand the complex interplay between upper-tropospheric dynamics and other aspects of synoptic-scale meteorology.

Acknowledgments. We are grateful to the chief editor, D. M. Schultz, for his encouragement and expert advice throughout the process of writing this review. We sincerely thank Brian Hoskins as well as four anonymous reviewers for their constructive criticism, helping us to streamline the text and clarify several issues. We also thank J. Eichhorn for his help with some of the figures. V. Wirth and M. Riemer acknowledge the German Research Foundation (DFG) for funding through the research unit “Predictability and Dynamics of Weather Systems in the Atlantic–European Sector” (PANDOWAE) and the transregional collaborative research center SFB/TRR165 “Waves to Weather” (W2W). O. Martius acknowledges support by SNSF Grant 200021_137543. E. Chang acknowledges support by NOAA Grant NA13NWS468002.

REFERENCES


