Assessing the skill of high-impact weather forecasts in southern South America: a study on Cut-off Lows

~~M.~~ Belén Choquehuanca[1,2,3], Alejandro A. Godoy[4,5], Ramiro I. Saurral[1,2,3,6]

[1]Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales, Departamento de Ciencias de la Atmósfera y los Océanos, Buenos Aires, Argentina
[2]CONICET-Universidad de Buenos Aires. Centro de Investigaciones del Mar y la Atmósfera, Buenos Aires, Argentina
[3]CNRS-IRD-CONICET. Instituto Franco-Argentino para el Estudio del Clima y sus Impactos (IRL 3351 IFAECI), Buenos Aires, Argentina
[4]Servicio Meteorológico Nacional (SMN), Buenos Aires, Argentina
[5]Facultad de Ciencias Astronómicas y Geofísicas (UNLP), La Plata, Argentina
[6]Barcelona Supercomputing Center (BSC), Barcelona, Spain

Correspondence: ~~M.~~ Belén Choquehuanca. (belen.choquehuanca@cima.fcen.uba.ar)

Abstract. Cut-off Lows (COL) are mid-tropospheric cyclonic systems that frequently form over southern South America, where they can cause high-impact precipitation events. However, their prediction remains a challenging task, even in state-of-the-art numerical weather prediction systems. In this study, we assess the skill of the Global Ensemble Forecasting System (GEFS) in predicting COL formation and evolution over the South American region where the highest frequency and intensity of such events is observed. The target season is austral autumn (March to May), in which the frequency of these events maximizes. Results show that GEFS is skillful in predicting the onset of COLs up to 3 days ahead, even though forecasts initialized up to 7 days ahead may provide hints of COL formation. We also find that as the lead time increases, GEFS is affected by a systematic bias in which the forecast tracks lay to the west of their observed positions. Analysis of two case studies provide useful information on the mechanisms explaining the documented errors. ~~These are mainly related to the depth and the intensity of the cold core, which affect the thermodynamic instability patterns (thus shaping precipitation downstream) as well as the horizontal thermal advection which can act to reinforce or weaken the COLs.~~ These are mainly related to inaccuracies in forecasting the vertical structure, including their cold core and associated low-level circulation. These inaccuracies potentially affect thermodynamic instability patterns (thus shaping precipitation downstream) as well as the horizontal thermal advection which can act to reinforce or weaken the COLs. These results are expected to provide not only further insight into the physical processes at play in these forecasts, but also useful tools to be used in operational forecasting of these high-impact weather events over southern South America.

1 Introduction

Severe weather phenomena can significantly impact densely populated regions (e.g. Curtis et al., 2017; Newman and Noy, 2023; Sanuy et al., 2021). Over southern South America, these are frequently associated with heavy precipitation events triggered by low-pressure systems known as Cut-off Lows (COLs; Campetella and Possia 2007; Godoy et al., 2011a; Muñoz and Schultz, 2021). COLs are synoptic-scale weather systems that originate from elongated cold troughs in the middle troposphere, which subsequently detach ('cut off') from the main westerly current (Palmén and Newton, 1969). This segregation from the main flow explains the isolated and erratic behavior of these systems, which ~~suppose~~ pose a significant challenge in operational weather forecasting, even for state-of-the-art numerical weather prediction (NWP) systems (Muofhe et al., 2020; Yáñez-Morroni et al., 2018). Naturally, this can have an impact on the reliability of weather forecasts and early warnings which may be particularly relevant for southern South America considering the remarkable affectation from COLs (Godoy et al., 2011a).

Previous studies have focused on quantifying the explicit forecast errors associated with COLs in NWP systems. Gray et al. (2014) examined forecast ensembles from three operational forecast centers in the Northern Hemisphere and found that forecast errors were systematically larger in COL compared to no-COL events for the same prediction time. Similarly, Saucedo (2010) conducted an assessment of the prediction skill of the Global Forecast System (GFS) and Weather Research & Forecasting (WRF) models in southern South America for three COL events. His results indicated that forecast accuracy varies significantly depending on the individual COL cases and emphasized the need for an accurate representation of the COL center position during initialization to achieve better forecast results.

Other studies, such as those from Muofhe et al. (2020) and Binder et al. (2021), have linked errors in precipitation forecasts with inaccuracies in the location of the COL centers. In their evaluation of Météo-France forecasts, Binder et al. (2021) analyzed a single COL event and documented an eastward shift in both precipitation and COL position, primarily due to an initial underestimation of the COL intensity. Meanwhile, Muofhe et al. (2020) assessed the skill of the NWP model currently used operationally at the South African Weather Service to simulate five COL events. They observed variations in the predictive skill of COL-related precipitation across different development stages of the COLs, attributing these differences to inaccurate positioning of their centers. Moreover, studies by Bozkurt et al. (2016), Yáñez-Morroni et al. (2018) and Portmann et al. (2020) have underscored the influence of the COL-induced circulation on extreme precipitation events, emphasizing the complexity and challenge of predicting these phenomena. In particular, Portmann et al. (2020) noted that uncertainties in the COL genesis position substantially affect the vertical thermal structure of a surface cyclone development as well as its subsequent evolution.

While previous studies have examined the skill of NWP systems in forecasting COLs, they usually cover a short period of time and do not address a compound evaluation of positional and intensity errors. For instance, the recent paper by Lupo et al. (2023) has quantified biases in COL forecasts globally, but for the operational version of the GFS model in a 7-year period running from 2015 to 2022. In this context, there is a necessity to deepen our comprehension of COL predictive skill, given

the close linkage with heavy rainfall events. Our study tries to fill this gap, focusing on southern South America, a hotspot region for COL development (e.g., Reboita et al., 2010; Godoy, 2012 henceforth GD12; Pinheiro et al., 2017).

Our main goal is to assess the prediction skill of COLs in the National Centers for Environmental Prediction (NCEP)'s Global Ensemble Forecasting System (GEFS). This is achieved through quantifying forecast errors using an objective feature-tracking methodology which involves the identification and tracking of COLs along the forecast trajectories to produce a set of forecast versus observed COLs.

In this study, we specifically address three aspects of COLs: their onset time, their central position and their intensity. In particular, we seek to respond the following questions:

1. What is the temporal scale at which GEFS can reliably predict the initiation phase of COLs, and how precise are these forecasts?
2. After formation, can GEFS accurately predict the subsequent trajectories of the COLs?
3. Can errors in COL forecasts impact those of precipitation further downstream?

It should be noted that this study can be considered as a first step towards a full characterization of the physical mechanisms controlling the forecast skill of COLs and how the associated errors in state-of-the-art NWP systems are transferred into other associated variables such as precipitation, atmospheric instability and winds. The rest of the paper is organized as follows: the datasets and methodology are described in Section 2. The results on the forecast skill of the GEFS in both COL onset and their evolution stages are included in Section 3, followed by a summary and the concluding remarks in Section 4.

2 Data and methodology

2.1 The GEFS Reforecast dataset

Daily averages from the GEFS Reforecast version 2 dataset (Hamill et al., 2013) are used as a representative sample of the GEFS model for the purpose of this study. This dataset consists of 11 ensemble members - one control run alongside 10 perturbed members - and covers a prediction horizon of 16 days after initialization. During the first week, data is saved at 3-hourly intervals considering a horizontal resolution of T254 (roughly 40 km x 40 km at 40° latitude) and 42 vertical levels. ~~In the second week, the intervals increase to 6-hourly and the horizontal resolution decreases to T190 (around 54 km x 54 km at 40° latitude) with no changes in the number of vertical levels.~~ The GEFS Reforecast dataset can be freely downloaded from ftp://ftp.cdc.noaa.gov/Projects/Reforecast2, where the reforecasts have been saved at 1ºx1º horizontal resolution from the native resolution data ~~using bilinear interpolation with wgrib2 software~~. It is worth noting that for all calculations within the paper, we considered the ensemble mean as the basis for analysis and comparisons (i.e., no assessment is performed on individual ensemble members). To validate the GEFS skill, we use the fifth version of the ECMWF Reanalysis Dataset (ERA5; Hersbach et al. 2020) as a representation of the real-world conditions. The ERA5 data, with the original resolution of approximately 0.25° x 0.25°, were coarsened to the same resolution of the reforecast to ease comparison.

Our analysis focused on the forecast verification of atmospheric variables at the 300 hPa level. This level was chosen because it hosts both the largest frequencies and intensities of COLs within the Southern Hemisphere (e.g., Reboita et al., 2010; Pinheiro et al., 2021). To detect COLs, we analyzed the geopotential height and the zonal wind component at 300 hPa as well as the 300/850 thickness. We also evaluated other variables of interest such as the geopotential height at 850 hPa and the total accumulated precipitation to represent the lower-level circulation and related impacts of COLs.

2.2 Temporal domain and study area

The temporal domain of our study is based on the availability of reforecast data, ranging from 1985 to 2020. Specifically, we focus on the austral autumn season, covering the months of March, April, and May, which is the season with the highest frequency of COLs in South America (Reboita et. al., 2010; Pinheiro et. al., 2017; Muñoz et al., 2020). Regarding the spatial domain, we focused on the area of greatest occurrence of COLs, which encompasses the western side of southern South America (Reboita et al., 2010; Campetella and Possia, 2007; GD12). Specifically, we utilized the area situated between latitudes 37.6° and 29.9° S and longitudes 77.6° and 68.75° W, as illustrated in Figure 1. This region has been extensively studied in the past by GD12, who found that the COLs in this area are particularly strong and can often cross the Andes Mountain range, leading to conditions prone to high-impact weather events over the continent further downstream (Godoy et al., 2011a).

2.3 COL identification and tracking algorithm

~~The methodology used to build the COL dataset from GEFS and ERA5 data aligns with the approach outlined by GD12 and underpinned by a detection algorithm grounded in the conceptual framework put forth by Nieto et al. (2005). The methodology looks for local minima in the 300 hPa geopotential height field by simply comparing the local height with neighboring grid points under certain restrictions of size (i.e. number of surrounding points) and intensity. When a minimum is detected, a second requirement is it being associated with a cold core, for which the 850/300 hPa layer thickness is considered as a proxy of the mean layer temperature. Finally, points that successfully passed criteria 1 and 2 must also be accompanied by easterly winds to their polar side to be labeled as a COL.~~ The COLs dataset from GEFS and ERA5 is built following the approach outlined by GD12 and based on the conceptual framework of COL by Nieto et al. (2005). This conceptual model characterizes a COL as a closed cyclonic circulation isolated from the main westerly current and characterized by a cold core at mid-levels. To detect COLs, the tracking algorithm uses the geopotential height and the zonal wind component at 300 hPa as well as the 300/850 hPa thickness, following a series of steps to classify potential grid points as COLs: 1) In order to detect the closed circulation, the algorithm looks for local minima in the 300 hPa geopotential height field. It selects a grid point that is at least 5 geopotential meter (gpm) lower than six of the eight surrounding grid points to ensure a higher geopotential height. If this condition is not met, the algorithm checks that fourteen out of the sixteen surrounding grid points have a higher or equal value within 20 gpm of the candidate grid point. 2) To ensure that the system is isolated from the westerly current, the algorithm requires changes in wind direction in at least six grid points located south of the candidate grid point. 3) Finally, to confirm

the presence of a cold core, the algorithm employs the 850/300 hPa thickness as an indicator of temperature. It searches for a local minimum in thickness at the candidate point, following a procedure similar to the one used in the initial detection step. If a cold core is not found, the algorithm iterates through the eight surrounding grid points, accounting for possible displacements of the cold core relative to the geopotential minimum, as described in previous studies.

For validation purposes, we performed a visual inspection of the ERA5 COLs outputs. This visual check confirmed that each event aligns with the conceptual model proposed by Nieto et al. (2005). Additionally, we stipulated that each COL should be identifiable for a minimum of two days in the reanalysis data. A total of 34 events met all the established criteria.

~~Once we identified the COLs, we tracked them using the nearest neighbor method in the GEFS and ERA5 datasets and determined their trajectories. Only those forecasted COLs that fill on specific matching criteria were retained for the subsequent statistical analysis. First, the trajectories were considered matched if at least one point (corresponding to one day) matched in time along the life cycle of the individual systems. Second, we state that the distance between the predicted and the observed first point of the trajectories should not exceed 800 kilometers. This distance choice corresponds to the typical diameter of COL systems, which is between 600 and 1200 kilometers (Kentarchos and Davies, 1998). In agreement with Froude et al. (2007), our spatial matching approach focuses primarily on the starting point of the predicted trajectories rather than the entire trajectory. This emphasis is because, although the trajectories may initially closely match the observed trajectories, they are likely to diverge as the forecast lead time increases.~~ Following the identification of the COLs, we validated the GEFS COL dataset by comparing it with the ERA5 COL dataset. A GEFS COL was considered to correspond to the same system as in the ERA5 COL dataset if their initial positions and respective trajectories satisfied predefined spatial and temporal criteria. The forecasted COL trajectories that met these criteria were used to generate diagnostics, quantifying errors in predicted positions, intensities, and other properties of the COLs. The spatial criterion required that the distance between the forecasted and reanalysis trajectories did not exceed 800 kilometers — this threshold was chosen based on the typical diameter of COL systems, which ranges between 600 and 1200 kilometers (Kentarchos and Davies, 1998). Notably, our spatial criterion primarily focuses on the initial segment of the forecast trajectories rather than the entire track, consistent with the methodology of Froude et al. (2007). This approach is justified by the expectation that forecast accuracy is generally higher at the start of the trajectory, where GEFS trajectories are likely to be more closely aligned with their ERA5 counterparts. Regarding the temporal criterion, a match was considered valid if at least one point along the system's life cycle coincided in time (i.e., within a 24-hour period).

2.4 Verification metrics

For the quantification of the model skill, we used a Lagrangian perspective to derive error statistics. This methodology has been previously employed to build position and intensity error statistics in previous investigations on tropical and extratropical cyclones such as in Froude et al. (2007) and Hamill et al. (2011). The validation metrics used in this study are sketched in Figure 2 and are as follows:

- Direct Positional Error (DPE): This metric is defined as the horizontal distance between the observed and forecast positions at the same forecast time.
- Cross-Track Error (CTE): This metric represents the component of DPE that is perpendicular to the observed track. It provides information on the bias to the left or right of the observed track.
- Along-Track Error (ATE): This metric represents the component of DPE that is along the observed track. It provides information on the directional bias along the track, indicative of whether the forecasts predict a faster or slower motion of the system compared to the reanalysis.

We adopted the convention that a positive (negative) value of CTE indicates a bias to the right (left) of the observed track, while a positive (negative) value of ATE indicates that the model has a fast (slow) bias in its forecast track. ~~forecast position is biased fast (slow)~~. It is important to note that CTE and ATE cannot be calculated for the first analyzed position of a COL since they depend on the existence of an observed position the day before the valid time. For a more detailed explanation of these metrics, see Heming (2017).

3 Results

As a first step to determine the temporal horizon at which the GEFS model can forecast COLs, we analyze the central position of the COLs and their intensity ~~(given by the Laplacian of the geopotential height field)~~. The intensity of COLs is defined by the maximum value of the Laplacian of the geopotential height field, where this maximum corresponds to the location of the COLs center. ~~We show results only for the seven days before the observed onset stage of COLs events since no COLs were detected beyond this period in the preliminary analysis.~~ We present results for forecasts initialized up to seven days prior to the observed onset of COL events, as the preliminary analysis indicated that no COLs were forecasted beyond this lead time. It should be noted that hereafter "onset stage" or "onset" of the COL refers to the beginning of the segregation stage, also known as stage 2 of the COL life cycle as defined by Nieto et al. (2005). We organized each forecast into eight groups based on their initialization day, namely init 0, init 1, init 2, init 3, init 4, init 5, init 6, and init 7. Forecasts labeled as init 0 correspond to those initialized at the onset day of the COL, while forecasts labeled as init 1 to init 7 indicate forecasts initialized one to seven days before the onset day of the COL, respectively.

3.1 Predictive skill of COL onset time in GEFS

Figure 3 shows the percentage of detected COLs as a function of their initialization day, i.e. how many days in advance could these systems be forecasted in the GEFS dataset. During initializations closest to the onset days (init 0 to init 2), over 94% of the total events (32 out of 34 COLs) were accurately predicted by the GEFS. However, this accuracy decreases significantly from init 3 onwards: 71% at init 3, 56% at init 4 and down to only 9% at init 7. It is interesting to highlight, still, that the reforecasts were able to correctly predict most COLs on the same date they were observed, even when the initializations were farthest from the onset days (i.e. init 4 and init 5), indicating the accuracy of GEFS for predicting the timing of the events.

Figure 4 illustrates the quartile distribution of the DPE and intensity error in the GEFS model for the onset day of the COLs where each boxplot represents a different initialization day. The boxes represent the interquartile range (IQR), which comprises 50% of the error distribution, with the median value indicated by a bold black line. Initially, a gradual increase in the median of DPE can be observed as the number of days before the onset of COL increases (Fig. 4a). The DPE increase varies from 140 kilometers at the first initialization (init 0) to about 300 kilometers at init 3. At the same time, the IQR expands from 300 kilometers at init 1 to 900 kilometers at init 3, indicating a widening spread of DPE with increasing forecast time. In contrast, the median of the intensity error exhibits a negative trend: it decreases from -2.5 gpm/m^2 at init 1 to -8 gpm/m^2 at init 3, with an IQR that varies significantly with the day of initialization. For subsequent initializations (init 5 to init 7), we observe a continuous increase in DPE from 400 kilometers to approximately 600 kilometers, alongside a consistent negative trend in intensity errors, with values around -13.0 gpm/m^2. However, it is important to note that these results are based on a smaller sample size than previous initializations and caution should be exercised when generalizing these results.

Figure 5 shows eight polar scatter plots illustrating the errors in the position of the predicted COLs in comparison to the reanalysis, with each plot corresponding to a particular initialization day. During the early initializations, the GEFS exhibits errors contained within a radius of 3° (approximately 300 km) around the observed positions and shows no discernible directional deviation. This indicates that the position errors are randomly distributed and show no systematic bias, which is particularly clear up to init 2. Meanwhile ~~Conversely~~, initializations from init 3 to init 5 show a larger spread, with more points deviating significantly from the observed cyclone positions. While we detected a southward deviation, the zonal (i.e. east-west) behavior was less uniform, as init 3 showed a southern bias, init 4, a southwestern bias, and init 5, a slight southwestern deviation. This indicates overall a slight deviation towards the south (on average between 1° and 3°), even if there is no clear longitudinal ~~direction~~ bias. Forecasts initialized with a larger lead time showed a larger spread, partly due to a smaller number of predicted COLs, but also revealing a predominant southwesterly bias of the model.

3.2 Predictive skill of COL intensity and tracks in GEFS

In this section, we investigate whether there is any bias in predicting cyclone intensity, propagation speed, and trajectory. We focused on the forecasts initialized up to 3 days before the segregation date since the number of detected cases is significantly lower for forecasts initialized beyond that point (i.e. init 4 to init 7), as explained in Figure 3. ~~Also, considering that most COLs have a duration of 4–5 days or less (not shown), we restricted our analysis to forecast lead times within 3 days of the detection of the COLs in the ERA5 reanalysis.~~ Given that a preliminary study shows that a large portion of COLs in the study region have lifespans of 3–4 days or more, with nearly 80% lasting beyond 3 days (not shown), we have focused our analysis on forecast lead times of up to 3 days following the initial detection of these COLs in the ERA5 reanalysis.

Figure 6 shows the quartile distribution of track errors, including DPE, ATE, CTE and the intensity error between the GEFS and ERA5 trajectories for init 0 to init 3. Regarding DPE error, e~~E~~ach initialization shows similar sensitivity. For init~~: in the case of~~ init 1 and init 2 (Fig. 6b,c), errors increase from 166 to over 320 kilometers within two or three days after COL detection in the ERA5 reanalysis. The situation is similar for init 0 (Fig. 6a), where the error increases from 144 to over 275 kilometers

in the same period. Not surprisingly, init 3 (Fig. 6d) has the largest mean error, with a linear increase from 290 to 550 kilometers. As regards IQR, it shows a linear increase, indicating that the dispersion of the position errors increases along the cyclone forecast period.

Conversely, a negative trend is observed in the intensity error~~difference between the matched GEFS trajectories~~ and the corresponding ERA5 reanalysis trajectories~~(Figure 7)~~. The magnitude of the error for init 0 and init 1 (Fig. 6~~7~~a,b) initially increases from -2.0 to over -4.3 gpm/m^2 within two to three days after COL detection in the ERA5 reanalysis. For init 2 and init 3 (Fig. 6~~7~~c,d), however, a further ~~increase in~~ escalation of the error can be observed. While init 2 shows an increase in the magnitude of error from -4.9 to -11.68 gpm/m^2, init 3 shows an even more pronounced initial error of -8.14, which subsequently increases in their magnitude to -9.0 gpm/m^2. Regarding the dispersion of the error, it is noteworthy that init 1 and init 2 (Fig. 6~~7~~b,c) show a slightly positive trend, indicating an increase in the uncertainty of the predicted system intensity. In contrast, the last initialization (Fig. 6~~7~~d) shows a significantly larger dispersion and a more variable behavior during the analyzed period. Despite the observed variability, however, a trend towards greater dispersion is discernible.

~~Given that the DPE may stem from biases in either the translation speed of the COL (ATE) or from its direction of motion (CTE), as shown graphically in Fig. 2, we disaggregate their relative contributions in Figs. 8 and 9, respectively. In general, the ATE distribution exhibits a negative bias towards the later stages of the forecast tracks, except for init 2 (Fig. 8c) which shows slightly positive values.~~ The ATE distribution exhibits a negative bias towards the later stages of the forecast trajectories, except for init 2 and init 0 (Fig. 6c) which show slightly positive values. Both init 1 and init 3 (Fig.6~~8~~b,d) exhibit negative biases with median distances of around 200 and 300 kilometers, respectively. This negative bias in ATE may indicate that GEFS tends to underestimate the translational speeds of COL towards the latter stages of the forecast lead times. Regarding the CTE distribution (Fig. 6~~9~~), no clear bias is observed; however, there are some noticeable trends in different initializations. In particular, init 2 (Fig. 6~~9~~c) shows negative values at around 100 kilometers. On the other hand, init. 3 (Fig. 6~~9~~d) displays predominantly positive values, representing a poleward bias according to its definition.

3.3 Case studies

In this subsection, we focus on two COLs that exhibited very different levels of prediction performance during their onset stage (Fig. 4a). The first case study, from March-April 2013, is characterized by small DPE values, below the first quartile in Fig. 4a, indicative of a forecast with high accuracy in the GEFS dataset. In contrast, the second case study, from March 2019, was associated with remarkably larger DPE values, with errors ranging between the median and the third quartile. This represents a scenario in which the prediction has a suboptimal performance. It is important to note that the selection of the case studies was based also on the impact model errors had on the associated precipitation downstream. For the analysis of precipitation, we considered as the area of influence of the COLs approximately 7 degrees (about ~700 km radius) from the geopotential height minimum at 300 hPa. Before exploring the associated errors in the GEFS dataset, we provide a brief description of the synoptic environment around each COL during its segregation stage.

3.3.1 Case study 1: COL development on March 31st, 2013

On March 31st, 2013, a COL formed to the west of the Andes Mountains at 36°S and 75.5°W. Its lifespan lasted for six days, covering a distance of over 2,000 kilometers into the Atlantic Ocean (not shown). This event was associated with severe weather conditions which resulted in unprecedented flash floods in the region, leading to loss of lives, significant infrastructural damage and economic losses of USD 1.3 billion (Pink, 2018).

During the segregation phase of the COL, the main atmospheric features included an amplified ridge upstream ~~downstream~~ of the system, the presence of two jet streaks - one to the north and one to the south of the COL - and a well-defined cold-core in the middle levels (Fig.7a,c)~~, and a cyclonic system off the central coast of Chile at lower levels.~~ The COL extended towards the lower troposphere where a closed cyclonic circulation can be observed, as indicated by the closed circulation at 850 hPa, directly beneath the COL at 300 hPa (Fig. 7c). ~~The circulation of the cyclonic system (Fig 10c) fostered cold air advection underneath the COL center which helped to sustain and intensify the COL itself (not shown).~~ Regarding to the precipitation field, ~~During this its early development stages,~~ this COL led to high amounts of ~~record breaking~~ rainfall of over 25 mm per day with peaks in excess of 50 mm in certain areas over south-central South America (Fig. 7~~10~~b).

Forecast-wise, it is found that the location of the COL formation was accurately predicted 1 and 3 days ahead and even 5 days ahead with a bias of less than 200 kilometers northwest of its observed position (init 1,~~ and~~ init 3 and init 5 ; second, ~~and~~ third and fourth rows in Fig. 7~~10~~). However, these~~, but both~~ initializations underestimated its intensity by -6 gpm/m^2, -11 gpm/m^2 and -14 gpm/m^2 in init 1, ~~and~~ init 3 and init 5, respectively. The GEFS model accurately predicted the strength and extent of the upper-level strong winds associated with the COL (jet-split structure) and the upstream ridge of the COL for init 1, init 3 and init 5 (Fig. 7d, g, j). Particularly, during init 5 (Fig. 7j) it predicted better the intensity of the jet streak on the polar side of COL than the jet on the equatorial side. At mid-levels, the model successfully captured the cold core during init 1 and init 3, although with slightly less strength compared to ERA5 reanalysis. However, it failed to capture the cold core during init 5. ~~The GEFS accurately predicted the strength and extent of upper high winds associated with the COL (Fig. 10d,g).~~ Additionally, the cyclonic circulation at lower levels was displaced to the north relative to the observation (Fig. 7c,f,i), leading to the COL and lower-level cyclones being out of phase. This results in a different vertical structure in the forecasts with regard to the observations, which is consistent with the underestimation of the COLs intensity in the model. As discussed by Pinheiro et al. (2021), the intensity of the COL directly affects its vertical structure. In this case, the incorrect forecast position of the cyclone at low levels likely weakened the upward vertical motion and low-level moisture convergence, both of which are key factors for precipitation development. This implies a weaker vertical coupling in the forecast, resulting from the discrepancy in the intensity of the COL ~~However, it underestimated the strength of the cold core in middle levels and misplaced the location of the cyclonic circulation at lower levels, which shifted to the north of the observation site (Fig. 10f,i). This suggests that the vertical coupling with the COL was affected, potentially impacting the intensity of the system. Regarding precipitation forecasts, in both init 1 and init 3, the regions with significant rainfall were located southeast of their actual position and amounts were overall underestimated, particularly in init 3. On the other hand, init 5 exhibited even less skill, with intensity~~

~~and location errors of around 14 gpm/m^2 and 200 kilometers northwest of its observed position, respectively. The GEFS also encountered difficulties in predicting the jet split structure, inadequately represented the low-level circulation, and failed to capture the cold core at mid-levels, which naturally had an impact on the predicted precipitation amounts as well (Fig. 10k). Rainfall forecasts located the highest precipitation in the northeast of the country, outside the area affected by the COL system. This suggests that GEFS may not perform well in producing precipitation associated with COLs.~~ Regarding precipitation forecasts, GEFS performs well in predicting the location of precipitation associated with COL (with a slightly southeast bias), but it underestimates the amount of precipitation, especially during init 3 and init 5, with underestimations around 20 mm/day (Fig. 7h,k).

3.3.2 Case study 2: COL development on March 9th, 2019

On March 9th, 2019, another COL formed off the coast of Chile, at 33°S and 74°W (first row of Fig. 8~~11~~). This system was weaker than the one described in case 1~~the previous COL~~ and lasted four days. It caused some weak precipitation in south-central South America, but the amounts were lower than those associated with the first COL.

The synoptic environment during the segregation stage of this COL in the ERA5 reanalysis (first row of Fig. 8~~11~~) included an upper-level ridge with a NW-SE axis to the southwest of the COL, a split jet structure, a strong low-level cyclone positioned just beneath the COL center off the coast of Chile, and a small cold core at middle levels. Although this COL had a smaller structure than the first COL, the cyclonic system extended into the lower levels, as evidenced by the accompanying low-level cyclone identified in Fig. 8~~11~~c. In the precipitation field, two distinct maxima were identified: one located northeast of the analysis domain, associated with a decaying frontal zone in that area, which is linked to a surface cyclone positioned over the South Atlantic Ocean (not shown), and another maximum over western Argentina, directly related to the ascent zone east of the COL. The frontal system mentioned here is separated from the COL and its associated dynamics. ~~In the precipitation field, two distinct maxima were observed: one situated northeast of the domain of analysis, probably linked to a decaying frontal zone over that area (not shown), and another one over western Argentina related to the ascent zone at the east of the COL.~~ The subsequent validation of the GEFS forecast focuses only on this second feature as it was the one directly associated with (or triggered by) the COL.

The GEFS forecasts for March 9th, 2019 initialized 1, 3 and 5 days ahead are shown in Fig. 8~~11~~ (second to fourth rows). ~~In init 1 (Fig. 11, second row) the forecasted COL was approximately 15 gpm/m^2 shallower and located around 210 kilometers to the southeast compared to ERA5. Regarding the circulation at upper levels, GEFS predicted well the strength and extent of high winds associated with the COL.~~ Forecasts showed that the predicted position and intensity of the COL were consistently inaccurate across the three initializations. The COL was predicted to be shallower and displaced to the southeast, the system was shifted approximately 210 km and 430 km from its observed location for init 1 and init 3, and it could not be even captured in init 5. Meanwhile, the intensity was underestimated by approximately 15 to 17 gpm/m2. ~~However, the circulation at low and middle levels was less accurate. GEFS predicted the 850hPa cyclone to be located further north than expected, and at middle levels, it failed to represent the cold core. For init 3 (Fig. 11, third row), the forecasted COL was approximately 17~~

~~gpm/m^2 shallower and 430 kilometers southeast of its actual intensity and location. In this case, while GEFS predicted well the strength of the winds associated with COL, their position was predicted wrong, eastward compared to its actual position. At low and mid levels, the forecast was also inaccurate; the 850 hPa cyclone was weaker and displaced more northward than observed, and the strength of the cold trough at middle levels was underestimated and displaced towards the east.~~ With respect to the upper-level winds associated with the COL, the GEFS demonstrated a good skill in forecasting both their intensity and their spatial positioning, particularly in relation to jet streaks on the polar flank of the COL. However, the model exhibited notable challenges in accurately representing the cold-core structure at mid-levels, with a complete absence of this feature in init 5. At lower levels, the representation of the closed cyclone at 850 hPa was similarly problematic, with the system being consistently displaced northward and exhibiting weaker intensity than observations, especially in inits 3 and 5. ~~Regarding rainfall amounts, both initializations underestimated the rainfall within the ascent zone of the COL and predicted to be northeast of their observed position, over the central and northeastern parts of the country (Figure 11e, h). As for the last initialization (Fig. 11, fourth row), the model failed to predict the COL. GEFS displaced the upper circulation towards the southeast, including the jets and associated upper ridge. At low levels, GEFS also failed to predict the cyclone off the coast of Chile. Meanwhile, the thickness field showed a small, less intense cold trough, resulting in a lack of rainfall amounts over the zone influenced by the COL, as shown in Fig. 11k.~~ In terms of precipitation, GEFS underestimated rainfall amounts in all initializations and was not able to represent the observed precipitation at the lee side of the Andes mountains (Fig. 8e,h,k), displacing the predicted precipitation northeast of the observed location, particularly over central and northeastern Argentina. However, while the GEFS model generally underestimated rainfall amounts across all initializations, it is important to note that this behavior is expected given the model's relatively coarse resolution (1x1 degree), especially at the lee side of the Andes where the complex features of COLs usually difficult the simulation of precipitation even in high-resolution regional models like WRF (Yañez-Morroni et al., 2018).

Based on these results, a wrongly positioned and less intense COL can lead to a poor forecast of the vertical structure of the two case studies, including their cold core and associated low-level circulation~~the cold core~~, subsequently affecting dynamical processes such as horizontal temperature advection, thermodynamic instability, vorticity advection and associated ascent which are ingredients for precipitation production downstream. Such errors may be related to the inadequate representation of diabatic effects or interaction with the Andes Cordillera (Garreaud and Fuenzalida 2007). Even though the characterization of such processes are beyond the scope of this study, they will be addressed in future work.

4 Discussion and Conclusions

This study explored the prediction skill of cut-off lows (COLs) in the NCEP Global Ensemble Forecasting System (GEFS) with a focus on the region with the highest frequency of COL occurrence in South America during austral autumn (March to May). The analysis made use of a verification framework centered on the individual systems. These were identified and tracked using a feature-based approach applied to the 300 hPa level geopotential height as the primary variable.

The main conclusions can be built on the questions posed at the Introduction of the study:

- What is the temporal scale at which GEFS can reliably predict the initiation phase of COLs, and how precise are these forecasts?

  The GEFS model is highly accurate in predicting the start of the segregation stage of COLs up to three days in advance, but this accuracy drops significantly as the lead time increases beyond four days. The percentage of COLs detected by the model decreases to 56% and 29% for predictions initialized four and seven days ahead of the segregation, respectively. Our analysis also revealed that COL centers diverge by an approximate distance of 200 km relative to the observations up to three days in advance. However, this error increases to 600 kilometers for forecasts more than four days ahead. Also, it has been shown that forecasts initialized up to two days in advance have no directional deviations while forecasts initialized at least three days ahead of COL formation have a~~have a predominant~~ southerly bias. At the same time, the intensity errors show a consistent increase in magnitude, with values ranging from -2.5 gpm/m^2 in init 1 to approximately -13.0 gpm/m^2 at higher lead times.

- After formation, can GEFS accurately predict the subsequent trajectories of the COLs?

  From our results, we can conclude that the GEFS model has variable skill when forecasting the trajectories of COLs. Overall, errors in position increase from 200 to 400 kilometers in forecasts of one to two days of lead time. Within this time period, trajectories tend to be slower in comparison to the observed behavior. Even though this pattern of errors is also found for longer lead times, errors in predictions three days ahead increase substantially, and skill beyond four days is dramatically reduced. We can conclude that the trajectories of COLs can be relatively well predicted with lead times up to three days, and forecasts initialized beyond that threshold are significantly degraded and depict a poor representation of the actual paths. Intensity-wise, we found that GEFS forecasts are characterized by an increase in the magnitude of underestimation of COL intensity as the lead time increases.

- Can errors in COL forecasts impact those of precipitation further downstream?

  Although this study is based on only two case studies, our analysis suggests that the predictive skill of COLs, particularly regarding their formation location, intensity and trajectory, can influence precipitation forecasts downstream. ~~Even though in this study we have provided only partial evidence on this point from the analysis of two case studies, we can conclude that in these events the predictive skill of COLs (including their formation location, intensity and trajectory) had a significant impact on the precipitation forecasts downstream.~~ In particular, the errors in the location and depth of the COLs were linked to the mechanism sustaining these systems. In our case studies, the strength of the COLs cold core affects the thermodynamic instability patterns, potentially influencing vertical motion and precipitation formation downstream. This is sustained by the well-documented relationship between COLs cold-core and atmospheric instability response (Pinheiro et al., 2021; Hirota et al., 2016; Nieto et al., 2007; Porcu et al., 2007; Llasat et al., 2007; Palmen and Newton 1969), through which the dynamical ascent and atmospheric instability associated with the cold-core trigger and/or enhance precipitation events (Godoy et al., 2011; Nieto et al. 2007).~~For example, underestimating the strength of the cold core of COLs can significantly alter thermodynamic instability~~

~~patterns, affecting vertical motion and precipitation formation downstream.~~ Moreover, incorrectly forecasting the position of a low-level cyclonic system in association with COLs can significantly impact the vertical coupling of COLs, potentially influencing their intensity. This aligns well with Pinheiro et al. (2021), who suggested a possible relation between the intensity of COLs in South America and their vertical depth. These deficiencies, transferred into the higher levels, are able to shape the intensity of the system and, via this alteration, some of the mechanisms responsible for precipitation formation. As such, a weaker (stronger) COL will foster more (less) vorticity advection, resulting in favored (unfavored) ascent downstream. Therefore, predicted precipitation amounts will naturally be modulated by these errors (e.g. Saucedo, 2010).

Results from this study can be compared with similar recent studies. For instance, Lupo et al. (2023) have concluded that the operational GFS model has a systematic bias to move Southern Hemisphere troughs and COLs too quickly downstream, even though in our study region the identified bias is towards the west. (i.e. slower than observed). It should be noted, however, that the GEFS and the operational GFS share some common components but are different models, particularly regarding the horizontal resolution. As such, results from both studies are not directly comparable.

Regarding the case studies, previous authors analyzing the synoptic evolution and predictive skill of COLs in other regions of the world, such as Portman et al. (2022) and Moufhe et al. (2020), have concluded that a proper representation of the COL's vertical structure is crucial for an accurate prediction of these systems. Pinheiro et al. (2021) also argue that the intensity of the COLs affect the entire structure of these systems, and that errors in their intensity/position can easily affect their associated precipitation fields.

Although a detailed investigation of the physical mechanisms underlying these forecast errors was beyond the scope of this study, this issue is of great scientific importance for understanding the challenges typically found in predicting COLs. In this context, the GEFS bias, such as the westward bias and underestimation of intensity, likely arises from the model's inadequate representation of eddy-mean flow interactions, as explored by Nie et al. (2022, 2023) and Pinheiro et al. (2022). Moreover, in our study region, the positioning of the jet stream and the enhancement of transient wave activity over the South Pacific identified in previous work (GD12) are key to understanding these biases.Therefore, exploring the physical mechanisms underlying these forecast errors is essential. Future work exploring the simulation of jet streams and Rossby wave activity could provide crucial insights. Preliminary research has already shown that specific Rossby wave patterns preceding COLs can be predicted up to a week in advance, although with reduced confidence beyond that period (Choquehuanca et al., 2023).

It should be stressed once again that this study is proposed as a first step towards a full characterization of the physical processes responsible for COL formation, evolution and predictive skill in NWP systems. Several open questions remain, which will be addressed in future studies. Among them, it is unclear why the predicted trajectories are systematically slower than the observations. A negative correspondence between COL intensity and location was also observed in the GEFS dataset, suggesting that the most intense COLs seem to be associated with lower positional errors. However, the underlying mechanism sustaining such a relationship (if any) is not clear.

As a final note, future studies will dive into the relative contributions of COL intensity, location and speed on the resulting forecasted precipitation fields, as a deeper understanding of the interplay between these might bring useful information for operational weather predictions of high-impact events over southern South America.

Data availability. *All data is available from the authors upon request*

Author contributions. *BC prepared all analyses and the manuscript. AAG provided scientific advice throughout the whole project and assisted in setting up the tracking algorithm. RIS provided editing assistance, technical support and valuable suggestions for improving the manuscript.*

Competing interests. *The authors declare that they have no conflict of interest.*

Acknowledgements. *The authors would like to acknowledge the European Centre for Medium-Range Weather Forecasts (ECMWF) as well as NOAA's Earth System Research Laboratory (ESRL) for making available the ERA-Interim and GEFS datasets used in this study. Grant PICT2020-SerieA-03172 from the Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación (Scientific and Technological Ministry of Argentina) to the University of Buenos Aires partially funded this research.*

References

Awan, N.K., Formayer, H.: Cutoff low systems and their relevance to large-scale extreme precipitation in the European Alps. Theor. Appl. Climatol., 129, 149–158. https://doi.org/10.1007/s00704-016-1767-0, 2017

Bell, G. D., & Bosart, L. F.: A Case Study Diagnosis of the Formation of an Upper-Level Cutoff Cyclonic Circulation over the Eastern United States, Mon. Weather. Rev., 121(6), 1635-1655. doi: https://doi.org/10.1175/1520-0493(1993)121<1635:ACSDOT>2.0.CO;2, 1993.

Binder, H., Rivière, G., Arbogast, P., Maynard, K., Bosser, P., Joly, B., et al.: Dynamics of forecast-error growth along cut-off Sanchez and its consequence for the prediction of a high-impact weather event over southern France. Q J R Meteorol Soc, 147(739), 3263–3285. Available from: https://doi.org/10.1002/qj.4127, 2021.

Bozkurt, D., Rondanelli, R.,Garreaud, R., Arriagada A.: Impact of Warmer Eastern Tropical Pacific SST on the March 2015 Atacama Floods. Mon. Weather Rev. https://doi.org/https://doi.org/10.1175/MWR-D-16-0041.1, 2016.

Campetella, C.M., Possia, N.E.: Upper-level cut-off lows in southern South America. Meteorol. Atmos. Phys. 96, 181–191. https://doi.org/10.1007/s00703-006-0227-2, 2007.

Choquehuanca, B., Godoy, A., & Saurral, R. 2023: Evaluating cut-off lows forecast from the NCEP Global Ensemble Forecasting System (GEFS) in southern South America. Session 27: Hazards and Extreme events in The WCRP Open Science Conference (Kigali, Rwanda)

Curtis, S., Fair, A., Wistow, J., Val, D., Over, K.: Impact of extreme weather and climate change for health and social care systems. Environ. Health, 16, 128. https://doi.org/10.1186/s12940-017-0324-3, 2017

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. Roy. Meteor. Soc. https://doi.org/https://doi.org/10.1002/qj.828, 2011.

Froude, L. S. R., Bengtsson, L., & Hodges, K. I.: The Predictability of Extratropical Storm Tracks and the Sensitivity of Their Prediction to the Observing System, Mon. Weather Rev., 135(2), 315-333. doi: https://doi.org/10.1175/MWR3274.1, 2007a.

Froude, L. S. R., Bengtsson, L., & Hodges, K. I. The Prediction of Extratropical Storm Tracks by the ECMWF and NCEP Ensemble Prediction Systems, Mon. Weather Rev., 135(7), 2545-2567. doi: https://doi.org/10.1175/MWR3422.1, 2007b.

Funk, C., Peterson, P., Landsfeld, M. et al.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2, 150066. https://doi.org/10.1038/sdata.2015.66, 2015.

Garreaud, R., & Fuenzalida, H. A.: The Influence of the Andes on Cutoff Lows: A Modeling Study, Mon. Weather Rev., 135(4), 1596-1613. doi: https://doi.org/10.1175/MWR3350.1, 2007.

Godoy, A. A., Campetella M. C., Possia N. E.: Un caso de baja segregación en el sur de Sudamérica: Descripción del ciclo de vida y su relación con la precipitación. Revista Brasileira de Meteorologia. v.26, n.3, 491 - 502. doi: https://doi.org/10.1590/S0102-77862011000300014, 2011a.

Godoy, A. A.; Possia, N. E.; Campetella, C. M. y García Skabar, Y.: A cut-off low in southern South America: dynamic and thermodynamic processes. Revista Brasileira de Meteorología, vol. 26, n.4, pp. 503-514. doi: https://doi.org/10.1590/S0102-77862011000400001, 2011b.

Godoy, A. A.: Procesos dinámicos asociados a las bajas segregadas en el sur de Sudamérica. Ph.D. thesis. Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales, Argentina. https://hdl.handle.net/20.500.12110/tesis_n5602_Godoy , 2012.

Godoy, A. A.; Possia, N. E., Campetella, M. C., Seluchi, E. M.: Influence of the Andes over a cut-off low's life cycle. 9th ICSHMO Conference, 2009.

Gray, S.L., Dunning, C.M., Methven, J., Masato, G., Chagnon, J.M.: Systematic model forecast error in Rossby wave structure. Geophys. Res. Lett., https://doi.org/https://doi.org/10.1002/2014GL059282, 2014.

Hamill, T. M., Bates, G. T., Whitaker, J. S., Murray, D. R., Fiorino, M., Galarneau, T. J., Zhu, Y., Lapenta, W.: NOAA's Second-Generation Global Medium-Range Ensemble Reforecast Dataset. B. Am. Meteorol. Soc., https://doi.org/https://doi.org/10.1175/BAMS-D-12-00014.1, 2013.

Heming, J.T.: Tropical cyclone tracking and verification techniques for Met Office numerical weather prediction models. Met. Apps, 24: 1-8. https://doi.org/10.1002/met.1599, 2017.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horanyi, A.,Muñoz-Sabater, J. et al.: The ERA5 global reanalysis. Q. J. Roy. Meteor. Soc., 146, 1999–2049. Available from: https://doi.org/10.1002/qj.3803, 2020.

Huang, W., Li, C.: Contrasting Hydrodynamic Responses to Atmospheric Systems with Different Scales: Impact of Cold Fronts vs. That of a Hurricane. Journal of Marine Science and Engineering. https://doi.org/10.3390/jmse8120979, 2020.

Jeong, J., Lee, D., & Wang, C.: Impact of the Cold Pool on Mesoscale Convective System–Produced Extreme Rainfall over Southeastern South Korea: 7 July 2009, Mon. Weather Rev., 144(10), 3985-4006. doi: https://doi.org/10.1175/MWR-D-16-0131.1, 2016.

Kentarchos, A.S. and Davies, T.D.:A climatology of cut-off lows at 200 hPa in the Northern Hemisphere, 1990–1994. Int. J. Climatol., 18: 379-390. doi: https://doi.org/10.1002/(SICI)1097-0088(19980330)18:4<379::AID-JOC257>3.0.CO;2-F, 1998.

Lupo, K., Schwartz, C., Romine, G.: Displacement error characteristics of 500-hPa cutoff lows in operational GFS forecasts. Wea. Forecasting, 38, 1849-1871. https://doi.org/10.1175/WAF-D-22-0224.1 , 2023.

Muñoz, C., M. Schultz, D.: Cutoff Lows, Moisture Plumes, and Their Influence on Extreme-Precipitation Days in Central Chile. J. Appl. Meteorol. Clim., https://doi.org/https://doi.org/10.1175/JAMC-D-20-0135.1 , 2021.

Muñoz, C., Schultz, D., Vaughan, G.: A Midlatitude Climatology and Interannual Variability of 200- and 500-hPa Cut-Off Lows., J. Climate, https://doi.org/https://doi.org/10.1175/JCLI-D-19-0497.1, 2020.

Muofhe, T.P., Chikoore, H., Bopape, M.-J.M., Nethengwe, N.S., Ndarana, T., Rambuwani, G.T.: Forecasting Intense Cut-Off Lows in South Africa Using the 4.4 km Unified Model. Climate. https://doi.org/10.3390/cli8110129, 2020.

Ndarana, T., Waugh, D.: The link between cut-off lows and Rossby wave breaking in the Southern Hemisphere. Q. J. Roy. Meteor. Soc., https://doi.org/10.1002/qj.627, 2010.

Newman, R., Noy, I.: The global costs of extreme weather that are attributable to climate change. Nat. Commun., 14, 6103. https://doi.org/10.1038/s41467-023-41888-1, 2023.

Nie, Yu, Jie Wu, Jinqing Zuo, Hong-Li Ren, Adam Scaife, Nick Dunstone, and Steven Hardiman: Subseasonal Prediction of Early-summer Northeast Asian Cut-off Lows by BCC-CSM2-HR and GloSea5. *Adv. Atmos. Sci.* 40, 2127-2134. https://doi.org/10.1007/s00376-022-2197-9 , 2023.

Nie, Yu, Yang Zhang, Jinqing Zuo, Mengling Wang, Jie Wu and Ying Liu: Dynamical processes controlling the evolution of early-summer cut-off lows in Northeast Asia. *Clim. Dyn*., 60, 1103-1119., 2022

Nieto, R., Gimeno, L., Laura de la Torre, Ribera, P., Gallego, D., García-Herrera, R., José Agustín García, Nuñez, M., Redaño, A., Jerónimo Lorente: Climatological Features of Cutoff Low Systems in the Northern Hemisphere. J. Climate, https://doi.org/https://doi.org/10.1175/JCLI3386.1 , 2005.

Palmén, E.: Origin and Structure of High-Level Cyclones South of the: Maximum Westerlies. Tellus, 1: 22-31. doi: https://doi.org/10.1111/j.2153-3490.1949.tb01925.x, 1949.

Palmén, E., Newton, C.W.: Atmospheric Circulation Systems: Their structure and physical interpretation. New York, NY: Academic Press., 1969.

Pepler, A., Dowdy, A.: A Three-Dimensional Perspective on Extratropical Cyclone Impacts. J. Climate. https://doi.org/https://doi.org/10.1175/JCLI-D-19-0445.1, 2020.

Pinheiro, H., T. Ambrizzi, K. Hodges, M. Gan, K. Andrade, and J. Garcia: Are cut-off lows simulated better in CMIP6 compared to CMIP5? Climate Dyn., 59, 2117−2136, https://doi.org/10.1007/s00382-022-06200-9, 2022.

Pinheiro, H., Gan, M., Hodges, K.: Structure and evolution of intense austral cut-off lows. Q. J. Roy. Meteor. Soc., https://doi.org/https://doi.org/10.1002/qj.3900, 2021.

Pinheiro, H.R., Hodges, K.I., Gan, M.A.: Sensitivity of identifying cut-off lows in the Southern Hemisphere using multiple criteria: implications for numbers, seasonality and intensity. Clim. Dynam., 53, 6699–6713. https://doi.org/10.1007/s00382-019-04984-x, 2019.

Pinheiro, H.R., Hodges, K.I., Gan, M.A., Ferreira, N.J.: A new perspective of the climatological features of upper-level cut-off lows in the Southern Hemisphere. Clim. Dynam., 48, 541–559. https://doi.org/10.1007/s00382-016-3093-8, 2017.

Pink, R. M.: The Climate Change Crisis: Solutions and Adaptations for a Planet in Peril. Cham, Switzerland: Palgrave Macmillan, 2018.

Portmann, R., González-Alemán, J.J., Sprenger, M., Wernli, H.: How an uncertain short-wave perturbation on the North Atlantic wave guide affects the forecast of an intense Mediterranean cyclone (Medicane Zorbas). Weather and Climate Dynamics. https://doi.org/10.5194/wcd-1-597-2020, 2020.

Poveda, G., Espinoza, J.C., Zuluaga, M.D., Solman, S.A., Garreaud, R., van Oevelen, P.J.: High Impact Weather Events in the Andes. Front. Earth Sci., https://doi.org/10.3389/feart.2020.00162, 2020.

Reboita, M.S., Nieto, R., Gimeno, L., da Rocha, R.P., Ambrizzi, T., Garreaud, R., Krüger, L.F.: Climatological features of cutoff low systems in the Southern Hemisphere. J. Geophys. Res-Atmos., https://doi.org/https://doi.org/10.1029/2009JD013251, 2010.

Sanuy, M., Rigo, T., Jiménez, J., Llasat, M.: Classifying compound coastal storm and heavy rainfall events in the north-western Spanish Mediterranean. Hydrol. Earth Syst. Sci., 25, 3759-3781. https://doi.org/10.5194/hess-25-3759-2021, 2021.

Saucedo, M.: Contribución de bajas segregadas a la predictibilidad de la precipitación en el sur de Sudamérica: casos de estudio. Lic. thesis. Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales, Argentina. 117 pp., 2010.

Yáñez-Morroni, G., Gironás, J., Caneo, M., Delgado, R., Garreaud, R.: Using the Weather Research and Forecasting (WRF) Model for Precipitation Forecasting in an Andean Region with Complex Topography. Atmosphere. https://doi.org/10.3390/atmos9080304, 2018.

Figures

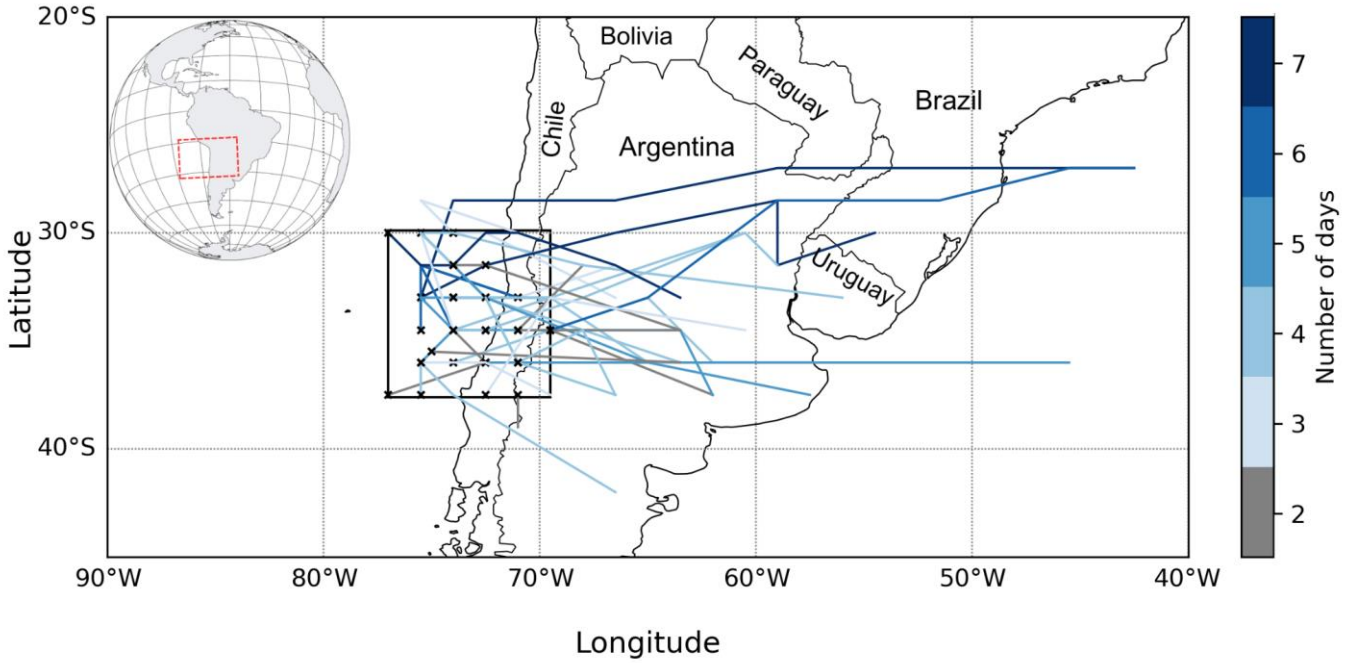


Figure 1: Spatial distribution of COLs in the region of highest COL frequency in southern South America from 1985 to 2020. Black crosses represent the start of trajectories of COLs detected in the study area (77.6°-68.75°W and 37.6°-29.9°S, solid black box) and lines represent their trajectories where colors represent the duration of each COL.

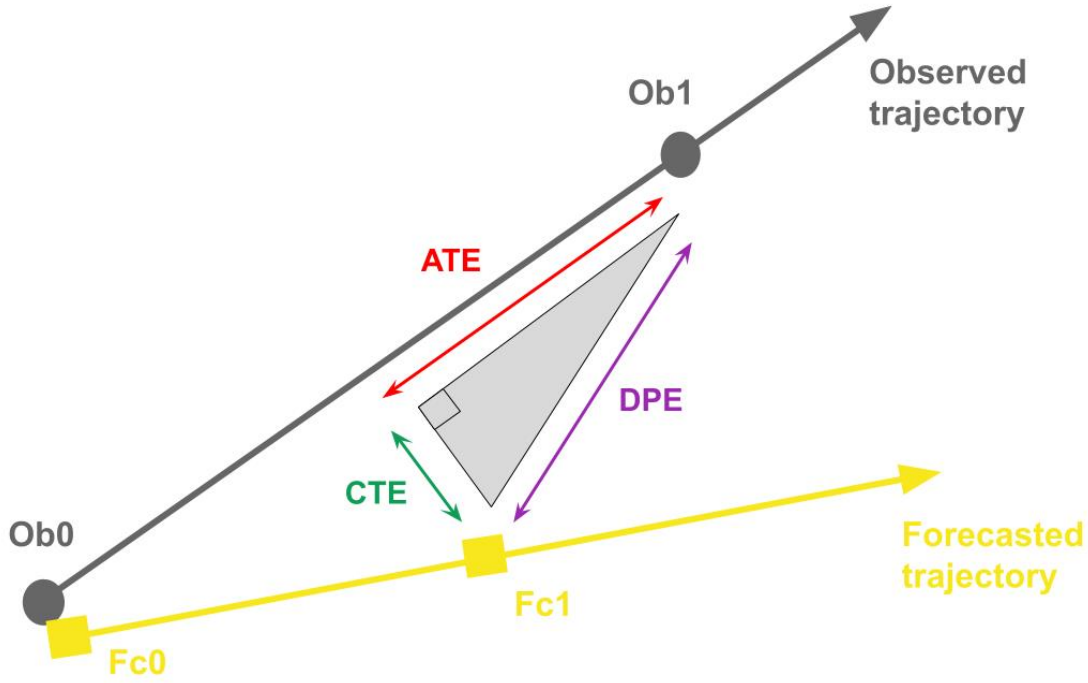


Figure 2: Measures of cyclone track forecast error: Direct Positional Error (DPE; violet arrow), Cross-Track Error (CTE; green arrow) and Along-Track Error (ATE; red arrow). Obs0 and Obs1 are observed positions at times 0 and 1, while Fc0 and Fc1 are their respective forecasted positions. The gray circles (yellow squares) represent the observations (the forecasts).

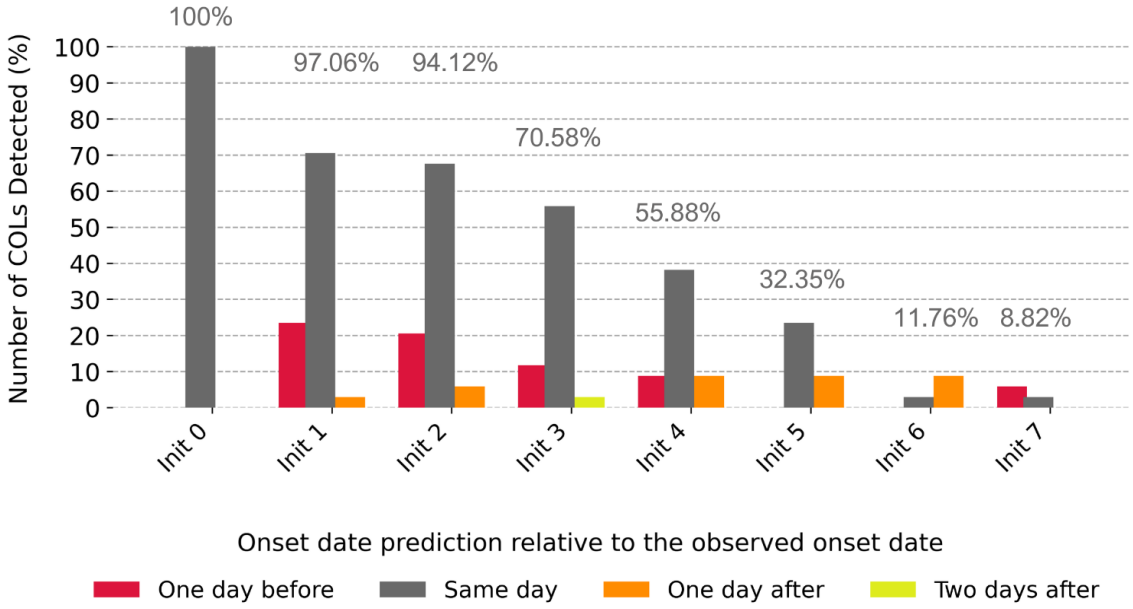


Figure 3: Percentage of forecasted COL initiations as a function of initializations, from init 0 (forecast initialized in the onset day) to init 7 (forecast initialized seven days before the onset of the COL). The red, gray, orange and yellow bars indicate the forecasted date of the onset day of COL relative to the observed date of onset day, from one day ahead of formation to two days after, respectively.

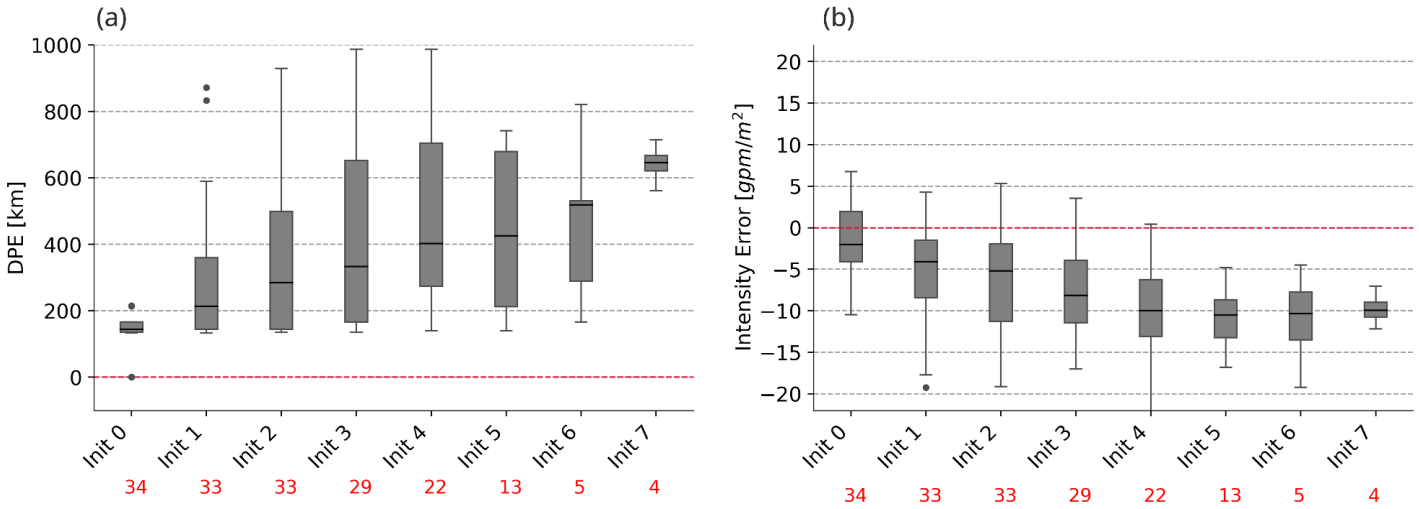
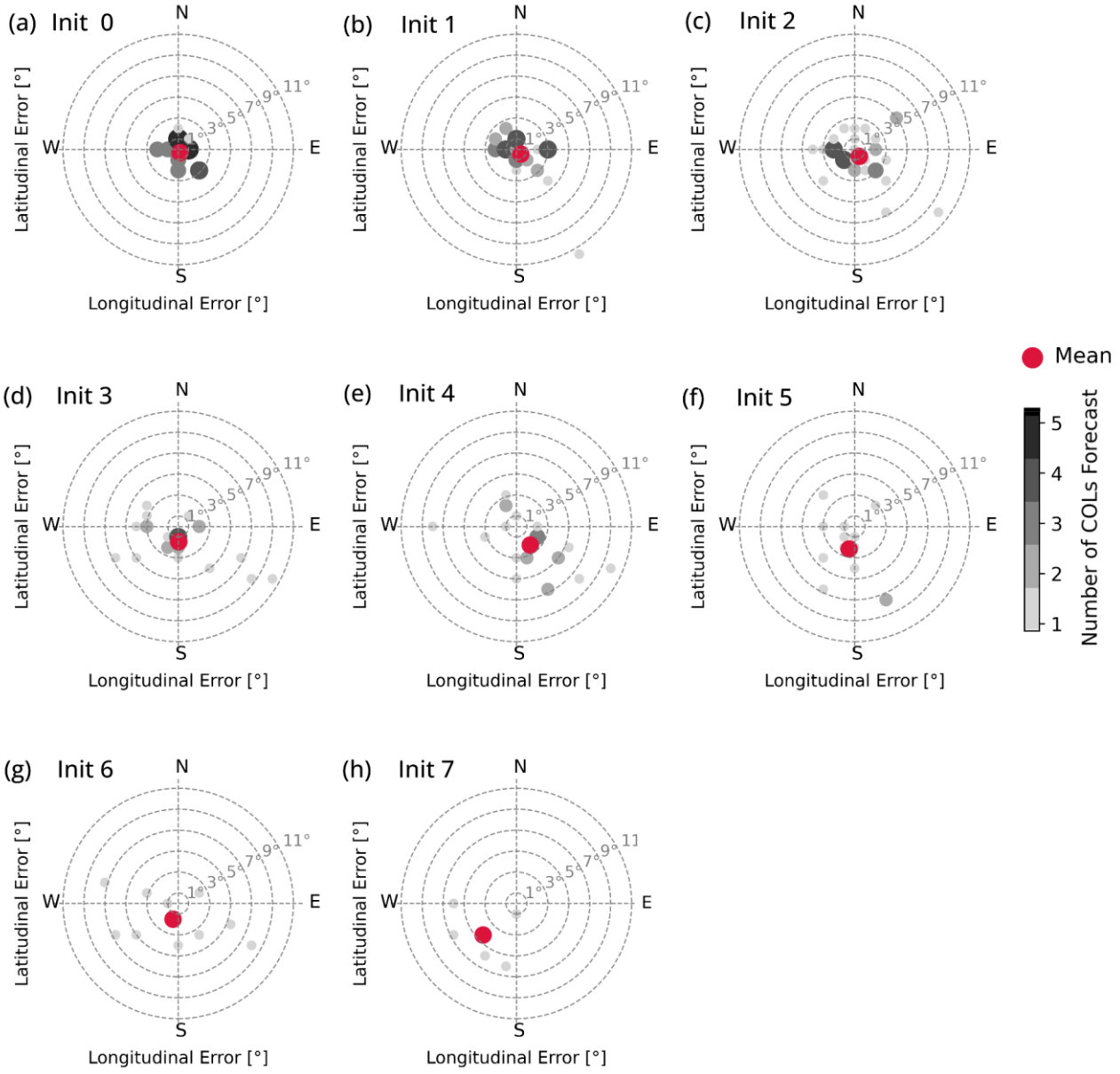
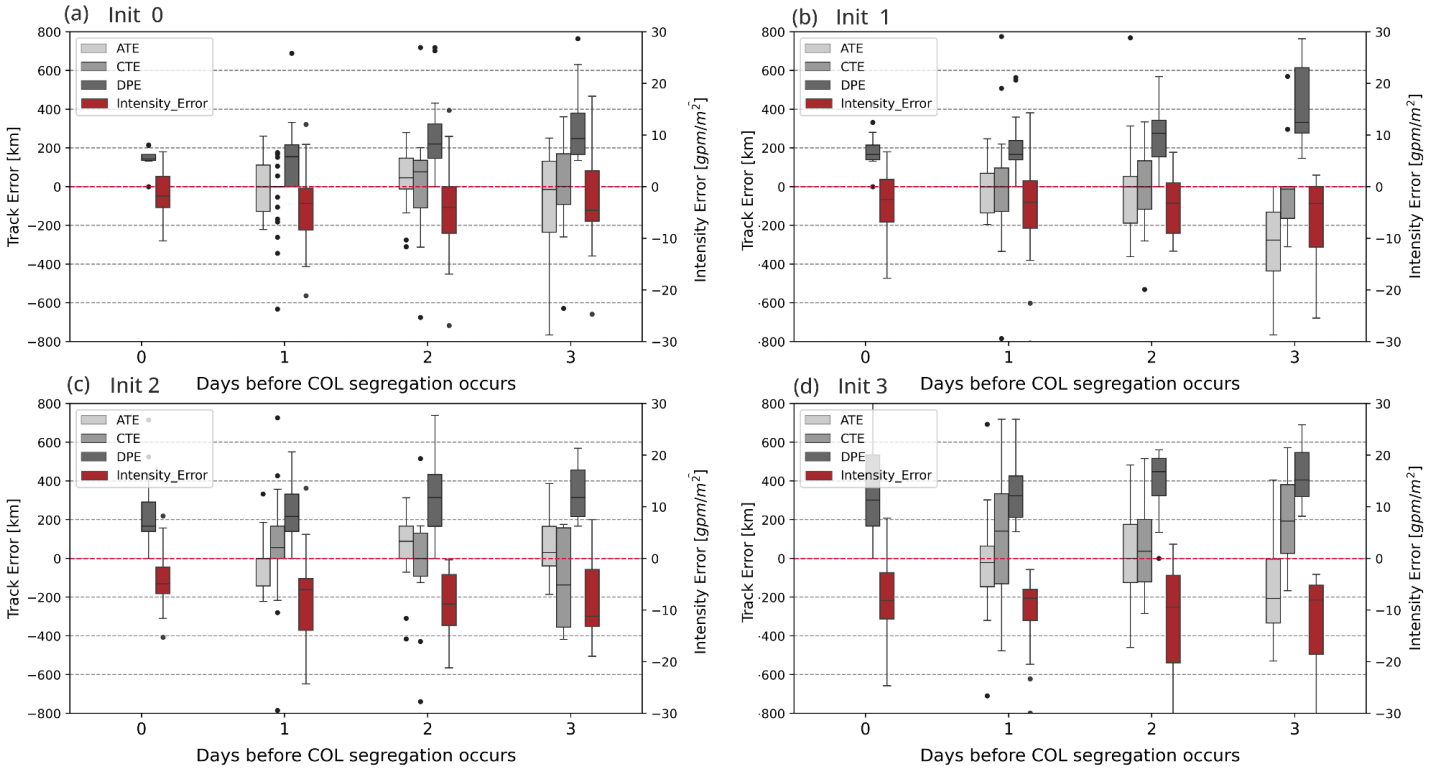


Figure 4: Variation in a) onset position (DPE) and b) the intensity error as a function of initializations. The whiskers at the top (bottom) of the boxes represent the error's 75th (25th) quantile. The black thick horizontal lines inside the boxes represent the median (the 50th quantile) and the points outside the whiskers are considered outliers. The red numbers at the bottom indicate the number of systems identified under each initialization.



565

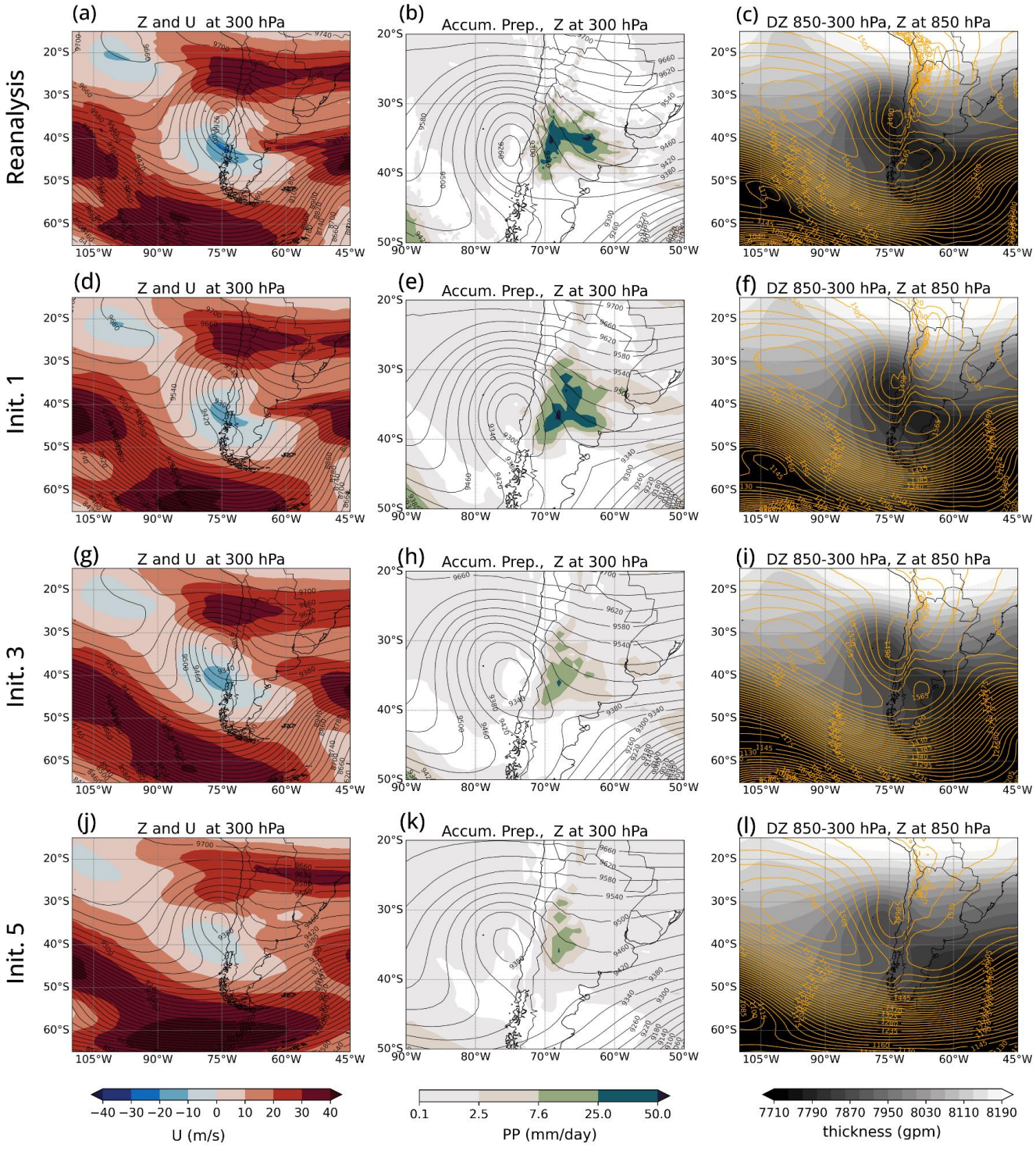
566 **Figure 5: Scatter diagrams of COL initial position deviation decomposed in longitudinal and latitudinal errors (in degrees), where**
567 **the central axis is the initial position observed. Each plot represents a different initialization: ranging from a) Init 0 (forecast**
568 **initialized in the onset day) to h) Init7 (7 days in advance). The gray/black dots indicate the location of the predicted COLs as a**
569 **function of the initialization day (see the color bar for reference on the number of predicted systems per day). The red dots show the**
570 **mean location after averaging all the COLs predicted in each initialization day.**



571

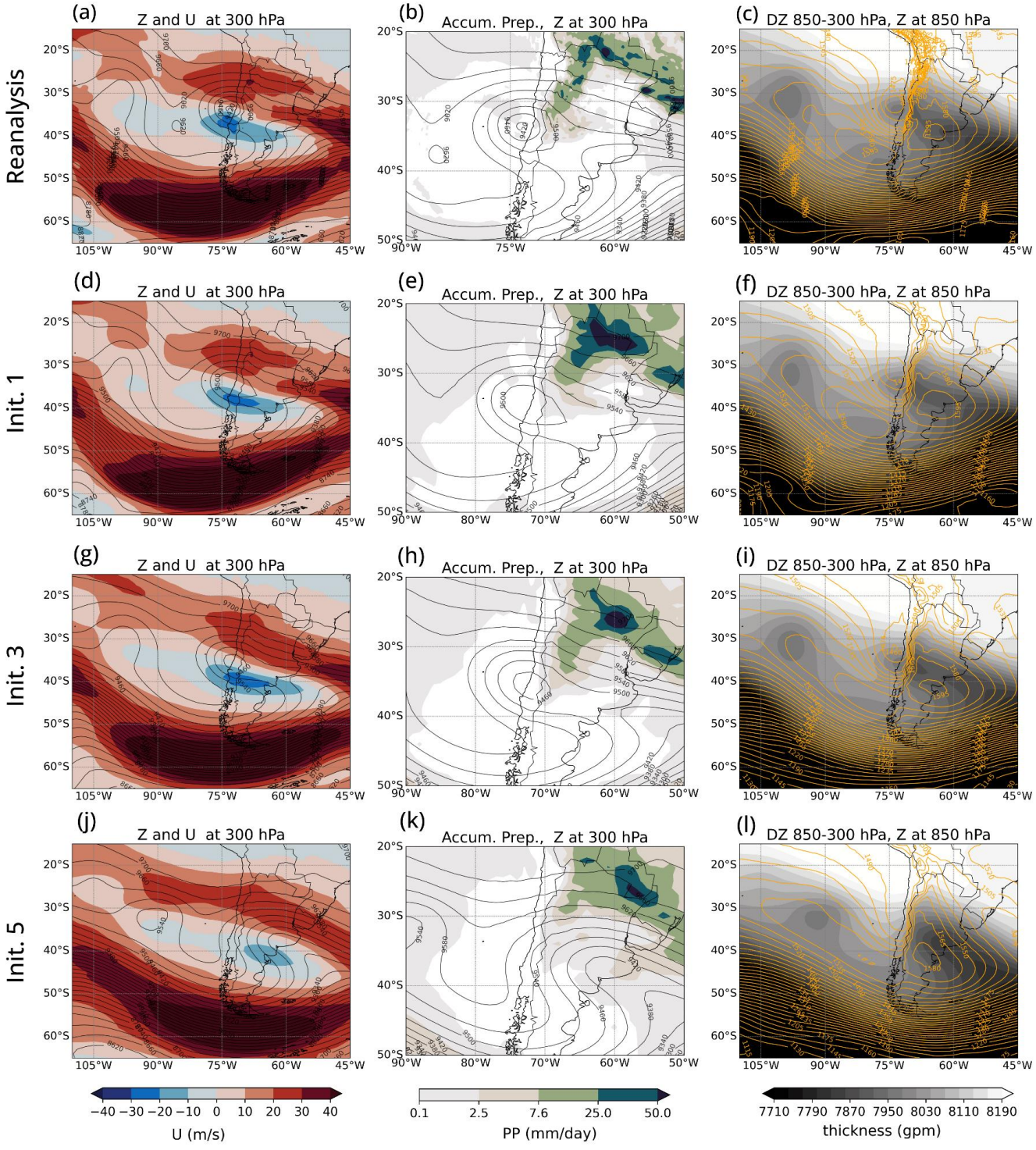
572 **Figure 6: Boxplots of errors in track forecasts for: DPE, ATE, CTE (on the left axis) and Intensity (on the right axis) along the life**
573 **cycle of the COLs. Each plot represents initializations at a) Init 0, b) Init 1, c) Init 2, and d) Init 3.**

574


Reanalysis
Init. 1
Init. 3
Init. 5
(a) Z and U at 300 hPa
(b) Accum. Prep., Z at 300 hPa
(c) DZ 850-300 hPa, Z at 850 hPa
(d) Z and U at 300 hPa
(e) Accum. Prep., Z at 300 hPa
(f) DZ 850-300 hPa, Z at 850 hPa
(g) Z and U at 300 hPa
(h) Accum. Prep., Z at 300 hPa
(i) DZ 850-300 hPa, Z at 850 hPa
(j) Z and U at 300 hPa
(k) Accum. Prep., Z at 300 hPa
(l) DZ 850-300 hPa, Z at 850 hPa
20°S
30°S
40°S
50°S
60°S
105°W
90°W
75°W
60°W
45°W
80°W
70°W
50°W
−40 −30 −20 −10 0 10 20 30 40
U (m/s)
0.1 2.5 7.6 25.0 50.0
PP (mm/day)
7710 7790 7870 7950 8030 8110 8190
thickness (gpm)

575

Figure 7: Segregation stage of the COL formed on March 31st, 2013. (Top) ERA5 and (rows 2 to 4) GEFS predictions of (first column) geopotential height (Z) and wind (U) at 300 hPa, (second column) geopotential height (Z) at 300 hPa and accumulated precipitation (Accum. prep.) over 24 hours, and (right column) geopotential height (Z) at 850 hPa alongside the 850/3000 hPa layer thickness (DZ) GEFS predictions correspond to init 1 (second row), init 3 (third row) and init 5 (fourth row), initialized on March 30th, March 28th and March 26th, 2013, respectively.


Reanalysis
Init. 1
Init. 3
Init. 5
(a) Z and U at 300 hPa
(b) Accum. Prep., Z at 300 hPa
(c) DZ 850-300 hPa, Z at 850 hPa
(d) Z and U at 300 hPa
(e) Accum. Prep., Z at 300 hPa
(f) DZ 850-300 hPa, Z at 850 hPa
(g) Z and U at 300 hPa
(h) Accum. Prep., Z at 300 hPa
(i) DZ 850-300 hPa, Z at 850 hPa
(j) Z and U at 300 hPa
(k) Accum. Prep., Z at 300 hPa
(l) DZ 850-300 hPa, Z at 850 hPa
−40 −30 −20 −10 0 10 20 30 40
U (m/s)
0.1 2.5 7.6 25.0 50.0
PP (mm/day)
7710 7790 7870 7950 8030 8110 8190
thickness (gpm)

583

584 **Figure 8: As in Figure 7 but for the COL formed on March 9th, 2019. In this case, the GEFS predictions corresponding to init 1**
585 **(second row), init 3 (third row) and init 5 (fourth row) were initialized on March 8th, March 6th and March 4th, 2019, respectively.**