Real-time identification of equatorial waves and evaluation of waves in global forecasts

Article

Accepted Version

Yang, G.-Y. ORCID: https://orcid.org/0000-0001-7450-3477, Ferrett, S. ORCID: https://orcid.org/0000-0003-4726-847X, Woolnough, S. ORCID: https://orcid.org/0000-0003-0500-8514, Methven, J. ORCID: https://orcid.org/0000-0002-7636-6872 and Holloway, C. ORCID: https://orcid.org/0000-0001-9903-8989 (2021) Real-time identification of equatorial waves and evaluation of waves in global forecasts. Weather and Forecasting, 36 (1). pp. 171-193. ISSN 0882-8156 doi: https://doi.org/10.1175/WAF-D-20-0144.1 Available at https://centaur.reading.ac.uk/94722/

To link to this article DOI: http://dx.doi.org/10.1175/WAF-D-20-0144.1

Publisher: American Meteorological Society

Real-time identification of equatorial waves and evaluation of waves in global forecasts

Gui-Ying Yang[1], Samantha Ferrett[1], Steve Woolnough[1],
John Methven[2], Chris Holloway[2],

[1]*National Centre of Atmospheric Science, and Department of Meteorology, University of Reading, Reading, UK*
[2]*Department of Meteorology, University of Reading, Reading, UK*

Submitted to Weather and Forecasting
Revised in November 2020

Corresponding Author: Gui-Ying Yang, g.y.yang@reading.ac.uk

ABSTRACT

A novel technique is developed to identify equatorial waves in analyses and forecasts. In a real-time operational context, it is not possible to apply a frequency filter based on a wide centred time-window due to the lack of future data. Therefore, equatorial wave identification is performed based primarily on spatial projection onto wave mode horizontal structures. Spatial projection alone cannot distinguish eastward from westward-moving waves, so a broad-band frequency filter is also applied. The novelty in the real-time technique is to off-centre the time-window needed for frequency filtering, using forecasts to extend the window beyond the current analysis. The quality of this equatorial wave diagnosis is evaluated. Firstly, the "edge effect" arising because the analysis is near the end of the filter time-window is assessed. Secondly, the impact of using forecasts to extend the window beyond the current date is quantified. Both impacts are shown to be small referenced to wave diagnosis based on a centred time-window of re-analysis data. The technique is used to evaluate the skill of the Met Office forecast system in 2015-2018. Global forecasts exhibit substantial skill (correlation > 0.6) in equatorial waves, to at least day 4 for Kelvin waves and day 6 for Westward Mixed Rossby-Gravity (WMRG), and meridional mode number n=1 and n=2 Rossby waves. A local wave phase diagram is introduced that is useful to visualise and validate wave forecasts. It shows that in the model Kelvin waves systematically propagate too fast and there is a 25% underestimate of amplitude in Kelvin and WMRG waves over the Central Pacific.

1. Introduction

Equatorial waves are fundamental components of the tropical atmosphere and are important for understanding its behaviour. A number of observational studies have shown that active deep convection and the location of convective systems are frequently observed to be associated with equatorial wave modes (e.g., Gruber 1974, Zangvil 1975; Zangvil and Yanai 1980, 1981; Liebmann and Hendon 1990; Hendon and Liebmann 19911 Takayabu and Nitta 1993; Takayabu 1994a, b; Redelsperger et al. 1998; Wheeler and Kiladis 1999; Wheeler et al. 2000; Straub and Kiladis 2002; Roundy and Frank 2004; Yang et al. 2007a, b, c; Roundy 2008; Kiladis et al. 2009). Understanding equatorial waves and their connection with tropical convective activity is important for the improvement of weather forecasting in the Tropics on time-scales beyond a few days, and is also likely to be crucial for climate prediction (e.g. Lin et al. 2006; Ringer et al.2006; Yang et al. 2009). However, global models used for numerical weather and climate prediction have difficulty in representing equatorially trapped waves with errors in phase speed, amplitude and structure (e.g. Slingo et al, 2003; Yang et al, 2009; Straub et al., 2010; Huang et al. 2013). These problems limit the ability to predict tropical wave activity and hence any modulation of high impact weather associated with them. Recently Ferrett et al. (2020) have shown that increases in the amount of precipitation and the likelihood of extreme precipitation are linked to Kelvin, westward-moving mixed Rossby-gravity (WMRG) and meridional wave number n=1 Rossby (R1) waves. Heavy precipitation can be up to four times more likely to occur during a period with high amplitude waves in Southeast Asia, indicating that the probability of extreme precipitation is highly dependent on equatorial wave activity. This suggests that these equatorial waves provide a potentially important source of predictability for tropical precipitation and high impact weather (HIW). Therefore, it is crucial to evaluate and improve model ability in representing and forecasting equatorial wave modes and their associated precipitation.

The importance of convectively coupled equatorial waves has recently drawn extensive attention and motivated a number of studies on the equatorial wave forecasting in operational models (e.g., Dias et al 2018; Janiga et al. 2018; Bengtsson et al 2019), and on equatorial wave predictability (e.g.,Ying and Zhang 2017; Judt 2020; Li and Stechmann 2020). However, since most of these studies focus on OLR or precipitation signals, they are not able to characterize the relationship between convection and the wind structures within waves (Yang et al. 2007a, b). The coherent circulation structures associated with equatorial waves are a are a major organising influence on tropical precipitation (Ferrett et al. 2020) and a potentially important source of predictability. Such analysis requires a methodology for identifying equatorial waves in analyses and forecasts in real-time.

Following the discovery of equatorial waves in the equatorial stratosphere (Yanai and Maruyama 1966; Wallace and Kousky 1968), the subsequent two decades of observational studies of equatorial waves in 1970-1990s mainly used time and/or space power spectral analysis to diagnose equatorial wave modes (e.g., Gruber 1974; Zangvil 1975; Zangvil and Yanai 1980, 1981; Liebmann and Hendon 1990; Hendon and Liebmann 1991; Takayabu 1994a, b; Magaña and Yanai 1995; Pires et al. 1997).

Since the late 1990s, there have been two main methods of identifying equatorial waves in observational data. In the first, following Wheeler and Kiladis (1999) which built on Takayabu (1994a, b), equatorial waves are isolated based on the Fourier transform of observed data (usually OLR) into zonal wavenumber and frequency space, isolating sectors of phase space defined about the dispersion curves from equatorial wave theory (on a resting basic state) and then transforming the data back to physical space (longitude-time) from the Fourier coefficients only within each sector.

The second method, following Yang et al. (2003), is to project global wind and geopotential height data onto an orthogonal basis defined by the horizontal equatorial wave structures obtained from the theory of disturbances to a resting atmosphere on the equatorial β-plane (Matsuno 1966). These structures are defined by sinusoidal waves in the zonal direction and parabolic cylinder functions in the meridional direction. The method does not assume that the dispersion relation and vertical structure from this theory apply to the real situation where these aspects would be sensitive to any background zonal flow that varies with height and time. Moreover, there is not a complete theory for equatorial waves in shear parallel flow, so the untilted modes on a resting atmosphere are used as basis structures, but are not expected to be exactly the same as normal modes of the real flow.

In addition to the two main methodologies mentioned above, there are some more recent techniques used for isolation of equatorial waves, such as those associated with spatial projection or 3-D normal mode projection (e.g, Gehne and Kleeman 2012; Žagar 2009 and Žagar et al et al. 2016; Castanheira and Marques 2015 and Marques and Castanheira2018), extended EOF projection (e.g., Roundy 2012) and wavelet-based filtering (e.g, Kikuchi 2014; Kikuchi et al. 2018).

It is possible to apply the spatial projection method of Yang et al (2003) without any time filter on the data. However, because the projection is done independently on three pressure-level variables, obtained from combination of the horizontal velocity components and geopotential height, some structure functions are not unique to one wave mode, appearing in both an eastward and westward mode from the theory. Also, it is found that there is often a strong projection of stationary features in the fields onto the structure functions. So better results have been obtained in examination of re-analysis and climate model data by applying a broad frequency filter using a wide time-window (e.g., Yang et al 2009, 2012). The purpose of this filter is to cut out the stationary features and to distinguish eastward and westward moving disturbances. The combination of spatial projection with this broad frequency filter does yield a unique attribution between projected fields and wave modes.

The challenge is that wave mode identification methods based on frequency filtering, combined with wavenumber filtering, require a wide time-window of data (usually much longer than 30 days to distinguish slower disturbances) but in a real-time operational context only forecast data is available beyond the current analysis and forecast data quality degrades quickly over the first week. The purpose of this research is to obtain an accurate method of equatorial wave identification that works for the current analysis and forecast data.

Wheeler and Weickmann (2001) adapted the wavenumber-frequency filtering method for real-time monitoring of equatorial waves by using the same time-window approach for the frequency filter, but filling unknown future values, at time-points in the window beyond the current analysis, with zeros (padding). The resulting anomalies can be used for monitoring wave modes up to the current day and can provide a “statistical” prediction of the evolution of these modes several days into the future. This methodology shows some predictive skill for the MJO and various equatorial waves, but the “padding method” results in rapid decay of amplitude near the end of the record and into the forecast. Patching real-time forecasts to analysis is also used by Gottschalk et al. (2010) for prediction of MJO in real-time forecast. Recently Carl Schreck extended the wavenumber-frequency filtering method for real-time applications (https://ncics.org/portfolio/monitor/mjo/) to include a rescaling of total variance to maintain the amplitude, or inclusion of 45-day OLR forecasts from the subseasonal-to-seasonal NCEP Climate Forecast System (CFS). These wavenumber-frequency methods show some skill in predicting tropical synoptic convective activity related to the preferred equatorially trapped modes. However, the methodology has two potential limitations. First, the pre-specified segments of wave-number frequency space used to partition “wave-modes” can be susceptible to errors induced by changes in wave frequency due to Doppler shifting by the background flow or effects of shear; or due to time-window edge effects introduced in the real-time filter methodology using padding. Secondly, the reliance on identifying the OLR signal can lead to the failure to identify equatorial modes in regions which may not be convectively active, and because they are identified from an OLR signal they cannot easily be used to relate the precipitation signal to the wave structure independently.

It is evident that a new methodology for real-time identification of equatorial wave modes that does not depend strongly on a time-window and frequency filtering approach could be very beneficial to identify equatorial waves and their associated tropical precipitation and HIW. The main aim of this study is to extend the methodology of Yang et al. (2003) by adapting the time-windowing approach for the broad frequency filter to provide a novel real-time technique for identification of equatorial waves in current analyses and forecasts. The methodology is used to identify horizontal winds, geopotential height and hence divergence and vorticity structures associated with distinct equatorial wave types.

This paper is organized as follows. Section 2 details the data used, briefly introduces the equatorial wave theory that is the basis for the diagnostic technique, and describes the spatial projection methodology used to identify equatorial waves. Section 3 presents the new time-window technique, combined with the spatial projection method, to identify equatorial waves in real-time and operational forecasts, and the evaluation of the methodology in terms of wave amplitude in horizontal wind using a 4-year UK Met Office operational global forecast dataset. A case study is given in Section 4 illustrating the identification of equatorial waves in real-time applications. Section 5 presents an evaluation of the skill of the forecast model in predicting wave behaviours, especially the phase and amplitude. Conclusions are made in Section 6

2. Equatorial wave theory and spatial projection of data onto a wave structure basis

a. Dataset

Operational global 6-hourly analysis and forecast data from the UK Met Office are used from the years 2015-2018. These forecasts use the Unified Model Global Atmosphere GA6.0 configuration (Walters et al. 2017) which was implemented operationally during 2014. The GA6.0 configuration includes the ENDGame dynamical core (Wood et al. 2014) which in climate simulations was shown to lead to a significant improvement in the representation of equatorial Kelvin waves (Walters et al. 2017). Whilst the atmospheric model version is consistent during this period there was a change from N768 (~17km) resolution to N1280 (~10km) resolution in mid-2017; an upgrade to the land-surface model in late 2018; and a number of changes to both the data-assimilation system and assimilated observations during the period. The horizontal wind components and geopotential height data are re-gridded onto a regular 1° × 1° degree grid before being projected onto equatorial wave structure functions. As a proxy for convection, use is made of NOAA interpolated daily Outgoing Longwave Radiation (OLR) on a 2.5° ×2.5° grid (Liebmann and Smith 1996).

b. Basic equatorial wave theory and methodology to identify equatorial waves

Equatorially trapped waves are obtained as solutions to the adiabatic, frictionless equations of motion on an equatorial β-plane, linearized about a state of rest. The solutions are separable in terms of vertical and horizontal structure functions (Matsuno 1966; Gill 1980). The horizontal and temporal behaviours of horizontal winds (u, v) and geopotential height (Z) are described by the linearized shallow water equations with gravity wave speed c_e, the separation constant from the vertical structure equation that must also satisfy relevant surface and upper boundary conditions. This is possible only for discrete values of the separation constant, c_e.. In an atmosphere with a constant buoyancy frequency with a rigid lid upper boundary condition, the vertical modes are sinusoidal in height, with corrections for the density variation.

For the horizontal equations, u, v and Z fields are taken to be of the form:

$$\{u, v, Z\}=\{U(y), V(y), Z(y)\} \exp[i(kx- \omega t)] \qquad (1)$$

where k is the zonal wavenumber and ω is the frequency. As in Gill (1980) the equatorial wave solutions are most easily formulated in terms of new variables, q, r and v where:

$$q = u + gZ/c_e, \; r = u - gZ/c_e \;, \qquad (2)$$

and the structures of equatorial waves in the meridional coordinate, y, can be described by parabolic cylinder functions:

$$D_n\left(\frac{y}{y_0}\right) = \exp\left[-\frac{1}{4}\left(\frac{y}{y_0}\right)^2\right] P_n\left(\frac{y}{\sqrt{2}y_0}\right) \;, \qquad (3)$$

where $$y_0 = \left(\frac{c_e}{2\beta}\right)^{1/2} \qquad (4)$$

is the meridional scale and P_n is proportional to a Hermite polynomial of order n. In this analysis y_0 =6^{o} is used, which is deduced from a best fit to data in observations, and the corresponding c_e is about 20 m s^{-1}.Three variables, $\{q, v, r\}$, can be projected onto the parabolic cylinder functions:

$$\{q, v, r\} = \sum_{n=0}^{n=\infty}\{q_n, v_n, r_n\} D_n. \qquad (5)$$

These functions form a complete and orthogonal basis and the projections in Eq. (5) are quite general. q_0D_0 describing the Kelvin wave, q_1D_1 and v_0D_0 describe n=0 mixed Rossby-gravity (MRG) wave which has both eastward (EMRG) and westward-moving (WMRG) solutions. $q_{n+1}D_{n+1}$, v_nD_n and $r_{n-1}D_{n-1}$ describe $n\geq1$ equatorial low frequency westward-moving equatorial Rossby waves, and both eastward and westward-moving high frequency gravity waves.

The theoretical horizontal structures of some of the gravest (lowest meridional wave number) equatorial waves are shown in Fig.1. The Kelvin wave is dominated by divergent zonal wind, greatest along the equator, and has zero meridional wind. The *n*=1 and 2 Rossby (R1 and R2) waves, are dominated by rotational flows, strongest off the equator. The westward mixed Rossby gravity (WMRG) wave has mixed rotational and divergent flow and has a dominant signature in meridional wind across the equator. If the low-level convergence provides the organization for convection, then we would expect this convection to occur in the blue shaded regions. If the low level cyclonic circulation is important for convection, then we would expect this convection to occur in the blue contour line regions, especially for R1 and R2. This relationship has been revealed in observational studies (e.g. Yang et al. 2007a, b and Ferrett et al. 2020).The key points of the analysis method developed in Yang et al. (2003) are summarised as follows:

1) FILTER DATA SPECTRALLY IN A BROAD WAVENUMBER AND FREQUENCY DOMAIN

Separate the equatorial wave solutions *v*, *q* and *r* (Eq.2) in the tropical belt (24^{o}N and 24^{o}S) into eastward and westward-moving components using a space-time spectral analysis which transforms data from the *x-t* domain into the *k-ω* domain by performing 2-D FFT in the zonal and time direction (Hayashi 1982). The data are filtered using a broad-band spectral domain with *k*=2-40 and period of 2-30 days which includes all equatorial waves except high frequency gravity waves. For analysis of historical data a taper is applied to the two ends of the time series. However, this taper is not applied in the real-time technique as the data of primary interest is at the end of the time record.

2) PROJECT FILTERED COMPONENTS ONTO THE HORIZONTAL STRUCTURES OF EQUATORIAL WAVES

On each pressure level, the Fourier coefficients (e.g., V(y) for each k and ω) of eastward or westward-moving v, q and r are separately projected onto the meridional structures of the equatorial waves as described below Eq. (5) to obtain the equatorial wave modes.

3) TRANSFORM THE FOURIER COEFFICIENTS FOR EACH WAVE MODE BACK INTO PHYSICAL SPACE

The projected v, q and r Fourier components for each wave mode are transformed back into physical space, and then u and Z are deduced for each wave mode from the projected q and r using Eq. (2).

3. Methodology for real-time identification of equatorial waves

As described in the Introduction, the key challenge is to develop a method which enables real-time identification of equatorial waves in current analyses and forecasts, using an off-centred time-window of data for frequency filtering without a strong dependence on future data. The sensitivity of the method is explored by using different sets of data in the “future window” beyond the current analysis time.

a. Real-time approach to frequency filter

Since the Met Office operational global NWP forecasts extended to a 7-day lead time over the years 2015-2018, for input to the real-time frequency filter a 90-day time-series is constructed from 83 days of analysis data and 7 days of global forecast data (see Fig.2a). A 90-day time window is chosen as it is three times the longest period (30 days) in the frequency filter for equatorial waves and it also corresponds to one season. We also explored the sensitivity to using 120,180, 360 day windows and found very little difference.

To evaluate the real-time analysis and forecast methodology we create several wave datasets each dealing differently with the data beyond the “current analysis” (T+0). To mimic the real-time methodology, all aspects of the method are the same, including the time-window length of 90 days, and only the data beyond T+0 differs (see Fig. 2). We label the days prior to the current analysis as T-1 day, T-2 days, … and days following the current analysis as T+1, T+2, …., T+7 days.

1) REAL-TIME ANALYSIS AND FORECAST WAVE DATASET

The REAL-TIME wave dataset is created using a sliding 90-day window. At each verification date, the current analysis is defined as T+0 and 83 days of analysis data before this date are concatenated with 7 days of forecast data initialised on this date. The frequency and zonal wavenumber filters are applied to the global data in this time window and the resulting filtered data is then projected onto the equatorial wave basis structures at the time of current analysis, but also at T-7, T-6 through T+0 to T+6 and T+7. The T+0 result is called the REAL-TIME ANALYSIS and for T > 0 the REAL-TIME FORECAST.

2) DIAGNOSTIC ANALYSISWAVE DATASE

This dataset is obtained using a 90-day time-window centred on the current analysis in the frequency filter (Fig.2b). Only analysis data is used in the window, which could not be achieved in near-real-time since it requires 45 days of analysis after the “current analysis”. This is used as the best estimate available for equatorial wave amplitude and phase, against which the other wave datasets will be evaluated.

3) PERFECT FORECAST WAVE DATASET

To isolate the impact of the edge effect associated with off-centering the time-window used for the frequency filter so that there is far less “future data” in the real-time methodology, we construct this wave dataset, by repeating the REAL-TIME methodology, but replacing the 7-day forecasts with analysis data (mimicking a perfect forecast, Fig.2a).

The difference between the PERFECT FORECAST and the DIAGNOSTIC ANALYSIS identifies the influence of the off-centered time window, and the difference between the PERFECT FORECAST and the REAL-TIME FORECAST wave dataset isolates the impact of forecast errors on the REAL-TIME ANALYSIS and the skill in NWP forecasts of the waves.

4) PADDED WAVE DATASET

To explore the value of using the forecast data in the REAL-TIME ANALYSIS of equatorial waves, we create one more additional wave dataset by repeating the REAL-TIME methodology but replacing the forecast data with zeros (see Figure 2a), referred to as the PADDED wave dataset following Wheeler and Weickman (2001).

b. Evaluation of amplitude and zonal propagation of equatorial waves in terms of horizontal winds

To examine the impact of the real-time filtering methodology we compare the waves identified from the PERFECT FORECAST dataset and from the REAL-TIME FORECAST dataset with those from the DIAGNOSTIC ANALYSIS. The amplitude and phase propagation of waves can be clearly demonstrated in a longitude-time Hovmöller diagram. A wind component is chosen to characterise the meridional structure of each wave type at a latitude where its amplitude is a maximum: Kelvin wave u on the equator, WMRG v on the equator, R1 v at 8° N and R2 v at 13° N. Since the meridional structure of each wave type is given by theory and therefore fixed (for the chosen equatorial trapping scale) we would obtain the same time series by showing the amplitude of projected wave components at any latitude (apart from differing magnitude by a constant factor). Results presented in this section will be illustrated on one year, 2016 (the other three years give very similar features).

Figures 3 a,b show Hovmöller diagrams of the Kevin wave *u* at 850hPa from the DIAGNOSTIC ANALYSIS (Fig.3a) and PERFECT FORECAST dataset at day T+2 (Fig.3b). By eye, the wave amplitude and zonal phase behaviours look very similar in the DIAGNOSTIC ANALYSIS and PERFECT FORECAST. This is confirmed by the difference between them shown in Fig.3d, being less than 0.5 m s^{-1} in most of the time and space domain. It is expected that the differences appear to be mainly on low frequencies, indicating the edge effect is small at high frequency. To examine the wave behaviours in the REAL-TIME FORECAST, the Kevin wave identified from the REAL-TIME FORECAST dataset at day T+2 (Fig.3c) shows that the NWP forecasts can capture strong waves reasonably well, for example, waves around the middle of April and the middle of June. However, it is clear that there are systematic errors which seem to be associated with the two highlands around 35°E and 280°E: over the East Africa and the Andes, respectively (Figs.3c, e). It should be noted that the difference between the REAL-TIME FORECAST and the DIAGNOSTIC ANALYSIS (Fig.3e) includes the impact of both the time-window methodology and the forecast skill.

Figure 4 shows Hovmöller diagrams for the WMRG wave *v* at 850hPa. As with the Kelvin wave, the WMRG wave amplitude and zonal propagation behaviours in the PERFECT FORECAST dataset (Fig.4b) are very similar to those of DIAGNOSTIC ANALYSIS (Fig.4a), with very small differences between them (Fig.4d). The REAL-TIME FORECAST wave dataset (Fig.4c) captures the WMRG waves well at T+2 day, especially high amplitude wave packets. However, there is also an orography-related bias (Fig.4e) though weaker than that for the Kelvin wave.

Similar analysis is also performed for R1 and R2 waves. It indicates that the two waves are well simulated by the Real-time methodology and their orography-related errors are smaller (Hovmöller plots not shown).

Since there is a strong orography-related component to the bias, the 12-month mean amplitudes are calculated for the Kelvin wave zonal wind, and for the WMRG, R1 and R2 meridional winds at the latitudes of their maxima. The results for the Kelvin waves are shown in Figure 5 for the PERFECT FORECAST wave dataset (Fig.5a) and REAL-TIME FORECAST wave dataset (Fig.5b) for a selection of lead times. For comparison, the 12-month mean of the waves in the DIAGNOSTIC ANALYSIS (black solid) is also shown in each panel. The blue lines are for T-2 and T+0 (current analysis), and three red lines for T+2, T+4 and T+6. It is seen that the time-mean zonal wind of the Kelvin wave is close to zero in the DIAGNOSTIC ANALYSIS (black line). For the PERFECT FORECAST wave dataset, the mean amplitudes of the Kelvin wave at T-2 and T+0 (blue) are close to that in the DIAGNOSTIC ANALYSIS, whereas the mean amplitude beyond day 0 (red) differs from zero mainly near 35°E and 280°E, close to the high orography.

On the other hand, time-mean Kelvin wave zonal wind amplitudes for the REAL-TIME FORECAST wave dataset (Fig.5b) have much larger departures, especially around the Andes, with an easterly bias to the west and westerly bias to the east. On close inspection, it is interesting to see that peaks of the wind bias shift eastward with lead time which may be an indication of spurious wave generation by processes in the vicinity of the orography. Errors in the three westward-moving wave fields are much smaller, especially for the R1 and R2 waves (not shown).

To remove the bias in the REAL-TIME FORECAST, for each lead time (T+L), the mean of the previous 30 days wave data (already obtained following the filtering and projection steps above) is subtracted. We choose a 30-day running mean for this bias correction because it matches the longest period retained by the pre-processing filter and minimizes the amount of rolling forecast data which needs to be stored to calculate the bias. Figure 5c shows the 12-month mean for REAL-TIME FORECAST Kelvin wave *u* after the 30-day time-mean bias correction. It is clear that the bias has been greatly reduced. From now all results for the REAL-TIME FORECAST are with bias corrected.

Figures 5 d, e show Hovmöller diagrams of the Kelvin waves and WMRG waves in the REAL-TIME FORECAST wave dataset with the lead-time dependent bias removed. It is seen that the forecast waves more closely resemble those of DIAGNOSTIC ANALYSIS (Fig.3a and Fig.4a) than those before the bias is removed (Fig. 3c and Fig.4c). The differences between the T+2 forecast and DIAGNOSTIC ANALYSIS for the Kevin wave and WMRG are much reduced and dominated by errors with spatial and temporal characteristics of the observed wave fields. An example for the WMRG wave being shown in Fig.5f.

After showing the wave behaviours for the PERFECT FORECAST and REAL-TIME FORECAST wave-datasets, it is of interest to evaluate the benefit of including the forecast data in the real-time methodology by examining the waves in the PADDED wave dataset. The result for the PADDED dataset is shown in Fig.6. It is seen that wave analysis (T+0) from the PADDED dataset differs greatly from that in the DIAGNOSTIC ANALYSIS, with little skill in capturing wave behaviours (Figs.6b,d), and at day 2 there are only some low frequency wave signals (Figs.6c,e). This suggests that the forecast data is indeed useful not only in providing future information about the equatorial waves, but also in an accurate REAL-TIME ANALYSIS.

c. Validation of wave variance, error and correlation

To examine the variability of equatorial waves, Figure 7 shows the standard deviation of the wind strength in the four wave components in the PERFECT FORECAST dataset (left), FORECAST dataset without removing previous 30-day mean (middle) and REAL-TIME FORECAST dataset with the 30-day mean removed for T+0 to T+7 (right). It is seen that for all waves in the PERFECT FORECAST, and R1 and R2 in the REAL-TIME FORECAST dataset, their standard deviation at all lead times is very close to those in the DIAGNOSTIC ANALYSIS (black), except for the westward moving waves at T+6 around 40°E where there is a spike. For Kelvin and WMRG waves in the REAL-TIME FORECAST dataset, their standard deviations are also close to those in DIAGNOSTIC ANALYSIS at days T-2 and T+0, but from day T+2 their variability is weaker than those in DIAGNOSTIC ANALYSIS. The REAL-TIME FORECAST dataset with 30-day mean removed (right) has very similar variability to that before the removal of the 30-day mean indicating that the 30-day time-mean is appropriate for removing the orography-related anomalies, without removing variability in wave field.

The root mean square errors (RMSE) relative to the DIAGNOSTIC ANALYSIS and correlations of each wave in the different datasets with those in DIAGNOSTIC ANALYSIS are shown in Fig.8. RMSEs for the four waves identified with different procedures are standardised by the standard deviation of the wave mode in the DIAGNOSTIC ANALYSIS. The correlations are calculated with samples at all longitudes and time (360 longitudes*366 days). For the PERFECT FORECAST (solid lines), the normalised RMSE for each wave is less than 0.2 at day -4 on both the 850 and 200 hPa levels, with the error at 850hPa being slightly smaller than that at 200hPa for all wave modes. It increases slowly with the lead time, to about 0.3 at day 4. After day 4 the errors increase faster but are still less than 0.5 at day 6. At the end of the time window used by the frequency filter (T+7) the errors jump to 0.75~ 0.95 due to the large edge effect of the filter. For the REAL-TIME FORECAST dataset (dashed lines) before day 0 the RMSEs are comparable to those in the PERFECT FORECAST dataset but increase more rapidly from day 0, reaching around 0.7 at day 4, 0.8

at day 5 and 1.0 at day 7. The REAL-TIME FORECAST RMSEs for WMRG and R1 waves at 200 hPa are larger than those at 850 hPa, as is the case for the PERFECT FORECAST. As expected, the errors in the PADDED dataset (dotted) are much larger than those of either of the other methods at all lead times, indicating that the forecast data is useful even in improving the Real-time analysis of equatorial waves.

The correlations (Figure 8, bottom two rows) convey similar information to the normalised RMSEs. The PERFECT FORECAST dataset (solid) correlations are quite high and drop slowly with lead time, remaining larger than 0.9 up to day 6. For the REAL-TIME FORECAST dataset (dashed), although the correlations for WMRG and Rossby waves drop steadily after day 1, they are still larger than 0.6 at day 6. The correlations for the Kelvin waves fall faster with lead time, reaching 0.6 at day 5 and 0.5 by day 6.

The correlations at the “current analysis time” T+0 for the PADDED dataset (dotted lines) are much lower (~0.75) than in REAL-TIME ANALYSIS at T+0 and decrease rapidly with lead time to 0.4 at T+1 showing that there is limited skill in the statistical interpolation associated with the wavenumber-frequency filter.

4. A case study illustrating real-time analysis of equatorial waves in January 2016

a. Hovmöller diagrams of wave propagation

As an example, the utility of real-time analysis of equatorial waves is illustrated in a case study. Figure 9 shows Hovmöller diagrams of horizontal wind components projected onto the different equatorial wave structures, averaged over 10°N-10°S. The REAL-TIME FORECAST initialised on 17 January 2016 (left) is compared with the PERFECT FORECAST (right) where analyses have been substituted for forecast fields in the diagnostic procedure. To examine the potential connection of waves with deep convection, NOAA OLR averaged over 10°N-10°S is also shown in each panel (colour shading).

Also to illustrate robustness of the spatial projection technique and attribution to different wave modes, the forecast winds projected onto the wave structures are contrasted with the winds that are subject only to the wavenumber-frequency filter (with the same off-centered time-window) but without the spatial projection step. Zonal wind (Fig.9a) is shown for the eastward-moving component (i.e., filtered winds for eastward-moving wavenumber – frequency domain) to be compared with the Kelvin wave (Fig.9b); and meridional wind (Fig.9c) is shown for the westward-moving component (filtered for westward-moving domain), to be compared with the WMRG waves (Fig. 9d), and the antisymmetric component of the meridional wind (Fig.9e) is to be compared with the R1 wave (Fig.9f).

Kelvin waves (Fig. 9b) dominate the eastward moving u (Fig. 9a) in both analyses and forecasts ($T > 0$). In the PERFECT FORECAST, there is a strong Kelvin wave signature to the east of the dateline which develops after the initialization of the forecast and is closely coupled to an eastward moving convection signal moving with the westerly flow. The forecasts certainly develop a signature of this propagating Kelvin wave, even though propagation only begins after the analysis time (T+0). However, the forecast amplitude is clearly too weak beyond T+2, consistent with the reduction in wave variance for Kelvin Waves seen in Figure 7. Strong Kelvin wave activity is also seen in the 60°E-140°E region in analyses before T+0, which is coupled with the convective activity there. However, after T+0, the forecast wave shows a weaker signal than that in the PERFECT FORECAST

In contrast, the forecasts seem to capture well the phase and amplitude of the westward moving waves since the REAL-TIME FORECAST resemble closely the PERFECT FORECAST out to lead times of 6 days. In this case a strong wave packet over the 40°E-120°E sector is well simulated in the forecast. The packet projects onto the WMRG and R1 wave components (Figs.9c-f). The phase propagation is westwards with the WMRG waves moving slightly faster than the R1 waves. The group speed appears to be eastward in both the WMRG and R1 waves, as anticipated from the theoretical dispersion relation (for WMRG and short wavelength R1 waves). Note that Yang et al. (2018) showed that WMRG and R1 waves frequently propagate westwards together over the Atlantic from West Africa. However, they established from composites of many events that the wave components do propagate at different speeds and have different group speeds, consistent with independent wave modes from linear theory, after accounting for the Doppler shift of frequencies by the background zonal flow.

In DJF, convection over the 40E-120E sector is biased to the Southern Hemisphere and is expected to dominate the average OLR signal from 10°S to 10°N. Fig. 9d indicates that low OLR (a proxy for deep convection) is coincident with $v < 0$ in the WMRG component in the centre of the wave packet (60-80E). Figure 1 shows that northerly winds across the equator in the WMRG wave are in phase with convergence in the Southern Hemisphere where the deep convection is occurring. In contrast, the same minimum in OLR is coincident with the phase of cyclonic circulation in the R1 wave component (on both sides of the equator). These wave-convection relations are consistent with those in previous observational studies (e. g., Wheeler and Kiladis 1999; Wheeler et al. 2000; Yang et al.2003, 2007a, b and Ferrett et al. 2020).

This case study demonstrates that another useful utility of real-time analysis of equatorial waves in operational forecasting is to use maps of the wave components overlain on satellite observations of convective activity (OLR or precipitation estimates). In this way, the complex structure in the OLR field could be interpreted as a superposition of different equatorial wave activity. This could help forecasters explore continuity within forecasts and between forecasts with updated lead times, given the anticipated propagation of the different wave modes.

b. Identification of wave amplitude and phase

Given the close relationship between precipitation and wave phase (e.g. Ferrett et al. 2020), for forecasting applications it can be useful to define a local wave phase and amplitude. Also, because the spatial projection onto equatorial wave structures is performed independently for each variable (v, q, r of Eqn. 2), it is important to consider temporal coherence of the variables when diagnosing propagation rate, phase and amplitude. For each wave type we define a local phase space diagram based on variables which from the theoretical horizontal structures we would expect to be in quadrature (Fig. 1). In contrast to the familiar RMM Phase Diagrams for the MJO (Wheeler and Hendon, 2004), where the phase refers to the longitude where the convection is most active, these wave phase diagrams refer to the passage of a wave over a fixed longitude. Some studies also use a local phase diagram on CloudSat data to study its relationship with the MJO or on the filtered precipitation to examine convectively couple equatorial waves (Riley 2011 and Yasunaga and Mapes 2012). We choose to define our wave phase diagrams such that in the mode structure (Fig. 1) variable-2 (W_2) has a positive maximum one quarter of a wavelength to the west of variable-1 (W_1). Consequently, from the perspective of an observer at a fixed longitude, W_1 is positive before W_2 for the eastward moving Kelvin wave and the propagation around the phase diagram is anti-clockwise (consistent with the eastward moving MJO in the Wheeler-Hendon Diagrams). The phase propagation is clockwise for the westward moving WMRG and R1 waves. For the Kelvin wave, which is dominated by divergent winds and has zero meridional wind, the two variables used are u and $\partial u/\partial x$ on the equator. For WMRG and R1 waves, which are dominated by rotational flows, the two variables are $-u$ and v at specified latitudes, these are in phase with vorticity and divergence, respectively. These variables and their latitudes are summarised in Table 1. Positive values of the first variable, W_1, combined with zero in the second variable, W_2, is used to define the zero phase angle.

Each variable is normalised by its standard deviation and averaged over a 5° longitude range. The wave amplitude, $A(t)$, is then defined as

$$\mathrm{A} = \sqrt{W_1{}^2 + W_2{}^2},$$

Phase angle (φ) is given by

$$\varphi = \arg(W_1 + iW_2)$$

where $\varphi=0$ corresponds to the positive x-axis and increases anti-clockwise.

Figure 10 gives examples of phase-amplitude diagrams for Kelvin wave at 200°E, and WMRG and R1 waves at 90° E in the period of the case discussed in Section 4a with the forecast initialised on 17 January. It is seen that the Kelvin wave moves anticlockwise and WMRG and R1 clockwise. Consistent with Fig.9, in this period the Kevin wave develops around T=0 (17 January) and propagates coherently eastward in the PERFECT FORECAST (black line), whereas the Kelvin wave in the REAL-TIME FORECAST (blue line) fails to grow after T=+ 2 (19 January) and moves too fast, with its phase on 20 January being ahead of that in the PERFECT FORECAST. On the other hand, WMRG and R1 waves over the east Indian Ocean show strong signals and propagate westward coherently both in the PERFECT FORECAST and the REAL-TIME FORECAST. However, it is seen that their amplitudes after T=+3 (20 January) are weaker than those in the PERFECT FORECAST, especially for the WMRG wave. This case indicates that the difference between the forecast (blue) and PERFECT FORECAST (black) at each day can be demonstrated clearly. The phase-amplitude indices will be used in the next section to evaluate forecast skill for equatorial waves.

5. Evaluation of forecast skill for equatorial waves

The analysis in section 3 focuses on the evaluation of the real-time technique. Errors are compared to the DIAGNOSTIC ANALYSIS, which includes errors due to both the off centred window and due to errors in the forecast waves. Here we use the PERFECT FORECAST dataset to evaluate the skill of the Met Office prediction system in forecasting equatorial wave modes.

a. Evaluation of combined errors in phase and amplitude

To isolate the effects of forecast error on equatorial wave identification and evolution we calculate the normalised RMSE and correlation of the REAL-TIME FORECAST compared with the PERFECT FORECAST dataset. The result is shown in Fig.11 (solid lines). For comparison, the corresponding RMSE and correlation of the REAL-TIME FORECAST with the DIAGNOSTIC ANALYSIS are repeated in Fig.11 (dashed lines). The RMSE relative to the PERFECT FORECAST is a measure of NWP forecast error alone whereas the RMSE relative to the DIAGNOSTIC ANALYSIS also includes errors due to the off-centered time-window used in the real-time wave identification method. For all waves, errors relative to PERFECT FORECAST (solid) increase slowly before day -1 where analysis data is used, then increase faster with lead time, to about 0.5~0.7 at day 4, depending on the wave type. A relatively larger difference between 850hPa and 200Pa is shown for the Kelvin wave beyond day 1, with the error at low level being larger and increasing more rapidly than that of the upper level. The RMSEs relative to the PERFECT FORECAST (solid) are smaller than those relative to the DIAGNOSTIC ANALYSIS (dashed), consistent with the fact that the latter also includes errors due to the time-window edge effect.

The correlations with the PERFECT FORECAST dataset are larger than those with the DIAGNOSTIC ANALYSIS (Fig.11, bottom two rows). Correlations with the PERFECT FORECAST before day T+0 have a very high value (larger than 0.95). This drops with lead time but still has a value of 0.8 around day 4 and 0.6-0.7 at day 6, except the Kelvin wave in the lower troposphere where correlation falls faster, consistent with the greater error at this level. It is clear from the separation of the solid and dashed curves that the effect of the off-centred time-window on real-time forecast error increases as the end of the window is approached (at T+7) but that this is a much smaller effect than the growth of NWP forecast error with lead time.

b. Evaluation of phase and amplitude errors of equatorial waves in forecasts

To separate the errors in phase and amplitude, the technique described in Section 4 is used to create a dataset of wave phase and amplitude for the PERFECT FORECAST and REAL-TIME FORECAST (2015-2018) for the wave types, at each date, longitude and lead time.

We calculate a phase difference ($\Delta\varphi = \varphi_f - \varphi_{pf}$) between the forecast phase (φ_f) and the PERFECT FORECAST phase(φ_{pf}), which due to the periodic nature of the phase diagram can be wrapped into the range of $-\pi$ to π.

RMSEs of wave phase and amplitude in the forecast relative to the PERFECT FORECAST in 2015-2018 are shown in Figs.12 a, b. As expected, errors in phase and amplitude increase with lead time. The phase error has a typical magnitude of 0.1π (18^o) at day 0, increasing to $0.3\ \pi$ at day 6 for the westward waves and the larger error of $0.4\ \pi$ for the Kelvin waves. Error in amplitude is about 0.2 at day 0, increasing to 0.6 at day 6 for the Rossby R1 wave, while Kelvin waves reach this high level of amplitude error in only 2.5 days and WMRG waves in 5 days. Note that as mentioned in Section 4, each variable is normalised by its standard deviation, so that amplitude of 1 corresponds to the amplitude error in that wave mode matching the RMS (root mean square) amplitude of that wave mode.

It is also informative to characterize the mean forecast bias in phase and amplitude (Figure 12 c, d). The forecast Kelvin wave shows large positive departures in phase and a much weaker amplitude (a bias of almost 20% of the RMS magnitude). Since Kelvin waves propagate anticlockwise in the phase diagram, the mean positive phase difference indicates that the forecast Kelvin wave is to the east of that in the PERFECT FORECAST, and hence implies a faster phase speed. It is noted that the phase departure does not always increase with lead time, with the largest systematic departure in phase occurring at day 2. The westward-moving WMRG and R1 waves show much smaller phase errors, but the amplitude of the WMRG is also weaker in the forecast (by 10% of average amplitude).

Figure 13 shows that the average phase and amplitude forecast errors depend strongly on longitude. For the Kelvin wave, forecasts at day 2, 4 and 6 consistently show the large positive phase difference (eastward shift of forecast waves) over the Maritime Continent and West Pacific regions and the error develops in the first 2 days of the forecast. On the other hand, the forecast shows weaker Kelvin wave amplitudes over the central and eastern Pacific. In contrast to the Kelvin wave, the WMRG wave does not show a systematic phase difference. However, forecasts have too weak amplitude over the central and eastern Pacific, similar to the Kelvin wave forecasts. Among the three waves, the R1 wave has smallest errors in both phase and amplitude. The inherent forecast errors in Kelvin wave phase speed and Kelvin and WMRG amplitude in these regions imply that understanding the cause of these errors would be crucial for improving the model's ability to predict the equatorial waves, and their associated HIW.

A possible cause of the Kelvin wave errors may be that the model fails to simulate observed tropical eastward-moving convective activity coupled with Kelvin waves, as found in an earlier version of the MetUM by Ringer et al. (2006) and Yang et al. (2009). The latter shows that observed equatorial convection tends to appear in the region of low-level wave-enhanced near-surface westerlies in Kelvin waves crossing the eastern hemisphere warm water region (where there is westerly ambient flow), while the older versions of MetUM (HadAM3 and HaGAM1) tends to place convection closer to the maximum in the low level convergence. This suggests that wind-dependent energy fluxes may play an important role in triggering/organising equatorial convection, which can then modify and possibly amplify the

Kelvin waves. The models also do not capture the observed vertical tilt structure and signatures of energy conversion in the Kelvin waves. It is worth investigating if these issues have been improved in the current version of the model.

6. Summary and discussion

In this study a novel technique for real-time identification of equatorial wave modes in analysis and forecast data has been developed. Most existing methods for identifying equatorial waves from global analysis data rely on the application of a frequency, as well as a wavenumber filter, followed by different approaches to partitioning into different wave modes. In the context of historical data or climate model analysis, a much longer time-window (one year or more) is used as input to the frequency filter. However, in an operational forecasting context, the time-window cannot extend far into the future beyond the current date due to lack of data. Therefore, the key challenge that has been addressed here is to identify equatorial waves objectively at the current date and in near-range forecasts without a strong dependence on future data.

The method relies on identifying equatorial waves through the spatial projection of global data onto the horizontal structures of equatorial waves derived from theory (Yang et al. 2003). A broad-band frequency filter is nevertheless required to filter out stationary disturbances and to partition eastward-moving and westward-moving disturbances. In our approach, the phase speed of the waves identified is not tightly constrained by the spectral filter. The Fourier transform is conducted across a broad range in zonal wavenumber ($2 \leq k \leq 40$) and frequency (with periods in the range $2 < \tau < 30$ days). This means that features with different characteristic scales, such as dispersive wave packets, can be represented well. The technique has been evaluated using four-years of Met Office operational global analysis and forecast data (2015-18).

The new methodology where the time-window used in the frequency filter is off-centred, with 83 days before the current analysis and 7 days after, was first evaluated using analysis data only. The PERFECT FORECAST dataset, created using the off-centered time-window method, is compared against the DIAGNOSTIC ANALYSIS dataset which uses the same methodology except that the time-window is centred on the current analysis (with knowledge of the atmospheric state 45 days into the future). RMSE shows the impact of the time-window and frequency filtering procedure to be small at the current analysis time (T+0) with error growing only slowly to T+6. The final day of the PERFECT FORECAST (T+7 days in this case) should not be used due to large errors from the edge effects of the time-window on the frequency filtered data.

A REAL-TIME FORECAST dataset is then constructed using the same off-centred time-window technique, but with 83 days of analyses before the current analysis (T+0) concatenated with 7 days of global forecasts initialised from the analysis at T+0. In a real-time forecasting context, the T+0 analysis would be the latest available. The DIAGNOSTIC ANALYSIS is used as a reference for truth, assumed to be the best estimate of observed wave amplitude and phase. However, the effect of forecast error can be partitioned from the effects of the off-centered time-window by comparing the REAL-TIME FORECAST with the PERFECT FORECAST.

Forecast skill in wave amplitude and phase is appreciable to day T+6. The forecast error is much larger than the diagnostic error associated with the off-centered time-window technique (excluding the last day in the window, T+7, which suffers from the edge effect of the diagnostic method). The correlation between the REAL-TIME FORECAST and the DIAGNOSTIC ANALYSIS exceeds 0.6 for the Rossby and WMRG waves out to day T+6. The skill in Kelvin wave forecasts is lower, with the correlation dropping below 0.6 on average by day T+5. These results are encouraging, indicating that the real-time technique is able to identify the waves in the operational forecasting context and furthermore that the Met Office prediction system (the global high resolution “deterministic” forecast) has demonstrable skill in forecasting the equatorial waves.

For comparison, the real-time technique is modified with the data in the 7 days beyond the current analysis T+0 being over-written with zeros. This PADDED wave dataset shows that even the representation of the current analysis (T+0) is affected detrimentally by this padding approach – the correlation with the DIAGNOSTIC ANALYSIS fields dropping to 0.75. Moreover, moving into the forecast window the skill drops rapidly with the correlation falling below 0.5 at T+1 and below 0.3 at T+2. The existence of skill in this range arises as a form of statistical forecast propagating wave information forward from the preceding 83 days of analysis as a result of the wavenumber-frequency filter. However, the correlations are much lower than the REAL-TIME FORECAST dataset, even at T+0, demonstrating the value of NWP forecast information.

Local wave phase diagrams are constructed using two variables that are known to be in quadrature for each wave mode structure (a fundamental property of the propagation mechanism). The variables are chosen so that variable-2 (W_2) has a maximum one quarter of a wavelength to the west of the maximum in variable-1 (W_1). The local wave phase diagram constructed using $W_1(X, t)$ and $W_2(X, t)$ as its axes then relates to the wave amplitude and phase that would be seen by an observer at a fixed longitude, X. In this way eastward-moving waves progress anti-clockwise in the phase diagram and westward moving waves progress clockwise. The trajectories of forecasts can be compared quantitatively with the sequence of analysis states in the local wave phase space (e.g., Fig.10).

The local wave phase space is used to quantify both systematic forecast bias and RMS forecast error. The characteristic features identified are:

- RMS error grows steadily with lead time for the WMRG and R1 waves to T+6.

- RMS error grows much faster for Kelvin waves reaching a similar level of amplitude error (0.6 of the RMS wave amplitude) by day T+2.
- There is a systematic eastward shift in wave phase for the Kelvin wave associated with too fast propagation in the model.
- The Kelvin wave phase error is greatest from the Indian Ocean, across the Maritime Continent to the Dateline.
- There is not a large systematic phase error for the WMRG and R1 waves.
- Kelvin and WMRG wave amplitude decays over the first 4 days of the forecast. Systematic bias is 20% of wave RMS amplitude for Kelvin waves, 10% for WMRG waves.
- This amplitude under-estimate is dominated by the contribution from waves across the Pacific east of 150°E where the bias is 25% for Kelvin and WMRG waves.

The inherent phase speed error for the Kelvin waves over the Maritime Continent to central Pacific and also the under-estimate of the Kelvin wave and WMRG activity in the central and eastern Pacific merit further investigation since these systematic errors can have a major impact on forecast skill in the Tropics. In addition to the possible error in modelling the coupling between convection and the dynamical structure of the waves as discussed in the last section, another possible cause of phase speed error could be due to the basic zonal flow error in the model. The basic flow results in a Doppler shift of equatorial wave frequencies (e.g, Yang et al 2011, 2012, 2018 and Dias and Kiladis 2014) and zonal variation of the zonal flow would also affect the energy dispersion of equatorial waves (e.g., Hoskins and Yang 2016).

The new diagnostic technique will enable more detailed investigation of the representation of equatorial wave structure and evolution by forecast models, and comparison with observed behaviour as represented in global analyses. The case study shown here (Section 4) illustrates application of the diagnostics in an operational forecast context. The continuity of equatorial waves from the recent history of analyses into the forecasts is immediately apparent in a Hovmöller plot over a two-week window (centred at the current analysis). The systematic tendency for the model to decay amplitude in the eastward-moving disturbances, and also to propagate too quickly, is important information for forecasters – especially in Southeast Asia where the systematic model errors have been shown to be largest. Forecasters could construct their advice taking into account the systematic error.

Equatorial waves are central to the occurrence of HIW associated with widespread heavy precipitation. There are major ramifications from errors in equatorial wave forecasts for severe weather warnings and advice to emergency responders, even on the short-range from lead times of one day to a week. On the positive side, there is appreciable skill in forecasts of equatorial wave structures (in the wind field) out to day 6, offering much higher predictability than associated with isolated convective systems. Because forecasting HIW associated with equatorial waves depends not only the ability to forecast the waves, but also on the ability of the model to capture the correct relationship between the waves and HIW, if improvements can be made in the relationship between equatorial waves, deep convection and precipitation rate, there is great scope for capitalising on this potential predictability.

Acknowledgments.

We would like to acknowledge the very helpful reviews we received from the two anonymous reviewers. This work and its contributors were supported by the Weather and Climate Science for Services Partnership (WCSSP) Southeast Asia as part of the Newton Fund. SJW and G-YY were also supported by the National Centre for Atmospheric Science ODA national capability programme ACREW (NE/R000034/1), which is supported by NERC and the GCRF.

REFERENCES

Bengtsson, L., and Coauthors, 2019: Convectively coupled equatorial wave simulations using the ECMWF IFS and the NOAA GFS cumulus convection schemes in the NOAA GFS model. *Mon. Wea. Rev.*, 147 4005-4025, doi:10.1175/MWR-D-19-0195.1.

Castanheira, J. M., and C. A. F. Marques, 2015: Convectively coupled equatorial waves diagnosis using 3-D normal modes. *Quart. J. Roy. Meteor. Soc.* **141**, 2776–2792. doi: 10.1002/qj.2563.

Dias, J., M. Gehne, G. N. Kiladis, N. Sakaeda, P. Bechtold, and T. Haiden, 2018: Equatorial waves and the skill of NCEP and ECMWF numerical weather prediction systems. Mon. Wea. Rev., 146, 1763-1784, doi:10.1175/MWR-D-17-0362.1.

Dias, J., and G. N.Kiladis, 2014: Influence of the basic state zonal flow on convectively coupled equatorial waves. Geophys. Res. Lett., 41, 6904–6913

Ferrett, S., G-Y. Yang, S. Woolnough, M. S. Methven, K**.** Hodges, and C. Holloway, 2020: Linking Extreme Precipitation in Southeast Asia to Equatorial Waves. *Quart. J. Roy. Meteor. Soc.,* **146***,* 665-684.

Gehne M, and R. Kleeman, 2012. Spectral analysis of tropical atmospheric dynamical variables using a linear shallow-water modal decomposition. *J. Atmos. Sci.*, **69**, 2300–2316.

Gill, A. E., 1980: Some simple solutions for heat induced tropical circulations. *Quart. J. Roy. Meteor. Soc.,* **106,** 447–462.

Gottschalk J, M. Wheeler, K. Weickmann, F. Vitart , N. Savage, H. Lin. H. Hendon, D. Waliser, K. Sperber, C. Prestrelo,M. Nakagawa, M. Flatau, W. Higgins, (2010): A framework for assessing operational model MJO forecasts: A project of the CLIVAR Madden—Julian Oscillation Working Group. *Bull. Am. Meteorol. Soc.* **91**: 1247– 1258.

Hayashi, Y., 1982: Space-time spectral analysis and its applications to atmospheric waves. *J. Meteor. Soc. Japan*, **60**, 156–171.

Hendon, H. H., and B. Liebmann, 1991: The structure and annual variation of antisymmetric fluctuations of tropical convection and their association with Rossby–gravity waves. *J. Atmos. Sci.*, **48,** 2127–2140.

Hoskins, B. J., and G.-Y. Yang, 2016: The longitudinal variation of equatorial waves due to propagation on a zonal varying flow. *J. Atmos. Sci.*, **73**, 605–620.

Huang, P., C. Chou, and R. Huang, 2013: The activity of convectively coupled equatorial waves in CMIP3 global climate models. *Theoretical and Applied Climatology*, **112**, 697-711.

Janiga, M. A., J. C. Schreck, J. A. Ridout, M. Flatau, N. P. Barton, E. J. Metzger, and C. A.Reynolds, 2018: Subseasonal Forecasts of Convectively Coupled Equatorial Waves and the MJO: Activity and Predictive Skill. *Mon. Wea. Rev.*, **146**, 2337-2360. doi: 10.1175/MWR-D-17-0261.1

Judt, F., 2020: Atmospheric predictability of the tropics, middle latitudes, and polar regions explored through global storm-resolving simulations. *J. Atmos. Sci.*, **77**, 257-276, doi:10.1175/ JAS-D-19-0116.1.

Li, Y., and S. N. Stechmann, 2020: Predictability of tropical rainfall and waves: Estimates from observational data. *Quart. J. Roy. Meteor. Soc.*, **146**, 1668–1684. doi:10.1002/qj.3759.

Kikuchi, K. 2014: An introduction to combined Fourier–wavelet transform and its application to convectively coupled equatorial waves. *Climate Dynamics*. **43**, 1339–1356.

Kikuchi, K., G. N. Kiladis, G. N., J. Dias, and T. Nasuno, 2018: Convectively coupled equatorial waves within the MJO during CINDY/DYNAMO: slow Kelvin waves as building blocks. *Climate Dynamics*, **50**, 4211-4230.doi:10.1007/s00382-017-3869-5.

Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Rev. Geophys.*, **47**, RG2003, doi:10.1029/2008RG000266.

Liebmann, B., and H. H. Hendon 1990: Synoptic-scale disturbances near the equator, *J. Atmos. Sci.*, **47**, 1463– 1479.

Liebmann, B., and C.A. Smith, 1996: Description of a Complete (Interpolated) outgoing longwave radiation dataset. *Bull. A. Met. Soc.*, **77**, 1275-1277.

Lin, J.-L., and Coauthors, 2006: Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: Convective signals. *J. Climate*, **19**, 2665–2690.

Magaña, V., and M. Yanai 1995: Mixed Rossby-gravity waves triggered by lateral forcing, *J. Atmos. Sci.*, **52**, 1473–1486.

Marques, C.A.F., and J. M. Castanheira, 2018: Diagnosis of Free and Convectively Coupled Equatorial Waves. *Math Geosci* **50,** 585–606. doi: 10.1007/s11004-018 -9729-y.

Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *J. Meteor. Soc. Japan,* **44,** 25–43.

Redelsperger, J.-L., and Coauthors, 1998: Review of convection in TOGA-COARE. Proc. CLIVAR/GEWEX COARE98 Conf., Boulder, CO, WCRP, 16–42.

Riley, E.M., B. E. Mapes, and S. N. Tulich: 2011: Clouds Associated with the Madden-Julian Oscillation: A New Perspective from CloudSat. J. *Atmos. Sci.*, **68**, 3032-3051. doi: 10.1175/JAS-D-11-030.1.

Ringer, M. A., and Coauthors, 2006: The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model (HadGEM1). Part II: Aspects of variability and regional climate. *J. Climate*, **19**, 1302–1326.

Roundy, P. E., and W. M. Frank 2004: A climatology of waves in the equatorial region. *J. Atmos.Sci.*,**6**, 2105–2132.

Roundy, P. E., 2008: Analysis of Convectively Coupled Kelvin Waves in the Indian Ocean MJO. *J. Atmos. Sci.,* **65**, 1342–1359.

Roundy, P. E. 2012: Tracking and prediction of large-scale organized tropical convection by spectrally focused two-step space–time EOF analysis. *Quart. J. Roy. Meteor. Soc.*, **138**, 919–931.doi:10.1002/qj.962

Slingo, J. M., E. Guilyardi, K. I. Hodges, B. J. Hoskins, P. M. Inness, D. M. Lawrence, R. Neale, T. M. Osborne, H. Spencer, and G.-Y. Yang, 2003: How good is the Hadley Centre climate model? Research at CGAM on identifying and understanding model systematic errors: *1999-2002 - CGAM/NCAS Report*.

Straub, K. H., and G. N. Kiladis 2002: Observations of a convectively coupled Kelvin wave in the eastern Pacific ITCZ. *J. Atmos. Sci.*, **59**, 30–53.

Straub, K.H., P. T. Haertel, and G. N. Kiladis, 2010: An analysis of convectively coupled Kelvin Waves in 20 WCRRP CMIP3 Global Climate Models. *J. Climate*, **23**, 3031-3056.

Takayabu, Y. N. 1994a: Large-scale cloud disturbances associated with equatorial waves. Part I: Spectral features of the cloud disturbances. *J. Meteorol. Soc. Jpn.,* **72**, 433–448.

Takayabu, Y. N. 1994b: Large-scale cloud disturbances associated with equatorial waves. Part II: Westward propagating inertio-gravity waves, *J. Meteorol. Soc. Jpn.*, **72**, 451–465.

Takayabu, Y. N., and T. S. Nitta 1993: 3–5 day-period disturbances coupled with convection over the tropical Pacific Ocean, J. *Meteorol. Soc. Jpn.*, **71**, 221–246.

Wallace, J. M., and V. E. Kousky, 1968: Observational evidence of Kelvin waves in the tropical stratosphere. *J. Atmos. Sci.*, **25**, 900–907.

Walters, D., I. Boutle, M. Brooks, T. Melvin, R. Stratton, S. Vosper, H. Wells, K. Williams, N. Wood, T. Allen, A. Bushell, D. Copsey, P. Earnshaw, J. Edwards, M. Gross, S. Hardiman, C. Harris, J. Heming, N. Klingaman, R. Levine, J. Manners, G. Martin, S. Milton, M. Mittermaier, C. Morcrette, T. Riddick, M. Roberts, C. Sanchez, P. Selwood, A. Stirling, C. Smith, D. Suri, W. Tennant, P. L.Vidale, J. Wilkinson, M. Willett, S. Woolnough, and P. Xavier,2017.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, Geosci. Model Dev., 10, 1487–1520, https://doi.org/10.5194/gmd-10-1487-2017, 2017.

Wheeler, M. C. and H. H. Hendon 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917-1932.

Wheeler, M. and G.N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber– frequency domain. *J. Atmos. Sci.*, **56**, 374–399.

Wheeler, M., G. N. Kiladis, and P. J. Webster 2000: Large-scale dynamical fields associated with convectively coupled equatorial waves, *J. Atmos. Sci.*, **57**, 613–640.

Wheeler, M. C., and K. M. Weickmann, 2001: Real-time monitoring and prediction of modes of coherent synoptic to intraseasonal tropical variability. *Mon. Wea. Rev.*, **129**, 2677–2694.

Wood, N., A. Staniforth, A. White, T. Allen, M. Diamantakis, M. Gross, T. Melvin, C. Smith, S. Vosper, M. Zerroukat, and J. Thuburn, 2014: An inherently mass-conserving semiimplicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations, *Q. J. Roy. Meteorol. Soc.*, **140**, 1505–1520.

Yanai, M., and T. Maruyama, 1966: Stratospheric wave disturbances propagating over the equatorial Pacific. *J. Meteor. Soc. Japan*, **44**, 291–294.

Yang G-Y, B. J. Hoskins, 2013: ENSO impact on Kelvin waves and associated tropical convection. *J. Atmos. Sci.,* **70**, 3513-3532.

Yang, G-Y., and B. J. Hoskins, 2016: ENSO-related variation of equatorial MRG and Rossby waves and forcing from higher latitudes. *Quart. J. Roy. Meteor. Soc.* **142**, 1488-2504.

Yang, G-Y., B. J. Hoskins and L. Gray, 2012: The Influence of the QBO on the propagation of equatorial waves into the stratosphere. *J. Atmos. Sci.*, **69**, 2959–2982.

Yang, G-Y., B. J. Hoskins and J. M. Slingo, 2003: Convectively coupled equatorial waves: A new methodology for identifying wave structures in observational data. *J. Atmos. Sci.*, **60**, 1637–1654.

Yang, G-Y., B. J. Hoskins and J. M. Slingo, 2007a: Convectively coupled equatorial waves: Part I: Horizontal structure. *J. Atmos. Sci.*, **64**, 3406– 3423.

Yang, G-Y., B. J. Hoskins and J. M. Slingo, 2007b: Convectively coupled equatorial waves: Part II: Zonal propagation. *J. Atmos. Sci.*, **64**, 3424–3437.

Yang, G-Y., B. J. Hoskins and J. M. Slingo, 2007c: Convectively coupled equatorial waves: Part III: Synthesis structures and extratropical forcing. *J. Atmos. Sci.*, **64**, 3438–3451.

Yang, G-Y., B. J. Hoskins and J. M. Slingo, 2011: Equatorial waves in opposite QBO phases**.** *J. Atmos. Sci.*, **68**, 839-862

Yang, G-Y., J. Methven, S. J. Woolnough, K. Hodges, and B. J. Hoskins, 2018: Linking African Easterly Wave activity with equatorial waves and the influence of Rossby waves from the Southern Hemisphere. *J. Atmos. Sci.,* **75***,* 1783-1809.

Yang, G-Y., J. M. Slingo and B.J. Hoskins, 2009: Convectively coupled equatorial waves in high resolution Hadley centre climate models. *J. Climate*, **22**, 8, 1897-1919.

Yasunaga, K., and B. E. Mapes, 2012: Differences between More Divergent and More Rotational Types of Convectively Coupled Equatorial Waves. Part II: Composite Analysis based on Space–Time Filtering. *J. Atmos. Sci.*, 69, 17-34.

Ying, Y. and F.Zhang, 2017: Practical and intrinsic predictability of multiscale weather and convectively coupled equatorial waves during the active phase of an MJO. *J. Atmos. Sci*, **74**, 3771– 3785.

Žagar, N., Kasahara, K. Terasaki, J. Tribbia, and H. Tanaka, 2015: Normal-mode function representation of global 3-D data sets: open-access software for the atmospheric research community. *Geoscientific Model Development*, **8**, 13411169–1195. doi: 10.5194/gmd-8-1169-2015.

Žagar, N., J. Tribbia, J. L. Anderson, and K. Raeder, 2009: Uncertainties of estimates of inertio-gravity energy in the atmosphere. Part I: intercomparison of four analysis datasets. *Mon. Wea. Rev.*,**137**, 3837-3857. Corrigendum, *Mon. Wea. Rev.*, **A.**, 2476-2477.

Zangvil, A. 1975: Temporal and spatial behavior of large-scale disturbances in tropical cloudiness deduced from satellite brightness data, *Mon. Weather Rev.*, **103**, 904–920.

Zangvil, A., and M. Yanai 1980: Upper tropospheric waves in the tropics. Part I: Dynamical analysis in the wavenumber-frequency domain, *J. Atmos. Sci.*, **37**, 283–298.

Zangvil, A., and M. Yanai 1981: Upper tropospheric waves in the tropics. Part II: Association with clouds in the wavenumber-frequency domain, *J. Atmos. Sci.*, **38**, 939–953.

Table 1. The two variables used to define the local wave phase diagrams. In each wave structure, variable-2 has a positive maximum one quarter of a wavelength to the west of the location where variable-1 is positive (see Fig.1).

Wave type	W_1	Latitude (W_1)	W_2	Latitude (W_2)
Kelvin	u	0	du/dx	0
WMRG	$-u$	10°S	v	0
R1	$-u$	0	v	8°N

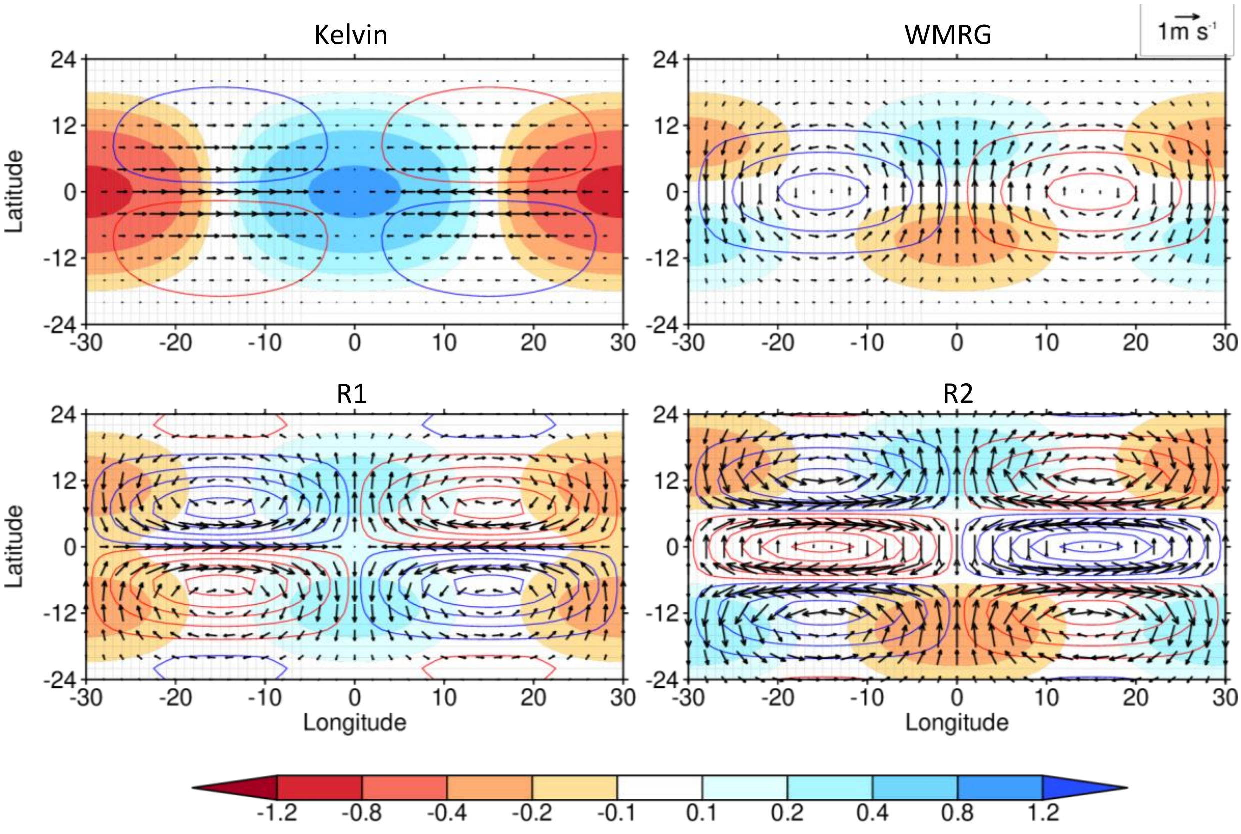
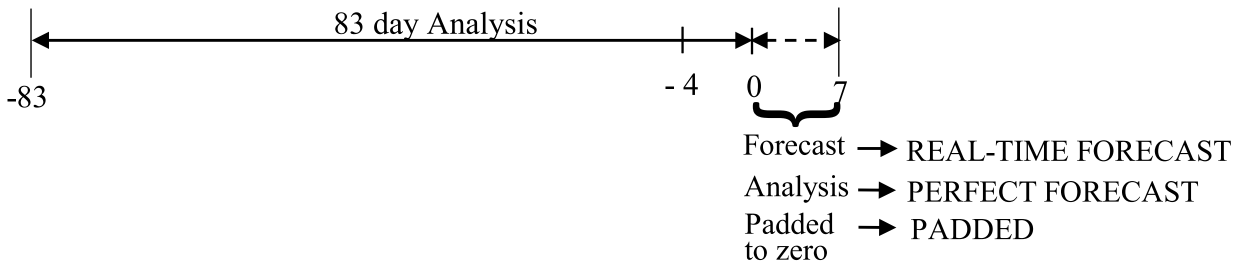


Figure 1. The theoretical horizontal structures of some of gravest equatorial wave modes in the resting atmosphere: the Kelvin wave, the *n*=0 westward-moving mixed Rossby-gravity (WMRG) the *n*=1 and 2 westward-moving Rossby (R1 and R2) waves. The meridional trapping scale y_0 has been taken to be 6° and the zonal wave number *k*=6. Vectors indicate horizontal wind. Colours shadings indicate divergence (10^{-6} s^{-1}) with convergence set to be positive. Colour contours lines are vorticity (10^{-6} s^{-1}) with blue lines for positive vorticity and red lines for negative vorticity: the contour interval is 0.6 starting from +/- 0.2 for Kelvin, WMRG and R1, and the contour interval for R2 is doubled. The amplitude of the wave is determined by setting the appropriate $(q_0, v_0, v_1,\ v_2)$ to 1.

(a) Real-time wave datasets



(b) Diagnostic analysis dataset

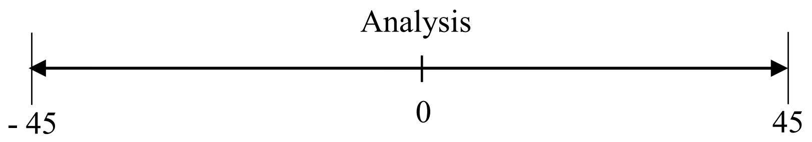


Figure 2. Schematic of procedures to create different wave datasets for developing and testing the real-time analysis technique. (a) Three 90-day time series used as input to the frequency filter. The REAL-TIME FORECAST data consists of 83 analyses and 7 days of forecast data. In the PERFECT FORECAST data, the last 7 days are replaced by analysis data. In the PADDED series the last 7 days are replaced by zeros. (b) The DIAGNOSTIC ANALYSIS uses data from a 90-day time-window, centred on each date, as input to the frequency filter.

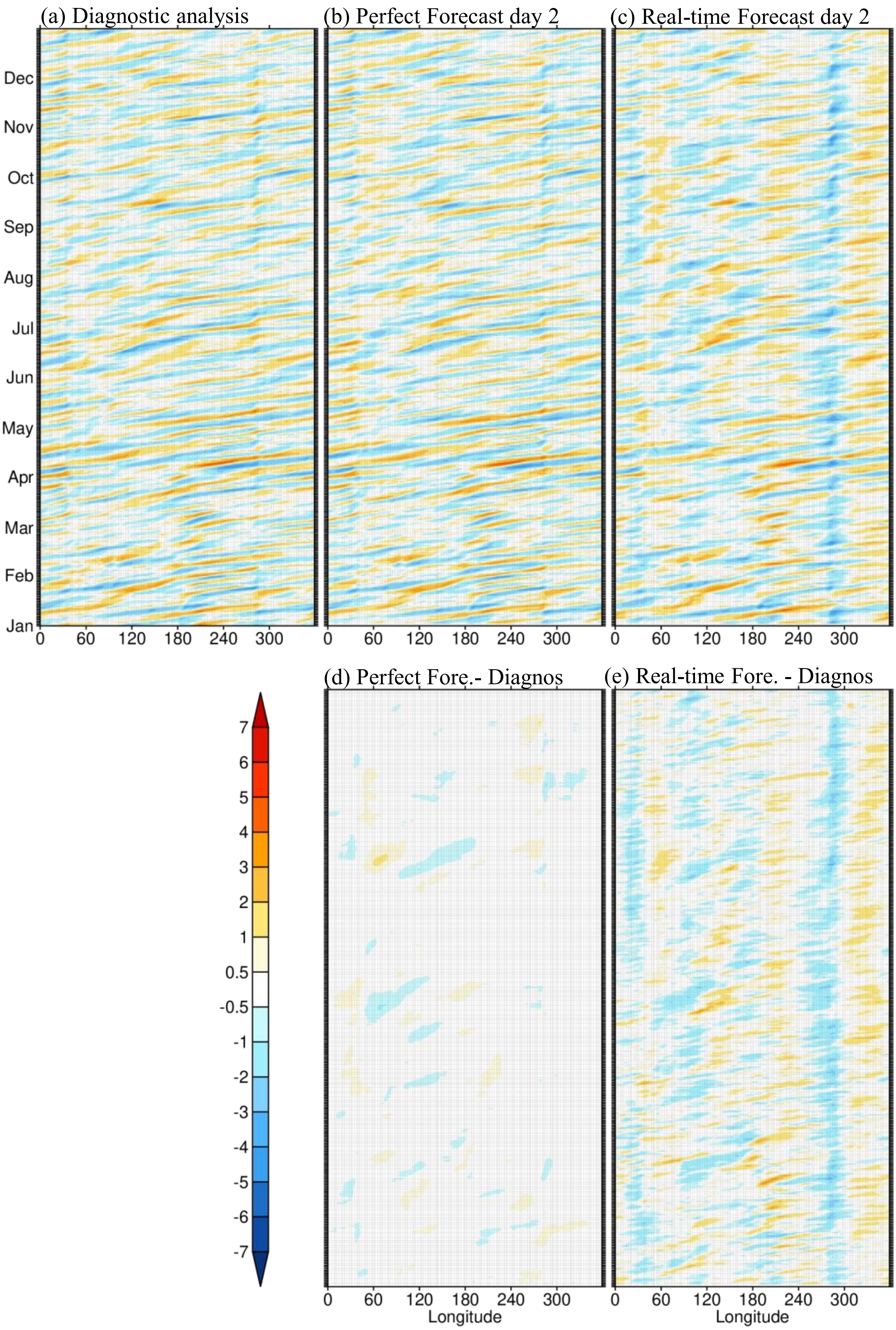


Figure 3 Longitude-time diagrams of 850-hPa Kelvin wave equatorial u (m s^{-1}) in 2016 identified from Met Office analysis data for (a) DIAGNOSTIC ANALYSIS, (b) PERFECT FORECAST dataset day T+2, (c) REAL-TIME FORECAST dataset day T+2, (d) difference (b minus a) between PERFECT FORECAST day T+2 and DIAGNOSTIC ANALYSIS. (e) Difference (c minus a) between REAL-TIME FORECAST day T+2 and DIAGNOSTIC ANALYSIS.

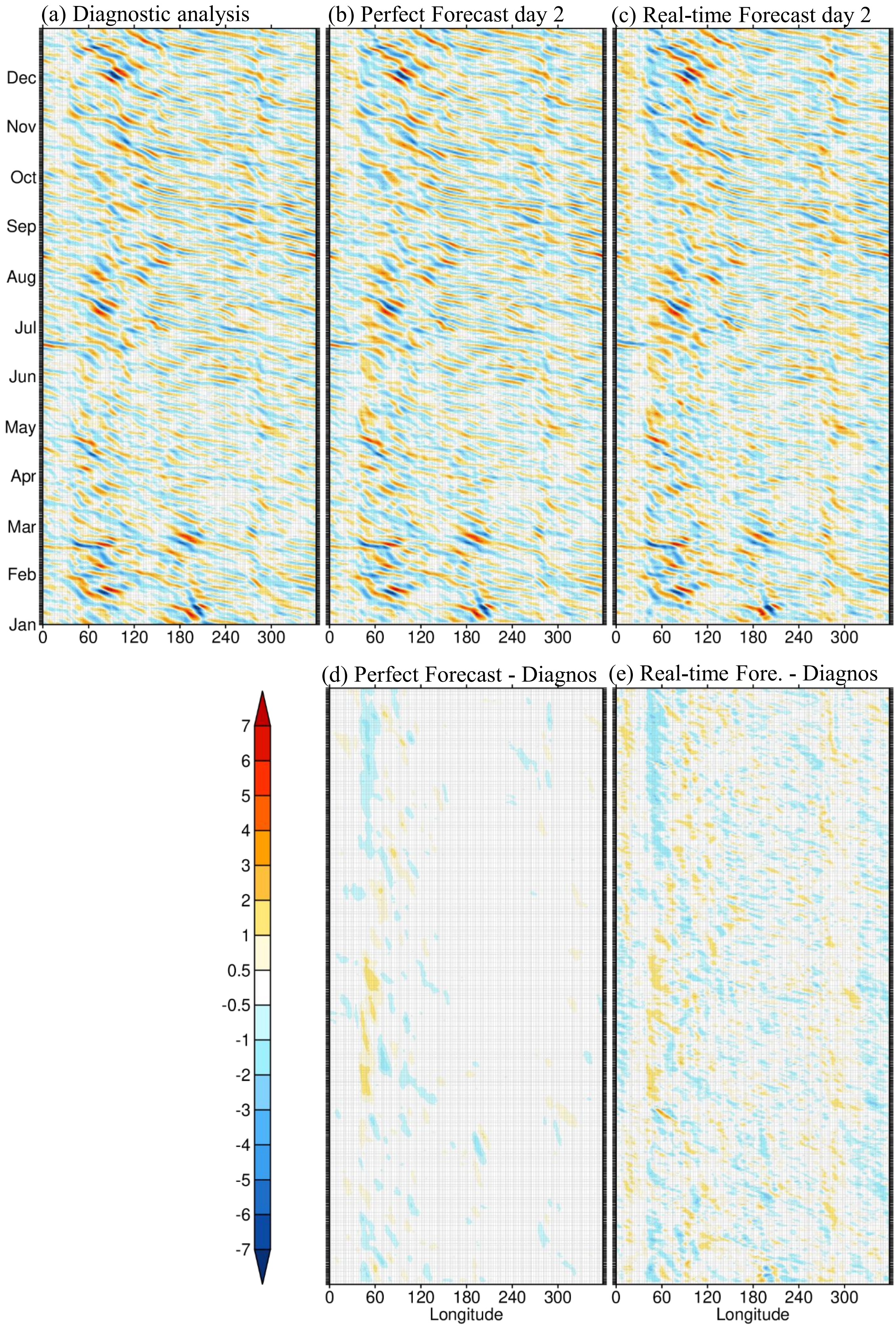


Fig.4 As for Fig.3 but for 850-hPa WMRG equatorial *v*.

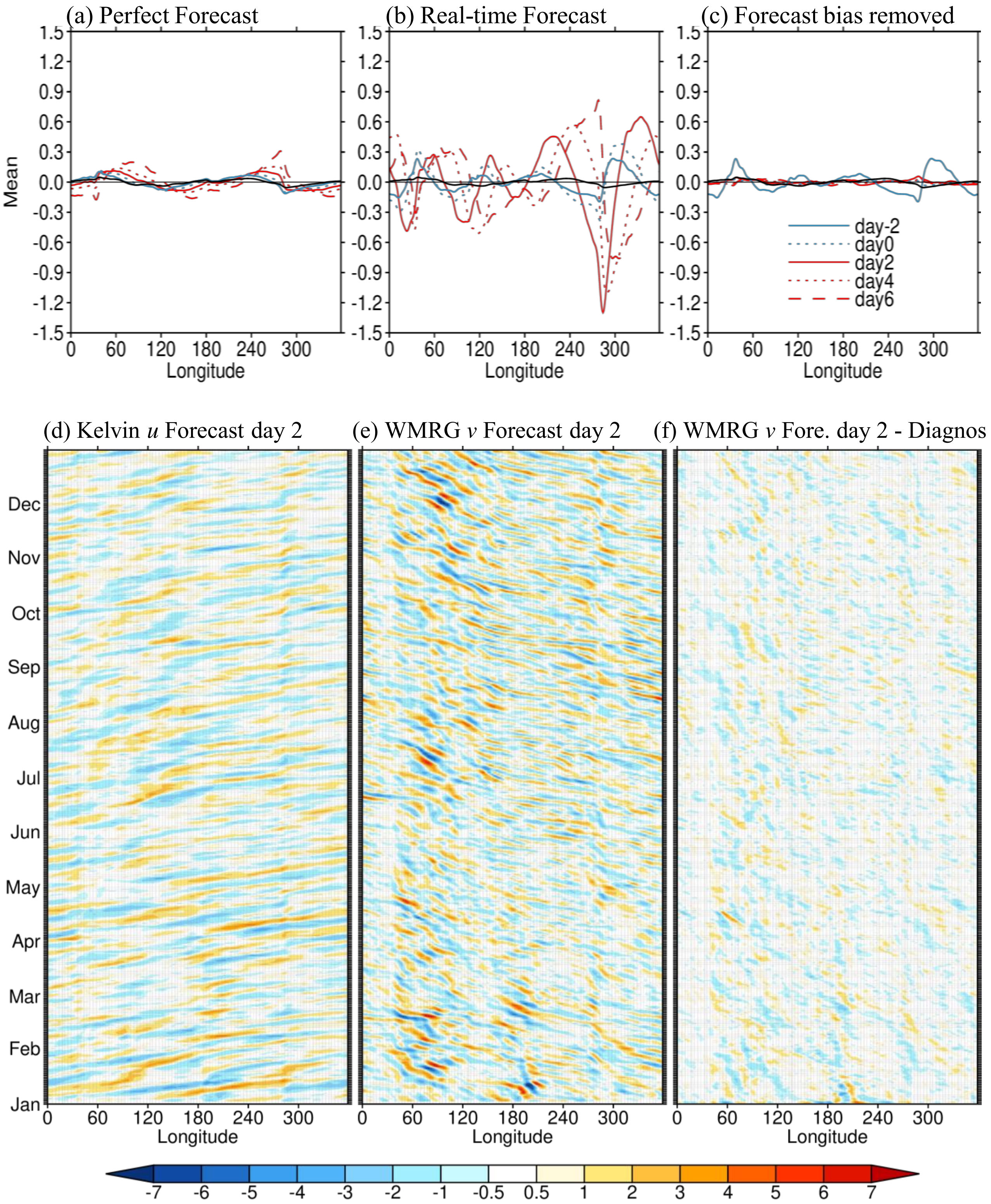


Figure 5 Top row: 12-month mean winds for Kelvin wave in 2016, in (a) PERFECT FORECAST dataset, (b) Real-time Forecast dataset and (c) REAL-TIME FORECAST dataset with previous 30-day mean removed for day T+0 to T+7. Black solid line in each panel is 12-month mean of DIAGNOSTIC ANALYSIS shown for comparison. (d)- (e) REAL-TIME FORECAST Kelvin wave *u* and WMRG *v* at day 2 with the bias correction. (f) Difference of WMRG *v* between REAL-TIME FORECAST day 2 and DIAGNOSTIC ANALYSIS. The label ‘day’ refers to lead time.

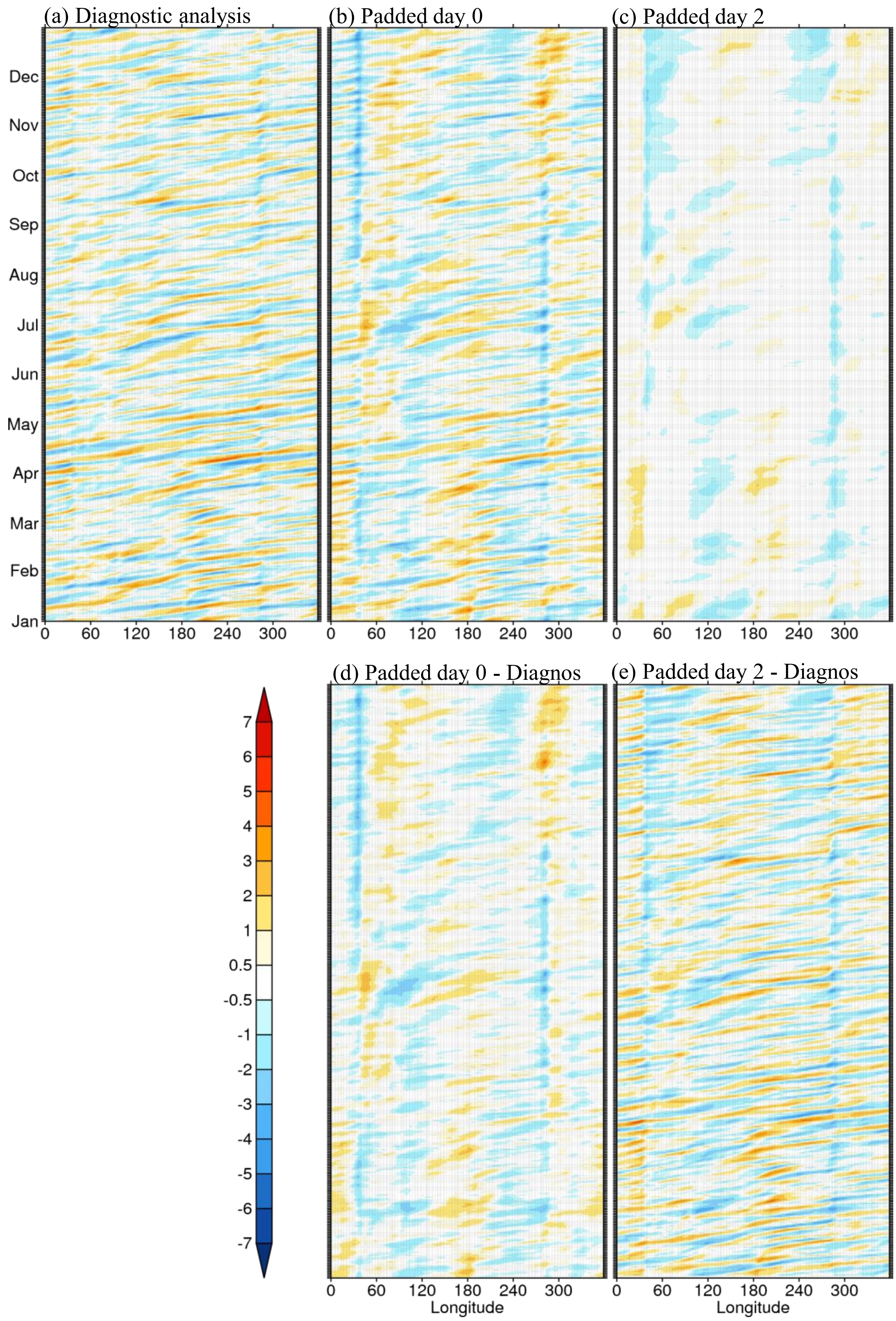


Fig.6. As for Fig.3 but for Kelvin wave identified from PADDED dataset at day 0 and day 2. (a) DIAGNOSTIC ANALYSIS, (b) PADDED at day 0, (c) PADDED at day 2, (d) Difference between PADDED day 0 and DIAGNOSTIC ANALYSIS. (e) Difference between PADDED day 2 and DIAGNOSTIC ANALYSIS.

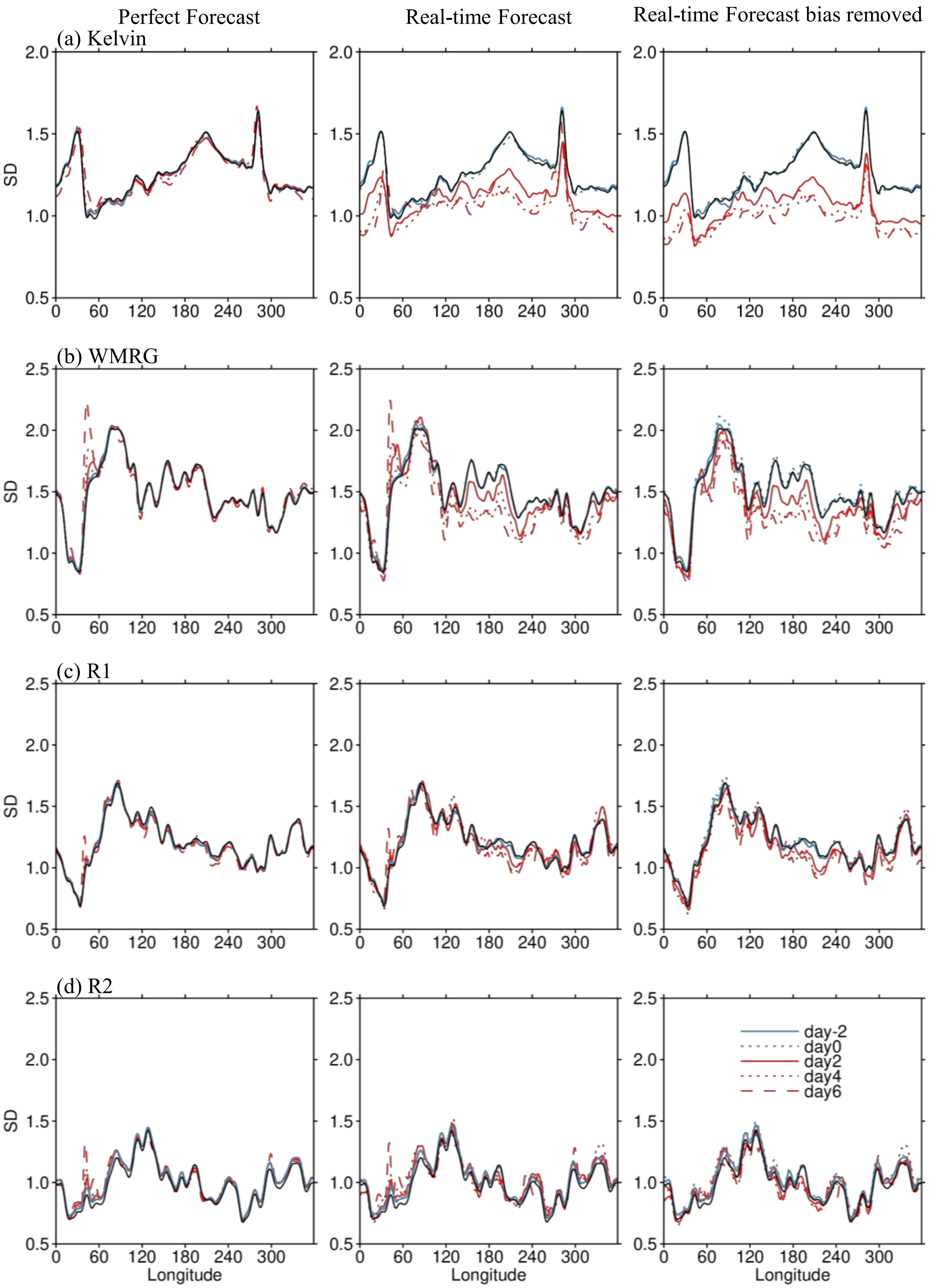


Fig.7 Standard deviations of winds at 850 hPa with respect to the 2016 time-mean for (a) Kelvin wave, (b) WMRG, (c) R1 and (d) R2, in (left) PERFECT FORECAST dataset, (middle) REAL-TIME FORECAST dataset and (right) REAL-TIME FORECAST dataset with previous 30-day mean removed. Black solid line in each panel is standard deviation of DIAGNOSTIC ANALYSIS shown for comparison. The label ‘day’ refers to lead time.

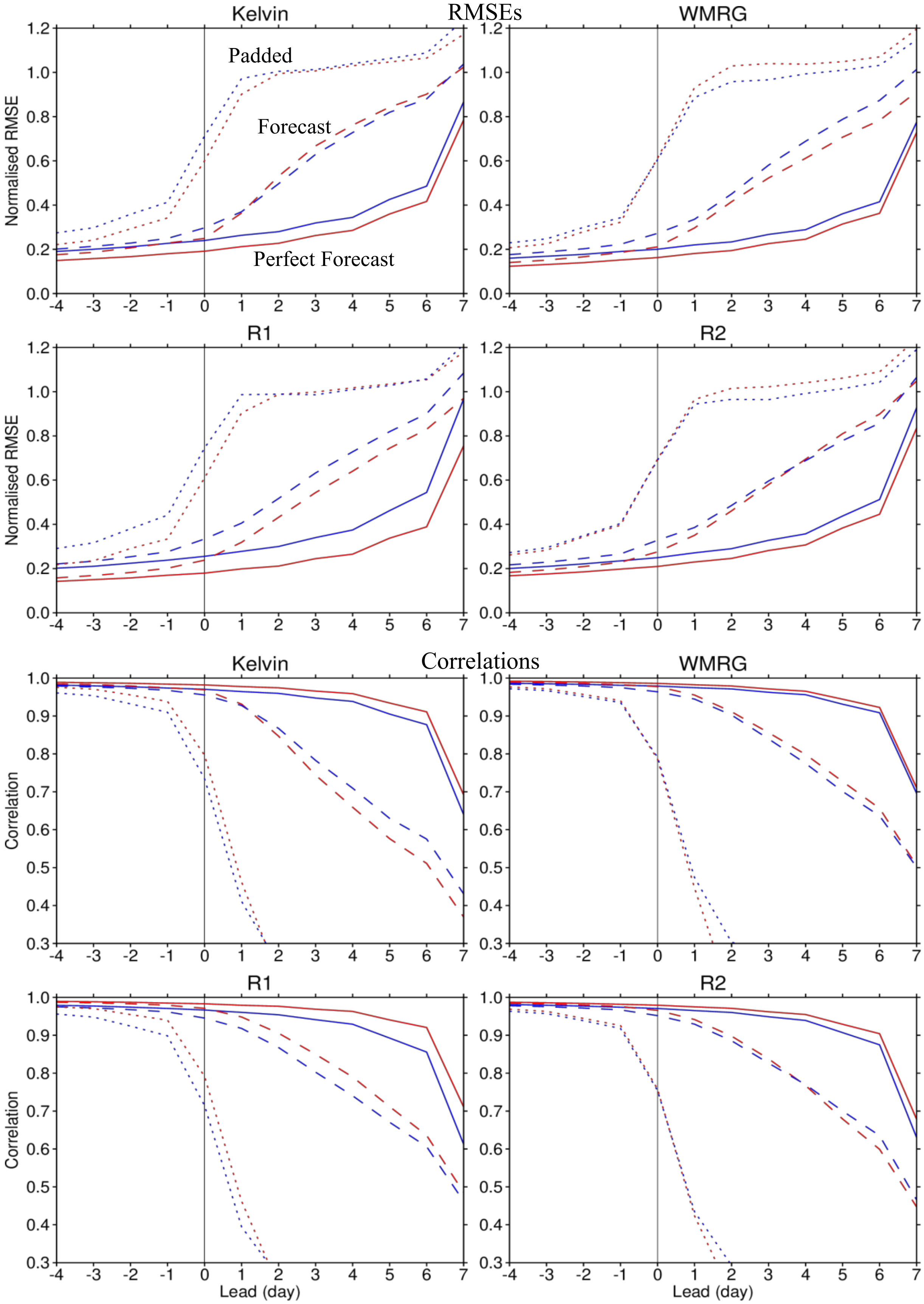
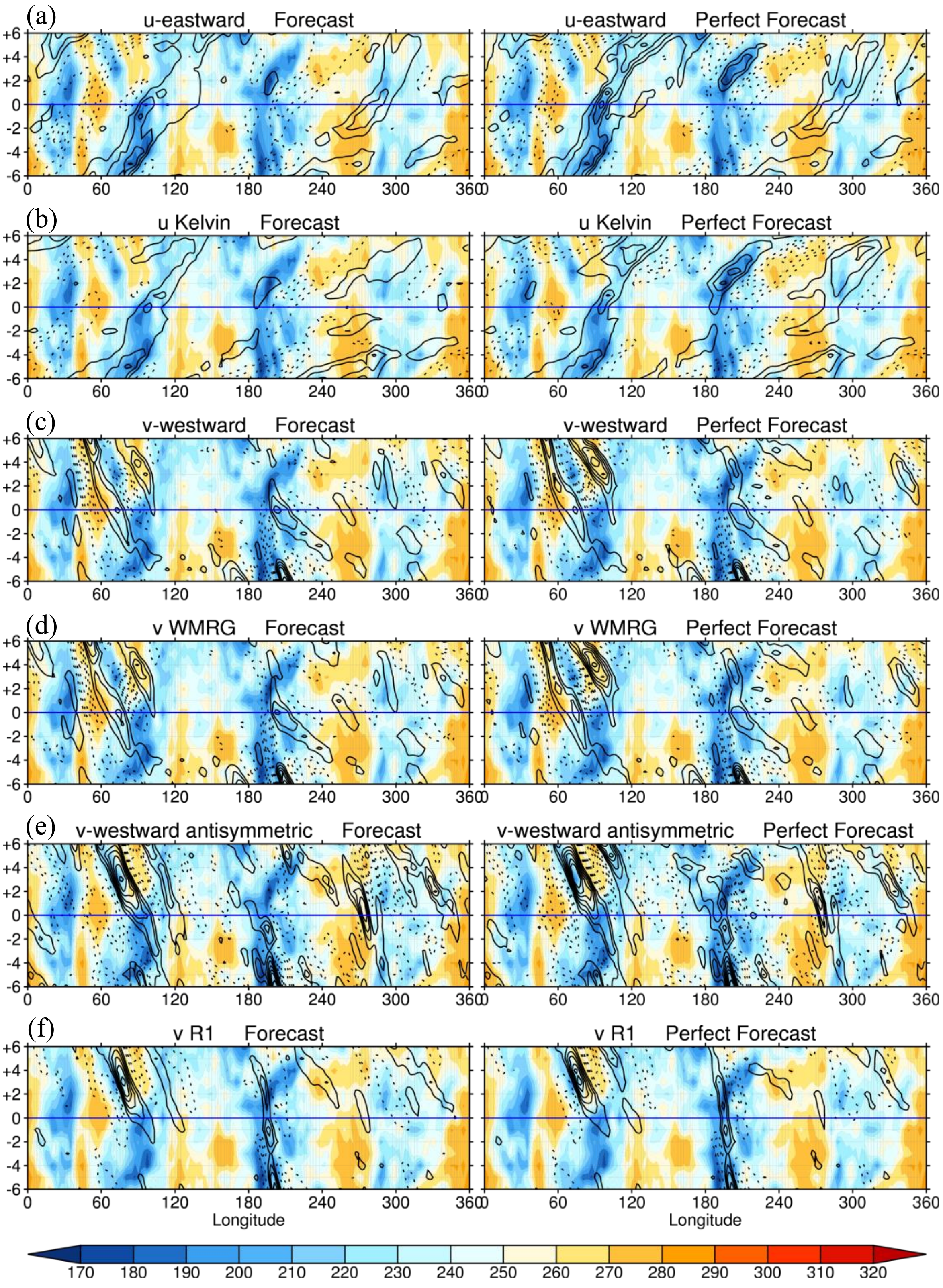


Figure 8 RMSEs normalised by the standard deviations from the DIAGNOSTIC ANALYSIS (top two rows) and correlations (bottom two rows) of waves at 850 hPa (red) and 200 hPa (blue) in 2016. PERFECT FORECAST dataset (solid), REAL-TIME FORECAST dataset (dashed) and PADDED dataset (dotted) all referenced to the DIAGNOSTIC ANALYSIS.



(a) u-eastward Forecast; u-eastward Perfect Forecast

(b) u Kelvin Forecast; u Kelvin Perfect Forecast

(c) v-westward Forecast; v-westward Perfect Forecast

(d) v WMRG Forecast; v WMRG Perfect Forecast

(e) v-westward antisymmetric Forecast; v-westward antisymmetric Perfect Forecast

(f) v R1 Forecast; v R1 Perfect Forecast

Figure 9 Waves forecast on 17 Jan 2016, indicated as ‘day 0’ for (left) REAL-TIME FORECAST and (right) PERFECT FORECAST, for NOAA OLR (colour) and various winds (contours) averaged over 10° N-10° S. (a) eastward-moving *u*, (b) Kelvin wave *u*, (c) westward-moving *v* symmetric about the equator. (d) WMRG *v*, (e) westward-moving *v* antisymmetric about the equator and (f) R1 *v* antisymmetric about the equator. The contour interval is 0.8 m s^{-1} for *u* and 1.0 m s^{-1} for *v*.

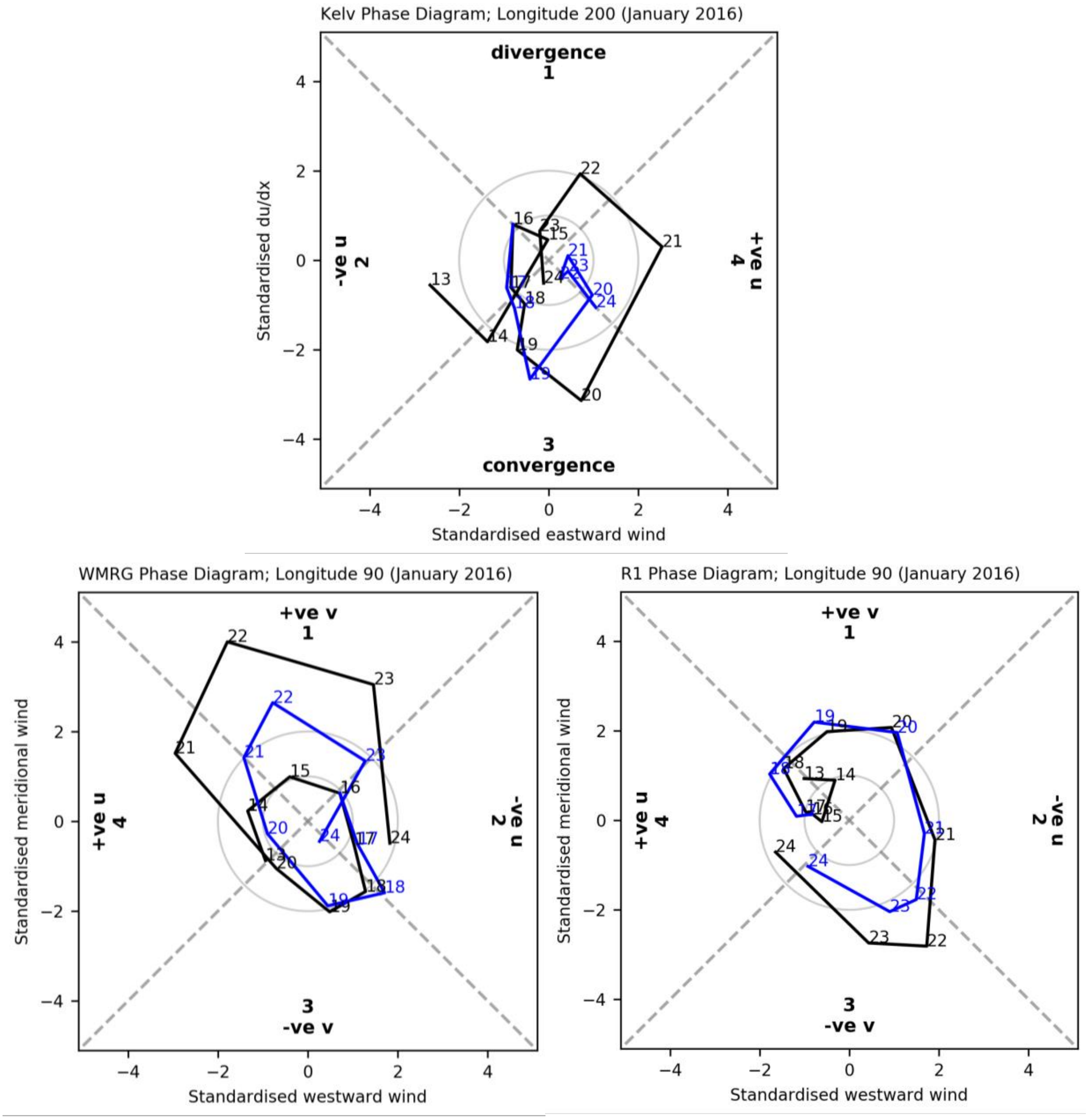


Figure 10 Phase and amplitude diagrams for Kelvin wave at 200° E, WMRG and R1 waves at 90° E. The latitude of the variables is indicated in Table 1. Blue lines are for the Met Office operational REAL_TIME FORECAST and black lines are for the PERFECT FORECAST. Forecast start date is 17th January 2016. The previous 4 days of REAL-TME FORECAST are also shown for the operational forecast. Quadrants 1-4 of wave phase are labelled in direction of propagation for each wave.

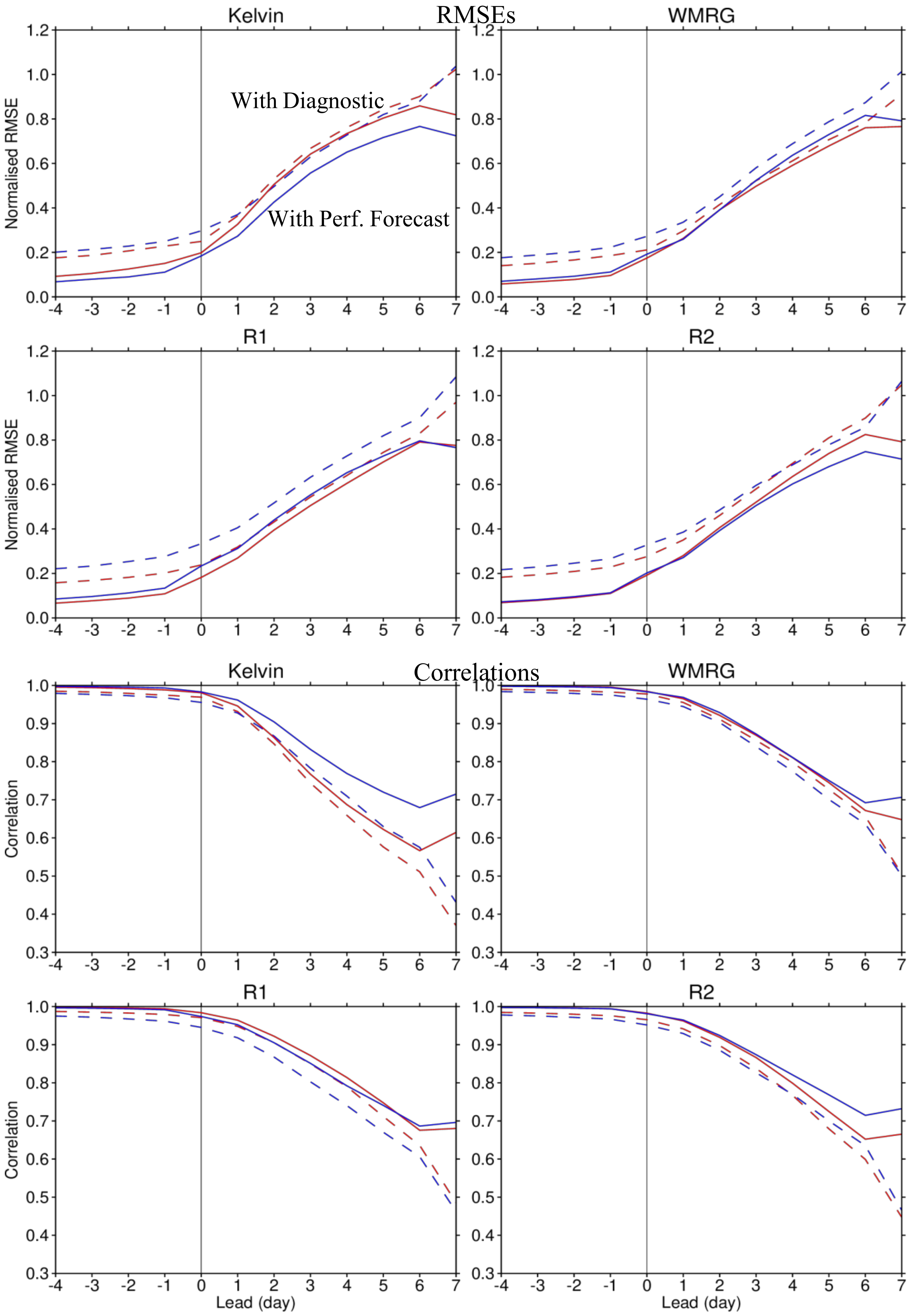


Figure 11 Normalised RMSEs (top two rows) and correlations (bottom two rows) of waves in REAL-TIME FORECAST, relative to PERFECT FORECAST dataset (solid), and relative to DIAGNOSTIC ANALYSIS (dashed) at 850 hPa (red) and 200 hPa (blue) in 2016.

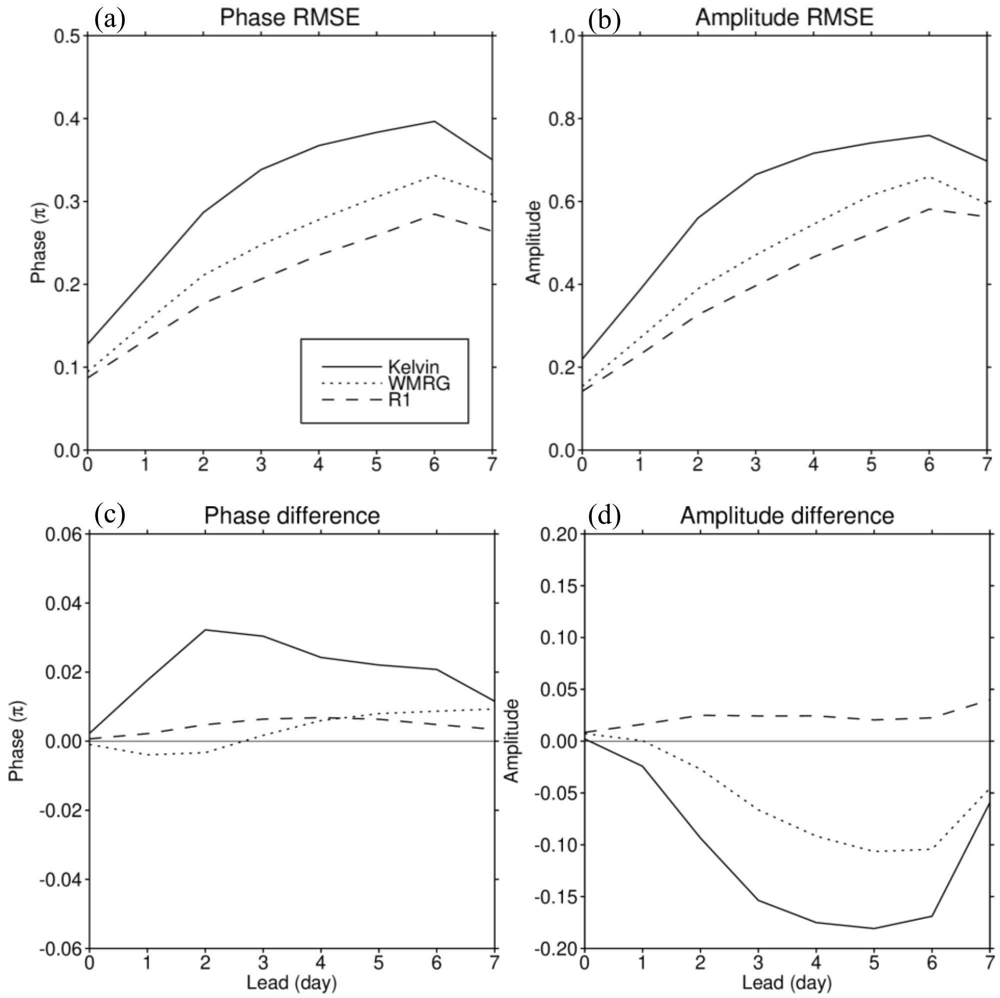


Figure 12 Forecast RMSEs in (a) phase and (b) amplitude at 850 hPa, relative to PERFECT FORECAST in 2015-2018. Mean difference of phase (c) and amplitude (d) between the REAL-TIME FORECAST and PERFECT FORECAST. Phase units are π radians.

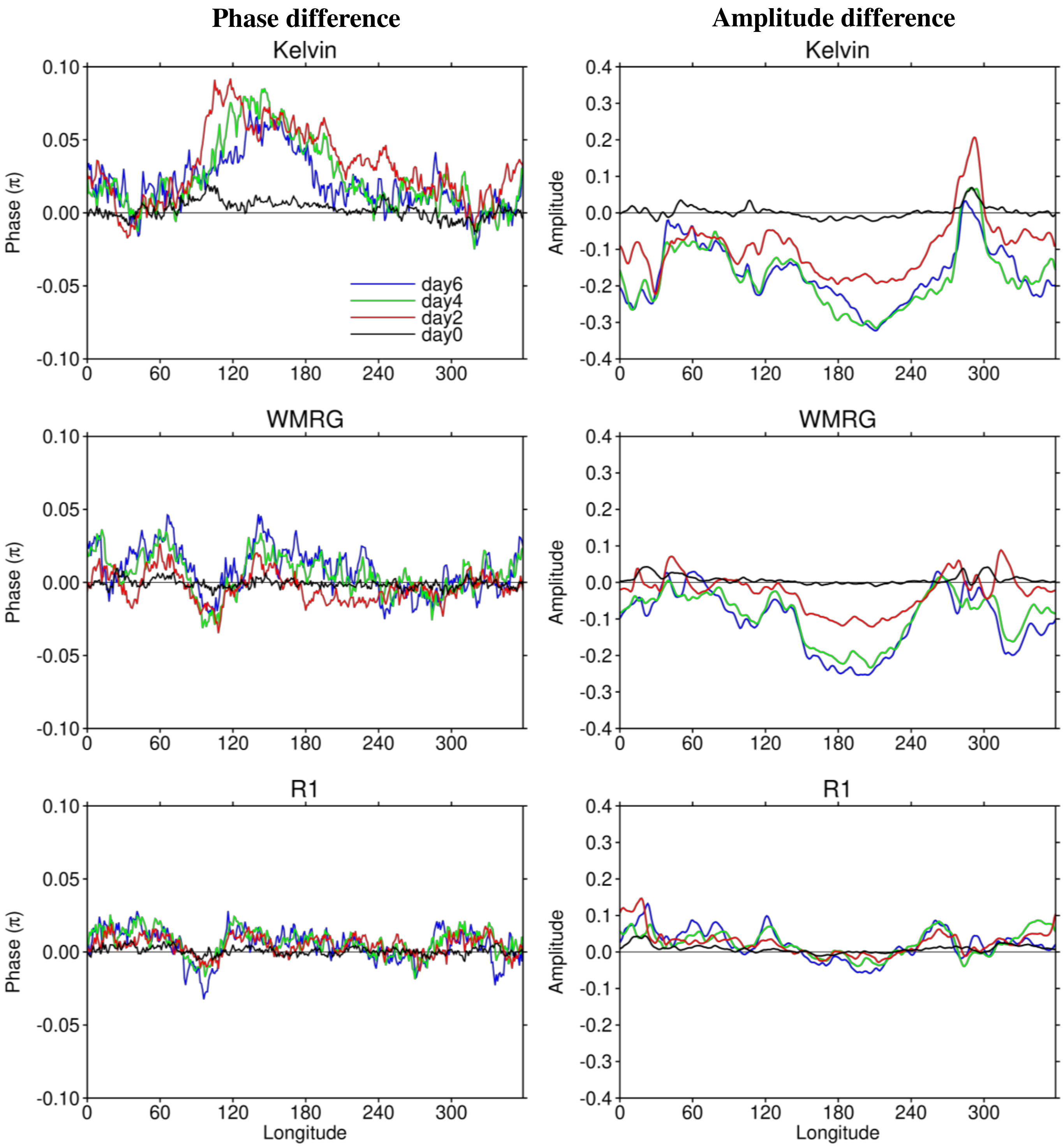


Figure 13 Difference in phase (left) and amplitude (right) between REAL-TIME FORECAST and PERFECT FORECAST in 2015-2018. For eastward-moving Kelvin wave, positive difference indicates faster phase speed, for westward-moving WMRG and R1 waves, positive difference indicates slower phase speed. The ‘day’ refers to lead time here.