Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

IAC Institute for Atmospheric and Climate Science

Master's thesis

Master's degree programme in Environmental Sciences

Predictability of extreme Mediterranean cyclones in past and current ECMWF models

Author: Gabriel Vollenweider (15-971-989)
Institute for Atmospheric and Climate Science, ETH Zürich

Supervisors: Dr. Dominik Büeler (main supervisor)
Institute for Atmospheric and Climate Science, ETH Zürich

Alexander Scherrmann
Institute for Atmospheric and Climate Science, ETH Zürich

Prof. Dr. Heini Wernli
Institute for Atmospheric and Climate Science, ETH Zürich

03.04.2023

Acknowledgements

First and foremost I want to thank my supervisors Dr. Dominik Büeler, Alexander Scherrmann, and Prof. Dr. Heini Wernli. You have guided me very well throughout the working process, and always provided me with valuable feedback and answers to my questions. Our meetings were always stimulating and motivating for me.

I also want to thank my fellow master students in the office. In particular, Franziska Schnyder who provided me with an initial script for the statistical analysis of cyclones, and Effi Perger who helped me design many figures in this thesis.

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Predictability of extreme Mediterranean cyclones in past and current ECMWF models

Authored by (in block letters):

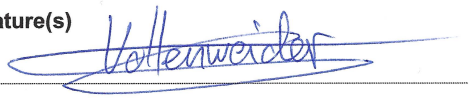
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
Vollenweider	Gabriel

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**
Wetzikon ZH, 03.04.2023	

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.

Abstract

Cyclones are one of the main factors that drive the variability of Mediterranean weather. If they reach high intensities, they pose a major environmental risk to the densely populated coasts along the sea. Although progress has been made in modelling the dynamics and physical processes relevant for the formation of Mediterranean cyclones, the prediction of extreme Mediterranean cyclones still remains a challenge.

While forecast performance is often verified in a climatological context based on years of forecast data, this thesis takes a novel approach and aims to understand how model developments actually affect forecast performance for individual extreme weather events. To this end, we compare three different forecast types from the ECMWF: deterministic forecasts (higher resolution), ensemble forecasts (lower resolution), and ensemble hindcasts (more modern model versions). On the one hand, this allows us to study the relation between forecast accuracy and model version, and on the other hand, this shows how well a historic event would have been predicted with more recent model versions. In addition, we compare different initialisation times to analyse how forecast accuracy depends on forecast lead time. Our evaluation is based on mean sea level pressure, horizontal wind speed, and precipitation. We focus on both the spatial distribution of forecast errors, and the accuracy of spatially and/or temporally averaged quantities.

We find three main errors for forecasts at long lead times: an uncertain position of the cyclone, a timing error of the mature stage (at least 0.5 d), and an underestimated cyclone intensity (winds too weak, precipitation too scarce). These uncertainties become smaller with decreasing lead time, and we notice a distinct jump in forecast accuracy at lead times between 4.5 d and 5 d before the mature stage of the cyclone. However, even the short lead time forecasts do not predict strong winds and heavy precipitation with sufficient accuracy to infer the societal impact of the Mediterranean cyclones.

Overall, the ensemble hindcasts are the most accurate forecasts, followed by the deterministic forecasts, and the ensemble forecasts. The advantages of the hindcasts are smaller positional uncertainties, and better accuracy at longer lead times (extended skill horizon). Since the hindcasts are based on more recent model versions, this highlights the benefits of the model improvements over time. Still, the forecast accuracy is very variable and each event shows unique error patterns. This reflects the large case-to-case variability of extreme Mediterranean cyclones. Our case study approach indicates potential sources of forecast errors, but specifically designed experiments would be needed in future research, to obtain a detailed analysis of forecast uncertainties.

Contents

List of Figures

List of Tables

Symbols

symbol	description	SI unit
λ	longitude, $\lambda \in [-180, 180)$	°E
φ	geodetic latitude, $\varphi \in [-90, 90]$	°N
p	atmospheric pressure	Pa
P	precipitation rate	$\mathrm{m\,s^{-1}}$
θ	potential temperature, $\theta = T \cdot (p_0/p)^{R/c_p}$	K
$\boldsymbol{v}$	horizontal wind velocity, $\boldsymbol{v} = u\,\boldsymbol{e}_1 + v\,\boldsymbol{e}_2$	$\mathrm{m\,s^{-1}}$
s	horizontal wind speed, $s = \lVert\boldsymbol{v}\rVert = \sqrt{u^2 + v^2}$	$\mathrm{m\,s^{-1}}$
ω	vertical component of velocity in pressure coordinates	$\mathrm{Pa\,s^{-1}}$

Chapter 1

Introduction

This chapter provides scientific background on the topics of Mediterranean cyclones (climatology and dynamics), forecast techniques (deterministic forecasts and ensemble predictions), and challenges in the prediction of Mediterranean cyclones. Finally, a brief outline of the thesis is given.

1.1 Mediterranean cyclones

The Mediterranean basin is an environment with unique features. As seen in Figure 1.1, it is enclosed by extended mountain ranges. Regions close to these mountains are climatological hotspots for cyclogenesis (Lionello et al., 2016), and the orography has the potential to intensify both associated convection and (heavy) precipitation (Raveh-Rubin and Wernli, 2016). The proximity to the Saharan desert further enables the transport of Saharan dust into the basin. Over the eastern Mediterranean, up to 25 % of all days with dust correspond to cyclonic activity (Flaounas et al., 2015). Climatologically, the Mediterranean sea is less baroclinic than the North Atlantic (Jones et al., 1999; Nakamura and Yamane, 2009), and Mediterranean cyclones are typically less intense than cyclones within the North Atlantic storm track (Trigo, 2006; Čampa and Wernli, 2012). However, it is still possible that strong cyclones form within the Mediterranean basin. As Argence et al. (2008) note, higher sea surface temperatures provide the necessary heat and moisture to form intense and heavily precipitating systems. If a strong cyclone develops in the Mediterranean, the factors above can lead to environmental hazards such as floods, wind storms, and dust storms. The impact of such events is particularly high considering the dense population along the coasts of the Mediterranean sea.

1.1.1 Climatology

Cyclones occur frequently in the Mediterranean region. Lionello et al. (2016) note that, out of all northern hemispheric cyclones, between 3 % and 10 % pass over the Mediterranean at some point along their track. Additionally, they find that for 76 % of these, cyclogenesis actually occured in the Mediterranean basin. This amounts to a yearly number of Mediterranean cyclogenesis events between 46 and 401, with a mean of 173. Cyclogenesis is most frequent in winter, and least frequent in summer, while autumn and spring are transitional periods (Campins et al., 2011; Lionello et al., 2016). The winter season is not only characterised by frequent cyclogenesis, but also by enhanced low-level baroclinity, which may favour the development of strong cyclones (Trigo et al., 2002). When considering only the most intense Mediterranean cyclones (about 30 per year), their mature stage locations show a distinct spatial distribution (Figure 1.2). Accumulated over 45 years, the hotspots for intense cyclones are regions near the coast of Italy (30–45 cyclones), northwestern Africa (30–40 cyclones), and southern Greece (20–30 cyclones).

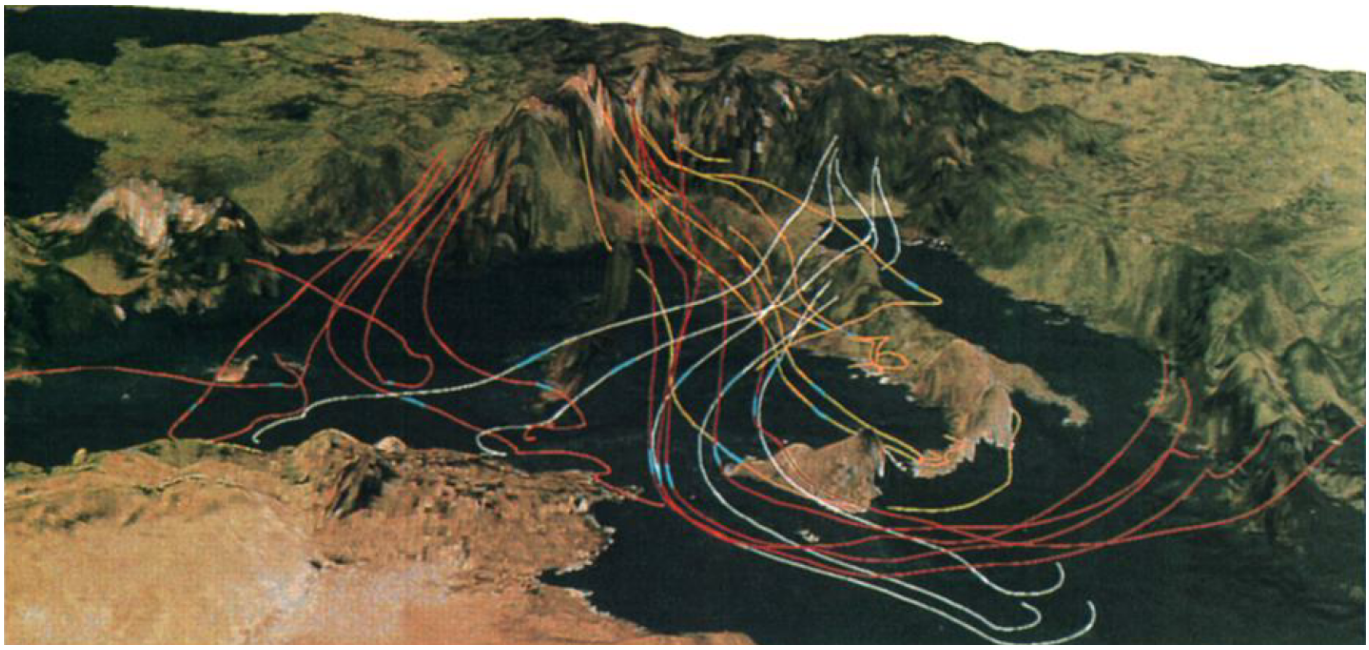


Figure 1.1: Three-dimensional map of the Mediterranean basin. The lines are 72 h backward trajectories associated with four different extreme cyclone events. Red trajectories on the left are for the event at Brig on 23. September 1993, pink trajectories in the middle correspond to Vaison-la-Romaine on 22. September 1992, yellow trajectories denote the event at Piedmont on 06. November 1994, and white trajectories are for the event in South Ticino on 14. September 1995. The figure is taken from Massacand et al. (1998).

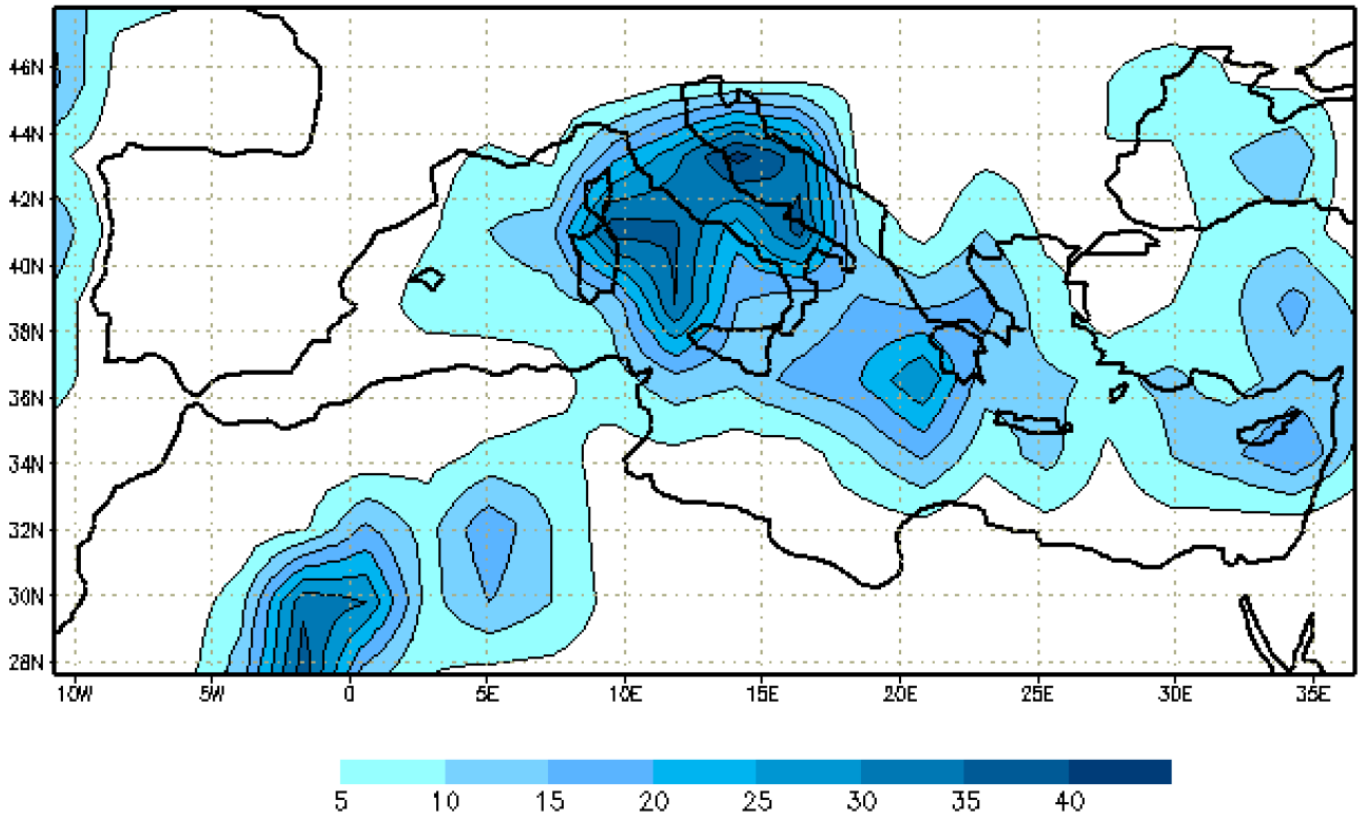


Figure 1.2: The number of intense cyclones at their mature stage over the 45 years from 1957 to 2002. The figure is taken from Homar et al. (2007).

1.1.2 Dynamics

Many studies (e.g., Flocas, 2000; Trigo et al., 2002; Nicolaides et al., 2006; Fita et al., 2007) indicate that Mediterranean cyclones are often triggered by a pre-existing upper-level trough, which can be identified as a streamer or cutoff of potential vorticity (PV). Massacand et al. (1998) note that, between 1993 and 1996, every heavy precipitation event in the Alpine region was associated with an upper-level PV anomaly. The movement, extent, and intensity of these upper-level anomalies strongly influence the evolution of Mediterranean cyclones (Portmann et al., 2020). As such, Argence et al. (2008) mention that uncertainties in the initial configuration of the upper-level anomaly (position, extent, and intensity) are a significant limitation for the predictability of extreme Mediterranean cyclones. Perturbations of the initial trough may propagate and intensify, which introduces uncertainty into the forecast of the surface cyclone. A displaced upper-level trough offsets the location of cyclogenesis, and if the cyclone position is slightly off, the evolution of associated convection and heavy precipitation is found to be drastically modified.

If forecasts are initialised before the formation of an upper-level trough, then additional uncertainties are introduced. As mentioned in Raveh-Rubin and Flaounas (2017), the upper-level troughs triggering the 200 most intense Mediterranean cyclones are typically formed due to Rossby wave breaking over the North Atlantic. For 90.5 % of these cases, the wave breaking could be associated with warm conveyor belts (slantwise ascending diabatic air streams, Harrold, 1973; Carlson, 1980; Browning, 1986) corresponding to North Atlantic cyclones. The upper-level outflow of these air streams can cause ridge amplification, which may lead to Rossby wave breaking, the formation of an upper-level trough further downstream, and subsequent cyclogenesis in the Mediterranean. To accurately predict Mediterranean cyclones associated with this mechanism, all processes in this chain need to be modelled with high precision. The challenge to predict extreme cyclones is thus considerably greater for forecasts beyond a few days. Portmann et al. (2020) emphasise that even a small scale perturbation on the North Atlantic wave guide can significantly affect the forecast skill for an intense Mediterranean cyclone. In particular, they investigate the resulting error in the position of an upper-level trough, as a function of forecast initialisation time. The error is only found to strongly decrease, when forecasts are initialised less than 3 d before cyclogenesis. Di Muzio et al. (2019) find a similar result, with forecast errors rapidly decreasing for initialisations less than 5 d to 7 d before the mature stages of Mediterranean cyclones. These jumps in forecast skill indicate the difficulty of predicting cyclogenesis in the Mediterranean basin. This is in contrast to extratropical cyclones in general, where jumps in forecast skill are less apparent, and errors are found to gradually decrease with later initialisation times (Froude, 2012).

1.1.3 Challenges in the prediction of Mediterranean cyclones

Even though numerical weather prediction drastically improved over the last decades (Bauer et al., 2015), the accurate prediction of Mediterranean cyclones still remains a challenge. Advancements in model resolution generally lead to improved forecasts. In fact, Cioni et al. (2018) find that a Mediterranean cyclone with structural similarities to a hurricane (often referred to as a medicane) could only be forecasted accurately with a convection resolving model, having a resolution of several kilometres. However, it has not yet been demonstrated, that there is a general benefit of kilometre-scale resolution for forecasts of Mediterranean cyclones (Flaounas et al., 2022). Improvements of parameterisations are also necessary to reduce forecast uncertainties. Flaounas et al. (2022) emphasise that microphysics and turbulence are particularly important. This is because they are linked to convection, which on the small scale determines precipitation patterns, and on the large scale influences Rossby wave breaking.

The forecast skill for Mediterranean cyclones thus depends on processes over a wide range of scales, from convective scale diabatic processes to synoptic scale Rossby wave breaking. Flaounas et al. (2022) also mention that geographical features in the Mediterranean basin (such as mountain chains or sea–land transitions) occur on very different scales, making cyclone prediction more difficult. Another challenging aspect is data assimilation. Using remote sensing data to improve predictions of such a multiscale problem has proven to be very difficult (Flaounas et al., 2022). Employing a data targeting system, the Mediterranean experiment (MEDEX, Jansà et al., 2014) helped to identify specific areas where improved observations are likely to reduce forecast errors. However, Campins et al. (2013) note that such data targeting systems strongly depend on the specific characteristics of a given forecasting problem. Flaounas et al. (2022) mention that it remains a challenge to identify targeting strategies, which improve forecasts of Mediterranean cyclones in general.

To better understand the evolution of Mediterranean cyclones, and particularly to improve the prediction of extreme events, the European Cooperation in Science and Technology (COST) has therefore initiated a project uniting researchers, scientists, and stakeholders form 30 countries around the Mediterranean (COST, 2022). This thesis is embedded within this overarching effort.

1.2 Forecast techniques

Numerical weather prediction is a challenging problem, given the chaotic nature of the atmosphere. For local and short-term weather, the first forecasts from 1981 showed useful skill up to 5 d into the future (Bauer et al., 2015). Today, the short-term weather forecasts are useful up to lead times of about 7 d to 9 d (Haiden et al., 2018). However, the forecast performance depends strongly on the atmospheric variable of interest, and the time scales of the associated physical processes. There are still many uncertainties, especially with regards to small scale processes like convection, cloud microphysics, or turbulence.

On the one hand, it is thus desirable to run forecasts with particularly high resolution, such that small scales are treated as accurately as possible. But on the other hand, it is also necessary to run forecasts that inform about the remaining uncertainties. To meet these different requirements, several specialised forecasting systems have been developed.

1.2.1 Deterministic forecasts

The deterministic forecasts are predictions for one possible evolution path of the atmosphere. They are run at the highest computationally feasible resolution, providing the most detailed picture of the possible future weather. However, since they predict a single state of the atmosphere, they do not inform about any uncertainties.

1.2.2 Ensemble forecasts

The ensemble forecasts are predictions for multiple possible evolution paths of the atmosphere. They are a collection of deterministic forecasts, where each member is started with perturbed and thus slightly different initial conditions. Over the last decades, the method to compute these perturbed initial states has greatly improved. The perturbations were optimised to simulate the uncertainties of both the observational data and the atmospheric model. Since the ensemble forecasts involve many independent forecasts, they are typically run at a smaller resolution than the deterministic forecasts.

Ensemble forecasts are run for the same time range as the deterministic forecast (10 d at the European Centre for Medium-Range Weather Forecasts, ECMWF). But they may also be run for longer time ranges, to predict the weekly changes in weather, and to evaluate the long-term performance of the atmospheric model. Since the uncertainties grow with time, it is challenging to calibrate the extended-range forecasts. At ECMWF, this problem is addressed by running additional forecasts of past weather situations, to obtain a detailed probability distribution of the forecasts (a forecast climatology). The additional forecasts are called "hindcasts" or "re-forecasts". The ECMWF currently runs these hindcasts twice a week. To obtain a large enough calibration sample, 20 different hindcasts are run on the same day. The first one with an initialisation time 20 yr in the past, the second one 19 yr in the past, and so on, until the last one which reaches back 1 yr into the past.

1.2.3 Notion of predictability

In the context of chaotic systems, predictability is related to the growth of uncertainties as a system evolves over time. Given a collection of slightly perturbed initial states, we say that a situation is unpredictable, when the different states quickly diverge form each other and evolve very differently. Conversely, we speak of a predictable situation, when the different states continue to be close to each other and evolve very similarly. However, there is no clear threshold to differentiate between these two regimes.

For the atmosphere, we refer to this chaotic systems definition as intrinsic predictability. Given that ensemble forecasts show the evolution of slightly perturbed initial atmospheric states, it might be tempting to link the spread between all ensemble members to intrinsic predictability limits. However, diverging ensemble members do not necessarily indicate low intrinsic predictability. This is because the ensemble spread also depends on the overall performance of the forecast model. If the model cannot accurately predict the evolution of a given initial state, then this may introduce errors into the ensemble spread, meaning that it does not accurately reflect the intrinsic predictability. Therefore, we cannot make definitive statements about intrinsic predictability, and instead focus on forecast skill, which is defined as the accuracy of the forecast relative to a reference data set (usually based on observations).

1.3 Outline

In this thesis, we investigate several cases of extreme Mediterranean cyclones. For each case, we study the synoptic situation, and then evaluate how well the cyclone was predicted by different types of forecasts: deterministic forecasts, ensemble forecasts, and ensemble hindcasts. Our focus is on the differences between these forecast types, in relation to mean sea level pressure, surface wind speed, and precipitation. The goal is to assess how well the forecasts predict these fields for each individual case, and whether the forecasts show similar patterns across the different cases, despite the large case-to-case variability. By comparing the different forecast types, we investigate the potential improvements of better resolution (deterministic forecast) or newer model versions (ensemble hindcasts).

The thesis is structured as follows. Chapter 2 first introduces the data upon which we base our analyses. This includes the different types of forecast data, and the reference data for the evaluation. Next, we explain the methods applied to the forecast data to investigate their accuracy. We present the results in chapter 3, for the different cases of extreme Mediterranean cyclones. Chapter 4 discusses the general performance of the forecasts based on the results of the case studies, together with possible explanations for their strengths and weaknesses. Finally, chapter 5 summarises the main results and conclusions, while also giving an outlook into further research topics.

Chapter 2

Data and methods

2.1 Data

All data used in this thesis is taken from the European Centre for Medium-Range Weather Forecasts (ECMWF). The data includes operational forecast data, hindcast data, and reanalysis data.

2.1.1 Operational forecasts

Operational forecasts are predictions for a future period of time, using the best available models at the time of forecast initialisation. The forecast data is retrieved from the ECMWF's Meteorological Archival and Retrieval System (MARS, ECMWF, 2018). We analyse both medium-range deterministic forecasts and medium-range ensemble forecasts. MARS provides daily deterministic 10 d forecasts starting from 1. January 1985. The ensemble forecasts are also available as 10 d forecasts, from 24. November 1992 onwards. Before 11. December 1996, the ensemble forecasts consist of 32 members, and afterwards the ensembles include 50 members.

For this thesis, we use forecasts of five specific weather events between 1993 and 2011. Throughout this period, the ECMWF continuously updated both the atmospheric model and the forecasting system. The full details on these updates are given in Tables 2.1 to 2.3.

2.1.2 Hindcasts

Hindcasts are predictions for a past period of time. They use the operational model at the time of forecast initialisation ("the present"), and take the initial conditions for the past from reanalysis data. The hindcasts include multiple members, and are run at the same reduced resolution as the operational ensemble forecasts. The ECMWF produces hindcasts to detect biases in their extended-range forecasts, as mentioned in section 1.2.2. The hindcasts are also retrieved from MARS.

The hindcasts are part of the extended-range forecasting system, which also went over many updates in the last 30 years. Originally, the extended-range forecasts were run as 32 d forecasts, and they were initialised once a week (every Thursday) starting from 13. March 2008. The first ensemble hindcasts included only four members, but for a given initialisation day (for example 13. March 2008), 18 different hindcasts were produced. One for an initial time 18 yr earlier (13. March 1990), one for an initial time 17 yr earlier (13. March 1991), and so on, until the finial initial time 1 yr earlier (13. March 2007). On 21. June 2012, the hindcasts were upgraded to include the last 20 yr. On 11. May 2015, the hindcast range was extended to 46 d, the ensemble was upgraded to 10 members, and the initialisation frequency was increased to two runs per week (Monday and Thursday). Note, however, that we only use the hindcasts up to the medium-range of 10 d.

Since a single hindcast run produces forecasts for multiple years, there may be many hindcasts that predict the same extreme weather event. For the five events of this thesis, the number of available hindcasts varies between 4 and 23. Out of all hindcasts for a given event, some are initialised a few days apart from each other, but usually they are initialised at least 1 yr apart from each other. For our five cases, this means that we have hindcasts from 2 to 15 different years. Given that the hindcasts improve over time, we then have a set of hindcasts with mixed model versions for the same extreme weather event. This enables us to study possible forecast improvements over time.

2.1.3 Reanalysis

As our reference data, we use the fifth generation ECMWF reanalysis (ERA5, Hersbach et al., 2020). This reanalysis is based on cycle Cy41r2 of the Integrated Forecasting System (IFS, ECMWF, 2022), which was operational at the ECMWF in 2016. The ERA5 data includes atmospheric fields from 1950 onwards, with a timestep of 1 h and a spatial resolution of about 31 km. Since the whole reanalysis uses the same atmospheric model and data assimilation system, the data set is coherent in both space and time throughout the whole time period. This is in contrast to operational analyses, which are based on assimilation systems that evolve over time. The ERA5 reanalysis is thus well suited as a reference data set. It can be used as a consistent basis to evaluate and compare different forecasts that may be produced with different model versions. However, it should be noted that both the forecast data and the reference data are based on the IFS model from the ECMWF. The reference data is thus not a completely independent data set. This may lead to estimates of forecast accuracy that are too optimistic.

2.1.4 Data availability for a given extreme event

This thesis deals with specific extreme events. These typically last between 2 d and 4 d. For the evaluation of forecasts, we choose initialisation times between 6 d and 0 d before the start of the event. Given the varying forecast initialisation frequencies, there may be missing data. One example is the first event in 1993, between 22. and 25. September. Here, the deterministic forecast is available every day from 18. to 22. September at 12 UTC. But the ensemble forecast is only available from Friday 18. September to Sunday 20. September at 12 UTC. For the ensemble hindcast, the data availability is less straightforward, and we need to check the hindcast initialisation times manually. In general, the data availability varies depending on the specific extreme event and the kind of forecast. An overview of the data for each case is given in Table A.1.

In addition to the varying initialisation frequencies, the output timestep may also be different for different forecasts. In the case of September 1993, the deterministic forecast has a variable timestep (3 h for the first 0.5 d, 6 h for the next 4.5 d, 12 h for the last 5 d). The ensemble forecast has a timestep of 12 h for 10 d, and the ensemble hindcasts have a timestep of 06 h for 32 d. In addition, the timestep may vary depending on the meteorological field or the vertical model level. In general, we simply select the smallest possible timestep for the whole forecast range of 10 d. For September 1993, this would be 12 h for the deterministic forecast and the ensemble forecast, but 6 h for the ensemble hindcasts. However, this is a problem for accumulated fields like precipitation. In that case, we need to choose a homogeneous timestep across all forecast types (12 h for September 1993).

It is also possible that the resolution of the model output varies among the forecasts and the reanalysis data. In order to compare the data, we thus retrieve all forecast data on a regular horizontal grid with a spacing of 1°. The reanalysis data, originally available on a regular $0.5° \times 0.5°$ grid, is also remapped to the regular $1° \times 1°$ grid[1]. Overall, the issues in data availability make it challenging to obtain a fair and systematic forecast verification for a specific event.

[1] This is done using the *remapcon2* command from the Climate Data Operators (CDO, Schulzweida, 2022).

Table 2.1: Changes in the forecasting systems at ECMWF. Listed are the deterministic forecast (FC), the ensemble forecast (EF), and the ensemble hindcast (EH). The updates include the frequency of forecast initialisation (f_{ini}), the forecast output timestep (Δt), and the past period for the ensemble hindcast (Y_{past}). The output timestep may vary with forecast lead time. We thus list the different timesteps chronologically, with the corresponding intervals of forecast lead time in parentheses. Note that the table does not indicate the dependence of the timestep on either the meteorological field or the vertical model level. The updates listed here are thus not necessarily true for all possible cases.

date	FC	EF	EH
01.08.1980	f_{ini}: every day at 12 UTC Δt: 12 h (10 d)	-	-
01.07.1985	Δt: 06 h (for 5 d) 12 h (for 5 d)	-	-
15.11.1990	Δt: 03 h (for 0.5 d) 06 h (for 4.5 d) 12 h (for 5 d)	-	-
30.09.1992	f_{ini}: every day at 00 UTC[2], 12 UTC	-	-
24.11.1992	-	f_{ini}: Fri-Sat-Sun at 12 UTC Δt: 12 h (for 10 d)	-
01.05.1994	-	f_{ini}: every day at 12 UTC	-
16.01.1997	f_{ini}: every day at 00 UTC, 12 UTC	-	-
14.10.1997	f_{ini}: every day at 12 UTC	-	-
20.01.1999	Δt: 03 h (for 0.5 d) 06 h (for 9.5 d)	-	-
13.01.2000	-	Δt: 06 h (for 10 d)	-
12.09.2000	Δt: 06 h (for 10 d)	-	-
24.10.2000	Δt: 03 h (for 3 d) 06 h (for 7 d)	-	-

[2] Forecasts started at 00 UTC have a range of 3 d instead of 10 d.

Table 2.2: Changes in the forecasting systems at ECMWF (continued).

date	FC	EF	EH
28.03.2001	f_{ini}: every day at 00 UTC, 12 UTC	-	-
25.03.2003	-	f_{ini}: every day at 00 UTC, 12 UTC	-
29.06.2005	-	Δt: 03 h (for 5.25 d) 06 h (for 4.75 d)	-
20.10.2005	-	Δt: 03 h (for 6 d) 06 h (for 4 d)	-
15.03.2006	Δt: 03 h (for 4 d) 06 h (for 6 d)	-	-
13.09.2006	Δt: 03 h (for 6 d) 06 h (for 4 d)	Δt: 03 h (for 6 d) 06 h (for 9 d)	-
13.03.2008	-	-	f_{ini}: every Thu at 00 UTC Δt: 06 h (for 32 d) Y_{past}: 18 yr
16.11.2011	Δt: 01 h (for 3.75 d) 03 h (for 2.25 d) 06 h (for 4 d)	-	-
21.06.2012	-	-	Y_{past}: 20 yr
14.05.2015	-	-	f_{ini}: every Mon/Thu at 00 UTC Δt: 06 h (for 46 d)
22.06.2015	-	f_{ini}: every day at 00 UTC, 06 UTC[3] 12 UTC, 18 UTC[4]	-
23.11.2016	-	Δt: 01 h (for 3.75 d) 03 h (for 2.25 d) 06 h (for 9 d)	-

[3] Forecasts started at 06 UTC have a range of 6 d instead of 15 d.
[4] Forecasts started at 18 UTC have a range of 6 d instead of 15 d.

Table 2.3: Resolution upgrades to the ECMWF model. The deterministic forecast (FC) is characterised with the maximum horizontal wave number (k_{max}), and the number of model levels (n_{lev}). The ensemble forecast (EF) and the ensemble hindcast (EH) also include the number of members (n_{mem}). The labels in front of the maximum wave numbers indicate both the truncation method of the spherical harmonics (T for triangular), and the definition of the horizontal grids: L for linear, Q for quadratic, and Co for cubic-octahedral. These grids differ in the number of grid points that represent the smallest possible horizontal wavelength (2 for linear, 3 for quadratic, 4 for cubic).

	FC		EF			EH		
date	k_{max}	n_{lev}	k_{max}	n_{lev}	n_{mem}	k_{max}	n_{lev}	n_{mem}
21.04.1983	T_Q63 (320 km)	16	-	-	-	-	-	-
01.05.1985	T_Q106 (190 km)	-	-	-	-	-	-	-
13.05.1986	-	19	-	-	-	-	-	-
17.09.1991	T_Q213 (95 km)	31	-	-	-	-	-	-
24.11.1992	-	-	T_Q63 (320 km)	19	32	-	-	-
11.12.1996	-	-	T_L159 (120 km)	31	50	-	-	-
01.04.1998	T_L319 (60 km)	-	-	-	-	-	-	-
09.03.1999	-	50	-	-	-	-	-	-
12.10.1999	-	60	-	40	-	-	-	-
21.11.2000	T_L511 (40 km)	-	T_L255 (80 km)	-	-	-	-	-
01.02.2006	T_L799 (25 km)	91	T_L399 (60 km)	62	-	-	-	-
13.03.2008	-	-	-	-	-	T_L399 (60 km)	62	4
26.01.2010	T_L1279 (16 km)	-	T_L639 (35 km)	-	-	T_L639 (35 km)	-	-
25.06.2013	-	137	-	-	-	-	-	-
19.11.2013	-	-	-	91	-	-	91	-
11.05.2015	-	-	-	-	-	-	-	10
08.03.2016	$T_{Co}1279$ (9 km)	-	$T_{Co}639$ (18 km)	-	-	$T_{Co}639$ (18 km)	-	-
11.05.2021	-	-	-	137	-	-	137	-

2.2 Methods

2.2.1 Cyclone statistics and event selection

For the selection of extreme events, we first identify all Mediterranean cyclones between January 1992 and December 2021. This is done with a cyclone identification and tracking algorithm originally developed by Wernli and Schwierz (2006) and later adapted by Sprenger et al. (2017). The algorithm first identifies cyclones as local minima of sea level pressure, which are surrounded by enclosed pressure contours. In a second step, the identified cyclones are then concatenated in time to obtain the cyclone tracks. Using the cyclone tracking algorithm, we obtain time series of both the cyclone centre (longitude and latitude), and the minimum pressure in the cyclone region. The cyclone tracks are computed globally using the ERA5 reanalysis. To filter these tracks for Mediterranean cyclones, we require that at least 75 % of the track lies within the Mediterranean region ($-10\,°\text{E} \leq \lambda \leq 40\,°\text{E}$, $30\,°\text{N} \leq \varphi \leq 50\,°\text{N}$). We note here that the cyclone identification and tracking algorithm does not always produce optimal tracks in the Mediterranean region. One reason for this, is that the cyclones are rather small and thus hard to identify. Another issue is the jumpy behaviour of mean sea level pressure around the steep topography in the Mediterranean. This may lead to tracks that start too late (after cyclogenesis), or tracks where the cyclone centre shows unrealistic jumps.

After identifying all tracks in the Mediterranean region, we then compute a 30 yr climatology of several quantities that represent the whole cyclone track. The quantities are chosen to characterise high impact weather associated with strong cyclones. An overview is given in Table 2.4.

Table 2.4: Characteristic cyclone quantities for which we compute a climatology. The table is intended as an overview. Further details are provided in the text. Note that precipitation is only considered within a great-circle-distance of 200 km around the cyclone centre. This choice is motivated by Flaounas et al. (2018), who showed that rainfall of heavily precipitating cyclones is most likely within a distance of 2.0° to 2.5° around the cyclone centre (in both the longitudinal and the latitudinal directions).

variable	unit	description
lifetime	h	lifetime of the cyclone
$p_{\text{cen}}^{\text{min}}$	hPa	minimum central pressure
P_{cyc}	$\text{mm}\,\text{h}^{-1}$	mean precipitation rate within a radius $r = 200\,\text{km}$
AP_{cyc}	mm	mean accumulated precipitation within a radius $r = 200\,\text{km}$

In the following, some notation is introduced to give the precise definitions for the precipitation quantities. Let $P(\boldsymbol{x}, t)$ denote the ERA5 precipitation rate (in $\text{mm}\,\text{h}^{-1}$) at a certain location on the Earth's surface $\boldsymbol{x}$ (0.5° horizontal resolution) and a specific time t (1 h timestep). Given a single Mediterranean cyclone track, we write $\mathcal{T}$ for the set of all corresponding timesteps, and $N(\mathcal{T})$ for the total number of timesteps. For each time $t \in \mathcal{T}$ along the track, we let $\mathcal{R}(t)$ be the region around the cyclone centre within a great-circle-distance $r = 200\,\text{km}$. The number of grid points contained in that region is denoted by $N(\mathcal{R}(t))$. The mean precipitation rate can then be defined as

$$P_{\text{cyc}} = \frac{1}{N(\mathcal{T})} \sum_{t \in \mathcal{T}} \frac{1}{N(\mathcal{R}(t))} \sum_{\boldsymbol{x} \in \mathcal{R}(t)} P(\boldsymbol{x}, t). \tag{2.1}$$

The corresponding mean accumulated precipitation is then obtained by multiplication with the number of timesteps $N(\mathcal{T})$. After computing these quantities for all Mediterranean cyclones between January 1992 and December 2021, it is possible to statistically quantify the extremeness of a certain event. Figure 2.1 shows the histograms for all characteristic quantities. The minimum central pressure (Figure 2.1b) roughly follows a symmetric distribution around the most common value of about 1010 hPa. But the lifetime (Figure 2.1a) and the precipitation quantities (Figures 2.1c to 2.1h) are not distributed symmetrically. Instead, the smallest values occur most often, and the larger values quickly become very scarce.

The top 20 extreme values are emphasised for each quantity (0.141 % of all cyclones). The smallest central pressures are between 970 hPa and 981 hPa, and the longest lifetimes range from about 5.5 d to over 10 d. The most extreme mean precipitation rates are between 1.56 mm h^{-1} and 2.29 mm h^{-1}. Note that these rates are averaged over both space and time. The rates are thus far more extreme on a local scale and over shorter time intervals. The extreme values of mean accumulated precipitation range from 82.8 mm to 224 mm. These quantities are also averaged over space, meaning that they are more extreme on a local scale.

An example of a particularly extreme cyclone is the storm Ianos (between 14. September 2020 and 20. September 2020). It reaches rank 18 in terms of lifetime (137 h) and rank 13 in terms of mean precipitation rate (1.63 mm h^{-1}). Together this also makes it the cyclone with the highest accumulated precipitation along its track (224 mm). Note that high accumulated precipitation can be the result of a long lifetime, even with a small precipitation rate. Or it can be a short lifetime, with a high precipitation rate. But for Ianos, both quantities are very extreme.

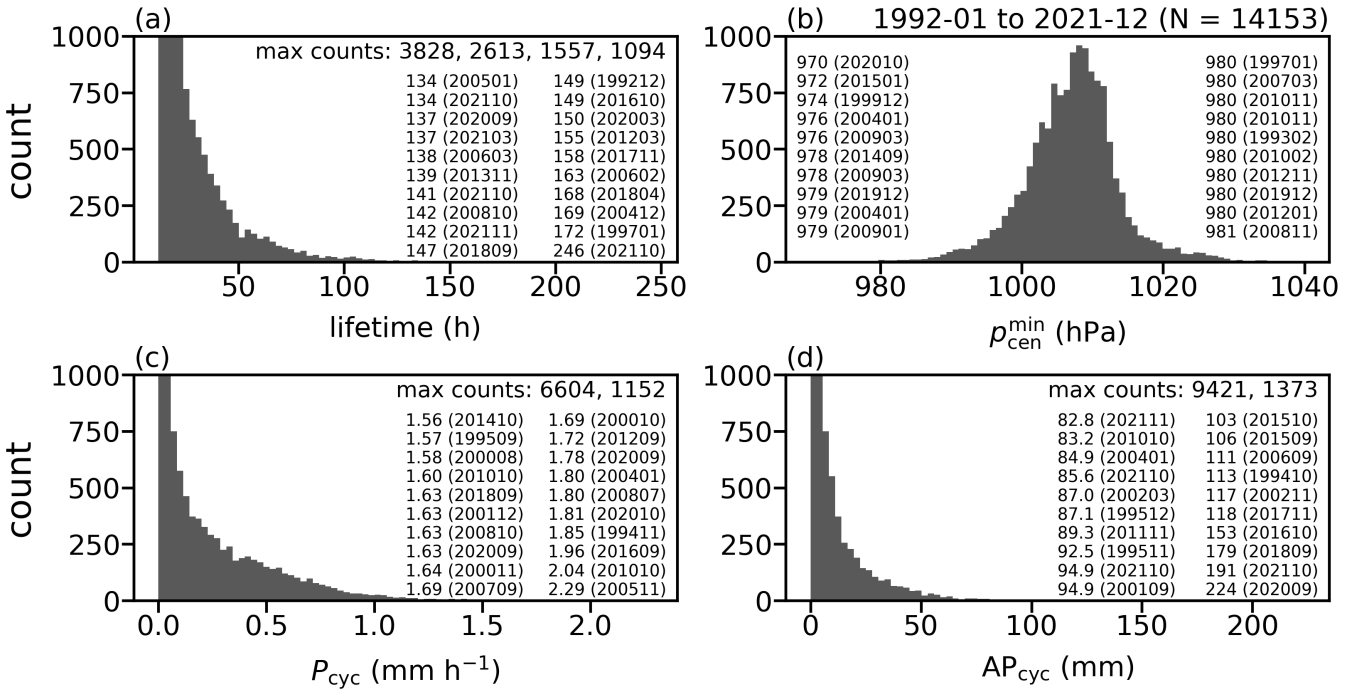


Figure 2.1: Mediterranean cyclone statistics for 1992 to 2021. (a) cyclone lifetime, (b) minimum central pressure, (c) mean precipitation rate, (d) mean accumulated precipitation. Each histogram consists of 80 equally sized bins. For better visibility, the largest counts are not fully displayed. Their counts are instead shown with labels in the top right corners. The top 20 extreme values are emphasised, along with the year and month of the corresponding cyclone.

Since we are interested in events with a high impact on society, we cannot simply select events based on the statistical extremeness of the cyclone characteristics. The impact is also determined by other factors such as: cyclone stationarity, soil saturation, population density, available risk infrastructure, and others. Therefore, we first select the events subjectively, and add the quantification of statistical extremeness in a second step. Table 2.5 gives an overview of the subjectively chosen events.

Table 2.5: Selection of Mediterranean cyclone events. For each event, we briefly describe the associated impact, which guided us to choose the respective cyclone.

event	description of cyclone impact
Brig September 1993	Flash flood in Brig with enormous damage to infrastructure, two fatalities, and much attention in all of Switzerland.
Jerusalem March 1998	The worst dust storm for three decades in the eastern Mediterranean region. The cyclone brought strong winds, followed by hail and snowfall, which caused a general shutdown of the city.
Algiers November 2001	One of Algeria's worst flooding events with over 5000 families left homeless, and more than 700 deaths.
Apulia September 2006	A Mediterranean cyclone with a structure similar to a Hurricane, bringing high precipitation rates and very strong winds causing moderate damages throughout southeastern Italy.
Hyères November 2011	A long-lasting and stationary cyclone brought heavy rainfall over southern France and northwestern Italy, causing widespread landslides and flooding.

2.2.2 Evaluation metrics

To quantify the performance of the different forecasts for a specific event, we use several forecast quality metrics. We adopt the Eulerian point of view for all metrics. They can thus be computed without reference to the actual cyclone tracks. The metrics must be chosen such that they work for all cases of data availability. This means that the least modern operational forecasts, which use the coarsest grid and the largest output timestep, must serve as a basis for the choice of suitable forecast quality metrics. The focus is on extreme weather events with high impacts on society. We are thus interested in the forecasts of pressure, horizontal wind speed, and precipitation. For each of these fields, the evaluation is done according to the following procedure.

1. Visualise the field over the Mediterranean region, for all timesteps and all forecasts. This gives a first overview of the differences between the forecasts and the ERA5 reanalysis. The spatial pattern of these differences provides further insight into the possible dynamics behind forecast errors.
2. Define an evaluation region $\mathcal{R}$ for the field, based on the reanalysis data of the event. This region is subjectively chosen to include the main features that characterise the given extreme event. For example a low pressure anomaly corresponding to a cyclone, a strong wind anomaly leading to enhanced moisture transport, or a precipitation cell causing heavy flooding. Since these features may not overlap, we may select different regions for each field. The evaluation region is fixed and does not change in time. This is different from the cyclone region in section 2.2.1.
3. Plot the time series of the field averaged over the evaluation region $\mathcal{R}$. Different forecasts can be evaluated by comparing the regional means of the forecasts to the regional mean of ERA5. With the means shown over several timesteps, it is possible to see timing discrepancies, intensification offsets, or differences in the peak values.
4. Define an evaluation time window $\mathcal{T}$ based on the time series. This window is subjectively defined to include the most intense phase of the event, which corresponds to one or multiple of the following: the pressure minimum along the track, the timesteps with strong winds, or the phase of maximum precipitation. The evaluation time window takes the role of the set of timesteps in section 2.2.1.
5. Take the regional mean field, and average over the evaluation time window $\mathcal{T}$. This computation is analogous to equation (2.1) with the differences mentioned above, and reduces each forecast to a single number. It is then possible to compare all available forecasts simultaneously, in a comprehensive and concise manner. We can see whether forecasts improve with shorter lead time, or if such improvements vary between the different types of forecasts (deterministic forecasts, ensemble forecasts, ensemble hindcasts).

In the case of precipitation, we also compute the SAL indices proposed by Wernli et al. (2008). These indices measure the errors in the structure (S), amplitude (A), and location (L) of a given precipitation forecast. The main idea is to evaluate the precipitation field based on its typical features, rather than as a simple mean squared error. Consider for example the situation of a small but heavily precipitating storm cell. Now suppose the forecast correctly predicts the size and amplitude of the cell, but gets the location of the cell wrong. Then the mean squared error would be large for such a forecast. This is because it penalises the forecast both for underestimation at the true location, and for overestimation at the wrong location. The SAL indices allow a more detailed evaluation. In the above example, the forecast would have a good S score (for the correct size), a good A score (for the correct amplitude), but a poor L score (for the wrong location).

In the following, we briefly describe the definitions of the A and L indices. The S index is not defined in detail, because it is not as informative for our purposes as the other two indices. We use the same notation as in section 2.2.1, without reference to a specific time. First, we fix an evaluation region $\mathcal{R}$. Then we compute the mean precipitation within this evaluation region, denoted by $\overline{P}$. Doing this for both the forecast and the reference data, we can then define the amplitude index A as

$$A = \frac{\overline{P}_{\text{pred}} - \overline{P}_{\text{ref}}}{\frac{1}{2}\left(\overline{P}_{\text{pred}} + \overline{P}_{\text{ref}}\right)}. \tag{2.2}$$

The amplitude index is thus a scaled difference of regional mean precipitation. Negative values indicate that precipitation is too weak, and positive values indicate that precipitation is too strong. The index is bounded between -2 and 2, and a value of 0 represents a perfect forecast. Letting $\overline{P}_{\text{pred}}$ be a multiple of $\overline{P}_{\text{ref}}$, it can be shown that errors by factors of 2, 4, 8, and 16, correspond, respectively, to A values of ± 0.67, ± 1.20, ± 1.56, and ± 1.76.

The L index involves the computation of several precipitation centroids (centres of mass for precipitation). First, we determine the overall centroid $\overline{\boldsymbol{x}}$ within the whole evaluation region $\mathcal{R}$. In a second step, we divide the precipitation field into disjoint "cells" or "features". This is done by first choosing a threshold P_{crit}[5], and then identifying every single precipitation peak above this threshold. Some features may be small with little total precipitation, while others may be large with considerable total precipitation. We then compute the centroids for all individual precipitation features. For each feature, we can compute the displacement between the feature centroid and the overall centroid. Using the total precipitation within the features as weight factors, we can finally compute the mean feature displacement $\overline{D}$ (relative to the overall centroid). Writing D_{max} for the maximum distance between any two points in $\mathcal{R}$, the location index L can now be defined as

$$L = L_1 + L_2 = \frac{\|\overline{\boldsymbol{x}}_{\text{pred}} - \overline{\boldsymbol{x}}_{\text{ref}}\|}{D_{\text{max}}} + \frac{\left|\overline{D}_{\text{pred}} - \overline{D}_{\text{ref}}\right|}{\frac{1}{2} D_{\text{max}}}. \tag{2.3}$$

The interpretation of the location index is more complicated, because it is a sum of two terms. The first term (L_1) represents a scaled distance between the overall precipitation centroids. Small values correspond to close centroids. The second term (L_2) can be interpreted as a scaled absolute error of the mean feature displacement. Small values indicate similar feature displacements, and large values indicate that the features are either too close to the centroid or too far away from the centroid. The L index ranges between 0 and 2, since $0 \leq L_1 \leq 1$ and $0 \leq L_2 \leq 1$. Again, a value of 0 represents a perfect forecast.

[5] Wernli et al. (2008) define this as $P_{\text{crit}} = P_{\text{max}}/15$, where P_{max} is the maximum precipitation value within the evaluation region $\mathcal{R}$. We adopt the same definition here.

Chapter 3

Results

This chapter describes the forecast evaluations on a case by case basis. For each case, we give a brief overview of the synoptic situation, which is entirely based on the ERA5 data set. Then, we provide the details for the evaluation of deterministic forecasts, ensemble forecasts, and ensemble hindcasts. We consider the three fields: mean sea level pressure (p_{msl}), horizontal wind speed (s), and precipitation rate (P).

For each event, the forecast lead times are divided into long lead times, middle lead times, and short lead times. Relative to the most intense phase of the cyclone, the long lead times are mostly larger than 4.5 d, the middle lead times larger than 3 d, and the short lead times larger than 0.5 d. The exact lead times are mentioned separately for each event. As an initial overview, we consider the spatial distribution of forecast errors. In a second step, we analyse the temporal evolution of the regional means. This gives further details on the accuracy of each forecast. Finally, we take the regional means and average them over the evaluation time window (the most intense phase). This enables us to study the relation between forecast accuracy and forecast lead time.

3.1 Brig September 1993

Between 23. and 25. September 1993 Brig (Switzerland, 7.99 °E, 46.32 °N) was hit with heavy rainfall. Over 3 d more than 300 mm of precipitation fell over the alpine valley (Buzzi and Foschini, 2000). This amount of precipitation usually falls within a whole month. The Brig cyclone had a relatively long lifetime of 46 h (top 11.1 % of all 14 153 cyclones between 1992 and 2021, Figure B.1a), and reached a minimum central pressure of 1000 hPa (top 12.6 %, Figure B.1b). Its precipitation was extreme, with a very large precipitation rate of 1.06 mm h^{-1} (top 1.1 %, Figure B.1c), and considerable accumulated precipitation of 48.8 mm (top 1.5 %, Figure B.1d).

In the days leading up to the event, a large anticyclone ($p_{\mathrm{max}} > 1030\,\mathrm{hPa}$) extended over the North Atlantic. Together with a cyclone in the north-eastern Atlantic (between Iceland and the UK), this led to strong southward flow, which brought high PV air from the north towards the south, resulting in the formation of a far south-reaching PV streamer. On the 320 K isentrope, it covered the region from the northern UK all the way to southern Spain (Figure 3.1a). The streamer (PV $> 2\,\mathrm{PVU}$) reached down to an altitude of 6 km (black contour in Figure 3.1c). A Mediterranean cyclone formed on the eastern side of the streamer (Spanish east coast, Figure 3.1b). The cyclone deepened with a rate of about 8 hPa in 1 d, as the strong PV streamer (PV $> 8\,\mathrm{PVU}$ on the 320 K isentrope) started wrapping around the cyclone, thereby probably further contributing to its intensification.

The cyclone reached its minimum pressure while travelling eastwards towards Corsica. It caused surface winds stronger than 50 km h^{-1}, and was associated with a surface front over the Mediterranean ocean. The alpine cross section (Figure 3.1d) indicates the moist surface air (RH > 80 %), that was advected towards the Alps. The air was then lifted on the southern flank of the Alps with velocities of about −1 Pa s^{-1}. This value is comparable to warm conveyor belts, which are typically associated with heavy precipitation. Oertel et al. (2019) analyse such warm conveyor belts and find a mean ascent rate of −0.5 Pa s^{-1} (with a range between 0 Pa s^{-1} and −4 Pa s^{-1}). Given these strong updrafts, the ascending moist air then led to heavy precipitation over the Alps. The reason for the catastrophic floods in Brig was that the rainfall was intense and that it continued for several days. Over other regions, the rain would either be weaker or it would come and go, but over Brig it was strong and it stayed for about 3 d. This is mainly due to the stationarity of the Mediterranean cyclone (see track in Figure 3.1b), which led to the sustained moisture transport towards the Alps.

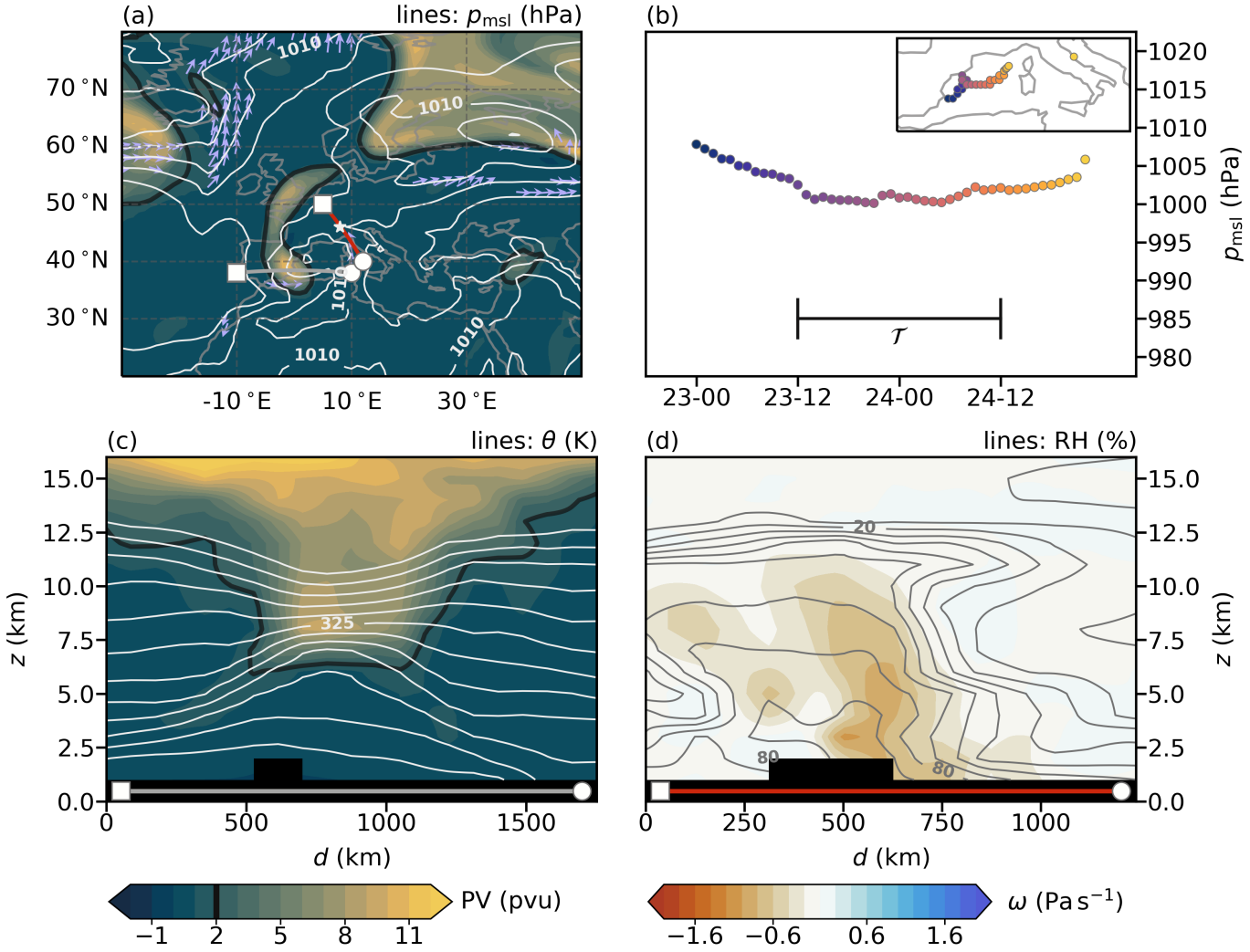


Figure 3.1: Synoptic situation on 23. September 1993, at 12 UTC. (a) PV (shading) and horizontal winds above 50 km h^{-1} (arrows) on the iso-surface $\theta = 320$ K. The black contour shows the 2 PVU iso-line, and the white contours indicate mean sea level pressure with an interval of 5 hPa. (b) Temporal evolution of the central pressure and the track of the Mediterranean cyclone that was associated with this event. The shading indicates time and the horizontal bar marks the evaluation time window ($\mathcal{T}$). (c) Vertical cross section of PV (shading) along the grey line in (a). White contours indicate potential temperature with an interval of 5 K. (d) Vertical cross section of vertical velocity (shading). The contours show relative humidity with 10 % intervals.

Spatial patterns

In Figure 3.2, we show the spatial patterns of mean sea level pressure at the time of cyclogenesis, for the ERA5 reference data and the long lead time forecasts. These are forecasts started 6 d to 4.5 d before the mature stage of the cyclone. Compared to ERA5, the forecasts predict lower pressure over northern Europe, particularly over the UK. This leads to stronger eastward winds over northern Europe. In addition, the pressure field over the northwestern Mediterranean ocean is too homogeneous. Over Spain it should be low but it is too high, and over Italy it should be high but it is too low. As a result the cyclonic winds in the northwestern Mediterranean ocean are too weak, also seen in Figure 3.3.

The location of the Brig cyclone is not well predicted by the long lead time forecasts (Figure 3.2). The deterministic forecasts (Figures 3.2a and 3.2b) and the ensemble hindcast (Figure 3.2f) predict a low pressure system with an eastward shift (over southeastern France and northern Italy). However, these systems are too weak as they are not isolated cyclones, but rather extensions of the falsely predicted large low pressure anomaly in the north. The picture is slightly different for the ensemble forecasts (Figures 3.2c and 3.2d). Some members actually get the location of cyclogenesis right, and the different contours of 1005 hPa mostly overlap. However, the cyclone forms 0.5 d too early, and the forecasted cyclones move too quickly towards the northeast (not shown). As they move, they quickly diverge from each other and they do not intensify as much as the ERA5 reference cyclone. The cyclone track is thus unpredictable for the long lead time forecasts. In addition to the uncertain track, the shorter lead time ensemble forecast (Figure 3.2d) also predicts lower pressure in northern Africa. In the following day, these low pressure anomalies develop into cyclones, that are not present in the ERA5 data, leading to very different circulation patterns in the Mediterranean (not shown).

Given that the forecasts underestimate the intensity of the cyclone, or predict it too far to the east, the subsequent wind speeds in the region south of the Alps are underestimated by more than $6\,\mathrm{m\,s^{-1}}$ (Figure 3.3). This implies weaker transport of warm and moist ocean air towards Brig. The shorter lead time deterministic forecast (Figure 3.3b) and the ensemble hindcast (Figure 3.3f) instead show a region of strong winds shifted either to the west or to the east. But all other forecasts simply underestimate the cyclonic winds. The strong winds are instead predicted in the north, in the region of the (false) low pressure anomaly (seen for all forecasts in Figure 3.3). As with pressure, the different ensemble members disagree on the location of strong winds (Figures 3.3c, 3.3d, and 3.3f), implying either low predictive skill of the ensemble or low intrinsic predictability.

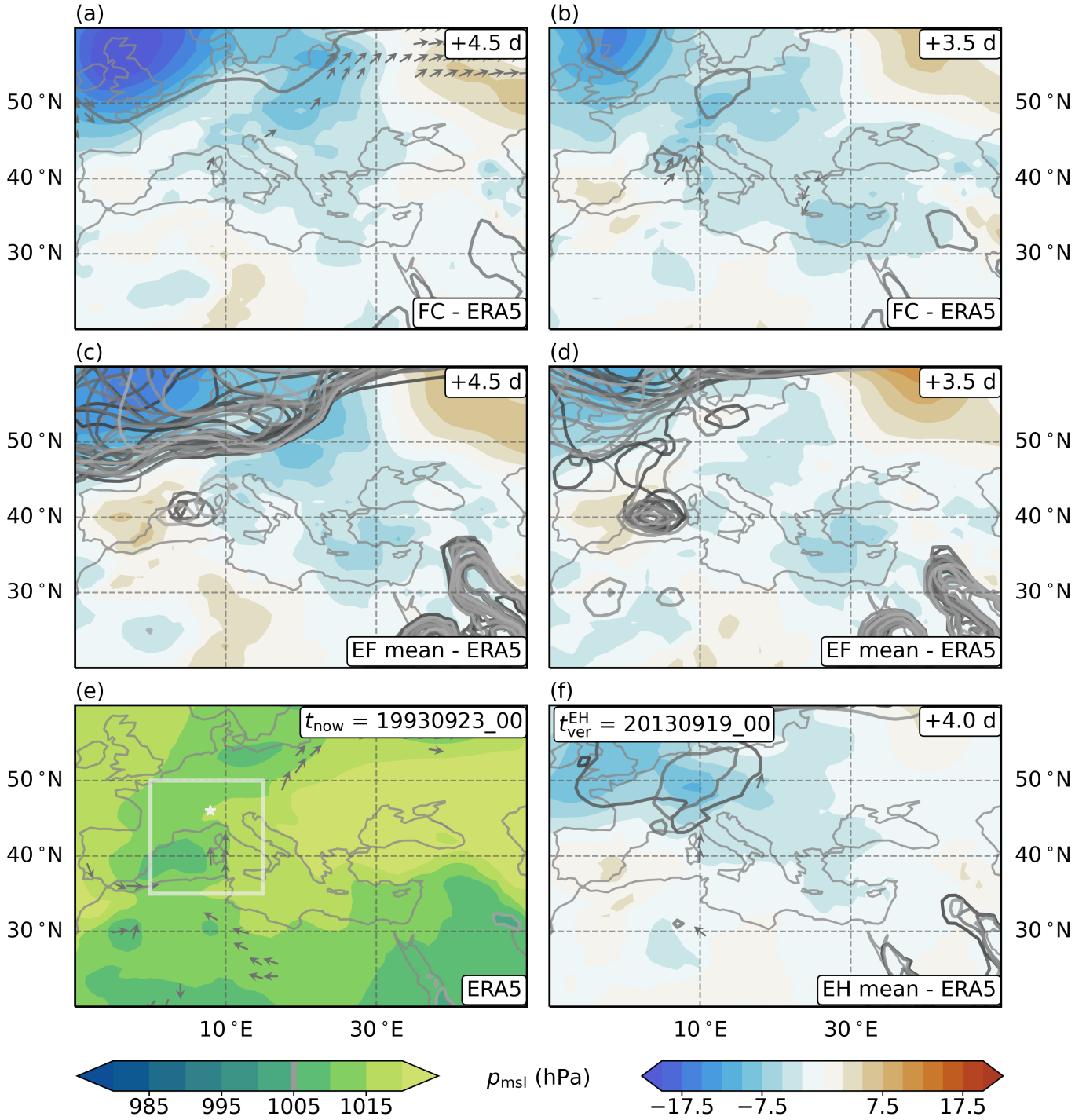


Figure 3.2: Long lead time forecasts of pressure for 1993, 1 d before the (middle of the) most intense cyclone stage. (a) and (b) deterministic forecasts (FC), (c) and (d) ensemble forecasts (EF), (e) ERA5 reference, and (f) ensemble hindcast (EH). We show the difference of the forecasts to the ERA5 reference data. The left column is for longer lead time forecasts, and the right column for shorter lead time forecasts, started 1 d later. The grey arrows indicate horizontal winds above 50 km h^{-1} on 900 hPa. The white markings in (e) indicate the evaluation region $\mathcal{R}$ as mentioned in section 2.2.2 (rectangle), and the nearest grid point to Brig (star). The pressure value of 1005 hPa is emphasised with a grey line, or multiple lines in different shades of grey (one for each ensemble member). Note that we show the mean field for the ensembles, which leads to smoother distributions compared to the deterministic forecasts. The ensemble forecasts include more members than the ensemble hindcasts (compare Table 2.3).

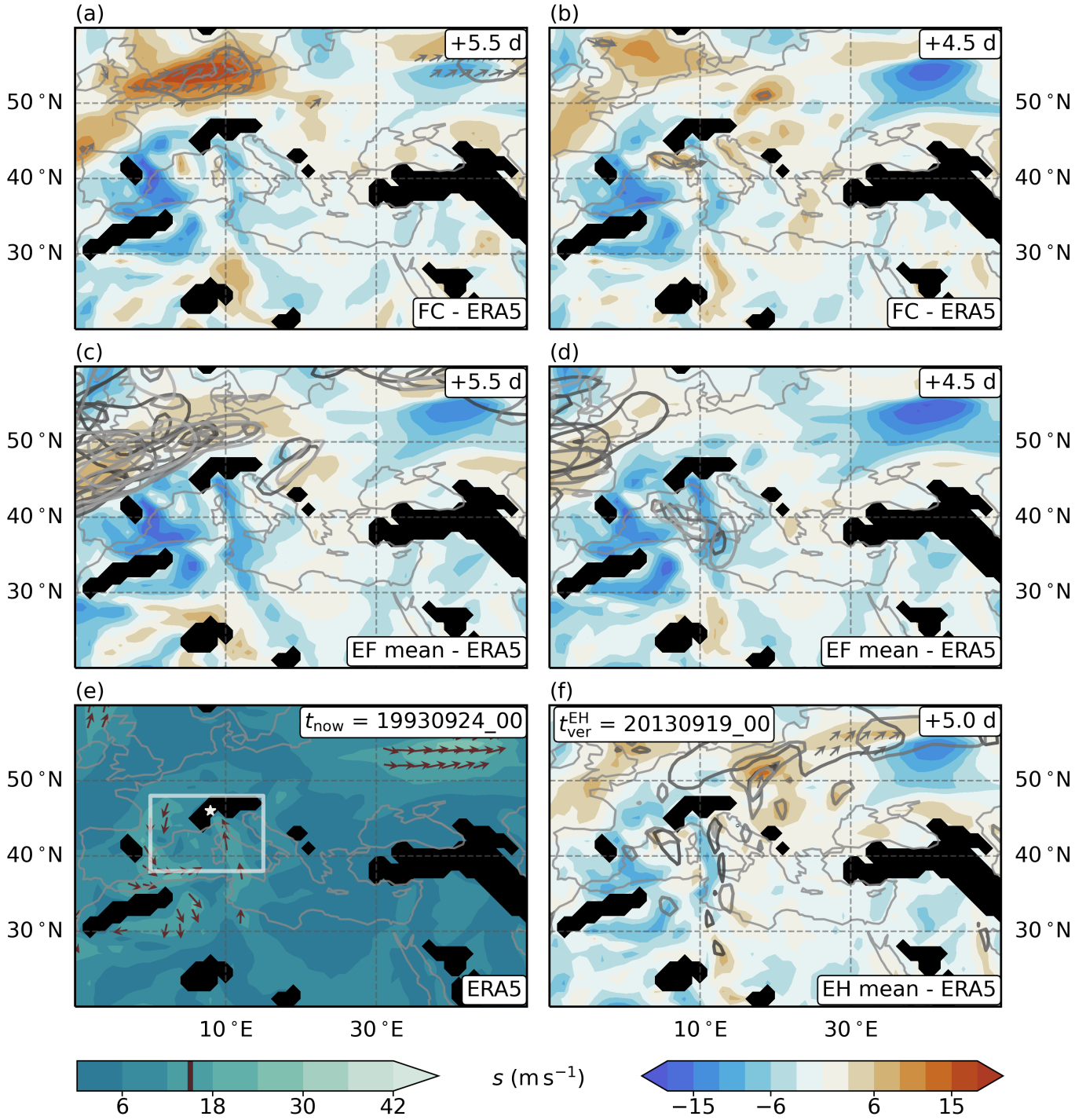


Figure 3.3: Long lead time forecasts of wind speed for 1993, in the middle of the most intense cyclone stage. Similar to Figure 3.2, but for a different reference time, and for horizontal wind speed on 900 hPa. The emphasised contour corresponds to 15 m s^{-1}.

Since the cyclones are too far east, the later precipitation also shows an eastward displacement, clearly seen in Figures 3.4a, 3.4b, and 3.4f. Most precipitation falls over central to northern Italy and over eastern Switzerland. In addition to the shift in location, the precipitation also extends over a smaller region than in the ERA5 reference data. These two factors combined result in considerable underestimation of precipitation near Brig, for the long lead time forecasts.

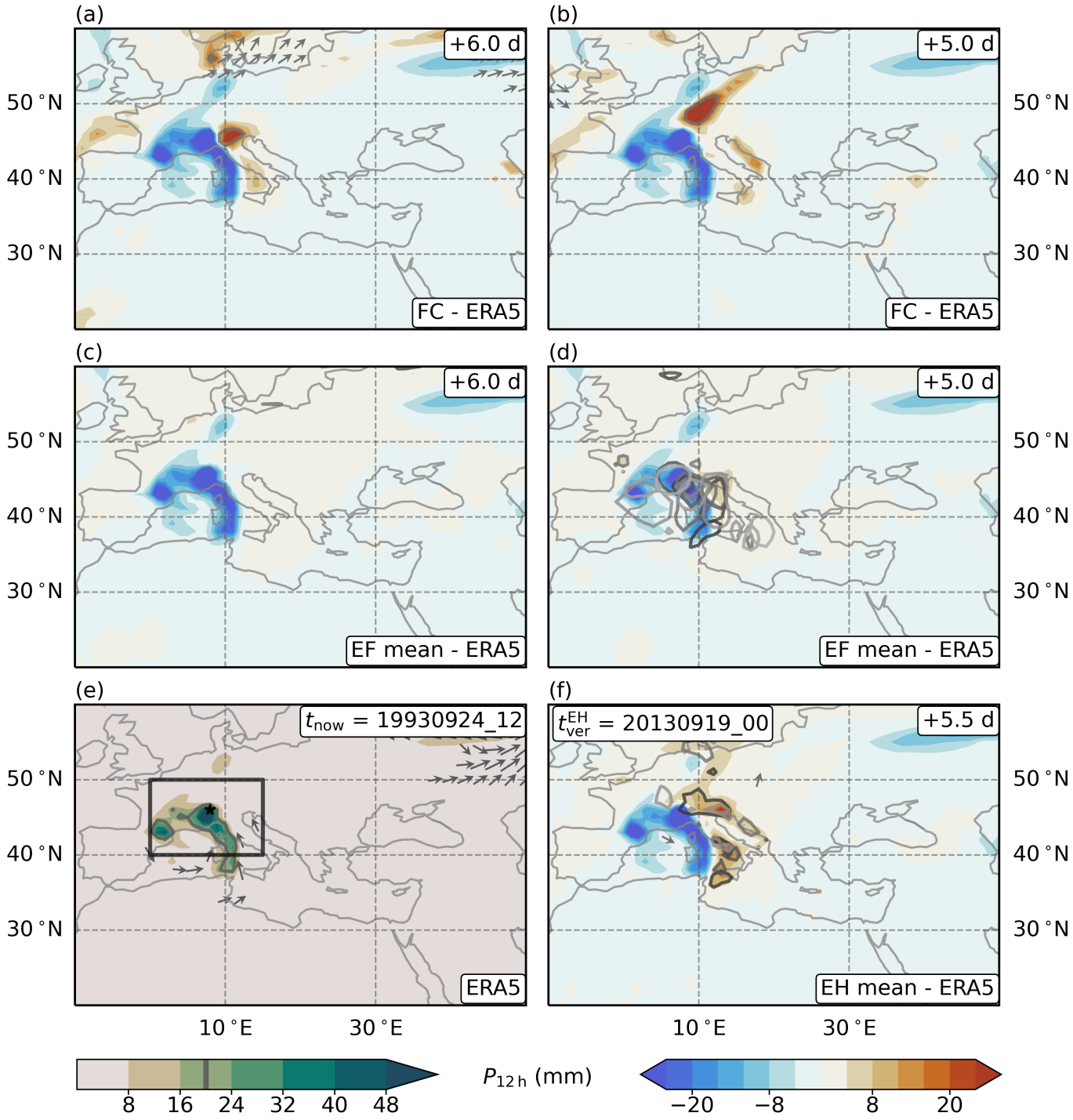


Figure 3.4: Long lead time forecasts of precipitation for 1993, at the end of the most intense cyclone stage. Similar to Figure 3.2, but for a different reference time, and for 12 h accumulated precipitation. The emphasised contour corresponds to 20 mm.

Similar to wind speed, the deterministic forecasts (Figures 3.4a and 3.4b), and the ensemble hindcast (Figure 3.4f) show a (north-) eastward shift of precipitation, while the ensemble forecasts (Figures 3.4c and 3.4d) tend to simply underestimate precipitation. The members of the ensemble forecasts also show high variability in the location of precipitation, from southern Italy to southern France (Figure 3.4d). In contrast, the members of the ensemble hindcast predict precipitation cells in more similar locations (Figure 3.4f). Still, it is clear that the long lead time forecasts do not manage to accurately predict the Mediterranean cyclone, and that there are large uncertainties.

The inaccuracies of the long lead time forecasts are less noticeable for the middle lead time forecasts (shown in Figures B.2 to B.4). These are forecasts started 4 d to 3.5 d before the mature stage of the cyclone. They no longer predict the same low pressure area over northern Europe (Figure B.2), and some ensemble members now correctly predict the stationarity of the Mediterranean cyclone (not shown). However, there are remaining uncertainties in the track, with different members predicting different tracks (not shown). The ensemble forecast is clearly less accurate than the ensemble hindcast, with cyclones that are too weak and too spread out in location. This is also true for wind speed. In fact, the ensemble forecast is the only forecast that does not predict any phases with high wind speeds south of the Alps (not shown). In terms of precipitation, the main benefit of the middle lead time forecasts is a generally smaller shift of the precipitation cells (Figure B.4). But the shift is still too large, so that the actual precipitation field is not accurately predicted.

For the short lead time forecasts (2.5 d to 0.5 d before the mature stage), both the time offset and the diverging cyclone positions are smaller issues (not shown). They generally predict the mean sea level pressure with high accuracy (Figure B.5), but the horizontal wind speed is either underestimated or overestimated south of the Alps (Figure B.6). The short lead time forecasts of precipitation are illustrated in Figure 3.5. They show a smaller and smaller eastward shift (compare Figures 3.5c and 3.5d to Figures 3.5e and 3.5f), but they still do not get the distribution entirely right. Precipitation is underestimated in some regions, while being overestimated it in others (blue and read areas in Figure 3.5). In addition, they predict smaller cells of rainfall compared to the ERA5 reanalysis. Overall, although pressure and wind speed are reasonably accurate, no forecast manages to correctly predict the main precipitation cells that caused the heavy flooding in Brig.

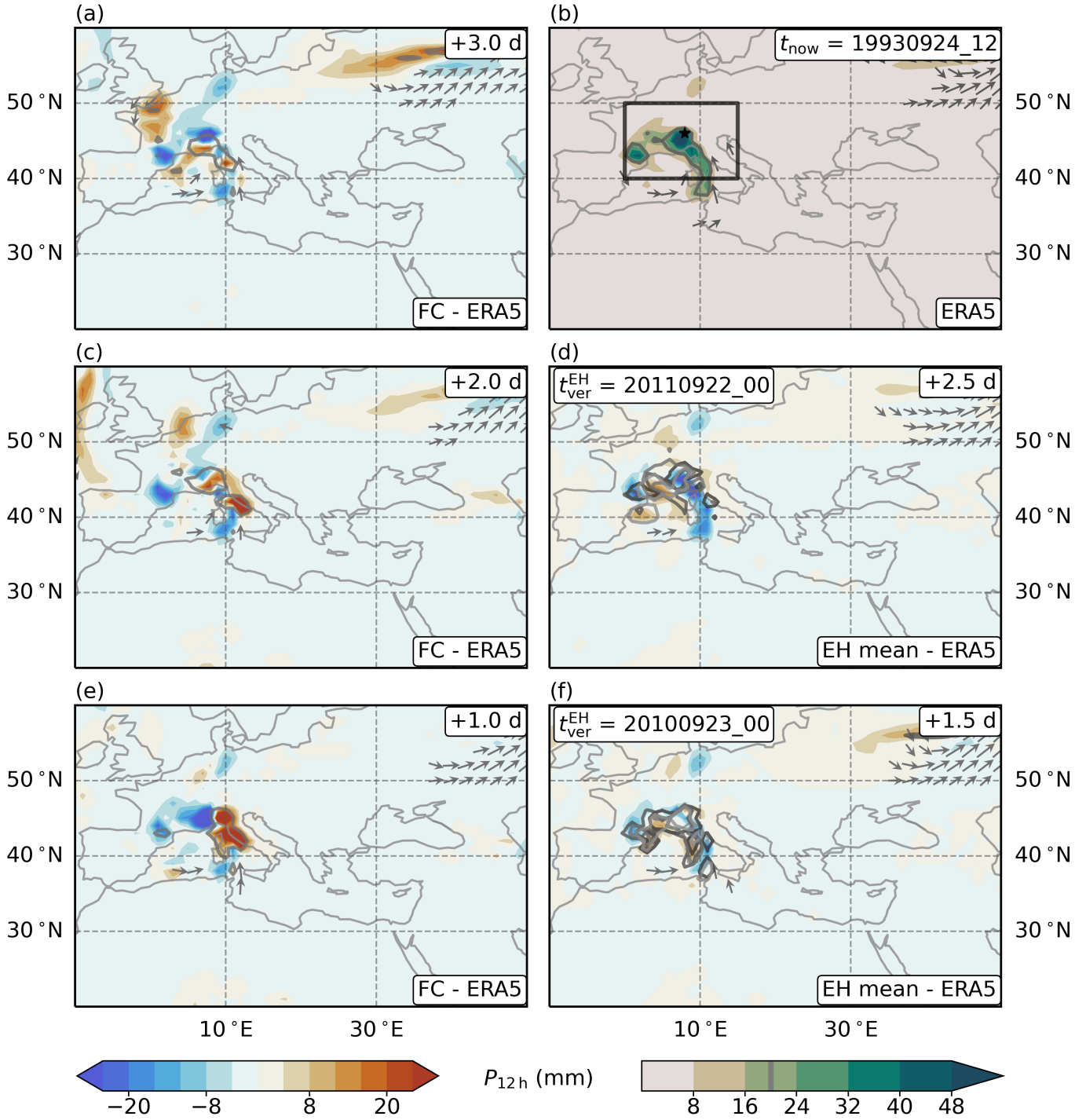


Figure 3.5: Short lead time forecasts of precipitation for 1993, at the end of the most intense cyclone stage. Similar to Figure 3.4, but with lead time decreasing from top to bottom. (a), (c), and (e) deterministic forecasts (FC), (b) ERA5 reference, (d) and (f) ensemble hindcasts (EH). There is no ensemble forecast available for this initialisation time.

Temporal evolution

To show the evolution of forecasts over time, we eliminate the spatial dimensions by taking a mean over the evaluation regions shown in Figures 3.2e, 3.3e, and 3.4e. It is then possible to show time series for mean sea level pressure, horizontal wind speed, and precipitation. Figure 3.6 illustrates this data for forecasts at a long lead time.

Even in the regional mean, it is clear that the long lead time forecasts predict the cyclogenesis of the Mediterranean cyclone too early. This is visible in all three fields, with the pressure minimum (Figure 3.6a) and the peaks in wind speed (Figure 3.6b) and rainfall (Figure 3.6c) occurring about 0.5 d too soon (shifted to the left). In addition, the amplitudes of the time series are mostly underestimated. For pressure, it is only the ensemble forecast (blue circles in Figure 3.6a), that does not predict the amplitude accurately. But for wind speed and precipitation (Figures 3.6b and 3.6c), all forecasts have an amplitude of at most 75 % relative to the ERA5 reference data[1]. The rates of intensification and dissipation are typically too fast. Especially the quick decay leads to large errors, since the low pressure and strong winds do not last long enough, causing precipitation to quickly die down.

In general, the spread of the ensembles (for both forecast and hindcast) increases during the phase of intensification, and then decreases again during the phase of dissipation. This indicates that the most intense phase of the event is particularly uncertain, and that the upstream atmospheric conditions are not yet settled enough to accurately predict the cyclone and its intensification. In terms of accuracy, the ensemble forecast preforms worst, then the deterministic forecast, and then (only slightly better) the ensemble hindcast. Because of the timing offset, and the quick intensification/dissipation rates, the ERA5 values are typically outside the middle 50 % of the ensemble values (outside the box). This is true for both the ensemble forecasts (blue boxplots) and the ensemble hindcasts (red boxplots). For wind speed and precipitation (Figures 3.6b and 3.6c), they are even mostly outside the middle 90 % (outside the whiskers).

The pattern for the middle lead time forecasts is similar (Figure B.7). The main differences are that the deterministic forecast is generally more accurate (smaller amplitude error, more accurate intensification rate), and that the ensemble hindcast no longer shows any timing errors in pressure or precipitation. Except for somewhat smaller spreads in pressure, the ensemble forecast does not noticeably improve.

The time series for the short lead time forecasts are shown in Figure 3.7. These forecasts are much more accurate. There are no timing errors, the amplitudes are mostly correct, and the intensification rate is well predicted. The only remaining issue is that the strong winds (Figure 3.7b) and intense precipitation (Figure 3.7c) are still underestimated. But the errors are less than half as large as before.

[1] Note that the amplitudes refer to the coloured markers presented in the legend of Figure 3.6. For the deterministic forecast, we refer to the distance between the smallest value and the largest value. But for the ensembles, we only consider the means (blue circles, red triangles), not the full data of all ensemble members.

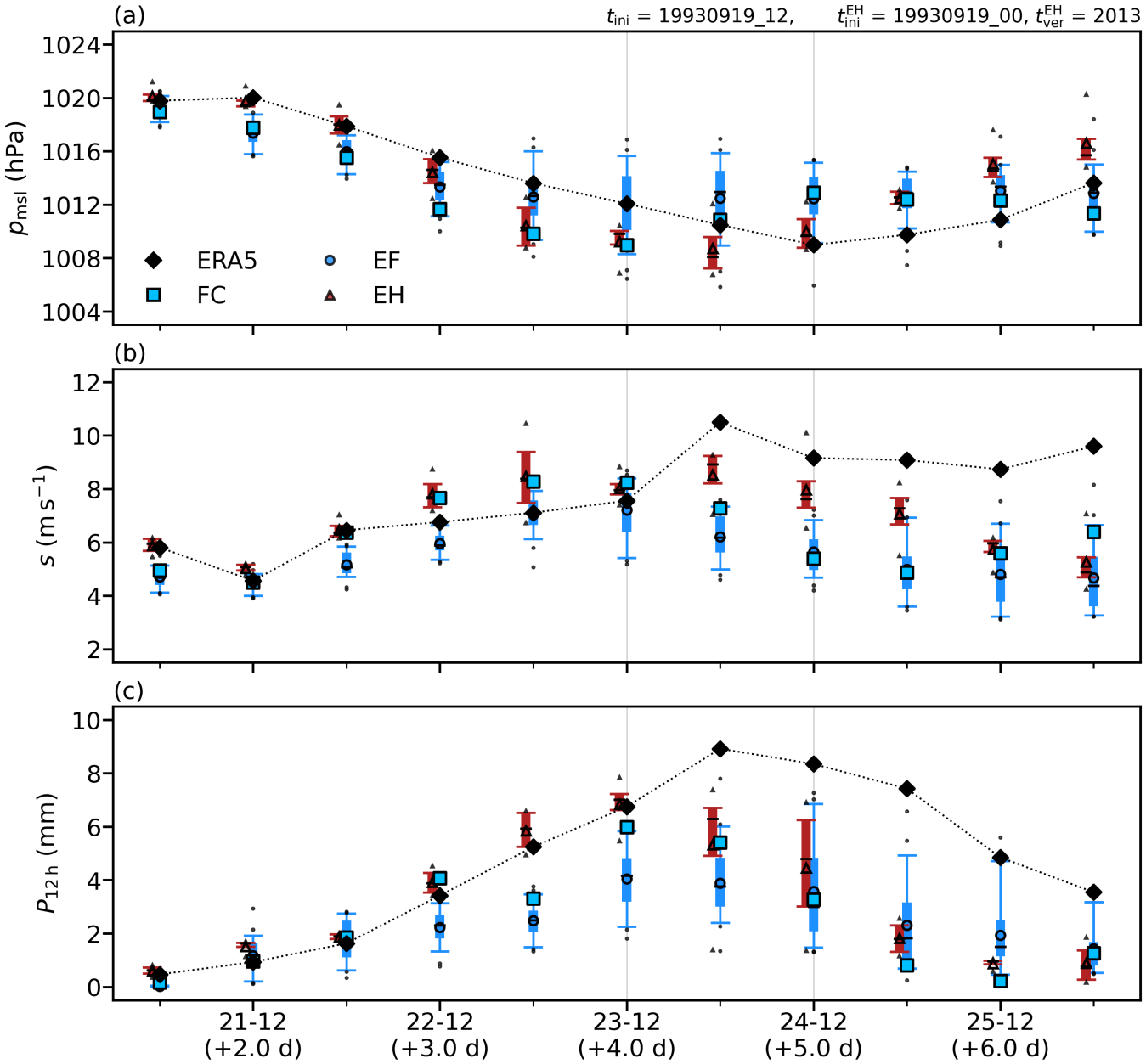


Figure 3.6: Time series of forecasts at a long lead time for 1993. (a) Mean sea level pressure, (b) horizontal wind speed on 900 hPa, (c) 12 h accumulated precipitation. All fields are spatial means over the evaluation regions shown in Figures 3.2e, 3.3e, and 3.4e. The forecasts were initialised on 19. September 1993, the ensemble hindcast (EH) at 00 UTC, and both the deterministic forecast (FC) and the ensemble forecast (EF) at 12 UTC. The lead times are listed relative to the later initialisation time of the operational forecasts (FC and EF). The ERA5 reference data is shown with black diamonds (and a dotted line), FC data with blue squares, EF means with blue circles, and EH means with red triangles. The boxplots indicate the following quantiles: 5 % (lower whisker), 25 % (lower box edge), 50 % (line inside box), 75 % (upper box edge), 95 % (upper whisker). The grey vertical lines indicate the chosen evaluation time window, over which we later take a temporal mean (as mentioned in section 2.2.2).

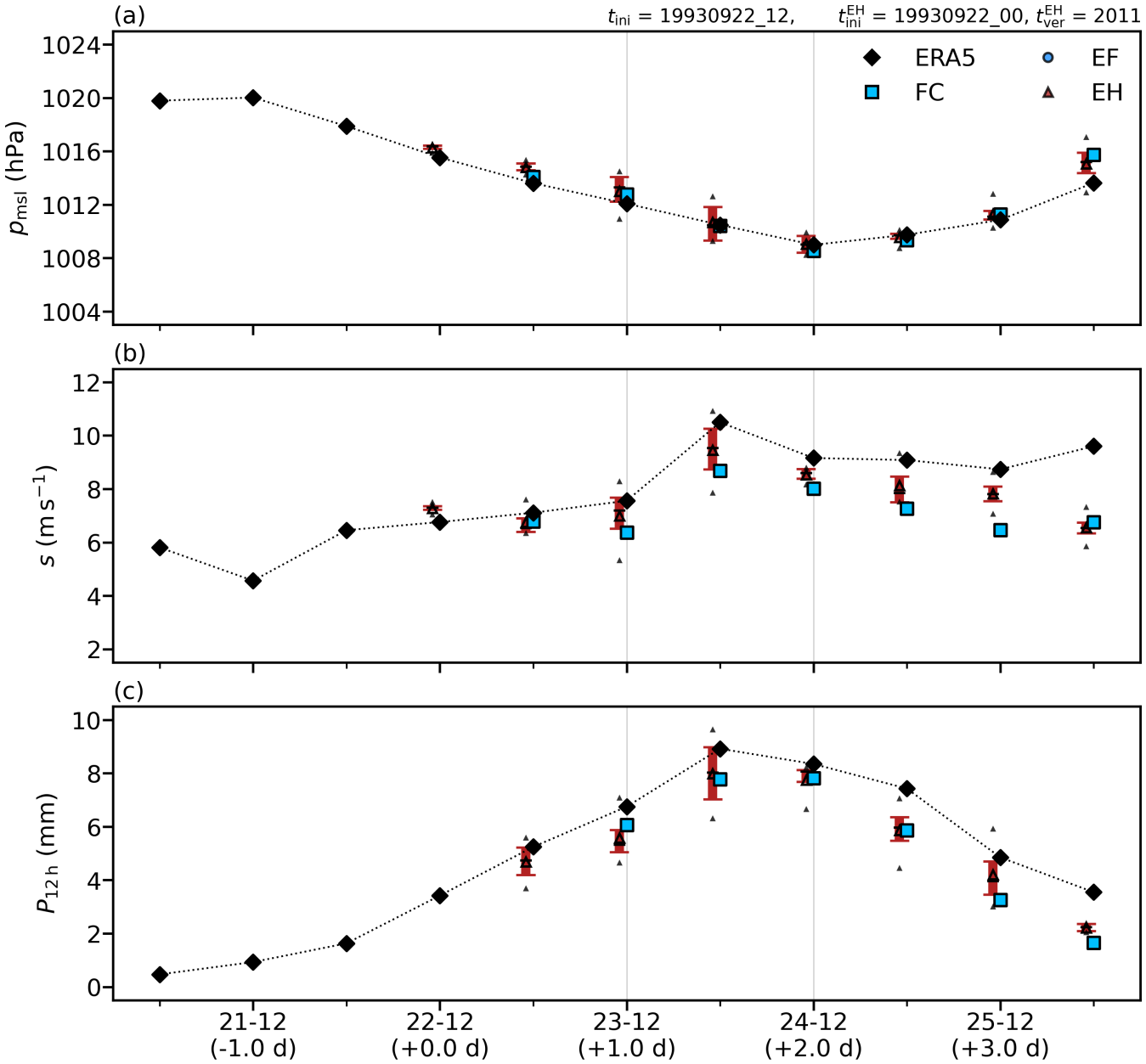


Figure 3.7: Time series of forecasts at a short lead time for 1993. Similar to Figure 3.6, but for forecasts at a short lead time. The forecasts were initialised on 22. September 1993, the ensemble hindcast (EH) at 00 UTC, and the deterministic forecast (FC) at 12 UTC. There is no ensemble forecast available for this initialisation time. The lead time of 0 d corresponds to the later initialisation time of the operational forecasts (FC and EF). Negative lead times indicate timesteps before the initialisation of the operational forecasts.

Lead time dependence

As a final step, we compute three representative values for each forecast, one for mean sea level pressure, one for horizontal wind speed, and one for precipitation. This is done by first choosing a specific forecast (for example the deterministic forecast initialised on 18. September 1993, at 12 UTC), then taking the regional means over the evaluation regions (shown in Figures 3.2e, 3.3e, and 3.4e), and finally averaging the regional means over the evaluation time window (shown in Figures 3.6 and 3.7). After computing these quantities for all forecasts, we can compare the different forecast types and initialisation times. Figure 3.8 illustrates all the available forecasts, with lead time on the horizontal axis, and the ERA5 reference value shown in black. The long lead time forecasts are on the left, and the short lead time forecasts are on the right.

The deterministic forecast correctly predicts the mean sea level pressure within 1 hPa, almost independent of forecast lead time (Figure 3.8a). The ensemble forecast is not as accurate, with pressure being overestimated by at least 1.5 hPa. It is only the ensemble spread that slightly improves with smaller lead times. Compared to the ensemble forecasts, the ensemble hindcasts are more accurate (errors within 1 hPa) and less spread out. The ensemble means improve with smaller lead time, but there is no clear reduction in spread.

For wind speed, the situation is slightly different (Figure 3.8b). The deterministic forecast generally underestimates the ERA5 reference value. Longer lead time forecasts show larger errors (about $2\,\mathrm{m\,s^{-1}}$), and shorter lead time forecasts show smaller errors (about $1\,\mathrm{m\,s^{-1}}$). The ensemble forecast is again less accurate, with an error beyond $2.5\,\mathrm{m\,s^{-1}}$. Only the ensemble spread slightly improves with smaller lead times. The ensemble hindcasts are the most accurate. The hindcast means agree with the ERA5 reanalysis within $1\,\mathrm{m\,s^{-1}}$, almost independent of lead time. The shortest lead time hindcast is a bit better than the other ones, but the spread does not noticeably improve.

For precipitation, there are clear improvements with smaller lead times (Figure 3.8c). The deterministic forecast approaches the true value asymptotically, with longer lead time forecasts showing an error of about 6 mm and shorter lead time forecasts being within 1 mm of the ERA5 value. The ensemble forecast underestimates the mean precipitation even more, and the available data shows less improvement with smaller lead times. Again, the ensemble hindcasts are the most accurate. The ones at longer lead times have errors of about 2.5 mm, and the one at the shortest lead time is again within 1 mm of the true value. Their spread is smaller compared to the ensemble forecasts, but it does not noticeably improve with smaller lead times.

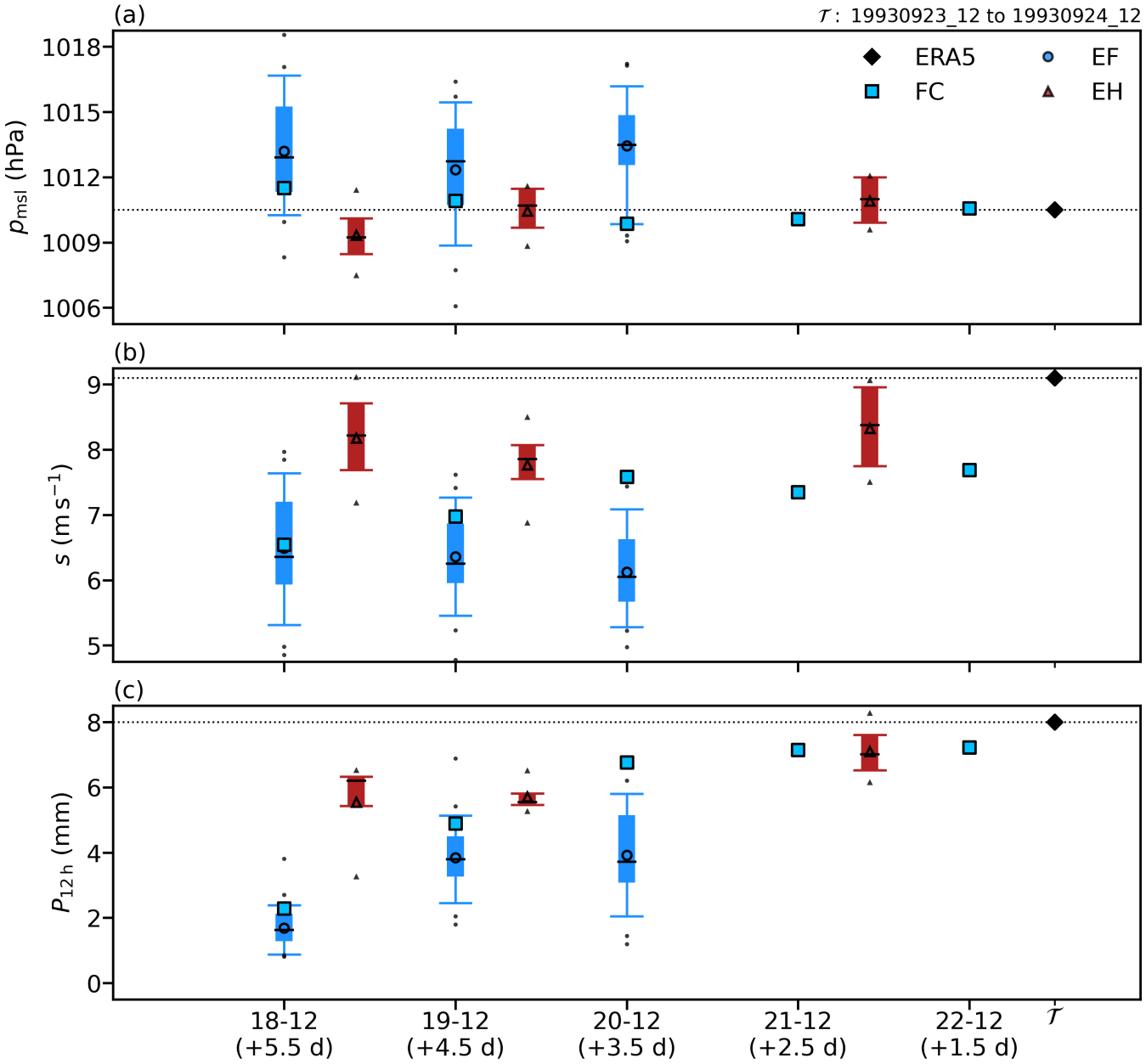


Figure 3.8: Lead time improvements for 1993. (a) Mean sea level pressure, (b) horizontal wind speed on 900 hPa, (c) 12 h accumulated precipitation. Similar to Figure 3.7, but with forecast lead time on the horizontal axis. The longest running forecasts are on the left, and the shortest running forecasts on the right. Note that the deterministic forecasts and the ensemble forecasts are always initialised at 12 UTC, while the ensemble hindcasts are always initialised at 00 UTC. The black diamond and the dotted line show the regional and temporal mean of the ERA5 reference data. The regions are as in Figures 3.2e, 3.3e, and 3.4e, and the evaluation time window is shown in Figures 3.6 and 3.7.

3.2 Jerusalem March 1998

A strong dust storm reached Jerusalem (Israel, 35.21 °E, 31.77 °N) on 16. March 1998. As reported by Struck (1998), it was the worst storm in three decades, bringing strong winds around 90 km h^{-1} followed by hail and about 100 mm of wet snow. The event led to a general shutdown of the city. With a minimum central pressure of 981 hPa, the Jerusalem cyclone was one of the most intense Mediterranean cyclones between 1992 and 2021 (top 0.2 %, Figure B.8b). It was also very persistent with a lifetime of 110 h (top 0.6 %, Figure B.8a). However, the mean precipitation rate was not as extreme (0.22 mm h^{-1}, top 25.8 %, Figure B.8c). Similarly, the accumulated precipitation (24.4 mm, top 6.5 %, Figure B.8d) was not particularly large, especially when considering the long lifetime of the cyclone, travelling from northwestern Africa all the way to Jerusalem (inset map in Figure 3.9b). Therefore, we do not discuss precipitation for this case, and only focus on mean sea level pressure and horizontal wind speed.

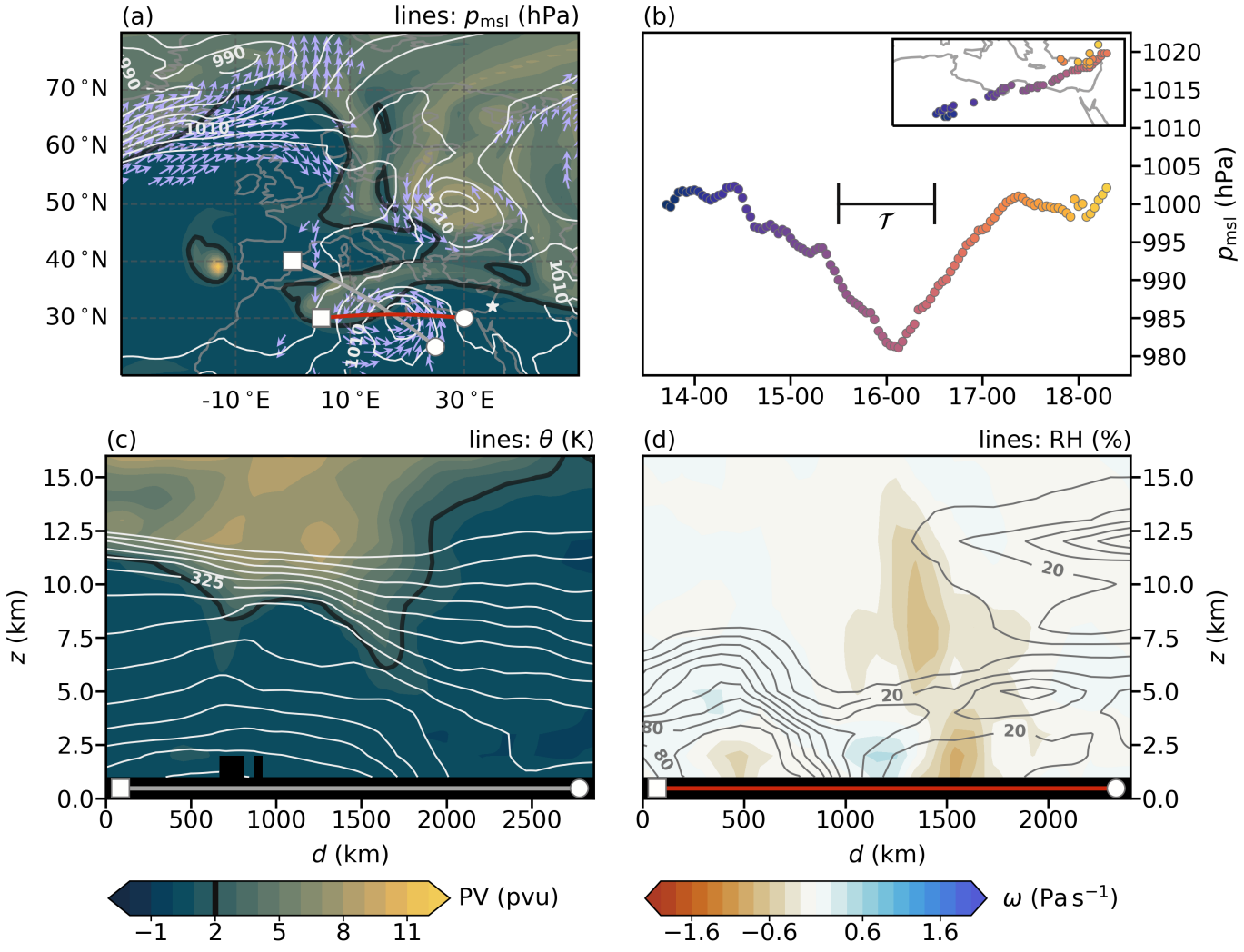


Figure 3.9: Synoptic situation on 15. March 1998, at 00 UTC. Similar to Figure 3.1, but for a different reference time, different vertical cross sections, and a different cyclone track.

Similar to the Brig event (section 3.1), there was a large anticyclone ($p_{max} > 1040$ hPa) over the North Atlantic which induced a far south-reaching PV streamer. Figure 3.9a shows a rather unusual east-west orientation of the streamer, reaching from eastern Europe down to northern Africa. As for the Brig event, the streamer reached a minimum altitude of about 6 km (Figure 3.9c). Although the PV values were smaller (mostly below 6 PVU), this does not necessarily imply that the streamer was

less intense. A possible reason for the smaller PV values is a generally higher surrounding tropopause, as seen in the top right of Figure 3.9c. However, Barnes et al. (2022) find that PV streamers reaching down to similar altitudes also lead surface wind fields at similar intensities, mostly independent of the tropopause height.

After the cyclone formed to the east of the streamer, it intensified rather quickly with a pressure drop of about 15 hPa within 1 d (Figure 3.9b). Following Sanders and Gyakum (1980), this rate corresponds to 1.08 Bergeron units[2], implying that the Jerusalem cyclone was a "bomb". As seen in Figure 3.9c, a clear frontal structure emerged at the surface. The west-east cross section in Figure 3.9d shows that the surface air was rather dry (RH $< 20\,\%$) and that the vertical winds were mostly weaker than $-0.8\,\mathrm{Pa\,s^{-1}}$. However, during the most intense phase of the cyclone, the vertical winds increased to $-1.5\,\mathrm{Pa\,s^{-1}}$.

Spatial patterns

In Figure 3.10, we show the long lead time forecasts (8 d to 6.5 d before the mature stage) for mean sea level pressure. As seen in Figure 3.10b, a low pressure system is starting to form over northern Africa. This is where the ERA5 cyclone forms. The long lead time forecasts of pressure have three main issues. The first issue is that the forecasts predict cyclogenesis at least 1 d too early. The only exception is the shorter lead time ensemble hindcast (Figure 3.10f), where there is not yet a closed pressure contour of 1000 hPa.

The second problem is that the region of cyclogenesis is uncertain. It is correctly predicted in northern Africa for the deterministic forecast (Figure 3.10a), and for some members of the ensemble forecasts (Figures 3.10c and 3.10d). But for the other members of the ensemble forecasts, and for all members of the ensemble hindcasts (Figures 3.10e and 3.10f), the location of cyclogenesis is falsely predicted over the Mediterranean ocean or central Europe. This northeastwards shift of the cyclone then remains an issue along most of the cyclone track (not shown).

The third issue is that the forecasts miss the low pressure system in northeastern Europe. This system is associated with eastward winds, which later accelerate the cyclone coming from northern Africa over the Mediterranean ocean (not shown). Due to its absence in the forecasts, the cyclones generally move too slowly towards the east (not shown). Overall the cyclones thus form too early and downstream of the ERA5 cyclone. But then they move too slowly allowing the ERA5 cyclone to "catch up". Nevertheless, the forecasts still do not accurately predict the final position of the cyclone as it approaches Jerusalem (not shown). This is because there is a lot of uncertainty in the cyclone track, leading to diverging cyclone locations.

The situation is very similar for wind speed (Figure B.9). In general, the strong winds are shifted northeastwards as the cyclone forms and intensifies, and they are shifted northwestwards as the cyclone reaches Jerusalem (not shown). The location of strong winds is also very uncertain, with a lot of variance between the different ensemble members.

The middle lead time forecasts (6 d to 4.5 d before the mature stage) perform better in some aspects (Figures B.10 and B.11). There are still some cyclones that form too early in northern Africa (not shown), and some that form at the right time but instead in the Mediterranean ocean (Figures B.10c to B.10e). But most forecasts now accurately predict the time and location of cyclogenesis. The only inaccurate ones are the longer lead time ensemble hindcasts (Figures B.12e and B.12g). The northeastward shift of low pressure also remains an issue. However, the shift is much smaller and the pressure is more accurate at the true location of the cyclone. Furthermore, the cyclone track is much less uncertain, with most ensemble members showing similar pressure contours of 1000 hPa

[2] A Bergeron unit is defined as an intensification rate of 24 hPa in 1 d, multiplied by the latitude correction factor $\sin(\varphi)/\sin\left(\frac{\pi}{3}\right)$, where φ is the geodetic latitude.

(Figures B.10c, B.10d, and B.10f). Similar improvements also hold for wind speed, with a smaller eastward shift and less uncertainty in the location of strong winds (Figure B.11).

The short lead time forecasts (4 d to 2.5 d before the mature stage) predict pressure with high accuracy. All three issues mentioned above are resolved (partly shown in Figure B.12). However, wind speed is still not entirely accurate (Figure B.13). There are still some shifts, which result in considerable underestimation of winds at the true locations. This problem is most noticeable for the ensemble forecasts and the ensemble hindcasts.

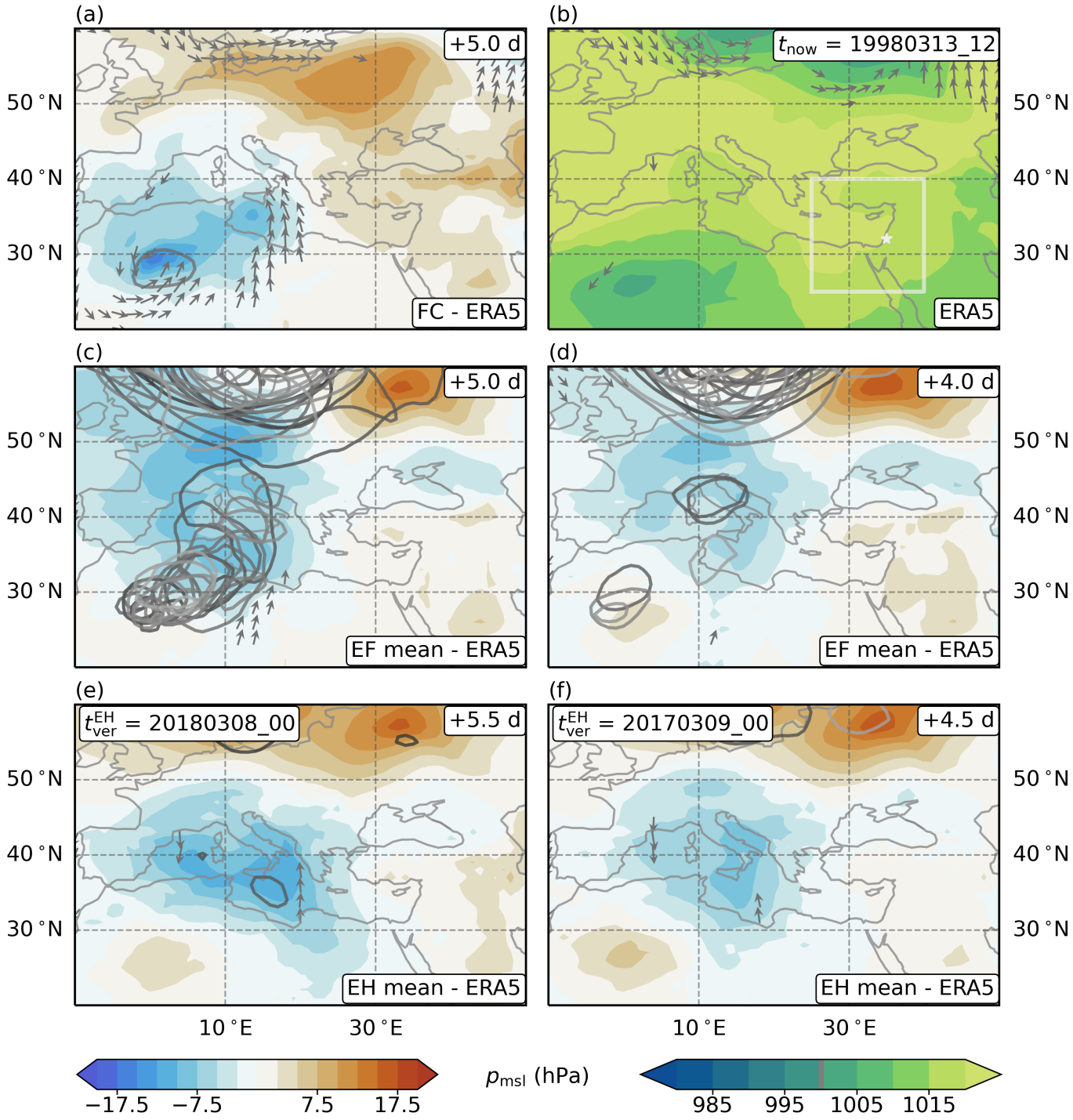


Figure 3.10: Long lead time forecasts of pressure for 1998, 2.5 d before the (middle of the) most intense cyclone stage. Similar to Figure 3.2, but for a different reference time.

Temporal evolution

The evaluation regions for the time series are defined over the eastern Mediterranean ocean including the coast near Jerusalem. The region for pressure is shown in Figure 3.10b, and the region for wind speed is very similar (Figure B.9b). This means that we mostly capture the later phase of the Jerusalem cyclone in the time series. As mentioned above, this later phase is different from the earlier phase. There are smaller shifts in the fields, since the ERA5 cyclone "catches up" to the forecasted cyclones, but there is generally high uncertainty in the precise location of the cyclone and the strong winds (not shown). These features are also seen after taking the regional mean (Figure 3.11).

The two time series in Figure 3.11 show that the long lead time forecasts underestimate the intensity of the cyclone, with all forecasts preforming similarly. The deterministic forecast generally overestimates pressure and predicts the pressure minimum 0.5 d too late (Figure 3.11a). This timing error is also present for wind speed (Figure 3.11b). The value of minimum pressure is accurately predicted, while the maximum wind speed is underestimated. The intensification rate is also accurate for pressure, but too small for wind speed.

For the ensemble forecast and the ensemble hindcast, both pressure amplitudes (Figure 3.11a) and wind speed amplitudes (Figure 3.11b) are too small. For pressure they are about 60 % to 80 % of the ERA5 amplitude, while for wind speed the amplitude is only about 40 % of the true value. During the most intense phase of the cyclone (on 16. March, at 00 UTC), the ERA5 reference values are outside the middle 90 % of the ensembles. Otherwise, the 90 % ranges typically include the ERA5 values. The ensembles generally show intensification rates that are too weak, and later phases that are too constant.

Similar to the Brig cyclone (section 3.1), the spread of the ensembles tends to increase during the intense phase of the cyclone. In the case of pressure (Figure 3.11a), this is clearly seen for both the ensemble forecast and the ensemble hindcast. For wind speed (Figure 3.11b), however, this is only true for the ensemble forecast. The ensemble hindcast instead shows large spreads already 2 d before the most intense phase.

The middle lead time forecasts show many improvements (Figure B.14). All forecasts are much more accurate for pressure (Figure B.14a). The deterministic forecast and the ensemble forecast only show notable errors during the last phase of the cyclone. The amplitudes and the intensification rates are close to the ERA5 data. In contrast, the ensemble hindcast now predicts the minimum about 0.5 d too early, and still underestimates the amplitude (about 80 % of the true value). Overall, there is a clear reduction in ensemble spread by about 50 %. In terms of wind speed (Figure B.14b), the improvements are less apparent. Although the deterministic forecast gets the time series of wind speed almost entirely right, and the ensembles are much more accurate before the phase of strong winds, both the ensemble forecast and the ensemble hindcast still underestimate the winds during the intense phase of the cyclone. Overall, there is also no clear reduction of ensemble spread during this phase.

The pressure evolution is predicted very accurately by the short lead time forecasts (Figure B.15a). But there are only small improvements for wind speed (Figure B.14b). The phase of strong winds is still most accurately predicted by the deterministic forecast, with both ensembles underestimating the true amplitude. There is only a slight reduction in the spread of the ensembles, and the main difference to the middle lead time forecasts is that the phase before the strong winds is more accurate, particularly for the ensemble hindcast.

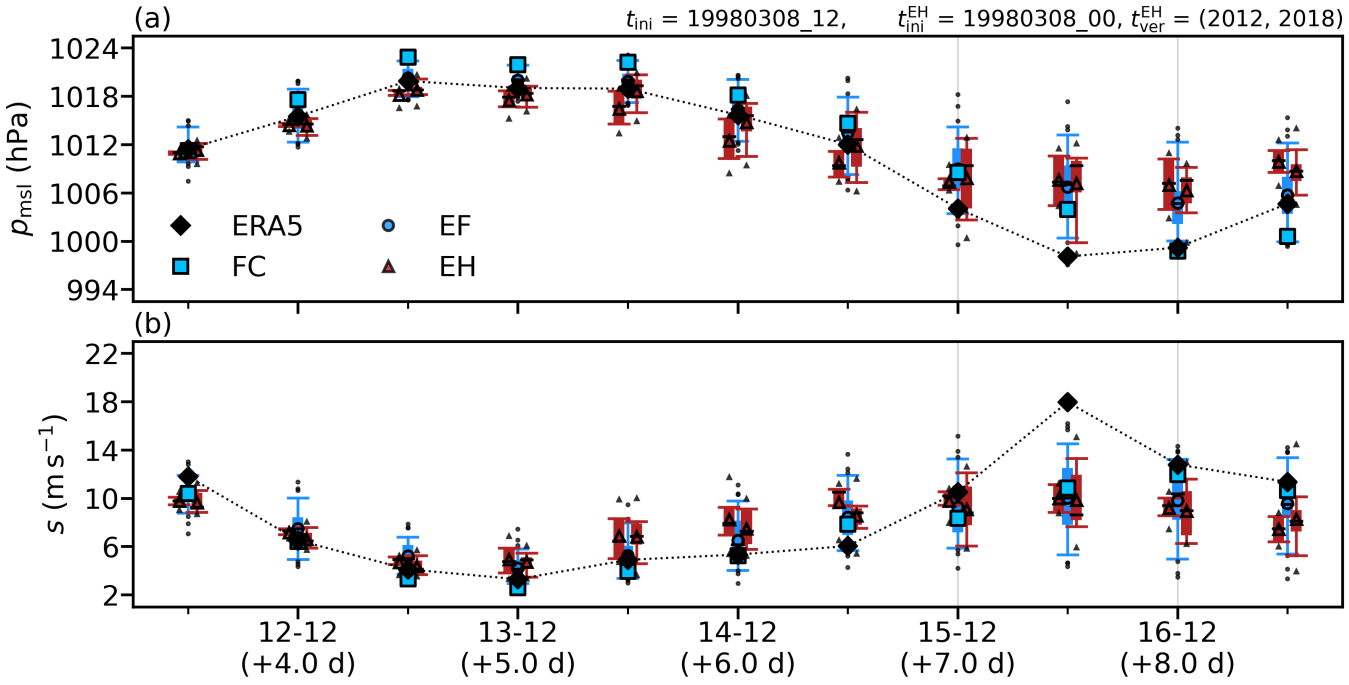


Figure 3.11: Time series of forecasts at a long lead time for 1998. Similar to Figure 3.6, but for a different reference time, and different evaluation regions shown in Figures 3.10b and B.9b. The forecasts were initialised on 8. March 1998, the ensemble hindcasts (EH) at 00 UTC, and both the deterministic forecast (FC) and the ensemble forecast (EF) at 12 UTC.

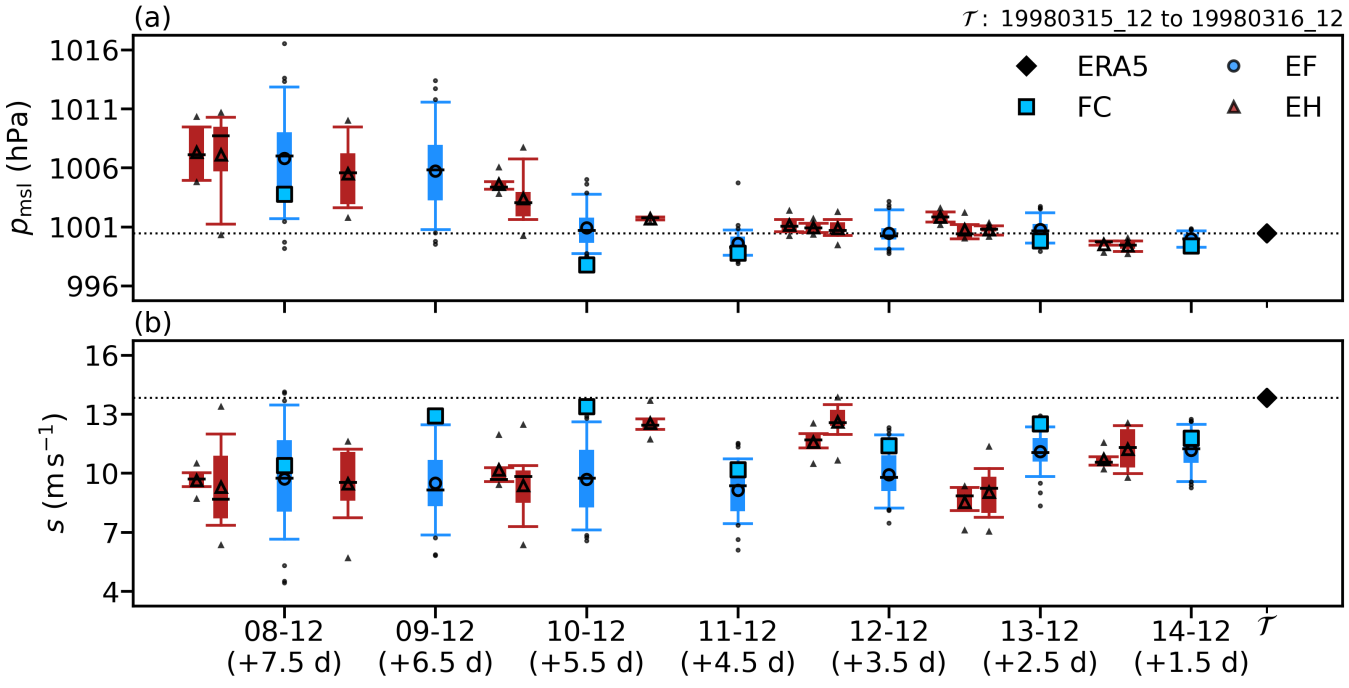


Figure 3.12: Lead time improvements for 1998. Similar to Figure 3.8, but for different evaluation regions (Figures 3.10b and B.9b), and for a different evaluation time window (Figure 3.11).

Lead time dependence

As seen in Figure 3.12a, the different forecasts asymptotically approach the reference value of mean sea level pressure. There is a distinct jump in accuracy at a lead time of 5 d. Most forecasts at longer lead times have errors of 3 hPa or more, while forecasts at shorter lead times are within 2 hPa of the ERA5 value. The ensemble spreads of the longer lead time forecasts are about 5 hPa. But the shorter lead time forecasts show spreads that are smaller by about 50 %. There is no clear best forecast type. The deterministic forecasts are not always better than the ensembles, and the ensemble forecasts are comparable to the ensemble hindcasts.

The situation is less straightforward for wind speed (Figure 3.12b). In general, there is no clear asymptotic approach to the ERA5 reference value with smaller lead time. But there is also a difference between the long lead time forecasts and the short lead time forecasts. Among the long lead time forecasts, the deterministic forecast is more accurate than the ensemble forecast and the ensemble hindcast. In contrast, the short lead time forecasts are more variable. First the hindcast is most accurate, then the hindcast is least accurate, and eventually all forecasts are comparable. The deterministic forecast generally predicts higher wind speeds than the ensemble forecast, being mostly in the upper 25 % of the ensemble forecast values. But, as for pressure, there is no clear best forecast, and the strong winds near Jerusalem appear to be hard to predict.

3.3 Algiers November 2001

Heavy rainfall flooded the city of Algiers (Algeria, 3.09 °E, 36.73 °N) between 10. and 12. November 2001. Tripoli et al. (2005) report that the associated Mediterranean cyclone brought winds of about 120 km h^{-1} and accumulated rainfall of about 285 mm. With more than 700 deaths, this was the worst flood to ever hit Algiers. The cyclone was among the most extreme ones between 1992 and 2021 (Figure B.16). During its rather long lifetime of 54 h (top 7.7 %, Figure B.16a), it reached a minimum central pressure of 989 hPa (top 1.1 %, Figure B.16b). It also ranked high in both the mean precipitation rate (1.19 mm h^{-1},top0.7 %,Figure B.16c) and the mean accumulated precipitation (64.2 mm, top 0.5 %, Figure B.16d).

The synoptic situation was very similar to the previous two cases (Figure 3.13). Again, a large anticyclone ($p_{\mathrm{max}} > 1030$ hPa) led to the formation of an extended and deep PV streamer, reaching PV values above 8 PVU on the 320 K isentrope (Figures 3.13a and 3.13c).

There was a lot of moist air coming from the western Mediterranean ocean towards northern Africa. Initially, the circulation was driven by the large anticyclone over the North Atlantic, bringing air from southern France towards Algeria (not shown). But as the cyclone formed in northern Africa, it brought even more moist air from the central Mediterranean ocean (Figure 3.13a). Since the cyclone intensified rather quickly (drop of about 8 hPa in 0.5 d, Figure 3.13b), the air was driven with strong cyclonic winds, causing a lot of moisture transport. At the surface front near the Atlas mountains (bottom of Figure 3.13c), the moist air (RH > 80 %) was then lifted with relatively strong velocities up to -1.8 Pa s^{-1} (Figure 3.13d). This led to the intense rainfall over Algiers.

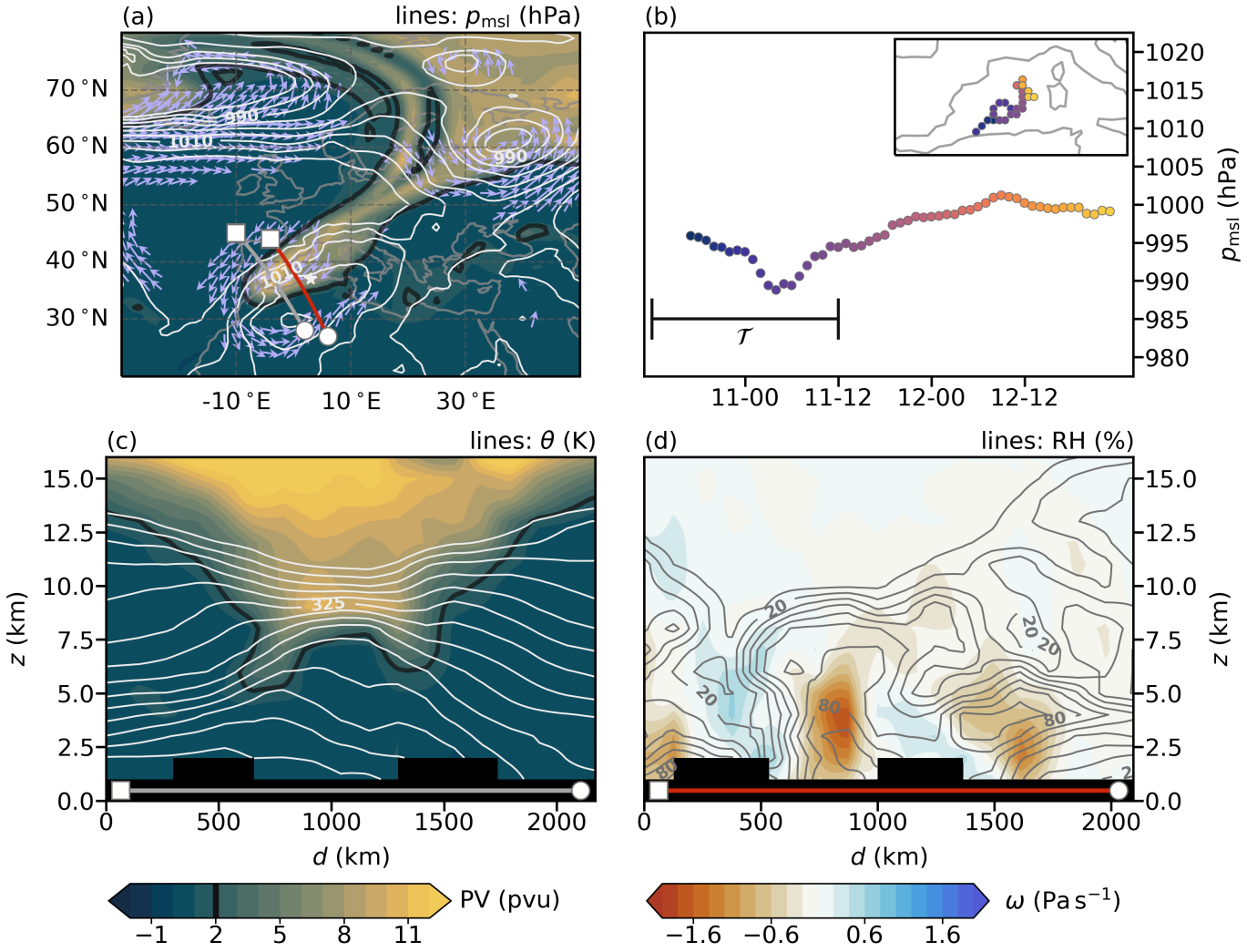


Figure 3.13: Synoptic situation on 10. November 2001, at 00 UTC. Similar to Figure 3.9, but for a different reference time, different vertical cross sections, and a different cyclone track.

Spatial patterns

The long lead time forecasts (8 d to 5.5 d before the mature stage) show similar problems as with the previous two cases. Looking at Figure 3.14, it is clear that they predict the cyclone with a north-eastward shift, and that the exact position of the cyclone is very uncertain. The forecasted cyclones typically form near the northern part of Italy, and then move towards the central Mediterranean ocean where they remain rather stationary until their decay (not shown). While some members of the ensemble forecasts only show small shifts (especially during the later phase of the cyclone, not shown), most members predict the cyclone somewhere to the east and Figures 3.14c and 3.14d show that there is a lot of variation in the cyclone position. The situation is similar for the ensemble hindcasts. However, some members of the hindcasts actually get the location of cyclogenesis correct, as seen by the small contours in Figures 3.14e and 3.14g. But these cyclones typically decay too quickly (not shown), and most members still show the eastwards shift in cyclone position (Figures 3.14e to 3.14h, and 3.14j). Over the whole cyclone track, the more modern hindcasts (newer model versions) tend to show smaller variations in cyclone position than the less modern hindcasts (older model versions), but the improvements are rather small (not shown).

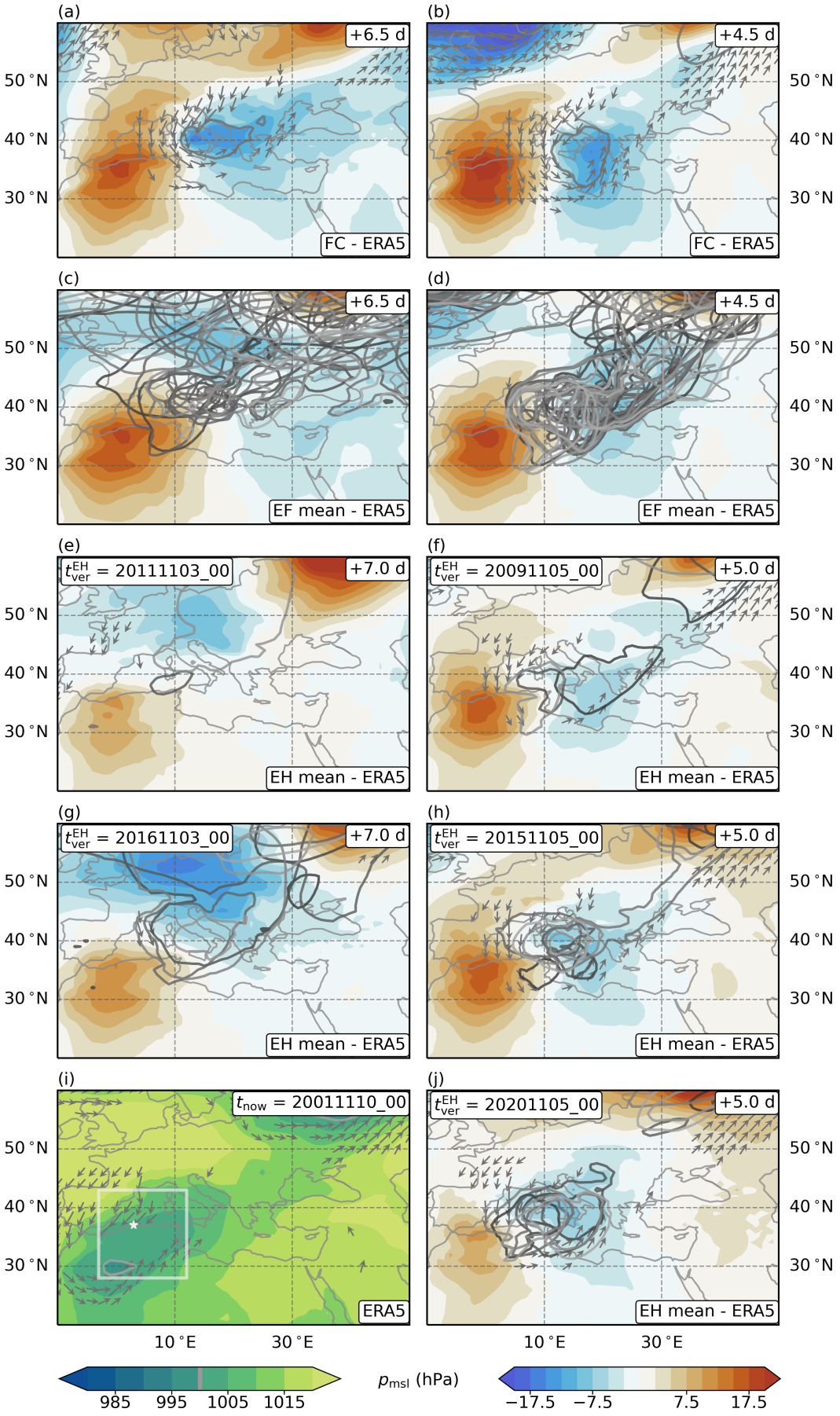


Figure 3.14: Long lead time forecasts of pressure for 2001, 1 d before the (middle of the) most intense cyclone stage. Similar to Figure 3.10, but for a different reference time.

Wind speed is also not well predicted by the long lead time forecasts (Figure B.17). The ensemble forecasts show highly variable locations of strong winds, with all members together covering the whole Mediterranean ocean (Figures B.17c and B.17d). The situation is unpredictable and chaotic. The ensemble hindcasts show less uncertainty in the position of strong winds, with all members covering only the central to western Mediterranean ocean (Figures B.17e to B.17h, and B.17j).

As seen in Figure 3.15, precipitation is mostly predicted over the central Mediterranean ocean rather than over Algeria. But some ensemble hindcasts also show precipitation at the correct location (Figures 3.15e to 3.15g, and 3.15j). Nevertheless, the ensemble variations are similar to wind speed, with the ensemble forecasts covering the entire Mediterranean ocean (Figures 3.15c and 3.15d), while the ensemble hindcasts only cover the central to western part (Figures 3.15e to 3.15h, and 3.15j).

The middle lead time forecasts (5 d to 3.5 d before the mature stage) predict pressure much more accurately (Figure B.18). Both the deterministic forecasts (Figures B.18a and B.18b) and the ensemble hindcasts (Figures B.18e, B.18f, and B.18h) now show cyclogenesis in northern Africa, as in the ERA5 reference data. The cyclones tend to move along the correct path towards the northern coast of Africa (not shown), and the eastwards shift is generally reduced. However, the ensemble forecast is less accurate (Figures B.18c and B.18d). Some members also get the location of cyclogenesis right, but others still develop with a northeastward shift (not shown). The eastward shift is larger compared to the other forecasts, but the uncertainty is much smaller than for the long lead time forecasts.

The improvements are less noticeable for wind speed (Figure B.19). The main benefit of the middle lead time forecasts is a smaller eastward shift. But this does not imply that winds are generally more accurate. Sometimes they are underestimated, and other times they are overestimated. There are still positional uncertainties, more so for the ensemble forecasts (Figures B.19c and B.19d) than the ensemble hindcasts (Figures B.19e, B.19f, and B.19h). However, all ensemble members together now cover a region that more closely resembles the region of strong cyclonic winds in ERA5. The more modern ensemble hindcasts tend to predict this region most accurately (Figure B.19h).

Similarly, the eastward shift is also reduced for precipitation (Figure B.20). Most forecasts now correctly predict precipitation near Algeria, and some even overestimate the precipitation of the ERA5 data. But the ensemble forecasts still tend to underestimate precipitation in northern Africa. This is related to the higher degree of uncertainty between the different members of the ensemble forecast.

Overall, the short lead time forecasts (3 d to 1.5 d before the mature stage) predict the cyclone accurately (Figure B.21). But they tend to predict central pressures that are (slightly) too low. In terms of wind speed, the short lead time forecasts (Figure B.22) are not much better than the middle lead time forecasts. The uncertainty in the position of strong winds is reduced, but the forecasts still show similar error patterns as the middle lead time forecasts, just with a smaller amplitude. Still, the winds near Algiers are underestimate by about 40 $\mathrm{km\,h^{-1}}$. The location of precipitation is also less uncertain, and the forecasts only show small shifts (Figure B.23). But the predicted cells are usually too small, resulting in alternating patterns of underestimation and overestimation within the large bands of ERA5 precipitation.

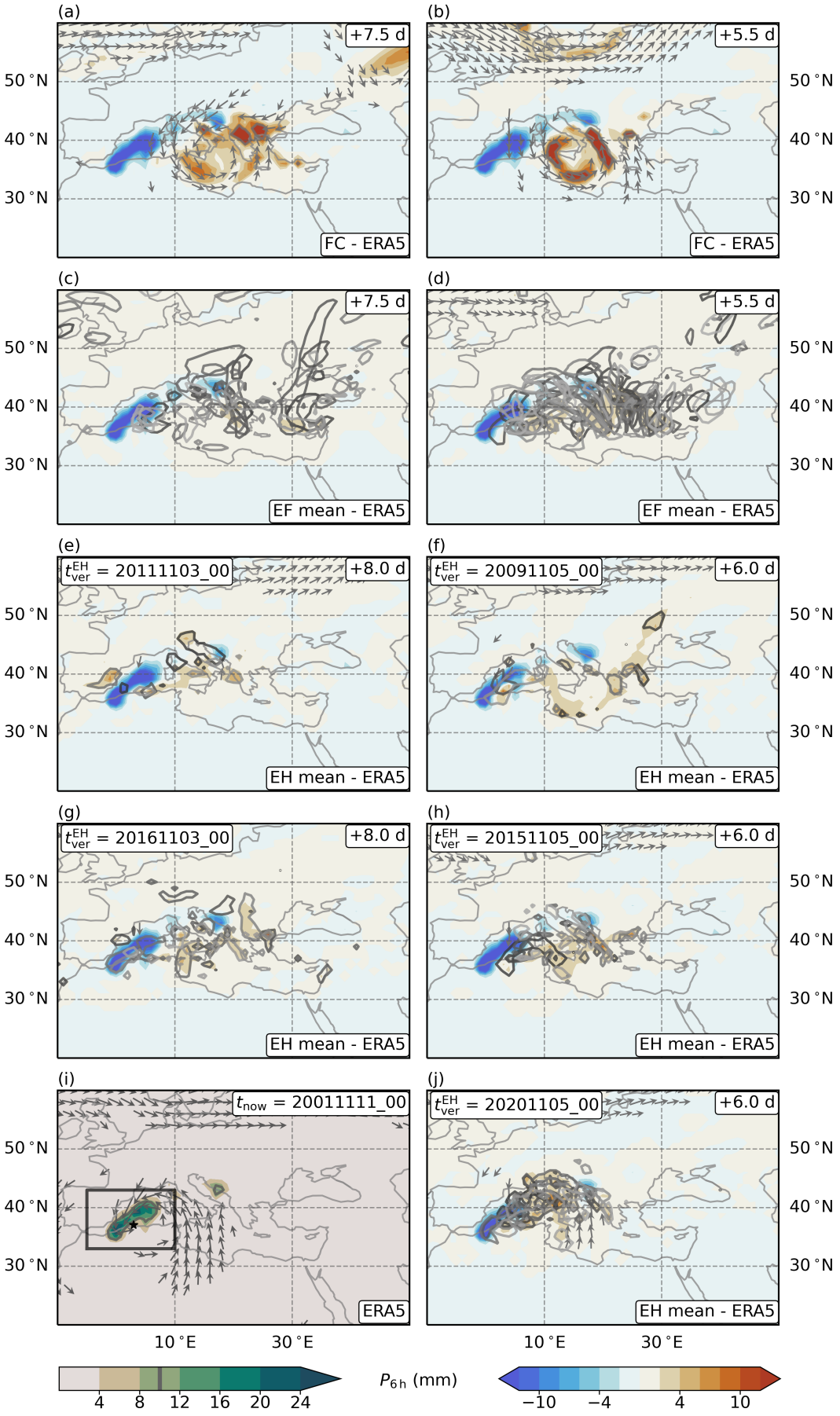


Figure 3.15: Long lead time forecasts of precipitation for 2001, in the middle of the most intense cyclone stage. Similar to Figure 3.14, but for a different reference time, and for 6 h accumulated precipitation.

Temporal evolution

The long lead time forecasts are generally inaccurate, as shown in Figure 3.16. The only exception is the most modern ensemble hindcast, which is considerably more accurate than all the other forecasts, especially for mean sea level pressure (Figure 3.16a). In fact, this hindcast is comparable to the middle ensemble hindcasts that are initialised 1 d later.

One problem with the long lead time forecasts is the time offset of the peaks. All forecasts predict the pressure minimum too early (by at least 0.5 d, Figure 3.16a). For the deterministic forecast and the ensemble forecast, the pressure minimum occurs together with the peaks in wind speed (Figure 3.16b) and precipitation (Figure 3.16c), meaning that these peaks also occur too early. However, this is not the case for the ensemble hindcasts. In terms of wind speed, they predict mostly constant values that start to decay later than in the other forecasts (Figure 3.16b). The situation is similar for precipitation, where the decrease also occurs later than in the other forecasts (Figure 3.16c). Nevertheless, most hindcasts still predict these decays too early, leading to underestimations during the evaluation time window (emphasised with the grey vertical lines).

Although all forecasts correctly predict the amplitude of the pressure time series (difference between maximum value and minimum value), they show pressure values that are considerably higher than in the ERA5 reference data (Figure 3.16a). Except for the most modern ensemble hindcast, the ERA5 values are typically outside the middle 90 % ranges. This is a consequence of the early peaks, while the dissipation rates are similar to the ERA5 reference data.

The situation is different for the other two fields (Figures 3.16b and 3.16c). While the deterministic forecast correctly predicts the wind speed amplitude, the other forecasts underestimate the amplitude, with the ensemble forecast and the ensemble hindcast showing about 75 % and 50 % of the true value (Figure 3.16b). In contrast, all forecasts underestimate the amplitude of precipitation (Figure 3.16c). The deterministic forecast and the ensemble forecasts predict the amplitude at about 20 % of the true value. The ensemble hindcasts are more accurate with amplitudes of about 50 %. Overall, the true values are considerably underestimated during and after the intense phase of the cyclone. The ERA5 values then typically lie outside the middle 90 % of the ensemble values. But again, the more modern hindcasts actually contain the ERA5 values within their middle 90 % ranges.

The middle lead time forecasts show noticeable improvements (Figure 3.17). The deterministic forecast is much better, now accurately predicting the evolution of both pressure (Figure 3.17a) and wind speed (Figure 3.17b). It also improves for precipitation, but it misses the peak before the evaluation time window, while actually overestimating the rainfall during and after the intense phase of the cyclone (amplitude about 160 %, Figure 3.17c).

The ensemble forecast and the ensemble hindcasts still predict the pressure minimum too early, but they also show generally lower pressure values (Figure 3.17a). The ERA5 values are now mostly within the middle 90 % range of the ensemble forecast. However, this is not true for the ensemble hindcasts, which show similar means but smaller spreads.

In terms of wind speed and precipitation (Figures 3.16b and 3.16c), the more modern ensemble hindcasts now correctly predict the timings and amplitudes of the peaks. But they also miss the early peak in precipitation. For the ensemble forecast, the timings are also correct, while the amplitudes are still too small (about 70 % for both fields). The ERA5 values of wind speed and precipitation are mostly within the middle 90 % of the ensemble values, sometimes even within the middle 50 %.

The issues from above are no longer present for the short lead time forecasts (Figure B.24). All forecasts accurately predict the evolution of pressure, wind speed, and precipitation. The only notable feature is that, compared to pressure, the ensemble spreads are larger for wind speed and precipitation. This is especially clear during the intense phase of the cyclone. The ensemble forecast generally shows larger spreads than the ensemble hindcasts.

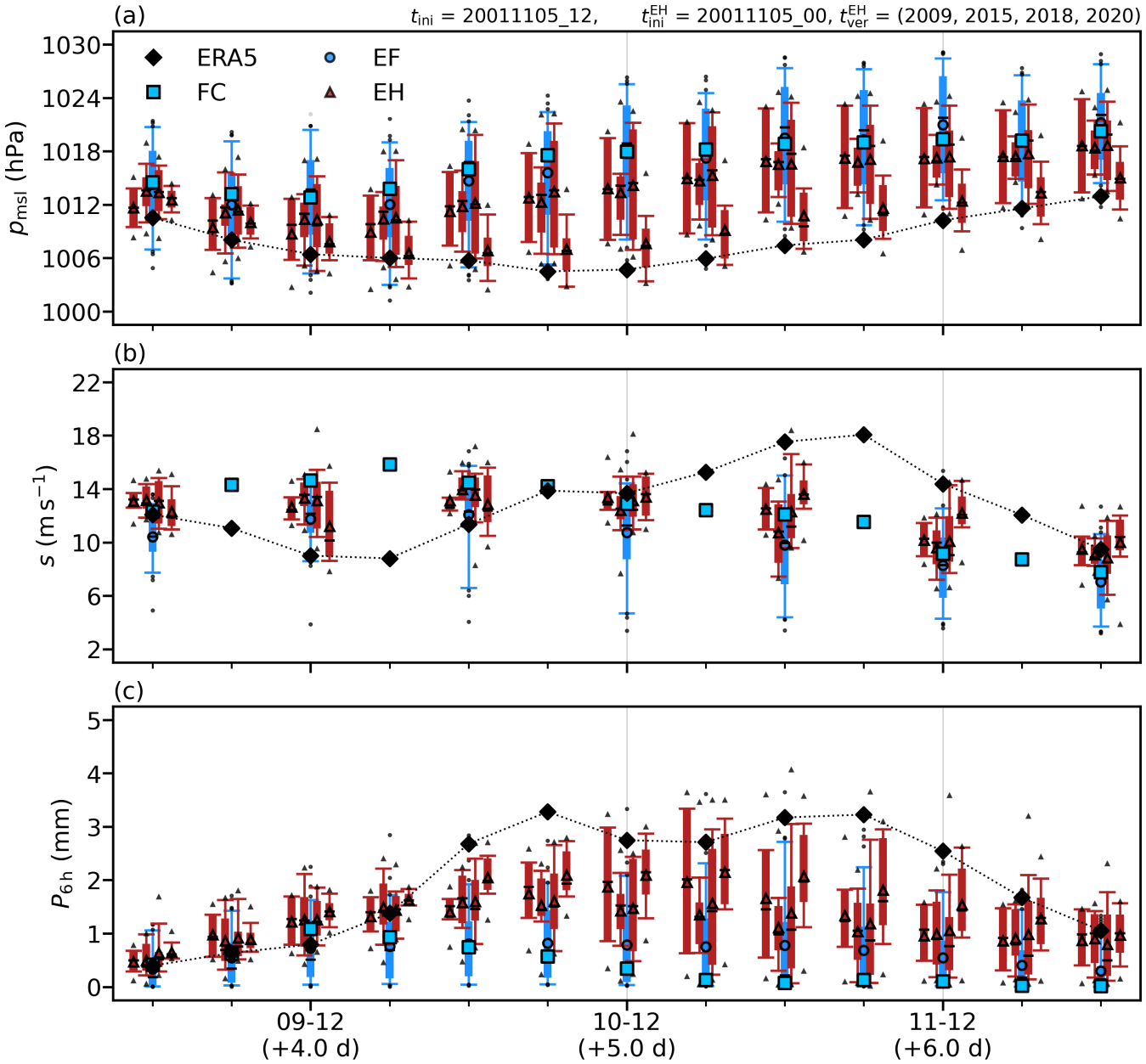


Figure 3.16: Time series of forecasts at a long lead time for 2001. Similar to Figure 3.11, but for a different reference time, and different evaluation regions shown in Figures 3.14i, B.17i, and 3.15i. The forecasts were initialised on 5. November 2001, the ensemble hindcasts (EH) at 00 UTC, and both the deterministic forecast (FC) and the ensemble forecast (EF) at 12 UTC.

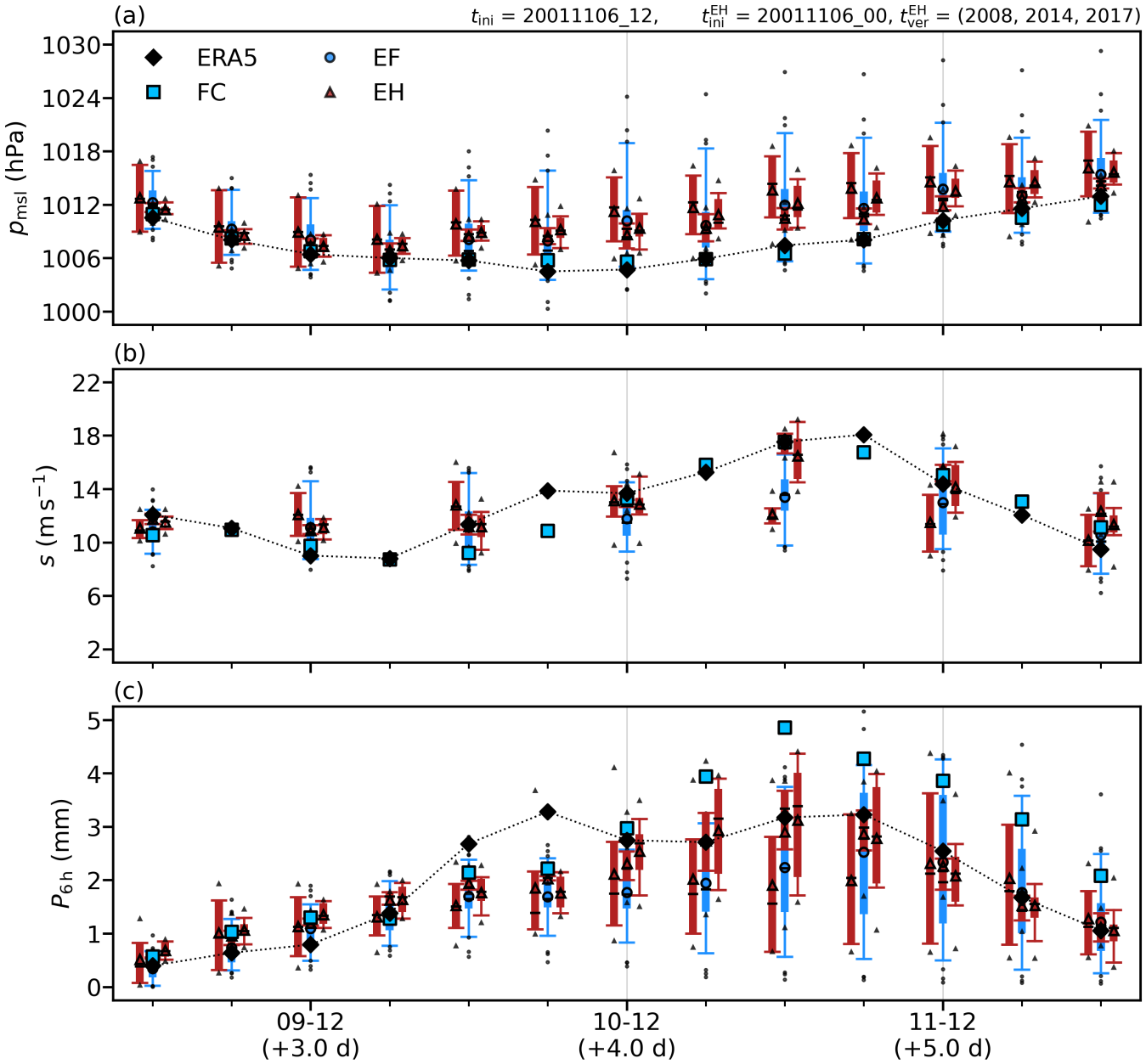


Figure 3.17: Time series of forecasts at a middle lead time for 2001. Similar to Figure 3.16, but for forecasts at a middle lead time. The forecasts were initialised on 6. November 2001, the ensemble hindcasts (EH) at 00 UTC, and the deterministic forecast (FC) at 12 UTC.

Lead time dependence

All available forecasts are compared in Figure 3.18. The deterministic forecast and the ensemble forecast (blue colours) show a clear forecast jump at a lead time of 5 d (similar to section 3.2, Jerusalem 1998). For longer lead times, the forecasts are much less accurate and they are mostly independent of lead time. But for shorter lead times, the forecasts are much more accurate and they start to converge towards the ERA5 reference value. For pressure and wind speed (Figures 3.18a and 3.18c), the forecast jump also shows a clear reduction in the spread of the ensemble forecasts.

The situation is somewhat different for the ensemble hindcasts (red colours). They are generally more accurate than the other forecasts at longer lead times, while being comparable at shorter lead times. In addition, they show spreads that are comparable to those of the ensemble forecasts (or even smaller). This makes the forecast jump less apparent for the hindcasts. Still, the lead time of 5 d marks a clear reduction of ensemble spread, along with a noticeable improvement in accuracy.

For this event, there are many ensemble hindcasts available. Each group of red ensemble hindcasts (between two blue ensemble forecasts) is sorted according to the corresponding model versions. The less modern ones are on the left of a group, while the more modern ones are on the right. From Figure 3.18, it cannot be stated that the most modern hindcast is necessarily the most accurate one. In fact, it is sometimes even the least modern hindcast that performs the best. Nevertheless, the ensemble hindcasts are clearly better than the ensemble forecasts from the past (similar to section 3.1, Brig 1993).

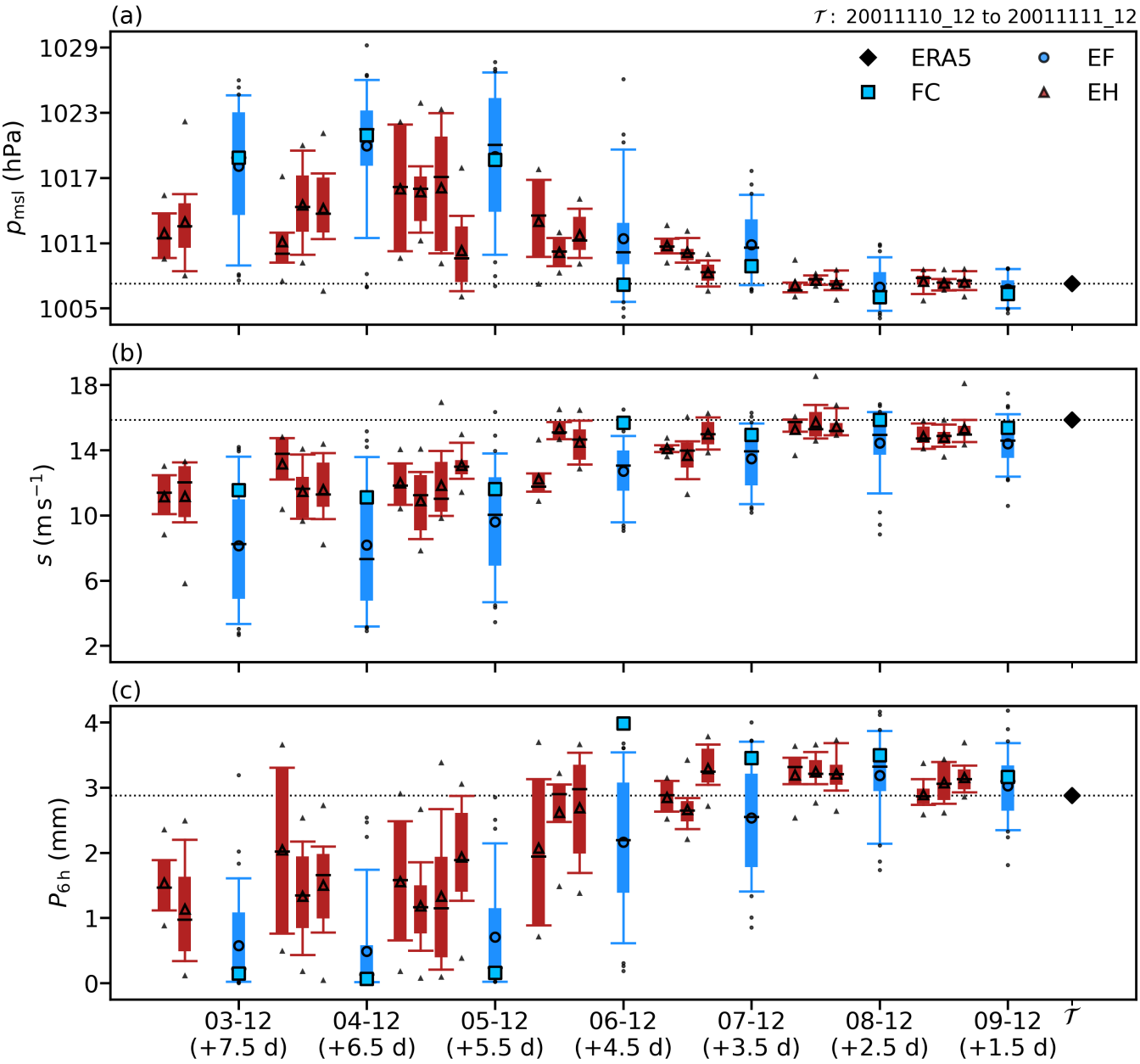


Figure 3.18: Lead time improvements for 2001. Similar to Figure 3.12, but for different evaluation regions (Figures 3.14i, B.17i, and 3.15i), and for a different evaluation time window (Figures 3.16 and 3.17).

3.4 Apulia September 2006

On 26. September 2006, strong winds and heavy rainfall hit the region of Apulia (Italy), first in the south (17 °E, 40.5 °N), and then in the north (15.5 °E, 41.5 °N). As analysed by Moscatello et al. (2008), the cyclone brought winds up to 140 km h^{-1} and reached precipitation rates up to 120 mm h^{-1}. It had a very long lifetime of 108 h (top 0.6 % between 1992 and 2021, Figure B.25a), a large mean precipitation rate of 1.03 mm h^{-1} (top 1.3 %, , Figure B.25c), and an extremely high mean accumulated precipitation of 111.1 mm (top 0.1 %, Figure B.25d). Its minimum central pressure of 997 hPa was also rather low (top 6.7 %, Figure B.25b).

The synoptic situation was somewhat different from the previous cases. There was no large anticyclone over the northern Atlantic, that caused the formation of an extended PV streamer (not shown). Instead, the days before the event already showed a PV streamer extending from the western UK down to Spain (still seen in Figure 3.19a). As the streamer started to weaken, a high pressure system approached Spain from the subtropical Atlantic (also seen in Figure 3.19a). The associated winds drove the tip of the decaying PV streamer further south. During this time, a Mediterranean cyclone formed at the surface and intensified while the streamer developed into the PV cutoff shown in Figure 3.19a. The cutoff was rather deep (6 km) and intense (PV up to 10 PVU), comparable to the other cases (Figure 3.19c). Similar to sections 3.2 (Jerusalem 1998) and 3.3 (Algiers 2001), the cyclone intensified rather quickly with a rate of about 8 hPa in 0.5 d (Figure 3.19b).

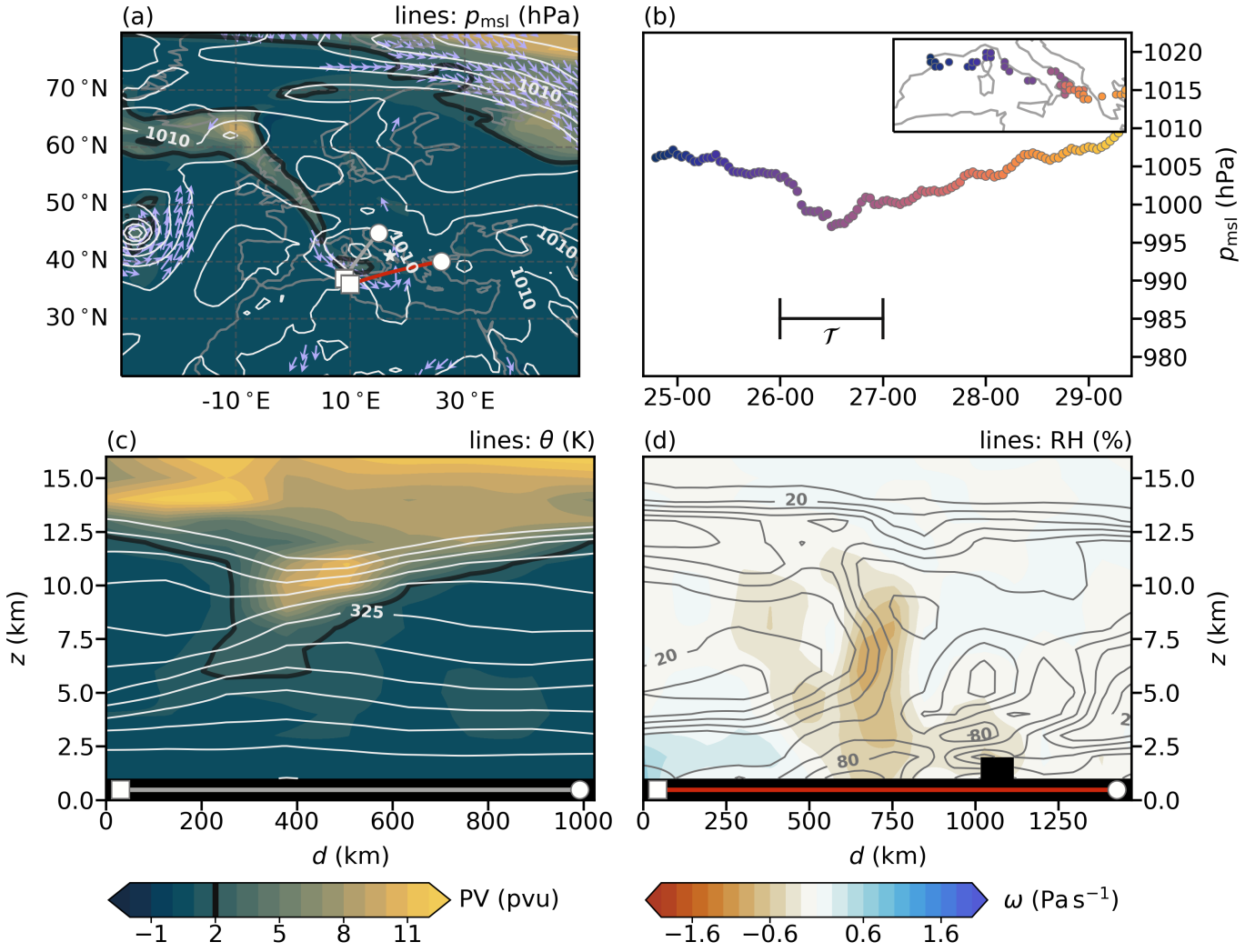


Figure 3.19: Synoptic situation on 26. September 2006, at 00 UTC. Similar to Figure 3.13, but for a different reference time, different vertical cross sections, and a different cyclone track.

About 12 h before the cyclone formed, a smaller cyclone developed over northern Africa (not shown). It brought strong winds that advected a lot of warm air over the central Mediterranean ocean. This formed a surface front, which was later picked up by the main cyclone to bring warm and moist air towards southeastern Italy (not shown). As shown in Figure 3.19d, the surface relative humidity was then mostly above 80 %, and the vertical velocities were around $-0.8\,\mathrm{Pa\,s^{-1}}$.

Spatial patterns

The long lead time forecasts (7.5 d to 6.5 d before the mature stage) generally predict the cyclones with a (north-) westwards shift (Figure 3.20). Most forecasts show a strong low pressure anomaly over northwestern Europe near the UK. The cyclones then cut off from this low pressure system, and are thus typically shifted towards the (north-) west.

The deterministic forecasts either show a cyclone with a northwards shift (Figure 3.20a), or they underestimate the cyclone strength (Figure 3.20b). Some members of the ensemble forecasts get the time and location of cyclogenesis right, while others predict the cyclone too late and at the wrong location (Figures 3.20c and 3.20d). The different members predict the cyclones at various positions in the northern Mediterranean ocean, indicating positional uncertainties. After cyclogenesis, the forecasted cyclones are not as stationary as the ERA5 cyclone, and they move too quickly towards the east (not shown). This is related to the (false) low pressure anomaly in the north, which leads to stronger eastward winds (similar to section 3.1, Brig 1993).

For the longer lead time ensemble hindcasts (Figures 3.20e, 3.20g, and 3.20i), there are also only a few members that get the timing of cyclogenesis right. Most members either show cyclones that develop too late, or not at all. In contrast, the shorter lead time ensemble hindcasts (Figures 3.20f and 3.20h) show more members with a cyclone, which leads to smaller errors in the mean. In general, the hindcasts tend to show smaller uncertainties in cyclone position. But the predicted cyclones also move too quickly towards the east (not shown).

Winds are generally underestimated, and instead shifted towards the false (north-) western low pressure anomaly (Figure B.26). No forecasts accurately predict the strong winds over northern Africa, bringing warm air towards southern Italy (not shown). Most forecasts also underestimate the cyclonic winds around Italy. Especially the deterministic forecasts (Figures B.26a and B.26b) and both the longer lead time ensemble forecast (Figure B.26c) and the longer lead time ensemble hindcasts (Figures B.26e, B.26g, and B.26i). The shorter lead time ensemble forecast (Figure B.26d) shows more members with strong winds near Italy, but the location and extent of the strong winds is rather uncertain. The shorter lead time ensemble hindcasts (Figures B.26f and B.26h) also show more members with strong winds. These show less uncertainty, mostly agreeing on the location and extent of the cyclonic winds.

In terms of precipitation, the main problem is again a shift in location (Figure B.27). The deterministic forecasts show shifts either towards the north, the west, or the east (Figures B.27a and B.27b). The ensemble forecasts also tend to predict precipitation cells with a westward shift, although some members also get the position right, or even show eastwards shifts (Figures B.27c and B.27d). The positional shifts are generally smaller for the ensemble hindcasts, with some precipitation too far in the west and some too far in the east (Figures B.27e to B.27i). However, the positional uncertainties are still large, resulting in underestimation of precipitation for the ensemble mean.

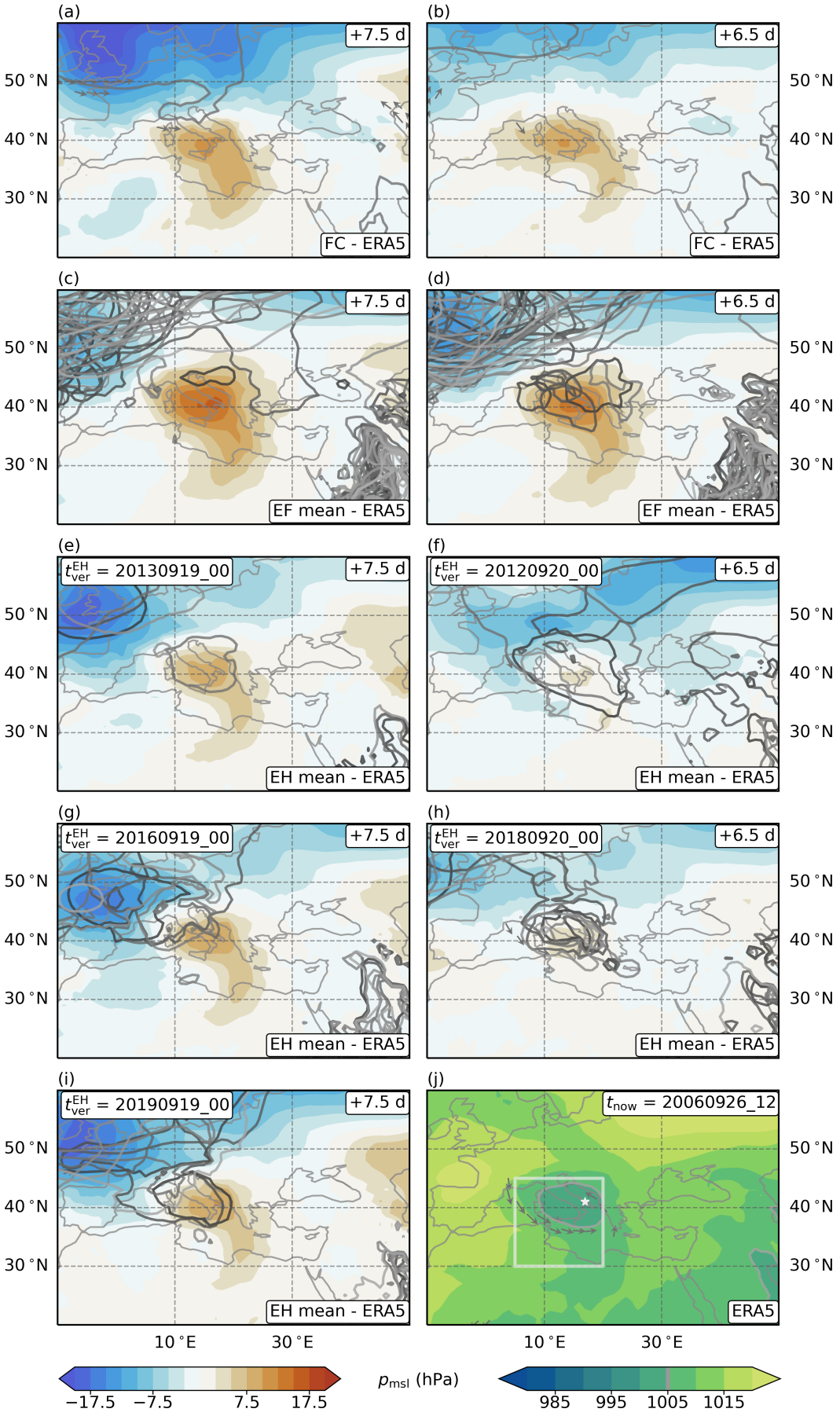


Figure 3.20: Long lead time forecasts of pressure for 2006, in the middle of the most intense cyclone stage. Similar to Figure 3.2, but for a different reference time.

The middle lead time forecasts (5.5 d to 3.5 d before the mature stage) improve on some issues (Figures B.28, B.29, and B.30). Although there is still a low pressure anomaly in the north, it is not as strong and more ensemble members now show cyclones that clearly cut off from this system (Figures B.28c and B.28d). While the deterministic forecasts still mostly underestimate the strength of the cyclone (Figures B.28a and B.28b), the forecasts now generally show smaller shifts. The positional uncertainties are not noticeably smaller for the ensemble forecasts, but most of the ensemble hindcasts are much more accurate with little uncertainty in the position of the cyclone (Figures B.28f, B.28h, and B.28i). Nevertheless, most forecasts still predict cyclones that move too quickly towards the west, and the cyclone position becomes uncertain along the track (not shown).

Similar to section 3.3 (Algiers 2001), wind speed does not improve as much (Figure B.29). The main added value is that more ensemble members now predict strong winds, and that the positional uncertainties are slightly reduced. But the winds are still generally too weak, both in northern Africa, and around the cyclone centre over Italy.

However, precipitation clearly improves (Figure B.30), as the positional shifts are much smaller. The deterministic forecasts no longer show westward shifts, but instead small eastward shifts (Figures B.30a and B.30b). In addition, most ensemble members now predict precipitation near Italy. Overall, there are still uncertainties in the exact location of precipitation, but the shorter lead time and more modern hindcast shows considerably smaller uncertainties (Figure B.30h).

The short lead time forecasts (2.5 d to 1.5 d before the mature stage) improve the a lot in terms of mean sea level pressure (Figure B.31). All forecasts now correctly predict the location of cyclogenesis (not shown) and the stationarity of the cyclone (cyclones still over Italy). The longer lead time forecasts still tend to underestimate the strength of the cyclone (left column in Figure B.31), and predict the decay too early (not shown). But the shorter lead time forecasts do not show these issues, generally predicting the cyclone track accurately and with little uncertainty (right column in Figure B.31). The situation is similar for wind speed, with the longer lead time forecasts still slightly underestimating the cyclonic winds, while the shorter lead time forecasts get the winds right (Figure B.32). The shifts in precipitation are reduced but still too large for the longer lead time forecasts, while the shorter lead time forecasts predict precipitation with higher accuracy (Figure B.33). The deterministic forecasts and the ensemble hindcasts predict precipitation with comparable accuracy. But the ensemble forecasts are clearly less accurate.

Temporal evolution

The time series of the long lead time forecasts are shown in Figure 3.21. The deterministic forecast and the ensemble forecast generally predict pressure values that are too high, with the ERA5 values mostly lying outside the middle 90 % of the ensemble values (Figure 3.21a). Although both forecasts get the amplitude of the pressure values right, they predict a continuous drop in pressure without a dissipation phase. In contrast, the ensemble hindcasts predict the evolution of pressure more accurately. But there are considerable differences between the different model versions. The least modern hindcast overestimates the amplitude (about 130 %) and predicts the pressure minimum too late, while the most modern hindcast underestimates the amplitude (about 60 %) and gets the timing right. But the intermediate hindcast (from 2018) accurately predicts the whole evolution of mean sea level pressure, performing similarly well as the middle ensemble hindcasts (Figure B.34a).

The situation is a bit different for wind speed (Figure 3.21b). The deterministic forecast accurately predicts the time of maximum wind speeds, but underestimates the amplitude (about 60 %). In contrast, the ensemble forecast shows continually increasing winds, while also underestimating the amplitude (about 30 %). As for pressure, the ensemble hindcasts are generally more accurate. They show no timing errors, but also tend to underestimate the amplitude of the wind speeds. While the least modern hindcast (with the low pressure values) predicts an amplitude of about 60 %, the two more modern hindcasts (with more accurate pressure values) show amplitudes of about 40 % compared to the ERA5 data. Overall, the ERA5 values mostly lie outside the middle 90 % of the ensemble values.

For precipitation, there are notable differences between the forecasts (Figure 3.21c). The overall amplitude is underestimated by all forecasts, about 60 % for the deterministic forecast, 10 % for the ensemble forecast, and 80 % for the ensemble hindcasts. While most forecasts miss the first peak in precipitation, they generally get the timing and amplitude of the second peak right. Only the ensemble forecast underestimates both peaks, and instead predicts mostly constant precipitation. Still, the middle 90 % of the ensemble values usually contain the ERA5 reference values. This is also true for the ensemble hindcasts, which are typically at least as good as the other forecasts. The only exception is the least modern hindcast, which largely overestimates precipitation after the second peak. The ERA5 values are then outside the middle 90 % of the ensemble values.

The middle lead time forecasts do not show considerable improvements (Figure B.34). Pressure is still generally overestimated by the deterministic forecast and the ensemble forecast, while the ensemble hindcasts are typically more accurate (Figure B.34a). However, the different hindcasts are now similarly accurate, all underestimating the pressure drop with amplitudes of about 60 %. For wind speed, both operational forecasts show continuously increasing winds, with the deterministic forecast getting the amplitude right, while the ensemble forecast underestimates it (about 50 %, Figure B.34b). The different ensemble hindcasts are comparable to each other, all underestimating the amplitude (about 50 %) but accurately predicting the timing of the peak. In terms of precipitation, the middle lead time forecasts are comparable to the long lead time forecasts (Figure B.34c). The one exception is the deterministic forecast, that now misses both peaks and is even worse than the ensemble forecast.

Most of these problems are resolved for the short lead time forecasts (Figure B.35). All forecasts accurately predict the evolution of both mean sea level pressure and horizontal wind speed. For precipitation, the forecasts tend to slightly overestimate the maximum values, especially the deterministic forecast and the ensemble forecast. But the forecasts are otherwise very accurate and the timing of both peaks is correct.

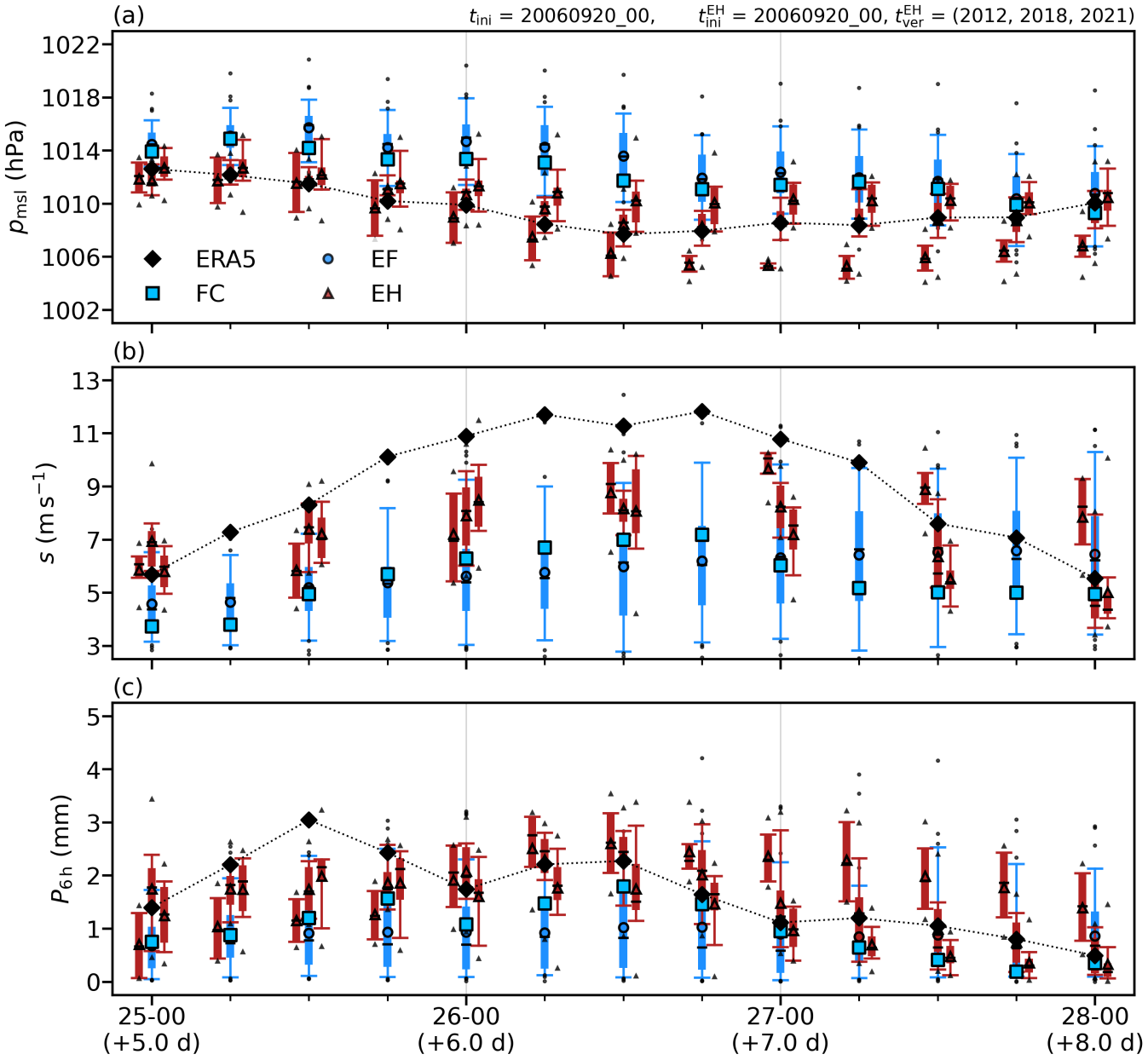


Figure 3.21: Time series of forecasts at a long lead time for 2006. Similar to Figure 3.16, but for a different reference time, and different evaluation regions shown in Figures 3.20j, B.26j, and B.27j. All forecasts were initialised on 20. September 2006, at 00 UTC.

Lead time dependence

As seen in Figure 3.22, there are notable differences among the various forecasts. The deterministic forecast shows an oscillatory dependence on lead time, before finally starting to converge towards the ERA5 reference value (Figures 3.22a to 3.22c). The reason for this is that forecasts initialised at 00 UTC typically show a low pressure anomaly in northwestern Europe (seen in Figures 3.20a and 3.20b), while the 12 UTC forecasts show a low pressure anomaly in northeastern Europe (not shown). A possible explanation for this variation could be the diurnal cycle of convection (see also the large precipitation values for 12 UTC forecasts, Figure 3.22c). The change between the oscillatory phase and the asymptotic phase occurs at a lead time of 4.5 d. This is reminiscent of the forecast jumps seen in sections 3.2 (Jerusalem 1998) and 3.3 (Algiers 2001), which occur at a lead time of 5 d.

In contrast, the ensemble forecast shows mostly monotonic convergence towards the true ERA5 value. This is the case for both pressure and wind speed (Figures 3.22a and 3.22b). For precipitation,

the ensemble forecast also shows some oscillations until a lead time of 4 d (Figure 3.22c). After that, the ensemble forecast behaves asymptotically.

The ensemble hindcasts are usually more accurate than the other forecasts, or comparable to them. Another advantage is that they already show values close to the ERA5 data at long lead times beyond 6 d (Figures 3.22a and 3.22c). This is similar to sections 3.1 (Brig 1993) and 3.3 (Algiers 2001), where the hindcasts are also less sensitive to lead time, while being at least as accurate as the other forecasts. As in section 3.3 (Algiers 2001), there is a clear jump in ensemble spread, occurring here at a lead time of 4 d (instead of 5 d). Another similarity is that there is no apparent relation between hindcast model version and forecast performance. In general, it cannot be said that the more modern hindcasts are also more accurate. Though, as in sections 3.1 (Brig 1993) and 3.3 (Algiers 2001), the ensemble hindcasts are clearly more accurate than the ensemble forecasts from the past.

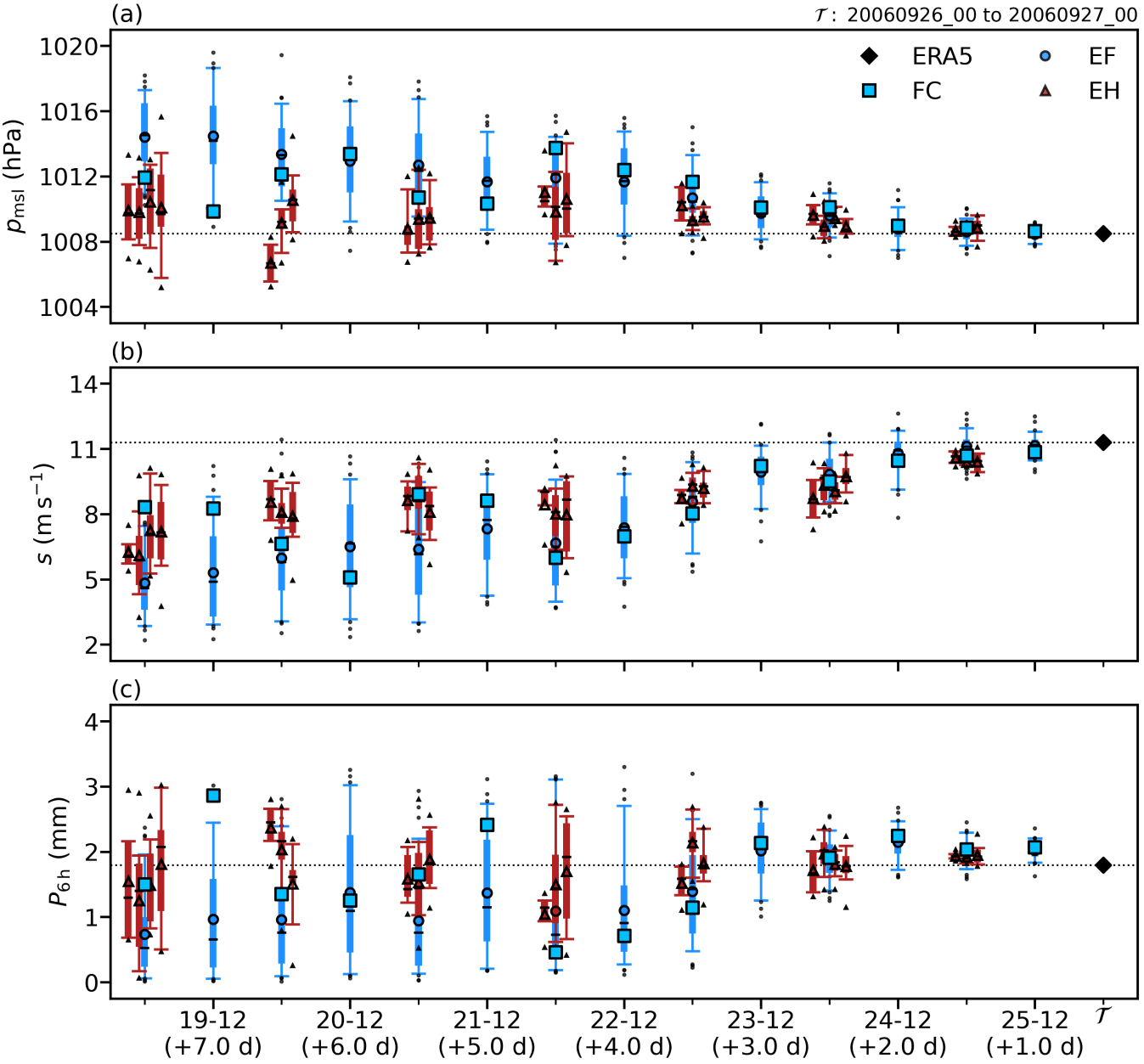


Figure 3.22: Lead time improvements for 2006. Similar to Figure 3.18, but for different evaluation regions (Figures 3.20j, B.26j, and B.27j), and for a different evaluation time window (Figure 3.21). Note that the deterministic forecasts and the ensemble forecasts are always initialised at 00 UTC and 12 UTC, while the ensemble hindcasts are always initialised at 00 UTC.

3.5 Hyères November 2011

Between 6. and 9. November 2011, the region near Hyères (France, 6.17 °E, 43.62 °N) was hit with recurring rainfall and strong winds. Ricchi et al. (2017) report wind speeds over 100 km h^{-1} causing wave heights up to 8 m, and intense precipitation near the coasts of France and Italy. The corresponding cyclone lived for 104 h, ranking in the top 0.8 % of all Mediterranean cyclones between 1992 and 2021 (Figure B.36a). Together with a strong precipitation rate of 0.86 mm h^{-1} (top 2.5 %, Figure B.36c), this led to the extremely high accumulated precipitation of 89.3 mm (top 0.1 %, Figure B.36d). The cyclone also reached a rather low central pressure of 996 hPa (top 5.6 %, Figure B.36b).

The synoptic situation was similar to the first three cases (sections 3.1 to 3.3). There was again a large anticyclone, but this time a bit weaker ($p_{max} > 1020$ hPa) and further south over the subtropical Atlantic (still seen in Figure 3.23a). Together with a cyclone near Iceland, this anticyclone led to strong southward winds. Although an extended PV streamer was already present (form the UK down to Algeria), these winds helped to keep it rather stationary over 2 d. During this time, a cyclone formed at its eastern side over southern France. As the cyclone persisted, the streamer started to develop into a PV cutoff in the following 2 d. After the cutoff formed, the upper-level PV anomaly was co-located with the surface cyclone.

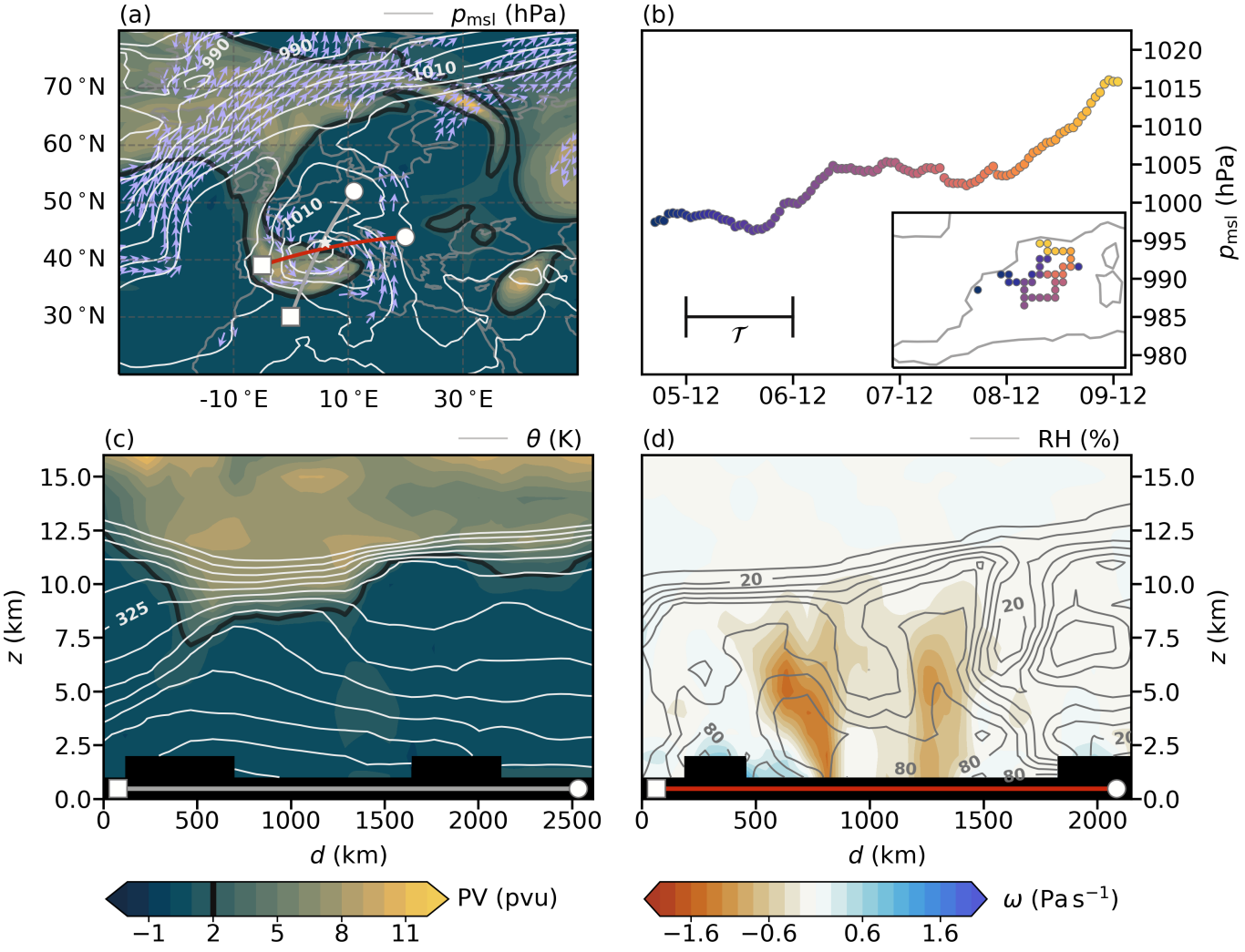


Figure 3.23: Synoptic situation on 6. November 2011, at 00 UTC. Similar to Figure 3.19, but for a different reference time, different vertical cross sections, and a different cyclone track.

As seen in Figure 3.23c, the PV anomaly was a bit weaker compared to the other cases. It only reached down to an altitude of about 7.5 km (instead of 6 km), and its PV values were mostly below 8 PVU (instead of up to 10 PVU). Another difference is that the cyclone did not intensify much after its genesis, seen in Figure 3.23b. Nevertheless, the strong horizontal winds (up to 90 km h^{-1}) brought a lot of moist air (RH $>$ 80 %, Figure 3.23d) towards southern France. The updrafts shown in Figure 3.23d (about -1 Pa s^{-1}) lifted this moist air, leading to heavy precipitation along the coasts of France and Italy.

Spatial patterns

Figure 3.24 shows the long lead time forecasts (5 d to 4 d before the mature stage) for mean sea level pressure. It is clear that most forecasts again predict the cyclone with a shift in position, this time towards the northwest. The cyclones correctly cut off from a low pressure anomaly over western Europe. However, the forecasts tend to predict either this cut off phase or the later track of the cyclone too far north. This is especially noticeable for the longer lead time deterministic forecast (Figure 3.24a) and the ensemble hindcasts (Figures 3.24e to 3.24i). The ensemble forecasts show more members with the correct cyclone position (Figures 3.24c and 3.24d), but most predicted cyclones decay about 0.5 d too early (not shown). This early decay is also an issue for the other forecasts (not shown).

Similar patterns are seen for horizontal wind speed (Figure B.37). Most forecasts underestimate the winds over the central to western Mediterranean ocean. The strong winds are instead predicted with a westward shift. Both the ensemble forecasts and the ensemble hindcasts show large positional uncertainties (Figures B.37c to B.37i). Due to these uncertainties, the ensemble means generally underestimate the winds seen in the ERA5 reference data.

Although precipitation shows smaller shifts, the forecasts still tend to underestimate the heavy rain at the critical locations along the southern coast of France (Figure B.38). This is because the precipitation is not as persistent as in the ERA5 data (not shown). Sometimes precipitation is overestimated, while other times it is underestimated. The persistent rainfall is instead predicted south of the Alps in northern Italy. Overall, there are large positional uncertainties, especially in the west-east direction.

The middle lead time forecasts (3 d to 2 d before the mature stage) are much more accurate than the long lead time forecasts (Figures B.39, B.40, and B.41). The northwestward shift of low pressure is no longer present for most members of the ensemble forecasts (Figures B.39c and B.39d), and most forecasts now correctly predict the decay of the cyclone (not shown). Although there are still some positional uncertainties, they are considerably smaller than for the long lead time forecasts. Wind speed is also much more accurate (Figure B.40). The positional uncertainties are reduced, and the shifts are no longer an issue. The middle lead time forecasts also improve in terms of precipitation (Figure B.41). The first phase of heavy rain is now mostly correct (not shown), while the second phase still shows some underestimations near southern France due to slight northwards shifts. The positional uncertainties are also smaller.

The short lead time forecasts (1 d to 0 d before the mature stage) mainly improve in terms of mean sea level pressure (Figure B.42). The evolution of the cyclones is now predicted with high accuracy, and positional uncertainties are very small. For wind speed and precipitation, the main benefit of the short lead time forecasts is the reduction of positional uncertainties (Figures B.43 and B.44).

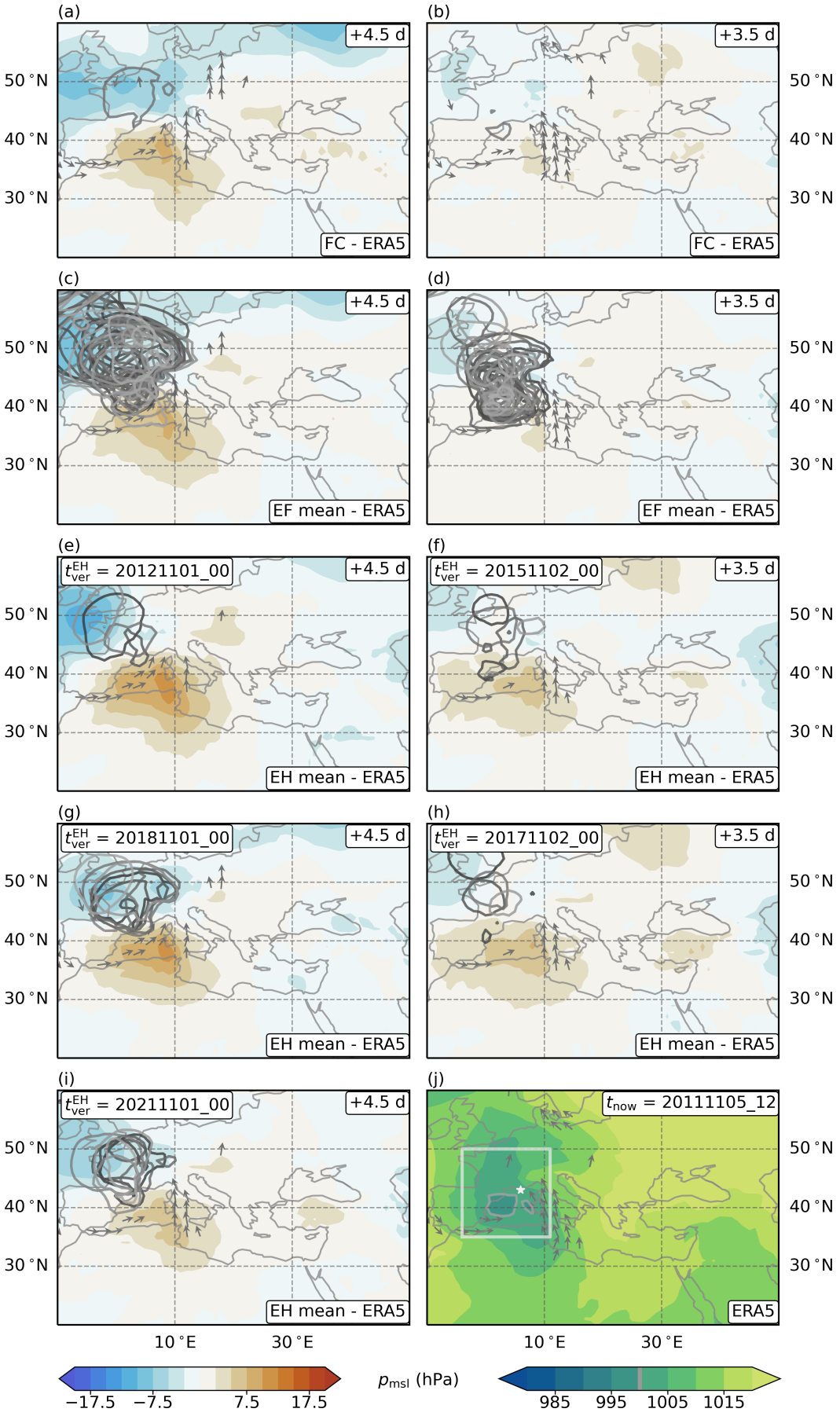


Figure 3.24: Long lead time forecasts of pressure for 2011, at the beginning of the most intense cyclone stage. Similar to Figure 3.2, but for a different reference time.

Temporal evolution

Figure 3.25 shows the time series of the long lead time forecasts. The first half of the pressure series is predicted rather accurately, while the second half is characterised by pressure values that are too high (Figure 3.25a). The forecasted amplitudes are then about 130 % of the true ERA5 amplitude. However, the ensemble forecast still includes the ERA5 values within the middle 90 % of the ensemble values. This is also true for the most modern hindcast. But the two less modern hindcasts are less accurate, with the ERA5 values lying mostly outside their middle 90 % ranges.

Wind speed is generally underestimated, especially during and after the evaluation time window (Figure 3.25b). Nevertheless, most forecasts predict the amplitude rather accurately (at least 90 % of the true value), and the ensemble forecast mostly includes the ERA5 values within its middle 90 % range. However, this is not true for the ensemble hindcasts.

Similar to the other two fields, the first half of the precipitation time series is rather accurate (Figure 3.25c). But during the second half, the forecasts show considerably smaller values than the ERA5 reference data. However, the different forecasts show rather distinct time series. The deterministic forecast predicts two relatively similar peaks in precipitation, with an amplitude of about 60 % relative to the ERA5 amplitude. In contrast, the ensemble forecast shows a single peak (about 0.5 d too early) with an amplitude of about 50 %. This is also the case for the least modern hindcast. The two more modern hindcasts show an amplitude of about 70 %, with the intermediate hindcast predicting the peak about 1 d too early, and the most modern hindcast showing the most accurate values overall (about half the errors of the ensemble forecast). The ERA5 values are mostly contained within the middle 90 % range of both the ensemble forecast and the most modern ensemble hindcast.

The middle lead time forecasts show noticeable improvements (Figure B.45). All forecasts now accurately predict the evolution of pressure, throughout all timesteps (Figure B.45a). The wind speeds are still underestimated during the second half, but the errors are reduced by about 50 % (Figure B.45b). Precipitation is predicted much more accurately, with all forecasts getting the time and amplitude of the peak right (Figure B.45c). The ERA5 values are now mostly within the middle 50 % of the ensemble values.

The short lead time forecasts predict the evolution of pressure even more accurately and with very small spreads (Figure B.46a). However, the improvements for wind speed and precipitation are comparatively small (Figures B.46b and B.46c). The main difference to the middle lead time forecasts is the reduction of the ensemble spreads.

Lead time dependence

All available forecasts are compared in Figure 3.26. Similar to section 3.1 (Brig 1993), the deterministic forecasts predict pressure close to the ERA5 reference value, mostly independent of lead time (Figure 3.26a). But different to the Brig event (section 3.1), this is also the case for the ensemble forecasts. The main changes with smaller lead times are the reduced ensemble spreads. At larger lead times, the ensemble hindcasts are typically less accurate for pressure than the operational forecasts. But this is no longer true for smaller lead times, where the hindcasts are comparable to the other forecasts, and show smaller ensemble spreads.

For wind speed, all forecasts asymptotically approach the ERA5 reference value (Figure 3.26b). However, even at the smallest lead times, all forecasts tend to underestimate the wind speeds. Similar to pressure, the large lead times show the ensemble hindcasts being mostly less accurate than the other forecasts, while all forecasts are comparable for smaller lead times. Note, however, that the ensemble forecasts are always more accurate than the deterministic forecasts.

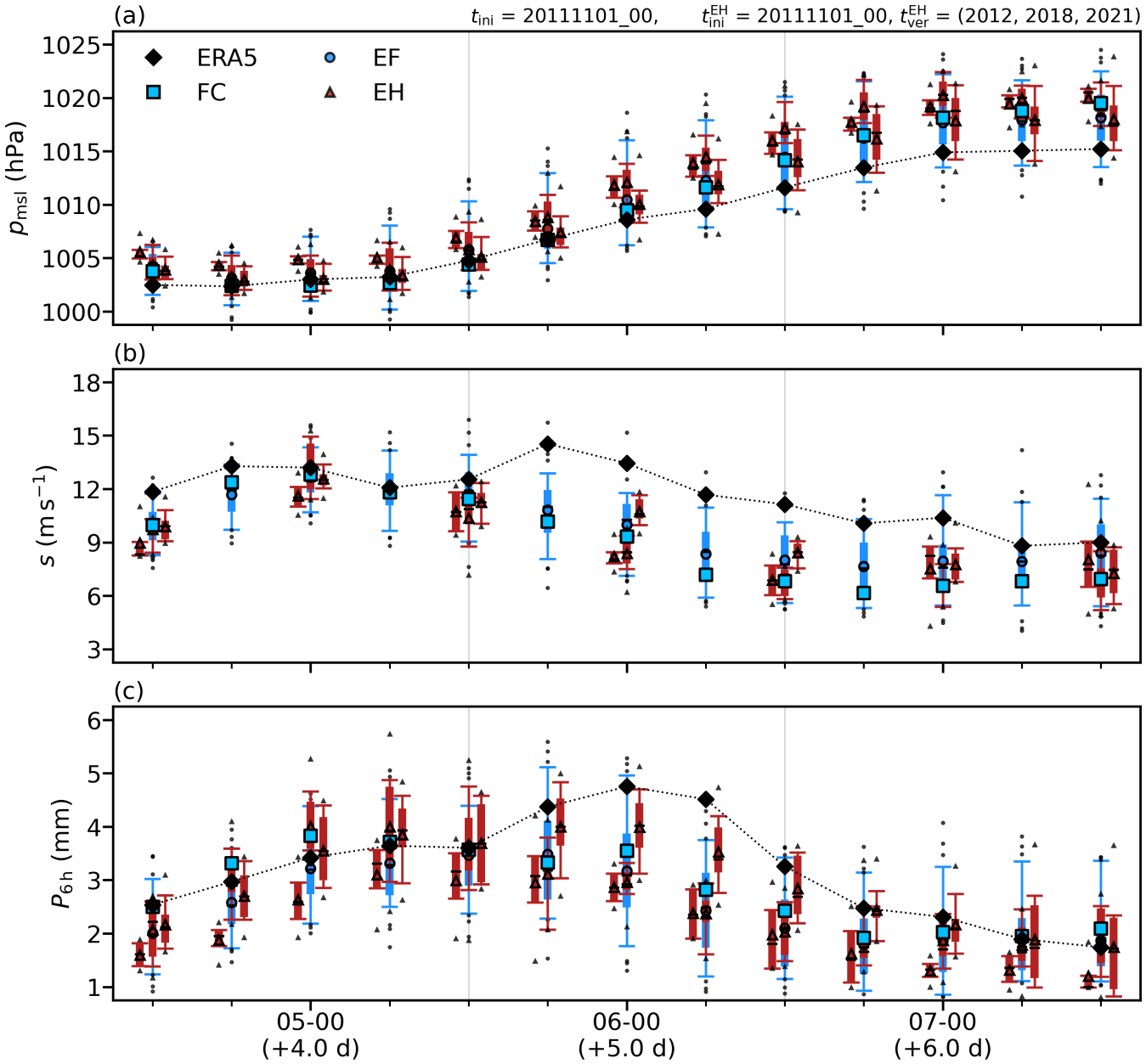


Figure 3.25: Time series of forecasts at a long lead time for 2011. Similar to Figure 3.21, but for a different reference time, and different evaluation regions shown in Figures 3.24j, B.37j, and B.38j. All forecasts were initialised on 1. November 2011, at 00 UTC.

In terms of precipitation, the forecasts already converge to the ERA5 reference value at large lead times, and then only marginally improve with smaller lead times (Figure 3.26c). The deterministic forecasts and the ensemble forecasts are comparably accurate, while the hindcasts are usually more accurate. At large lead times, the most modern hindcasts are also the most accurate ones. But for shorter lead times, all hindcasts are similarly accurate.

In contrast to sections 3.3 (Algiers 2001) and 3.4 (Apulia 2006), this event indicates that the most modern hindcasts are also the most accurate ones. This can be said for all three fields, at least for large lead times, where the different hindcasts show distinct forecast performances. Another feature of this event is that the ensemble hindcasts are not clearly better than the ensemble forecasts from the past. This is different from three of the other cases in sections 3.1 (Brig 1993), 3.3 (Algiers 2001), and 3.4 (Apulia 2006). This may be the result of model improvements, particularly the resolution upgrade on 26. January 2010 (Table 2.3), making the operational ensemble forecasts more comparable to the ensemble hindcasts.

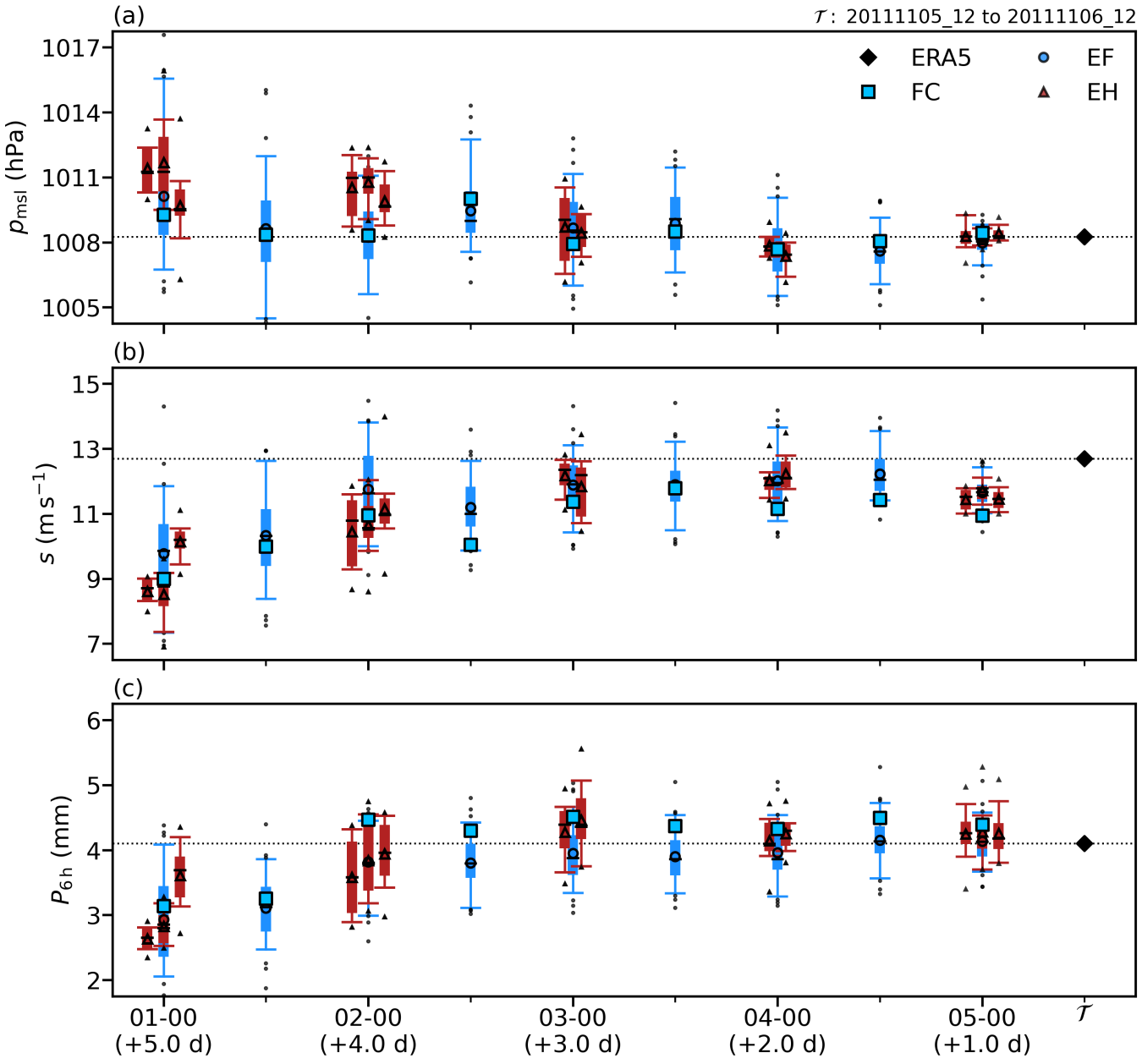


Figure 3.26: Lead time improvements for 2011. Similar to Figure 3.22, but for different evaluation regions (Figures 3.24j, B.37j, and B.38j), and for a different evaluation time window (Figure 3.25).

Chapter 4

Discussion

In chapter 3, we investigated the forecast performances for five different extreme weather events in the Mediterranean basin. In this chapter, we first compare all five cases and focus on both the synoptic situations and the forecast performances. Then, we discuss the general limitations of our methodology.

4.1 Comparison of all five cases

Synoptic situation

The typical precursor of the events is a large anticyclone ($p_{max} > 1020\,\mathrm{hPa}$) over the North Atlantic (4 cases). The associated southward winds then lead to the formation of a far south-reaching PV streamer. This streamer reaches down to an altitude of about 6 km (4 cases), and reaches PV values up to 10 PVU (3 cases). The Mediterranean cyclones then form on the eastern side of the upper-level PV anomaly, intensifying with rates up to $15\,\mathrm{hPa\,d^{-1}}$ (3 cases). At their most intense stage, the cyclones bring strong winds over $100\,\mathrm{km\,h^{-1}}$ (3 cases), which lead to intense moisture transport. At the surface, the relative humidity is mostly above 80 % (4 cases), and the updrafts are at least $-1\,\mathrm{Pa\,s^{-1}}$ (4 cases) leading to heavy precipitation. In some cases, the surrounding orography is crucial for the strong updrafts. As mentioned in section 1.1.2, this evolution is very typical for Mediterranean cyclones, and our results agree with past studies (Flocas, 2000; Trigo et al., 2002; Nicolaides et al., 2006; Fita et al., 2007).

Performance of long lead time forecasts

The most prevalent issue with the long lead time forecasts is a clear shift in the location of the Mediterranean cyclones. The forecasts typically predict the cyclones too far in the north (4 cases). On top of that, there is either a westward shift (2 cases) or an eastward shift (2 cases). The reason for this error is usually an overly strong low pressure system in northern Europe (3 cases). The cyclones either form near this system or cut off from it. Given that pressure gradients drive winds, and winds drive moisture transport, the shifts of the cyclones also lead to shifts of strong winds and heavy precipitation (5 cases).

The shifts are seen across all forecasts (deterministic forecasts, ensemble forecasts, and ensemble hindcasts), and there is no clear indication for a best forecast type. Therefore, the problem cannot be linked to a particular weakness in the ECMWF models. However, since the cyclones typically form near an upper-level trough, the cyclone shifts may be the result of a shifted PV streamer. Another indication for this, it that the shifts typically occur along the direction of the PV streamer, implying that the intensity of the streamer impacts the position of the surface cyclones. This would suggest that

the forecast errors originate from dynamical processes before the formation of the upper PV anomaly. As analysed by Portmann et al. (2020), this phase of Rossby wave breaking is particularly sensitive to small scale perturbations and may lead to considerable errors in the location of the subsequent PV streamer. This suggests that the small scale processes involved in Rossby wave breaking (like condensation, radiation, or turbulence) may be a source of positional shifts.

Related to the positional shifts, the long lead time forecasts also show large positional uncertainties. This is seen for pressure, wind speed, and precipitation. The uncertainties are either already present at the time of cyclogenesis where different ensemble members predict cyclones at different locations, or they develop along the track of the cyclone as the different ensemble members diverge from each other. The ensemble hindcasts typically show smaller uncertainties than the ensemble forecasts (3 cases). However, since these two ensembles have both different resolutions and different model configurations, we cannot infer that one is more important than the other in reducing these uncertainties. Still, the ensemble hindcasts are clearly better than the ensemble forecasts, indicating an extended skill horizon of forecasts with newer model versions.

Another common problem are timing errors. The characteristic peaks in mean sea level pressure, horizontal wind speed, and precipitation are more often than not offset by at least 0.5 d. This problem is more apparent for the deterministic forecasts and the ensemble forecasts, while the ensemble hindcasts show this error less often. The deterministic forecast and the ensemble forecast only differ in resolution. But the ensemble hindcasts are run with considerably different model versions. The changes include (among others): more accurate initial conditions (more modern reanalyses), updated methods of generating perturbed ensemble members, and better parameterisations of small scale processes. Since the timing errors are comparable among the operational forecasts, but a smaller issue for the ensemble hindcasts, it appears that these errors are less dependent on resolution, and more sensitive to other changes in the models. Given that the parameterisations are updated more frequently than other model components, this again hints at the relevance of small scale dynamical processes as a source of forecast errors.

Finally, we note that most forecasts underestimate the intense phases of the cyclones. This is partly due to the timing errors mentioned above, and partly due to underestimated amplitudes of the different time series. While the amplitude of the pressure time series is generally predicted with good accuracy, strong winds and heavy precipitation are typically underestimated. For wind speed, the amplitudes generally range between 30 % and 60 % relative to the true ERA5 value (3 cases). The deterministic forecast is at least as accurate as the other forecasts (5 cases). The ensemble hindcasts are comparably accurate (3 cases) or worse (2 cases), while the ensemble forecasts are the least accurate (2 cases). The situation is more variable for precipitation, with amplitudes between 10 % and 80 % (4 cases). Here, the ensemble hindcasts are the most accurate (4 cases), with the deterministic forecast being less accurate (4 cases), and the ensemble forecast again showing the largest errors (3 cases). Overall, the ensemble forecast tends to underestimate the intense phases of the cyclones the most. This indicates that the amplitudes of the extreme events are better predicted with higher resolution. Since the ensemble hindcasts are most accurate for precipitation, it can also be inferred that predictions of heavy rainfall particularly benefit from more accurate initial conditions or better parameterisations of small scale processes.

Improvements with middle lead time forecasts

Most problems seen for the long lead time forecasts are smaller issues with the middle lead time forecasts. The cyclones are forecasted with smaller shifts, and the positional uncertainties are reduced. However, these improvements are sometimes not seen for wind speed (2 cases). As with the long lead time forecasts, the ensemble forecast still shows much more uncertainty than the ensemble hindcasts (3 cases). But most forecasts show smaller timing errors (3 cases), and amplitudes that are closer to the ERA5 reference data (4 cases). It is difficult to identify a forecast type that generally improves the most. This is because the improvements vary across the different events, and even across the different atmospheric fields. It thus appears that the chosen weather events are governed by distinct dynamics, where one event may be more sensitive to resolution, while another may require precise parameterisations of small scale processes.

Remaining issues with short lead time forecasts

The short lead time forecasts show smaller shifts, reduced uncertainties, fewer timing errors, and more accurate amplitudes. The evolution of pressure is predicted with high accuracy (5 cases). However, this is generally not true for wind speed and precipitation (3 cases). The strong winds are mostly underestimated (3 cases). Similarly, the precipitation forecasts are not accurate enough to correctly diagnose the societal impact of the weather events (3 cases). This is typically due to remaining shifts or rain cells that are too small. At short lead times, the uncertainties are more comparable between the ensemble forecasts and the ensemble hindcasts.

SAL evaluation

In terms of forecast use for high impact weather, the most important fields are horizontal wind speed and precipitation. But these fields also tend to be less predictable than mean sea level pressure. It is thus particularly interesting to see up to which lead times the forecasts of these fields are still useful. For precipitation, the SAL indices, described in section 2.2.2, allow us to quantify the usefulness of a given forecast.

In Figure 4.1, we show the lead time dependence of the A and L indices for the four events with extreme precipitation. Note that we leave out the structure index S, since most events only show a few precipitation cells (or "features"), making the structure index less informative. Both scores clearly improve for smaller forecast lead times, with forecasts converging towards the optimal scores ($A = 0, L = 0$). For lead times larger than 3 d, most forecasts clearly underestimate the amplitude of precipitation. There are only a few exceptions seen in Figure 4.1c. It is also clear that the long lead time forecasts predict displaced precipitation fields, while the shorter lead time forecasts get the location right. In general the errors and uncertainties are larger for the amplitude than for the location of precipitation. In terms of amplitude, only the latest case (Figure 4.1d) shows much smaller errors and uncertainty ranges, which may be the result of model improvements between 2006 and 2011. In contrast, the location uncertainties are comparable across all four cases. We can also see that the shortest lead time forecasts still show some errors. For the oldest case (Figure 4.1a), the amplitude remains too small, while all other cases show forecasts that tend to overestimate the amplitude. Finally, we note that the ensemble hindcasts (triangles) are generally more accurate than the ensemble forecasts (circles), while the deterministic forecasts (squares) are more variable.

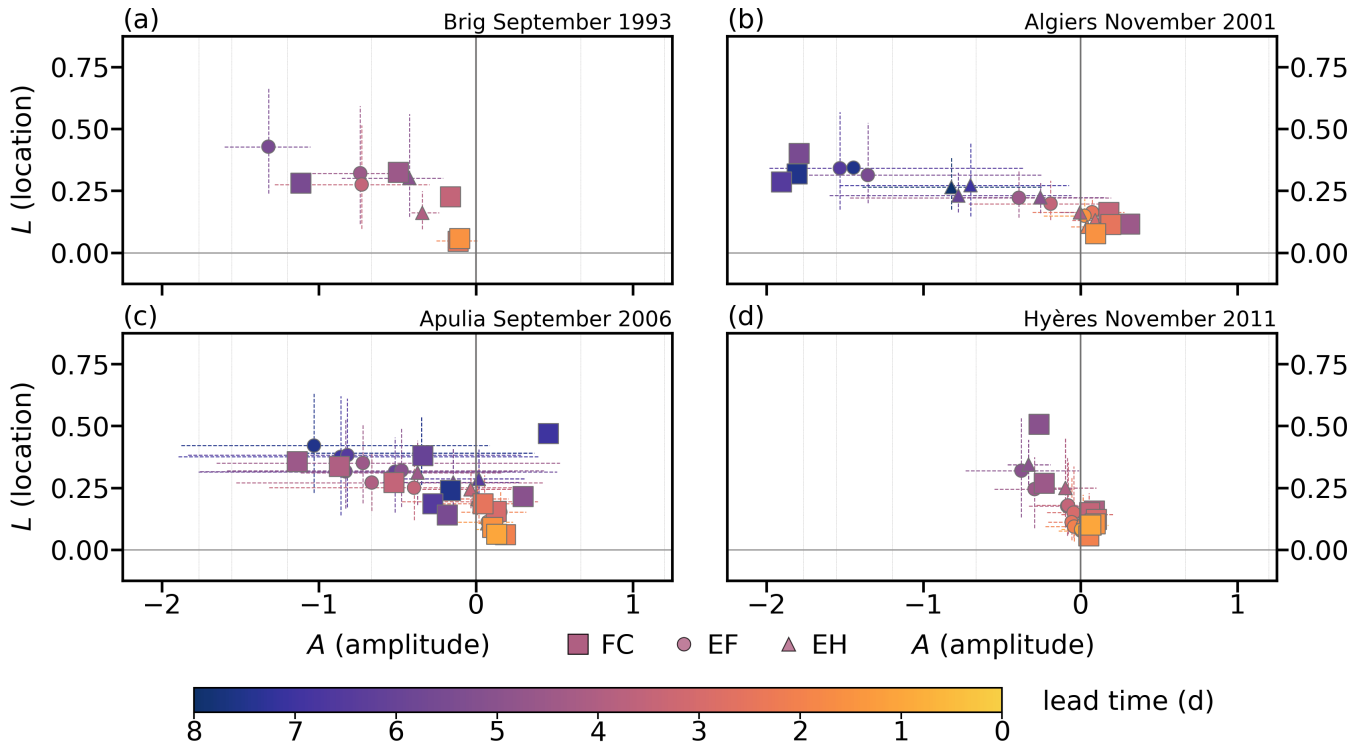


Figure 4.1: Comparison of A and L indices for four cases. The thick grey lines emphasise the optimal scores ($A = 0, L = 0$), and the thin vertical lines indicate amplitude errors by factors of 2 (± 0.67), 4 (± 1.20), 8 (± 1.56), and 16 (± 1.76). Each panel shows a different case, indicated by the labels in the top right corners. The Jerusalem cyclone from March 1998 is missing, because its precipitation is considerably weaker compared to the other cases. The scores are first computed for each timestep in the evaluation time window, and then averaged over this time window. This leaves the forecast lead time, emphasised with colours, as the only remaining forecast characteristic (similar to Figures 3.8, 3.12, 3.18, 3.22, and 3.26). The deterministic forecasts are shown with squares. The mean scores of the ensemble forecasts are indicated with circles, and the mean scores of the ensemble hindcasts use triangles. For both ensembles, we emphasise the middle 90 % range with dashed lines. Note that there may be multiple ensemble hindcasts for the same forecast lead time. In that case, we take an average over the different hindcasts, to increase visibility.

Ranking of different forecast types

Given that forecast performances are not consistent across the different weather events, there is no general ranking that holds for all cases. Nevertheless, the ensemble forecasts are typically the least accurate, followed by the deterministic forecasts, and finally the ensemble hindcasts (3 cases). The main problems with the ensemble forecasts are the underestimated amplitudes and the large uncertainties. The deterministic forecasts predict the amplitudes more accurately (2 cases), and the ensemble hindcasts additionally show much smaller uncertainties (2 cases). Sometimes the ensemble hindcasts are comparable to the deterministic forecast (1 case), but especially for precipitation, the more modern hindcasts are clearly the most accurate (2 cases). Another advantage of the ensemble hindcasts is that they are clearly more accurate at larger lead times compared to the operational forecasts (3 cases). Note, however, that the performance of the ensemble hindcasts is comparable to the ensemble forecasts for the remaining two cases (Jerusalem 1998 and Hyères 2011), with the deterministic forecast being more accurate. We therefore see the potential benefits of forecasts based on newer model versions, but from our limited number of cases, we cannot definitively say that they are the best forecasts overall.

4.2 Limitations of methodology

We compare forecasts for events between 1993 and 2011. In order to obtain comparable results, the methods need to be applicable to all forecasts, especially the oldest ones from 1993. Since these old forecasts only include the more fundamental atmospheric fields, we limit our analysis to mean sea level pressure, horizontal wind speed, and precipitation. These quantities are essential to predict with high accuracy, but they are also influenced by small scale processes which may introduce considerable uncertainties into the forecasts. However, we cannot investigate these small scales, given our limited horizontal resolution of 1°. Another caveat of the old forecasts is that they are only available with a coarse temporal resolution of 12 h. In fact, this is also the timestep for some of the newer forecasts. It is thus not possible to study processes that take place over shorter time scales.

It should also be noted that many factors influence forecast accuracy. These can be separated into initial condition uncertainties and model uncertainties. The initial conditions are limited by observational uncertainties and the accuracy of the data assimilation system, while the model is limited by scientific understanding of atmospheric processes and computational aspects like resolution in time and space or numerical artefacts. Considering all these sources of uncertainty, we can only hypothesise why different forecasts may be more or less accurate. An exact attribution of forecast errors would require controlled numerical experiments.

Another important point is that we only consider five cases. On the one hand, this is an important first step towards a systematic evaluation of more cases. On the other hand, it is possible that an extended analysis involving more cases might reveal different error patterns. It is thus not clear, if our identified forecast characteristics can be generalised to other cases of Mediterranean cyclones. A related limitation is that the data availability varies considerably between the cases. This hinders a robust comparison of the different events.

Lastly, we note that our methodology involves several subjective choices. On the one hand, we subjectively choose the five extreme weather events out of many possible Mediterranean cyclones. These choices are mainly guided by societal impact, but partly also by statistical extremeness. But finding more specific criteria may help to cluster different types of cyclones, improving comparability. On the other hand, we subjectively define the evaluation regions for the regional means, and the evaluation time windows for the temporal means. These definitions are based on the most intense phase of the cyclones, as identified in the ERA5 reference data. Nevertheless, the results that we obtain may depend on these choices, and more systematic definitions may yield more robust results. This is also true for the categorisation of forecasts into long lead time forecasts, middle lead time forecasts, and short lead time forecasts.

Chapter 5

Summary and conclusions

Objectives

In this thesis, we focused on five extreme Mediterranean cyclones that led to considerable damages for society. The main goal was to study the forecast accuracy for these events, and to investigate the origin of forecast errors. We studied the behaviour of three different forecast types run at the ECMWF. The first two being the deterministic forecasts and the ensemble forecasts, which were operational at the time of the events. The third forecast type are the ensemble hindcasts, which use more modern versions of the ECMWF model to re-forecast the events from the past. By comparing the operational forecasts to the hindcasts, we analysed the evolution of forecast quality with time.

Methods

We retrieved forecasts with different initialisation times and categorised them into long lead time forecasts (typically more than 4.5 d before the mature cyclone), middle lead time forecasts (more than 3 d), and short lead time forecasts (more than 0.5 d). By evaluating these forecasts against the ERA5 reanalysis provided by ECMWF, it is possible to study forecast improvements with smaller lead times. The evaluations were based on several approaches involving the fields: mean sea level pressure, horizontal wind speed, and precipitation. In a first step, we assessed the spatial patterns of the forecasts within the Mediterranean basin, focusing on phase shifts of the fields and positional uncertainties. For the second step, we averaged the fields over subjectively defined evaluation regions, and compared the time series of the different forecasts. This evaluation was based on timing errors of the mature stages, amplitude errors of the cyclone intensities, and general accuracy relative to the ERA5 reference data. For the third step, we further averaged the regional means over a subjectively defined 1 d time window, including the most intense phase of the cyclone. This enabled us to quantify the accuracy of each forecast with a single number, making it possible to compare all forecasts, depending on their different initialisation times.

Results

The long lead time forecasts typically predict the Mediterranean cyclones with a northwards shift, and there are considerable uncertainties in the cyclone position. A possible source of these errors are uncertainties in the process of Rossby wave breaking over the North Atlantic. Another issue is that the most intense phase of the cyclone is predicted at the wrong time. This error is mostly seen for the operational forecasts, and the offsets are typically larger than 0.5 d. In addition, the forecasts tend to underestimate the intensity of the Mediterranean cyclones. This is a smaller problem for pressure, and more apparent for wind speeds and precipitation, which are generally too weak in the forecasts. For our case studies, the operational ensemble forecast was the least accurate forecast type. Given that it has the coarsest resolution, it may be concluded that higher resolution helps to accurately predict the intensity of a Mediterranean cyclone.

The problems with the long lead time forecasts are gradually resolved, as the forecast lead time becomes smaller and smaller. We find a distinct jump in forecast accuracy at a lead time between 4.5 d and 5 d relative to the mature stage of the cyclone. However, even the shortest lead time forecasts show some remaining issues. Both wind speed and precipitation are not sufficiently accurate to predict the societal impact of the corresponding Mediterranean cyclones.

Conclusions

While we identify the ensemble hindcasts as the best forecast type, followed by the deterministic forecasts, and then the ensemble forecasts, this conclusion is only true for three out of five cases. The main advantages of the hindcasts are smaller positional uncertainties, and better accuracy at longer lead times (extended skill horizon). Since the hindcasts are based on more recent model versions, this highlights the benefits of the model improvements over time.

But overall, there is no single best forecast type for all events, at all lead times, and for all atmospheric fields. Instead, forecast accuracy is found to be very variable and each event shows unique error patterns. To answer the question, how specific performance improvements are related to upgrades at the ECMWF, it would thus be necessary to extend our methodology to a more systematic and climatological analysis. This could be an interesting topic for future research. However, the forecast data provided by ECMWF may not be sufficient to answer this question, and it may be necessary to run specifically designed numerical experiments to study the effects of better resolution, better initial conditions, or better parameterisations of small scale processes.

Another point to consider is that our five cases only include Mediterranean cyclones between 1993 and 2011. In this range, there are still noticeable differences between the operational forecasts and the hindcasts. But for more recent events, this difference among the forecast types is expected to become smaller. Therefore, it may be interesting to look at cases after 2011, and investigate whether the different forecast types perform similarly or not.

Appendix A

Data availability

As mentioned in section 2.1.4, we show in Table A.1 the available forecasts for each extreme event. The details on the model versions corresponding to the different forecasts are listed in Tables 2.1 to 2.3.

Table A.1: Available forecast data for each extreme weather event. The initialisation times are shown in the UTC format dd–hh. We distinguish between the deterministic forecasts (FC), the ensemble forecasts (EF), and the ensemble hindcasts (EH). In the case of the ensemble hindcasts, we also include the available model versions (in parentheses).

FC	EF	EH
Brig September 1993:		
18–12	18–12	18–00 (2008)
19–12	19–12	19–00 (2013)
20–12	20–12	20–00 (2012)
21–12	-	-
22–12	-	22–00 (2011)
23–12	-	23–00 (2010)
Jerusalem March 1998:		
08–12	08–12	08–00 (2012, 2018)
09–12	09–12	09–00 (2017)
10–12	10–12	10–00 (2011, 2016)
11–12	11–12	11–00 (2010)
12–12	12–12	12–00 (2009, 2015, 2018)
13–12	13–12	13–00 (2008, 2014, 2017)
14–12	14–12	14–00 (2013, 2016)

Table A.2: Available forecast data for each extreme weather event (continued).

FC	EF	EH
Algiers November 2001:		
03–12	03–12	03–00 (2011, 2016)
04–12	04–12	04–00 (2010, 2019, 2021)
05–12	05–12	05–00 (2009, 2015, 2018, 2020)
06–12	06–12	06–00 (2008, 2014, 2017)
07–12	07–12	07–00 (2013, 2016, 2019)
08–12	08–12	08–00 (2012, 2018, 2021)
09–12	09–12	09–00 (2015, 2017, 2020)
Apulia September 2006:		
19–00, 19–12	19–00, 19–12	19–00 (2013, 2016, 2019, 2022)
20–00, 20–12	20–00, 20–12	20–00 (2012, 2018, 2021)
21–00, 21–12	21–00, 21–12	21–00 (2015, 2017, 2020)
22–00, 22–12	22–00, 22–12	22–00 (2011, 2016, 2022)
23–00, 23–12	23–00, 23–12	23–00 (2010, 2019, 2021)
24–00, 24–12	24–00, 24–12	24–00 (2009, 2015, 2018, 2020)
25–00, 25–12	25–00, 25–12	25–00 (2008, 2014, 2017)
Hyères November 2011:		
01–00, 01–12	01–00, 01–12	01–00 (2012, 2018, 2021)
02–00, 02–12	02–00, 02–12	02–00 (2015, 2017, 2020)
03–00, 03–12	03–00, 03–12	03–00 (2016, 2022)
04–00, 04–12	04–00, 04–12	04–00 (2019, 2021)
05–00, 05–12	05–00, 05–12	05–00 (2015, 2018, 2020)
06–00, 06–12	06–00, 06–12	06–00 (2014, 2017)

Appendix B

Additional figures

As mentioned in chapter 3, this section includes all supplemental material on the five cases of extreme Mediterranean cyclones. As in chapter 3, we split the additional material into five sections, one for each case.

B.1 Brig September 1993

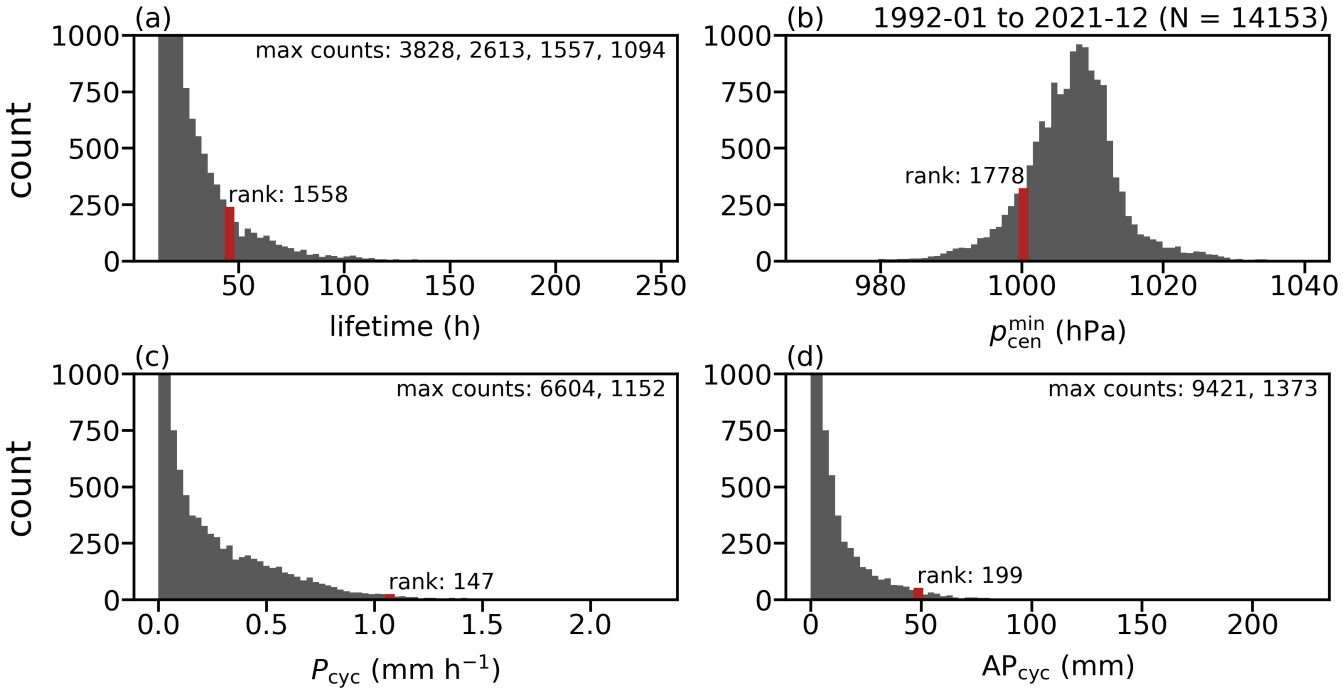


Figure B.1: Statistics for the September 1993 cyclone. Similar to Figure 2.1, but with the bins corresponding to the Brig cyclone emphasised in red, and the ranks listed for each quantity.

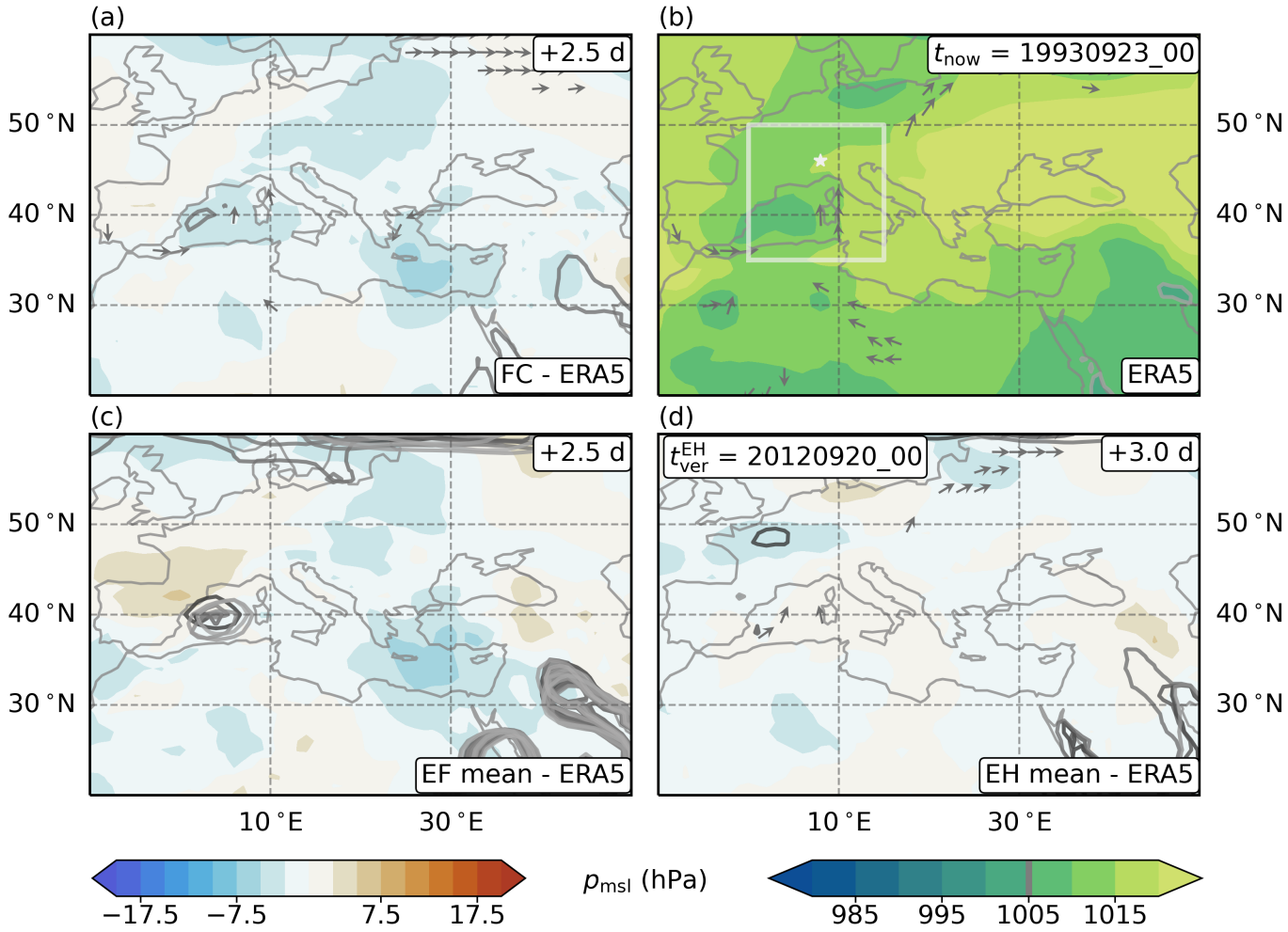


Figure B.2: Middle lead time forecasts of pressure for 1993, 1 d before the (middle of the) most intense cyclone stage. Similar to Figure 3.2, but for a single initialisation time. (a) Deterministic forecast (FC), (b) ERA5 reference, (c) ensemble forecast (EF), and (d) ensemble hindcast (EH).

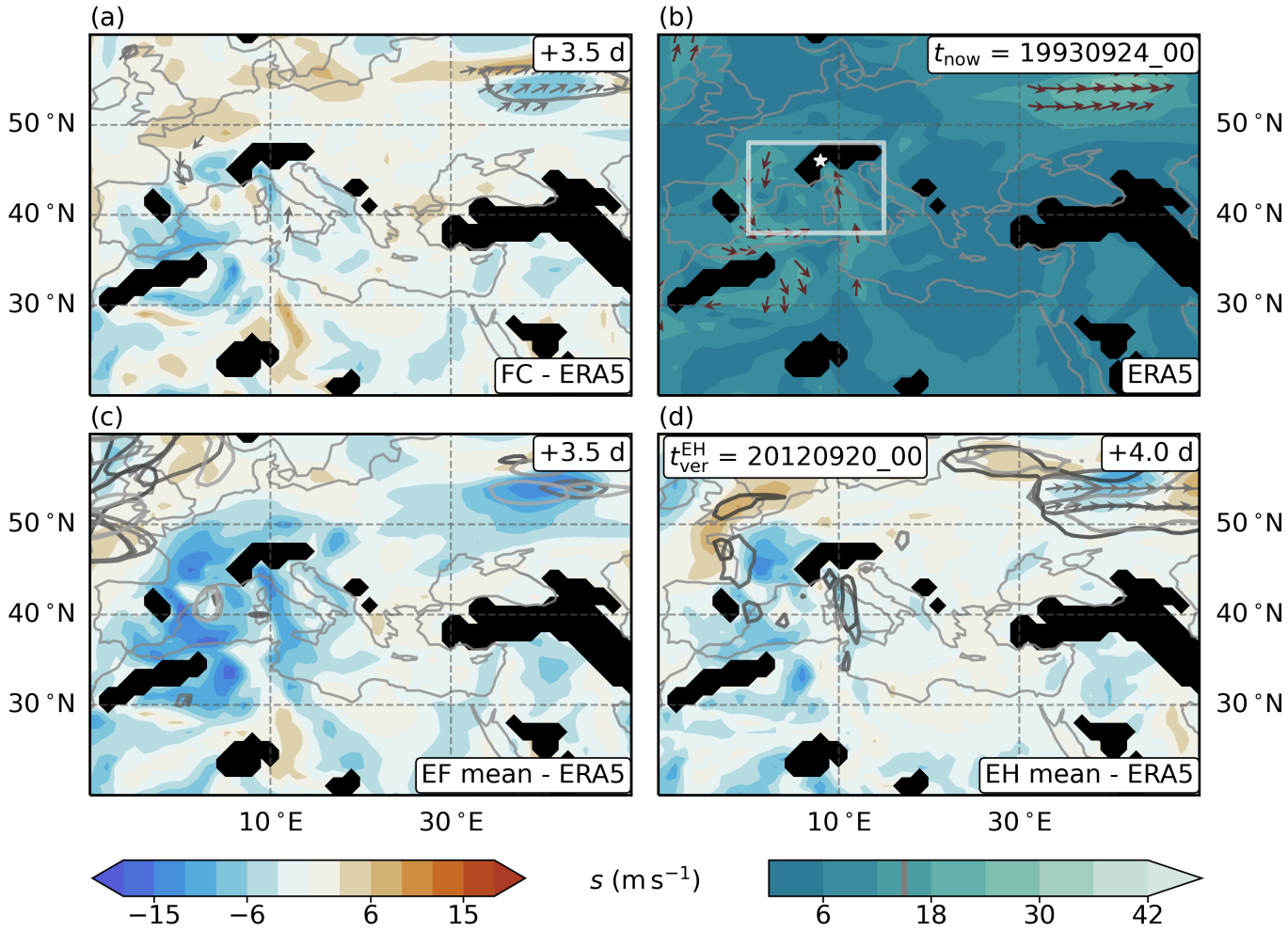


Figure B.3: Middle lead time forecasts of wind speed for 1993, in the middle of the most intense cyclone stage. Similar to Figure B.2, but for a different reference time, and for horizontal wind speed on 900 hPa.

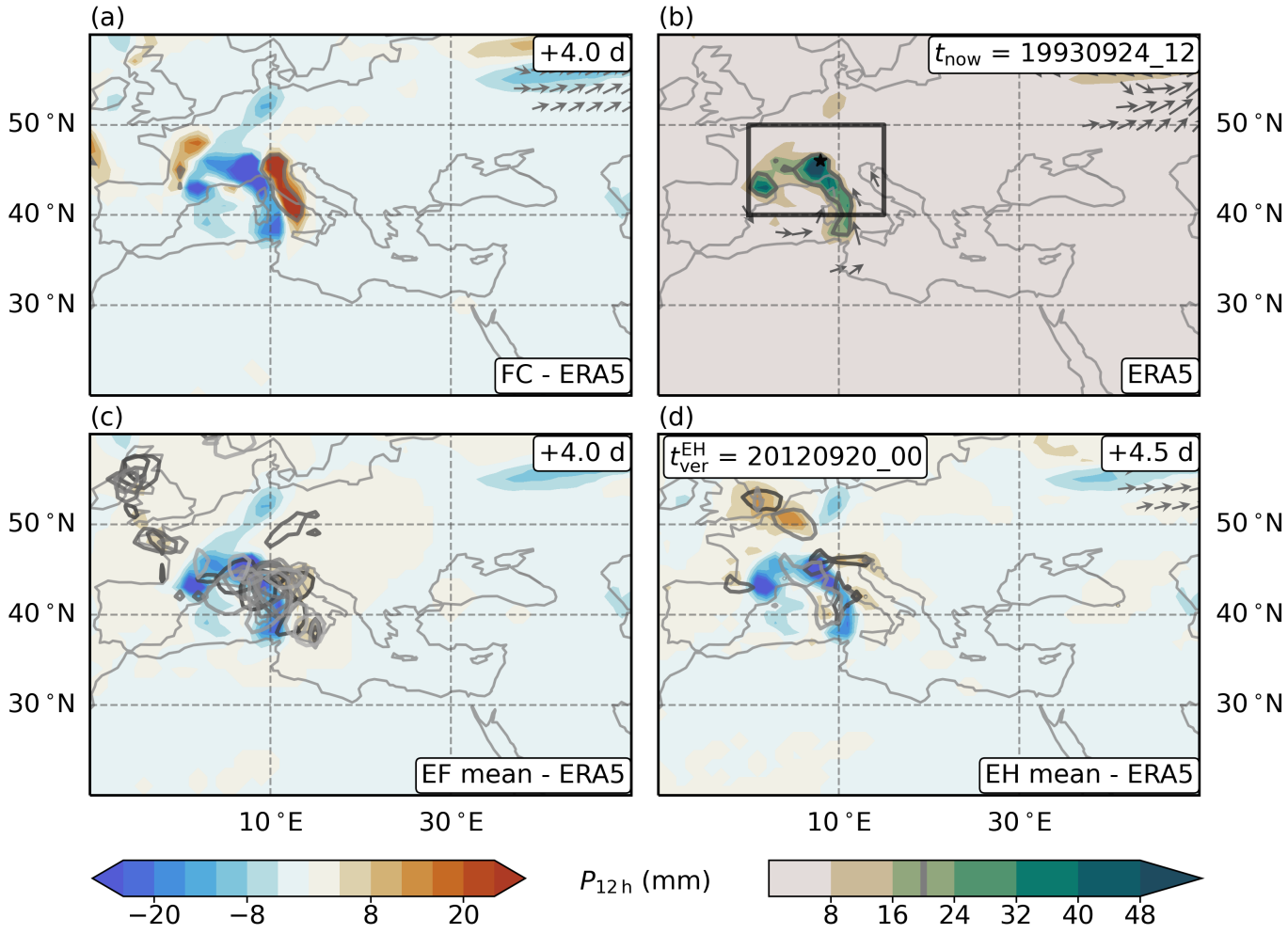


Figure B.4: Middle lead time forecasts of precipitation for 1993, at the end of the most intense cyclone stage. Similar to Figure B.3, but for a different reference time, and for 12 h accumulated precipitation.

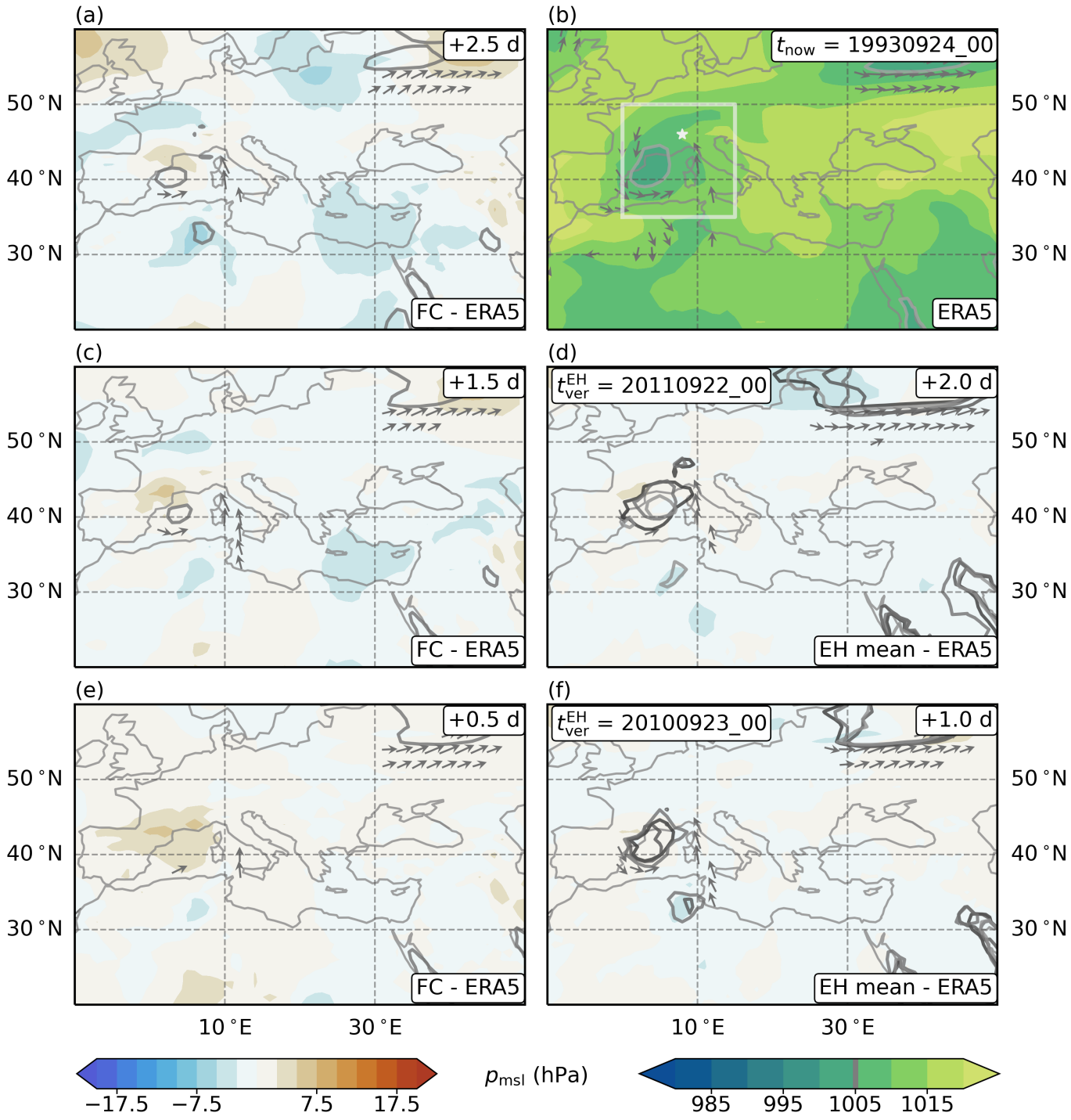


Figure B.5: Short lead time forecasts of pressure for 1993, in the middle of the most intense cyclone stage. Similar to Figure 3.5, but for a different reference time, and for mean sea level pressure. There is no ensemble forecast available for this initialisation time.

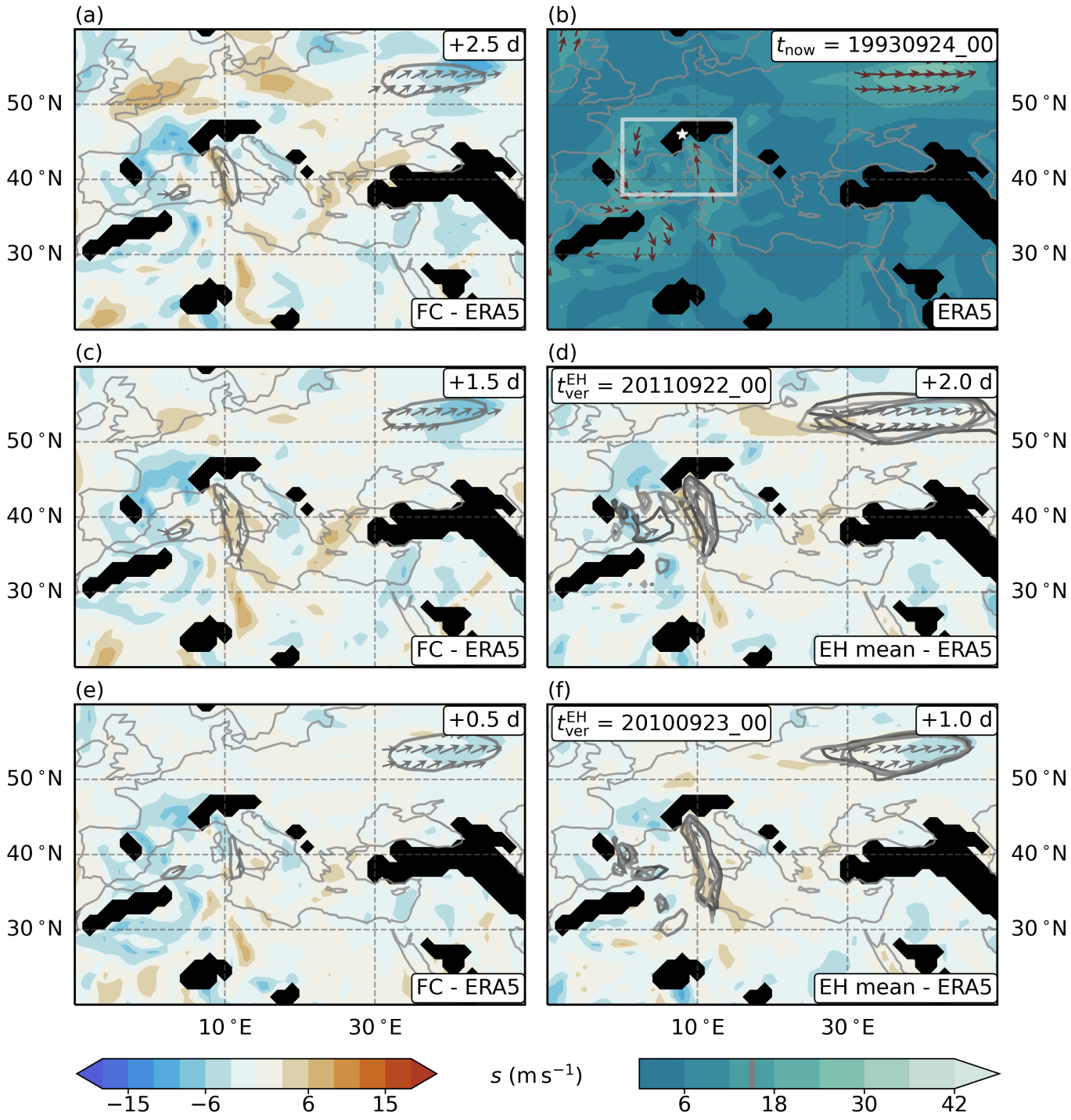


Figure B.6: Short lead time forecasts of wind speed for 1993, in the middle of the most intense cyclone stage. Similar to Figure B.5, but for horizontal wind speed on 900 hPa. There is no ensemble forecast available for this initialisation time.

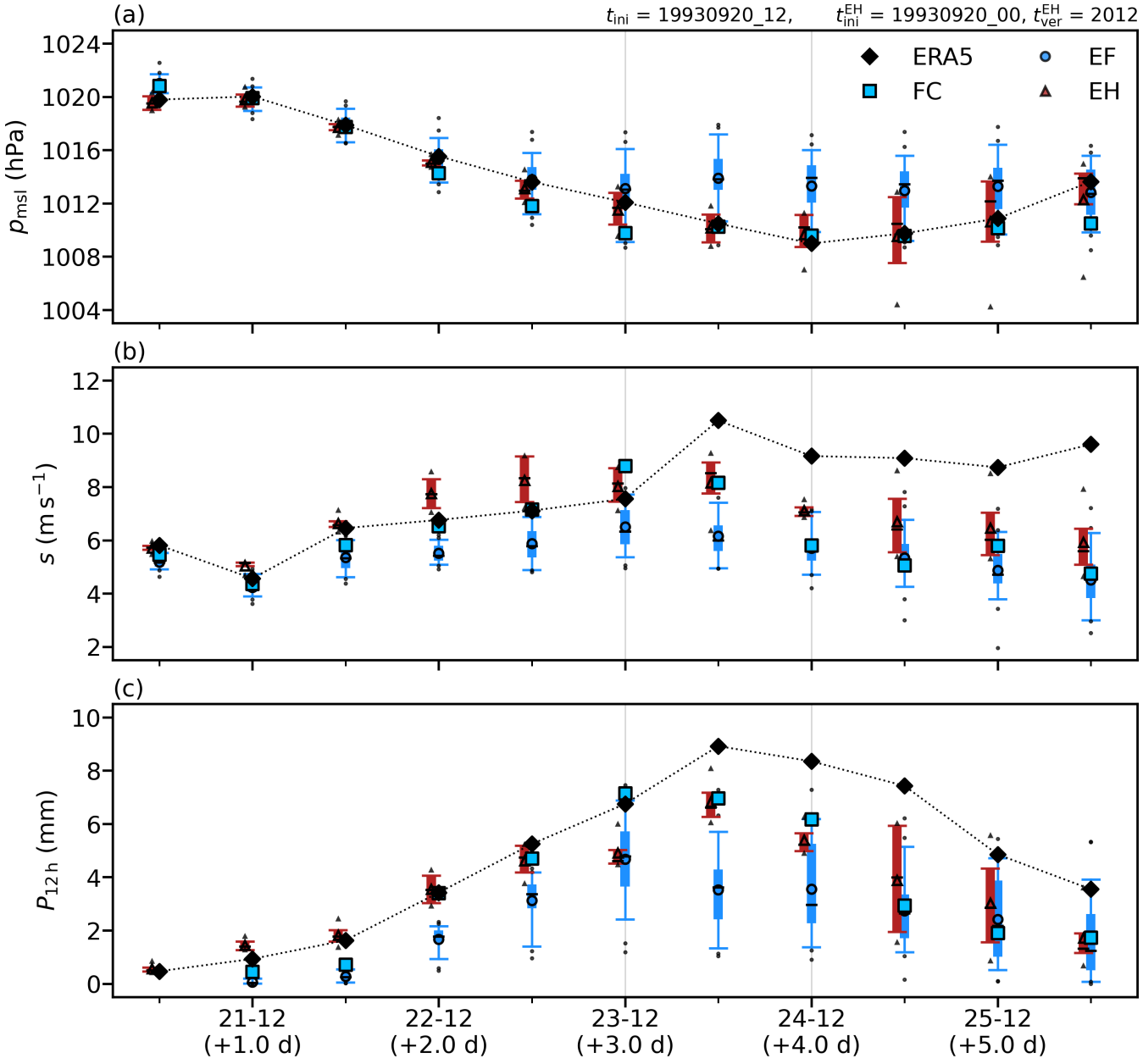


Figure B.7: Time series of forecasts at a middle lead time for 1993. Similar to Figure 3.6. The forecasts were initialised on 20. September 1993, the ensemble hindcast (EH) at 00 UTC, and both the deterministic forecast (FC) and the ensemble forecast (EF) at 12 UTC.

B.2 Jerusalem March 1998

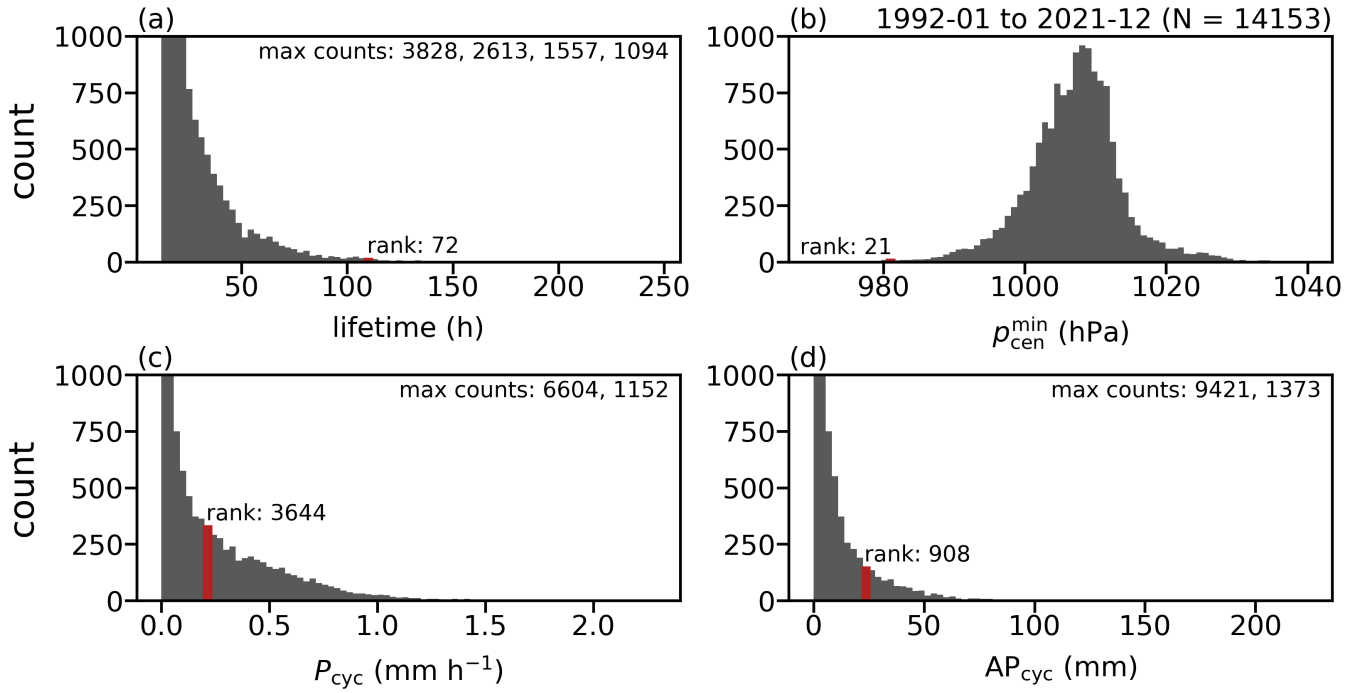


Figure B.8: Statistics for the March 1998 cyclone. Similar to Figure B.1, but for a different reference cyclone.

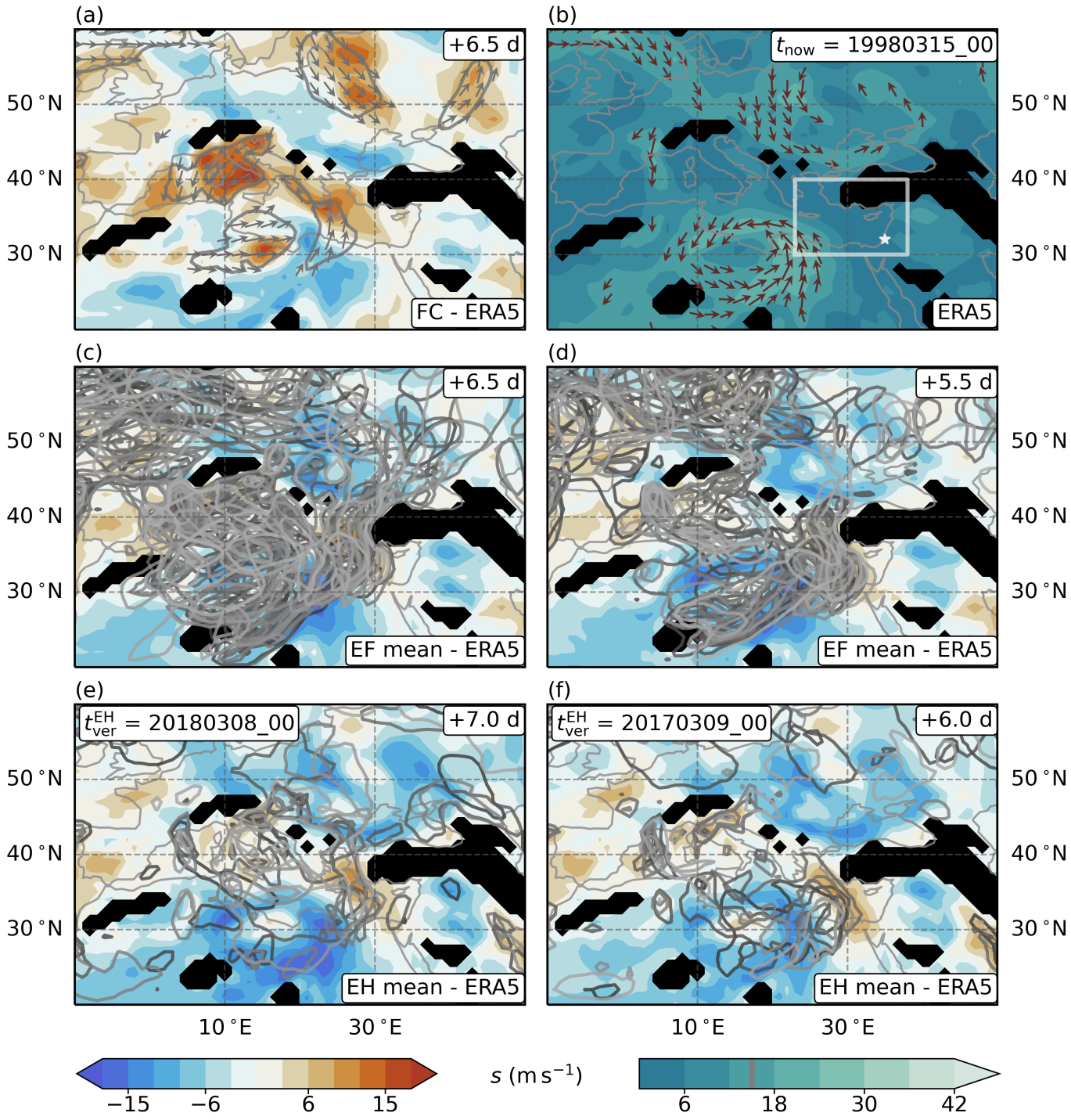


Figure B.9: Long lead time forecasts of wind speed for 1998, 1 d before the (middle of the) most intense cyclone stage. Similar to Figure 3.10, but for a different reference time, and for horizontal wind speed on 900 hPa.

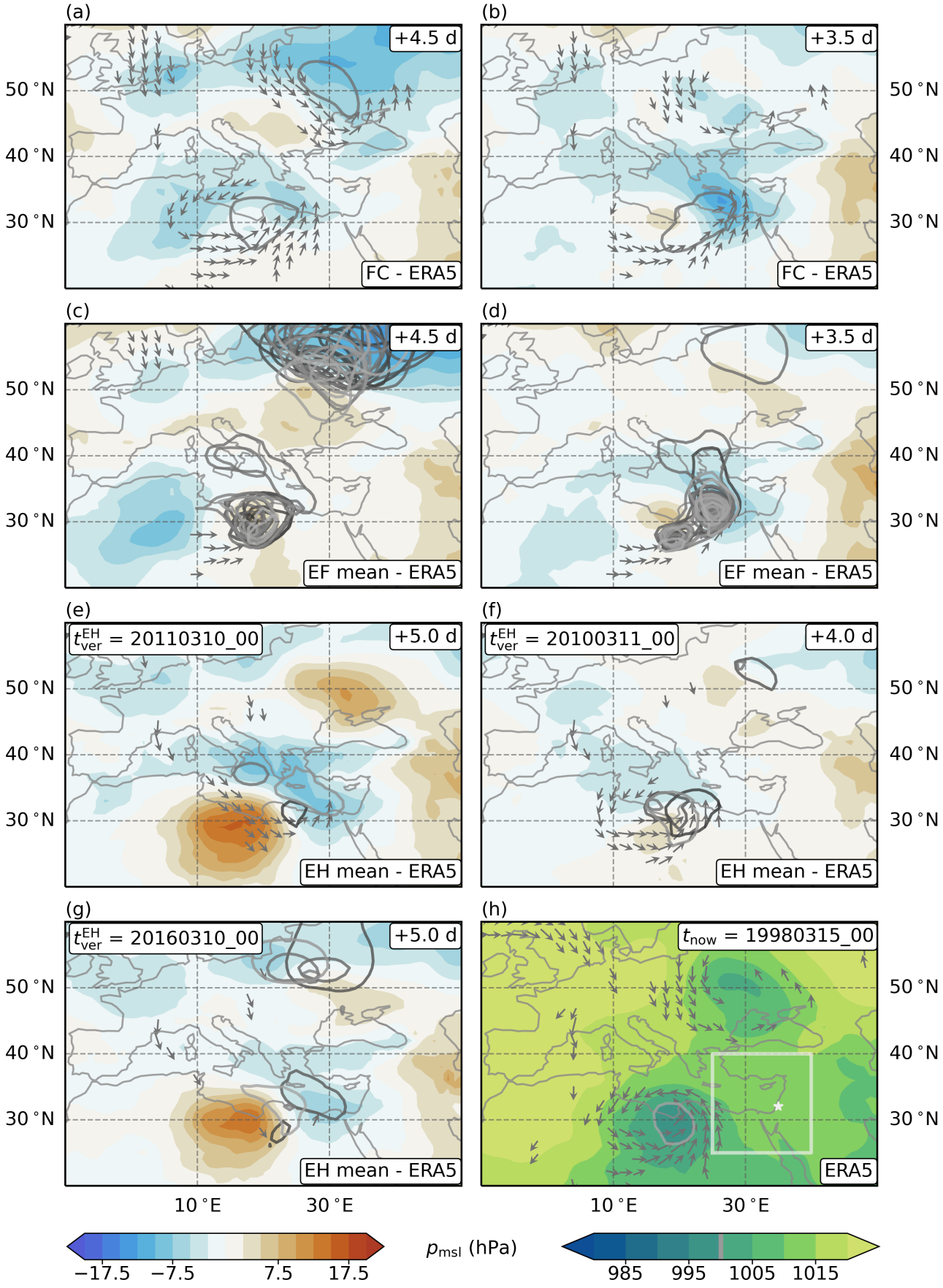


Figure B.10: Middle lead time forecasts of pressure for 1998, 1 d before the (middle of the) most intense cyclone stage. Similar to Figure 3.10, but for a different reference time. (a) and (b) deterministic forecasts (FC), (c) and (d) ensemble forecasts (EF), (e), (f), and (g) ensemble hindcasts (EH), and (h) ERA5 reference.

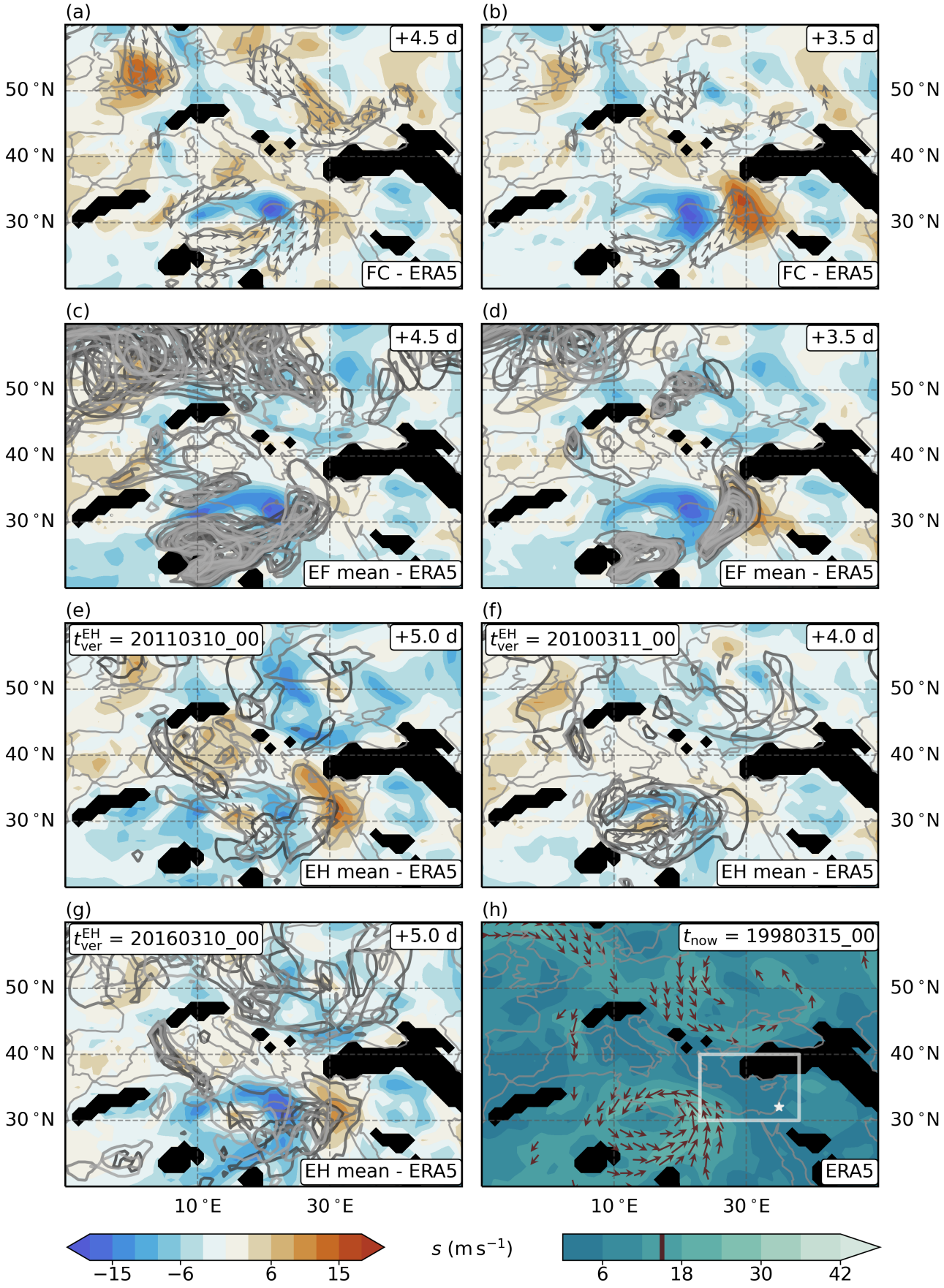


Figure B.11: Middle lead time forecasts of wind speed for 1998, 1 d before the (middle of the) most intense cyclone stage. Similar to Figure B.10, but for horizontal wind speed on 900 hPa.

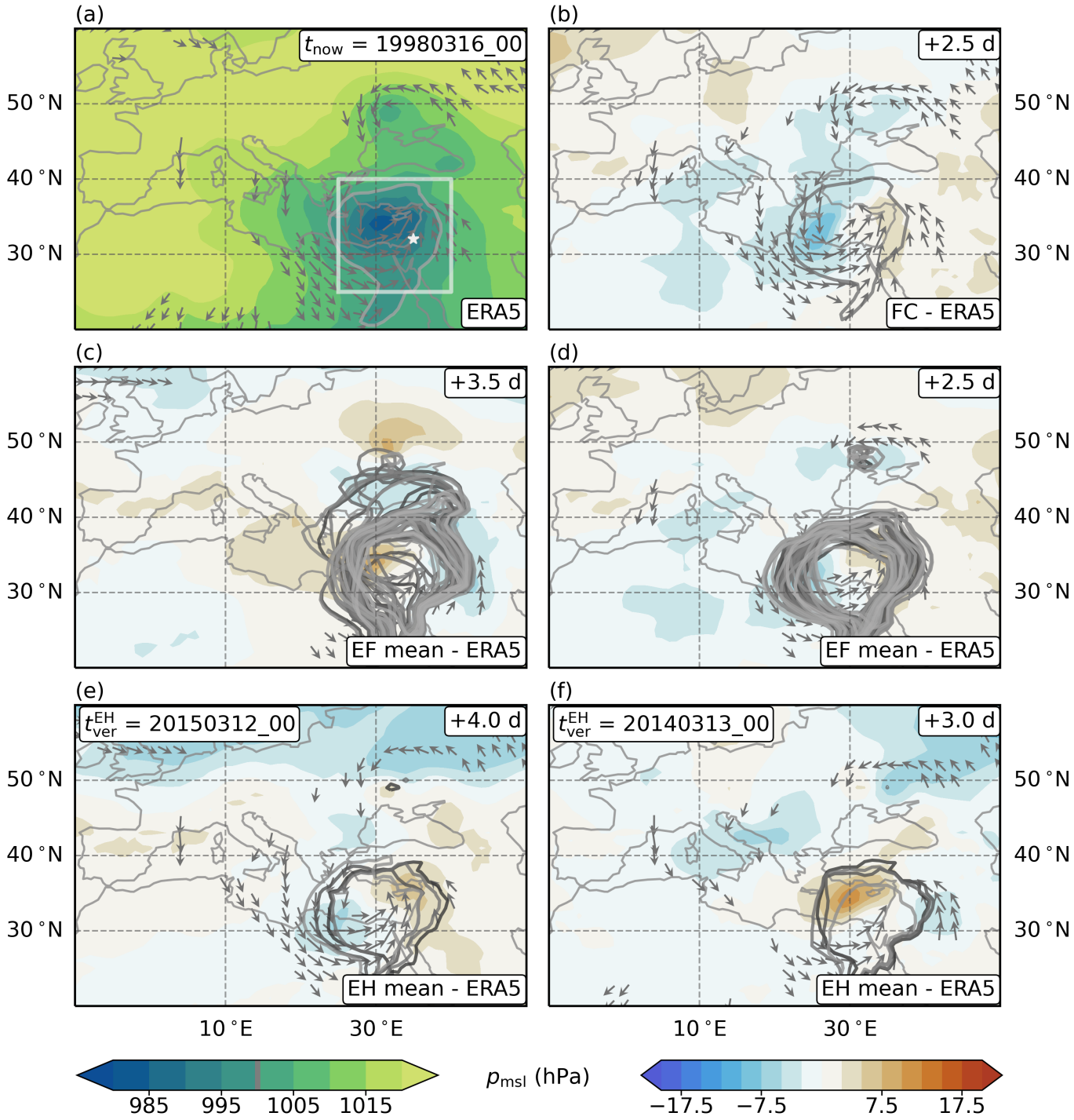


Figure B.12: Short lead time forecasts of pressure for 1998, in the middle of the most intense cyclone stage. Similar to Figure 3.10, but for a different reference time. (a) ERA5 reference, (b) deterministic forecast (FC), (c) and (d) ensemble forecasts (EF), (e) and (f) ensemble hindcasts (EH).

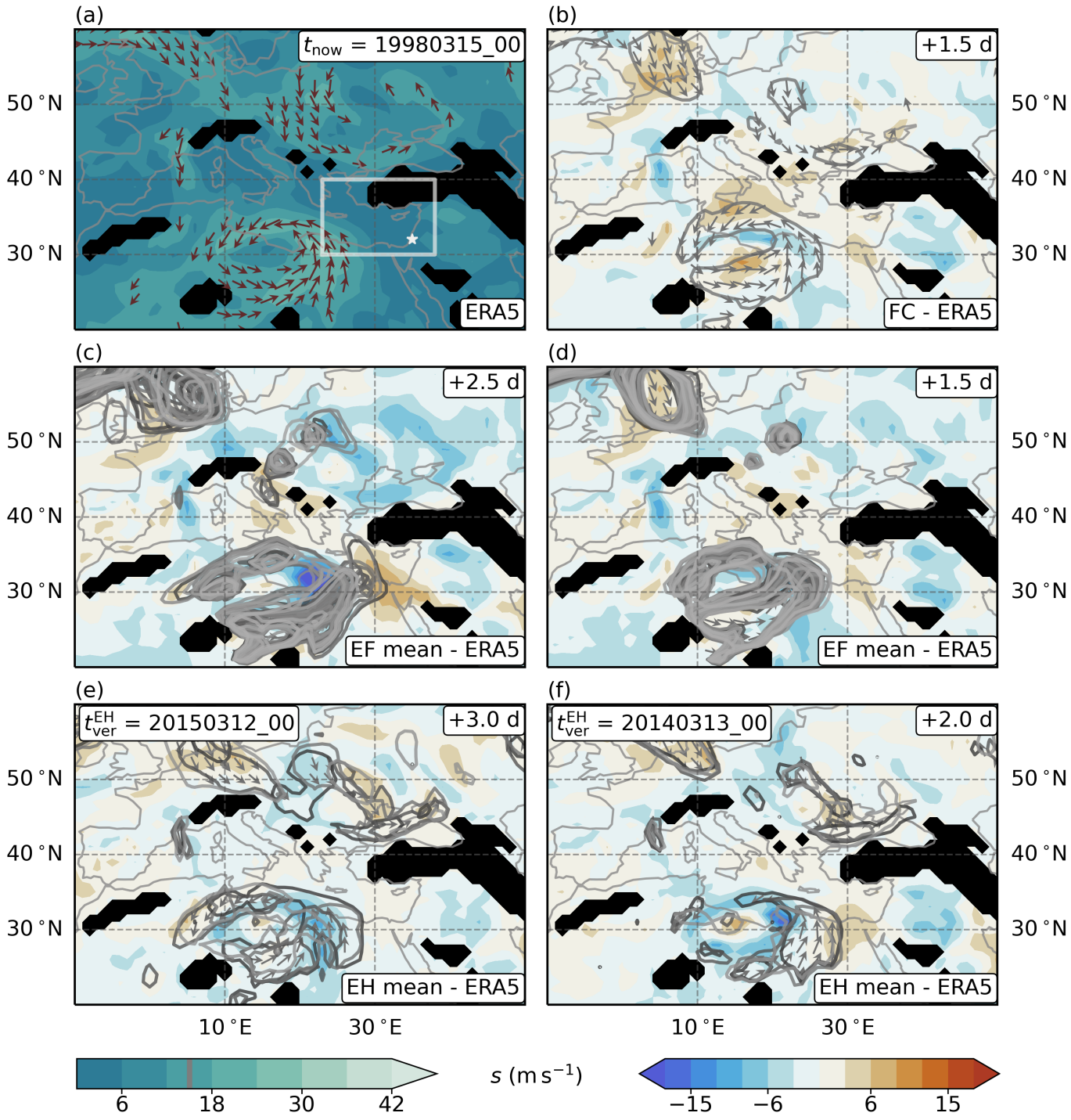


Figure B.13: Short lead time forecasts of wind speed for 1998, 1 d before the (middle of the) most intense cyclone stage. Similar to Figure B.12, but for a different reference time, and for horizontal wind speed on 900 hPa.

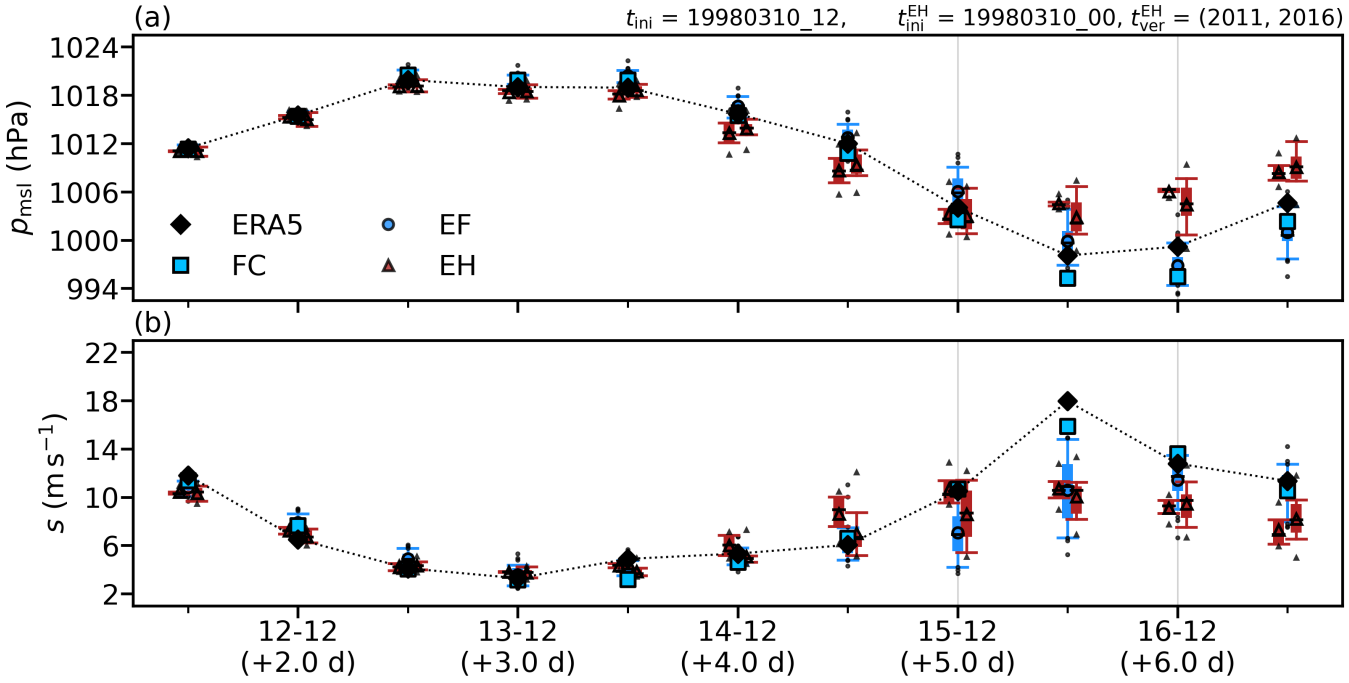


Figure B.14: Time series of forecasts at a middle lead time for 1998. Similar to Figure 3.11. The forecasts were initialised on 10. March 1998, the ensemble hindcasts (EH) at 00 UTC, and both the deterministic forecast (FC) and the ensemble forecast (EF) at 12 UTC.

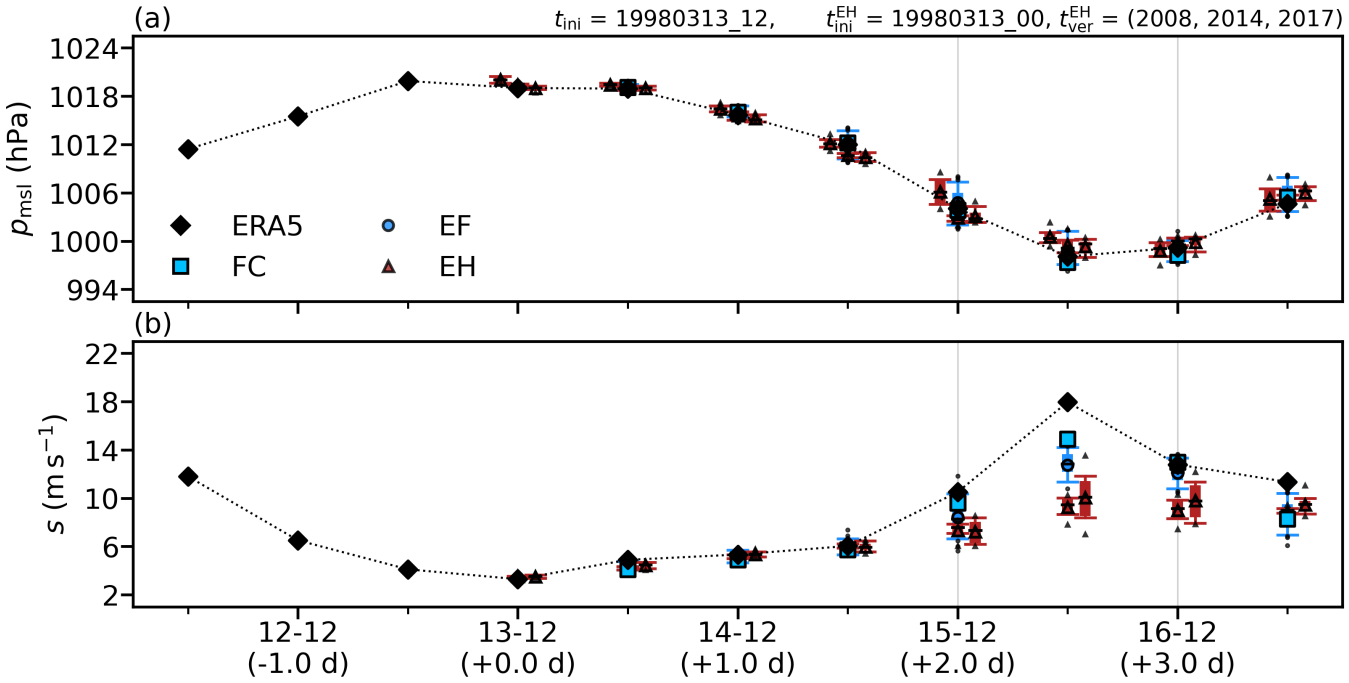


Figure B.15: Time series of forecasts at a short lead time for 1998. Similar to Figure B.14. The forecasts were initialised on 13. March 1998, the ensemble hindcasts (EH) at 00 UTC, and both the deterministic forecast (FC) and the ensemble forecast (EF) at 12 UTC.

B.3 Algiers November 2001

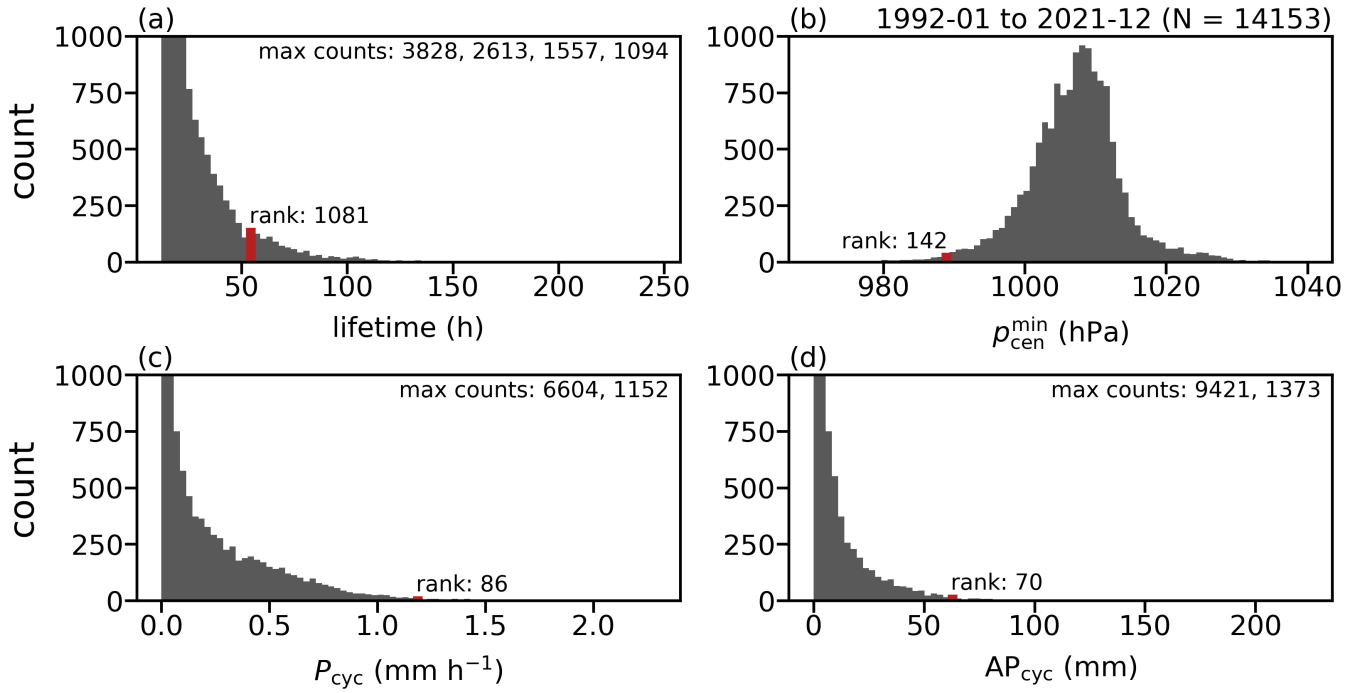


Figure B.16: Statistics for the November 2001 cyclone. Similar to Figure B.8, but for a different reference cyclone.

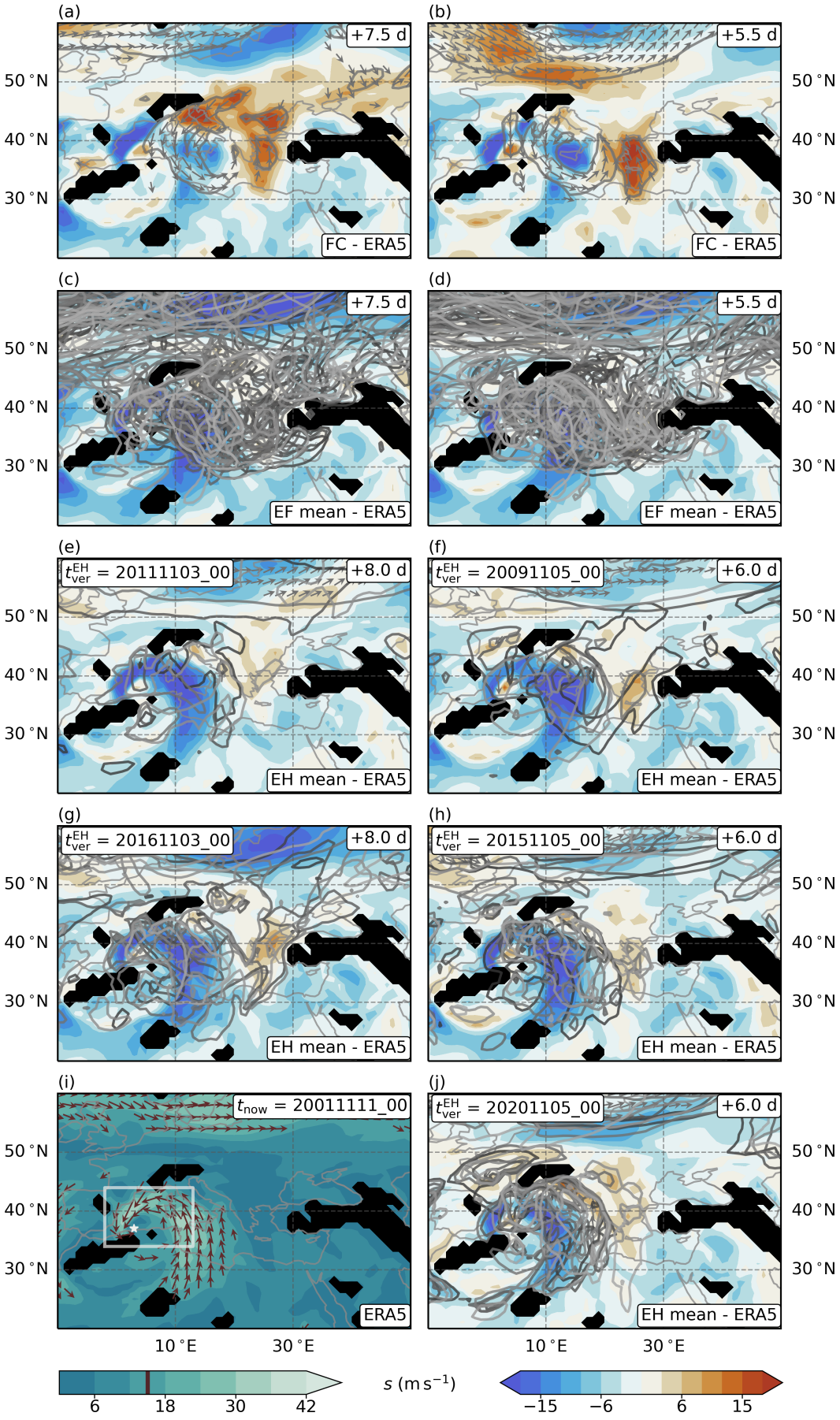


Figure B.17: Long lead time forecasts of wind speed for 2001, in the middle of the most intense cyclone stage. Similar to Figure 3.14, but for a different reference time, and for horizontal wind speed on 900 hPa.

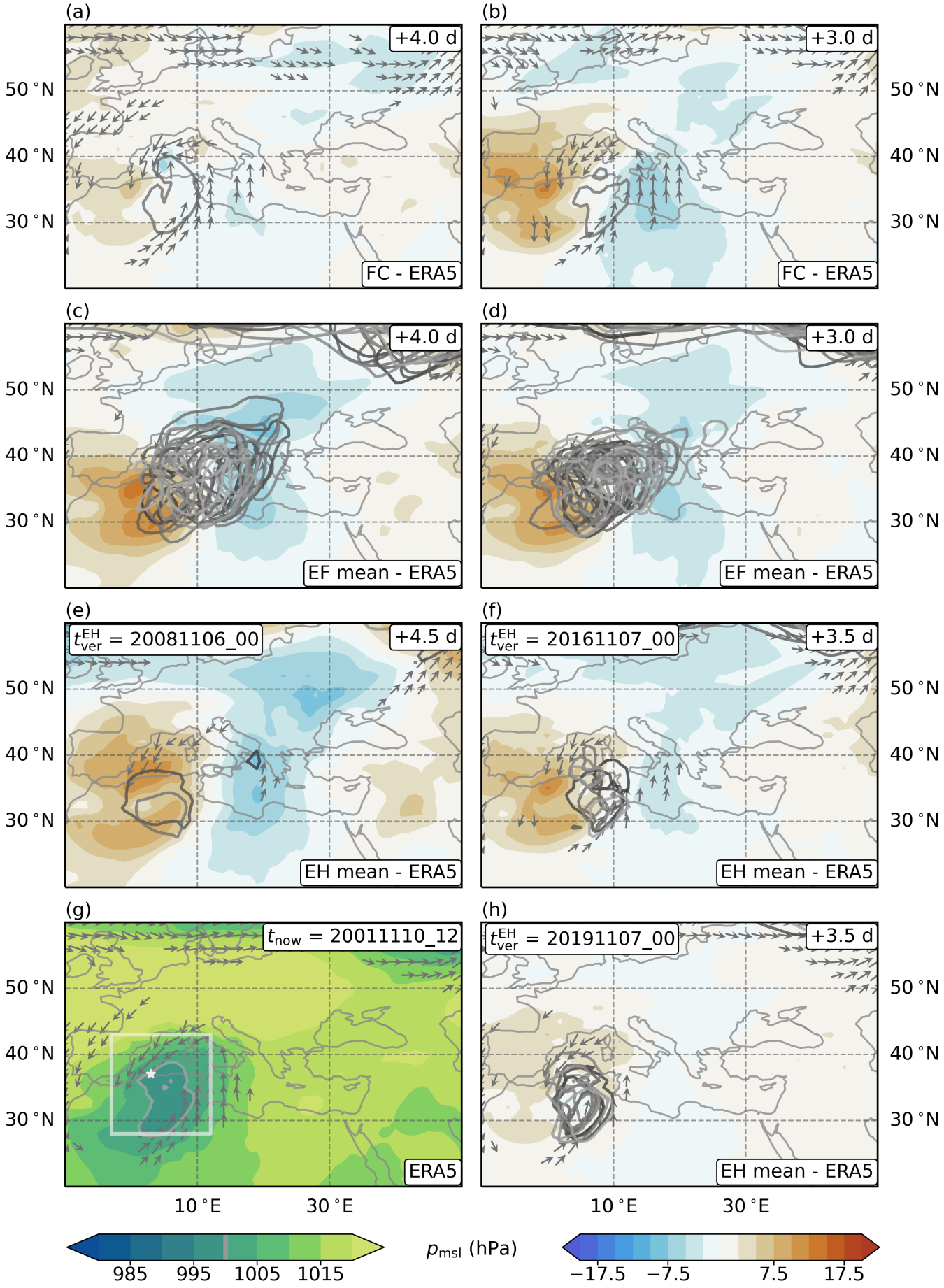


Figure B.18: Middle lead time forecasts of pressure for 2001, at the beginning of the most intense cyclone stage. Similar to Figure 3.14, but for a different reference time.

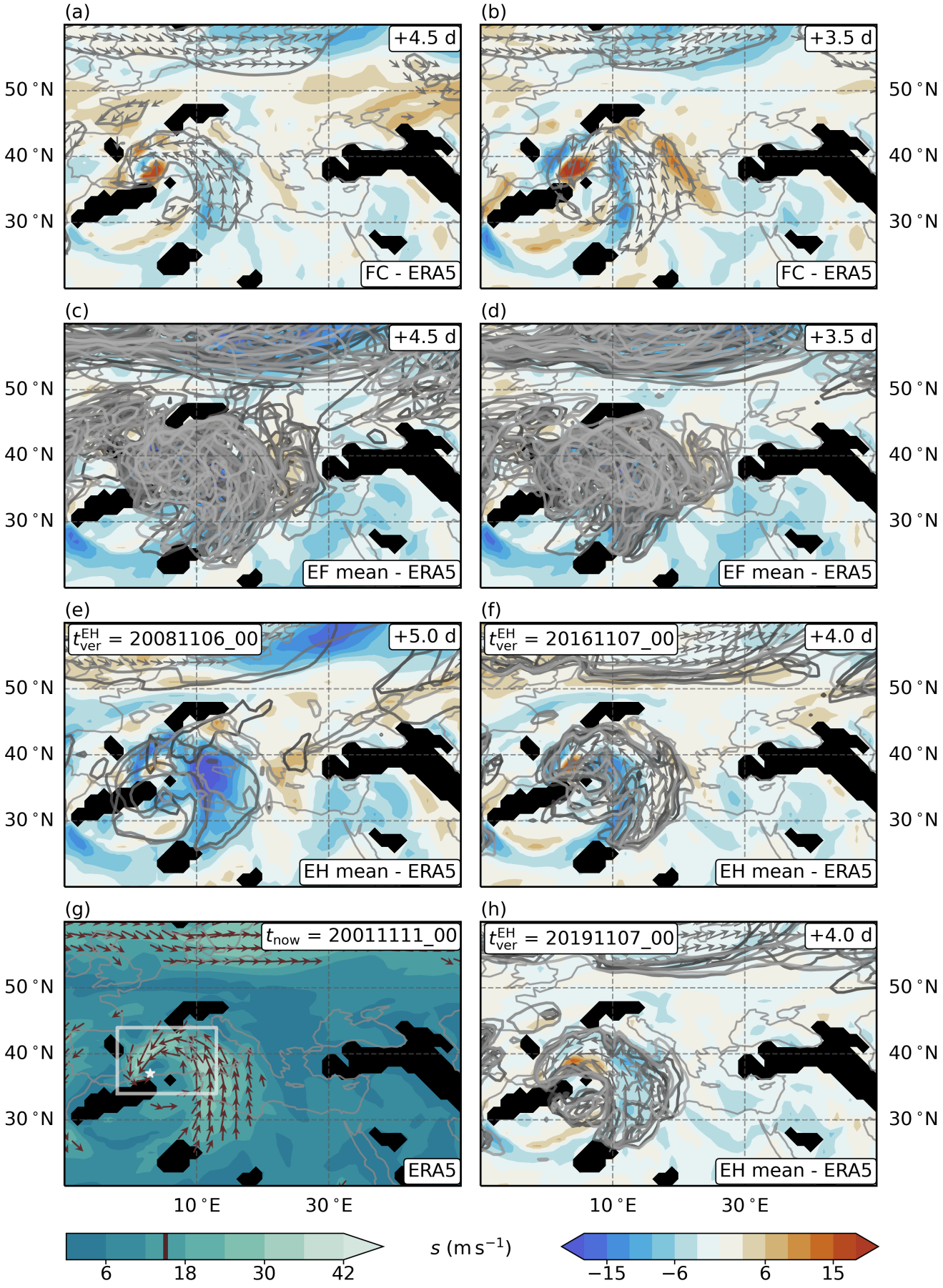


Figure B.19: Middle lead time forecasts of wind speed for 2001, in the middle of the most intense cyclone stage. Similar to Figure B.18, but for a different reference time, and for horizontal wind speed on 900 hPa.

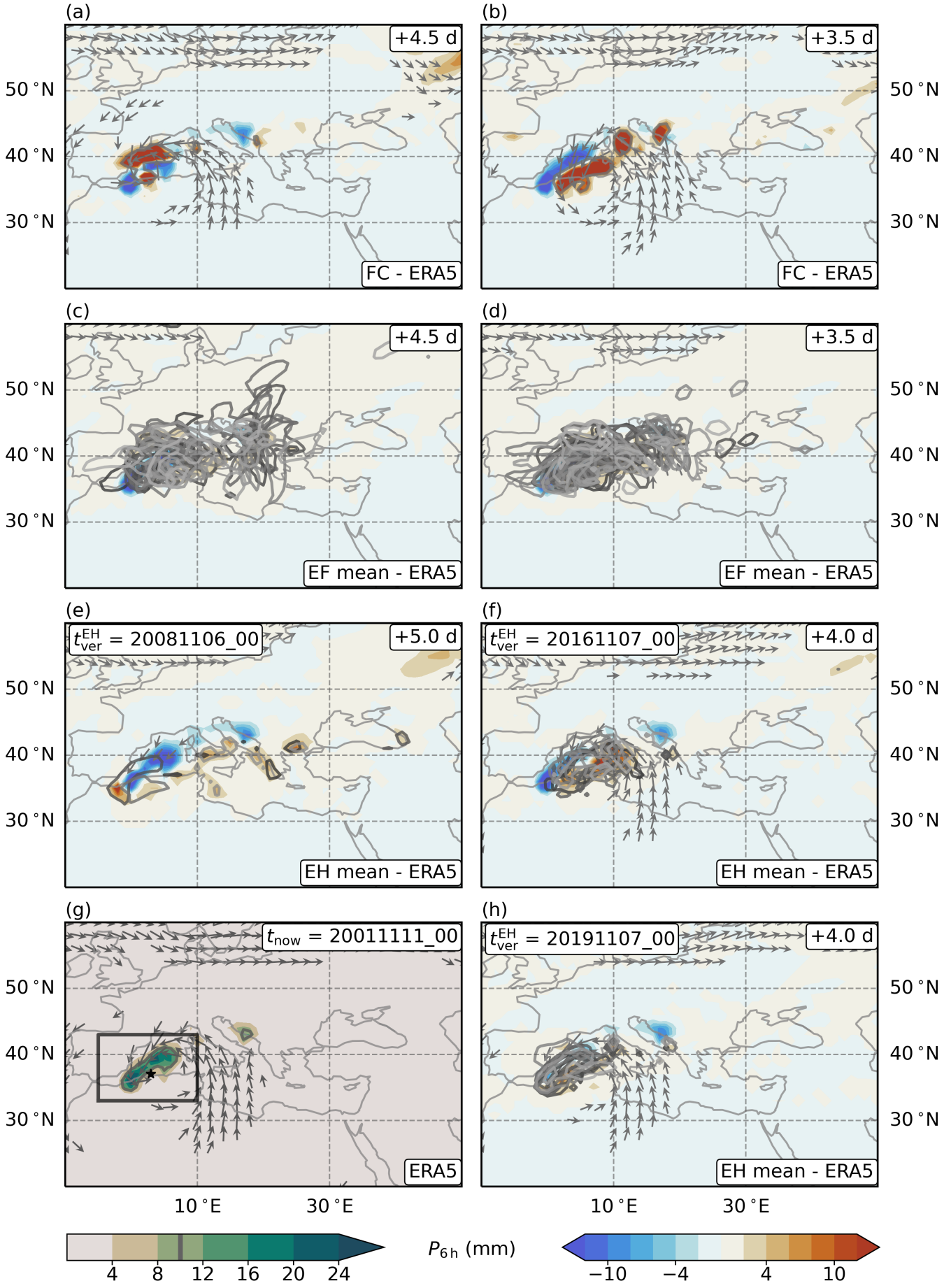


Figure B.20: Middle lead time forecasts of precipitation for 2001, in the middle of the most intense cyclone stage. Similar to Figure B.19, but for 6 h accumulated precipitation.

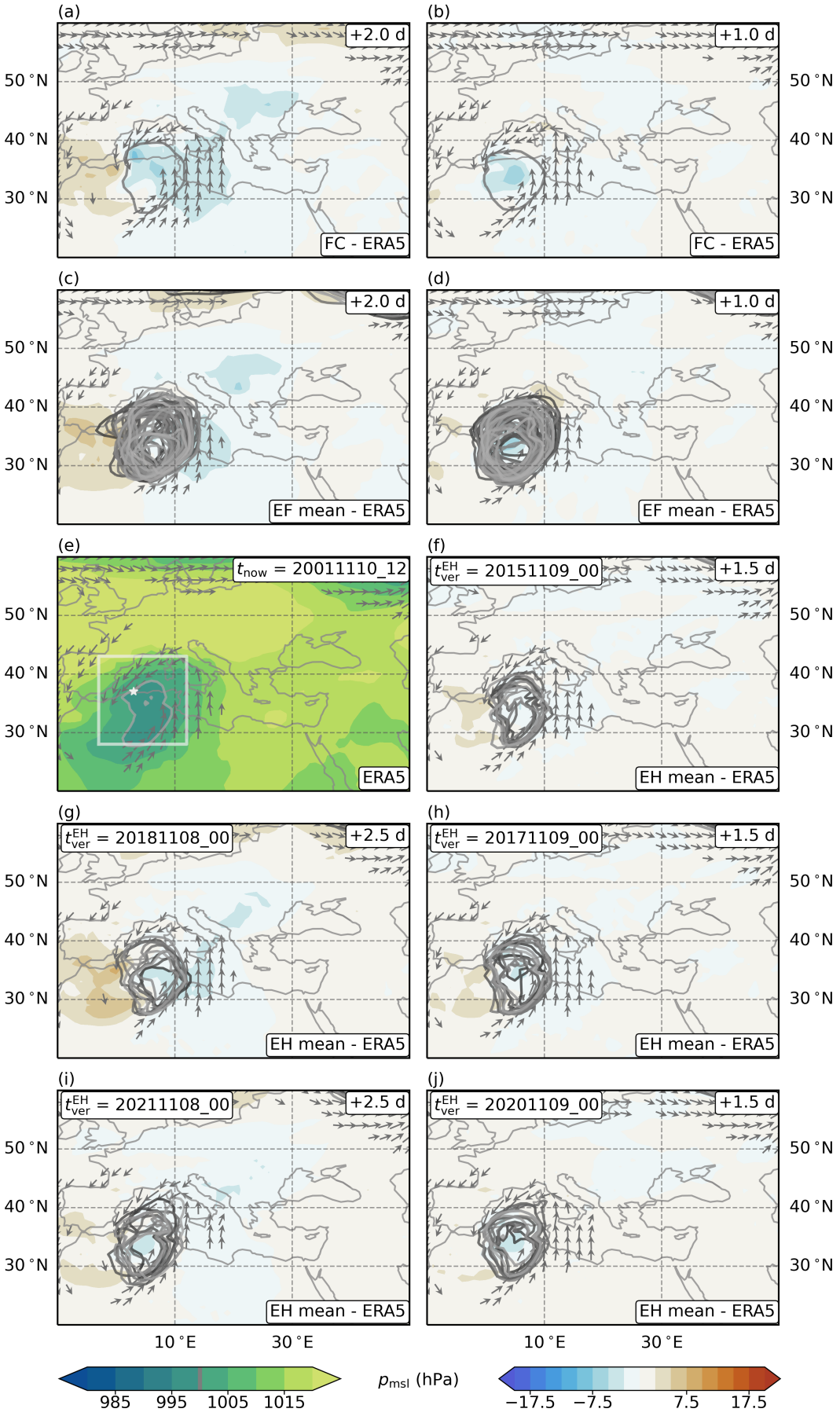


Figure B.21: Short lead time forecasts of pressure for 2001, at the beginning of the most intense cyclone stage. Similar to Figure 3.14, but for a different reference time.

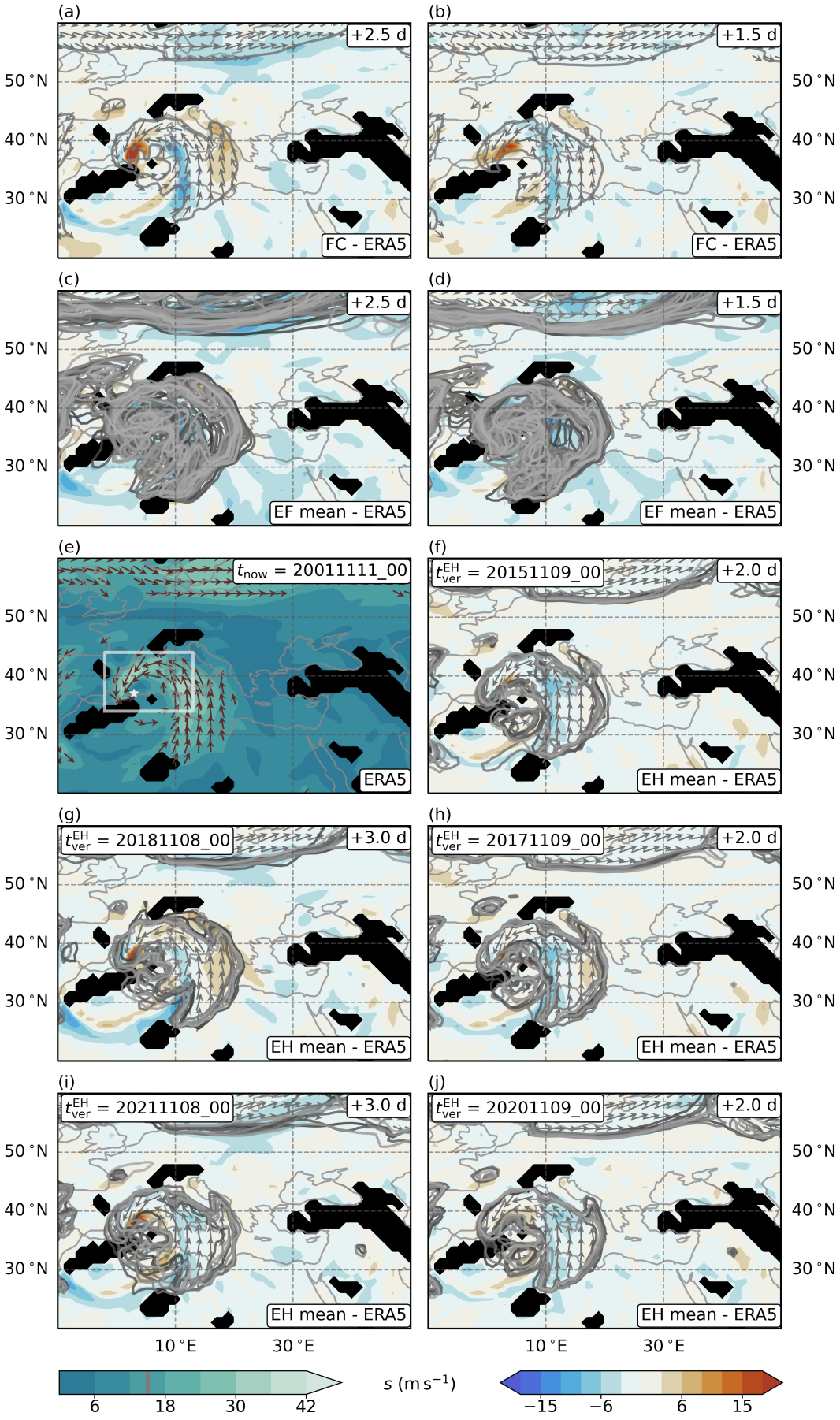


Figure B.22: Short lead time forecasts of wind speed for 2001, in the middle of the most intense cyclone stage. Similar to Figure B.21, but for a different reference time, and for horizontal wind speed on 900 hPa.

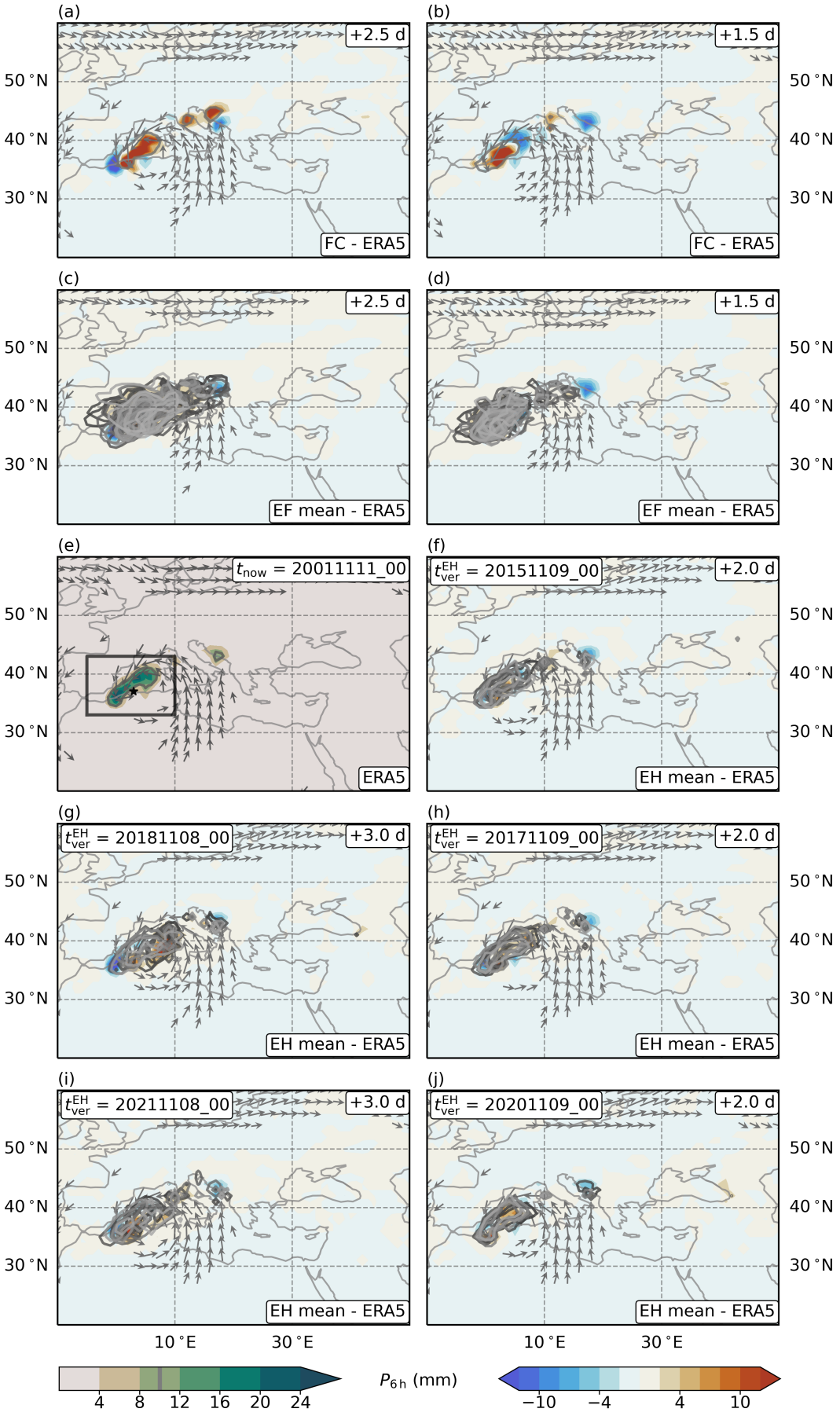


Figure B.23: Short lead time forecasts of precipitation for 2001, in the middle of the most intense cyclone stage. Similar to Figure B.22, but for 6 h accumulated precipitation.

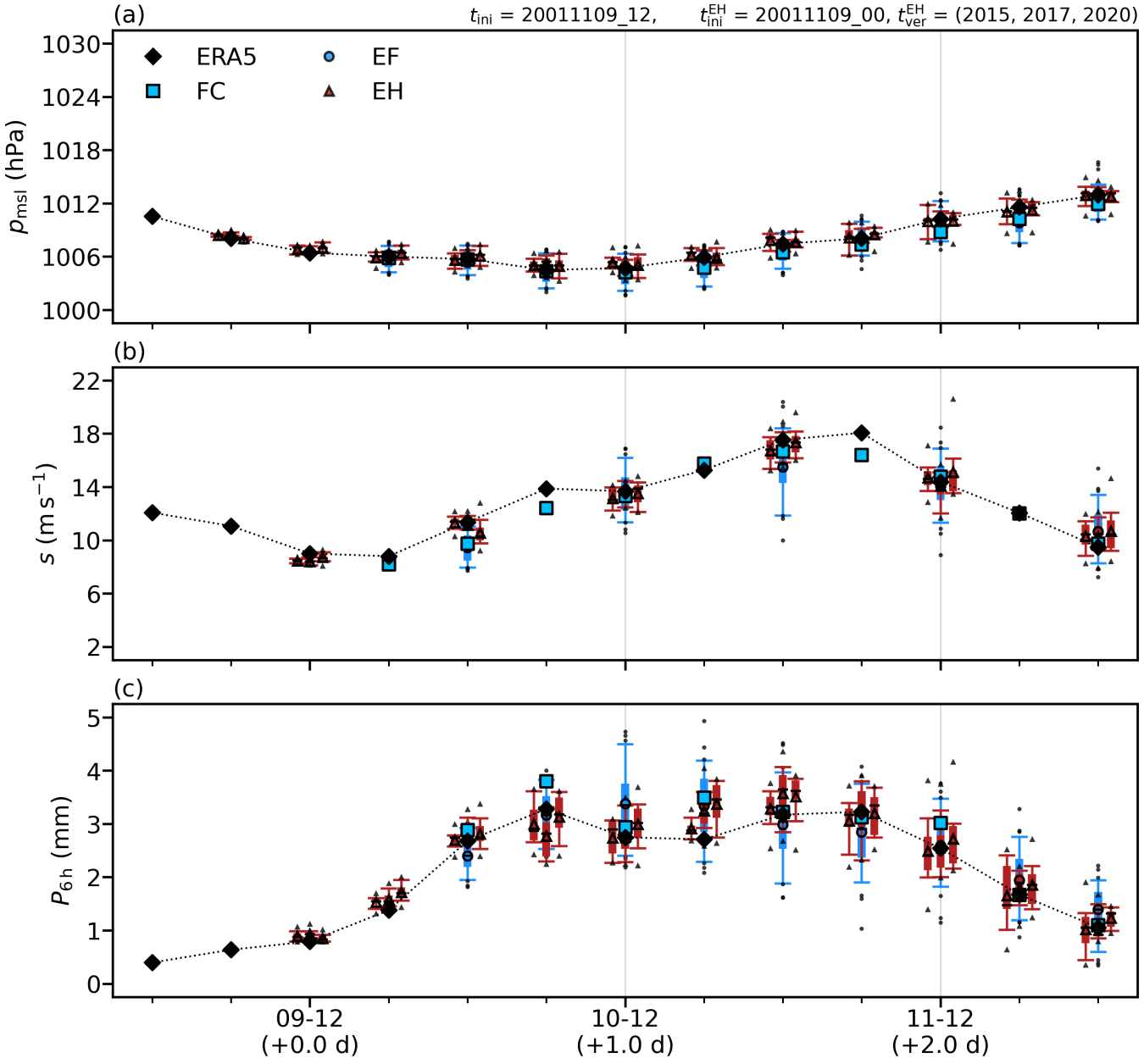


Figure B.24: Time series of forecasts at a short lead time for 2001. Similar to Figure 3.17. The forecasts were initialised on 9. November 2001, the ensemble hindcasts (EH) at 00 UTC, and both the deterministic forecast (FC) and the ensemble forecast (EF) at 12 UTC.

B.4 Apulia September 2006

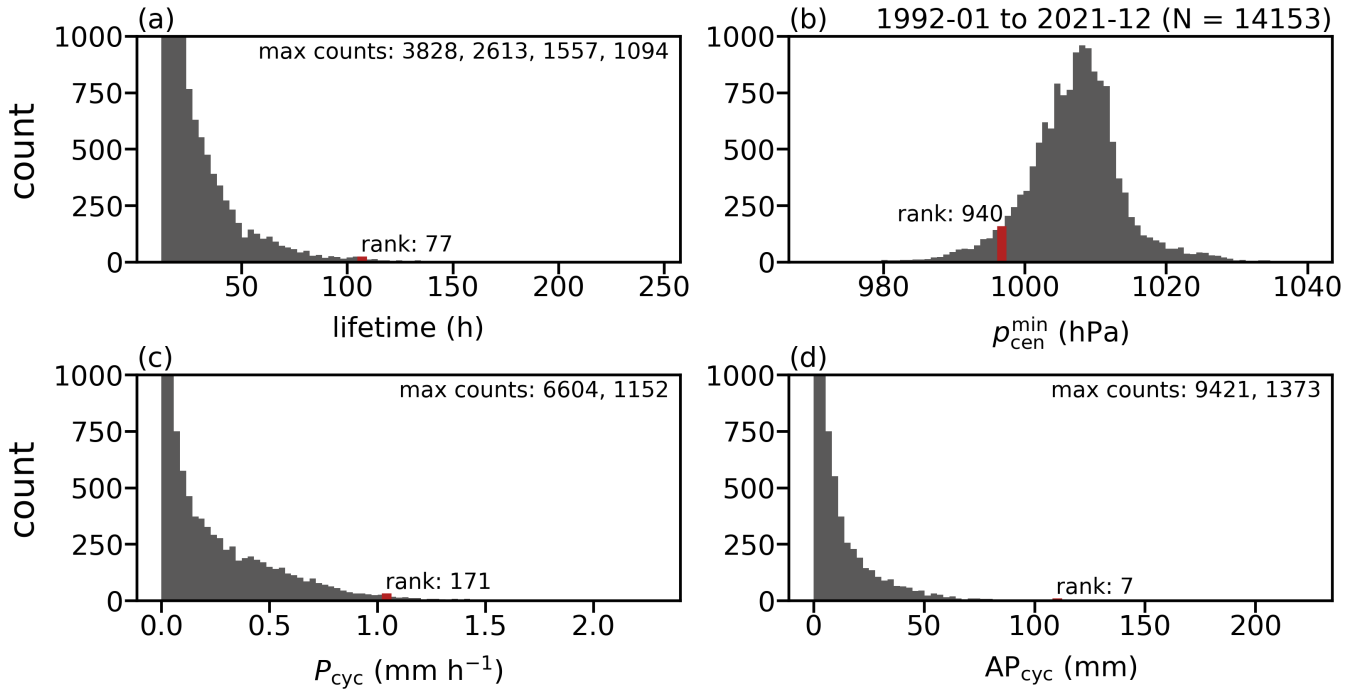


Figure B.25: Statistics for the September 2006 cyclone. Similar to Figure B.16, but for a different reference cyclone.

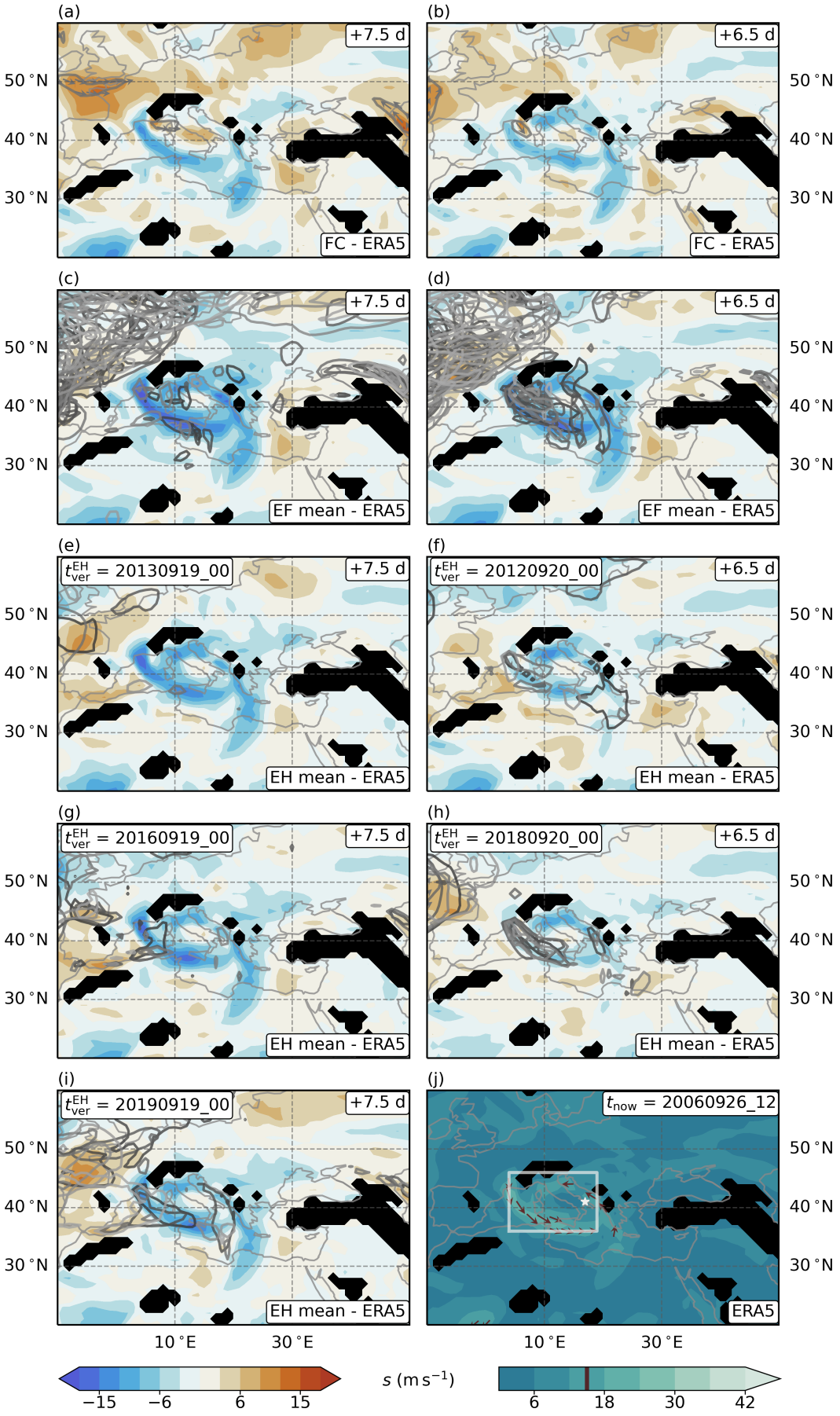


Figure B.26: Long lead time forecasts of wind speed for 2006, in the middle of the most intense cyclone stage. Similar to Figure 3.20, but for horizontal wind speed on 900 hPa.

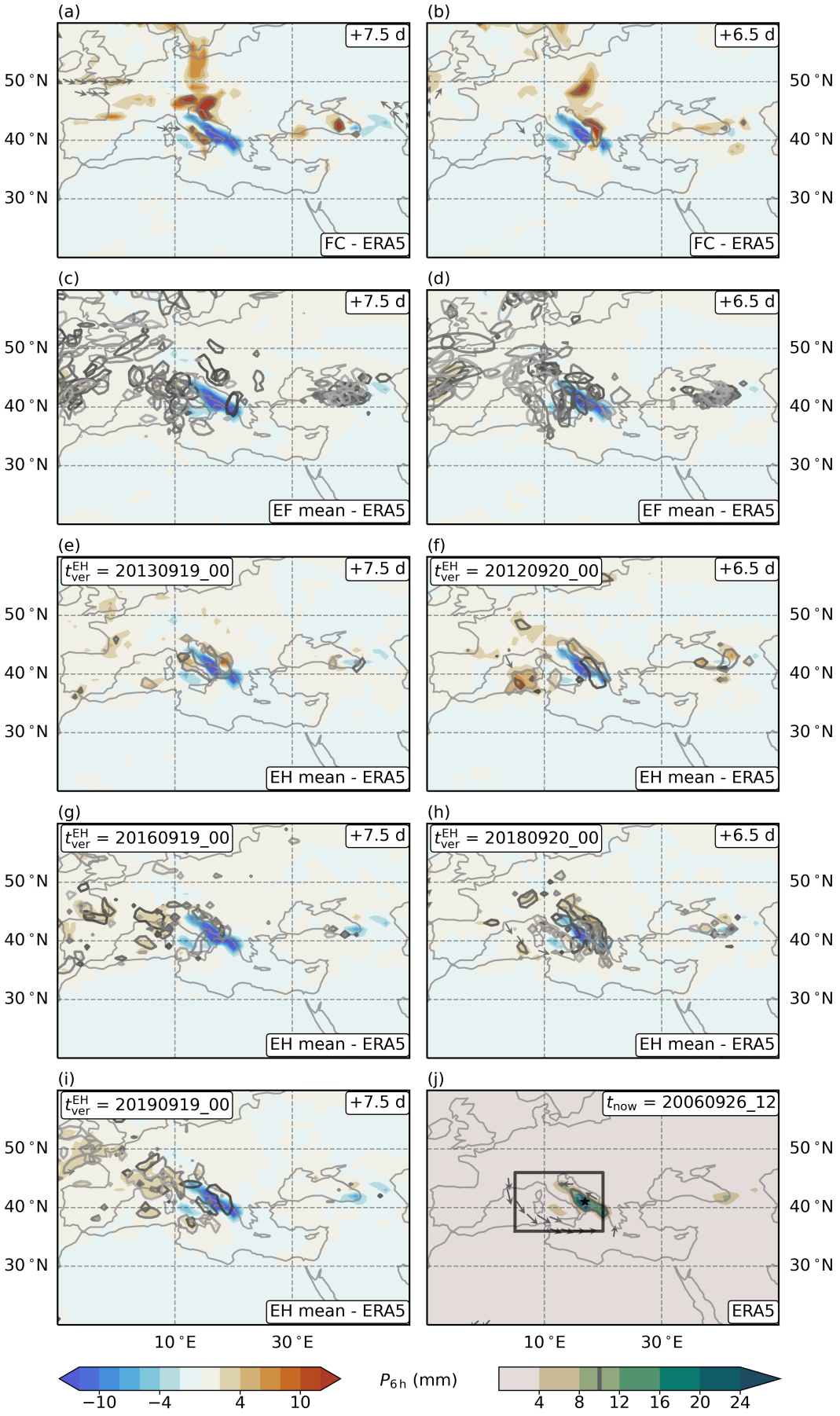


Figure B.27: Long lead time forecasts of precipitation for 2006, in the middle of the most intense cyclone stage. Similar to Figure B.26, but for 6 h accumulated precipitation.

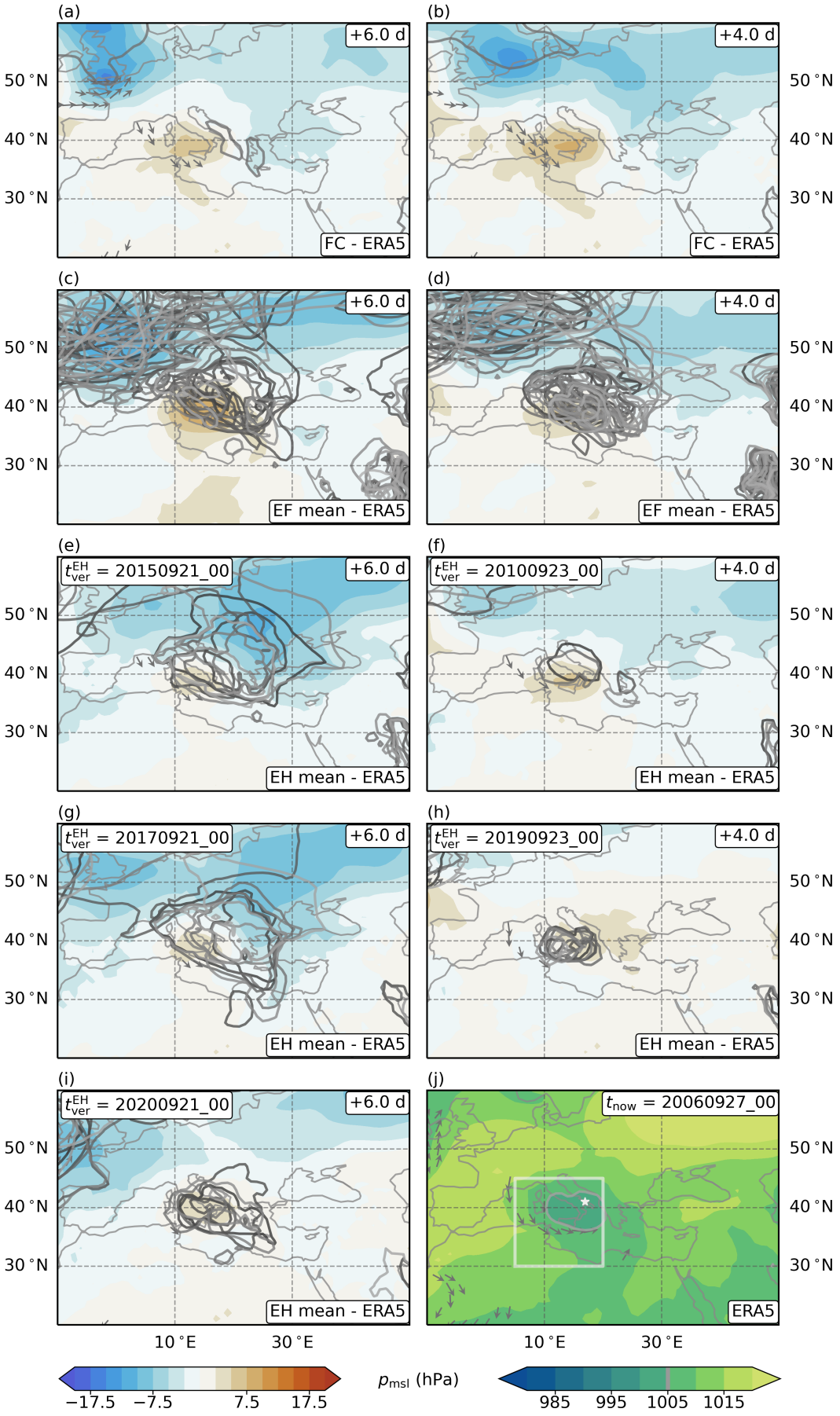


Figure B.28: Middle lead time forecasts of pressure for 2006, at the end of the most intense cyclone stage. Similar to Figure 3.20, but for a different reference time.

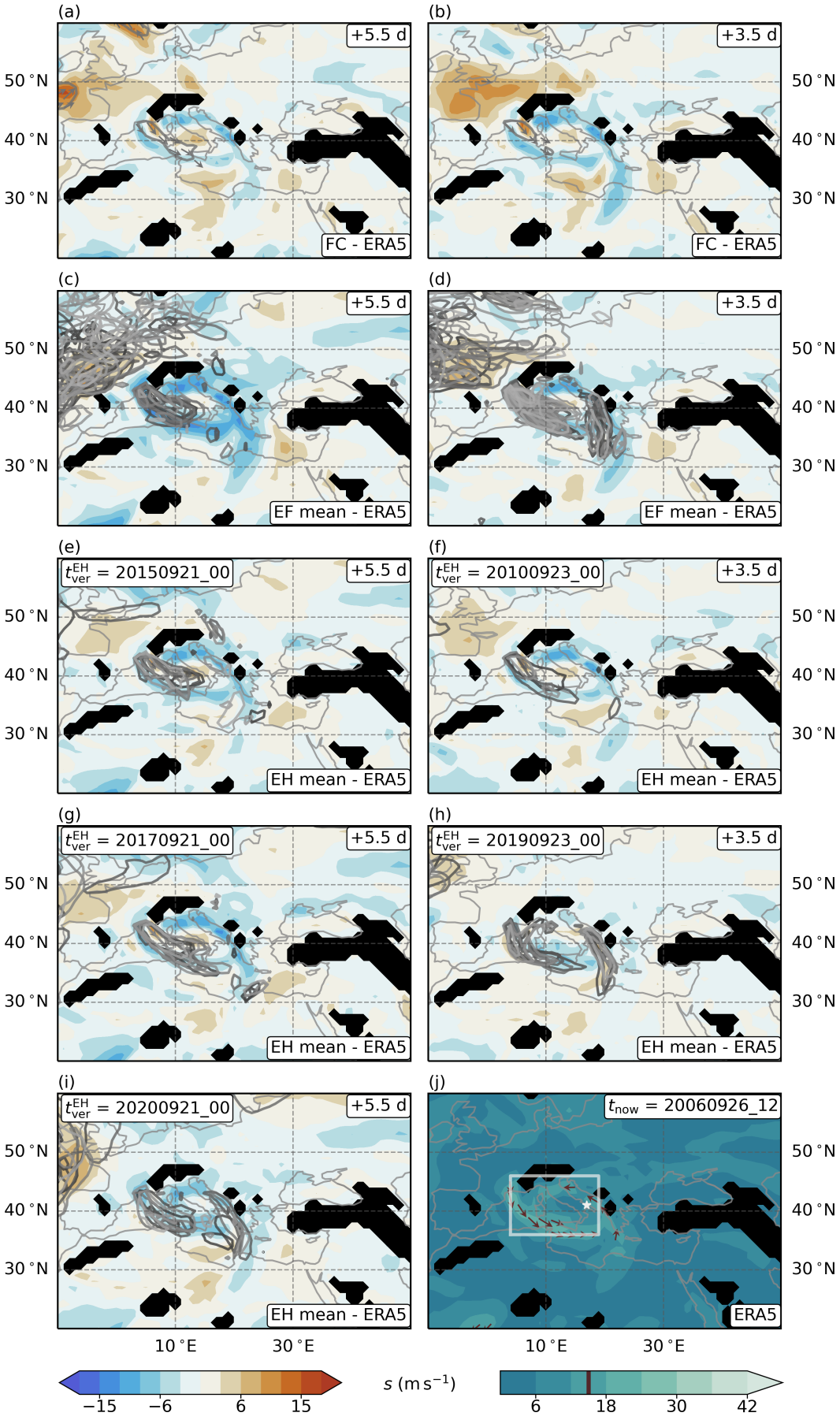


Figure B.29: Middle lead time forecasts of wind speed for 2006, in the middle of the most intense cyclone stage. Similar to Figure B.28, but for a different reference time, and for horizontal wind speed on 900 hPa.

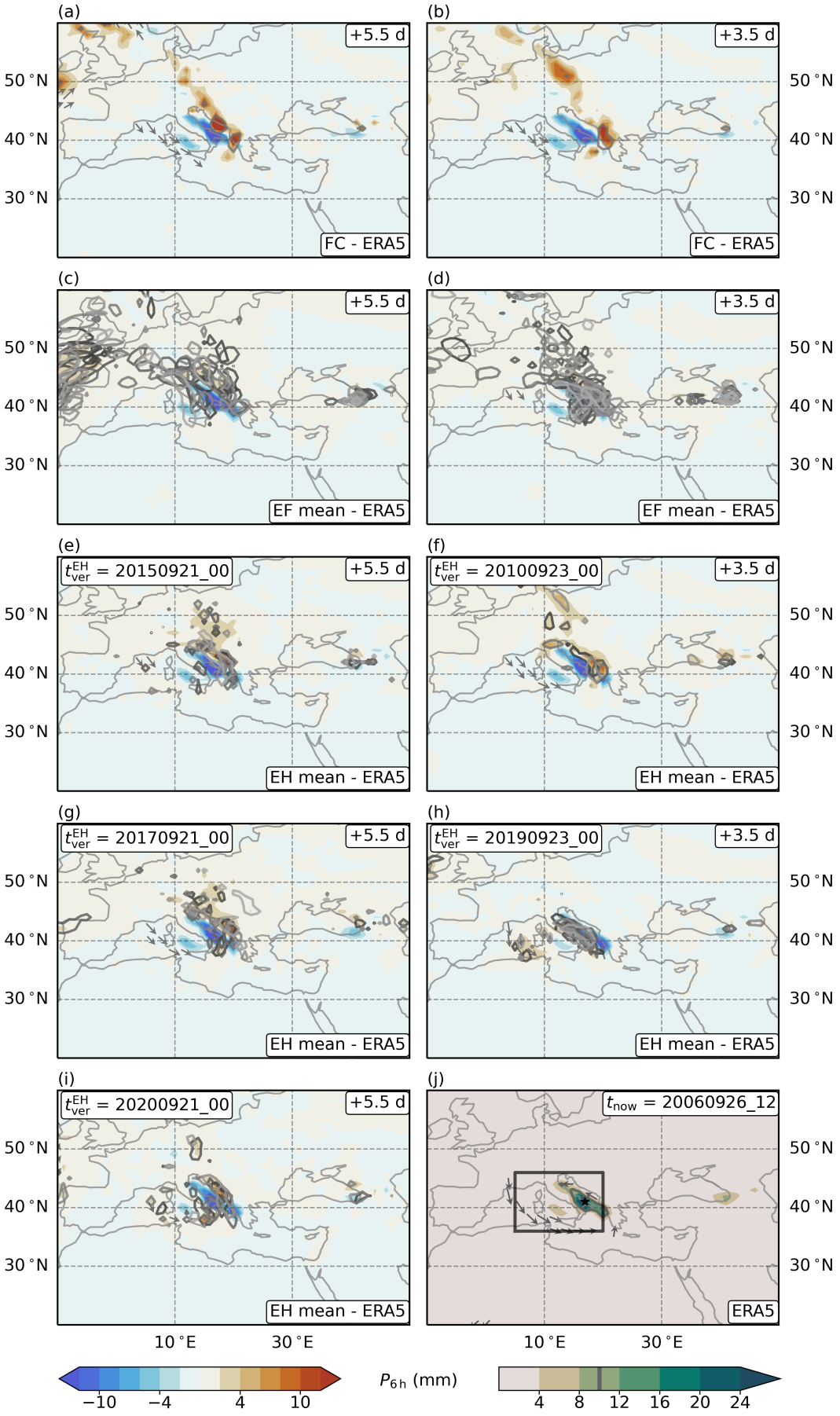


Figure B.30: Middle lead time forecasts of precipitation for 2006, in the middle of the most intense cyclone stage. Similar to Figure B.29, but for 6 h accumulated precipitation.

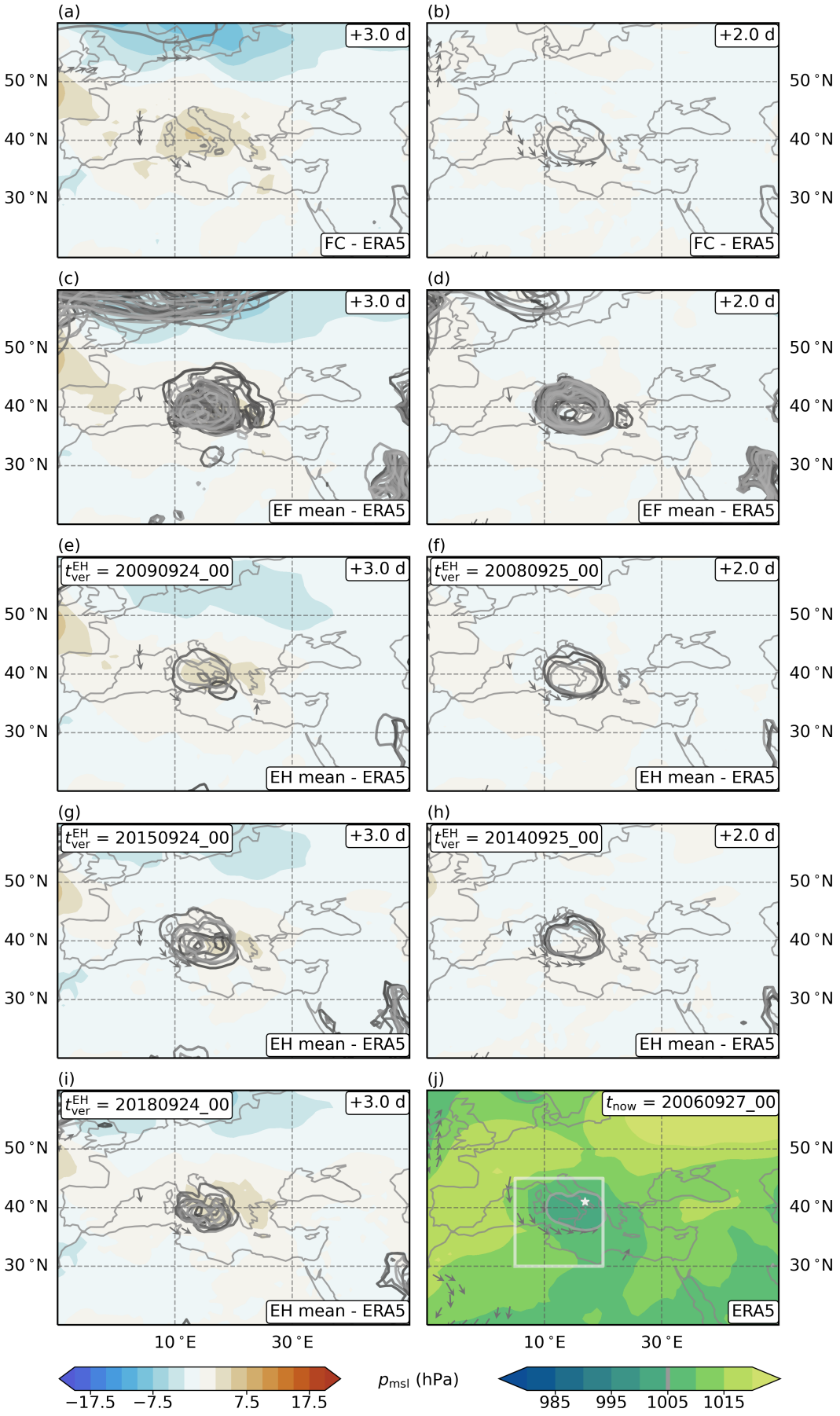


Figure B.31: Short lead time forecasts of pressure for 2006, at the end of the most intense cyclone stage. Similar to Figure 3.20, but for a different reference time.

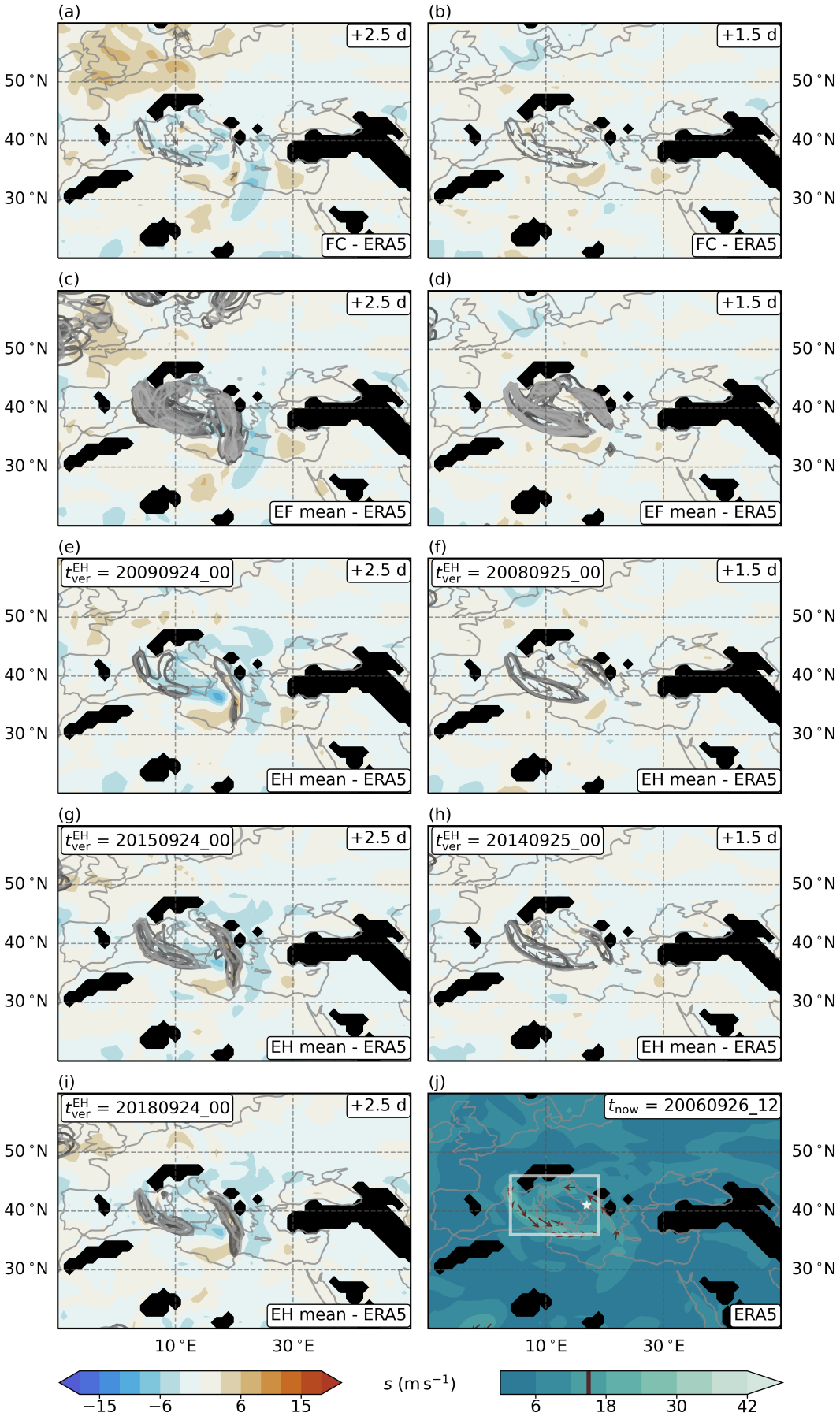


Figure B.32: Short lead time forecasts of wind speed for 2006, in the middle of the most intense cyclone stage. Similar to Figure B.31, but for a different reference time, and for horizontal wind speed on 900 hPa.

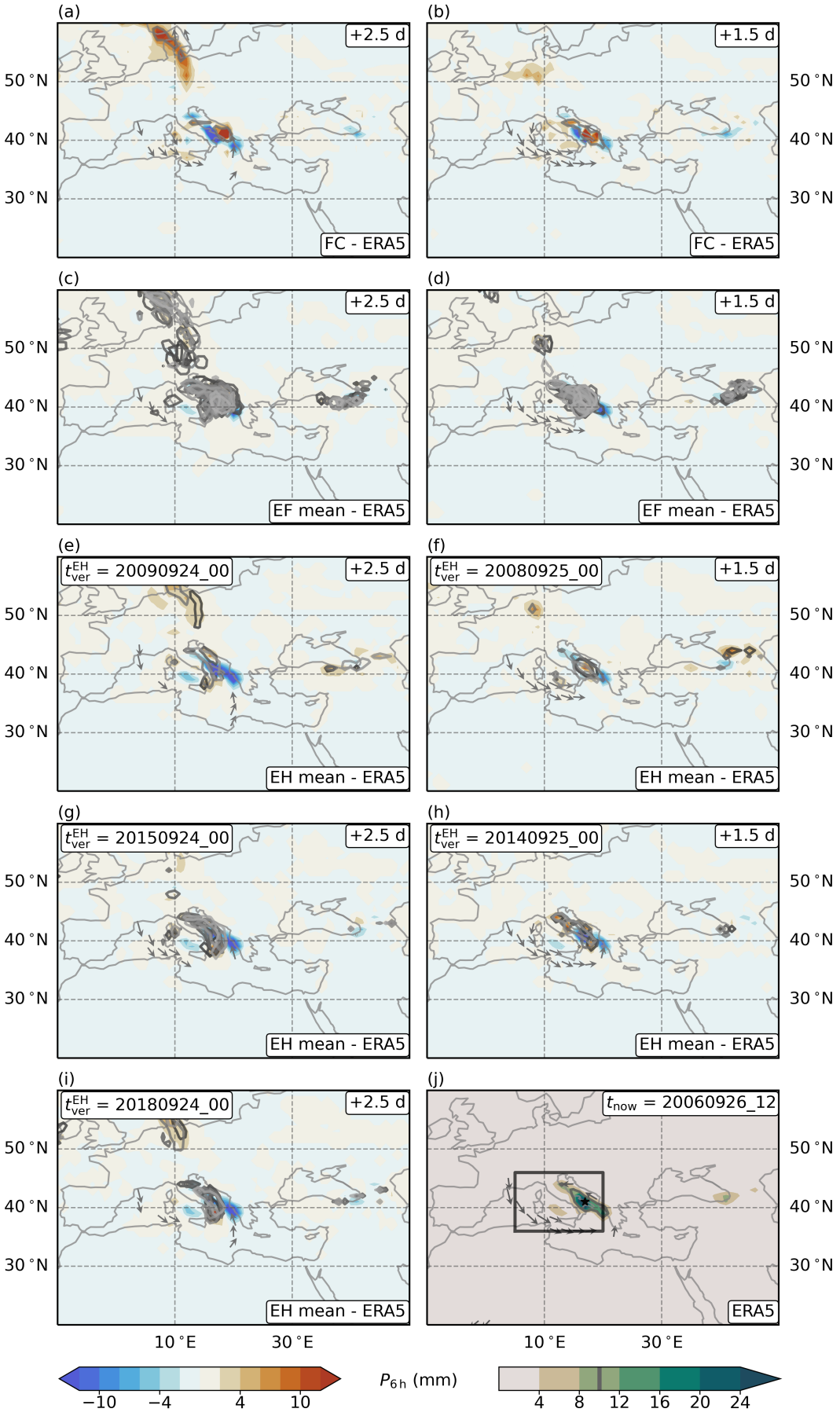


Figure B.33: Short lead time forecasts of precipitation for 2006, in the middle of the most intense cyclone stage. Similar to Figure B.32, but for 6 h accumulated precipitation.

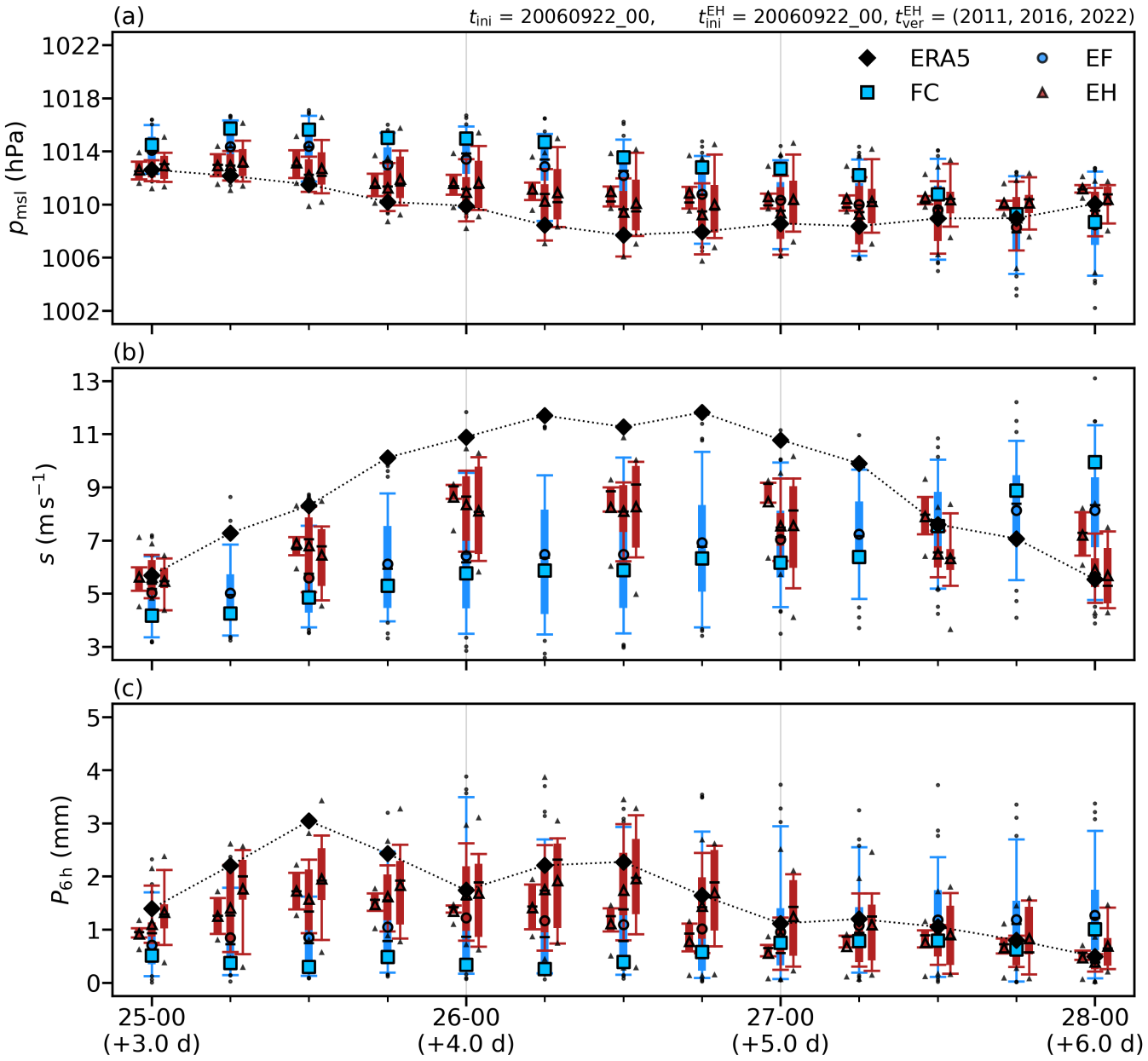


Figure B.34: Time series of forecasts at a middle lead time for 2006. Similar to Figure 3.21. All forecasts were initialised on 22. September 2006, at 00 UTC.

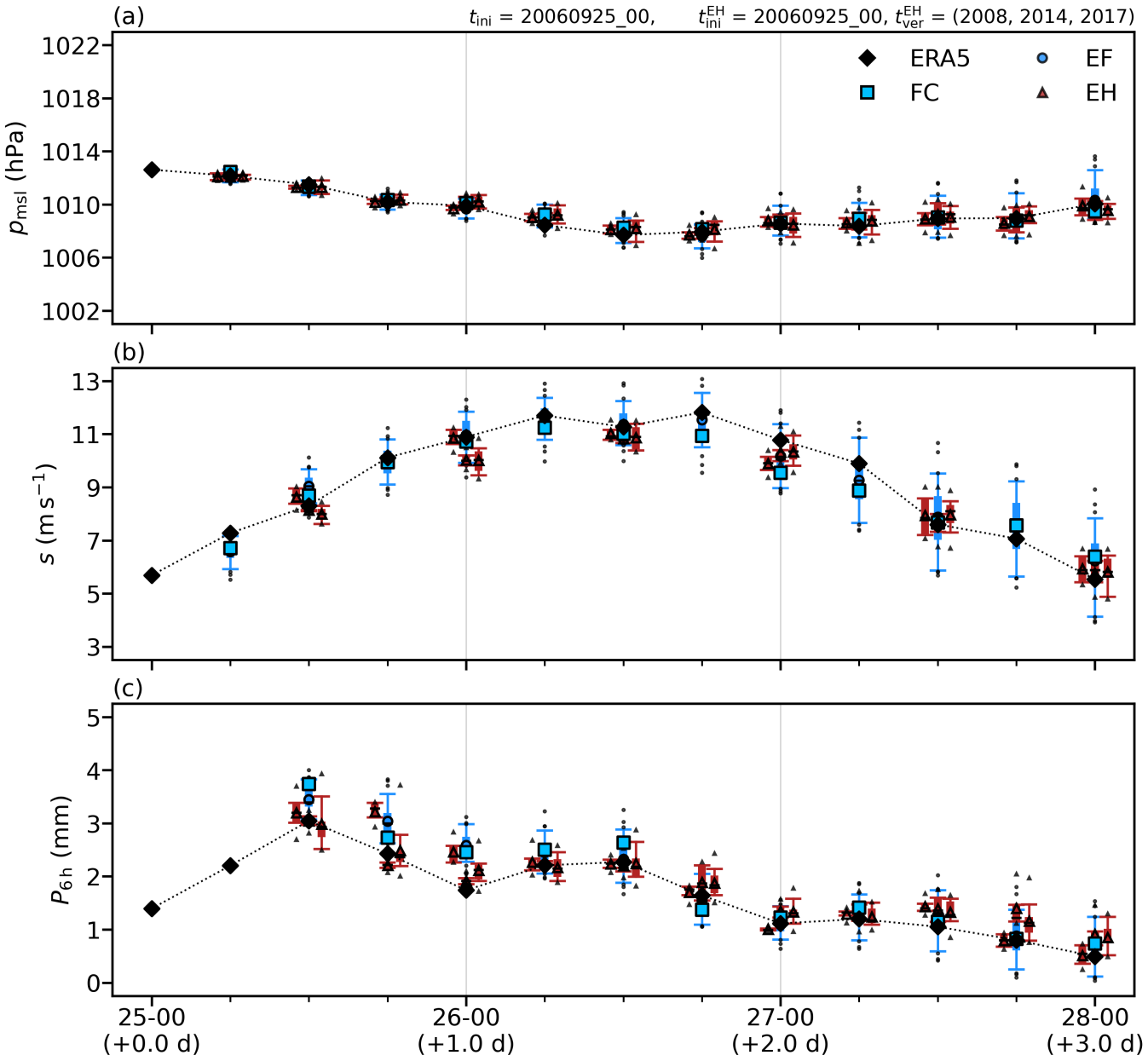


Figure B.35: Time series of forecasts at a short lead time for 2006. Similar to Figure B.34. All forecasts were initialised on 25. September 2006, at 00 UTC.

B.5 Hyères November 2011

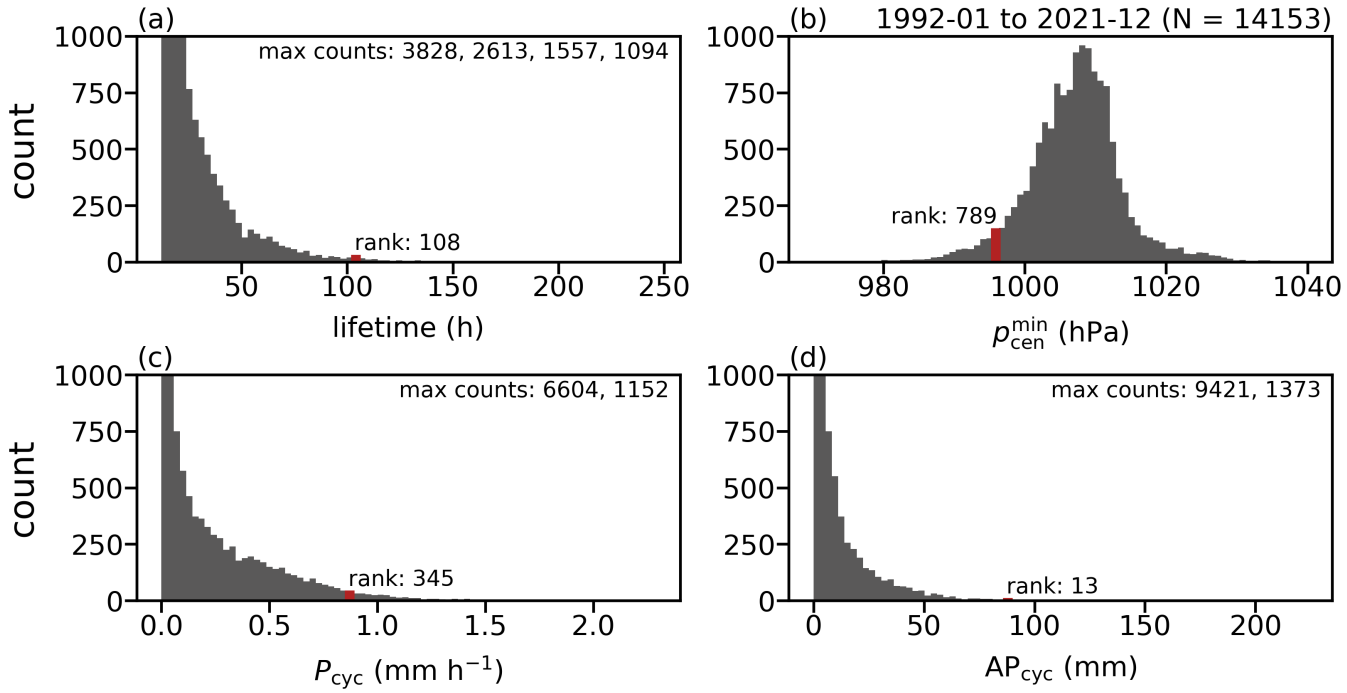


Figure B.36: Statistics for the November 2011 cyclone. Similar to Figure B.25, but for a different reference cyclone.

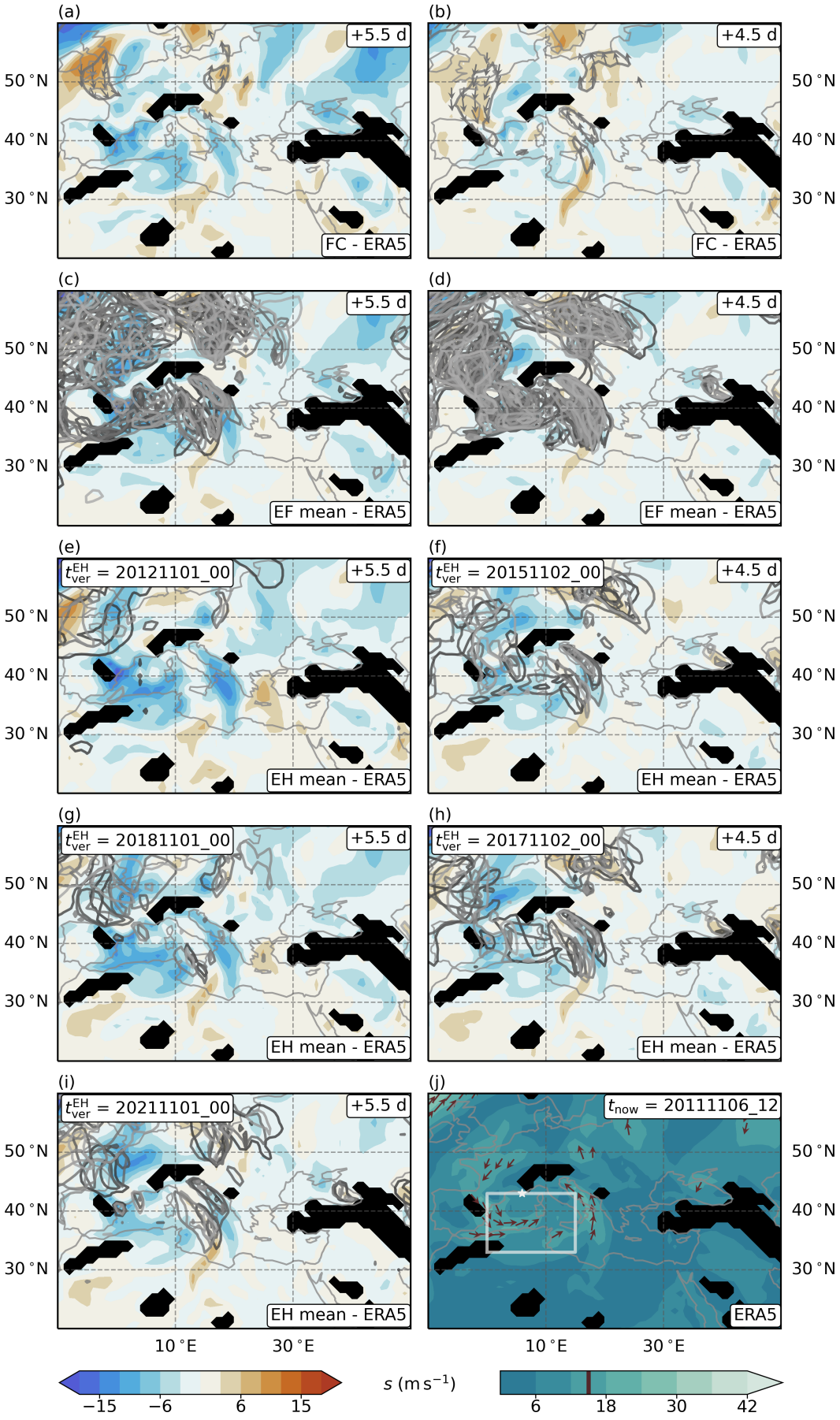


Figure B.37: Long lead time forecasts of wind speed for 2011, at the end of the most intense cyclone stage. Similar to Figure 3.24, but for a different reference time, and for horizontal wind speed on 900 hPa.

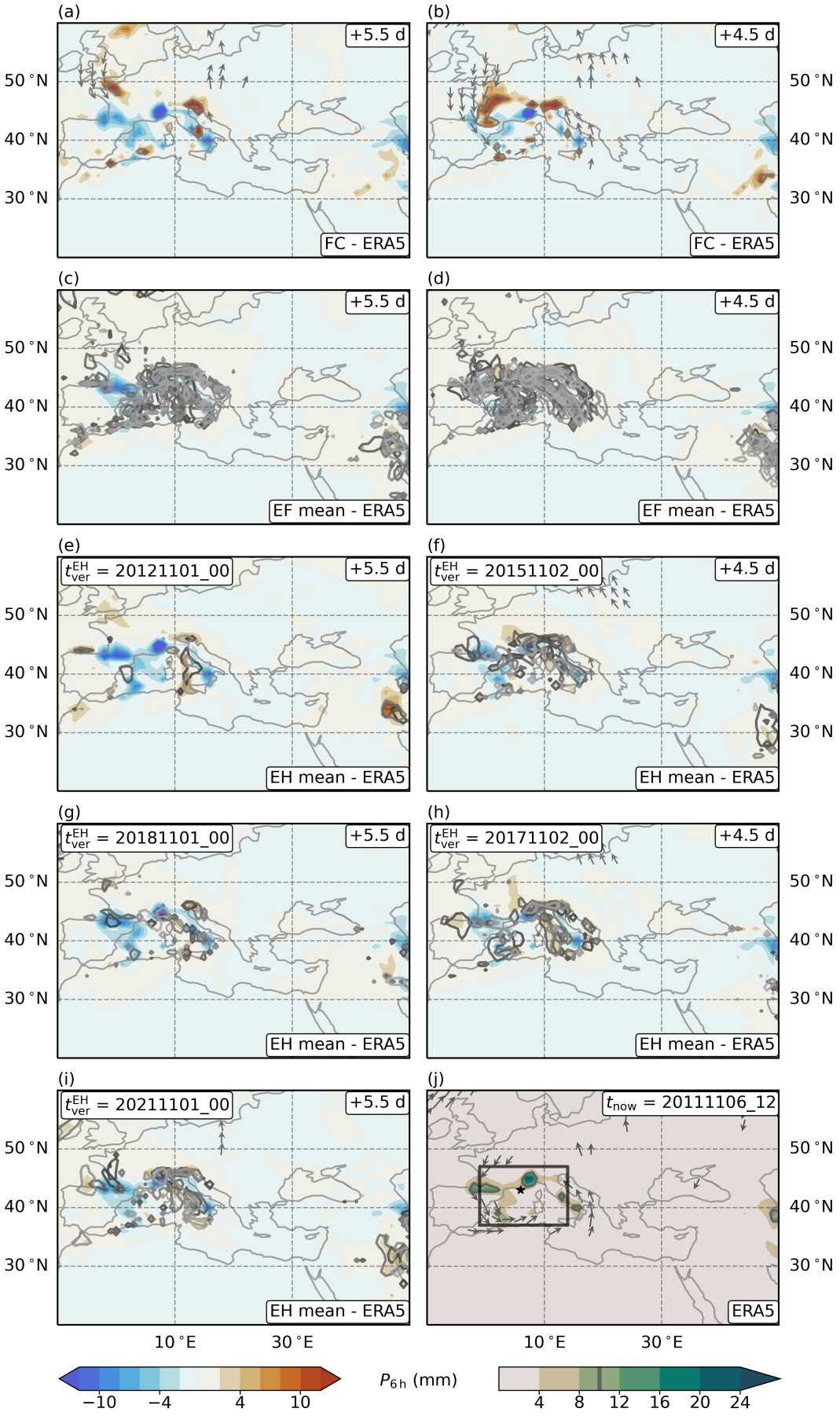


Figure B.38: Long lead time forecasts of precipitation for 2011, at the end of the most intense cyclone stage. Similar to Figure B.37, but for 6 h accumulated precipitation.

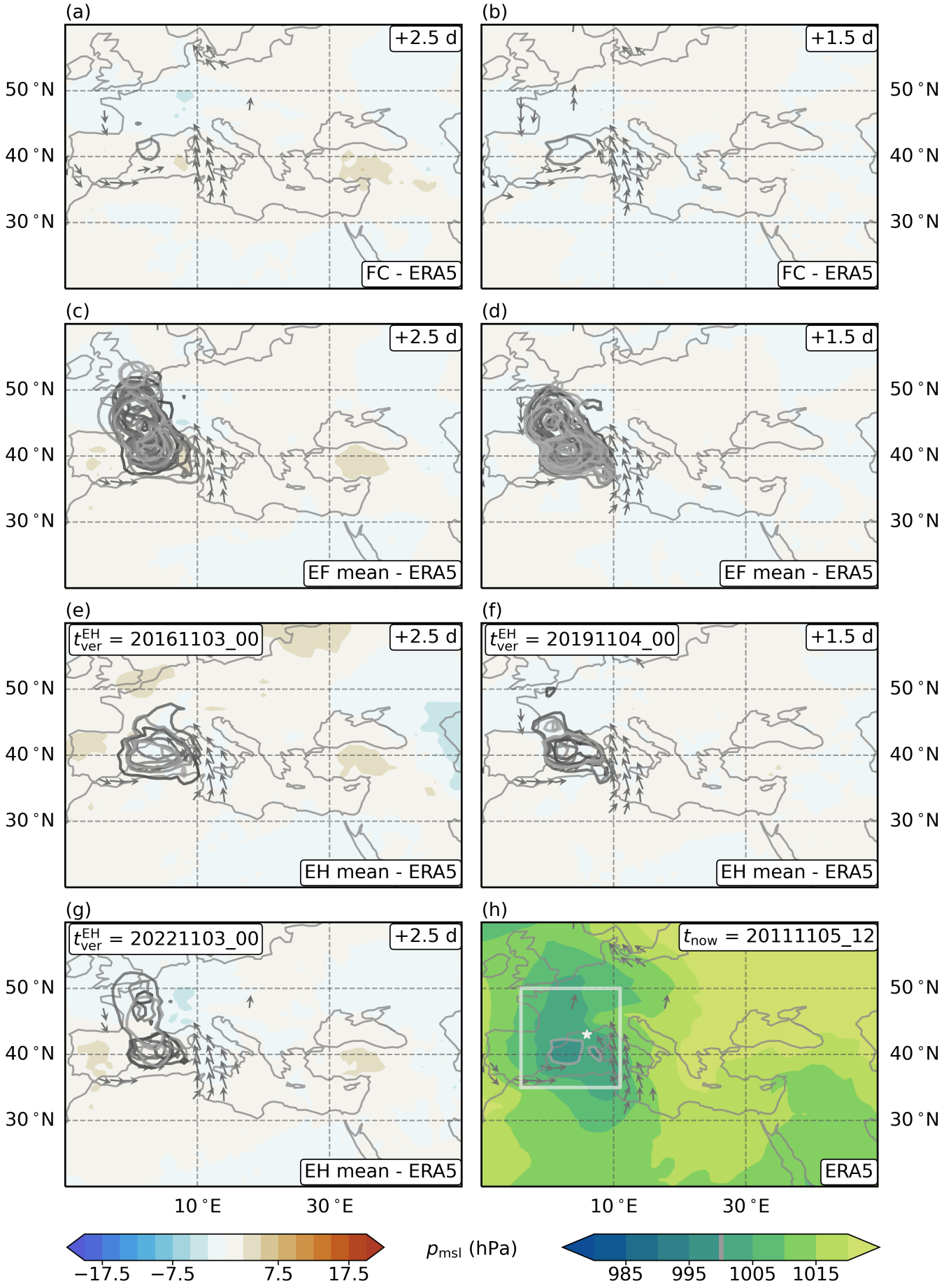


Figure B.39: Middle lead time forecasts of pressure for 2011, at the beginning of the most intense cyclone stage. Similar to Figure 3.24.

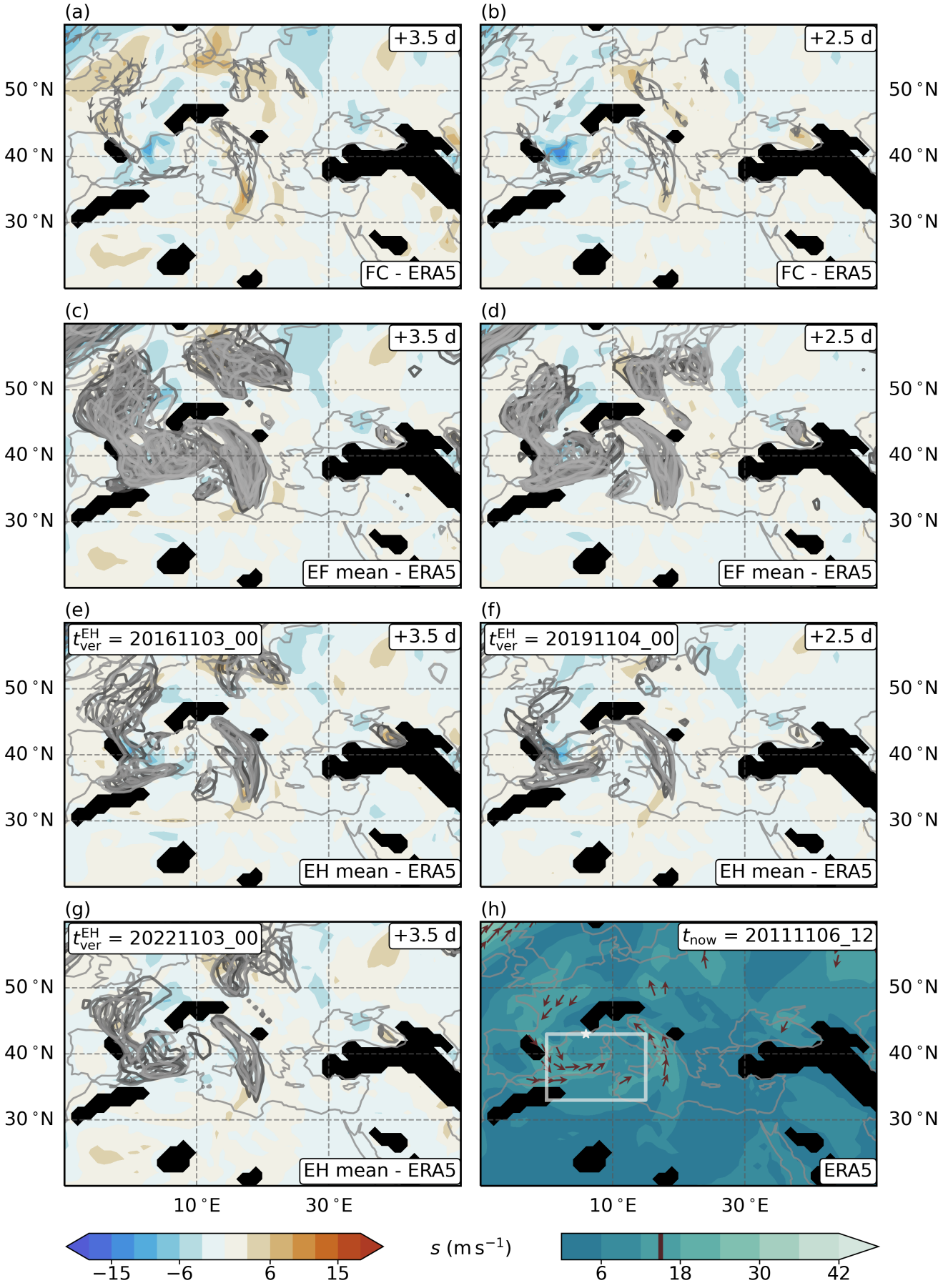


Figure B.40: Middle lead time forecasts of wind speed for 2011, at the end of the most intense cyclone stage. Similar to Figure B.39, but for a different reference time, and for horizontal wind speed on 900 hPa.

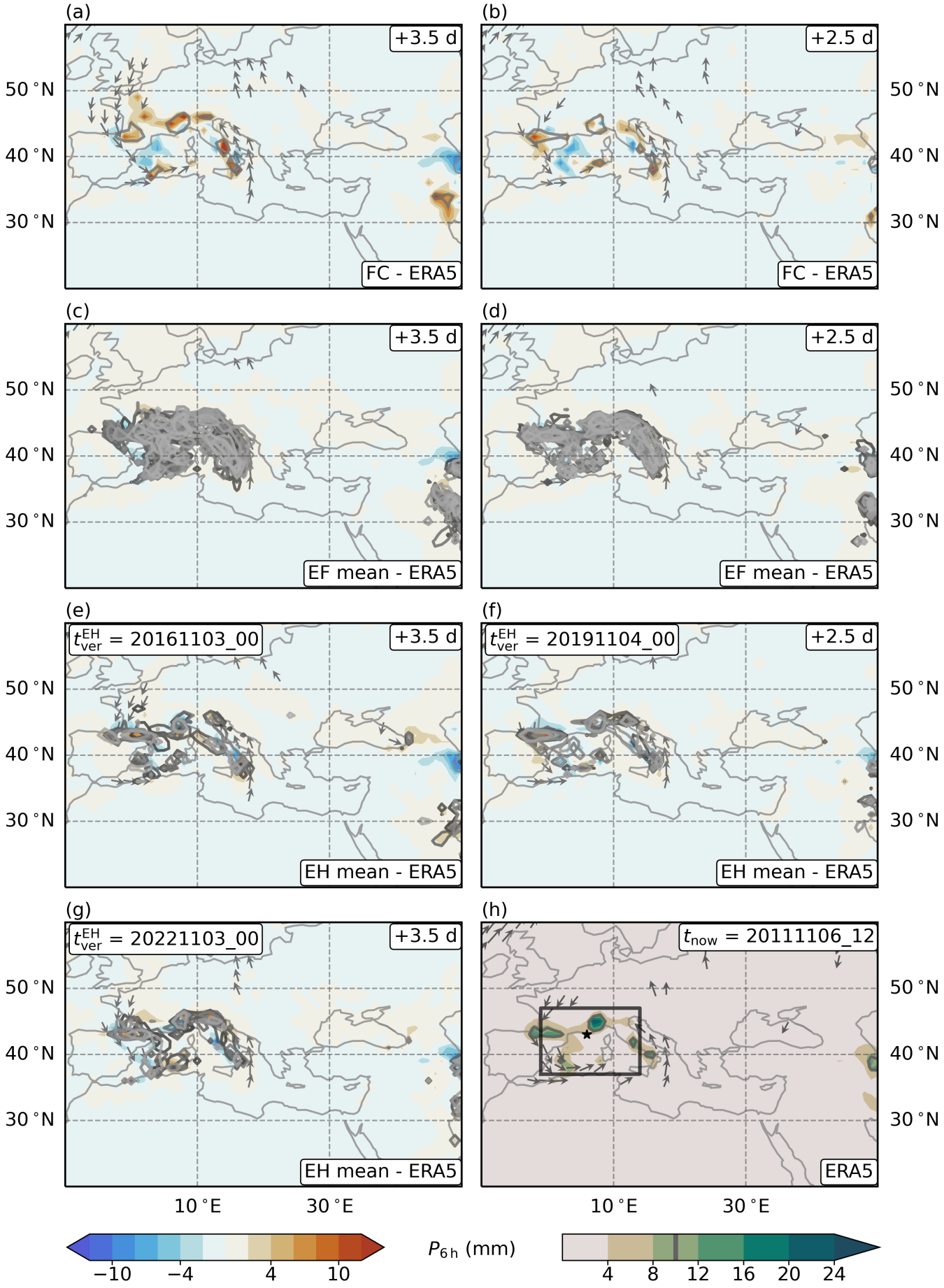


Figure B.41: Middle lead time forecasts of precipitation for 2011, at the end of the most intense cyclone stage. Similar to Figure B.40, but for 6 h accumulated precipitation.

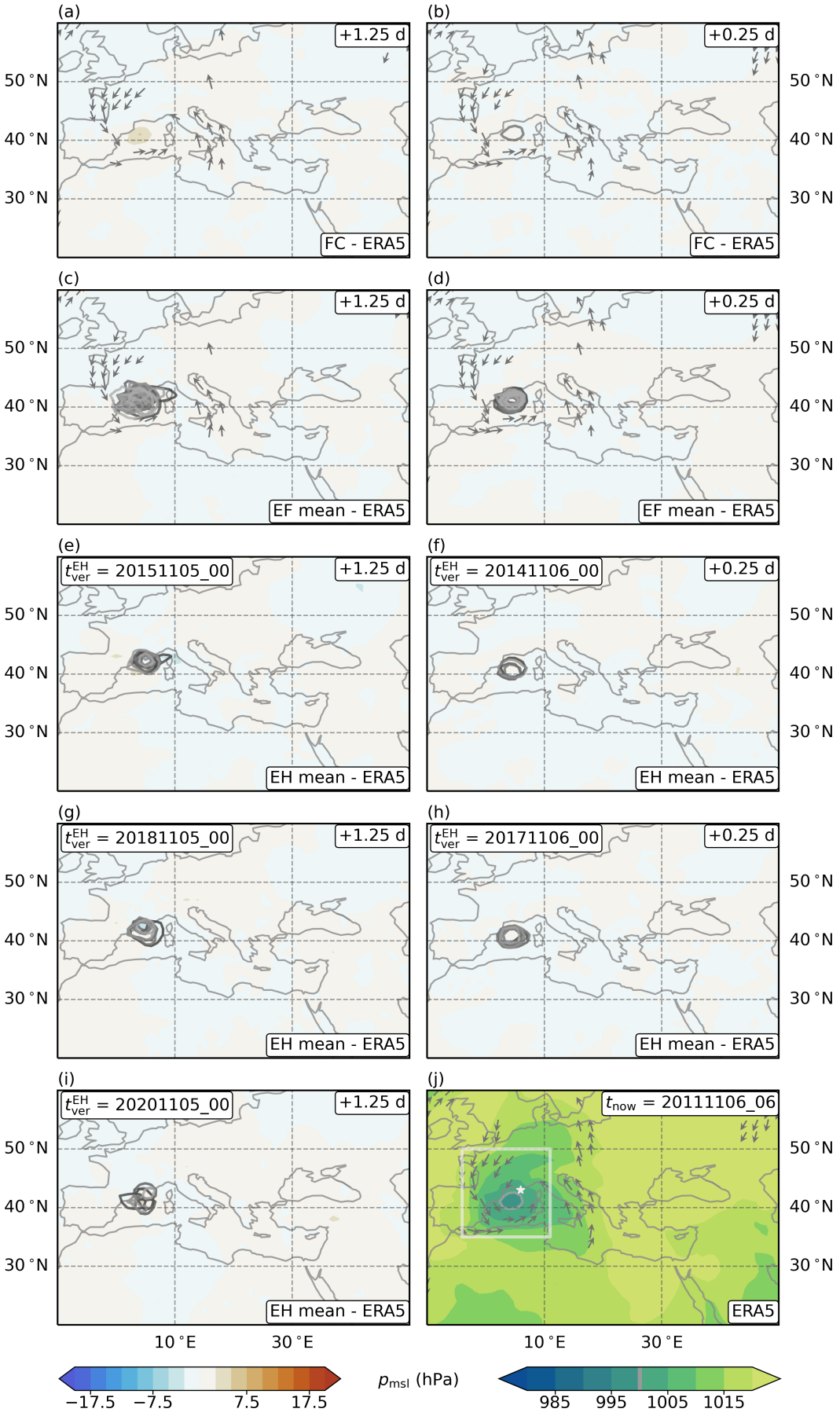


Figure B.42: Short lead time forecasts of pressure for 2011, towards the end of the most intense cyclone stage. Similar to Figure 3.24, but for a different reference time.

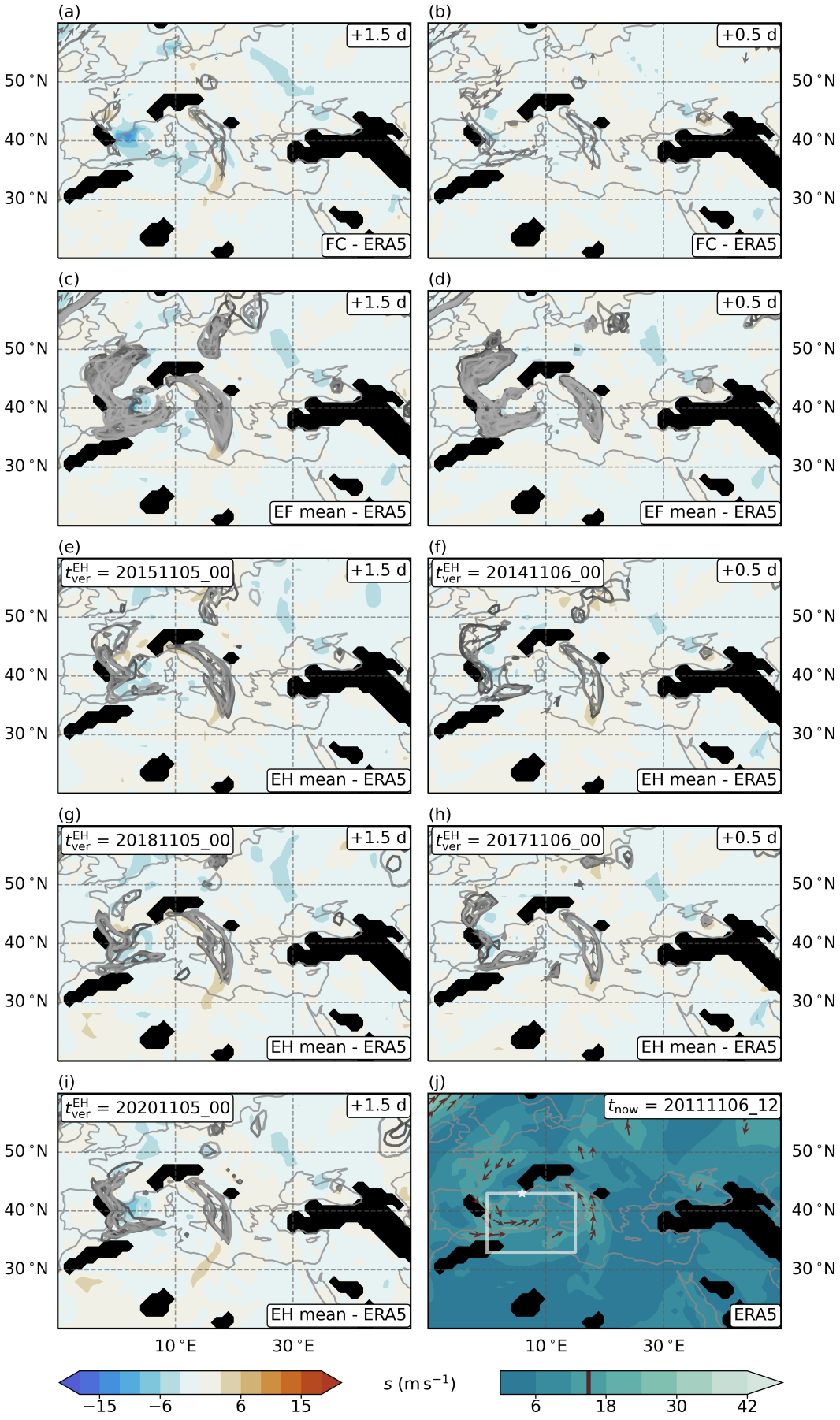


Figure B.43: Short lead time forecasts of wind speed for 2011, at the end of the most intense cyclone stage. Similar to Figure B.42, but for a different reference time, and for horizontal wind speed on 900 hPa.

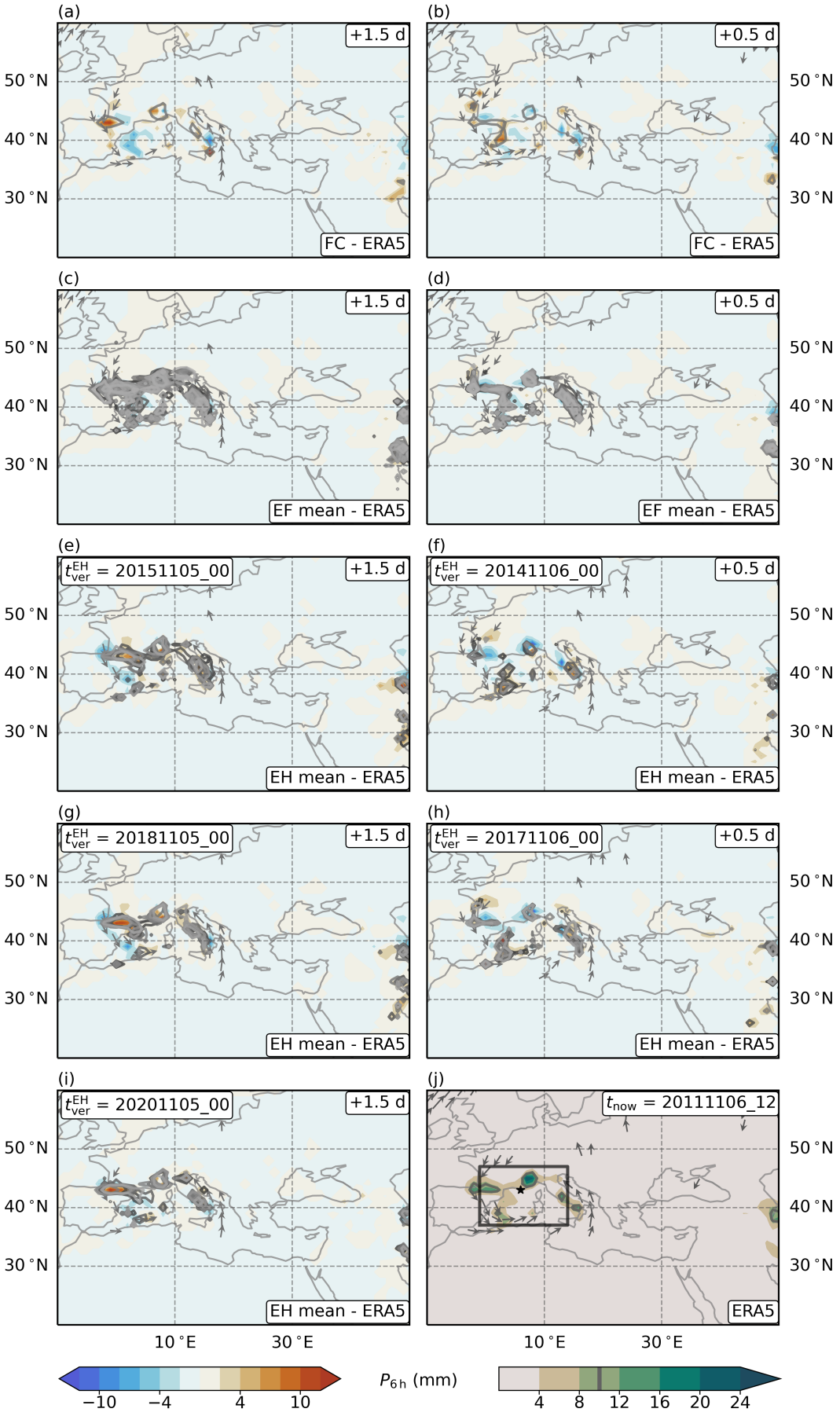


Figure B.44: Short lead time forecasts of precipitation for 2011, at the end of the most intense cyclone stage. Similar to Figure B.43, but for 6 h accumulated precipitation.

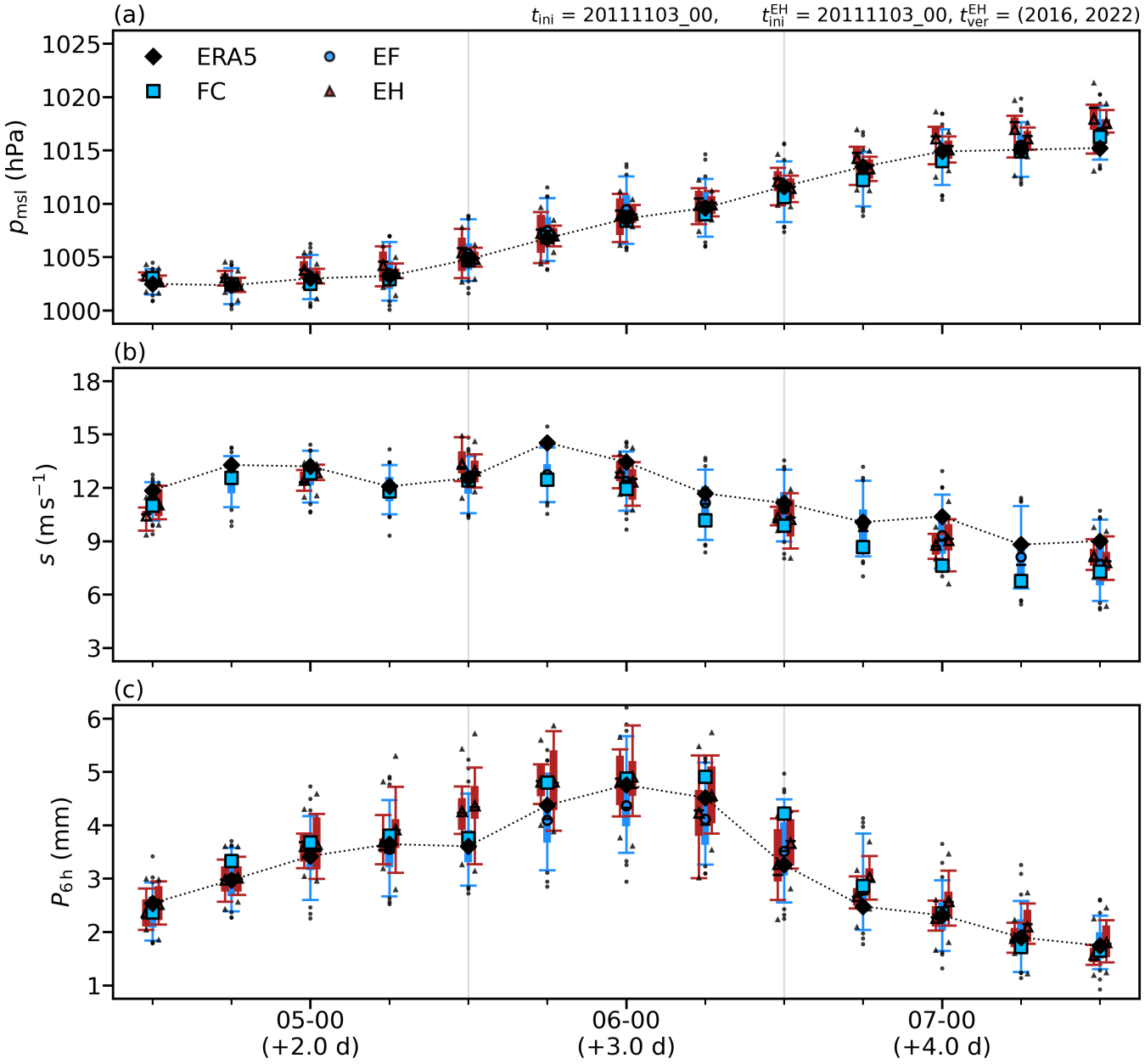


Figure B.45: Time series of forecasts at a middle lead time for 2011. Similar to Figure 3.25. All forecasts were initialised on 3. November 2011, at 00 UTC.

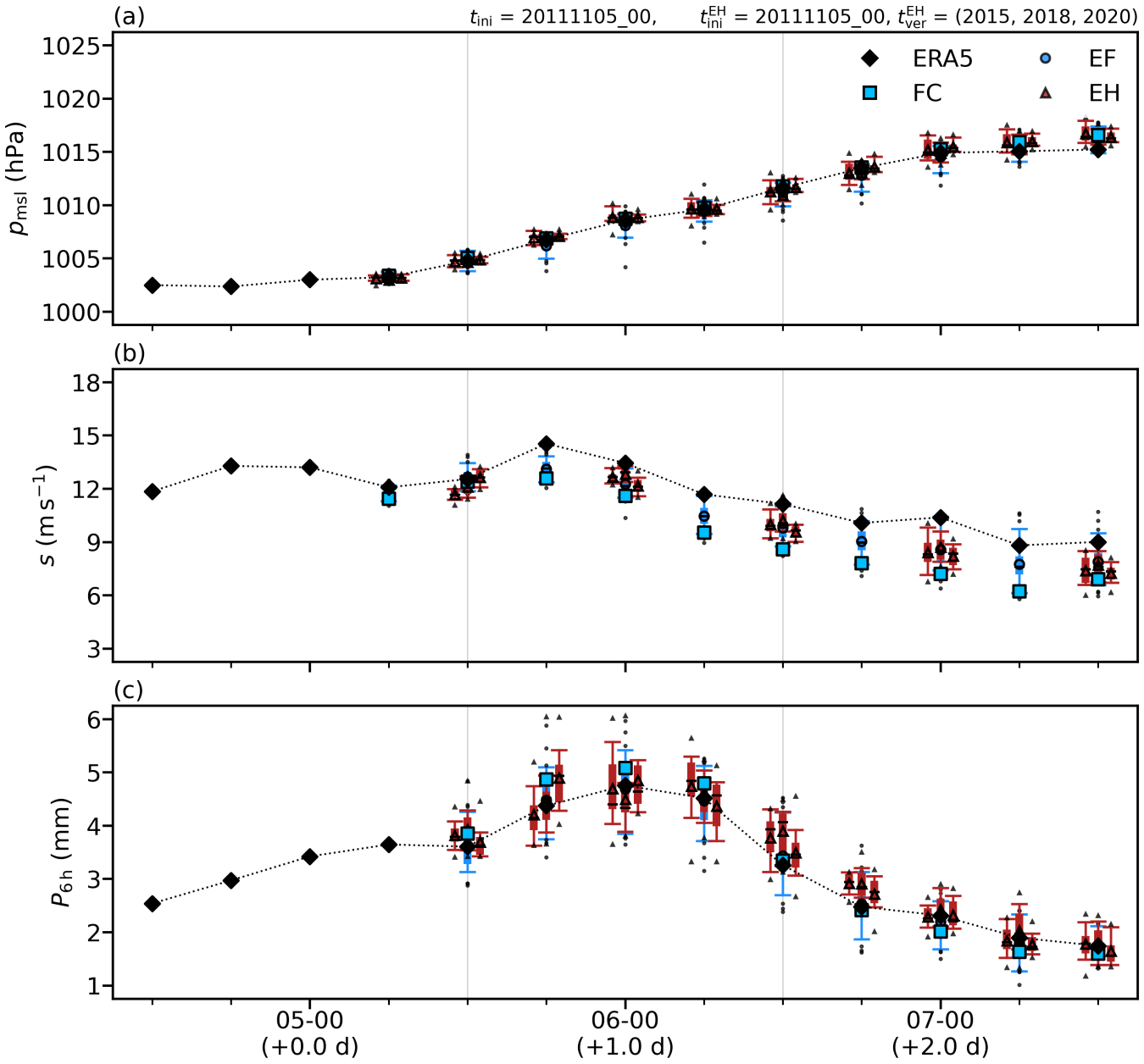


Figure B.46: Time series of forecasts at a short lead time for 2011. Similar to Figure B.45. All forecasts were initialised on 5. November 2011, at 00 UTC.

References

Argence, Sébastien et al. (2008). 'Impact of initial condition uncertainties on the predictability of heavy rainfall in the Mediterranean: a case study'. In: *Quarterly Journal of the Royal Meteorological Society* 134 (636), pp. 1775–1788. DOI: 10.1002/qj.314.

Barnes, Michael A. et al. (2022). 'Stratospheric intrusion depth and its effect on surface cyclogenetic forcing: an idealized potential vorticity (PV) inversion experiment'. In: *Weather and Climate Dynamics* 3 (4), pp. 1291–1309. DOI: 10.5194/wcd-3-1291-2022.

Bauer, Peter, Alan J. Thorpe and Gilbert Brunet (2015). 'The quiet revolution of numerical weather prediction'. In: *Nature* 525 (7567), pp. 47–55. DOI: 10.1038/nature14956.

Browning, Keith A. (1986). 'Conceptual Models of Precipitation Systems'. In: *Weather and Forecasting* 1 (1), pp. 23–41. DOI: 10.1175/1520-0434(1986)001<0023:CMOPS>2.0.CO;2.

Buzzi, Andrea and Luigi Foschini (2000). 'Mesoscale meteorological features associated with heavy precipitation in the southern Alpine region'. In: *Meteorology and Atmospheric Physics* 72 (2-4), pp. 131–146. DOI: 10.1007/s007030050011.

Čampa, Jana and Heini Wernli (2012). 'A PV perspective on the vertical structure of mature midlatitude cyclones in the northern hemisphere'. In: *Journal of the Atmospheric Sciences* 69 (2), pp. 725–740. DOI: 10.1175/JAS-D-11-050.1.

Campins, Joan et al. (2011). 'Climatology of Mediterranean cyclones using the ERA-40 dataset'. In: *International Journal of Climatology* 31 (11), pp. 1596–1614. DOI: 10.1002/joc.2183.

Campins, Joan et al. (2013). 'Influence of targeted observations on short-term forecasts of high-impact weather events in the Mediterranean'. In: *Natural Hazards and Earth System Sciences* 13 (11), pp. 2891–2910. DOI: 10.5194/nhess-13-2891-2013.

Carlson, Toby N. (1980). 'Airflow Through Midlatitude Cyclones and the Comma Cloud Pattern'. In: *Monthly Weather Review* 108 (10), pp. 1498–1509. DOI: 10.1175/1520-0493(1980)108<1498:ATMCAT>2.0.CO;2.

Cioni, Guido, Diego Cerrai and Daniel Klocke (2018). 'Investigating the predictability of a Mediterranean tropical-like cyclone using a storm-resolving model'. In: *Quarterly Journal of the Royal Meteorological Society* 144 (714), pp. 1598–1610. DOI: 10.1002/qj.3322.

COST (2022). *CA19109 - European network for Mediterranean cyclones in weather and climate*. URL: https://www.cost.eu/cost-action/european-network-for-mediterranean-cyclones-in-weather-and-climate/ (visited on 26/05/2022).

Di Muzio, Enrico et al. (2019). 'Assessing the predictability of Medicanes in ECMWF ensemble forecasts using an object-based approach'. In: *Quarterly Journal of the Royal Meteorological Society* 145 (720), pp. 1202–1217. DOI: 10.1002/qj.3489.

ECMWF (2018). *MARS - the ECMWF meteorological archive*. URL: https://www.ecmwf.int/node/18124 (visited on 09/03/2023).

ECMWF (2022). *IFS documentation*. URL: https://www.ecmwf.int/en/publications/ifs-documentation (visited on 30/03/2023).

Fita, L. et al. (2007). 'Analysis of the environments of seven Mediterranean tropical-like storms using an axisymmetric, nonhydrostatic, cloud resolving model'. In: *Natural Hazards and Earth System Sciences* 7 (1), pp. 41–56. DOI: 10.5194/nhess-7-41-2007.

Flaounas, Emmanouil et al. (2015). 'The dynamical structure of intense Mediterranean cyclones'. In: *Climate Dynamics* 44 (9), pp. 2411–2427. DOI: 10.1007/s00382-014-2330-2.

Flaounas, Emmanouil et al. (2018). 'Heavy rainfall in Mediterranean cyclones. Part I: contribution of deep convection and warm conveyor belt'. In: *Climate Dynamics* 50 (7-8), pp. 2935–2949. DOI: 10.1007/s00382-017-3783-x.

Flaounas, Emmanouil et al. (2022). 'Mediterranean cyclones: current knowledge and open questions on dynamics, prediction, climatology and impacts'. In: *Weather and Climate Dynamics* 3 (1), pp. 173–208. DOI: 10.5194/wcd-3-173-2022.

Flocas, H. A. (2000). 'Diagnostics of Cyclogenesis Over the Aegean Sea Using Potential Vorticity Inversion'. In: *Meteorology and Atmospheric Physics* 73 (1), pp. 25–33. DOI: 10.1007/s007030050061.

Froude, Lizzie S. R. (2012). 'The predictability of extratropical cyclones'. PhD thesis, pp. 31–45. URL: https://www.ecmwf.int/sites/default/files/elibrary/2012/9466-predictability-extratropical-cyclones.pdf (visited on 23/03/2023).

Haiden, Thomas et al. (2018). *Evaluation of ECMWF forecasts, including the 2018 upgrade*. DOI: 10.21957/ldw15ckqi.

Harrold, T. W. (1973). 'Mechanisms influencing the distribution of precipitation within baroclinic disturbances'. In: *Quarterly Journal of the Royal Meteorological Society* 99 (420), pp. 232–251. DOI: 10.1002/qj.49709942003.

Hersbach, Hans et al. (2020). 'The ERA5 global reanalysis'. In: *Quarterly Journal of the Royal Meteorological Society* 146 (730), pp. 1999–2049. DOI: 10.1002/qj.3803.

Homar, Victor et al. (2007). 'Towards a systematic climatology of sensitivities of Mediterranean high impact weather: A contribution based on intense cyclones'. In: *Natural Hazards and Earth System Science* 7 (4), pp. 445–454. DOI: 10.5194/nhess-7-445-2007.

Jansà, A. et al. (2014). 'MEDEX: a general overview'. In: *Natural Hazards and Earth System Sciences* 14 (8), pp. 1965–1984. DOI: 10.5194/nhess-14-1965-2014.

Jones, P. D. et al. (1999). 'Surface air temperature and its changes over the past 150 years'. In: *Reviews of Geophysics* 37 (2), pp. 173–199. DOI: 10.1029/1999RG900002.

Lionello, Piero et al. (2016). 'Objective climatology of cyclones in the Mediterranean region: a consensus view among methods with different system identification and tracking criteria'. In: *Tellus A: Dynamic Meteorology and Oceanography* 68 (1). DOI: 10.3402/tellusa.v68.29391.

Massacand, Alexia C., Heini Wernli and Huw C. Davies (1998). 'Heavy precipitation on the Alpine southside : An upper-level precursor'. In: *Geophysical Research Letters* 25 (9), pp. 1435–1438. DOI: 10.1029/98GL50869.

Moscatello, Agata, Mario Marcello Miglietta and Richard Rotunno (2008). 'Numerical analysis of a mediterranean "Hurricane" over Southeastern Italy'. In: *Monthly Weather Review* 136 (11), pp. 4373–4397. DOI: 10.1175/2008MWR2512.1.

Nakamura, Mototaka and Shozo Yamane (2009). 'Dominant anomaly patterns in the near-surface baroclinicity and accompanying anomalies in the atmosphere and oceans. Part I: North Atlantic basin'. In: *Journal of Climate* 22 (4), pp. 880–904. DOI: 10.1175/2008JCLI2297.1.

Nicolaides, K. A., S. C. Michalelides and T. Karacostas (2006). 'Synoptic and dynamic characteristics of selected deep depressions over Cyprus'. In: *Advances in Geosciences* 7, pp. 175–180. DOI: 10.5194/adgeo-7-175-2006.

Oertel, Annika et al. (2019). 'Convective activity in an extratropical cyclone and its warm conveyor belt – a case-study combining observations and a convection-permitting model simulation'. In: *Quarterly Journal of the Royal Meteorological Society* 145 (721), pp. 1406–1426. DOI: 10.1002/qj.3500.

Portmann, Raphael et al. (2020). 'How an uncertain short-wave perturbation on the North Atlantic wave guide affects the forecast of an intense Mediterranean cyclone (Medicane Zorbas)'. In: *Weather and Climate Dynamics* 1 (2), pp. 597–615. DOI: 10.5194/wcd-1-597-2020.

Raveh-Rubin, Shira and Emmanouil Flaounas (2017). 'A dynamical link between deep Atlantic extratropical cyclones and intense Mediterranean cyclones'. In: *Atmospheric Science Letters* 18 (5), pp. 215–221. DOI: 10.1002/asl.745.

Raveh-Rubin, Shira and Heini Wernli (2016). 'Large-scale wind and precipitation extremes in the Mediterranean: dynamical aspects of five selected cyclone events'. In: *Quarterly Journal of the Royal Meteorological Society* 142 (701), pp. 3097–3114. DOI: 10.1002/qj.2891.

Ricchi, Antonio et al. (2017). 'Sensitivity of a Mediterranean Tropical-Like Cyclone to Different Model Configurations and Coupling Strategies'. In: *Atmosphere* 8 (5). DOI: 10.3390/atmos8050092.

Sanders, Frederick and John R. Gyakum (1980). 'Synoptic-Dynamic Climatology of the "Bomb"'. In: *Monthly Weather Review* 108 (10), pp. 1589–1606. DOI: 10.1175/1520-0493(1980)108<1589:SDCOT>2.0.CO;2.

Schulzweida, Uwe (2022). *CDO User Guide*. Version 2.1.1.

Sprenger, Michael et al. (2017). 'Global climatologies of Eulerian and Lagrangian flow features based on ERA-Interim'. In: *Bulletin of the American Meteorological Society* 98 (8), pp. 1739–1748. DOI: 10.1175/BAMS-D-15-00299.1.

Struck, Doug (1998). *Freak Jerusalem storm dumps dust, hail, snow*. URL: https://www.washingtonpost.com/archive/politics/1998/03/19/freak-jerusalem-storm-dumps-dust-hail-snow/02b7abe6-4797-4dc6-b255-41d524c2a164/ (visited on 14/03/2023).

Trigo, Isabel F. (2006). 'Climatology and interannual variability of storm-tracks in the Euro-Atlantic sector: A comparison between ERA-40 and NCEP/NCAR reanalyses'. In: *Climate Dynamics* 26 (2-3), pp. 127–143. DOI: 10.1007/s00382-005-0065-9.

Trigo, Isabel F., Grant R. Bigg and Trevor D. Davies (2002). 'Climatology of Cyclogenesis Mechanisms in the Mediterranean'. In: *Monthly Weather Review* 130 (3), pp. 549–569. DOI: 10.1175/1520-0493(2002)130<0549:COCMIT>2.0.CO;2.

Tripoli, G. J. et al. (2005). 'The 9-10 November 2001 Algerian flood: A numerical study'. In: *Bulletin of the American Meteorological Society* 86 (9), pp. 1229–1235. DOI: 10.1175/BAMS-86-9-1229.

Wernli, Heini and Cornelia Schwierz (2006). 'Surface cyclones in the ERA-40 dataset (1958-2001). Part I: Novel identification method and global climatology'. In: *Journal of the Atmospheric Sciences* 63 (10), pp. 2486–2507. DOI: 10.1175/JAS3766.1.

Wernli, Heini et al. (2008). 'SAL—A Novel Quality Measure for the Verification of Quantitative Precipitation Forecasts'. In: *Monthly Weather Review* 136 (11), pp. 4470–4487. DOI: 10.1175/2008MWR2415.1.