



Research and Innovation action

H2020-SC5-2017

# **Benchmarking skill assessment of current sub-seasonal and seasonal forecast systems for users' selected case studies**

**Deliverable D4.1**

Version 2.0

Authors: Ilias Pechlivanidis, Louise Crochemore (SMHI), Albert Soret, Llorenç Lledó, Andrea Manrique-Suñén (BSC), David Brayshaw, Paula Gonzalez, Andrew Charlton Perez, Hannah Bloomfield (UREAD), Franco Catalano, Irene Cionni (ENEA)

## **Disclaimer**

The content of this deliverable reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

## Document Information

<b>Grant Agreement</b>	776787
<b>Project Title</b>	Sub-seasonal to Seasonal climate forecasting for Energy
<b>Project Acronym</b>	S2S4E
<b>Project Start Date</b>	01/12/2017
<b>Related work package</b>	WP 4. S2S Climate predictions
<b>Related task(s)</b>	Task 4.1: Benchmark assessment of the case studies using hindcast data
<b>Lead Organisation</b>	SMHI
<b>Submission date</b>	30/11/2018; 29/03/2019
<b>Dissemination Level</b>	PU

## History

<b>Date</b>	<b>Submitted by</b>	<b>Reviewed by</b>	<b>Version (Notes)</b>
21/11/2018	Ilias Pechlivanidis (SMHI)	Albert Soret (BSC)	Version 1.1
29/11/2018	Ilias Pechlivanidis (SMHI)	Harilaos Loukos (TCDF), Andrea Manrique (BSC)	Version 1.2
30/11/2018	Ilias Pechlivanidis (SMHI)	Llorenç Lledó (BSC), all authors	Version 1.3
28/02/2019	Franco Catalano, Irene Cionni, Alessandro Dell'Aquila (ENEA)	Albert Soret (BSC)	Version 2.1: solar radiation and solar capacity factor included in case study 2 and 8.
29/03/2019	Albert Soret (BSC)	Albert Soret (BSC)	Version 2.2: Annex 4 (factsheets on case studies)

## Table of content

About S2S4E .....	13
Summary .....	14
Keywords .....	15
Glossary .....	15
1. Introduction .....	17
2. Background on S2S forecasting .....	19
2.1. An overview of sub-seasonal to seasonal forecasting .....	19
2.2. Sectorial applications .....	20
2.2.1. Hydro .....	20
2.2.2. Wind .....	20
2.2.3. Solar and Demand .....	21
2.2.4. Country-scale supply demand balance indicators .....	21
3. Datasets and Models .....	24
3.1. The forecasting systems .....	24
3.1.1. Seasonal time-scale .....	24
3.1.2. Sub-seasonal time-scale .....	25
3.2. The hydrological model .....	26
3.3. Demand and wind-power models .....	28
4. Methodology and evaluation protocol .....	30
4.1. Climate models and bias-adjustment .....	30
4.1.1. Hydro bias-adjustment protocol .....	30
4.1.2. Wind speed bias-adjustment protocol .....	31
4.1.3. Supply-demand bias-adjustment protocol .....	31
4.2. ECVs and indicators .....	34
4.3. Benchmarking and metrics .....	34
4.4. The case studies .....	35
5. Europe-wide skill assessment .....	36
5.1. Precipitation .....	36
5.1.1. Sub-seasonal time-scale .....	36
5.1.2. Seasonal time-scale .....	41
5.2. Surface wind speed .....	45
5.2.1. Sub-seasonal time-scale .....	45
5.2.2. Seasonal time-scale .....	48



5.3.	Country-scale supply-demand balance indicators .....	53
5.3.1.	Sub-seasonal time-scale.....	53
5.3.2.	Seasonal time-scale.....	57
5.4.	Conclusions from the European-wide assessment.....	60
5.5.	Limitations from the assessment.....	61
6.	Case study assessment and evaluation.....	63
6.1.	Case study 1 – France, Germany 2017 .....	64
6.2.	Case study 2 – Germany 2013 .....	70
6.3.	Case study 3 – Spain 2016.....	78
6.4.	Case study 4 – Sweden 2015 .....	85
6.5.	Case study 5 – Romania 2014 .....	90
6.6.	Case study 6 – USA 2015.....	95
6.7.	Case study 7 – France, Europe 2018 .....	102
6.8.	Case study 8 – Spain 2018.....	107
7.	Conclusions.....	116
7.1.	Moving forward .....	117
	Bibliography.....	118
	Annex .....	122
	Annex 1 – Bias-adjustment of ENS-ER for hydro .....	122
	Annex 2 – European-wide assessment – Additional metrics .....	123
	Annex2.1. Seasonal time-scale .....	124
	Annex2.2. Sub-seasonal time-scale .....	127
	Annex 3 – Skill assessment of current S2S systems for the user’s selected case studies. ...	138
	Annex 4 – Factsheets on the case studies. ....	149

## List of figures

Figure 1: Domain of the E-HYPE hydrological model.....	27
Figure 2: Bias-adjustment procedure for the sub-seasonal forecasts.....	31
Figure 3: Comparison of MERRA-2 and ERA5 wind (scaled to 100m) and resultant capacity factors. Top: Sample one-month period of capacity factor. Bottom left: Frequency distribution of hourly country-average wind speeds. Bottom right: Frequency distribution of hourly country-average wind power capacity factors. In each plot, orange line denotes MERRA-2 and blue line denotes ERA5.....	33
Figure 4: Forecast week 1 monthly spatial variability of the CRPS score (CRPSS).....	37
Figure 5: Forecast week 2 monthly spatial variability of the CRPS score (CRPSS).....	38
Figure 6: Forecast week 3 monthly spatial variability of the CRPS score (CRPSS).....	39
Figure 7: Forecast week 4 monthly spatial variability of the CRPS score (CRPSS).....	40
Figure 8: Forecast month 1 monthly spatial variability of the CRPS score (CRPSS).....	42
Figure 9: Forecast month 2 monthly spatial variability of the CRPS score (CRPSS).....	43
Figure 10: Forecast month 3 monthly spatial variability of the CRPS score (CRPSS).....	44
Figure 11: FairCRPSS of 10 m wind speed from ECMWF monthly prediction system for January to June and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts are calibrated with the variance inflation method. (Note that the scale is not linear).....	46
Figure 12: FairCRPSS of 10m wind speed from ECMWF monthly prediction system for July to December and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts are calibrated with the variance inflation method. (Note that the scale is not linear).....	47
Figure 13: Fair CRPSS of ECMWF SEAS5 predictions of surface wind. Each row presents the skill for the forecasts started on January to June, with the columns representing one to three months of lead time. ....	49
Figure 14: Fair CRPSS of ECMWF SEAS5 predictions of surface wind. Each row presents the skill for the forecasts started on July to December, with the columns representing one to three months of lead time. ....	50
Figure 15: Fair RPSS of ECMWF SEAS5 predictions of surface wind. Each row presents the skill for the forecasts started on January to June, with the columns representing one to three months of lead time. ....	51

- Figure 16: Fair RPSS of ECMWF SEAS5 predictions of surface wind. Each row presents the skill for the forecasts started on July to December, with the columns representing one to three months of lead time. .... 52
- Figure 17: ECMWF sub-seasonal forecast skill for weekly-mean, national-aggregate demand-net-wind for validation weeks in January. Left four panels – ensemble mean correlation. Right four panels – CRPSS. In each set of four subfigures, the subtitles indicate lead time (weeks 1-4)..... 53
- Figure 18: As Figure 17 but for ECMWF sub-seasonal forecast skill for weekly-mean, national-aggregate demand-net-wind for validation weeks in July. .... 54
- Figure 19: ECMWF sub-seasonal forecast skill (ensemble mean correlation score) for weekly-mean, national-aggregate demand-net-wind for a selection of countries as a function of verification month. The subtitles indicate lead time (weeks 1-4). .... 55
- Figure 20: ECMWF sub-seasonal forecast skill (ensemble mean correlation score) for weekly-mean, national-aggregate demand, wind-power, and demand-net-wind for Norway and Germany by verification month. Lines: blue – forecast week 1, green – forecast week 2, red – forecast week 3, pink – forecast week 4. .... 57
- Figure 21: As Figure 17 but for ECMWF seasonal forecast skill for monthly-mean, national-aggregate demand-net-wind for validation month in January. In each set of four panels, the subtitles indicate lead time (months 0-3). .... 57
- Figure 22: As Figure 17 but for ECMWF seasonal forecast skill for monthly-mean, national-aggregate demand-net-wind for validation month in July. In each set of four panels, the subtitles indicate lead time (months 0-3). .... 58
- Figure 23: As Figure 19 but for ECMWF seasonal forecast skill for monthly-mean, national-aggregate demand-net-wind for a selection of countries as a function of verification month. The subtitles indicate lead time (months 0-3). .... 59
- Figure 24: Observed and surface meteorological conditions compared to ERA-Interim climatology during 17<sup>th</sup> Jan – 23<sup>rd</sup> Jan 2017. .... 65
- Figure 25: Observed and climatological surface air temperature (top left), 10m wind speed (top right) and precipitation (bottom) averaged over 5° W–12° E, 47–54° N during December 2016 to February 2017. .... 66
- Figure 26: Sub-seasonal forecast for 10m wind speed averaged over 5°W–12°E, 47–54°N for the week 17<sup>th</sup>–23<sup>rd</sup> January 2017. From right to left corresponds to forecasts launched from lead times week 1 to 4 (forecasts times 5-11 days, 12-18 days, 19-25 days, 26-32 days). Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS)..... 67
- Figure 27: Sub-seasonal forecast for temperature averaged over 5°W–12°E, 47–54°N for the week 17<sup>th</sup>–23<sup>rd</sup> January 2017. From left to right corresponds to forecasts launched from

lead times week 4 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).....	68
Figure 28: Sub-seasonal forecast for demand averaged over 5°W–12°E, 47–54°N for the week 17 <sup>th</sup> –23 <sup>rd</sup> January 2017. From left to right corresponds to forecasts launched from lead times week 4 to 1. Methodology: lead time dependent mean bias correction, applied to both temperature and demand (once converted), calibrated to ERA5, based on a 17-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).....	69
Figure 29: Standardized anomalies of temperature, precipitation, wind and radiation for July 2013, obtained from ERA-Interim. ....	71
Figure 30: Weekly evolution of the observed ECVs in the region 6–15°E and 47–54.5°N during July 2013 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis. ....	72
Figure 31: Precipitation forecasts for July 2013 issued three, two and one months in advance for Germany.....	73
Figure 32: Inflows forecasts for July 2013 issued three, two and one months in advance for the Rhine river, Germany.....	73
Figure 33: Inflows forecasts for July 2014 issued three, two and one months in advance for the Neckar river, Germany. ....	74
Figure 34: Inflows forecasts for July 2014 issued three, two and one months in advance for the Elbe river, Germany.....	74
Figure 35: Surface solar radiation forecasts for July 2013 issued three, two and one months in advance for domain (6E-15E, 47N-54N). Methodology: mean inflation calibration to ERA-Interim, based on 38 year hindcast. ....	75
Figure 36: Photovoltaic capacity factor forecasts for July 2013 issued three, two and one months in advance for domain (6E-15E, 47N-54N). Methodology: mean inflation calibration to ERA-Interim, based on 38 year hindcast.....	75
Figure 37: Wind speed forecasts for July 2013 issued three, two and one months in advance for the domain of study: 6–15°E and 47–54.5°N. ....	76
Figure 38: Temperature forecasts for July 2013 issued three, two and one months in advance for Germany.....	77
Figure 39: Demand forecasts for July 2013 issued three, two and one months in advance for Germany.....	77
Figure 40: Standardized anomalies of temperature, precipitation, wind and radiation for week starting on 30 of August 2016, obtained from ERA-Interim.....	79

Figure 41: Weekly evolution of the observed temperature in the region 8.5°W–3°E and 36.5–43.5°N from August to October compared to the climatological distribution. Values obtained from ERA-Interim reanalysis. ....	79
Figure 42: Weekly evolution of the observed surface wind in the region 8.5°W–3°E and 36.5–43.5°N from August to October compared to the climatological distribution. Values obtained from ERA-Interim reanalysis. ....	80
Figure 43: Sub-seasonal forecast for 10m wind speed averaged for the week 30 <sup>th</sup> August – 5 <sup>th</sup> September 2016 for 0E, 40.5N. From right to left corresponds to forecasts launched from lead times week 1 to 4 (forecasts times 5-11 days, 12-18 days, 19-25 days, 26-32 days). Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).....	81
Figure 44: Sub-seasonal forecast for 2 m temperature averaged for the week 30 <sup>th</sup> August – 5 <sup>th</sup> September 2016 for 0E, 40.5N. From right to left corresponds to forecasts launched from lead times week 1 to 4 (forecasts times 5-11 days, 12-18 days, 19-25 days, 26-32 days). Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).....	82
Figure 45: Demand forecasts for 30 August 2016 issued four, three, two and one weeks in advance over Spain.....	83
Figure 46: Demand net wind forecasts for 30 August 2016 issued four, three, two and one weeks in advance over Spain. ....	84
Figure 47: Standardized anomalies of temperature, precipitation, wind and radiation for July 2015, obtained from ERA-Interim. ....	86
Figure 48: Weekly evolution of the observed precipitation in the region 13–21.5°E and 63–66°N during May and June 2015 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis. ....	87
Figure 49: Weekly evolution of the observed temperature in the region 13–21.5°E and 63–66°N during May and June 2015 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis. ....	87
Figure 50: Precipitation forecasts for May-July 2015 issued three, two and one months in advance for the Umeälven river basin.....	89
Figure 51: Snow water equivalent forecasts for May-July 2015 issued three, two and one months in advance for the Umeälven river basin. ....	89
Figure 52: Inflows forecasts for May-July 2015 issued three, two and one months in advance for the Umeälven river basin.....	90

Figure 53: Observed and surface meteorological conditions compared to climatology during 28th Jan – 3rd Feb 2014.....	92
Figure 54: Observed and climatological surface air temperature (left) and 10m wind speed (right) in the region 24–29.5°E and 43.5–47°N during January and February 2014.....	92
Figure 55: Sub-seasonal forecast for temperature for a specific grid point (27.5 E, 46.6 N) for the week 28/01/2014-03/02/2014. From left to right corresponds to forecasts launched from lead times week 4 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS). .....	93
Figure 56: Sub-seasonal forecast for minimum temperature for a specific grid point (27.5 E, 46.6 N) for the week 28/01/2014-03/02/2014. From left to right corresponds to forecasts launched from lead times week 4 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS). .....	94
Figure 57: Standardized anomalies of temperature, precipitation, wind and radiation for the first quarter of 2015, obtained from ERA-Interim.....	96
Figure 58: Weekly evolution of the observed surface wind in the region 124°W-95°W and 26°N-44°N during Q1 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis.....	96
Figure 59: Surface wind speed forecasts for Jan-Mar 2015 issued three, two and one months in advance for the western USA region. ....	98
Figure 60: Capacity factor forecasts for IEC1, IEC2 and IEC3 class turbines, for the region and period of interest.....	99
Figure 61: ENSO and NPM patterns (left) and its evolution during the last years for Q1 (extracted from Lledó et al. 2018). .....	100
Figure 62: SST and surface wind speed anomalies for the first months of 2015, expressed as the number of standard deviations away from the 1979–2014 mean for the same month, drawn from ERA-Interim reanalysis. Purple contours show precipitation anomalies, and green contours show sea level pressure anomalies. Contour intervals are 5 mm/day and 2 hPa, respectively, with zero contour omitted (extracted from Lledó et al. 2018). .....	101
Figure 63: Standardized anomalies of temperature, precipitation, wind and radiation for the first quarter of 2018, obtained from ERA-Interim.....	103
Figure 64: Weekly evolution of the observed precipitation in the region 10°W-30°E and 36°N-65°N during February and March 2018 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis. ....	104

Figure 65: Weekly evolution of the observed temperature in the region 10°W-30°E and 36°N-65°N during February and March 2018 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis. ....	104
Figure 66: Temperature forecasts for 27 Feb 2018 issued four, three, two and one weeks in advance for the domain (5W-12E, 47N-54N). ....	106
Figure 67: Electricity demand forecasts for 27 Feb 2018 issued four, three, two and one weeks in advance for France. ....	106
Figure 68: Observed and surface meteorological conditions compared to climatology during March 2018. ....	108
Figure 69: Observed and climatological surface air temperature (top left), 10m wind speed (top right) and precipitation (bottom) averaged over 8.5°W-3E, 36.5-43.5°N during February and March 2018. ....	109
Figure 70: Seasonal forecast for 10m wind speed averaged over 5°W-12°E, 47-54°N for March 2018. From left to right corresponds to forecasts launched from lead times month 3 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. ....	110
Figure 71: Seasonal forecast for capacity factors derived from the IEC2 class turbine (Gamesa G87 2MW), averaged over 5°W - 12°E, 47-54°N for March 2018. From left to right corresponds to forecasts launched from lead times month 3 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS). ....	111
Figure 72: Sub-seasonal forecasts for nationally aggregated wind power for Spain for March 2018. From left to right corresponds to forecasts launched from lead times week 4 to 1. Methodology: lead time dependent mean bias correction, applied to both 10m wind speed and national wind power generation (once calculated from the wind), calibrated to ERA5, based on a 17-year hindcast. Note that the ERA5 climatology used to calibrate the forecast leads to unrealistically low wind power generation estimates (in GW). ....	112
Figure 73: Precipitation forecasts for March 2018 issued three, two and one months in advance for Tajo and Mino regions in Spain. ....	113
Figure 74: Seasonal forecast for temperature averaged over 5°W-12°E, 47-54°N for March 2018. From left to right corresponds to forecasts launched from lead times month 3 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. ....	114
Figure 75: Surface solar radiation forecasts for July 2013 issued three, two and one months in advance for domain (6E-15E, 47N-54N). Methodology: mean inflation calibration to ERA-Interim, based on 38 year hindcast. ....	115

Figure 76: Photovoltaic capacity factor forecasts for July 2013 issued three, two and one months in advance for domain (6E-15E, 47N-54N). Methodology: mean inflation calibration to ERA-Interim, based on 38 year hindcast..... 115

## List of tables

Table 1: Characteristics of ECMWF SEAS5 seasonal forecasting system. ....	25
Table 2: Characteristics of the climate models used for sub-seasonal and seasonal forecasting. ....	26
Table 3: Data sources and characteristics of the E-HYPE setup. ....	28
Table 4: List of ECVs and indicators requested by users. ....	34
Table 5: List of case studies. ....	35
Table 6: Region, period, forecast type and main interest for case study 1. ....	64
Table 7: Region, period, forecast type and main interest for case study 2. ....	70
Table 8: Region, period, forecast type and main interest for case study 3. ....	78
Table 9: Region, period, forecast type and main interest for case study 4. ....	85
Table 10: Start dates, lead times, valid period and variables for each of the forecasts presented.....	88
Table 11: Region, period, forecast type and main interest for case study 5.....	90
Table 12: Region, period, forecast type and main interest for case study 6.....	95
Table 13: Start dates, lead times, valid period and variables for each of the forecasts presented.....	97
Table 14: Region, period, forecast type and main interest for case study 7.....	102
Table 15: Start dates, lead times, valid period and variables for each of the forecasts presented.....	105
Table 16: Region, period, forecast type and main interest for case study 8.....	107



## About S2S4E

The project seeks to improve renewable energy variability management by developing a tool that for the first time integrates sub-seasonal to seasonal climate predictions with renewable energy production and electricity demand. Our long-term goal is to make the European energy sector more resilient to climate variability and extreme events. Large-scale deployment of renewable energy is key to comply with the emissions reductions agreed upon in the Paris Agreement. However, despite being cost competitive in many settings, renewable energy diffusion remains limited largely due to seasonal variability. Knowledge of power output and demand forecasting beyond a few days remains poor, creating a major barrier to renewable energy integration in electricity networks. To help solve this problem, S2S4E is developing an innovative service to improve renewable energy variability management. The outcome will be new research methods exploring the frontiers of weather conditions for future weeks and months and a decision support tool for the renewable industry.

More information: [www.s2s4e.eu](http://www.s2s4e.eu)

## Summary

Renewable energy (i.e. wind, solar, hydro, electricity demand) is the fastest growing source of electricity globally. Although renewable energy diffusion still faces important challenges related to large-scale integration in the energy system, there is already high potential over the European energy market. The S2S4E project aims to: 1) foster renewable energy deployment while also maintaining energy security by providing sub-seasonal to seasonal (S2S) climate forecasts; 2) enable the energy industry and policy makers to assess how well different renewable energy sources will meet demand over extended time horizons (weeks to months), focusing on the impact of climate variables on energy outputs and needs; 3) make the European energy sector more resilient to climate variability and high impact events by exploring the frontiers of what can be achieved by using S2S predictions offering a new decision support tool based on S2S climate predictions; 4) contribute for the expansion of climate services to users and markets as it will base its development on a user-centric framework for co-design and co-development.

To achieve the projects objectives, an exploration of the scientific frontiers (multi-modelling, downscaling, bias-adjustment etc.) of sub-seasonal predictions and their synergies with seasonal predictions is needed, allowing providing useful information for decision-making process. In addition, statistical assessment of the predictions and their economic added-value from S2S predictions based on a historical period and historical case studies is important for improved communication with the users of the energy sector. Prior to a quantification of the added-value, a benchmarking assessment is necessary, in which a single model approach (i.e. analysing forecasts of individual climate systems from different data/service providers) is employed, based also on simple BA methods to downscale the forecasted information at the local scale. The outcomes of this benchmarking assessment allow the users to have a first vision of the potential application of S2S climate predictions for decision-making.

This report presents results from the S2S benchmarking assessment using a set of essential climate variables (ECVs) and indicators identified by the users in order to maximize the utility of S2S forecasts and better inform the design of a real-time forecasting system. The assessment is conducted at two different spatial extents, i.e. large- (European-wide) and local-scale, and using two different ECMWF-based climate systems, i.e. ECMWF Extended Range (ENS-ER) and ECMWF SEAS5 for sub-seasonal and seasonal assessment respectively. The large-scale complements the “deep” knowledge from local scale assessment, whilst giving an overview of the predictability to users with large-scale interests. Large-scale assessment has the potential to encompass many regions, cross-regional and international boundaries and represent a number of different climatic zones and hence it allows for exploration of emerging patterns and facilitation of comparative analysis. In addition to the

pan-European assessment, a number of case studies have been selected based on the simulation and analysis of real data obtained from past periods with an unusual climate behaviour affecting the energy market. These case studies provide in depth understanding of S2S climate service application in the energy sector.

## Keywords

Forecasting; Sub-seasonal; Seasonal time-scale; Energy; Benchmarking; Wind; Hydro; Solar; Demand

## Glossary

<b>Bias-Adjustment (BA)</b>	Process aiming at removing systematic errors in the output of a model. Methods include: linear scaling, distribution-based scaling, quantile mapping.
<b>Downscaling</b>	General name for a procedure to take information known at large scales to make predictions at local scales. The two main approaches to downscaling climate information are dynamical and statistical. Statistical methods include: linear scaling, distribution-based scaling, quantile mapping.
<b>ECMWF</b>	European Centre for Medium-Range Weather Forecasts
<b>ECV</b>	Essential Climate Variable
<b>ENS-ER</b>	ECMWF Extended Range (sub-seasonal forecasts) ensemble prediction system
<b>Forecast time</b>	The time between the initiation and completion of a forecast.
<b>Forecast quality</b>	How well a forecast compares against a corresponding observation of what actually occurred, or some good estimate of the true outcome.
<b>Initial conditions (IC)</b>	The hydrological states (soil moisture, snow cover, water already in the river, among others) at or close to the start

	of the forecast run.
<b>Seasonal climate forcing (SCF)</b>	The seasonal meteorological forecast used as input to a hydrological model.
<b>S2S</b>	Sub-seasonal to seasonal forecasts
<b>S2S4E</b>	The project "Sub-seasonal to seasonal climate predictions for energy"
<b>SEAS5</b>	ECMWF long-range System 5 (seasonal forecasts)
<b>SS</b>	Sub-seasonal forecasts

# 1. Introduction

Climatic information from sub-seasonal (6 weeks ahead) to seasonal (7 months ahead) time-scales is needed for decision-making in a number of sectors. Compared to the short-to-medium-range (up to 10 days ahead) forecasts, S2S time-scales hold the potential for being of great value for a wide range of users who are affected by variability in climate, water and energy and who would benefit from understanding and better managing climate-related risks (Bruno Soares et al., 2017; Stoft, 2002; Green, 2005). Wind, solar, hydro and energy demand are examples of renewable energy applications in which S2S information can affect decision-making. In Europe, there has been relatively little uptake and use of S2S forecasts by users for decision making, compared to other parts of the world, such as the USA and Australia, probably due to the relatively limited inherent predictability and limited quality of models and observations (Bennett et al., 2017; Mendoza et al., 2017; Arnal et al., 2018). However, recent advances in our understanding and forecasting of climate have resulted in skillful and useful climatic predictions, which can consequently increase the confidence of energy-related prognoses, and improve awareness, preparedness and decision-making from a user perspective (Bruno Soares and Dessai, 2016).

The predictability of S2S forecasts is subject to multiple sources of error and uncertainty, which are present in the various components of the production chain going from climate models (their parameterization, initialization, bias-adjustment etc.) to the service provides impact indicators (impact model setup, structure and parameterization). Consequently, to improve the predictability of the climate services, each component has to be evaluated to assess its contribution to the overall forecasting accuracy. In addition, predictability is characterized by strong spatial variation and commonly a temporal degradation of its skill in longer time-scales (Arnal et al., 2017; Greuell et al., 2018).

In order to address the scientific and technical challenges for improved climate services, S2S4E aims to: 1) foster renewable energy deployment while also maintaining energy security by providing sub-seasonal to seasonal (S2S) climate forecasts; 2) enable the energy industry and policy makers to assess how well different renewable energy sources will meet demand over extended time horizons (weeks to months), focusing on the impact of climate variables on energy outputs and needs; 3) make the European energy sector more resilient to climate variability and high impact events by exploring the frontiers of what can be achieved by using S2S predictions offering a new decision support tool based on S2S climate predictions; 4) contribute for the expansion of climate services to users and markets as it will base its development on a user-centric framework for co-design, co-development and co-evaluation.

This report sets the benchmark performance of existing climate systems and current services allowing further the quantification of improvements with the scientific advancements from this project. It is therefore focused on the analysis of the current skill of state-of-the-art S2S forecasts for a number of Essential Climate Variables (ECVs) and indicators at the European scale and local scale. The aim of large-scale analysis is to provide an overview of the predictive skill beyond cross regional and international boundaries, allowing also the

consideration of different climatic zones and regimes; hence it can provide a deeper understanding of the underlying sensitivities in the forecasting skill. On the other hand, local scale and case study assessments provide a deep understanding of the drivers causing an event, whilst such assessments are better communicated due to generally higher interest to the users.

Here we make a first step forward by benchmarking the spatial patterns of S2S predictability at the large scale and link this to the characteristics of the local systems for specific events. Two systems are used to benchmark the S2S forecasts: the ECMWF Extended Range (ENS-ER) and SEAS5 to provide sub-seasonal (4 weeks ahead) and seasonal (3 months ahead) information respectively. We pose the following questions: (1) What are the limits of predictability for S2S forecasting systems over Europe? (2) Which is the predictability of forecasting systems for a set of user identified events and which could be the drivers affecting the S2S forecasting skill? To address these questions, we assess the forecasting accuracy across Europe's climatic gradient for all initialization weeks and months and different lead times. In addition, together with the S2S4E partners, 8 case studies are selected (find more details in Deliverable 2.1) addressing needs and high interest for wind, solar, hydro and energy demand. A detailed presentation of the S2S forecasts from the two systems is provided for better understanding of the drivers affecting the forecasting skill. The S2S forecast data are adjusted for biases based with simple bias-adjustment (BA) methods for each ECVs, and the assessment are conducted for a hindcast period using a number of skill metrics.

A background of the S2S forecasting systems and their potential in the energy sector is given in Section 2. The datasets and models used here are presented in Section 3, followed by a description of the methodology and evaluation protocol in Section 4. The European-wide evaluation for a number of ECVs is presented in Section 5, whilst Section 6 presents the evaluation for the case study events that were identified by the users. Finally Section 7 states the conclusions.

## 2. Background on S2S forecasting

### 2.1. An overview of sub-seasonal to seasonal forecasting

Sub-seasonal and seasonal forecasts provide predictions of how the average atmospheric conditions over particular periods of time are likely to be different from the long-term average. Forecasts at such time-scale are useful to a number of sectors, including energy, and can allow better decision-making from the users or even improved preparedness in the case of extreme conditions (Bazile et al., 2017; Meißner et al., 2017; Sene et al., 2018). Moreover, the prediction at the monthly and seasonal timescales is potentially very useful in order to: (i) produce products which extend the medium-range forecast horizon (up to 10 days ahead), (ii) benefit from hindcasts for pre- and post-processing to produce output of higher quality (e.g., model-based return periods), and (iii) re-design decision support frameworks complementing them with early information.

Recently, there has been attention to sub-seasonal ranges, probably because the extended lead time provided by numerical weather predictions provides added value in very early planning and decision-making. The main motivation for sub-seasonal forecasting is to provide meteorological information in the time range of 10 to 46 days and hence address relative user needs. Forecasting services using sub-seasonal information could close the identified gap between forecasts on the medium and seasonal range (White et al., 2017). Although the objective of such forecasting service is not to capture short extreme events, e.g. floods, they are able to detect anomalous conditions on events lasting about weeks, such as droughts (Dutra et al., 2014).

However, there is a major challenge. Medium-range weather forecasting is mostly an atmospheric initial value problem, whilst seasonal forecasting is justified by the long predictability of the atmospheric boundary conditions (ocean/land/ice), which can be of the order of several months and by their impacts on the atmospheric circulation. In other words, the sub-seasonal time range is probably short enough that the atmosphere still has a memory of its initial condition and long enough that the ocean variability could have an impact on the atmospheric circulation.

A number of forecasting centres provide predicted atmospheric variables at ranges from a month to 13 months ahead, e.g. ECMWF SEAS5, ECMWF ENS-ER, MeteoFrance, GloSEA5 from the UK MetOffice, NCEP CFSv2 etc. Such services rely on coupled atmospheric-ocean-land global circulation models; however, the configuration (i.e. spatial resolution, ensemble members etc.) differs between the existing services and hence the forecasting skill varies as a function of geographical domain, week/month, and lead time. Consequently it is expected that a multi-model approach could “fill in the gaps” at occasions with limited skill.

Together with forecasts, re-forecasts are provided to assess the skill and reliability of the forecasting system. Skill generally measures the ability of the re-forecast to reproduce past intra-annual and inter-annual variability, and whether the forecast probability of events (e.g. a warm summer in central Europe) matches the observed occurrence of those events. These evaluations are provided together with forecast products, yet the reference dataset might vary from one service provider to another, and hence a re-evaluation of forecasting skill is

usually conducted. The quality of observations used for model initialisation is a key component for the quality of the forecasts and re-forecasts. As a consequence, re-forecasts for the period before the 80s are not generally used to assess sub-seasonal and seasonal forecast systems and most state-of-the-art service providers create from 10 to 40 years of re-forecasts. This lack of long record of past data limits the number of (sub-) seasons available to more thoroughly and robustly assess the forecasting service.

## 2.2. Sectorial applications

### 2.2.1. Hydro

Reservoirs used for hydropower generation have commonly been constructed at mountainous basins, which are highly influenced by snow processes (accumulation and melting). In order to maximize the energy (GWs) production, the river water is regulated particularly during the snow melting season respecting the environmental laws and meeting the energy needs of the current and coming season. Information on the snowmelt runoff and volumes is of high importance particularly in river systems influenced by multi-reservoir operation that a multi-optimization system is required. Seasonal forecasts of snowmelt inflows (runoff volumes entering the reservoir) together with ground information (e.g. from in situ measurements and/or information from earth observations) are key inputs to the optimization models operated by the hydropower companies and hence important for decision making, when planning the production for the current and coming winter season.

A common approach to provide long-range information for decision-making is based on climatology. In particular for hydropower, reservoir operators are interested in accumulated forecasts of inflows over the spring flood period, which is from April to July. As an example in Sweden, seasonal forecasts for the April-July accumulated runoff are issued once a month from January until the start of the melt season in April. The targeted objective is related to the reservoir level at the end of summer, when a trade-off is sought between water usage for hydropower production during the spring period and high water levels at the end of the summer is present. Therefore, release of water from the reservoir that is not used for energy production is considered as a loss of production and hence economic value. An example of reservoir water spilling is when the remaining spring flood volume is underestimated and reservoirs are filled up too early.

### 2.2.2. Wind

Wind power generation relies on wind, which is an intermittent source. Prediction of the variability of wind energy resources has been identified as a challenge to the grid integration of wind energy systems (Najafi et al. 2016; Füss et al. 2015). During periods of low wind the electricity demand has to be met with other types of energy production. It is therefore very important for energy producers to have an estimate of wind production in order to ensure a good planning of other sources. Moreover, the energy mix determines the price of electricity, with the variability of wind being an important factor. Currently, the wind industry exploits short term predictions; 6h - 2-3 days forecast are used by transmission system operators for power system management (scheduling reserves, planning, congestion management); 2-3 days up to a week are used by energy producers for operations & maintenance planning of



wind farms. The increase in skill in sub-seasonal and seasonal time scales opens possibilities for using probabilistic predictions for operations and logistics substituting the present approach (based on climatology).

### 2.2.3. Solar and Demand

Given the necessity of ensuring the balance between electricity production and demand, an accurate estimation of future weather state could improve the efficiency and reliability of energy management at local and national scales. In fact, weather is a crucial element both for the generation and demand of electricity. The relationship between temperature and demand is well known and it has been already investigated in many works focused on Europe. Valor et al. (2001) and Pardo et al. (2002) first recognized the strong coupling between electricity demand and temperature. Furthermore Bessec and Fouquau (2008), analysing 15 European countries over twenty years, put the emphasis on the increasing sensitivity of electricity demand with respect to temperature during the recent years. The effectiveness of seasonal climate forecasts for electricity demand forecasting has been analyzed both considering deterministic and probabilistic approaches by De Felice et al. (2015a) which demonstrated the potential of the use of seasonal climate forecasts for electricity demand over Italy in boreal summer.

The main problem concerning the use of solar energy is its continuous spatial and temporal variability. Power generation of photovoltaic (PV) can fluctuate significantly with negative consequences on the management of the electrical distribution and stability and on the kWh production costs (Antonanzas et al., 2016). Solar power generation depends on the amount of solar radiation available, making the power supply particularly vulnerable to clouds and to the occurrence of low-pressure systems. In addition, solar power generation is influenced by other atmospheric variables related to the efficiency of the PV panels, i.e. the efficiency of the photovoltaic cells diminishes as their temperature increase. Therefore, for an appropriate management of solar power, it is essential to have reliable and accurate information about weather/climate conditions that affect the production of electricity. Short term forecasting (hours to few days) of solar power generation has been considered necessary for the scheduling of non-renewable power plants and or decision-making processes within the energy market (De Felice et al., 2015b). However, information about the incoming months can be relevant to support and inform operational and maintenance activities as well as to help secure future power contracts. Climate forecasts at the sub-seasonal to seasonal time scale can play an important role in supporting long-term decision-making processes.

### 2.2.4. Country-scale supply demand balance indicators

A key feature of current large-scale power systems is that the demand for electricity must be met by electricity generation on a near instantaneous basis. Traditionally, this is achieved by "scheduling" a set of generation resources (typically coal, gas and nuclear power stations) to meet a forecasted demand. The growing use of renewables (particularly wind and solar-PV), however, leads to new challenges. The generation output from renewables is determined by meteorological conditions, and thus cannot be controlled to the same extent as the

generation from traditional power plants. Both demand and generation-potential therefore contain strongly weather-sensitive components.

The impact of a shift to weather-sensitive generation has implications not only for the owners/operators of renewable resources, but across the power system. It is beyond the scope of this document to discuss the complex and detailed mechanics of specific power systems or markets and the issues surrounding renewables integration in detail. The interested reader is referred to, e.g., Stoft (2002) and Green (2005), for a more comprehensive but still introductory discussion.

The marginal output costs associated with renewable generation are typically very low in contrast to traditional generators (which require costly fuel to operate and produce carbon emissions). As such, it is almost always preferable for renewables to operate at their maximum possible output, with traditional resources then making up the remaining generation output to meet demand. Thus, the volume of electricity required to be produced by traditional power generators corresponds not to demand itself, but to the *residual demand* once renewable generation is accounted for. Therefore, although traditional generators are not themselves directly exposed to day-to-day weather risk, they are indirectly exposed to weather via its impact on both demand and renewable resources. In a power-system where wind-power is the dominant renewable, this is often referred to as “demand-net-wind”.

As noted above, the volume of generation required from traditional generators can be strongly impacted by weather through both renewable resource availability and demand. In a market setting, this has consequences for pricing. For renewables, the low marginal cost of generation leads to “price taking” behaviour: it almost always makes economic sense for a renewable generator to produce the maximum output possible. The price for power is therefore set by other generators where non-zero marginal costs mean that they are only willing to generate up to the level where the electricity price matches the marginal cost of production. If demand is assumed to be price-insensitive, then the power price is typically set by the most-expensive traditional generator required to meet demand, in a process referred to as the “merit order” curve. In the case of a power system with renewables, one can consider the demand to be replaced by the *residual demand* once renewable generation is removed. The price received by all generators (including renewables) is thus strongly influenced by the residual demand. For a power system where the dominant renewable is wind-power only, residual demand can again be considered as demand-net-wind.

The present description is an illustration of extremely complex processes, the details of which vary across Europe. Moreover, it is noted that many of the market-processes described above often occur on shorter time-intervals than the daily-, weekly- and monthly-averages presented in the forecasts described in this Deliverable. Nevertheless, skillful forecasts of country-aggregated demand, renewable generation and – crucially – residual demand (once renewable generation is removed) are believed to provide valuable contextual information to a variety of energy system stakeholders: from individual traders, power plant operators and owners to national transmission system operators.

For simplicity, this Deliverable focuses on forecast skill for country-aggregate wind, demand and residual demand-net-wind (i.e., solar PV generation is not considered). This is consistent with the analysis presented in Deliverable D3.2, which suggests that wind (and therefore wind-power) variability dominate over solar PV. For brevity, the main body of the text presents results for January and July only using two skill-scores (ensemble mean correlation and CRPSS), though additional months/metrics are provided in Annex 2.

## 3. Datasets and Models

### 3.1. The forecasting systems

#### 3.1.1. Seasonal time-scale

As benchmarking for seasonal forecasting we selected the latest long-range forecasting system from ECMWF, known as SEAS5. SEAS5 is the fifth generation system for seasonal forecasting from ECMWF, including a number of upgrades in comparison to the previous System4, in particular in the ocean model, atmospheric resolution, and land surface initialization. SEAS5 uses IFS (Integrated Forecasting System) Cycle 43r1 and represents six years of IFS development in terms of physics, modules, new Earth system components and initialization methods. SEAS5 forecasts show substantial improvements in the tropics, in particular for sea-surface temperature in the equatorial Pacific. SEAS5 is upgraded only occasionally, at intervals of four to six years, since the resources needed to complete the large re-forecast sets required for calibration are large, whilst a slow refresh cycle offers users a more stable service.

SEAS5 uses the community ocean model NEMO (Nucleus for European Modelling of the Ocean). The resolution is  $0.25^\circ$  and 75 layers (ocean model configuration ORCA025z75). The vertical resolution is particularly high in the uppermost part of the ocean (18 levels), whilst the high horizontal resolution improves the representation of sharp fronts and ocean transports. An important innovation in SEAS5 is the inclusion of prognostic sea-ice. The sea-ice model is LIM2, part of the NEMO modelling framework. The prognostic sea-ice model allows sea-ice cover to respond to changes in the atmosphere and ocean states. The intention is to capture inter-annual variability and trends in the sea-ice cover.

SEAS5 ocean and sea-ice initial conditions are provided by the new ocean analysis and reanalysis ensemble (ORAS5), which uses the same ocean model and sea-ice as the coupled forecasts in SEAS5 and is driven by ocean observations from floats, buoys, satellites and ships. Compared to its predecessor ORAS4, which was used for System4 (the predecessor of SEAS5), ORAS5 has higher resolution and updated data assimilation and observational datasets. It provides sea-ice initial conditions by assimilating sea-ice concentration. In addition, a perturbation scheme is used to generate the ensemble of ocean re-analyses, consisting of perturbations to the assimilated observations (both profiles and surface observations), and perturbations to the surface forcing fields.

Horizontal resolution in the atmospheric component of SEAS5 is also significantly high (TCo319), and the corresponding grid-point resolution is 36 km. Although the spectral resolution increase is less dramatic than the change in grid-point resolution, the ability of the cubic grid to better represent the smallest spectral scales and the energy within them more than makes up for this. The wave model resolution is  $0.5^\circ$  to maintain a match to the atmosphere. The vertical resolution remains at L91.

For SEAS5 a new offline recalculation of the land surface initial conditions was made, at the required TCo319 resolution and with a revised precipitation forcing. Comparison of test forecasts made using this dataset with others using the operational analysis for a recent overlap period showed a generally very good degree of consistency, while also demonstrating the superiority of the operational analysis in terms of the impact on 2 m temperature forecast anomalies. For more details see Table 1.

Characteristics	ECMWF SEAS5
IFS Cycle	43r1
IFS horizontal resolution	TCO319
IFS Gaussian grid	O320 (35 km)
IFS vertical resolution (TOA)	L91 (0.01 hPa)
IFS model stochastic physics	3-lev SPPT and SPBS
Ocean model	NEMO v3.4
Ocean horizontal resolution	ORCA 0.25
Ocean vertical resolution	L75
Sea ice model	LIM2
Atmosphere initialization (Re-forecast/Forecast)	ERA-Interim/Operations
Land Initialization (Re-forecast/Forecast)	ERA-Interim land (43r1)/Operations
Ocean initialization	OCEAN5
Forecast ensemble size	51 (0-7m) 15 (8-13m)
Re-forecast years	36 (1981-2016)
Re-forecast ensemble size	25 (0-7m) 15 (8-13m)
Calibration period	1993-2016

**Table 1: Characteristics of ECMWF SEAS5 seasonal forecasting system.**

### 3.1.2. Sub-seasonal time-scale

With regard to sub-seasonal forecasting, the ECMWF ENS-ER system was selected as benchmark. ECMWF issues the ENS-ER twice a week (Monday and Thursday) by extending the integration time beyond day 15 up to day 46. The monthly forecasting system is run 51

times from slightly different initial atmospheric and oceanic conditions, which are designed to represent the uncertainties inherent in the operational analyses. Each ENS-ER run comes with an 11-member hindcast set produced for the same dates as the forecast date over the previous 20 years. For instance, if the first starting date of the real-time forecast is 27 March 2013. The corresponding re-forecast is an 11-member ensemble starting on 27 March 2012, 27 March 2011, ..., 27 March 1993. The 11-member ensemble is thus integrated with 20 different starting dates.

This hindcast set provides identical integrations as the current operational forecast with the difference that ERA-Interim reanalysis (ERA-Interim; Dee et al., 2011) and ERA-Interim land reanalysis (Balsamo et al., 2015) are used to provide the initial conditions, whereas the operational ensemble forecast uses the operational analysis (Wetterhall and Di Giuseppe, 2018).

In terms of atmospheric initial conditions, the real-time forecasts are initialized from the operational analysis, whilst the re-forecasts are initialized from ERA-Interim, except for the soil initial conditions (soil temperature, soil moisture, snow initial conditions) that are provided by an offline soil reanalysis. As for the oceanic initial conditions, these are provided by the real-time suite of NEMOVAR. As a summary Table 2 presents comparatively the characteristics of SEAS5 and ENS-ER.

System	Time res.	Spatial res.	Horizon	Ensemble size	Issue frequency	Hindcast set	Hindcast ensemble size
ENS-ER	6 h	18 / 36 km	46 days	51	Twice weekly	20 yrs	11 members
SEAS5	6 h	36 km	7 / 13 months	51	Monthly	36 yrs	25 members

**Table 2: Characteristics of the climate models used for sub-seasonal and seasonal forecasting.**

### 3.2. The hydrological model

For the water-related applications that require hydrological modelling, e.g. hydro indicators, we are benefited from a setup of the HYPE hydrological model at the continental scale (Figure 1). The HYPE model has been setup for the pan-European region (8.8 million km<sup>2</sup>) (Hundecha et al., 2016) with a spatial resolution of about 35400, i.e. in average 215 km<sup>2</sup>, and is referred to as E-HYPE v3.0.



**Figure 1: Domain of the E-HYPE hydrological model.**

The model runs at a daily time step and its setup is based on open data (see Table 3). River networks and sub-catchments were delineated using WWF's Hydrosheds data (Lehner et al., 2008). Hydrological response units (HRUs) were derived from land use and soil data obtained from different sources. Land use was derived from the CORINE land use data (Bartholomé et al., 2002). Lakes and reservoirs were extracted from GLWD (Lehner and Döll, 2004) and Grand (Lehner et al., 2011) data sets respectively. Irrigated areas were identified from GMIA (Siebert et al., 2010) and MIRCA (Portmann et al., 2010) data sets. Soil types were derived from the Harmonised World Soil Database (HWSD) (Nachtergaele et al., 2012). Daily discharge data at more than 3000 gauging stations were obtained from different sources. The GFD meteorological forcing dataset (daily mean precipitation and temperature), an observation corrected reanalysis (Berg et al., 2018) for the period 1981 – Today, is used as reference for model calibration and verification. In total, 8 soil types and 15 land use classes were used and based on their combination, a maximum of 75 HRUs were identified. A subset of the gauging discharge stations was used – 115 discharge stations used for calibration while 538 independent stations used for model validation (for details about the selection procedure see Hundecha et al., (2016)).

Characteristic/Data type	Info/Name	Provider
Topography (routing - delineation)	hydroSHEDS (15 arcsec), Hellenic Water Framework Directive (WFD)	Lehner et al. (2008), Hellenic Water Framework Directive (WFD)
Soil characteristics	Harmonised World Soil Database	Nachtergaele et al. (2012)
Land use characteristics	CORINE	Bartholomé et al. (2002)
Reservoir and dam	Global Reservoir Dam database	Lehner et al. (2011)
Lake and wetland	Global Lake and Wetland Database	Lehner & Döll (2004)
Irrigation	Global Map of Irrigation Areas	Siebert et al. (2010)
Discharge	GRDC, EWA and others (2690 stations)	<a href="http://www.bafg.de/GRDC">http://www.bafg.de/GRDC</a>

Precipitation (calibration and initialisation)	Global Forcing Dataset (0.5° x 0.5°)	Berg et al. (2018)
Temperature (mean, min, max) (calibration and initialisation)	Global Forcing Dataset (0.5° x 0.5°)	Berg et al. (2018)
Snow cover area	GlobSnow	<a href="http://www.globsnow.info">http://www.globsnow.info</a>

**Table 3: Data sources and characteristics of the E-HYPE setup.**

The model was calibrated to secure usefulness to the potential users and applications; hence an adequate model performance in terms of discharge and other hydrological variables is important. EHYPE v3.0 parameters were calibrated using observed discharge stations, in-situ data and satellite data. Yet here we focus on the results using discharge data. The implemented regionalization scheme is found to allow transferability of parameters from a limited set of calibration stations to other locations without a need to calibrate a large number of catchments in a multi-basin setup of large scale modelling and to further enable modelling in ungauged catchments. The model performance in catchments that are not used for the derivation of the regional parameters is comparable with that of the catchments used for model calibration and parameter regionalization. Parameters are linked to catchment descriptors with good transferability, with median Nash-Sutcliffe Efficiency (NSE; Nash and Sutcliffe, 1970) of 0.54 and 0.53, and median volume error of -1.6% and 1.3% in the calibration and verification stations, respectively.

Note that in this investigation, bias-adjusted seasonal meteorological forecasts from ECMWF ENS-ER and SEAS5 are used to force the E-HYPE hydrological model. The reference dataset is the HydroGFD, which is a global data set of adjusted reanalysis data for daily minimum, mean and maximum temperature, as well as daily mean precipitation. It was constructed mainly for large hydrological modelling, and therefore has the main aim of improving hydrological water balance simulations.

### 3.3. Demand and wind-power models

Two “weather to energy” models are used to convert gridded meteorological data to corresponding estimates of country-aggregate demand and wind power. Each model converts meteorological data to energy at the highest possible time-and-space resolution (sub-seasonal - 1.5° resolution daily (00Z) snapshots for wind, daily mean temperature for demand; seasonal – 1° resolution 6-hourly snapshots for wind, daily mean temperature for demand). A full description and validation of the demand and wind-power models is provided in Deliverable D3.2, though in this case the methods are applied to the ERA5 reanalysis rather than MERRA-2. As such, only a brief recap for each model is provided here.



**Demand:** A multiple linear regression model is trained on recorded daily demand in each country (2016-2017) as follows:

$$Demand(t) = \alpha_1 LTT(t) + \sum_{i=2}^7 \alpha_i Weekday(t) + \sum_{i=8}^9 \alpha_i Weekend(t) + HD(t) + CD(t)$$

Where HD and CD correspond to heating and cooling degrees (days where daily-mean temperature is below 15.5 °C and above 22 °C respectively) calculated on a country-average basis. The remaining terms correspond to dummy variables used to model human behavioral patterns such as long-term trends in GDP or technology (LTT) and day-of-week effects. Once the model is trained, these dummy variables are all discarded (i.e., set to constant or zero values). This reduced model is then applied to the whole of the ERA-5 reanalysis period, producing a synthetic demand time-series – dependent only on weather (rather than human behavioral patterns or long term socio-techno-economic trends).

**National wind power generation:** A physics-based model is used to produce estimates of wind power. Grid-point near-surface (10m) wind speeds are first scaled using a power law to a typical turbine height (100m), assuming a constant exponent of 1/7. The resulting wind-speeds are then converted to wind-power capacity factors using an appropriately selected wind-turbine power curve (the selection of power curve is based upon the prevailing long-term mean wind speed in each grid box, such that less windy grid boxes have turbines installed which respond better at lower wind speed). The resulting capacity factors are then multiplied by the estimated installed capacity in each grid box and aggregated over the country domain.

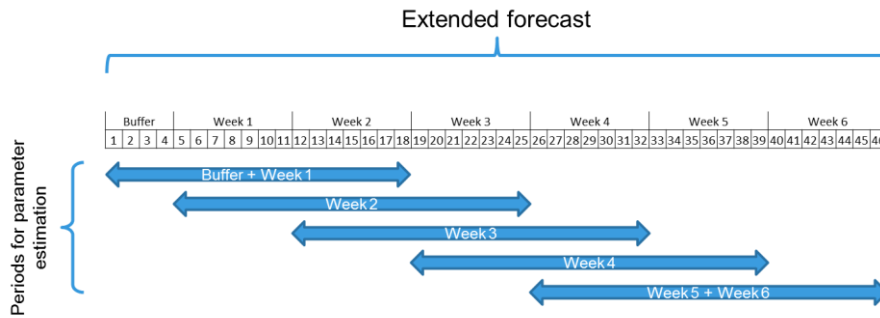
## 4. Methodology and evaluation protocol

### 4.1. Climate models and bias-adjustment

Benchmarking was set based on two different systems, the ECMWF Extended Range (ENS-ER; available at 1.5 degrees) and the ECMWF SEAS5, provided from the Copernicus Data Store (available at 1 degree) for the sub-seasonal and seasonal investigation respectively. The period of re-forecast data availability varies between the two systems. For the ENS-ER the period 1996-2015 (1999-2010 for hydro) was used as hindcast and is associated to the forecasts launched during the year 2016, whilst for SEAS5 the period 1993-2015 was used as hindcast. Due to differences in experience between the partners and familiarity with the bias-adjustment approaches, different bias-adjustment methods were applied depending on the investigation.

#### 4.1.1. Hydro bias-adjustment protocol

For the hydro applications, the forecasts are bias-adjusted using a modified version of the Distribution Based Scaling (DBS) method (Yang et al., 2010) to account for drifting conditioning the bias adjustment on the lead time. It was adapted from the quantile-mapping method for application in (sub-) seasonal forecasting. Bias-adjustment is conducted on all members of SEAS5 and ENS-ER using the hydroGFD data (Berg et al., 2017) as reference. After bias adjustment, the cumulative distribution of daily precipitation and temperature values closely follows that of the hydroGFD data. In brief, DBS aims to map the quantile distributions of precipitation and temperature in the forecasts to those of the reference data. For precipitation, a two-step procedure is applied: 1) adjustment of the wet-day frequency by applying a wet-day cut-off threshold (in case of wet frequency bias) or by adding wet-days to pre-existing wet-spells (if dry frequency bias), and 2) quantile-mapping of the precipitation data using a double-gamma distribution to accurately represent both normal and extreme precipitation intensity ranges. For temperature, a quantile-mapping adjustment based on a Gaussian distribution is used. The temperature adjustment model is dependent on the wet/dry state of the corresponding precipitation. This means that DBS takes into account different biases on wet and dry days (see details in Yang et al., 2010). For SEAS5 the adjustment was done for each month; however for ENS-ER a 3-week moving window was used in the adjustment and the estimated BA parameters were applied in the middle week of the window (see Figure 2). The biases prior and after the bias-adjustment of precipitation and temperature are presented in Annex 1.



**Figure 2: Bias-adjustment procedure for the sub-seasonal forecasts.**

### 4.1.2. Wind speed bias-adjustment protocol

The variance inflation bias-adjustment method, which was used here, produces predictions that will have interannual variance that is equivalent to that of a reference dataset. The method is described in Doblas-Reyes et al. (2005). It was tested against a simple bias correction for the case of wind in Torralba et al. (2017). This method was also tested for temperature and precipitation in Manzanas et al. (in review). The inflation of the ensemble spread ensures that predictions will have reliable probabilities.

The evaluation is performed on the hindcasts corresponding to the forecasts issued in 2016, therefore spanning the period 1996-2015. Hindcasts are issued on Mondays and Thursdays. The anomalies of weekly averages are computed per start-date taking all years and members (20x11). Skill scores are computed separately each grid point and forecast time: for days 5-11, days 12-18, days 19-25 and days 26-32. In order to obtain a long enough series of pairs of corresponding values of model and reference to produce robust verification scores, start dates of each month were combined. This way, each monthly verification measure is calculated from 80-100 pairs of hindcast and reference values, resulting from the 4-5 weekly start dates. The fair version of CRPSS, FairCRPSS (Ferro, 2014) is used since it compensates for the effect of the number of members on the score because it rewards ensembles with members that behave as if they and the verifying reference are sampled from the same distribution. The weekly averages were calibrated using the variance inflation method (Doblas-Reyes et al., 2005; Torralba et al., 2017). This method ensures that the predictions will have interannual variance that is equivalent to that of a reference dataset.

### 4.1.3. Supply-demand bias-adjustment protocol

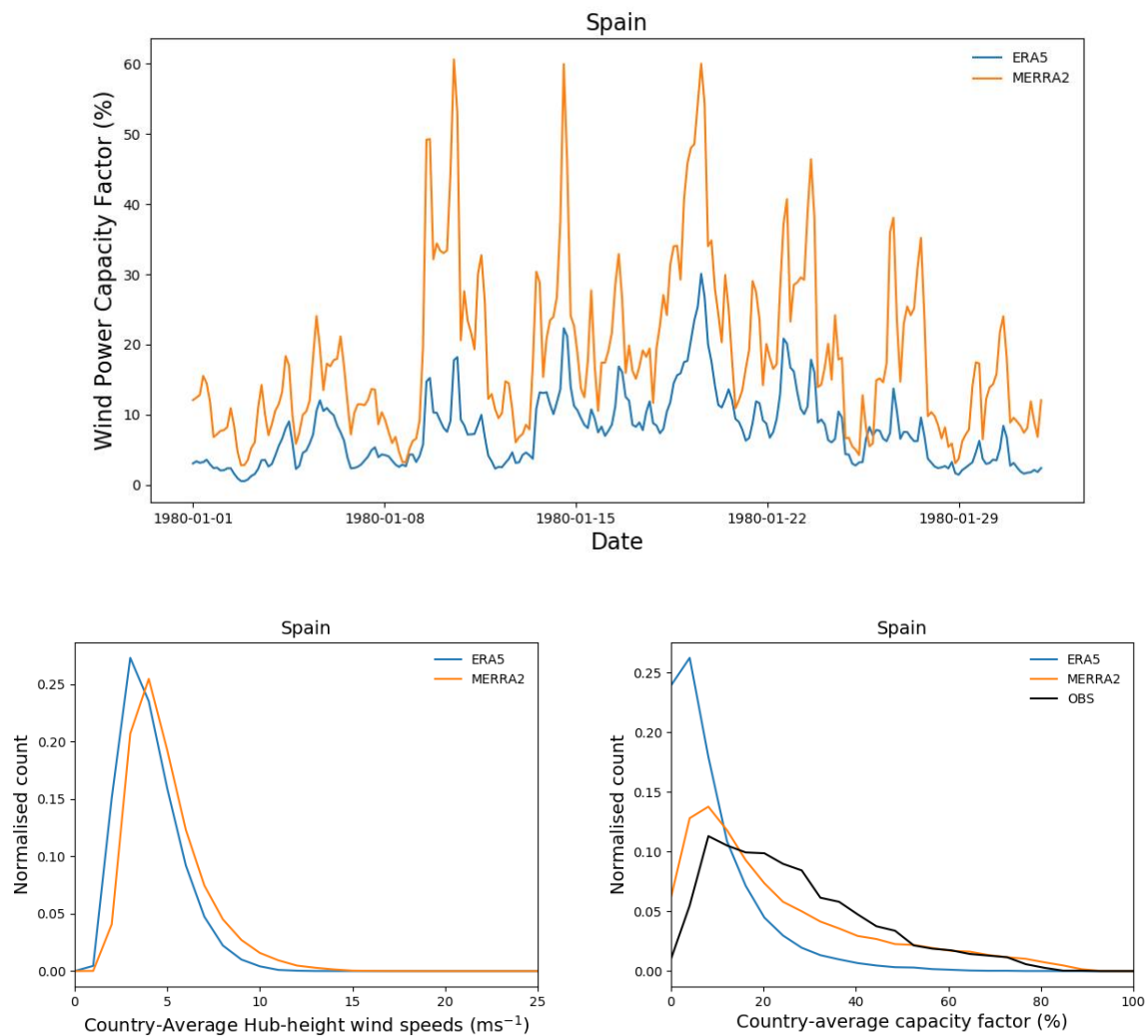
Sub-seasonal forecast from the ECMWF monthly forecasting system have been analysed. For the country-level skill indicators, hindcasts of the model with 11 ensemble members, at 1.5x1.5 degree resolution have been analysed. All hindcasts matching forecasts initialised during calendar year 2016 are included in the analysis and are downloaded from the S2S project website. Note that this means that three separate model cycles are included in the analysis. For hindcasts in January and February and prior to March 8<sup>th</sup>, CY41R1 is used. For hindcasts from March 8<sup>th</sup> to November 22<sup>nd</sup>, CY41R2 is used. For hindcasts from November

22<sup>nd</sup> and all of December, CY43R1 is used. ERA5 data from 1<sup>st</sup> January 2000 is used as the reference value. Model bias is calculated as a function of start date and lead time using a leave-one-out approach for each year and subtracted from forecasts further to prior processing. Bias corrected forecasts are then used to calculate energy indicators using the models described above. A second bias correction is applied to the country-level indicators, comparing against their best estimates from the native resolution ERA-5 dataset. Verification statistics are calculated using the easyVerification package.

Seasonal hindcasts from the ECMWF System5 seasonal forecasting system (SEAS5) were considered for the period January 2000 to December 2016. The data was considered at a one-degree resolution and the full 25-member ensemble was used in the analysis. Daily mean temperature at 2m was used to forecast energy demand and 6-hourly wind speeds at 10m to forecast wind power. In each case, a simple bias correction was performed to remove the ensemble-mean lead-dependent mean bias from each forecast start month, using a leave-one-out approach on the start years (with ERA5 data used as the reference value). These bias-corrected climate variables were then used to force the energy models and produce daily mean demand and wind power at the country level. These were subsequently aggregated to monthly values and a second bias-correction is applied (analogous to the first one, but using the best estimates of the equivalent energy indicators derived from the native resolution ERA-5 dataset). The verification assessment was performed against these same values using the R packages easyVerification and SpecsVerification. The verification metrics were calculated both for the full 17-year period (2000-2016) and for a reduced 11-year period (2000-2010), though only analysis for the full period is presented here (equivalent analysis for the reduced period is provided in an Annex).

It is emphasised that the energy indicator data derived from ERA5 is assumed to be “truth” for the purposes of the seasonal and sub-seasonal forecast assessment. As both forecast systems are calibrated to ERA5, errors in the ERA5 reconstruction of the energy data will therefore be reproduced in the forecast systems. These errors will, however, not be penalised in any of the skill measures used as the ERA5-derived data is also used for verification.

In general, reanalysis-based reconstructions of energy indicators perform well at reproducing the impact of meteorological variability in the energy system (e.g., see discussion in Deliverable D3.2 and previous work such as Cannon et al 2015). ERA5 does, however, contain significant differences relative to MERRA-2 (used in Deliverable D3.2), which are not yet fully understood: e.g., 10m wind speeds over Iberia tend to be slightly lower in ERA5 compared to MERRA-2, which leads to a much more noticeable difference in the resulting capacity factor estimates (Figure 3). Spanish wind power generation is therefore, in general, strongly underestimated from both the ERA5 “observations” and the forecast systems, even though its time-variations – i.e., the aspects the meteorological forecast seeks to predict – are well captured.



**Figure 3: Comparison of MERRA-2 and ERA5 wind (scaled to 100m) and resultant capacity factors. Top: Sample one-month period of capacity factor. Bottom left: Frequency distribution of hourly country-average wind speeds. Bottom right: Frequency distribution of hourly country-average wind power capacity factors. In each plot, orange line denotes MERRA-2 and blue line denotes ERA5.**

Errors in the average magnitude of any individual power system indicator is likely to have limited impact on forecast skill for each individual power system indicator when assessed by any of the metrics presented below. Misestimates of the average value of any single indicator may, however, have more subtle consequences for compound power system indicators (e.g., demand-net-wind): for example, in Spain an underestimate of the average wind capacity factors, may lead to the role of wind variability in determining demand-net-wind forecast skill being under-represented. This emphasises the need for an in-depth understanding and improvements in the weather-to-energy conversion process and the uncertainties it produces.

## 4.2. ECVs and indicators

The ECVs and indicators are selected based on user requests (see Deliverable 2.1). The ECVs are analysed for the entire hindcast period and presented for the entire European domain, whilst the indicators are presented for the case studies (see Table 4).

ECV/Indicator type	Metadata
<b>Demand</b>	<b>Demand (national aggregate, units GW day):</b> Daily mean demand, averaged over weekly or monthly periods.
<b>Demand-net-wind</b>	<b>Demand-net-wind (national aggregate, units GW day):</b> Daily mean demand-net-wind, averaged over weekly or monthly periods.
<b>Precipitation</b>	<b>Precipitation (mean, mm/day):</b> Calculated as the mean values of daily precipitation averaged over weekly or monthly periods.
<b>Temperature</b>	<b>Temperature (mean, C°):</b> Calculated as the mean annual values of daily mean temperature (at 2m height) averaged over weekly or monthly periods.
<b>Wind speed</b>	<b>Wind speed (mean, m/s):</b> Calculated as the mean values of daily wind flow velocity averaged over weekly or monthly periods.
<b>Wind power</b>	<b>Wind power (national aggregate, units GW day):</b> Daily mean wind power, averaged over weekly or monthly periods.

**Table 4: List of ECVs and indicators requested by users.**

## 4.3. Benchmarking and metrics

Overall performance of the forecasts is evaluated based on the fair version of Continuous Ranked Probability Skill Score (CRPSS), FairCRPSS. FairCRPSS compensates for the effect of the number of members on the score because it rewards ensembles with members that behave as if they and the verifying reference are sampled from the same distribution. Other skill scores (anomaly correlation and fairRPSS for the terciles and the Brier Skill Score (BSS) for the 90<sup>th</sup> and 10<sup>th</sup> percentiles) are included in the Annex. While the CRPSS measures the quality of the whole distribution of ensemble member values, the RPSS measures the quality of the forecast product that is presented as tercile probabilities.

The forecast performance is commonly compared to a benchmark to translate quality into gain or loss. This benchmark can be a previous forecasting system or a common forecast ensemble whose performance is known. Here, we use the simulated climatology as benchmark, which is the historical simulated climatology of the variable of interest. A number of historical time-series (equal to the number of member in the ECMWF SEAS5 and ENS-ER systems) simulated from the targeted forecast date are randomly selected from the model climatology to build an ensemble of possible outcomes.

In terms of the temporal scale analysis, we assess the skill for forecast months 1, 2 and 3 (forecast month 1 is one month after initialization, so for a forecast launched in November, forecast month 1 is December) when using ECMWF SEAS5. When analysing sub-seasonal forecasts, we assess the skill for forecast week 1, 2, 3 and 4 (with the week starting at day 5 after the launch of the forecast, so forecast week 1 is from day 5 till 11).

## 4.4. The case studies

During the user interviews which was part of Deliverable 2.1, respondents were invited to think about anomalies that were of particular interest for them. Eight relevant case studies are the result of experiences they had to manage behind the lack of a skillful S2S decision-support tool. The benchmarking is done for those identified case studies (Section 6). The complete list of case studies is reported in Table 5.

#	Period	Time horizon	Region	Implications
Case 1	17-23 Jan 2017	Sub-seasonal	France, Germany	Wind, demand
Case 2	July 2013	Seasonal	Germany	Demand, solar, hydro, wind
Case 3	30 Aug – 5 Sept 2016	Sub-seasonal	Spain	Wind, demand
Case 4	May-Jul 2015	Seasonal	Sweden	Hydro
Case 5	28 Jan – 3 Feb 2014	Sub-seasonal	Romania	Cold spell, impact wind power
Case 6	Jan – Mar 2015	Seasonal	USA	Wind
Case 7	27 Feb – 5 Mar 2018	Sub-seasonal	Europe / France	Demand
Case 8	Mar 2018	Seasonal	Spain	Wind, solar, hydro, demand

**Table 5: List of case studies.**

## 5. Europe-wide skill assessment

We first analyse the ECVs over the entire European domain in order to understand the limits and emerging patterns of predictability. We expect that patterns are related to Europe's climatic gradient and their predictability will depend on time and lead time, therefore we explore the skill for all initialization weeks/months, and different lead times (from week 1 up to month 3).

The evaluation of monthly predictions is performed on the hindcasts corresponding to the forecasts issued in 2016, therefore spanning the period 1996-2015. Forecasts (and hindcasts) are issued on Mondays and Thursdays. The anomalies of weekly averages are computed per start date taking all years and members (20 x 11). Skill scores are computed separately each grid point (or E-HYPE's subbasin resolution) and forecast time: for days 5-11, days 12-18, days 19-25 and days 26-32. To produce robust verification scores, start dates of each month were combined to obtain long enough series of pairs of hindcast and reference values. This way, each monthly verification measure is calculated from 80-100 pairs of hindcast and reference values, resulting from the 4-5 weekly start dates.

Seasonal predictions from ECMWF SEAS5 are compared to ERA-Interim for wind speed and supply-demand and to hydroGFD for precipitation observations for the period 1993-2015 over Europe. The one, two and three months ahead predictions are verified for each start date, using different metrics.

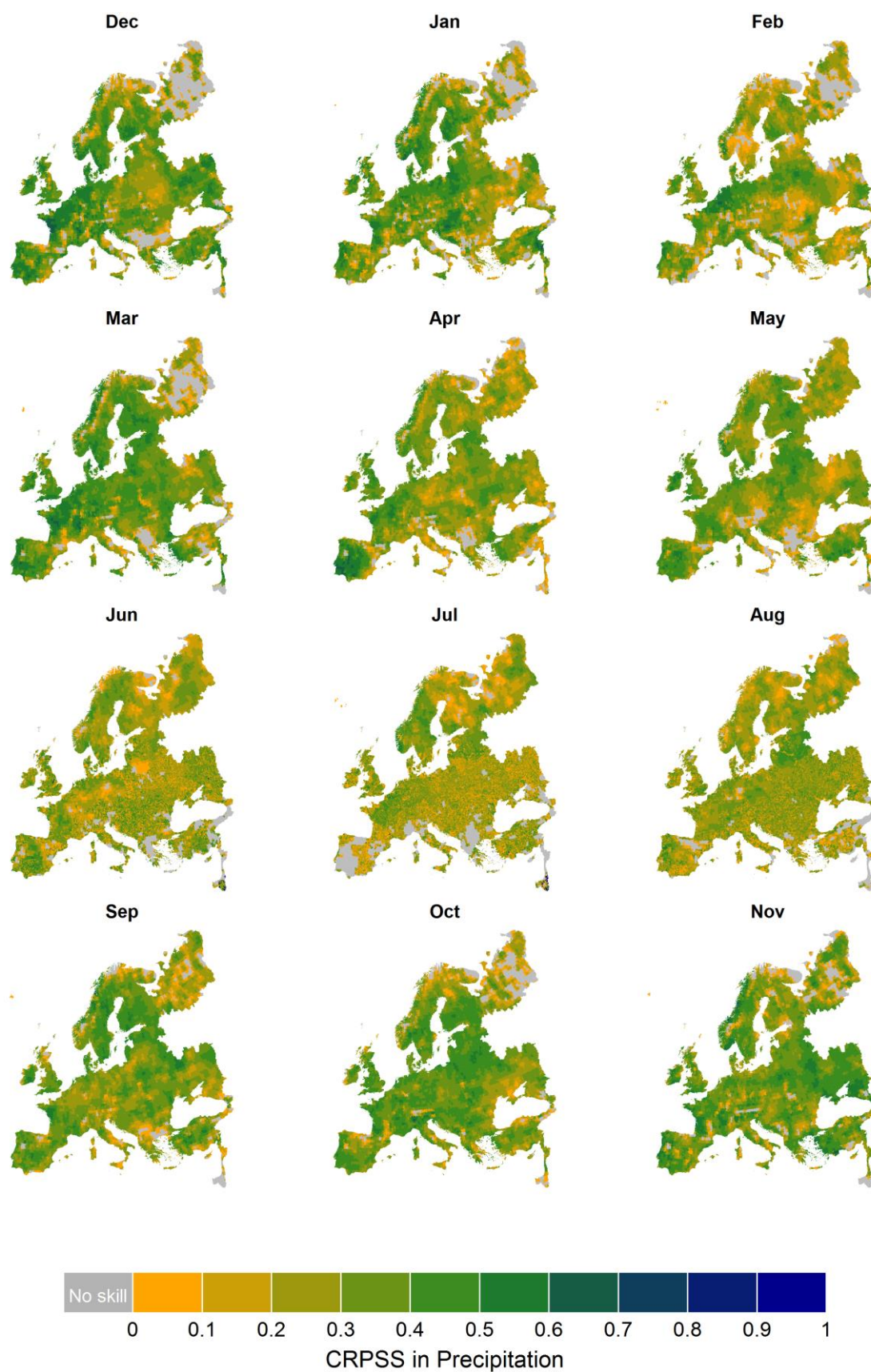
### 5.1. Precipitation

#### 5.1.1. Sub-seasonal time-scale

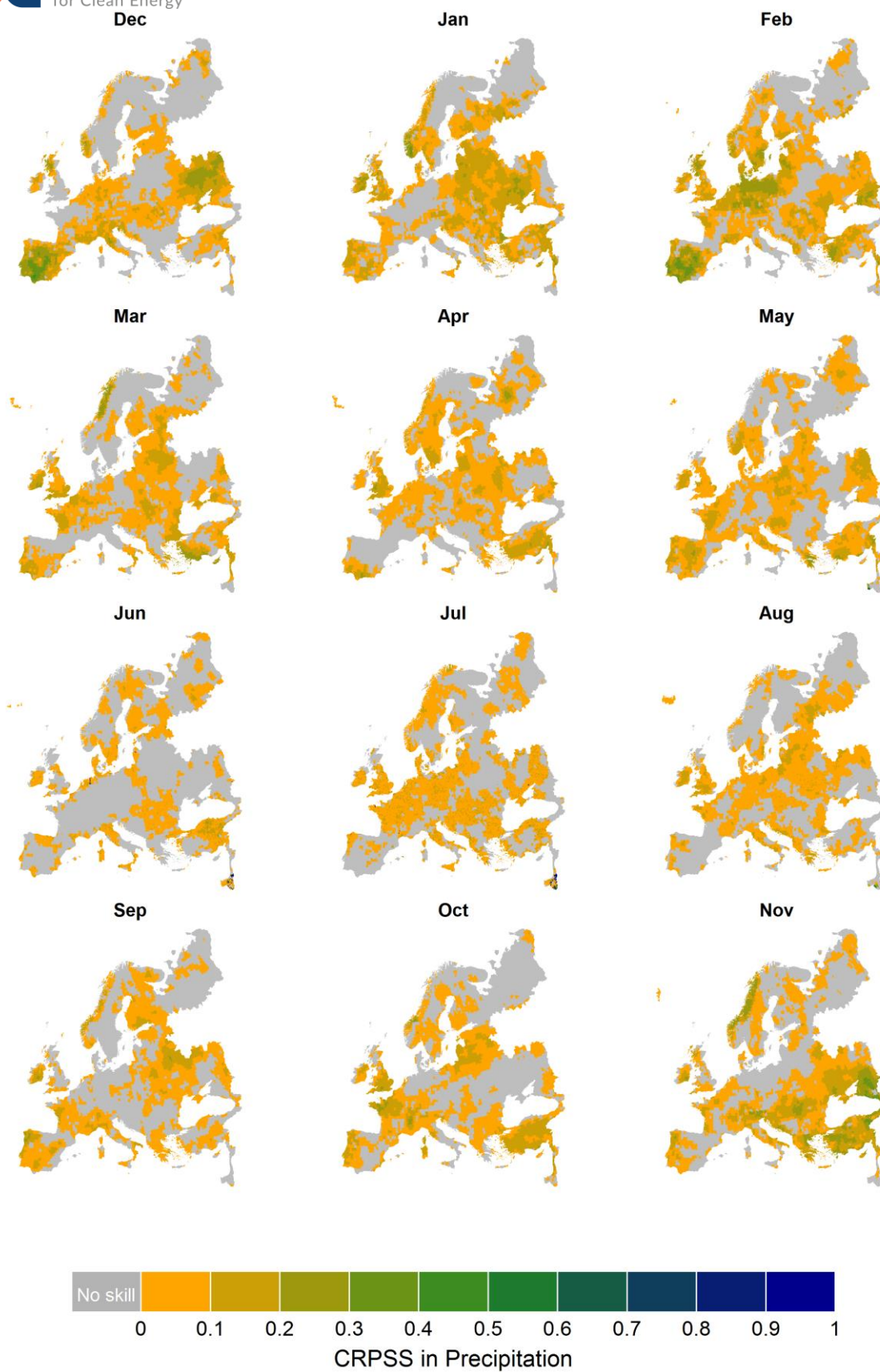
The weekly averages are based on daily aggregation, also using the modified version of DBS bias-adjustment approach to account for drifting conditioning the bias adjustment on the lead time. A 3-week moving window was used in the bias adjustment with the estimated BA parameters being applied in the middle week of the 3-week window.

Figure 4, Figure 5, Figure 6 and Figure 7 present the CRPS scores for all months and forecast weeks (forecast week 1 to forecast week 4 respectively). Overall, the skill varies both geographically and seasonally, with forecasts over most part of Europe being skilful for all months in the 1<sup>st</sup> forecast week. Although forecasts have skill during the summer months (June-August), this seems to be relatively weak in comparison to the other months. Overall the skill further deteriorates rapidly already from the 2<sup>nd</sup> forecast week, with only very few regions indicating skill in forecast week 4. These results indicate that ENS-ER can be important for decision-making up to about 18 days into the future, and later forecasts should be treated with cautious.

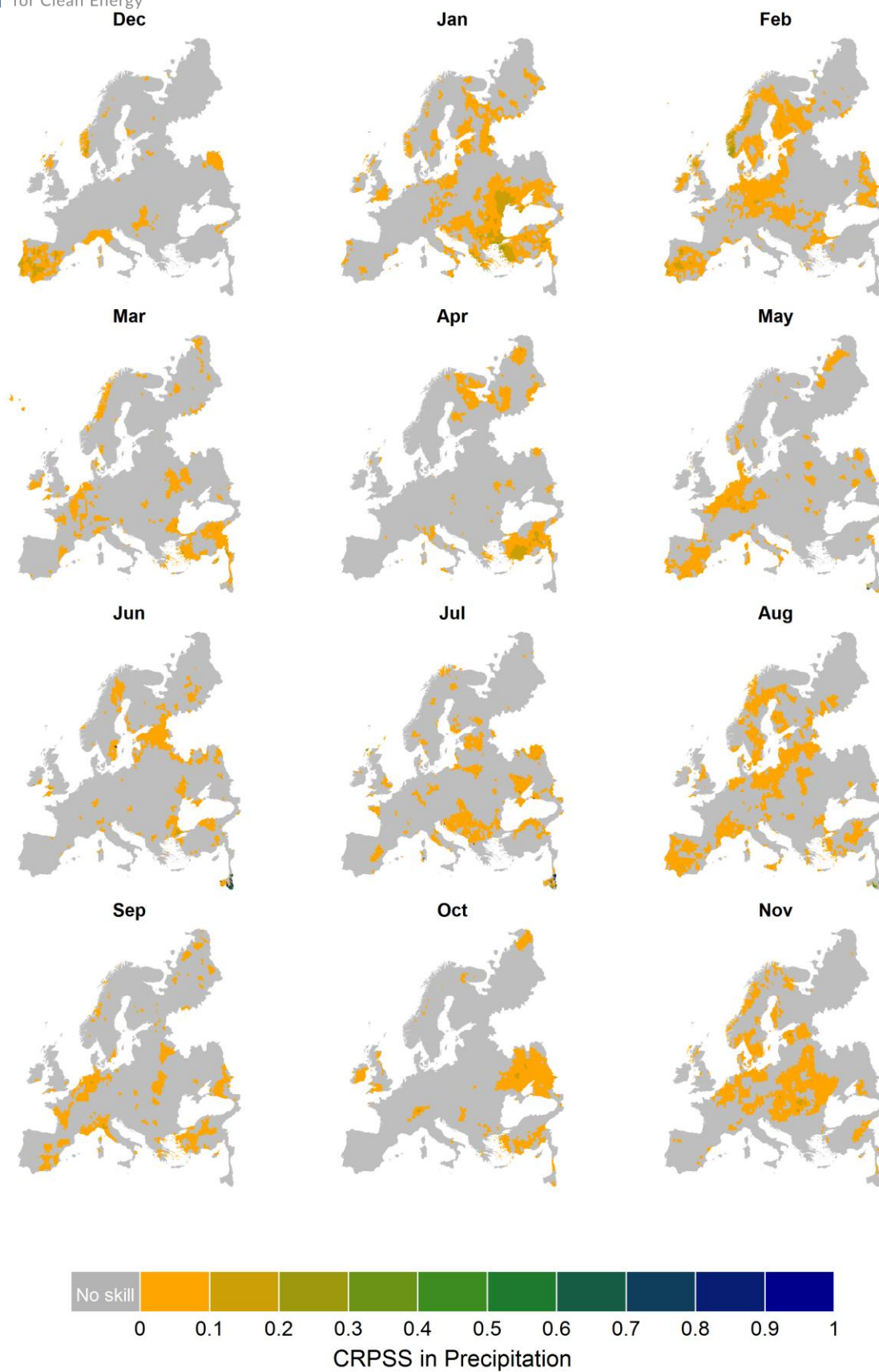




**Figure 4: Forecast week 1 monthly spatial variability of the CRPS score (CRPSS).**

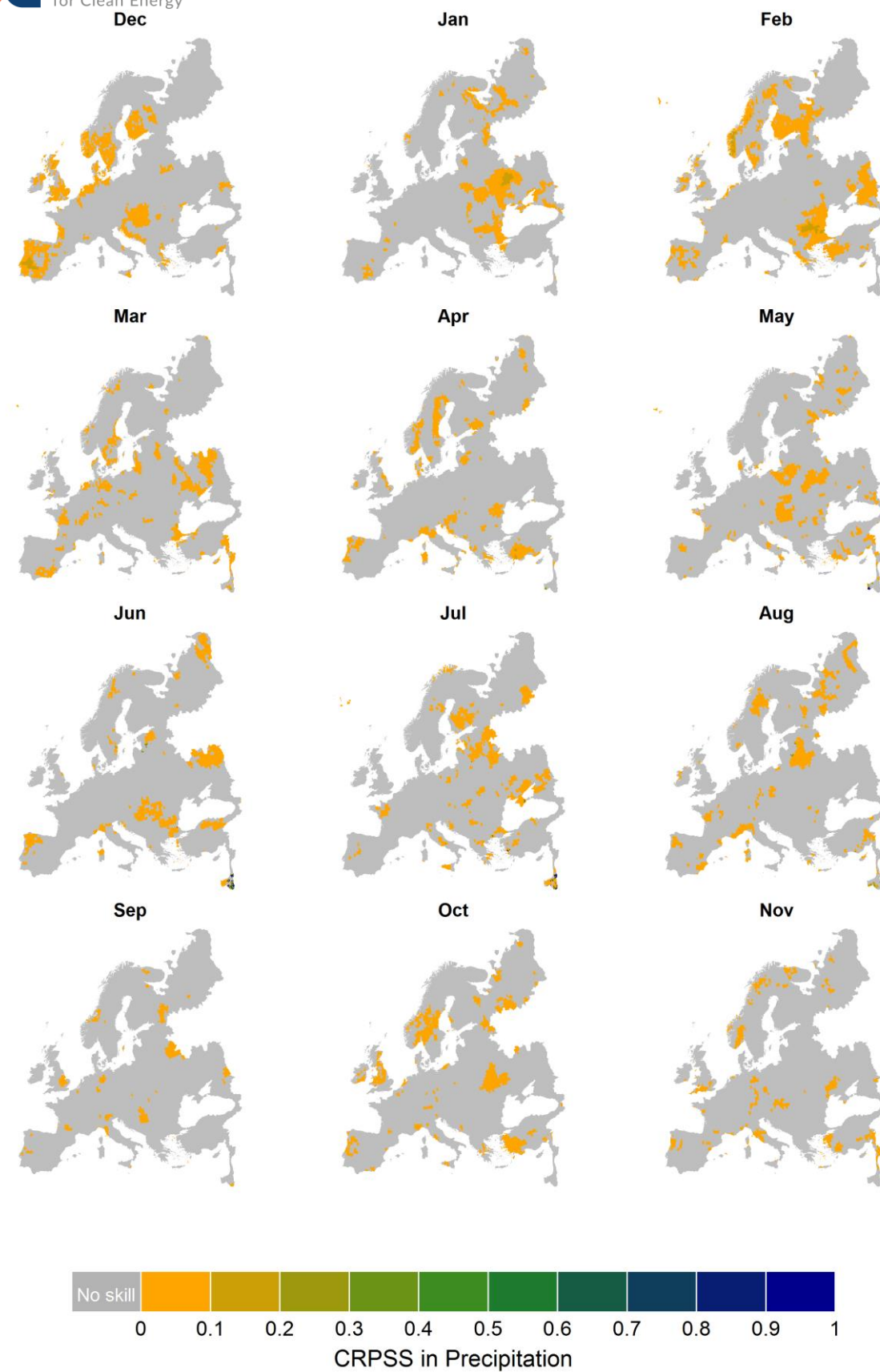


**Figure 5: Forecast week 2 monthly spatial variability of the CRPS score (CRPSS).**



**Figure 6: Forecast week 3 monthly spatial variability of the CRPS score (CRPSS).**





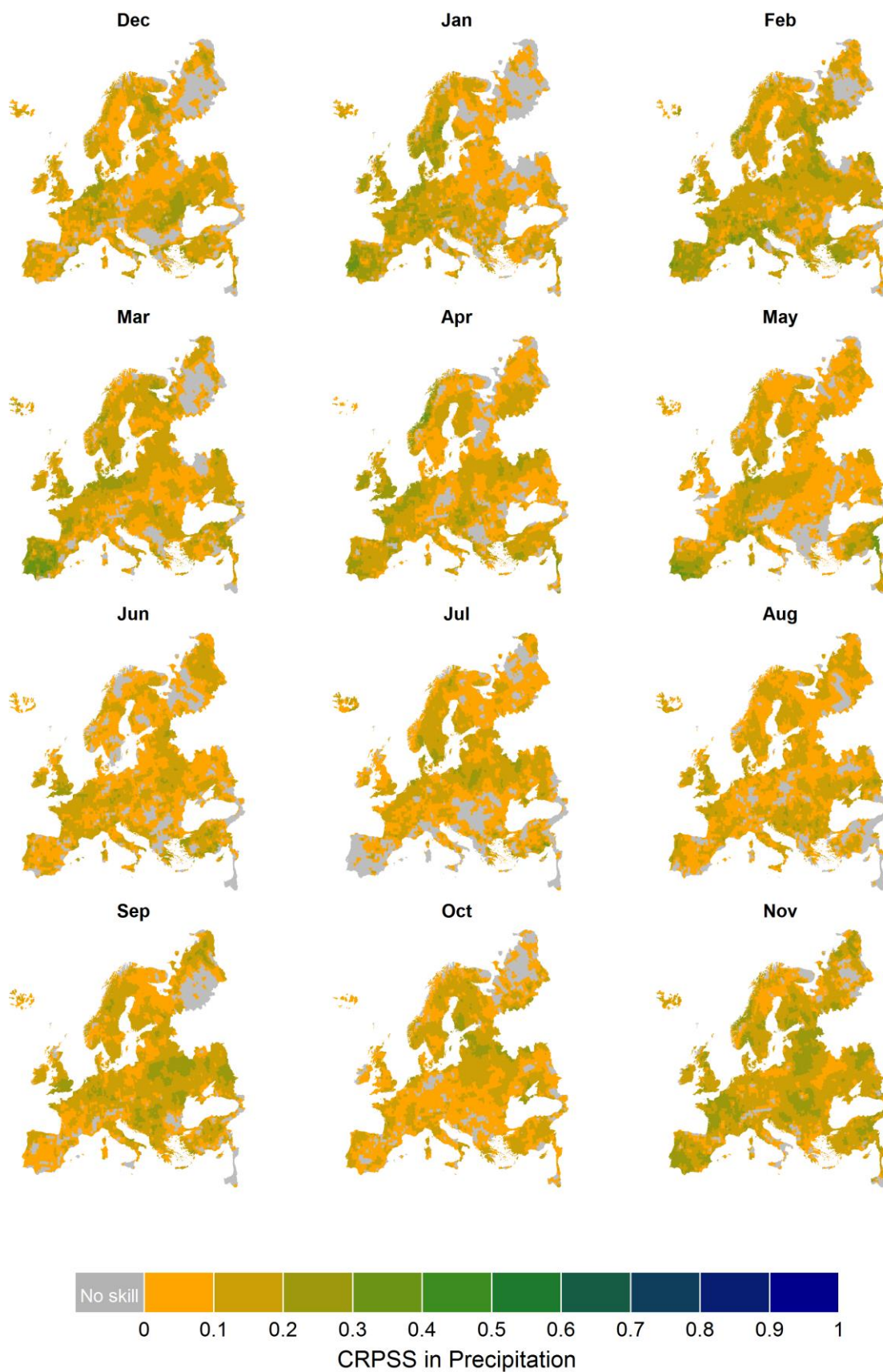
**Figure 7: Forecast week 4 monthly spatial variability of the CRPS score (CRPSS).**

### 5.1.2. Seasonal time-scale

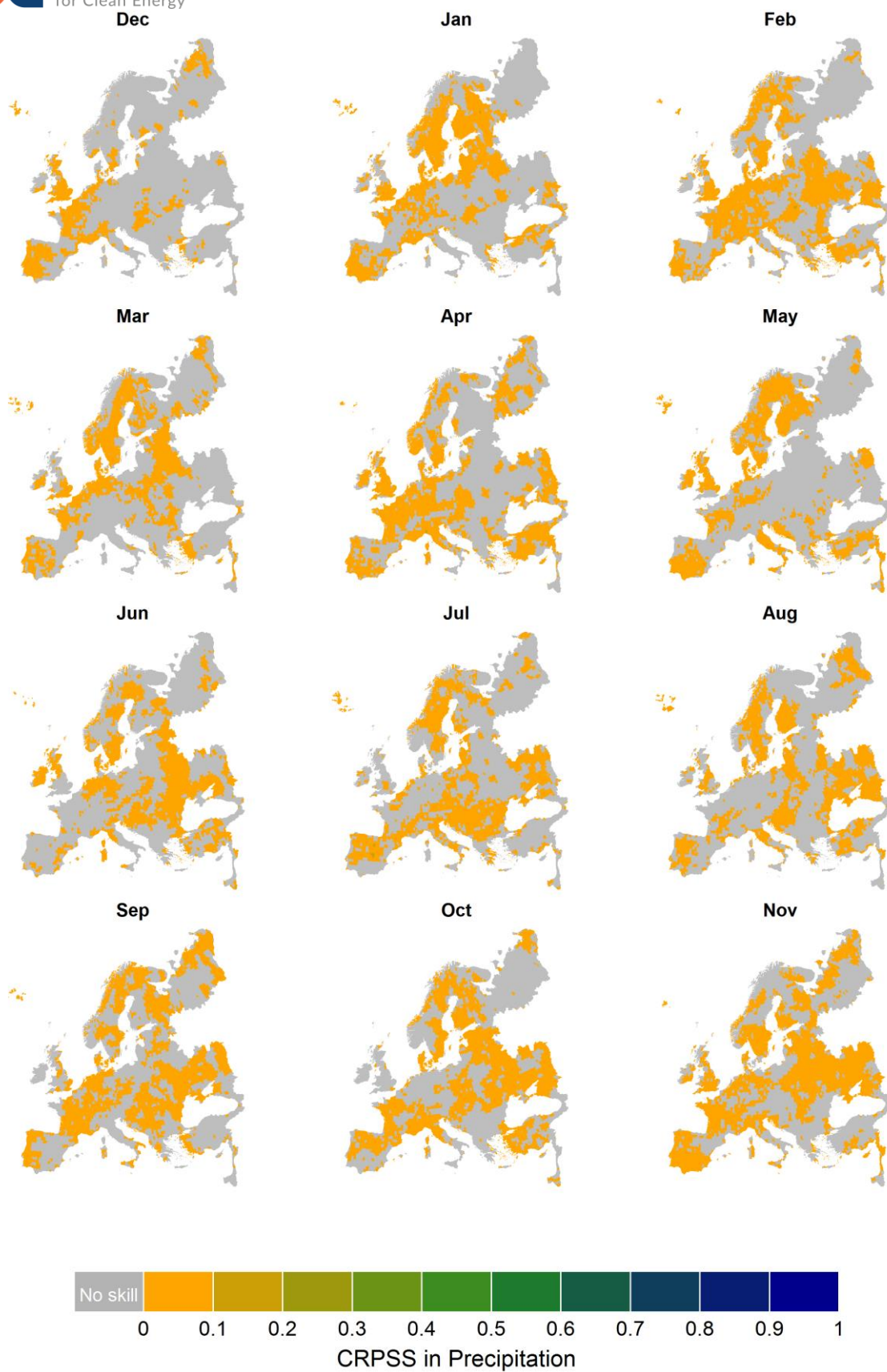
The monthly averages were based on daily aggregation, also using the modified version of DBS bias-adjustment approach to account for drifting conditioning the bias adjustment on the lead time. The bias-adjustment was done for each month with the estimated BA parameters being applied to each one of them.

Figure 8, Figure 9 and Figure 10 present the CRPS score for all months and forecast months (forecast month 1 to forecast month 3 respectively). As for the sub-seasonal precipitation forecasts, the skill varies both geographically and seasonally with acceptable skill in different regions for the 1<sup>st</sup> forecast month. Results show that the skill deteriorates rapidly already from the 2<sup>nd</sup> forecast month, however the Scandinavia and occasionally central Europe seem to retain the moderate skill for high forecast months.

It is interesting to note that although adequate forecasting skill is observed for forecast month 1, the skill value is less than the value observed in sub-seasonal forecasting. This could be due to the lack of skill in forecast week 3 and 4, which over the 1<sup>st</sup> forecast month is “averaged” into a single value. Nevertheless, seasonal precipitation forecasts are capable for downstream applications/services for the first forecast months.

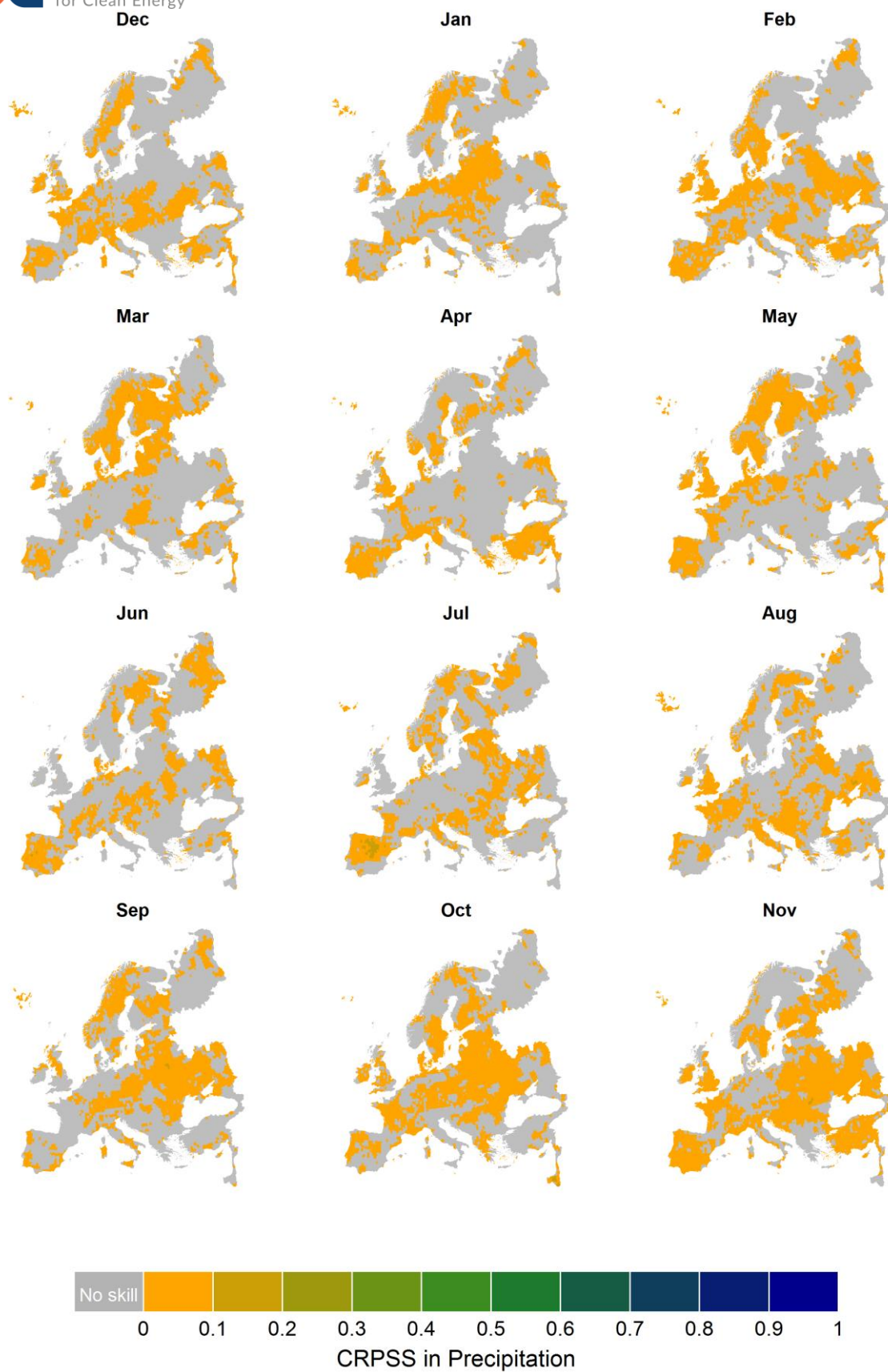


**Figure 8: Forecast month 1 monthly spatial variability of the CRPS score (CRPSS).**



**Figure 9: Forecast month 2 monthly spatial variability of the CRPS score (CRPSS).**





**Figure 10: Forecast month 3 monthly spatial variability of the CRPS score (CRPSS).**



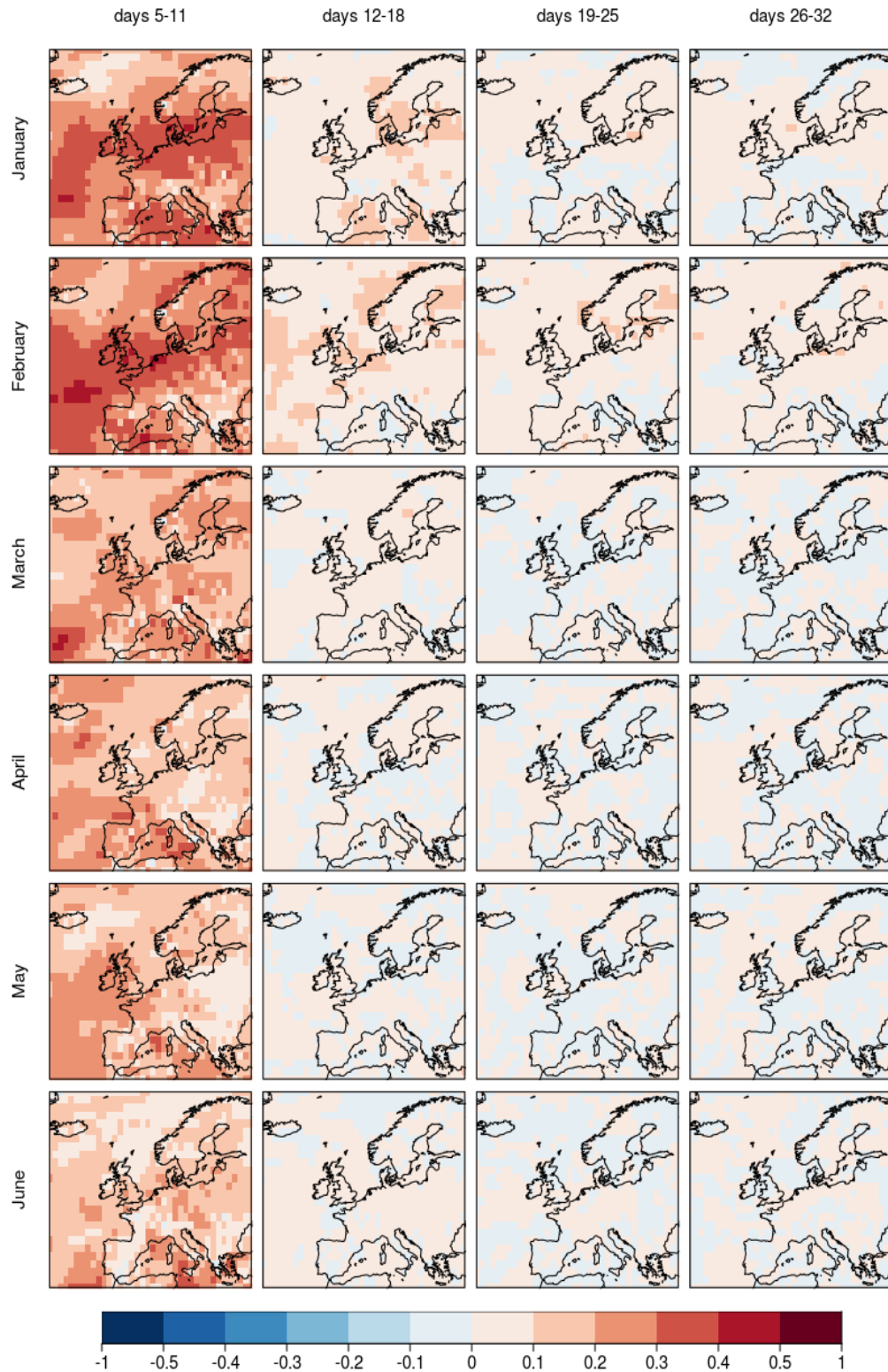
## 5.2. Surface wind speed

### 5.2.1. Sub-seasonal time-scale

The weekly averages were calibrated using the variation inflation method (Doblas-Reyes et al., 2005; Torralba et al., 2017). This method ensures that the predictions will have inter-annual variance that is equivalent to that of a reference dataset.

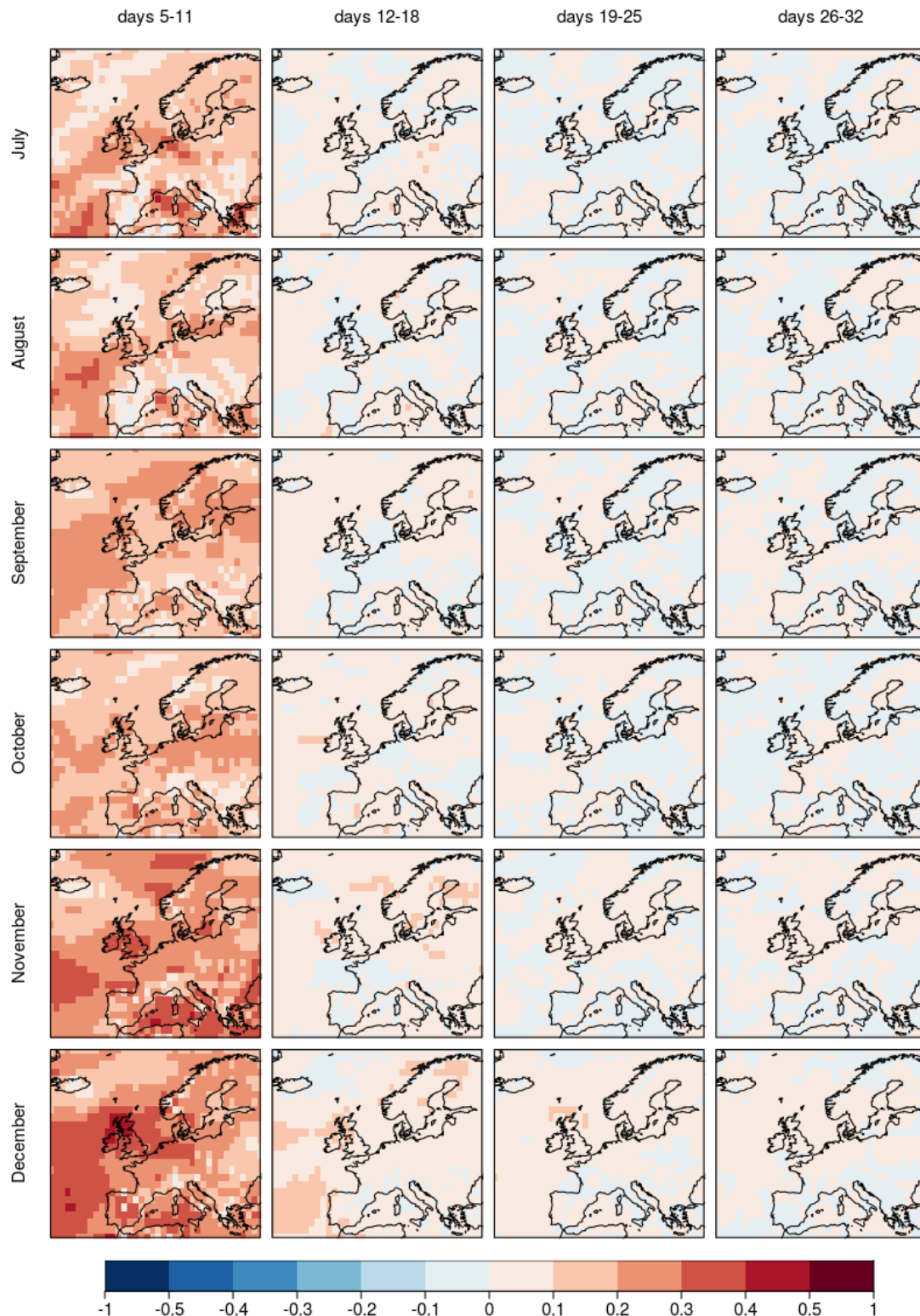
Figure 11 and Figure 12 show the maps of fairCRPSS for weekly 10 m wind speed over the European region computed per month and forecast time. The days are counted since the date the forecast is issued. The first thing to notice is that for week 1 forecasts (forecast time 5-11 days) there is positive fairCRPSS over the whole European domain. For longer lead times (weeks 2, 3 and 4) skill is lower and negative in some areas. During winter months (January, February, December) the fairCRPSS is higher for all lead times. In the Scandinavian region there are positive values up to week 4 in February. The areas that present highest skill are Northern Europe and Scandinavian countries. Over southern countries, beyond week 1 the skill is less consistent.

## sfcWind FairCrpss calibration



**Figure 11: FairCRPSS of 10 m wind speed from ECMWF monthly prediction system for January to June and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts are calibrated with the variance inflation method. (Note that the scale is not linear).**

## sfcWind FairCpss calibration



**Figure 12: FairCRPSS of 10m wind speed from ECMWF monthly prediction system for July to December and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts are calibrated with the variance inflation method. (Note that the scale is not linear).**

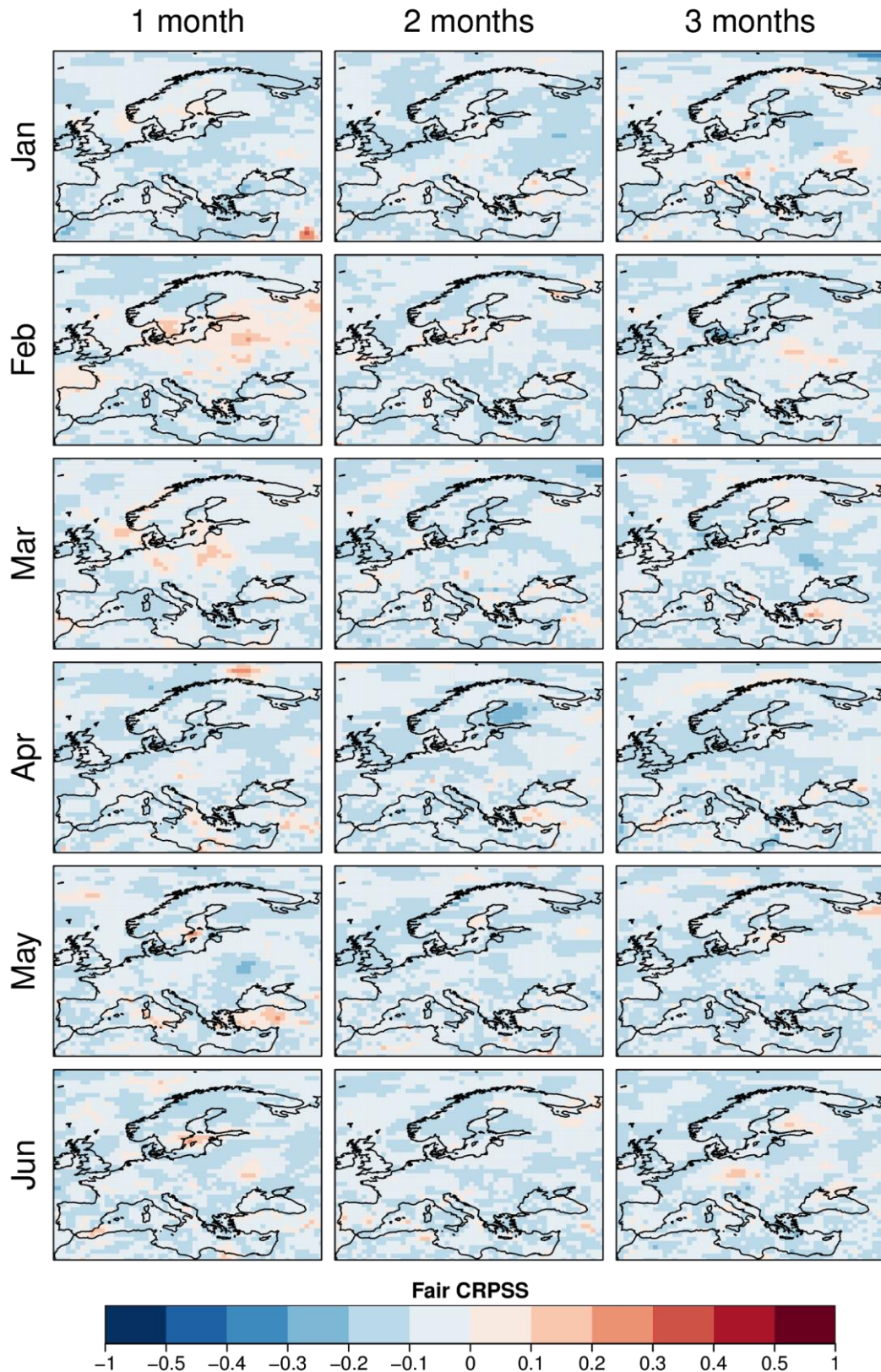
### 5.2.2. Seasonal time-scale

The seasonal predictions are bias-adjusted beforehand using the calibration method described in Torralba et al. 2018. The method adjusts the mean bias and inflates the spread of the distribution to produce unbiased and reliable forecasts. A leave-one-out setting is employed to adjust each year forecast without employing the corresponding observations for that year.

Figure 13 and Figure 14 present the results for the Fair CRPSS. For all start dates and lead times most of the values are close to zero, with negative values dominating the continent. Some spots with modest positive CRPSS values can be seen in some regions especially for the 1-month ahead lead time in February, March, May, August, October or December. However, the process of aggregating the ensemble members into probabilities for the tercile categories can provide some additional skill. The RPSS (Figure 13 and Figure 14), that measures the quality of the tercile probabilities itself, shows higher values than CRPSS. Therefore, there are chances of employing the tercile predictions and perform better than the climatology in many cases. Some windows of opportunity are highlighted below:

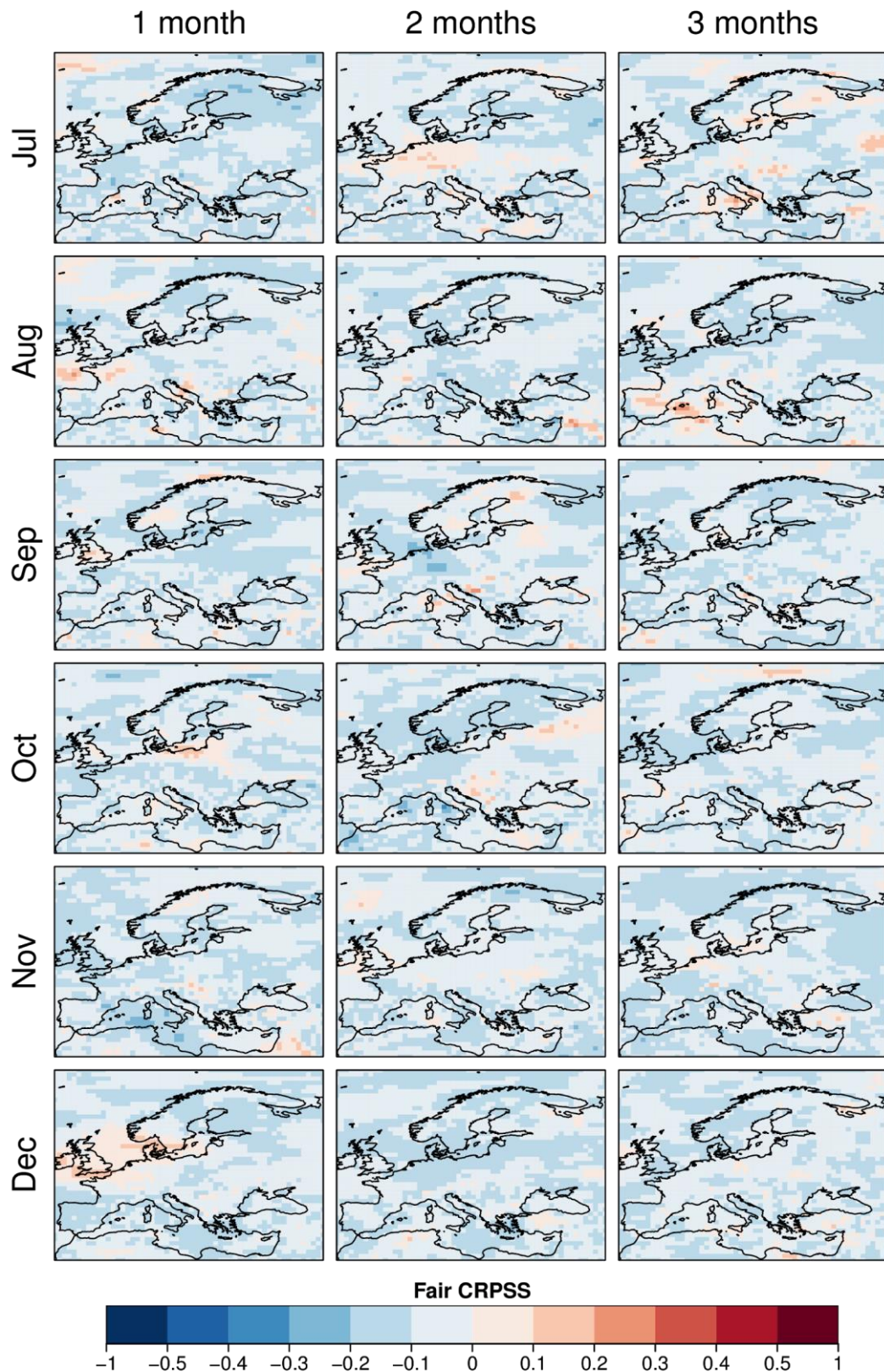
- wind speed forecasts issued in July and valid for September over France and Germany.
- wind speed forecasts issued in March, April and May and valid for June over Turkey.
- wind speed forecasts issued in August and valid for November over Spain.
- wind speed forecasts issued in February and March and valid one month ahead over central and eastern Europe, and around the Baltic sea.
- wind speed forecasts issued in December and valid for January over the British Isles, Denmark and northern France.





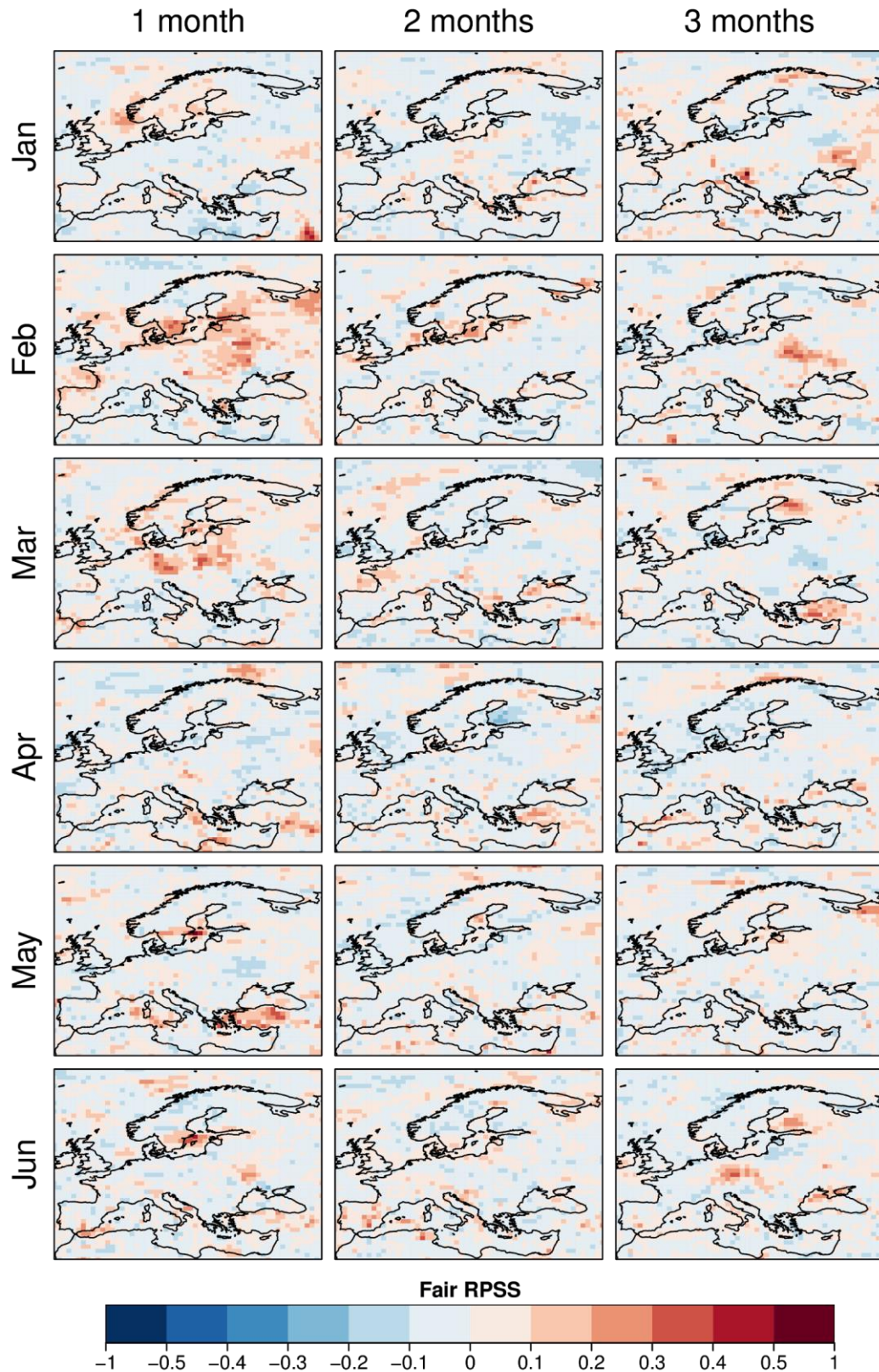
**Figure 13: Fair CRPSS of ECMWF SEAS5 predictions of surface wind. Each row presents the skill for the forecasts started on January to June, with the columns representing one to three months of lead time.**





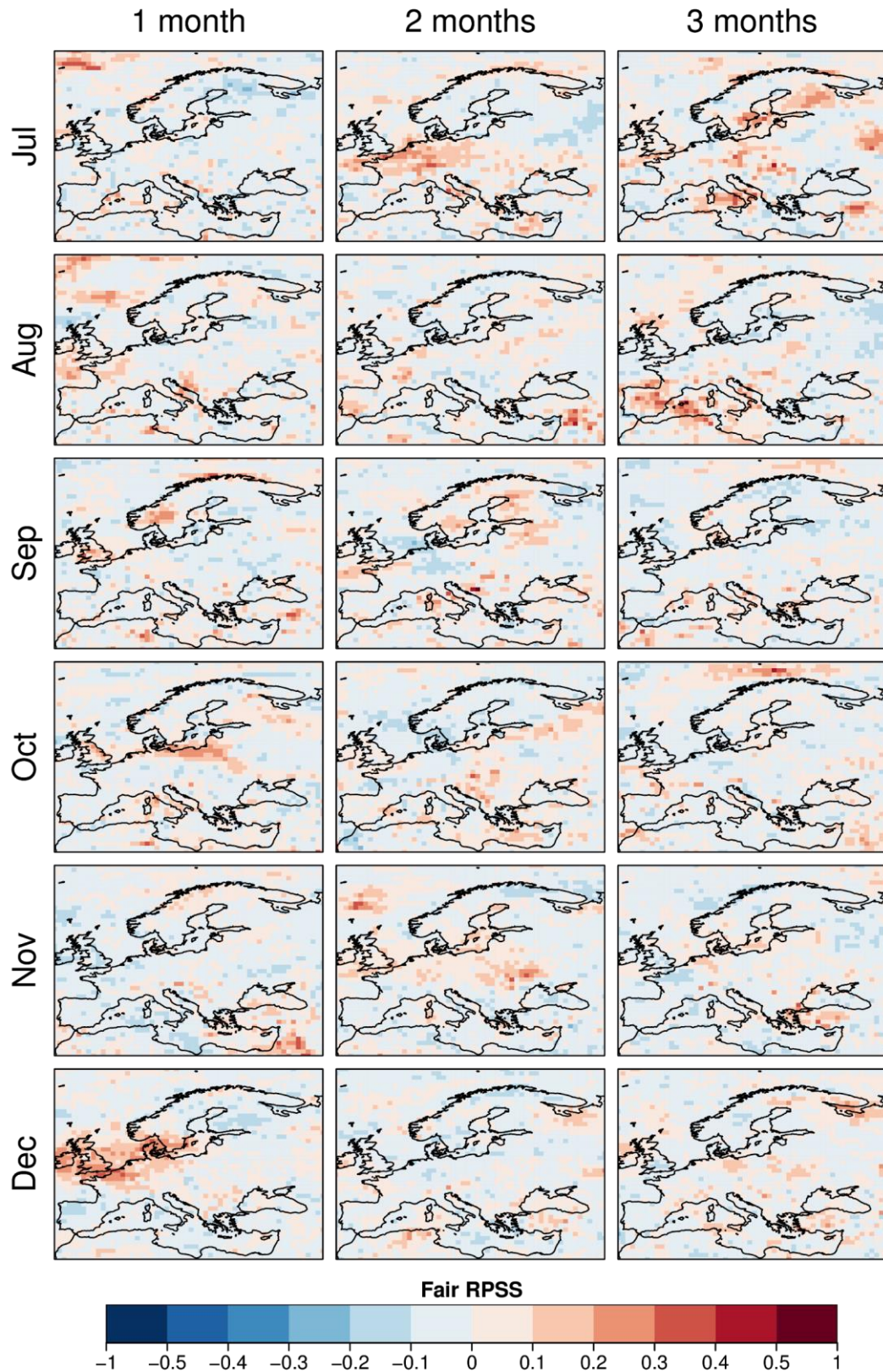
**Figure 14: Fair CRPSS of ECMWF SEAS5 predictions of surface wind. Each row presents the skill for the forecasts started on July to December, with the columns representing one to three months of lead time.**





**Figure 15: Fair RPSS of ECMWF SEAS5 predictions of surface wind. Each row presents the skill for the forecasts started on January to June, with the columns representing one to three months of lead time.**





**Figure 16: Fair RPSS of ECMWF SEAS5 predictions of surface wind. Each row presents the skill for the forecasts started on July to December, with the columns representing one to three months of lead time.**

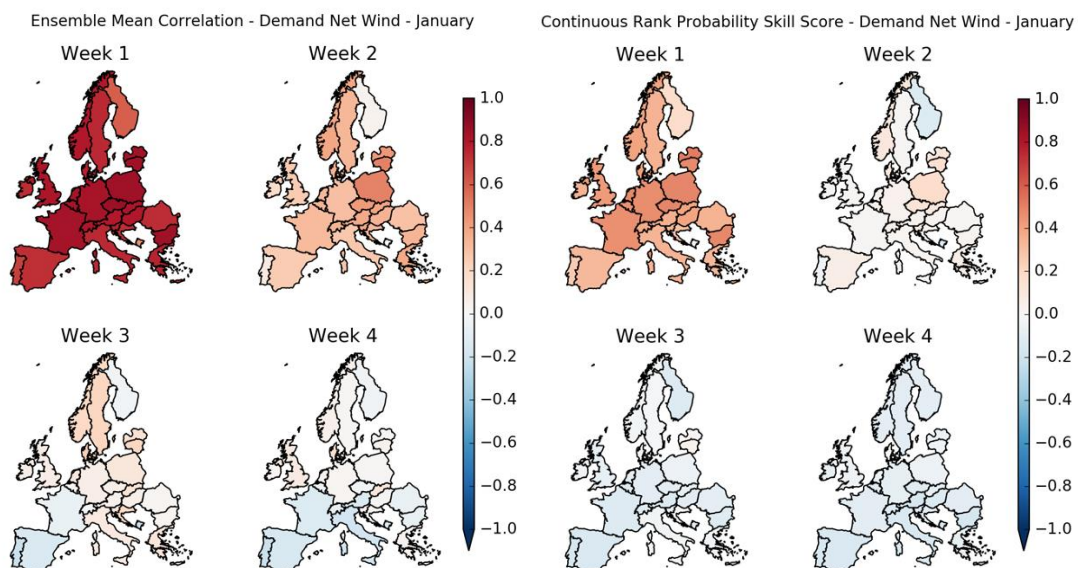


## 5.3. Country-scale supply-demand balance indicators

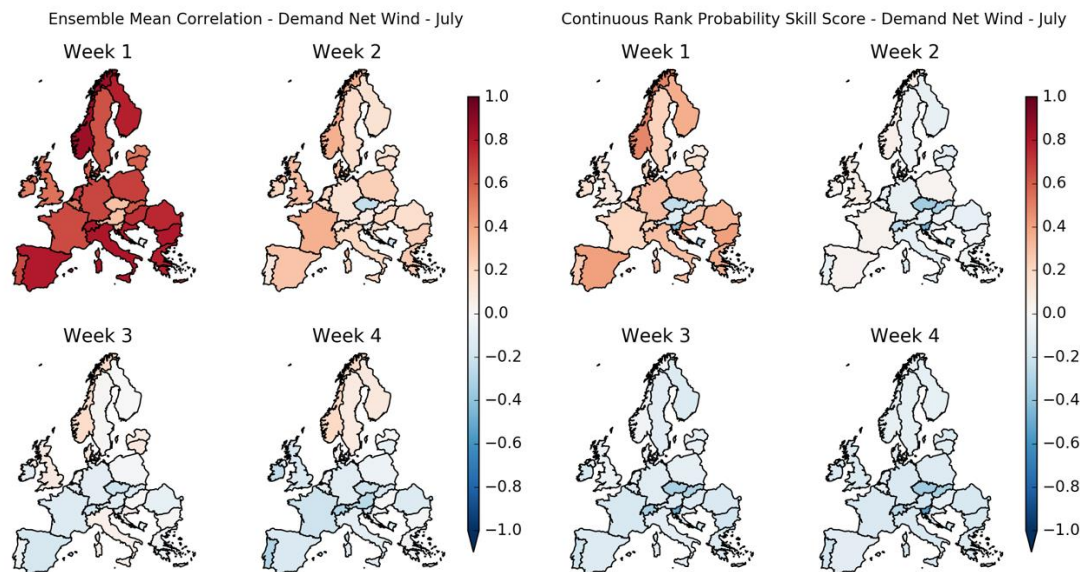
As discussed earlier, country-scale indicators of demand, renewable potential, and residual demand-net-renewables provide vital context to many stakeholders in the energy system. In this section, the skill in forecasting the residual demand-net-wind is presented for sub-seasonal and seasonal forecasts, and discussed in terms of its component parts (i.e., demand and wind-power individually). The discussion focuses on only two representative skill metrics (ensemble mean correlation, CRPSS) for two selected calendar months (January and July). Figures corresponding to other metrics (BSS P10, BSS P90, RPSS), calendar months, and a fuller presentation of skill in the component variables (demand and wind power) are presented separately in Annex 2.

### 5.3.1. Sub-seasonal time-scale

Figure 17 and Figure 18 show the forecast skill in predicting country-aggregate weekly-mean demand-net-wind for validation months January and July respectively.



**Figure 17: ECMWF sub-seasonal forecast skill for weekly-mean, national-aggregate demand-net-wind for validation weeks in January. Left four panels – ensemble mean correlation. Right four panels – CRPSS. In each set of four subfigures, the subtitles indicate lead time (weeks 1-4).**



**Figure 18: As Figure 17 but for ECMWF sub-seasonal forecast skill for weekly-mean, national-aggregate demand-net-wind for validation weeks in July.**

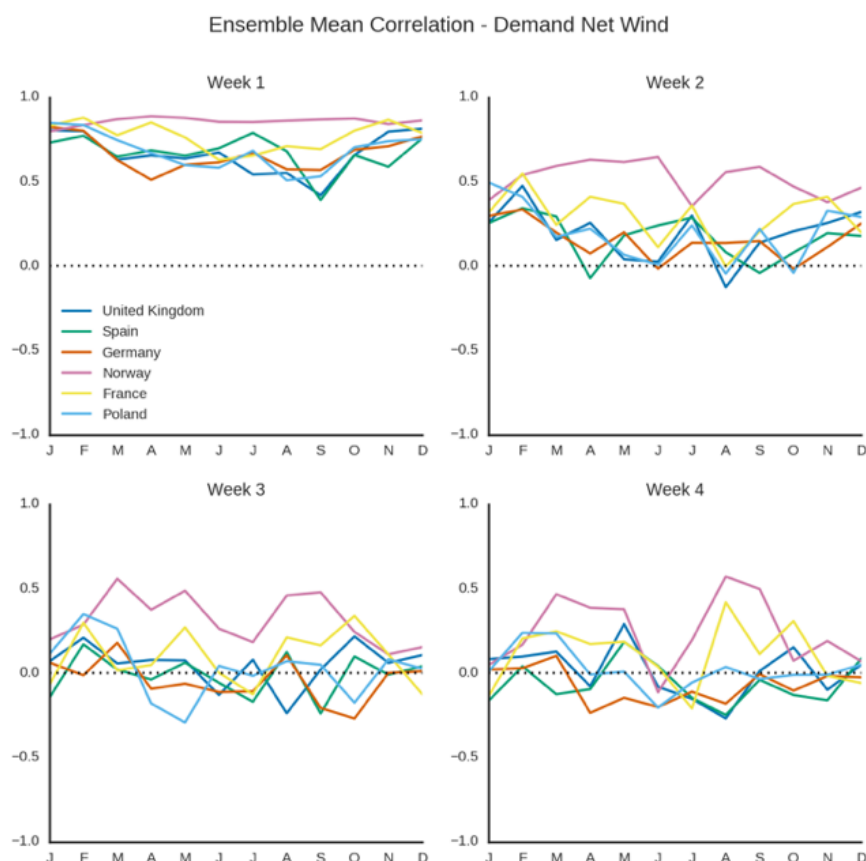
It is immediately clear from these figures that the sub-seasonal forecast generally contains a good level of skill for week 1 (i.e., averaging over a period 5-11 days ahead; top left plots in each 4-figure panel). This is true for both seasons, though somewhat weaker for July than January (compare corresponding subfigures in Figure 17 and Figure 18). It is speculated that this is consistent with the synoptic-scale atmospheric flow (which is generally well-predicted by the forecast system) exerting a stronger influence on surface properties in winter compared to summer. Interestingly, for both seasons, the skill is distributed somewhat uniformly over Europe: i.e., there is good skill in predicting country-aggregate demand-net-wind across most of the land area. There are, however, indications that this skill is not entirely uniform, particularly over Austria and the Czech Republic – suggesting that there are potentially difficulties in converting weather-variables to power-variables in some locations.

The improvement in skill compared to a climatological forecast is clearer for the ensemble mean correlation score compared to the CRPSS (e.g., compare the top-left plot in the left hand set of panels with the equivalent plot in the right hand set of panels in Figure 17). This difference in performance between the two metrics is consistent with the complexity of the forecast information: ensemble mean correlation essentially depends only on the first moment of the forecast ensemble distribution (i.e., its mean value), whereas the CRPSS is sensitive to higher moments of the forecast ensemble distribution (i.e., its shape). It is worth noting, that these differences in skill may impact some users differently to others (e.g., previous research suggests that an “idealised” user seeking profit maximization based on purely speculative trading may achieve it with a skillful forecast of the ensemble mean alone, whereas more advanced risk management strategies - such as hedging to reduce volatility –

may require skill in forecasting higher moments of the forecast distribution's shape (Lynch 2016).

At longer lead times, forecast skill reduces in both January and July. For week 2 (12-18 days ahead), there are indications of forecast skill in the ensemble mean correlation (left hand panels of Figure 17 and Figure 18) but CRPSS is minimal (right hand panels). Estimates of statistical uncertainty are not included here, but it is unlikely that the regions of weak positive CRPSS skill are statistically robust from sampling uncertainty. The skill in the ensemble mean correlation is strongest over eastern Europe in January (Figure 17), contrasting with a more even spread in July (Figure 18). At lead times beyond week 2 (19 days+), there is little evidence of skill for either month.

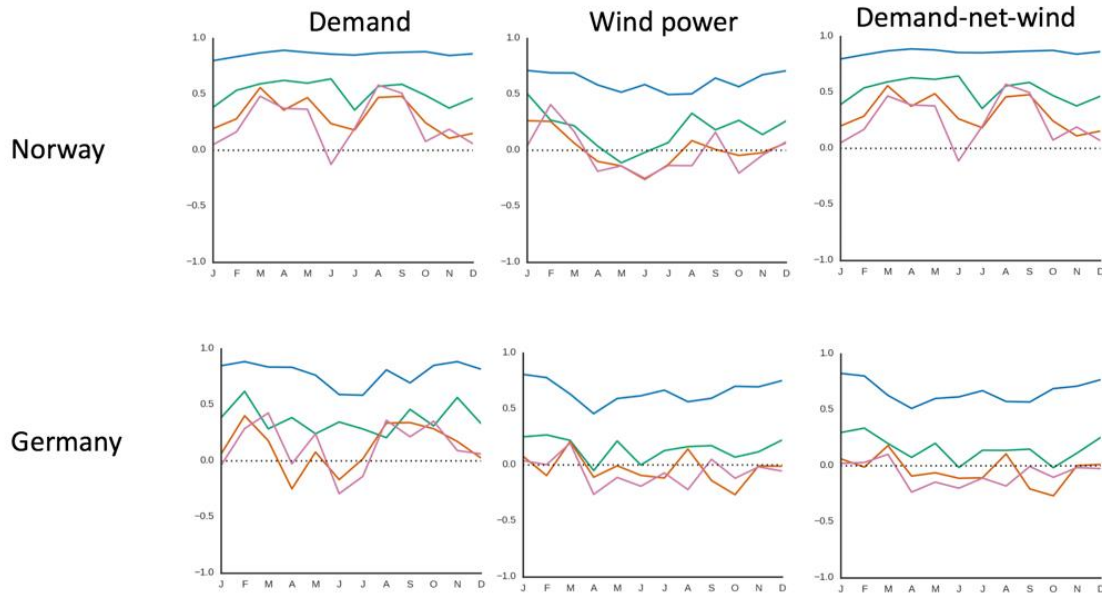
A broader year-round view of skill is provided in Figure 19 for a selection of countries. As discussed previously, skill clearly decays with lead time throughout the year and is lower for CRPSS than ensemble mean correlation (not shown). In general, the skill levels of different countries are consistent insofar as differences in skill between countries are smaller than the difference in skill between successive lead times. Intriguingly, however, there is perhaps some evidence of stronger skill at longer lead times for a few countries, particularly Norway (pink line in in Figure 19), as discussed below.



**Figure 19: ECMWF sub-seasonal forecast skill (ensemble mean correlation score) for weekly-mean, national-aggregate demand-net-wind for a selection of countries as a function of verification month. The subtitles indicate lead time (weeks 1-4).**

An in-depth discussion of the causes of the detailed structure of spatial differences in forecast skill is beyond the scope of this document, but their presence emphasises the need for a greater process-based understanding of both the meteorology and “energy conversion” aspects of the forecasting process. In particular, spatial differences in skill may arise from different sources, with differing implications for how predictive skill can be most readily improved. Differences may arise from a mixture of sampling uncertainty (in effect, random chance), fundamental differences in the predictability of different geographical regions or meteorological properties (temperature vs wind), or as a result of complex sensitivities/insensitivities introduced by the “weather to energy conversion” process (e.g., demand is insensitive to temperature variations between 15.5 and 22 °C, wind power is particularly sensitive to variations between  $\sim 3 \text{ ms}^{-1}$  and  $15 \text{ ms}^{-1}$ , see Deliverable 3.2).

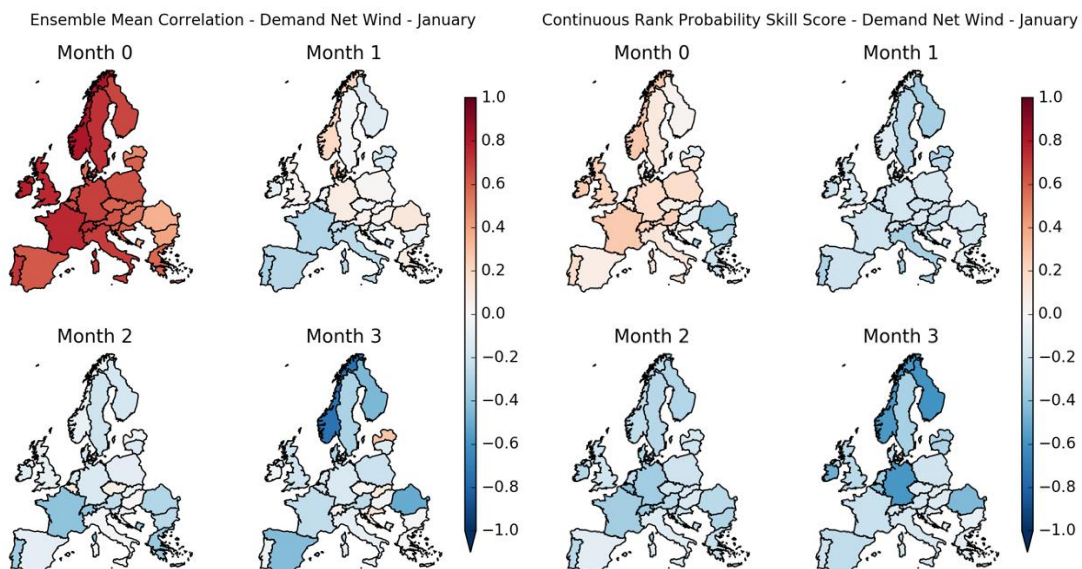
Focusing on the Norway case, however, emphasizes also the role of the underlying make-up of technologies used in the power system itself in determining the level of forecast skill for a variable dependent on a complex composite of meteorological inputs (such as residual demand-net-wind). Figure 20, for example, compares the forecasts skill for country-aggregated demand, wind power and demand-net-wind for Norway and Germany. In this case, the wind-power and demand forecasts exhibit qualitatively similar behaviour in both countries (i.e., some skill in weeks 1 and 2, decaying with lead time), but with the wind-power forecasts typically slightly less skillful than those for demand (this effect is most visible in week 2: comparing the green curves in the left hand column to the centre column suggests that the wind-power forecast has marginal skill whereas the demand forecast skill is more strongly positive). The reason for this difference in skill between demand vs wind-power skill requires further research to understand its origin (whether it is meteorological or connected with the “energy conversion”), but it is nonetheless clear that the forecast week-2 demand-net-wind forecast for Norway retains positive skill (i.e., similar to the demand forecast) while the skill over Germany is marginal (i.e., similar to the wind-power forecast). This emphasizes the role played by the inherent properties of the power system which is being predicted: in Norway, wind power plays a relatively modest role in meeting demand (0.8GW installed against  $\sim 15$ GW mean demand) and is therefore most sensitive to the skill of the demand forecast, whereas in Germany, wind power enjoys rather higher penetration levels (44 GW capacity against  $\sim 55$  GW mean demand) and hence is more sensitive to the skill of the wind-power forecast. This is consistent with changing power-system sensitivities in response to increasing renewables integration, as discussed in Deliverable D3.2.



**Figure 20: ECMWF sub-seasonal forecast skill (ensemble mean correlation score) for weekly-mean, national-aggregate demand, wind-power, and demand-net-wind for Norway and Germany by verification month. Lines: blue – forecast week 1, green – forecast week 2, red – forecast week 3, pink – forecast week 4.**

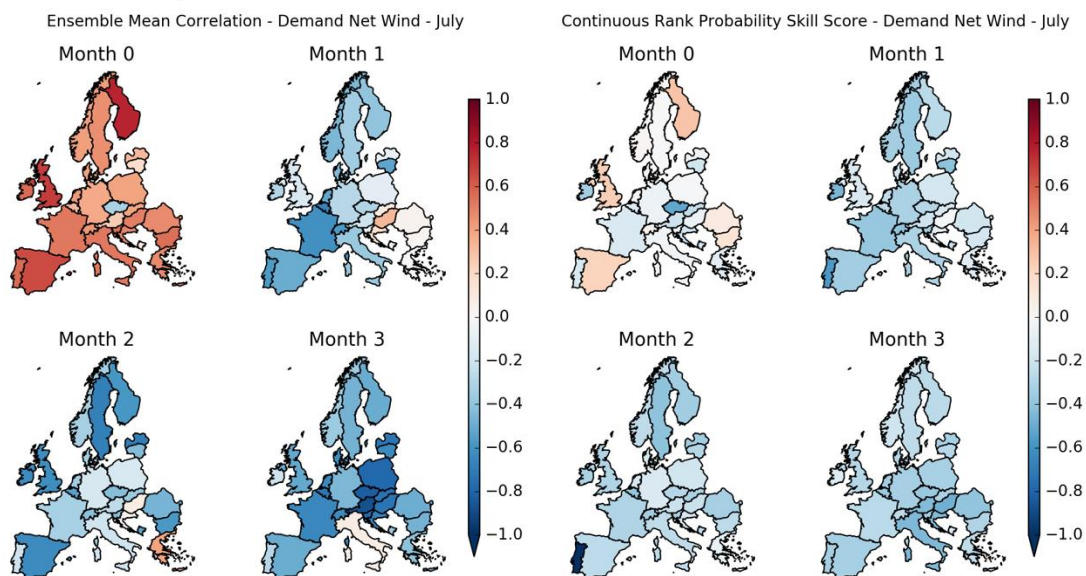
### 5.3.2. Seasonal time-scale

Figure 21 and Figure 22 show the forecast skill in predicting country-aggregate monthly-mean residual demand-net-wind for validation months January and July respectively.



**Figure 21: As Figure 17 but for ECMWF seasonal forecast skill for monthly-mean, national-aggregate demand-net-wind for validation month in January. In each set of four panels, the subtitles indicate lead time (months 0-3).**

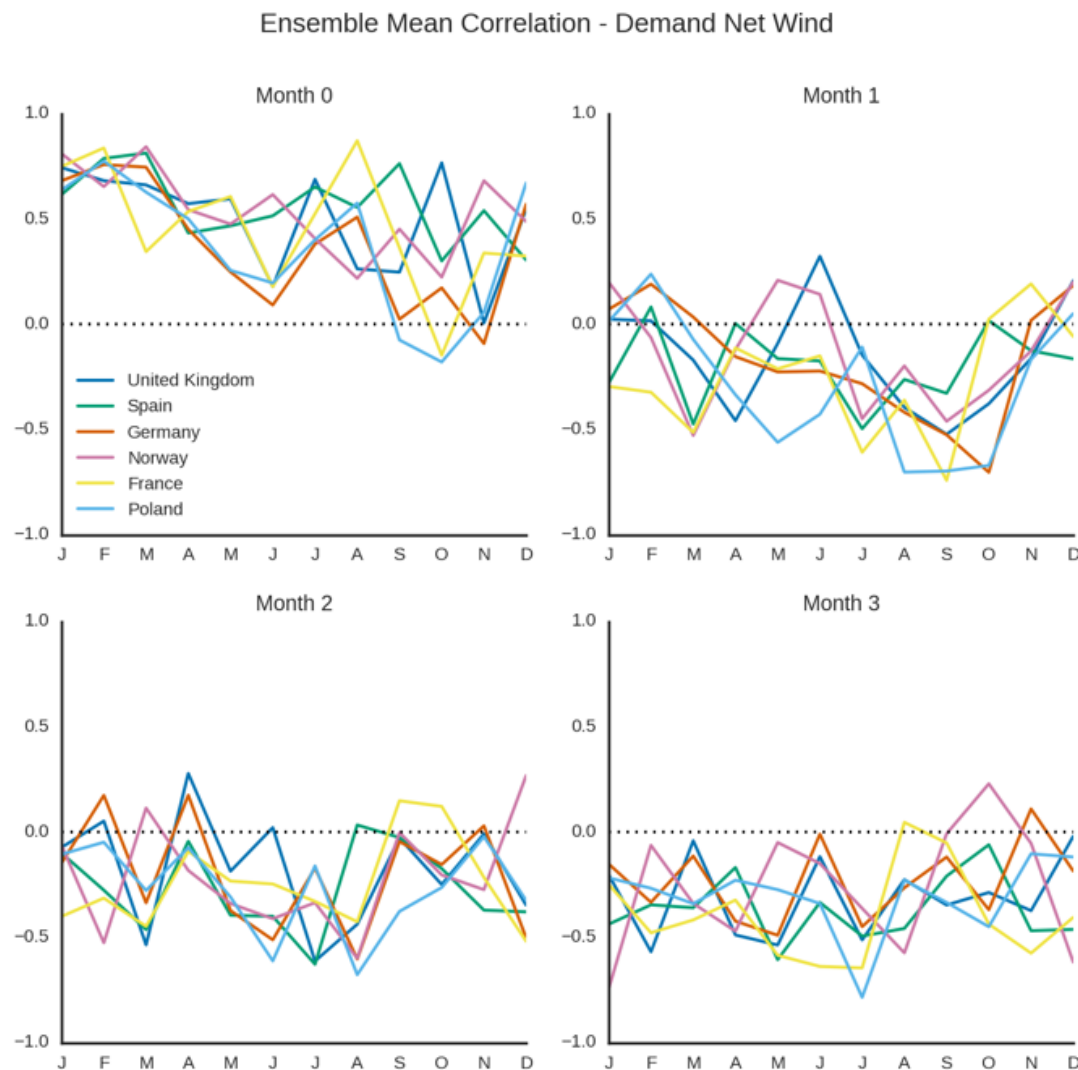




**Figure 22: As Figure 17 but for ECMWF seasonal forecast skill for monthly-mean, national-aggregate demand-net-wind for validation month in July. In each set of four panels, the subtitles indicate lead time (months 0-3).**

As in the sub-seasonal forecasts, there is a greater skill at short lead times (month 0), decaying rapidly with lead time: indeed little skill is found beyond the month in which the forecast is launched. As before, CRPSS skill is usually weaker than the ensemble mean correlation score, with many regions where CRPSS skill is low or negative (i.e., no better or even worse than climatology) even in month 0. Skill is marginally lower in July compared to January (compare corresponding panels in Figure 21 and Figure 22). The Czech Republic performs particularly poorly – especially in July, potentially highlighting issues with the weather-to-energy conversion process in this country.

The year-round and pan-European nature of this generally weak skill is confirmed in Figure 23, with the summer/autumn period typically worse than the winter/spring. It is interesting to note that for months 2-3, the forecast appears to perform systematically worse than climatology (i.e., skill scores less than zero), suggesting that improvements in forecast calibration may potentially improve these negative scores and – by extension – enhance the positive scores at shorter lead times.



**Figure 23: As Figure 19 but for ECMWF seasonal forecast skill for monthly-mean, national-aggregate demand-net-wind for a selection of countries as a function of verification month. The subtitles indicate lead time (months 0-3).**

A clear message from this analysis is that extreme caution must be taken in using forecasts of monthly-mean demand-net-wind. These results do not necessarily imply, however, that it is not possible to derive any skill from such forecasts. In particular, further work is required to understand the nature of the low-skill (e.g., whether it originates from a poor meteorological forecast of temperature or the complexity of the “conversion” to demand). Similarly, improved meteorological calibration (e.g., quantile matching and/or variance inflation), weather-to-energy conversion models, pattern-based techniques, or regime-dependent predictability (“windows of opportunity”) may offer routes to enhanced predictability. Finally, it is noted that the forecasts perform particularly poorly in some countries and time periods (e.g., Czech Republic in July) – suggesting that a more in-depth understanding of the

uncertainties in the weather-to-energy conversion process (and limitations in the training data available) is needed.

## 5.4. Conclusions from the European-wide assessment

### Precipitation

The analysis of precipitation forecasting skill for about 35000 points over Europe allowed understanding of emerging patterns along Europe's climatic gradient:

- Both ENS-ER and SEAS5 precipitation forecasts have biases that need to be adjusted to observations prior to impact assessment. These biases are not similar, and hence a different approach is applied for the two different systems (adjustment over a 3-week moving window and each month for sub-seasonal and seasonal systems respectively).
- Skill in the forecasts varies both geographically and (sub-) seasonally with the forecasts being skillful in most part of Europe up to two weeks/months after their initialization for sub-seasonal/seasonal forecasts respectively.
- Forecasts in the summer months (June-August) have less skill in comparison to the other months.
- Seasonal precipitation forecasts in the Scandinavia and occasionally central Europe seem to retain their moderate skill for higher than two forecast months ahead.
- Although seasonal forecasts in forecast month one are skillful, the skill value is less than the value observed in sub-seasonal forecasting due to the lack of skill in forecast week 3 and 4, which over the 1<sup>st</sup> forecast month is "averaged" into a single value. This could also be an artefact of the hindcast period used in the analysis; 11-12 years for ENS-ER and 23 years for SEAS5.

### Wind speed

The analysis of wind speed at 10 m for ENS-ER and SEAS5 systems at the European domain is done at both sub-seasonal and seasonal time-scales employing several skill scores. It is seen that there is higher skill in aggregating in terciles (fairRPSS) than in predicting the full distribution (fairCRPSS), this is particularly important in seasonal predictions. Nevertheless, the dependence of skill with lead time and region has revealed some windows of opportunity.

- The sub-seasonal forecast show good skill at week one (days 5-11). After week two skill is reduced and there are areas of positive and negative skill.
- During the winter months (January, February, March) the level of skill of sub-seasonal predictions is the highest.
- The areas with highest sub-seasonal skill are Scandinavia and Northern Europe.

For seasonal predictions, the windows of opportunity are found in:



- Forecasts issued in February and March is valid one month ahead over central and eastern Europe, and around the Baltic Sea. Forecasts issued in July is valid for September over France and Germany.
- Forecasts issued in March, April and May is valid for June over Turkey.
- Forecasts issued in August is valid for November over Spain.
- Forecasts issued in February and March is valid one month ahead over central and eastern Europe, and around the Baltic Sea.
- Forecasts issued in December is valid for January over the British Isles, Denmark and northern France.

### Country-scale supply-demand balance

The analysis of forecast skill for the country-scale balance indicators provided above allows several qualitative insights to be drawn relating to the ECMWF sub-seasonal and seasonal forecast systems:

- The sub-seasonal forecast can provide skill in many regions in forecast weeks 1 and 2 (when forecasting weekly-average blocks for days 5-11 and 12-18). There is occasional skill for some countries and variables at week 3.
- Skill in the seasonal forecast is only strongly evident in the launch month (when forecasting a monthly-average block).
- There is generally less skill in predicting higher order moments of the forecast distribution compared to lower order moments (e.g., its mean), though this may be improved with higher-order bias correction methods.
- In general, forecasts show more skill in winter than summer.
- In general, forecast skill is reasonably consistent across a wide geographical domain. However, it is possible for the skill in individual countries to be strongly modified, potentially consistent with sensitivities, errors and/or uncertainties in the weather-to-energy conversion process.
- Different “energy quantities” offer different levels of forecast skill. In general, wind power forecasts are slightly less skillful than demand (for the corresponding lead time). Depending on the make-up of a given county’s power system, it appears that the forecast skill for the residual demand-net-wind may respond more strongly to one or other of its constituent ingredients (i.e., wind-power or demand). Care must be taken, however, to ensure that the underlying reanalysis-derived power system indicators are a reliable representation of the underlying “real” power system.

## 5.5. Limitations from the assessment

While the qualitative findings discussed above are believed to be robust, caution must be taken in interpreting the details of the quantitative results. In particular, it is recognised that the sub-seasonal analysis depends over a relatively limited period of hindcast data (~12-17 years), based on a forecast systems current in ~2016/2017. This has two implications:

- **Sampling uncertainty.** Given the limited period of the hindcast-set, there is likely to be considerable sampling uncertainty. A deliberate decision was, however, taken not to include uncertainty estimates on the material provided in this section, primarily to *avoid* promoting overconfidence in the quantitative results. A limited test of sampling uncertainty (reducing the hindcast period to ~11 years rather than 17) confirms, however, the qualitative robustness of the patterns of skill discussed above, though qualitative changes occur. Caution is therefore recommended in interpreting “small” differences in skill – for example, values that are small but positive should not necessarily be taken as reliable indicators of forecast skill.
- **Current operational forecasts vs. archived hindcasts.** Operational forecast systems are continually updated and it is therefore reasonable to expect quantitative improvements in currently operational forecast systems (with updated physics and greater ensemble size) compared to the hindcast sets analysed here. It is therefore reasonable to view the results presented here as a potential lower bound for the skill of current operational forecasts.

## 6. Case study assessment and evaluation

To support the proof of concept phase, eight historical case studies pointed as the most relevant by industrial partners, i.e. periods with an unusual climate behaviour affecting the energy market, were defined in the participatory activities performed in WP2 (for further information see Deliverable 2.1). Those case studies will serve to:

- 1) Understand how potential users would have benefited from S2S forecasts in those contexts. It might be understood as an evaluation from the users' point of view that will be included in Deliverable 2.2.
- 2) Help defining the Decision Support Tool (WP5).

The complete list of case studies is reported in Table 5. Below is a detailed description of each of the eight case studies. This description includes a short explanation of the episode from the user's perspective, a description of the event by means of analysing reanalysis datasets and the forecast available for the most relevant variables of the case study at that time to illustrate the potential application of S2S forecast. In the discussion, relevant skill scores are also presented to better understand the potential application of S2S forecast. A complete skill assessment for the case studies is included in Annex 3.

## 6.1. Case study 1 – France, Germany 2017

With the increase of renewable energy sources in the electricity generation mix and the rapid reduction of generation capacity from traditional generation sources, the European energy system has become highly sensitive towards extreme weather and climate events. Cold events in winter have a strong impact on the power system. They are often due to blocking situations that combine cold temperatures, mince precipitation and low wind speeds. These conditions translate into high electricity demand and lower than usual hydro and wind power generation. Electricity demand and power generation forecasts are therefore potentially valuable for assessing the risks on the French and the European systems. Several mechanisms can be activated to face such situations, but they all need as accurate as possible forecasts to optimize decision-making.

Cold wave over Europe created a combination of large increase in electricity demand and lower than normal wind power generation.			
Region:	France, Germany	Period:	17-23 Jan 2017
Forecast type:	Sub-seasonal	Main interest:	Demand and wind
Forecast variables:	Wind speed, temperature and demand		

**Table 6: Region, period, forecast type and main interest for case study 1.**

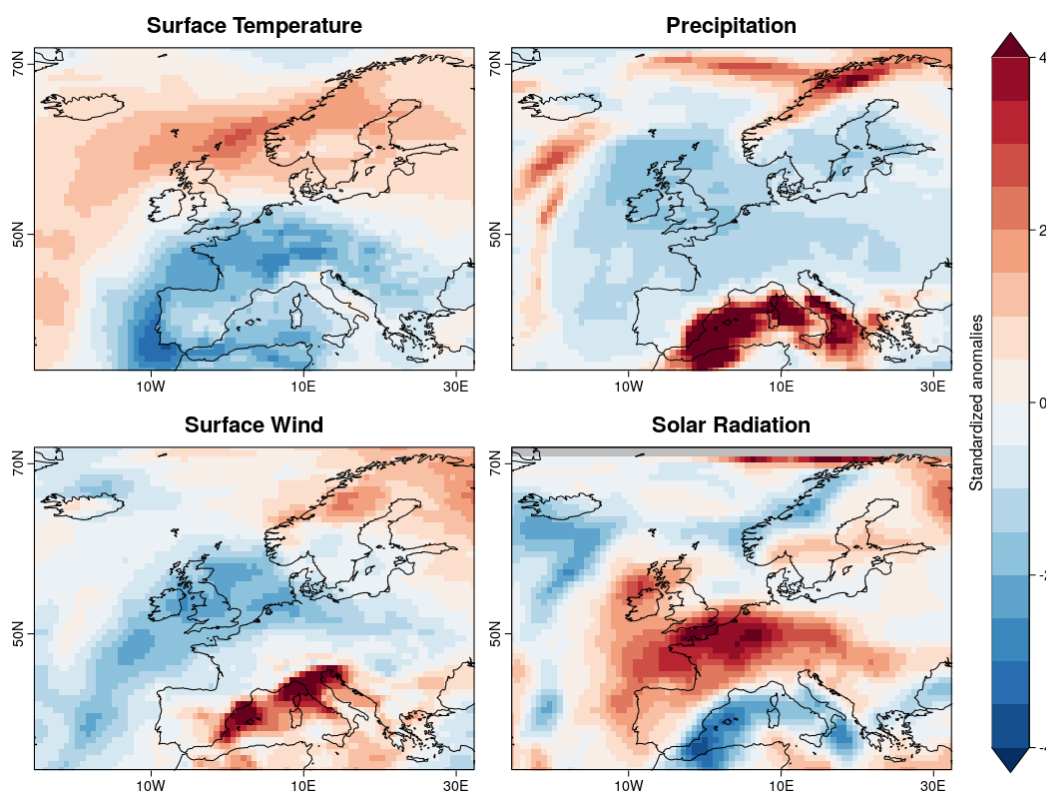
### Event description

In January 2017 a cold wave over Europe created a large increase in electricity demand, especially in France where most of the house heating is powered by electricity. During this period, there were also low wind speeds over Europe and therefore lower than normal renewable energy supply. In France, under the national regulation authority request, utilities had to stop several nuclear reactors in order to carefully check some components. This created a high-risk situation that could have been better managed with accurate forecasts.

The period 17 January to 23 January 2017 was characterized by a tripole of surface pressure, with a central high-pressure anticyclone stretching eastward from the UK towards the Black Sea, and low pressures to the south (in the western Mediterranean Sea associated with a Mediterranean cyclone), and north (north of Scandinavia). Overall, this is consistent with still, cold and dry conditions over most of central Europe (a band across southern UK, northern France, Germany, Austria and Czech Republic; see Figure 24). To the north, the strengthened westerly flow was associated with warmer and wetter conditions, as the westerly flow advects relatively mild and moist air from the Atlantic. Conversely, to the south the easterly flow

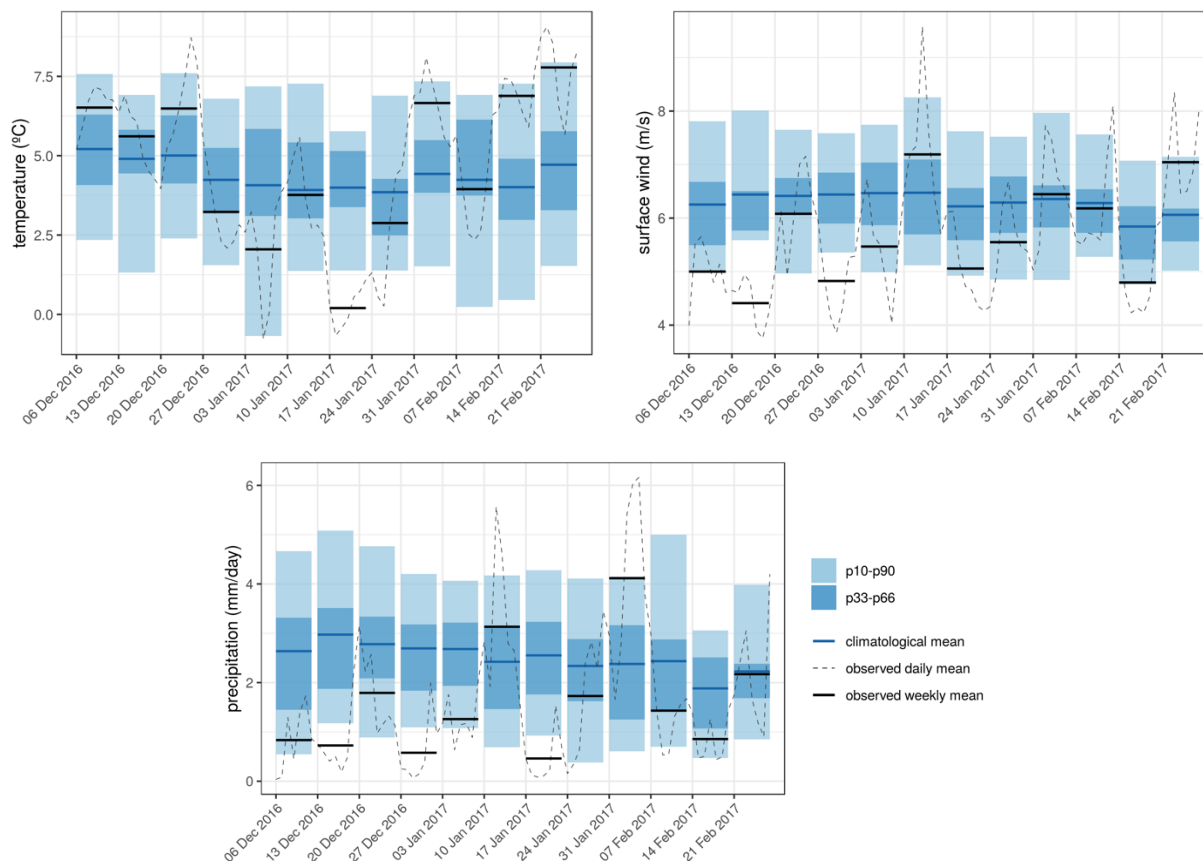
typically brings cold, dry continental air and is associated also with the strong precipitation over the Mediterranean (when the cold air meets the relatively hot sea surface produces strong precipitation episodes).

Temperatures over much of the central region (under the anticyclone) reach approximately 2 standard deviations below the climatological norm for this period, and similar deviations are found for wind speed (though the wind speed anomalies are strongest over the western/ocean part of the domain).



**Figure 24: Observed and surface meteorological conditions compared to ERA-Interim climatology during 17<sup>th</sup> Jan – 23<sup>rd</sup> Jan 2017.**

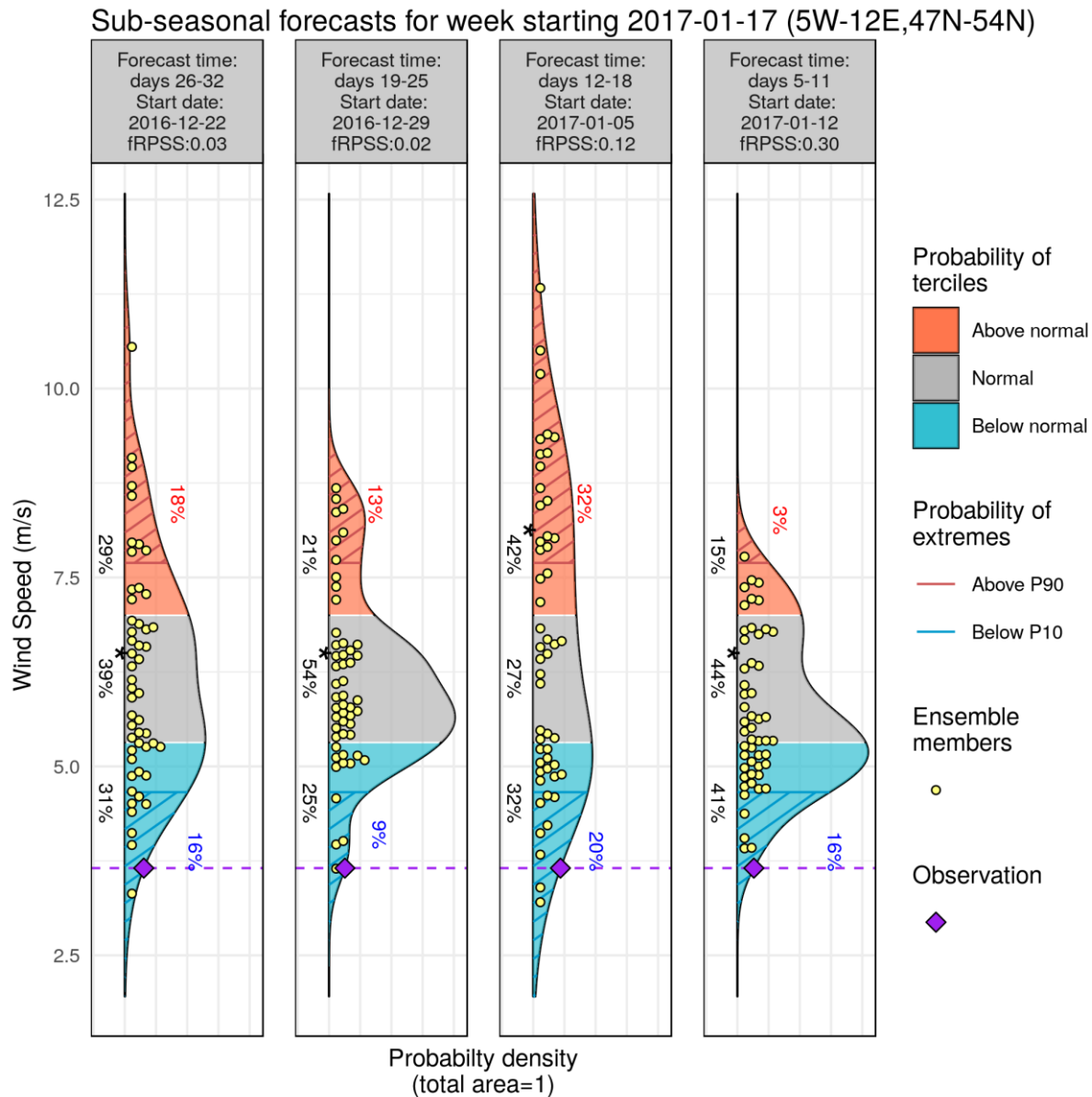
A wider view of the season (Figure 25) suggests that the entire period from late December to the end of January was generally cooler than usual (weekly values are typically in the lower tercile of climatology). 17<sup>th</sup>-23<sup>rd</sup> January represents the coldest week, with temperatures well below the 10<sup>th</sup> percentile of climatology, and low wind speeds (~10<sup>th</sup> percentile). More generally, it is noted that wind speeds and precipitation were both generally low throughout much of the December – suggesting an absence of cyclonic activity – though temperatures in much of early to mid December were generally mild or near-normal.



**Figure 25: Observed and climatological surface air temperature (top left), 10m wind speed (top right) and precipitation (bottom) averaged over 5° W–12° E, 47–54° N during December 2016 to February 2017.**

### Available forecasts

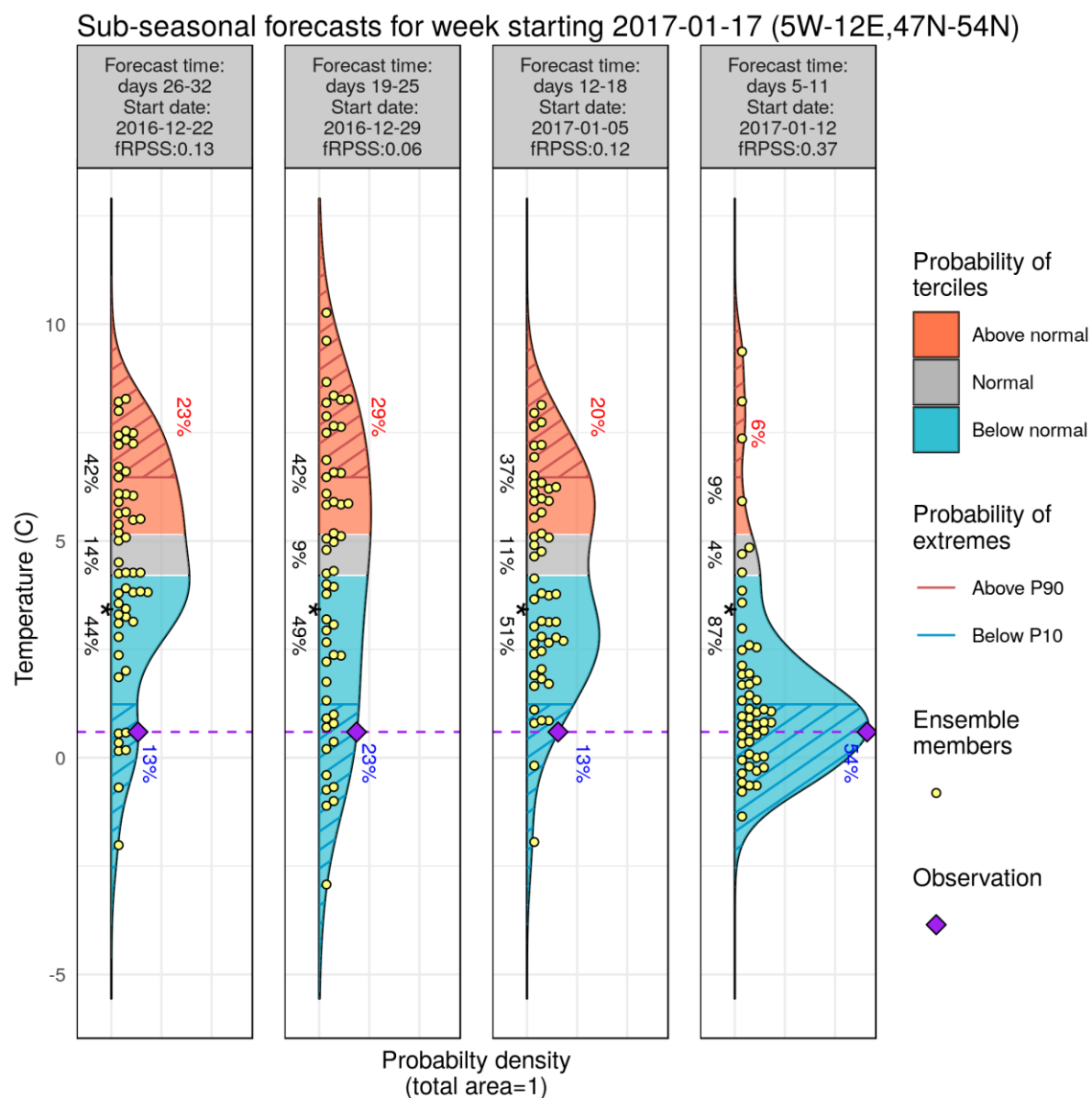
Figure 26, Figure 27 and Figure 28 show a sample of sub-seasonal forecasts that were available during this event. For wind-speed, Figure 26 demonstrates the behaviour of a forecast with limited skill. At lead times in excess of one week, the most dominant tercile is above- or near- normal – i.e., the weaker than normal winds are not well forecasted (consistent with the low fRPSS skill scores shown in the figure). At shorter lead times (1 week ahead) the forecast does, however, tend to suggest likely lower-than-normal wind speeds, though there all members underestimate the severity of the event and the near-normal tercile is estimated to be more probable than below-normal. Interestingly, the forecast at lead week 2 (start date 5<sup>th</sup> Jan) is much spread – more so than weeks 3 and 4 – suggestive of a high uncertainty in the circulation conditions.



**Figure 26: Sub-seasonal forecast for 10m wind speed averaged over 5°W–12°E, 47–54°N for the week 17<sup>th</sup>–23<sup>rd</sup> January 2017. From right to left corresponds to forecasts launched from lead times week 1 to 4 (forecasts times 5-11 days, 12-18 days, 19-25 days, 26-32 days). Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).**

The temperature during the week 17/01/2017-23/01/2017 in the area of study was close to 0 °C. The forecasts available 25 days in advance showed the lower tercile as the most likely (44%). The prediction on 29/12/2016 showed more spread in the members and less skill but was consistent in indicating the lower tercile for temperature as the most likely. The forecast on the 05/01/2017 gave a 51% probability to the lower tercile with greater skill (fairRPSS=0.12). Finally the forecast issued on 12/01/2017 with lead time of 4 days, shows

better agreement amongst members indicating temperatures close to 0 °C. The likelihood of the lower tercile was 87% and of the chances of the temperature falling below the 10<sup>th</sup> percentile were 54%.



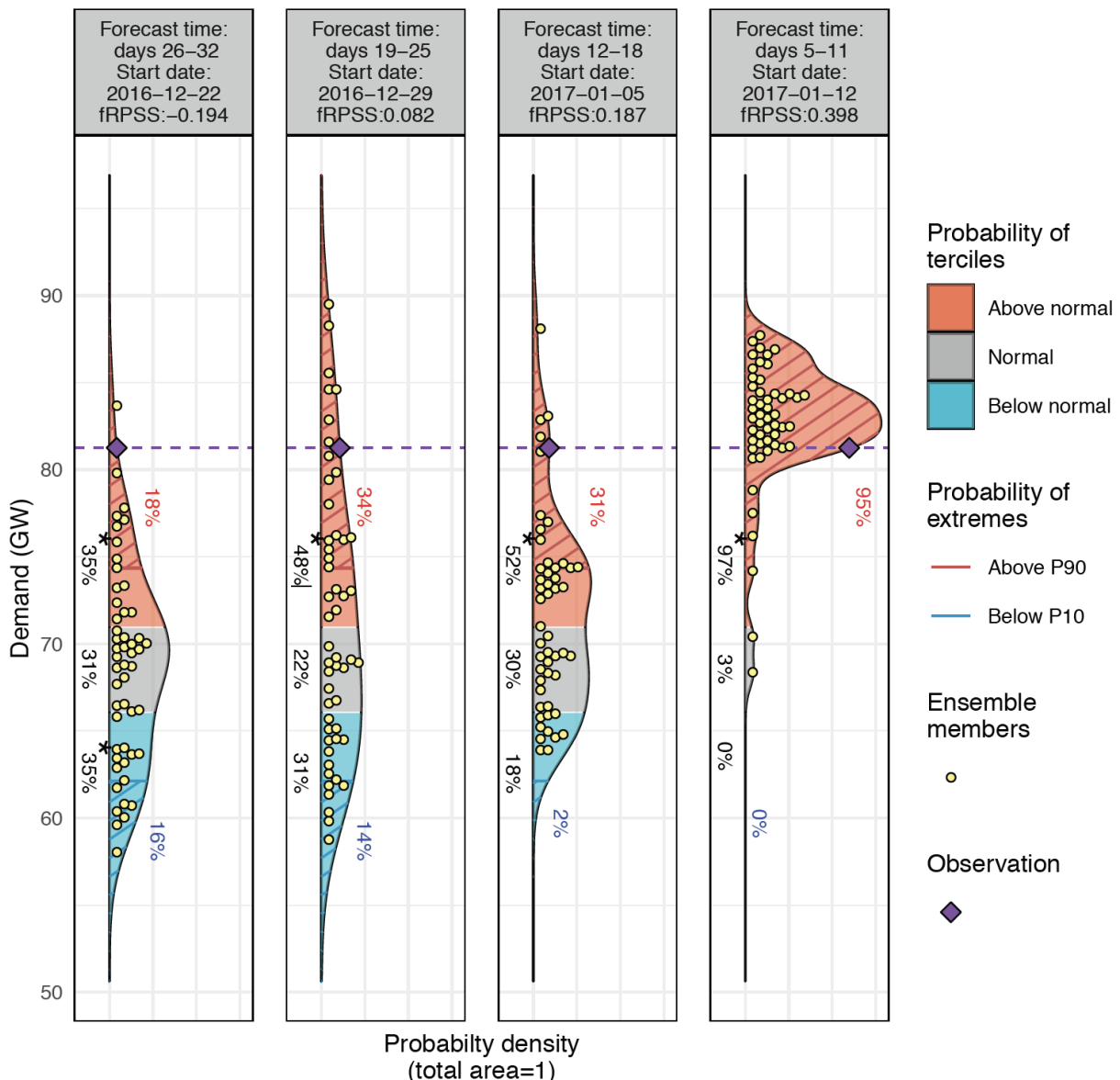
**Figure 27: Sub-seasonal forecast for temperature averaged over 5°W–12°E, 47–54°N for the week 17<sup>th</sup>–23<sup>rd</sup> January 2017. From left to right corresponds to forecasts launched from lead times week 4 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).**

For demand (which is strongly connected to cold temperatures in winter), Figure 28 shows that the forecasts indicate a preference to above normal demand conditions beginning in week 3. This preference increases slightly in week 2, but becomes particularly pronounced in



week 1. The generally positive fRPSS scores (indicated at the top of each forecast PDF) indicate that the forecasts have predictive skill at these time ranges.

### Sub-seasonal forecasts for week starting 2017-01-17 (France)



**Figure 28: Sub-seasonal forecast for demand averaged over 5°W–12°E, 47–54°N for the week 17<sup>th</sup>–23<sup>rd</sup> January 2017. From left to right corresponds to forecasts launched from lead times week 4 to 1. Methodology: lead time dependent mean bias correction, applied to both temperature and demand (once converted), calibrated to ERA5, based on a 17-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).**

In conclusion, this case study shows that S2S predictions of temperature and wind speed would have been relevant and beneficial, and therefore offer an added value over the current practice of using climatological forecasts. The skill increases considerably with shorter lead times and is higher for temperature than for wind speed. The forecasts of temperature and

electricity demand would have helped the energy companies to be prepared for the cold event, and the critical situation in France could have been avoided. There is therefore confidence in anticipating episodes of high electricity demand a few weeks in advance (heat waves or cold spells), although there is less confidence in ensuring the energy supply from wind energy can meet the demand.

## 6.2. Case study 2 – Germany 2013

With nearly 39 GW of installed photovoltaic capacity in Germany, periods of high solar radiation during summer may shift the energy mix considerably in this country. During these periods of elevated solar generation, expensive and polluting conventional power plants may be shut down, with important effects in the energy trading market. In this context, coal power plants are typically used as a backup to ensure security of supply. In Germany, coal supply is largely based on river transportation, which is critically affected by river water levels and navigability. This in turn is associated with precipitation amounts in the previous weeks. In this specific case, very low precipitation levels, may, if prolonged, restrict transportation capacity on major waterways like the Rhine and Neckar rivers. This is of particular relevance for energy companies in the region since they transport coal to several power plants along the aforementioned rivers for conventional power supply.

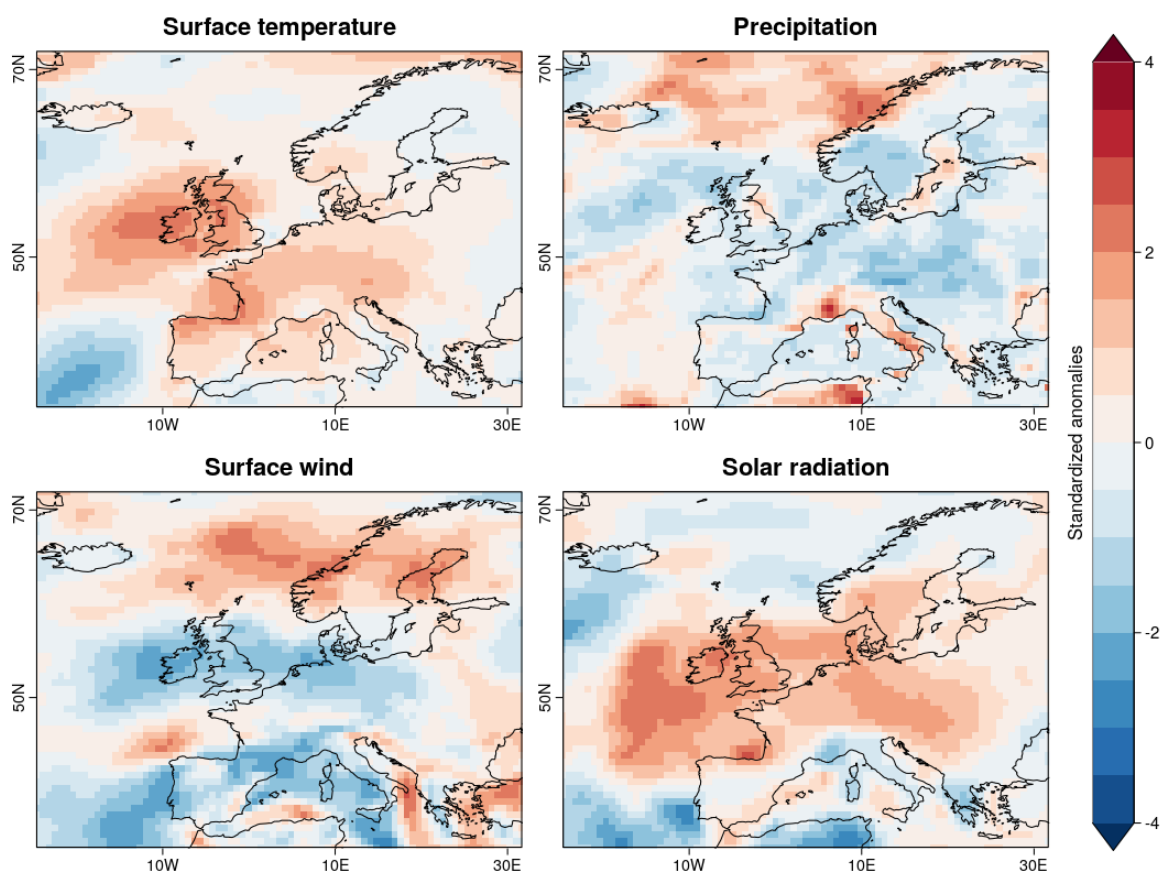
A high-pressure system over central Europe produced large electricity demand, higher than normal solar generation and low precipitation rates.			
Region:	Germany	Period:	July 2013
Forecast type:	Seasonal	Main interest:	Demand, solar, wind and hydro
Forecast variables:	Precipitation, inflows, solar radiation, solar capacity factor, wind speed, temperature and demand		

**Table 7: Region, period, forecast type and main interest for case study 2.**

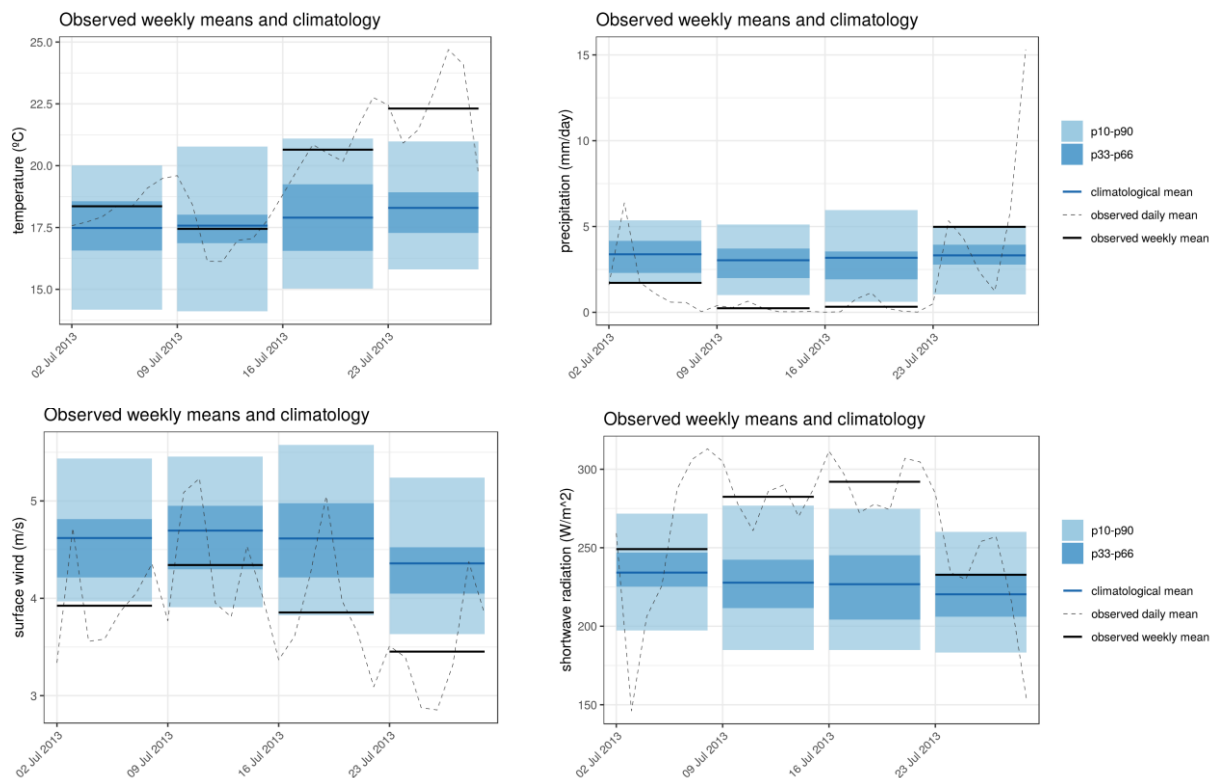
### Event description

Anomaly maps (Figure 29) for July 2013 show that solar radiation was up to two standard deviations higher than the climatological average (bottom right) and precipitation up to two standard deviations below the climatological average (top right). Temperature anomalies were more than one standard deviation higher than climatological average (top left) and wind anomalies (bottom left) were generally below the climatological average. The region that corresponds to Germany was the most affected by the combined effect of lack of precipitation and increase of solar radiation.

The black solid line in Figure 30 represents the weekly evolution of average solar radiation. The blue line and the blue shadows represent the climatological mean and distribution. During the second and third week of July 2013, observed solar radiation over Germany was higher than P90, especially during the third week. Since the beginning of the month, the observed precipitation was lower than normal. The observed value was around the P10 during the first week, but below P10 during the second and third week of July 2013. During the last week of the month, solar radiation value was within the climatological distribution of the solar radiation for the month of July, while the wind speed and consequently the wind farm production decreased to values below P10. The observed temperature was higher than P90, with an average weekly value greater than 22 °C and a potential impact on the energy demand due to electrical cooling needs.



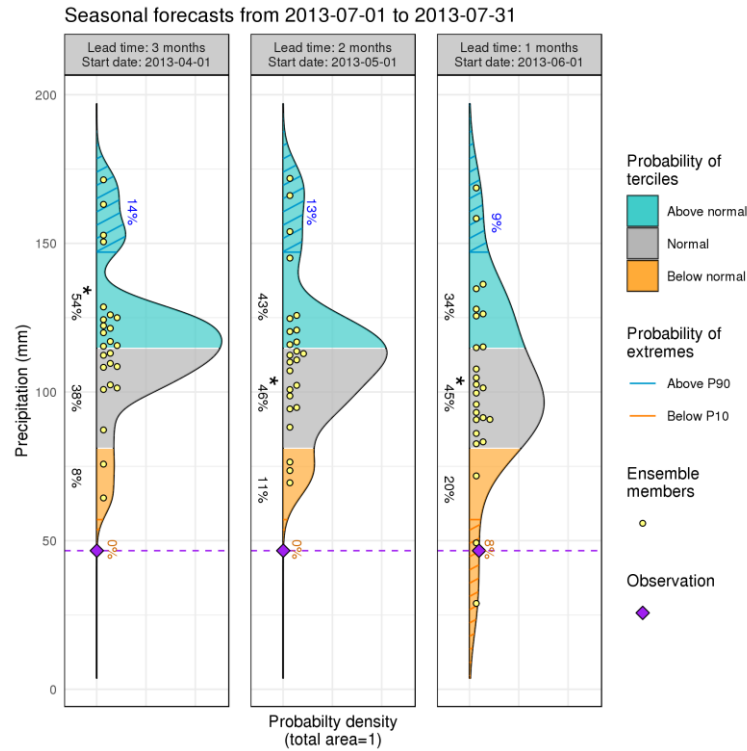
**Figure 29: Standardized anomalies of temperature, precipitation, wind and radiation for July 2013, obtained from ERA-Interim.**



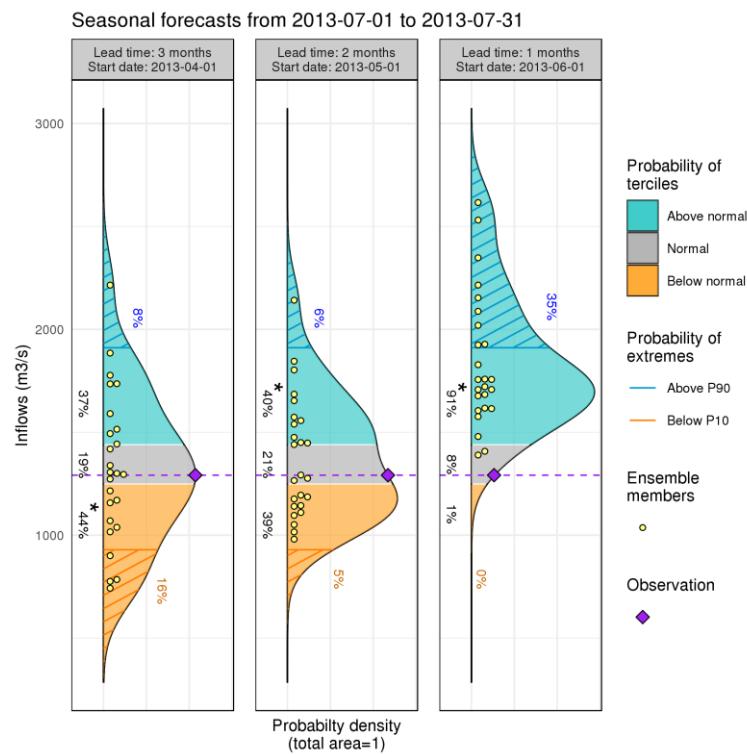
**Figure 30: Weekly evolution of the observed ECVs in the region 6–15°E and 47–54.5°N during July 2013 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis.**

### Available forecasts

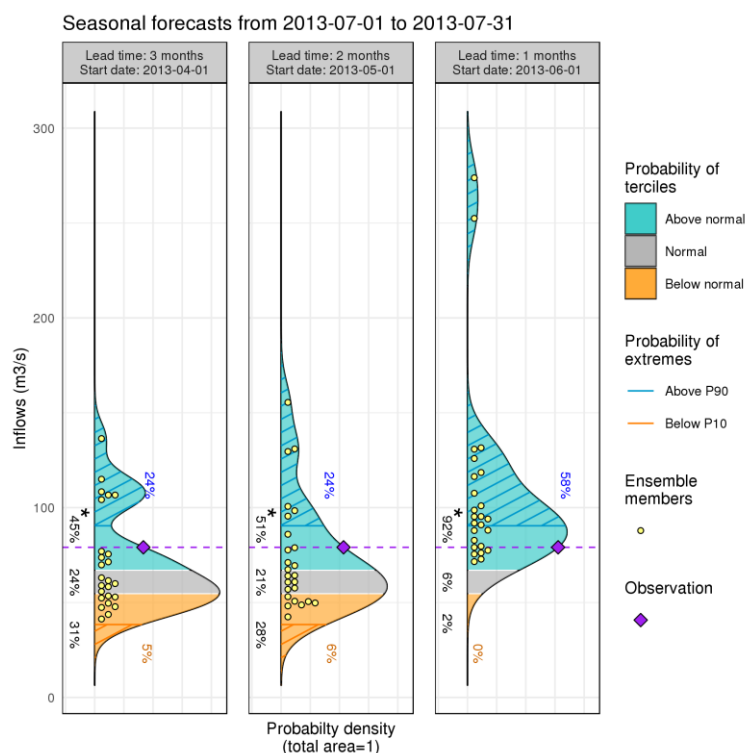
For this case study, the SEAS5 forecast for the analysed variables lacked skill (see Annex 3) with generally a very small probability of capturing the event in general terms. Regarding precipitation (Figure 31), forecast showed an almost equal probability of the event being above or within the normal conditions, and at short lead times the forecasts were tending towards an event within the normal conditions. However, the precipitation was below normal conditions, and none of the three initialisations was capable of capturing this. Interestingly, from a hydrological perspective, the forecasts of inflows at three different rivers showed some accuracy (see Figure 32, Figure 33 and Figure 34 for the Rhine, Neckar and Elber rivers respectively). Results for the Rhine river show that the inflows for this event were below normal conditions, and this was well predicted 3 and 2 months in advance. However, the forecast that initialised closer to the actual event indicated high probability for inflow being above normal conditions. For the Neckar river, the event was persistently predicted to be above normal conditions, as further justified by the observation. The probability of the inflow being above normal conditions increased with shorter lead times. Similar patterns were also observed for the Elbe river.



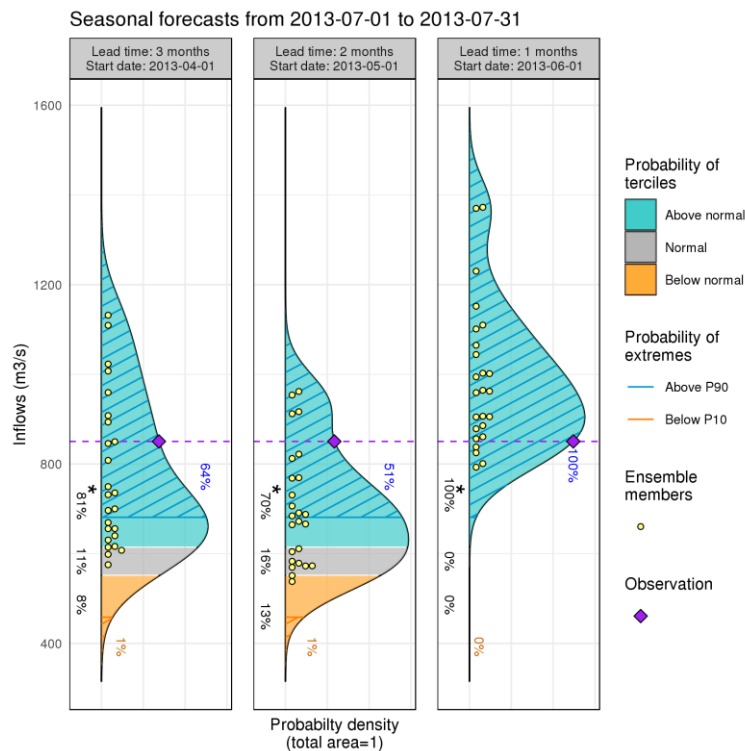
**Figure 31: Precipitation forecasts for July 2013 issued three, two and one months in advance for Germany.**



**Figure 32: Inflows forecasts for July 2013 issued three, two and one months in advance for the Rhine river, Germany.**

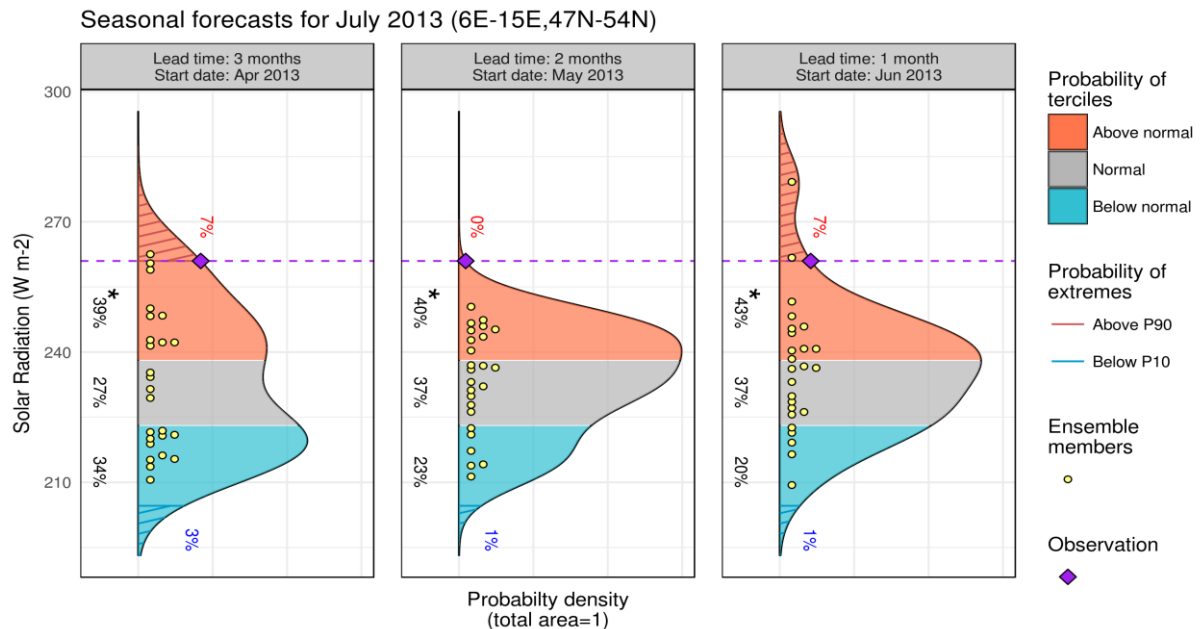


**Figure 33: Inflows forecasts for July 2014 issued three, two and one months in advance for the Neckar river, Germany.**

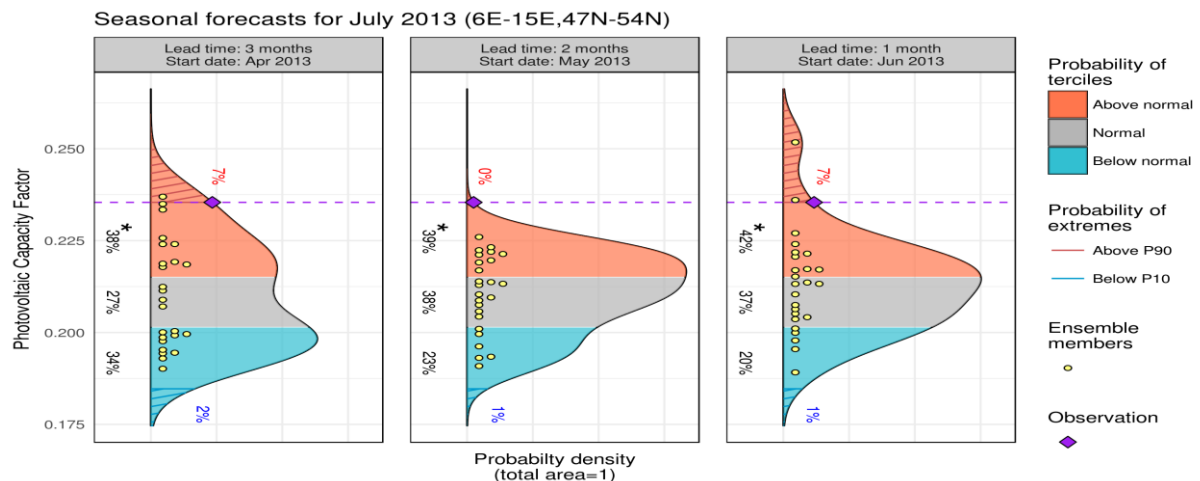


**Figure 34: Inflows forecasts for July 2014 issued three, two and one months in advance for the Elbe river, Germany.**

Surface solar radiation and photovoltaic capacity factor forecasts (Figure 35 and Figure 36) lacked skill (see Annex 3; negative skill for both probabilistic and deterministic metrics) with generally a very small probability of capturing the event. Forecasts at lead times 1 and 2 months showed a slightly larger probability of the event being above the normal conditions, in agreement with observations.

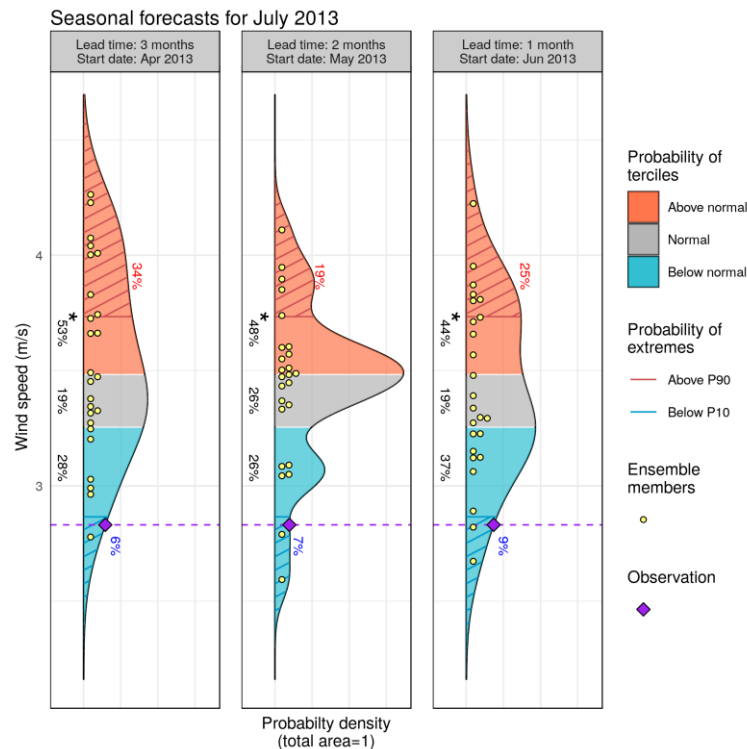


**Figure 35: Surface solar radiation forecasts for July 2013 issued three, two and one months in advance for domain (6E-15E, 47N-54N). Methodology: mean inflation calibration to ERA-Interim, based on 38 year hindcast.**



**Figure 36: Photovoltaic capacity factor forecasts for July 2013 issued three, two and one months in advance for domain (6E-15E, 47N-54N). Methodology: mean inflation calibration to ERA-Interim, based on 38 year hindcast.**

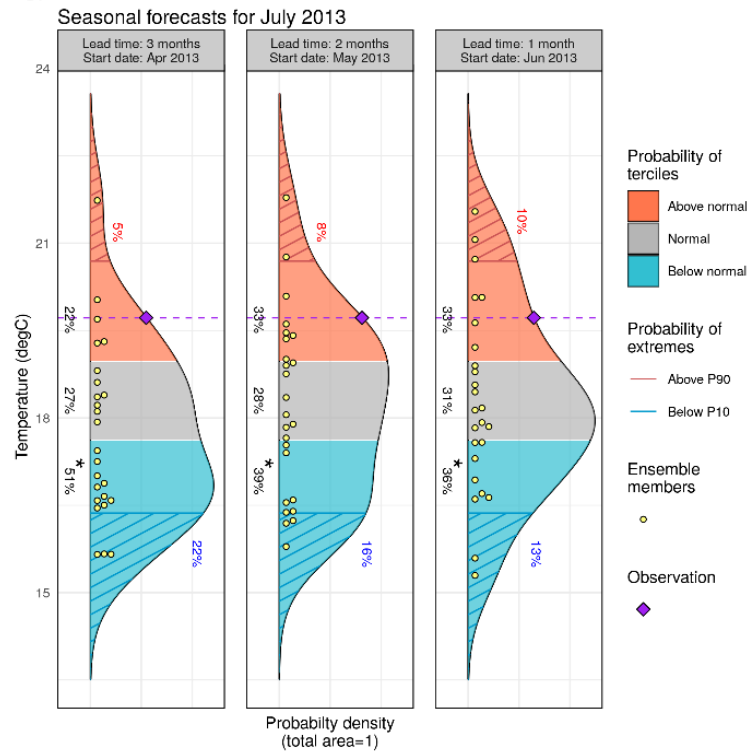
Wind speed predictions (Figure 37) showed a large spread among ensemble members with negative skill scores indicating a very small probability of capturing the wind drought. Forecasts for the three lead times (1 to 3 months) showed a slightly larger probability of above normal wind conditions, in contrast to observations.



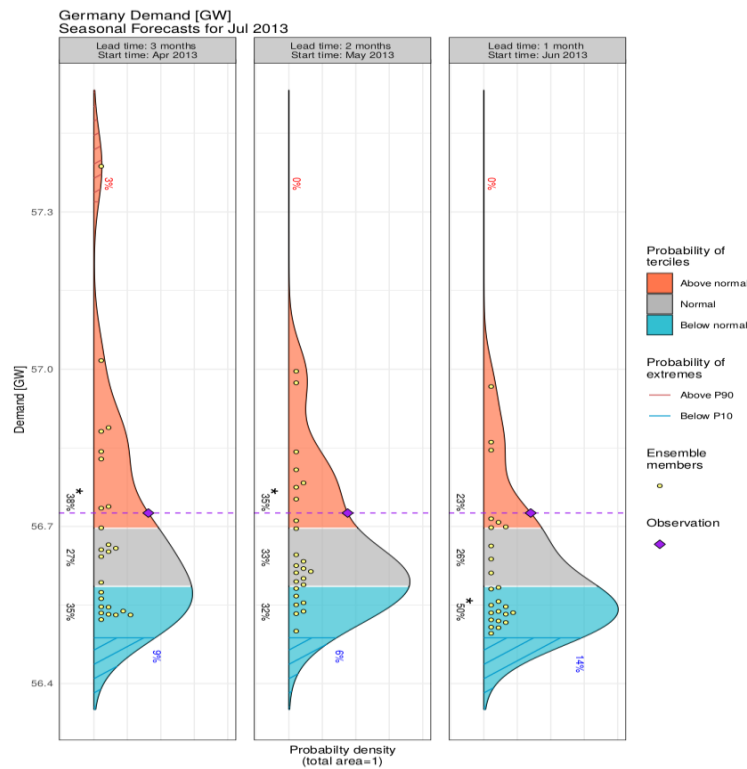
**Figure 37: Wind speed forecasts for July 2013 issued three, two and one months in advance for the domain of study: 6–15°E and 47–54.5°N.**

In line with the variables to predict renewable energy supply, temperature forecasts (Figure 38) could not capture the hot wave for July 2013. Predictions showed probabilities close to climatology in a context of near zero skill. Accordingly, the closely related demand forecast could not anticipate the increase in energy demand for cooling after the temperature increase (Figure 39). Skill scores associated to the energy demand forecast were negative for all lead times, a situation not encouraging the use of the predictions for decision making.





**Figure 38: Temperature forecasts for July 2013 issued three, two and one months in advance for Germany.**



**Figure 39: Demand forecasts for July 2013 issued three, two and one months in advance for Germany.**

### 6.3. Case study 3 – Spain 2016

According to the Spanish TSO, in 2016 the installed wind power capacity represented 22.8% of the total capacity for electricity generation in Spain, and wind energy supplied 19.2% of the demand. This high level of wind power penetration can have a significant impact on the energy market in periods with lower than normal wind power production. At the beginning of September 2016, the Iberian Peninsula was affected by a significant heat wave that combined with a notable decrease of wind speed through all central and southern Europe.

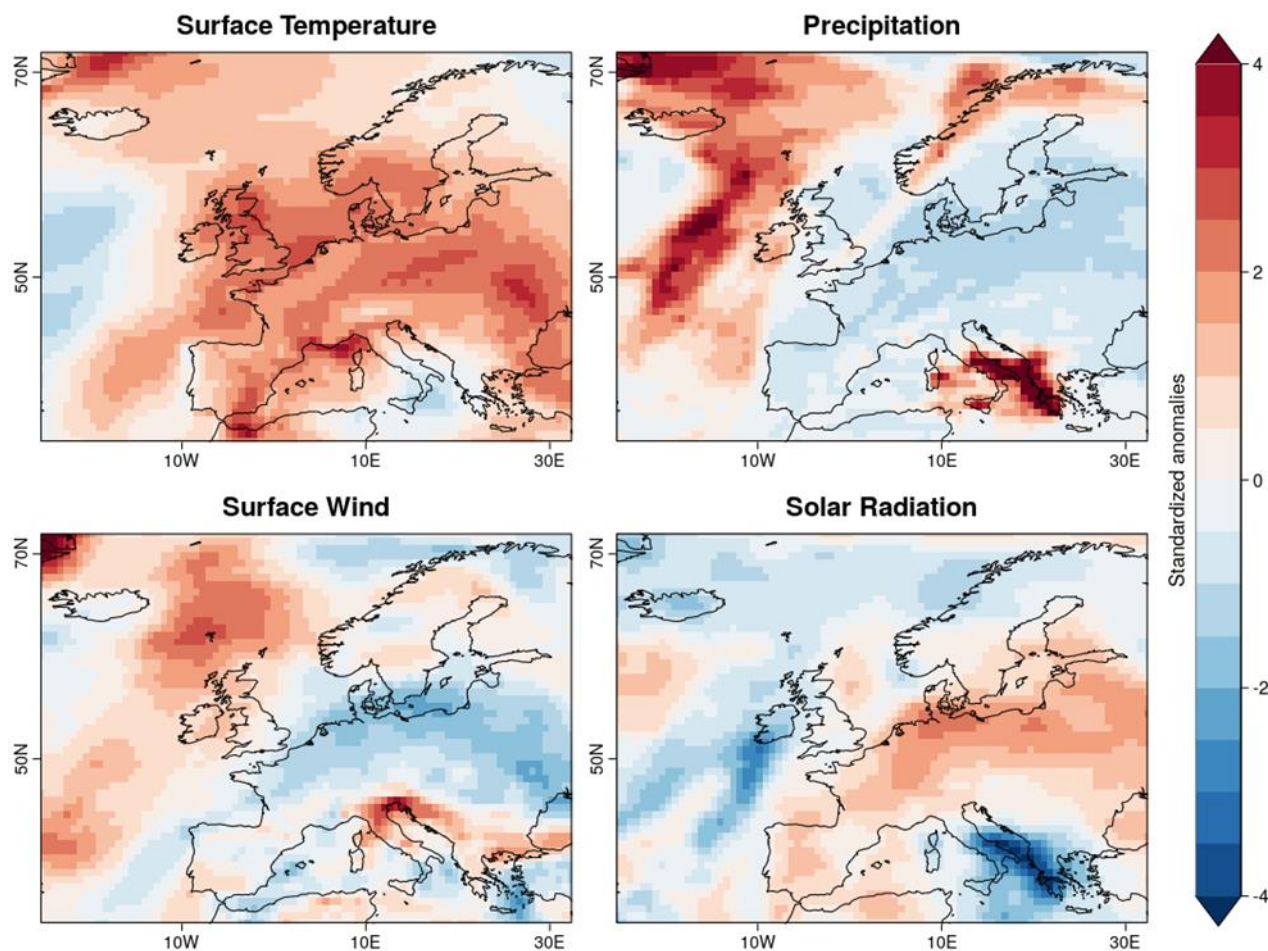
Heat wave and wind drought			
Region:	Spain	Period:	30 Aug-5 Sep 2016
Forecast type:	Sub-seasonal	Main interest:	Demand and wind
Forecast variables:	Wind speed, temperature, demand and demand net wind		

**Table 8: Region, period, forecast type and main interest for case study 3.**

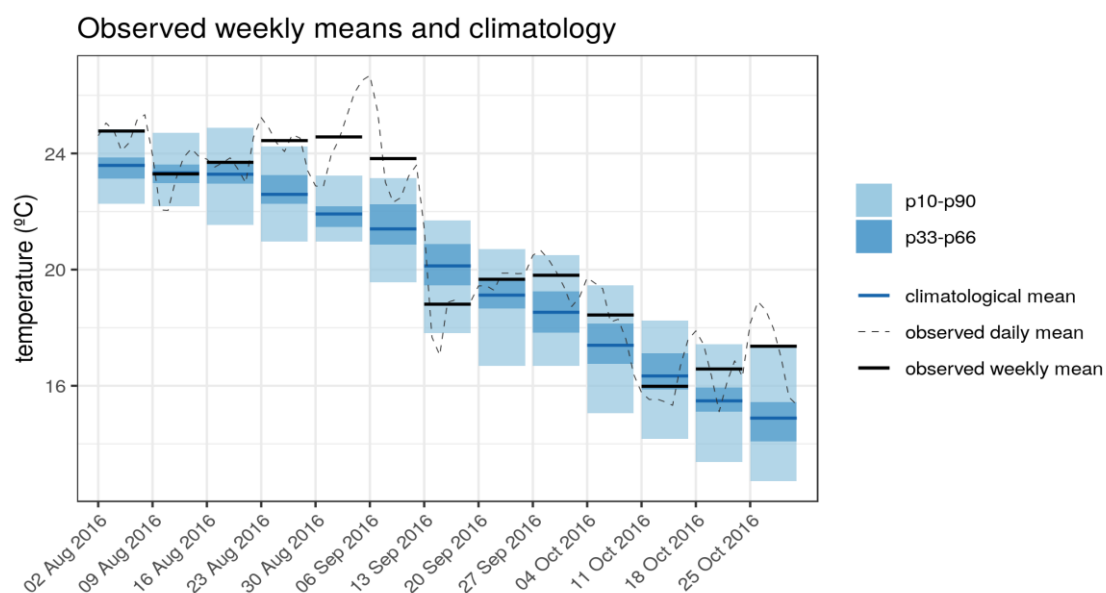
#### Event description

Anomaly maps for the week starting on 6th of September 2016 (Figure 40) show that over the Iberian Peninsula temperature was generally higher than the climatological average (top left) and surface wind (bottom left) was slightly below.

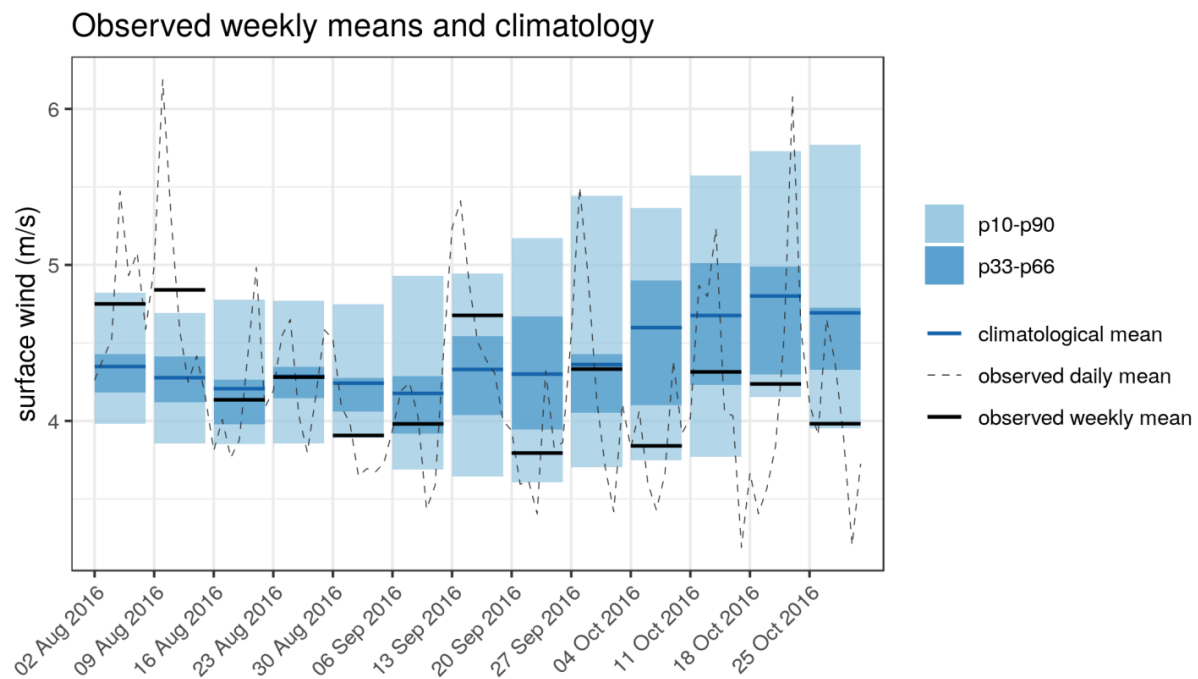
The weekly evolution of the average temperature, represented by the black solid line in Figure 41, has values higher than 90th percentile of the climatological distribution (light blue shadow) for three weeks from 23 of August to 12 of September. The average weekly temperatures were about 24 °C and, therefore, with an expected increase of the energy demand. Figure 42 shows the weekly evolution of the average surface wind. The values of wind are close to the P10 for the week that starts on 30th of August and close to P33 for the week that starts on 6th of September.



**Figure 40: Standardized anomalies of temperature, precipitation, wind and radiation for week starting on 30 of August 2016, obtained from ERA-Interim.**



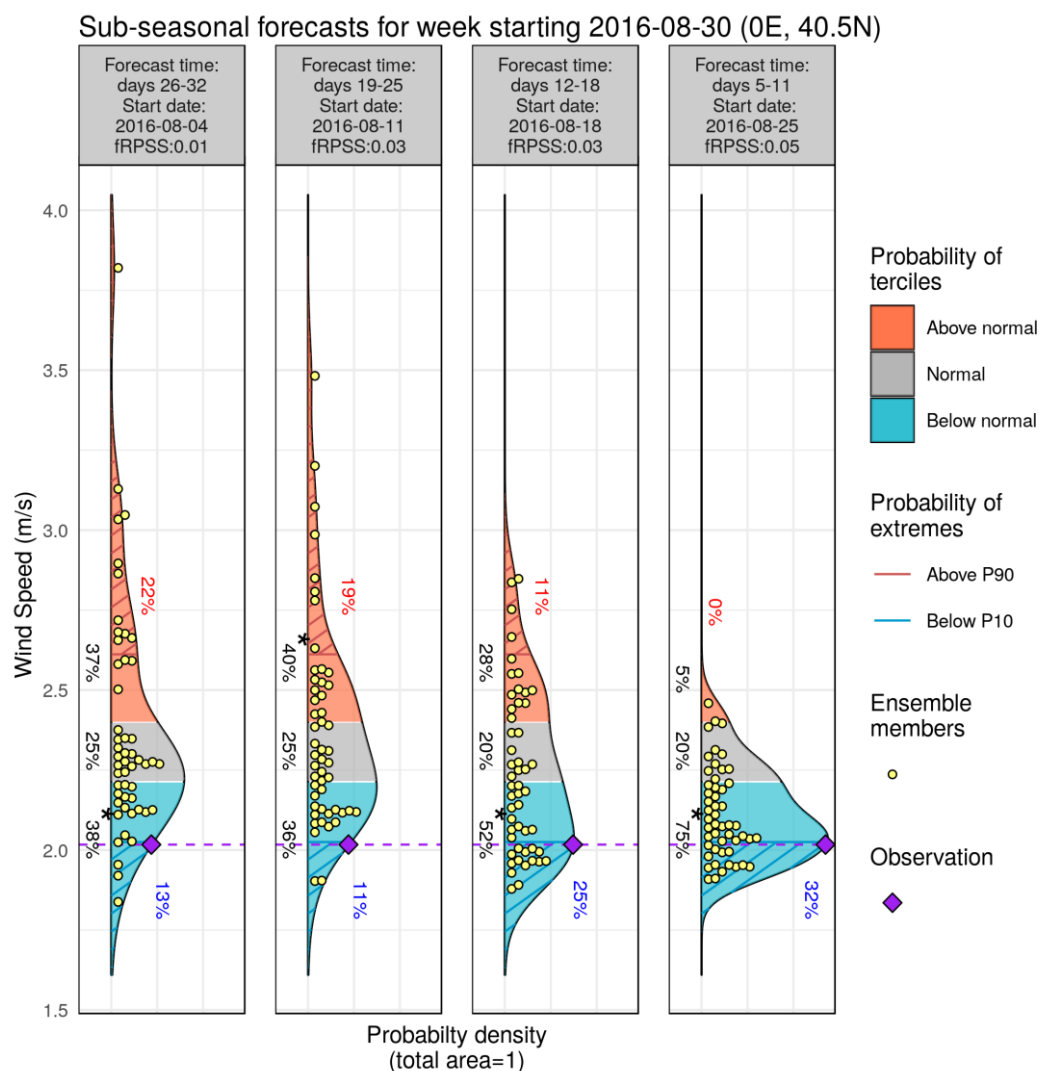
**Figure 41: Weekly evolution of the observed temperature in the region 8.5°W–3°E and 36.5–43.5°N from August to October compared to the climatological distribution. Values obtained from ERA-Interim reanalysis.**



**Figure 42: Weekly evolution of the observed surface wind in the region 8.5°W–3°E and 36.5–43.5°N from August to October compared to the climatological distribution. Values obtained from ERA-Interim reanalysis.**

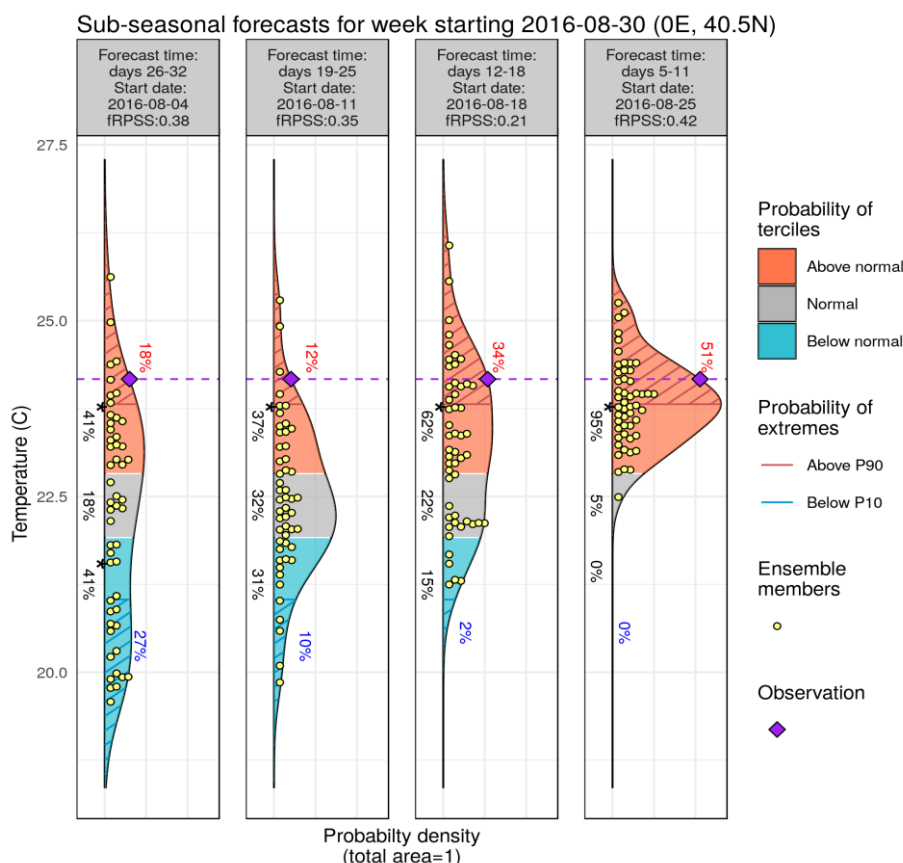
### Available forecasts

In this case, the forecasts are issued for a specific point located in the west of Spain (41°N, 0.5°E). The forecast for wind speed issued four weeks in advance (Figure 43) was indicating a 38% probability of wind speed being lower than normal ; however the forecast indicated a similar probability or the wind speed being in the upper tercile, and the fairRPSS associated to this forecast was only 0.01. The following week's forecasts also indicated the lower tercile as the most likely with increasing probability as approaching the target (except for 3 weeks in advance issued on 11/08/2016 for which the probability of the upper tercile was higher). As lead time decreases, it can be seen how the ensemble members present less spread and the fairRPSS increases, although it remains low for all forecasts. The last available forecast (4 days in advance) showed high agreement in the wind speed being in the lower tercile, and the peak of the distribution coincided with the observation. However, given the low degree of skill (fairRPSS=0.05), the confidence that could have been attributed to this forecast is limited.



**Figure 43: Sub-seasonal forecast for 10m wind speed averaged for the week 30<sup>th</sup> August – 5<sup>th</sup> September 2016 for 0E, 40.5N. From right to left corresponds to forecasts launched from lead times week 1 to 4 (forecasts times 5-11 days, 12-18 days, 19-25 days, 26-32 days). Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).**

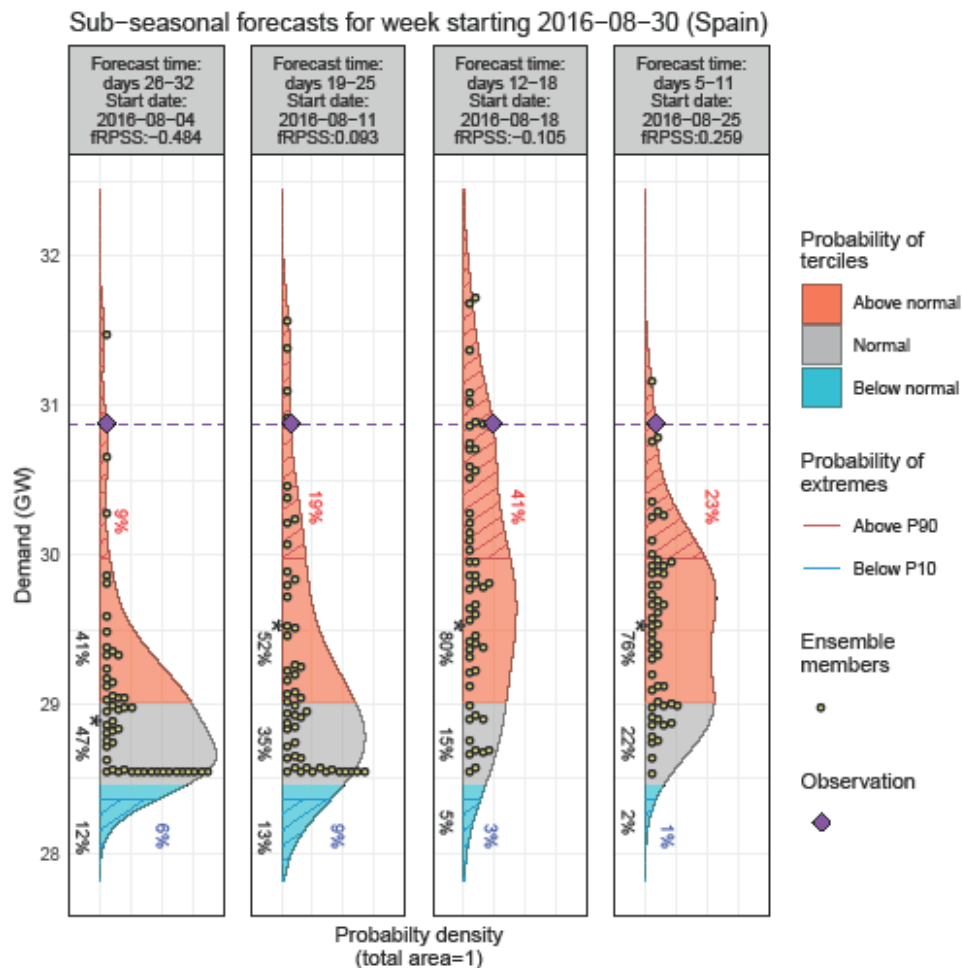
In the case of temperature (Figure 44), the forecasts issued four and three weeks in advance were not offering a clear indication of the heat wave, with comparable probabilities for all three terciles. From two weeks in advance, (lead time 11 days, forecast issued on the 16/08/2016) the probability of the higher tercile is 62% (fairRPSS=0.21) and the last available forecast gave a 95% (fairRPSS=0.42) chance of the temperatures being in the upper tercile. The hot conditions over Spain were predicted as early as 11 days in advance.



**Figure 44: Sub-seasonal forecast for 2 m temperature averaged for the week 30<sup>th</sup> August – 5<sup>th</sup> September 2016 for 0E, 40.5N. From right to left corresponds to forecasts launched from lead times week 1 to 4 (forecasts times 5-11 days, 12-18 days, 19-25 days, 26-32 days). Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).**

In conclusion, the heat wave was well predicted from week 2. In the case of wind, although the forecast issued a week in advance was correct, the low skill associated with it would have not provided enough confidence to base a decision on it.

The behaviour of the forecast for the demand (Figure 45) reflects the features of the forecast for temperature (Figure 44). Forecast issued 4 weeks in advance was indicating a 47% probability of demand being near normal and a 41% probability of demand being above normal. The probability of demand being above normal increased in forecast issued 3 weeks in advance (52%), 2 weeks in advance (80%) and slightly decrease in forecast issued 1 week in advance (76%). However, the confidence associated with forecast issued 1 week (fRPSS 0.259) was higher than the confidence (fRPSS -0.105) associated with forecast issued 2 weeks before.



**Figure 45: Demand forecasts for 30 August 2016 issued four, three, two and one weeks in advance over Spain.**

All the forecast for Demand Net Wind (Figure 46) indicated the tercile consistent with the observed value. The probability increased from 39% for forecast issued 4 weeks in advance to 69 % for forecast issued 1 week in advance. However, the confidence associated with those forecasts was low (fRPSS negative or lower than 0.2) and most of the members underestimated the observed value.



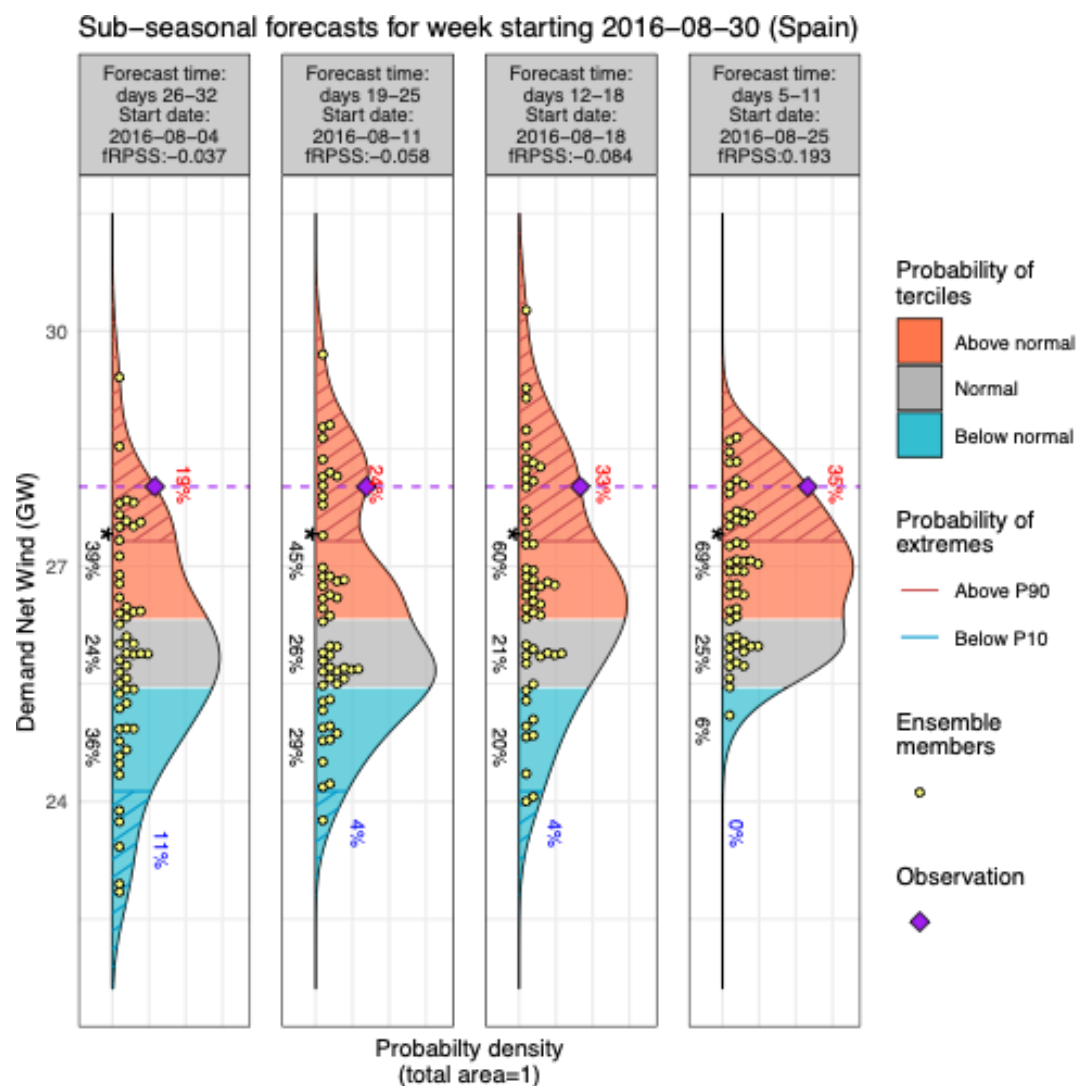


Figure 46: Demand net wind forecasts for 30 August 2016 issued four, three, two and one weeks in advance over Spain.



## 6.4. Case study 4 – Sweden 2015

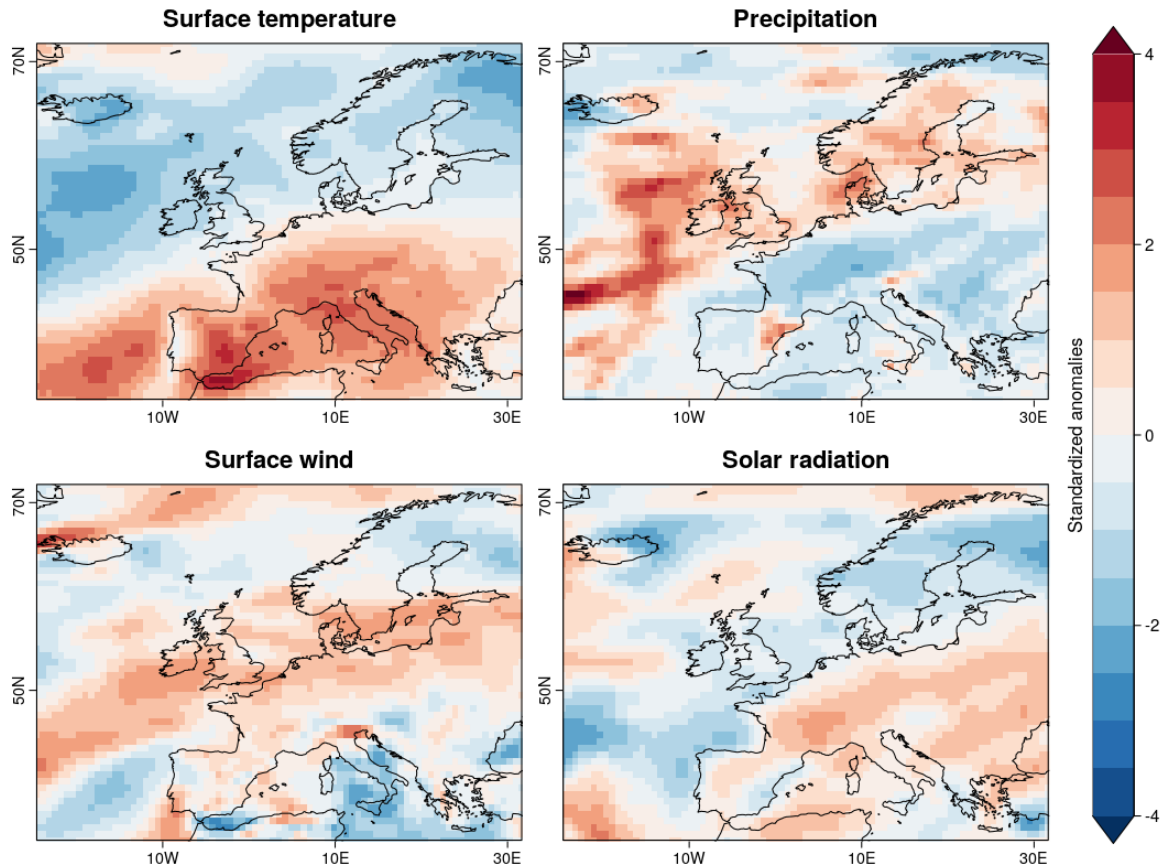
In July 2015, a combined snowmelt and rain caused a lot of unproductive release of reservoir water in the Umeälven river basin, Sweden. The reservoir was not managed appropriately without releasing enough water earlier in June/July. This was due to inaccurate hydrological forecasts that predicted a lot of remaining snow for melting. In the first weeks of July, the melting runoff stopped due to low temperatures. However, snow was still available which flowed to the reservoir later. The lack of accurate information about snow availability resulted in a significant economic loss for hydropower generators.

Spring flood in Sweden			
Region:	Sweden	Period:	May – Jul, 2015
Forecast type:	Seasonal	Main interest:	Hydro
Forecast variables:	Precipitation, snow water equivalent, and inflows		

**Table 9: Region, period, forecast type and main interest for case study 4.**

### Event description

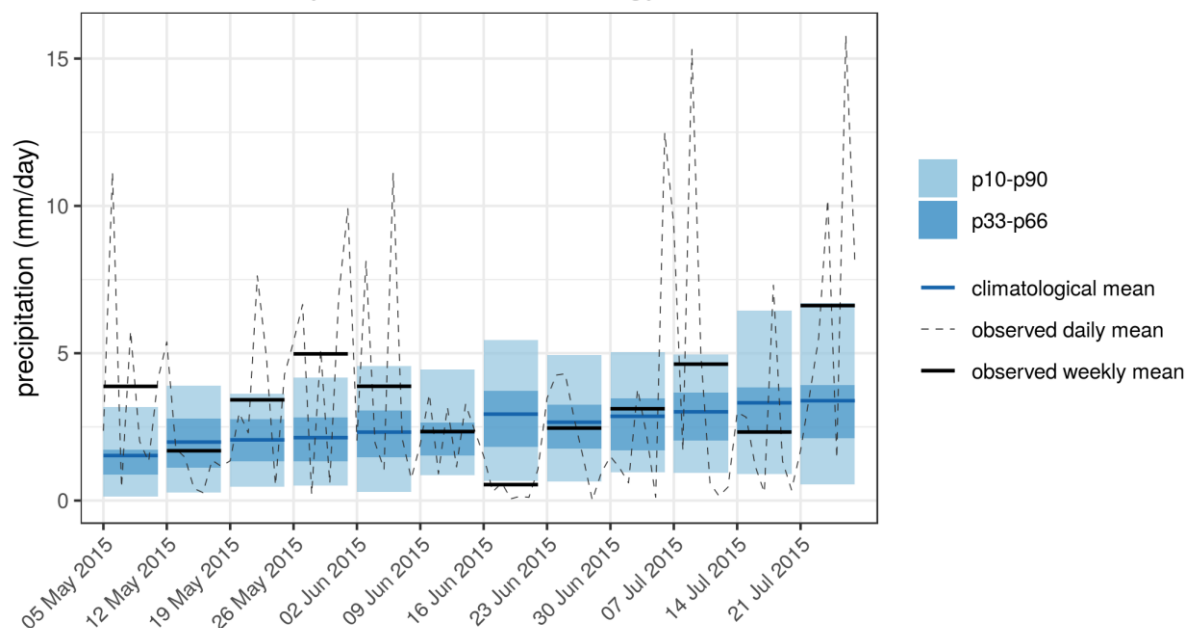
During the period May to July 2015, surface temperature was below normal over Scandinavia (Figure 47). Surface wind speeds were below normal in the central and northern part of Sweden, while the southern part of the country experienced above normal wind speed conditions. Precipitation was also above normal for the entire Sweden, which is one of the drivers for this case study. Finally, solar radiation was below normal for the entire country in agreement with precipitation patterns.



**Figure 47: Standardized anomalies of temperature, precipitation, wind and radiation for July 2015, obtained from ERA-Interim.**

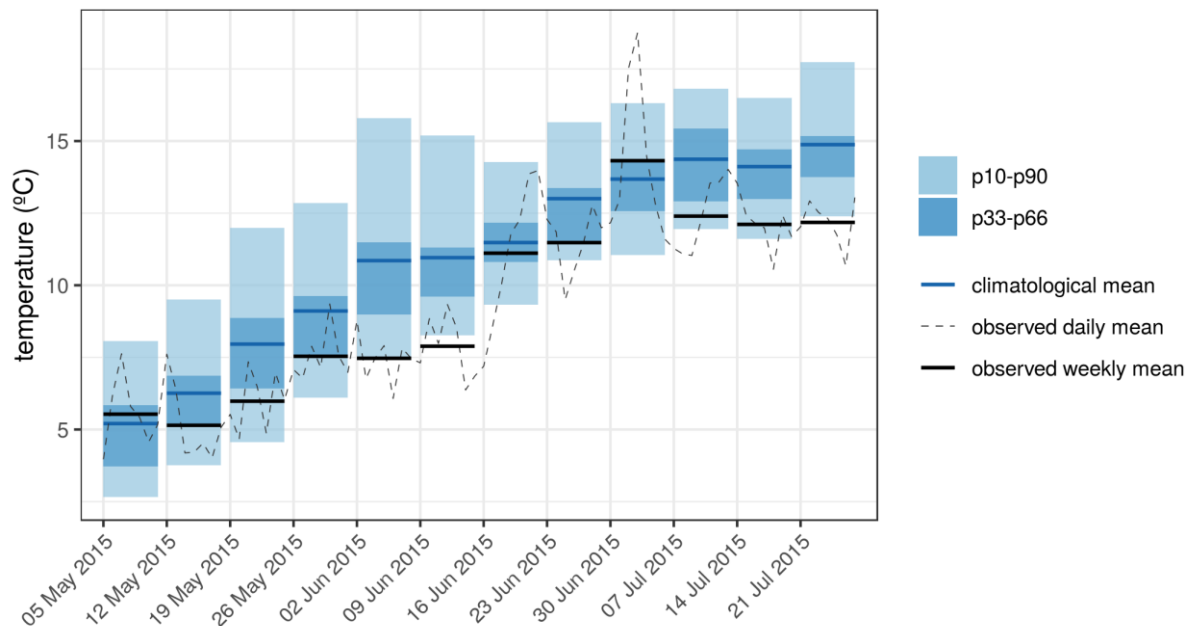
Analysis of the weekly anomalies for the period of interest (May – July) shows that for some weeks in May and July the precipitation was above 33<sup>th</sup> percentile with generally high daily and weekly temporal variability (Figure 48). Particularly during the end of May and beginning of June the variability was high and the mean close to be above the 90<sup>th</sup> percentile. Temperature on the other hand was below normal conditions for almost the entire May-July period. However during the beginning and mid-June, temperature was even lower than the 10<sup>th</sup> percentile (Figure 49). The daily variation was also quite strong with the observed temperature differing from the climatological mean.

### Observed weekly means and climatology



**Figure 48: Weekly evolution of the observed precipitation in the region 13–21.5°E and 63–66°N during May and June 2015 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis.**

### Observed weekly means and climatology



**Figure 49: Weekly evolution of the observed temperature in the region 13–21.5°E and 63–66°N during May and June 2015 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis.**

#### Available forecasts

Seasonal forecasts of precipitation, snow water equivalent and inflows have been produced for the Umeälven river basin using ECMWF SEAS5 (Table 10). Forecasts issued one, two and

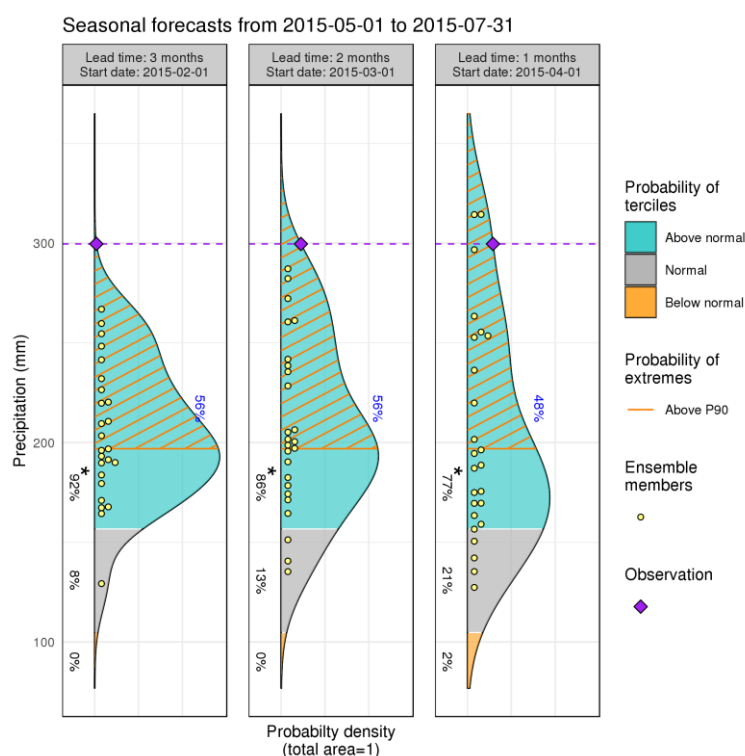
three months in advance are presented. The predictions have been bias-adjusted using the DBS quantile-quantile method with a hindcast covering 1993 to 2015 and hydroGFD meteorological observations for the same period.

Start date	Valid period	Lead time	Variables	Forecast system
Feb 2015	May-July 2015	3 to 5 months	<ul style="list-style-type: none"> <li>- precipitation</li> <li>- snow water equivalent</li> <li>- inflows</li> </ul>	SEAS5 from C3S 1 degree resolution 25 ensemble members
Mar 2015	May-July 2015	2 to 4 months		
Apr 2015	May-July 2015	1 to 3 months		

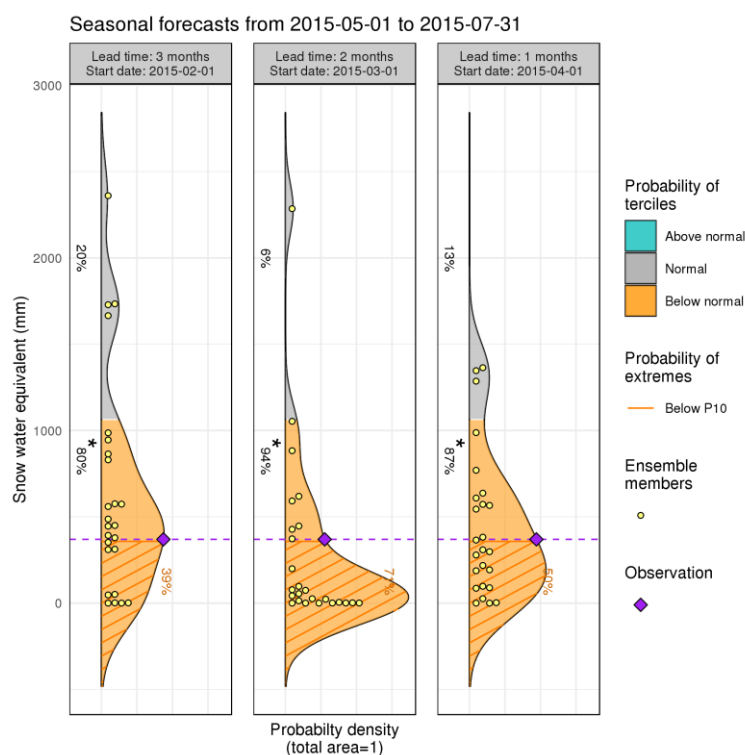
**Table 10: Start dates, lead times, valid period and variables for each of the forecasts presented.**

Figure 50, Figure 51 and Figure 52 present the precipitation, snow water equivalent and inflow forecasts for the three lead times respectively. The forecasts show that the precipitation amount occurred in May-July were significant and extreme, and quite above the normal range. Although the forecasts issued in February were capable of predicting with high probability the above normal conditions, the amount was not accurate. This lack of precipitation forecast accuracy is challenging in decision-making of the hydropower reservoirs. Inflows were also forecasted to be above the normal conditions (as was indeed the actual event); however similarly to precipitation, the actual inflows were much higher than the forecasted inflows. Forecasts of snow water equivalent on the other hand were indicated to be below normal conditions with the observation being different. As the river basin is snow dominated, lack of predictability in snow in terms of volumes and timing can have a strong impact on inflows and hence decision-making. For snow water equivalent over the period May-July, the forecasts were relatively better forecasted than the other variables, yet the temporal distribution of snow (which cannot be diagnosed here) can be of high importance.

Looking at skill scores for precipitation, snow water equivalent and inflows Annex 3), the metrics indicate different predictability potential between the variables. Inflows have a high skill for all lead months, whilst snow water equivalent achieves very high skill in the first lead time. Precipitation has the lowest skill, with skill for forecast month three being negative (i.e. climatology performs better than SEAS5). The metrics indicate that precipitation needs more caution when used in decision making particularly when predicting the far seasons. In general, as expected, the skill gets stronger with shorter lead times. Brier skill scores are high for snow water equivalent and inflows, which indicate potential for application in extreme conditions.

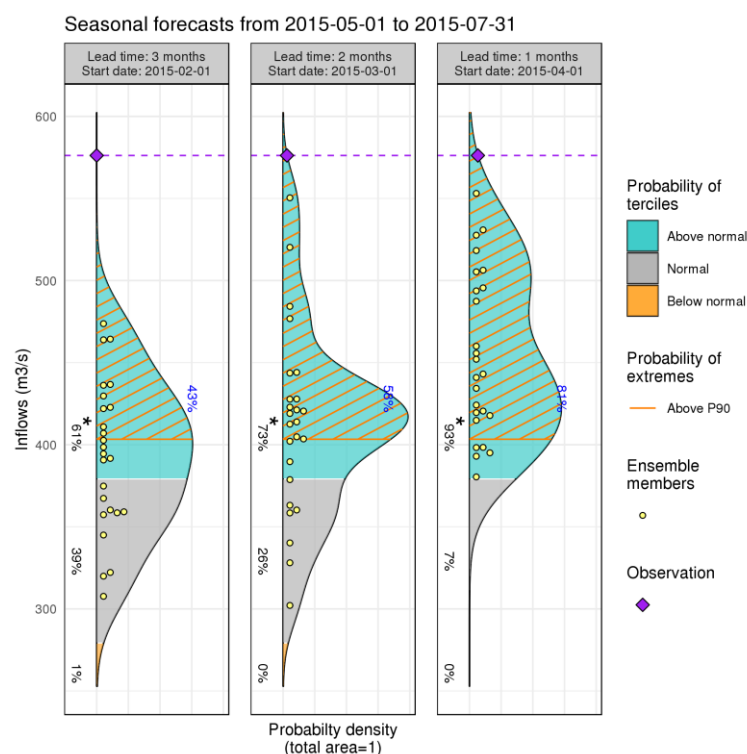


**Figure 50: Precipitation forecasts for May-July 2015 issued three, two and one months in advance for the Umeälven river basin.**



**Figure 51: Snow water equivalent forecasts for May-July 2015 issued three, two and one months in advance for the Umeälven river basin.**





**Figure 52: Inflows forecasts for May-July 2015 issued three, two and one months in advance for the Umeälven river basin.**

## 6.5. Case study 5 – Romania 2014

On 31 of January 2014, the European Commission's Directorate-General for European Civil Protection and Humanitarian Aid Operations, reported that severe weather conditions (heavy snowfalls, low temperatures and rainfall) in central and eastern Europe, particularly Romania, caused power outages, and transportation problems. In Prahova (center-east of Romania), 8,500 families suffered from power failures (Reliefweb, 2014).

Freezing event in Romania			
Region:	Romania	Period:	28 Jan- 3 Feb 2014
Forecast type:	Sub-seasonal	Main interest:	Wind energy
Forecast variables:	Temperature and minimum temperature		

**Table 11: Region, period, forecast type and main interest for case study 5.**

A very strong icing event occurred in January and February 2014 in Romania. In some of the wind farms in the country both the rotors and the road accesses were frozen for several days.

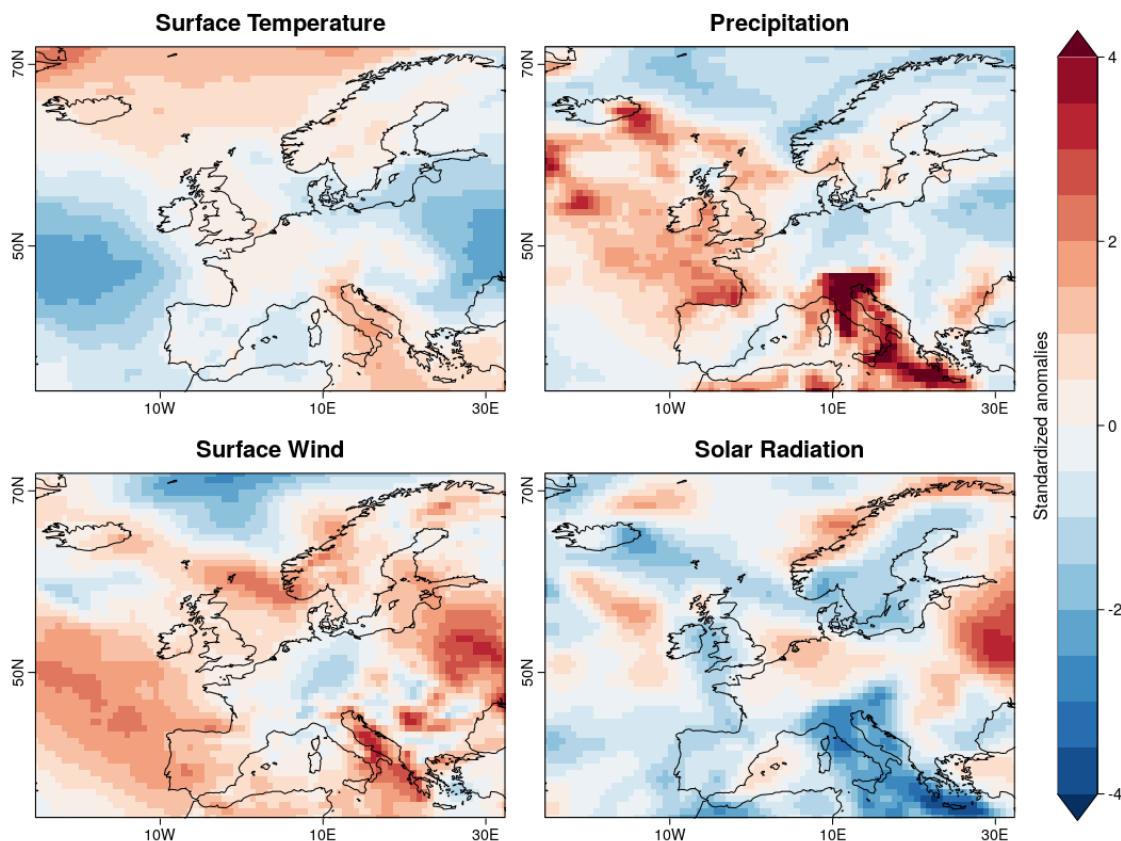
The parks, mostly in the center and south of the country, had to be stopped and, in some cases, they lost all communication receptions. Whenever the park manager could not access the site, the day ahead market offers had to be corrected manually. The impact of this event, in addition to the losses inherent to the energy sales, came from the cost of deviations of the manual correction, based on what had happened the previous day. The worst situations were due to the installations' transients between start and stop. In Romania, the legislation is quite restrictive in this sense, and these deviation costs per MWh are usually very high. The stakeholders of this case study indicated that event information one or two weeks ahead would have been useful, at least to prevent the control center and take action.

### **Event description**

28th January to 3rd February 2014 was characterized by an east-west pressure dipole across Europe, with a region of high pressure over western Russia and low pressure to the west of the UK. Overall, this is associated with generally southerly flow over most of Central Europe. Consistent with this flow, temperatures over most of Western Europe were near-normal for the time of year (Figure 53).

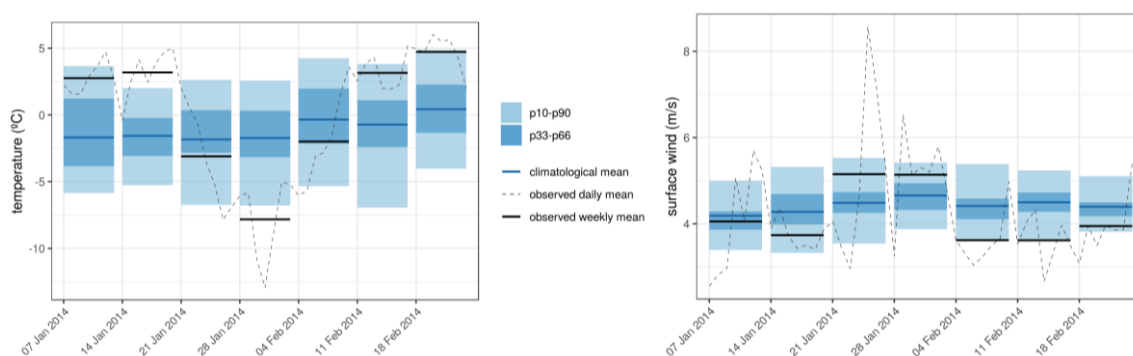
The general synoptic conditions, however, led to colder temperatures along the southern edge of the high-pressure anticyclone in the east of the region, particularly near Bulgaria and Romania. These cold conditions correspond to negative temperature anomalies exceeding 2 standard deviations below seasonal average in a region near the Black Sea, with weaker anomalies in the surrounding area.

Surface wind speeds over much of Europe are relatively unremarkable and, in the region around Romania, typically somewhat above seasonal norms (circa 1-2 standard deviations), with some local geographic variability.



**Figure 53: Observed and surface meteorological conditions compared to climatology during 28th Jan – 3rd Feb 2014.**

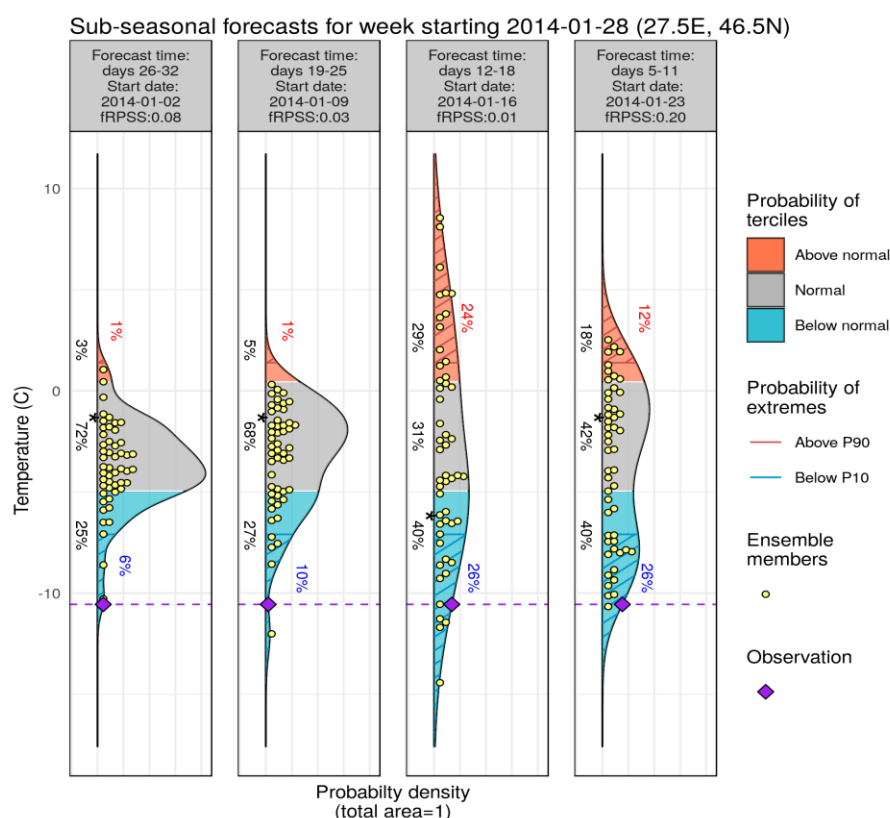
Figure 54 describes the meteorological evolution in the region of Romania over the period January to February 2014. The period as a whole contains three weeks of below seasonal normal temperature (28th January – 3rd February and one-week either side), with above seasonal norm temperatures either side of this period. Inspection of the daily time-series (dotted line, Figure 54), suggests below seasonal norm temperatures persisted from ~24th January to ~8th February. Wind-speeds, by contrast, are generally near-normal or above-normal throughout most of the period.



**Figure 54: Observed and climatological surface air temperature (left) and 10m wind speed (right) in the region 24–29.5°E and 43.5–47°N during January and February 2014.**

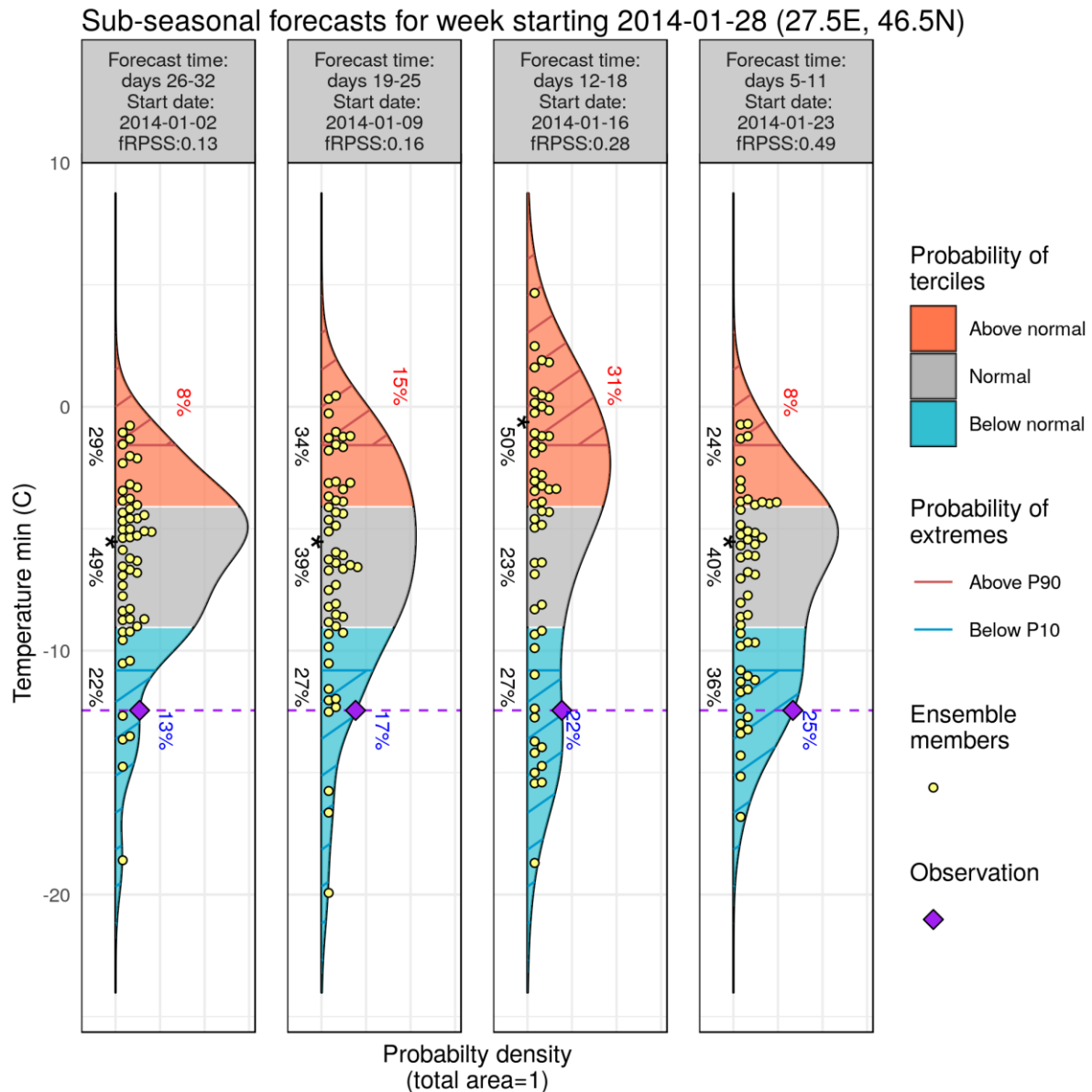
This event was characterised by very low temperatures in Eastern Europe during the week 28/01/2014-03/02/2014. We look at the temperature forecast for a specific grid point (27.5°E, 46.6°N). The weekly average temperature for that point for the target week from the ERA-Interim reanalysis was around -11°C. The forecasts were indicating high probabilities of the temperature being in the middle tercile 3 and 4 weeks in advance. In the forecasts issued 2 and 1 weeks in advance, the probabilities of the lower tercile increased, and the forecasts showed similar probabilities for the middle and lower terciles. A broad spread in the ensemble members is seen even for the last forecast. However, some individual members were close to the observation and thus, the probability attributed to the 10th percentile was 26% in both forecasts issued on the 16/01/2014 and the 23/01/2014.

It also must be noted that the model version of ECMWF monthly system employed in these forecasts is CY40R1, since this case study occurs in 2014. ECMWF monthly versions before CY41R1 (implemented in May 2015) have associated hindcast runs of five ensemble members (as opposed to eleven members thereafter). This lower number may affect the quality of the calibration process, as the calibration parameters are less robust.



**Figure 55: Sub-seasonal forecast for temperature for a specific grid point (27.5 E, 46.6 N) for the week 28/01/2014-03/02/2014. From left to right corresponds to forecasts launched from lead times week 4 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).**

The behaviour of the forecast for minimum temperature (Figure 56) is similar to the forecast for temperature (Figure 55). Forecast issued 4, 3 and 1 weeks in advance was indicating probability of minimum temperature near normal (49, 39 and 40%, respectively). The confidence in predicting minimum temperature for this location and period is high, especially 2 and 1 weeks before the event, however, in this case, forecast didn't predicted the observations.



**Figure 56: Sub-seasonal forecast for minimum temperature for a specific grid point (27.5 E, 46.6 N) for the week 28/01/2014-03/02/2014. From left to right corresponds to forecasts launched from lead times week 4 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).**



## 6.6. Case study 6 – USA 2015

During the first months of 2015 (January–March), surface wind speeds were well below normal in most of the contiguous United States, which reduced substantially the power generation of most of the wind farms in the western part of the country. This had severe implications for wind farm owners who saw an important reduction in revenues, making difficult regular cash-flow operations, and also produced a depreciation of the value of their assets.

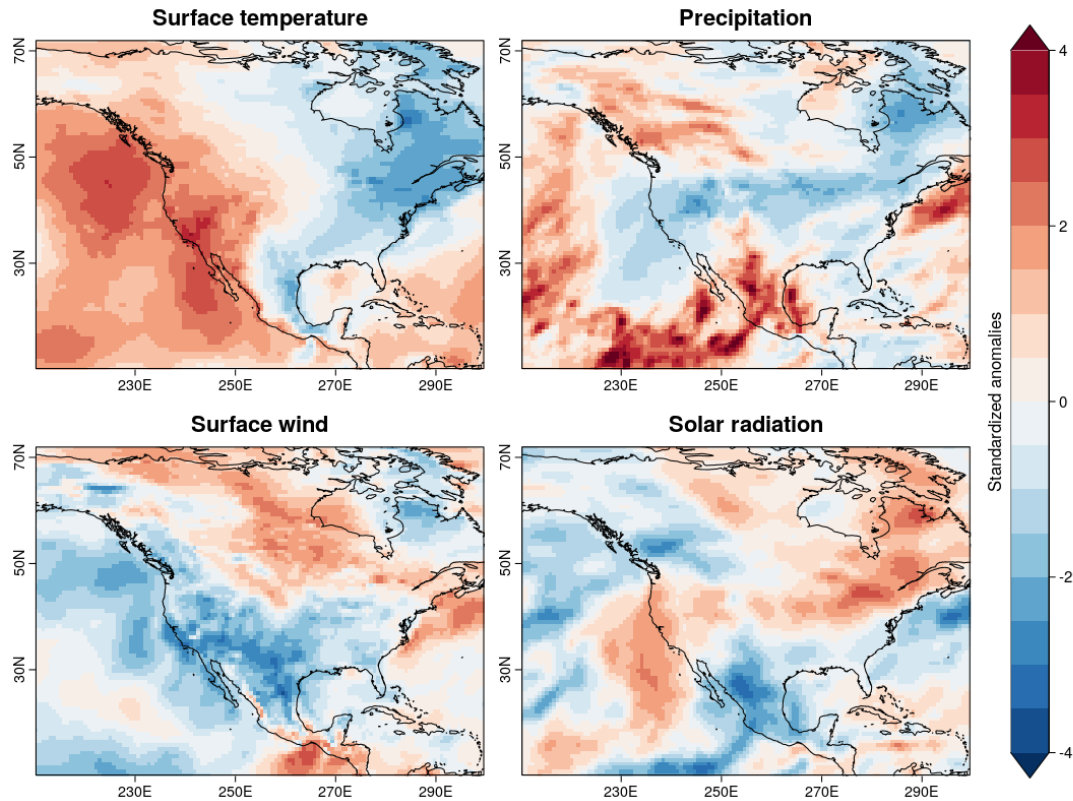
Wind drought in Western USA			
Region:	Western USA	Period:	Jan-Mar 2015
Forecast window:	Seasonal	Main interest:	Wind
Forecast variables:	Wind speed, wind power capacity factors		

**Table 12: Region, period, forecast type and main interest for case study 6.**

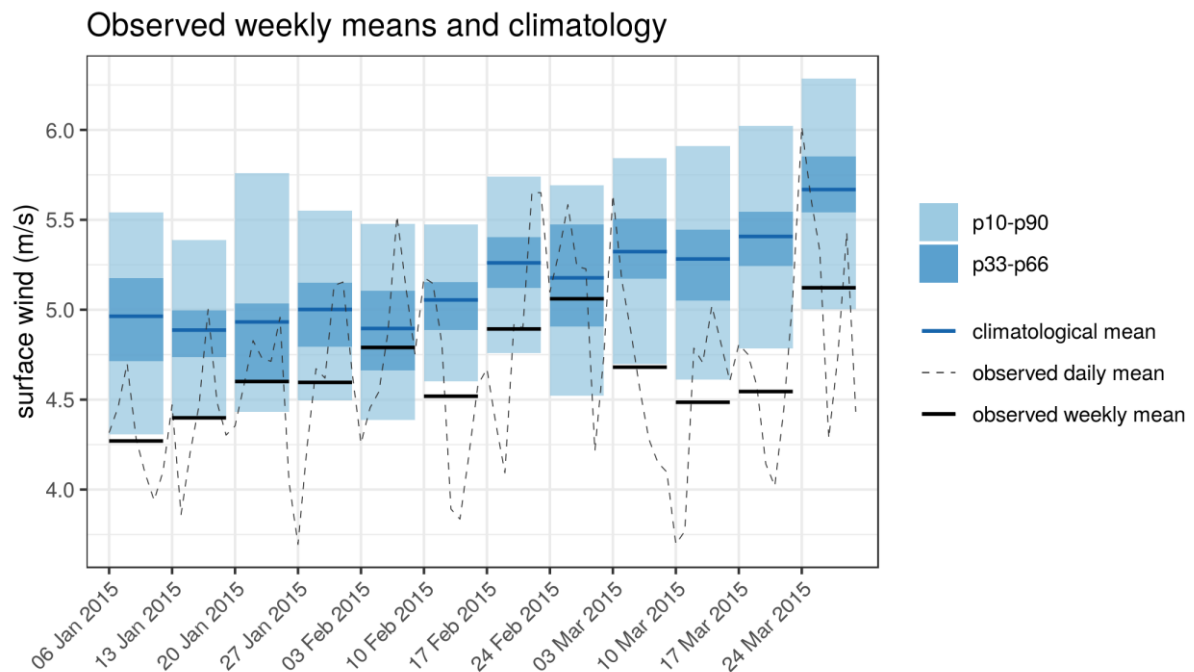
### Event description

Anomaly maps for the period show that surface winds in the south-western part of North America were more than three standard deviations below the climatological average (see Figure 57, bottom left panel). Simultaneous temperature anomalies were also recorded in the US, with very cold temperatures on the eastern coast, while strong precipitation events occurred in Mexico. The region most affected by the lack of wind episode (also known as wind drought) has been selected for this case study (124°W-95°W and 26°N-44°N).

The weekly evolution of average surface wind speed over the region during the three-month period is shown in Figure 58 (black solid line). The climatological mean and distribution for each week is also shown in blue. The observed values were below the P33 for ten out of the twelve weeks and below the P10 for six of them. The drought was especially strong during March. This continued and widespread lack of wind is exceptionally rare from the point of view of the climatological records in the past years. The possibility that a climate forcing is driving those anomalies seems very plausible in view of those anomalies.



**Figure 57: Standardized anomalies of temperature, precipitation, wind and radiation for the first quarter of 2015, obtained from ERA-Interim.**



**Figure 58: Weekly evolution of the observed surface wind in the region 124°W-95°W and 26°N-44°N during Q1 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis.**

## Available forecasts

Seasonal forecasts of surface wind speed have been produced for the region of interest using ECMWF SEAS5. Forecasts issued one, two and three months in advance are presented. The predictions have been bias adjusted with a hindcast covering 1993 to 2015 and ERA-Interim observations for the same period. The calibration method described in Torralba et al. (2018) has been used. This method adjusts the mean bias and inflates the spread of the distribution to produce unbiased and reliable forecasts. Seasonal averages have been used directly in the calibration procedure. A leave-one-out setting is employed to adjust the forecast without including the forecasts and observations of the period of interest.

Start date	Valid period	Lead time	Variables	Forecast system
Oct 2014	Jan-Mar 2015	3 to 5 months	-surface wind	SEAS5 from C3S 1x1deg resolution 25 ensemble members
Nov 2014	Jan-Mar 2015	2 to 4 months	-capacity factor iec1 -capacity factor iec2	
Dec 2014	Jan-Mar 2015	1 to 3 months	-capacity factor iec3	

**Table 13: Start dates, lead times, valid period and variables for each of the forecasts presented.**

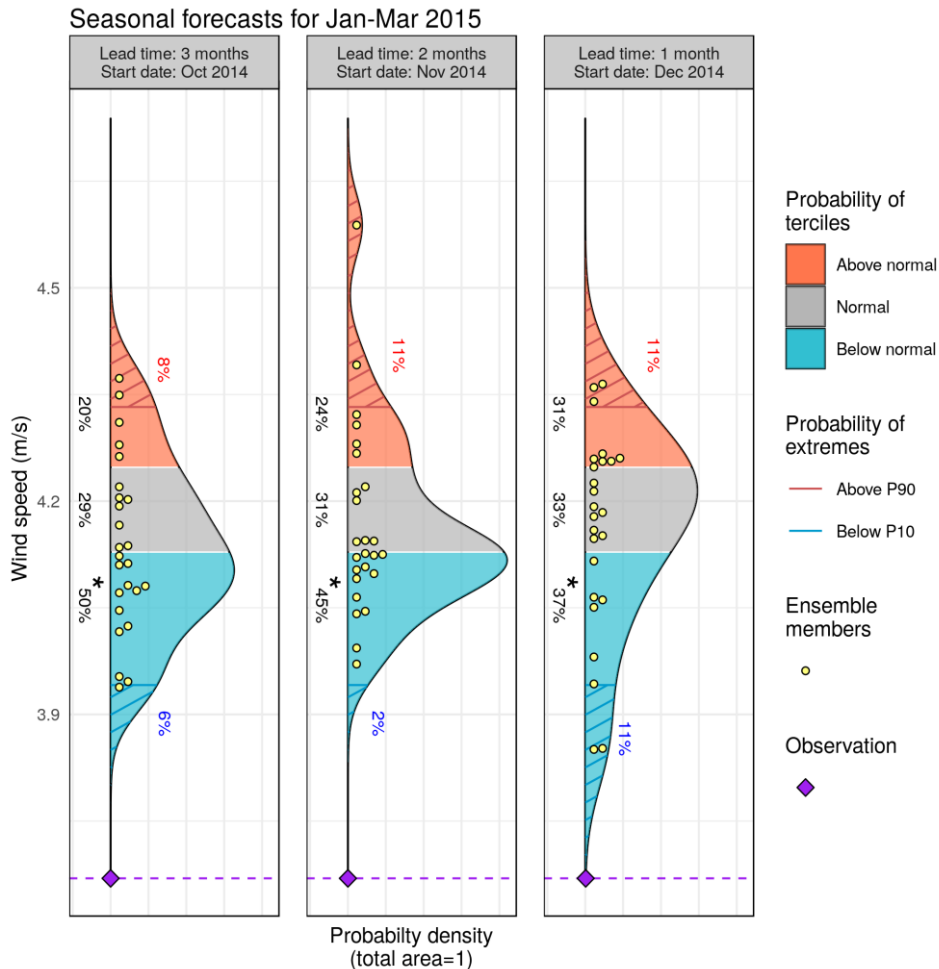
Additionally, capacity factor (CF) forecasts have been produced for three different turbines of IEC classes 1, 2 and 3. Those CF forecasts are produced from 6-hourly winds and bias adjusted with an empirical quantile mapping methodology. Once the wind speeds are corrected, a power curve is used to translate wind to electricity generation.

For each forecast, skill scores have been estimated using the hindcasts. This is key in order to understand the past performance of forecasts in this region. According to this skill estimates the forecasts can be trusted or need to be disregarded.

Figure 59 presents the surface wind forecasts for the three lead times. The forecasts show increased probabilities of below normal wind speed, with 50, 45 and 37% probability in the above normal tercile. But the forecasts do not indicate higher probabilities of an extreme event (below P10). The observed wind speeds were very low and fell below P10.

Looking at skill scores (Annex 3) we see that CRPSS (measuring the forecast distribution quality) is between 0.1 and 0.14 for the three lead times, while the RPSS (measuring the terciles forecast quality) are even higher (above 35% for all leads). Therefore, those tercile forecasts can be used for decision-making.

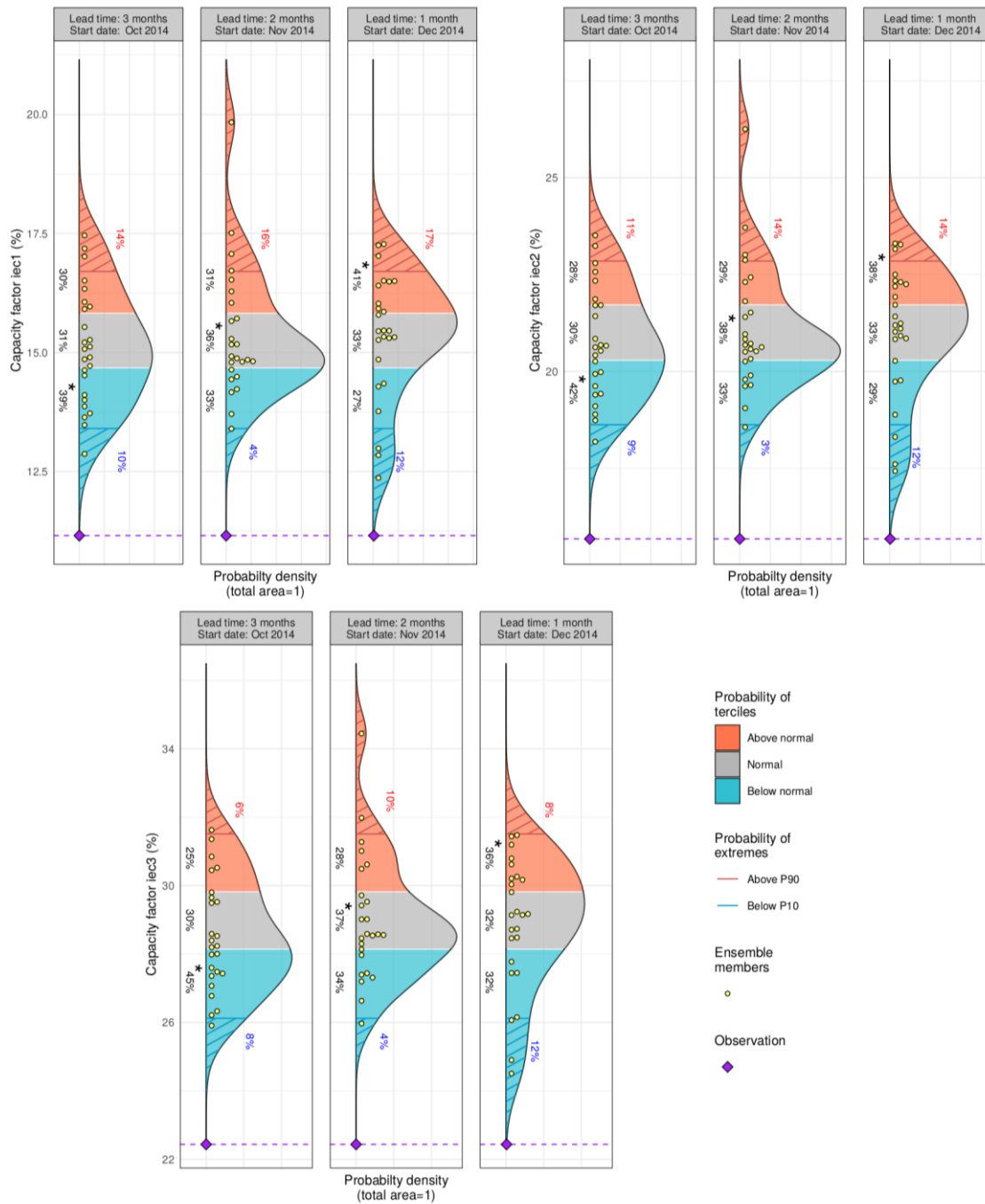
The quality of the extreme probabilities, evaluated with the Brier Scores, is slightly positive for P90, but negative for P10. Therefore, the low-probabilities of below-P10 winds shall not be trusted in those forecasts.



**Figure 59: Surface wind speed forecasts for Jan-Mar 2015 issued three, two and one months in advance for the western USA region.**

Forecasts for CF iec1, iec2 and iec3 (Figure 60) differ from wind forecasts in terms of tercile probabilities. For instance, the 1-month lead forecasts have higher probability of above-normal capacity factors, an opposite behaviour to what is found in wind speed forecasts. The skill scores for CF iec1, iec2 and iec3 are smaller than those for wind, but still quite good for its usage (Annex 3). The skill scores for P10 and P90 are negative in this case, advising against its usage.

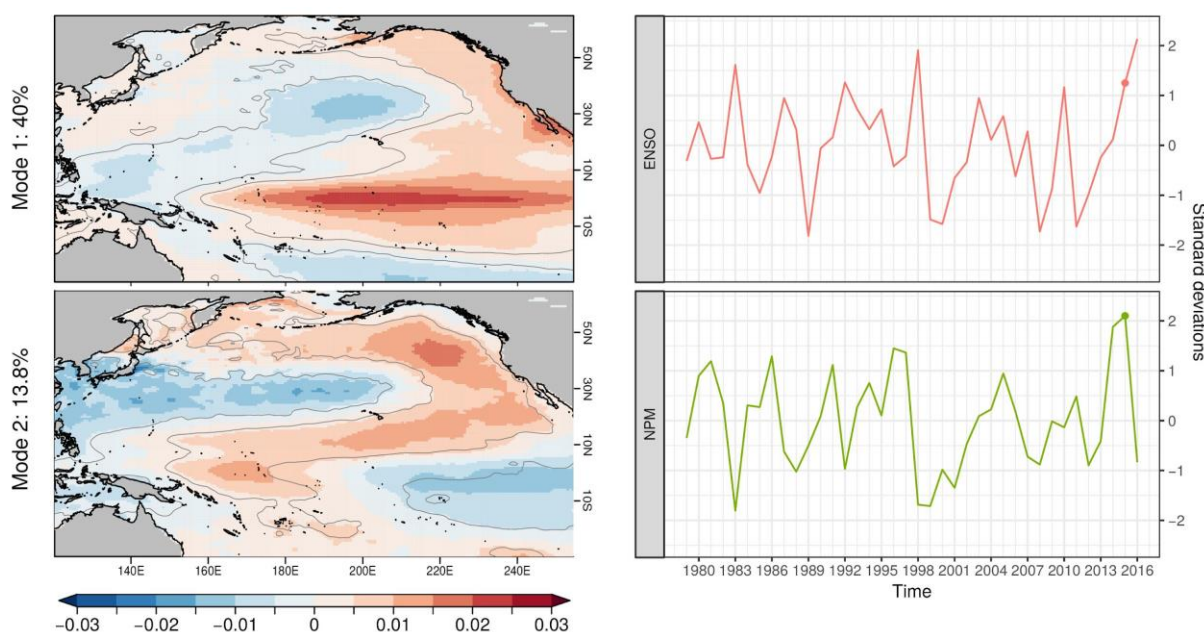
Seasonal forecasts for Jan–Mar 2015



**Figure 60: Capacity factor forecasts for IEC1, IEC2 and IEC3 class turbines, for the region and period of interest.**

## Climate drivers

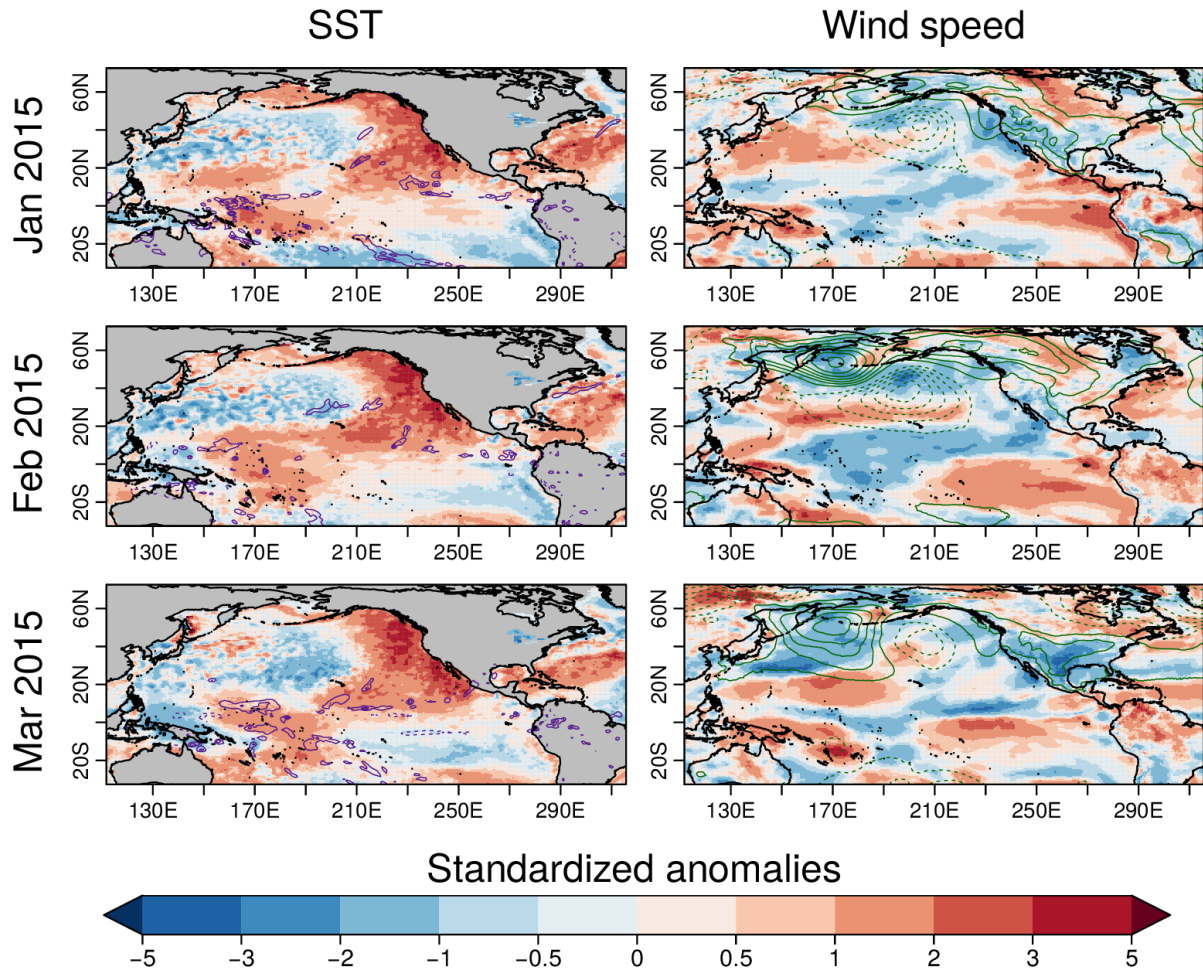
Although interannual variability of wind speeds in North America is known to be mainly dominated by ENSO state, a recent study by Lledó et al. (2018) has shown that another climate oscillation known as North Pacific Mode (NPM) can also influence wind speed variability in the south-western part of North America. During the first quarter of 2015, the NPM was in a highly positive phase (Figure 61 bottom right panel), reaching more than two standard deviations away from the mean in an unprecedented situation. Meanwhile ENSO was only in a moderate positive phase (Figure 61 top right panel).



**Figure 61: ENSO and NPM patterns (left) and its evolution during the last years for Q1 (extracted from Lledó et al. 2018).**

The ENSO state is known to influence the climate thousands of kilometres away from where it initially starts through Rossby waves that originate in the tropical Pacific Ocean and propagate towards the extratropics. The NPM teleconnections are similar in nature. The NPM positive phase is characterized by high sea surface temperatures (SST) in the western tropical Pacific Ocean. Those high SSTs enhance convective activity in the area, and the release of latent heat and the convective cells induce changes in the geopotential height that propagate towards extratropical latitudes. During the three Q1 2015 months, SSTs were high in the western tropical Pacific, and induced very high surface pressure over North America that resulted in the wind stalling (Figure 62).





**Figure 62: SST and surface wind speed anomalies for the first months of 2015, expressed as the number of standard deviations away from the 1979–2014 mean for the same month, drawn from ERA-Interim reanalysis. Purple contours show precipitation anomalies, and green contours show sea level pressure anomalies. Contour intervals are 5 mm/day and 2 hPa, respectively, with zero contour omitted (extracted from Lledó et al. 2018).**

## 6.7. Case study 7 – France, Europe 2018

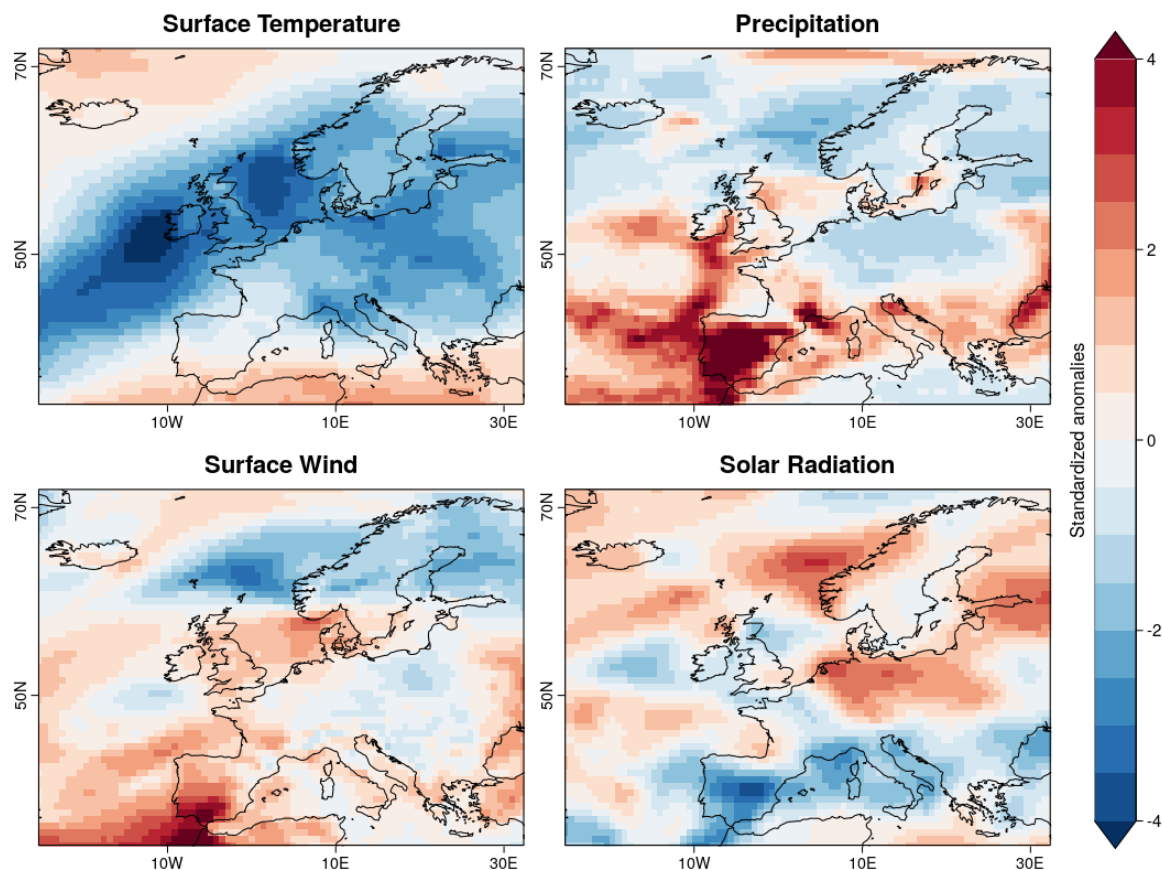
During 2017/18 winter, several weeks of extreme cold temperatures affected most of Europe. Due to the cold temperatures, energy consumption increased in the areas affected.

Cold wave over central Europe			
Region:	Europe/France	Period:	27 Feb–5 Mar, 2018
Forecast type:	Sub-seasonal	Main interest:	Energy demand
Forecast variables:	Temperature and demand		

**Table 14: Region, period, forecast type and main interest for case study 7.**

### Event description

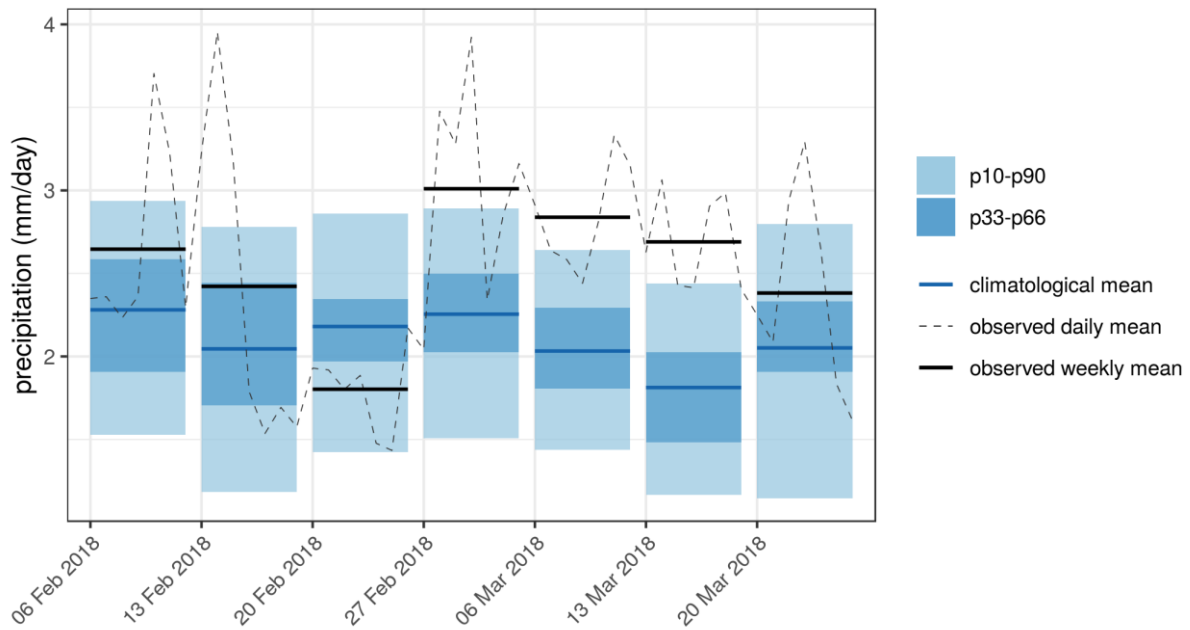
During the first months of 2018 (January–March), surface temperature was below normal over the entire European domain (Figure 63). Surface wind speeds were below normal in Scandinavia and central Europe, whilst western Europe and the Mediterranean region experienced above normal winds. The pattern for solar radiation was reversed in comparison to wind speed, with northern Europe being under above normal, and west Europe and the Mediterranean being below normal. Finally, precipitation was below normal in the central Europe and the Scandinavia, whilst quite above normal in west Europe and the Mediterranean.



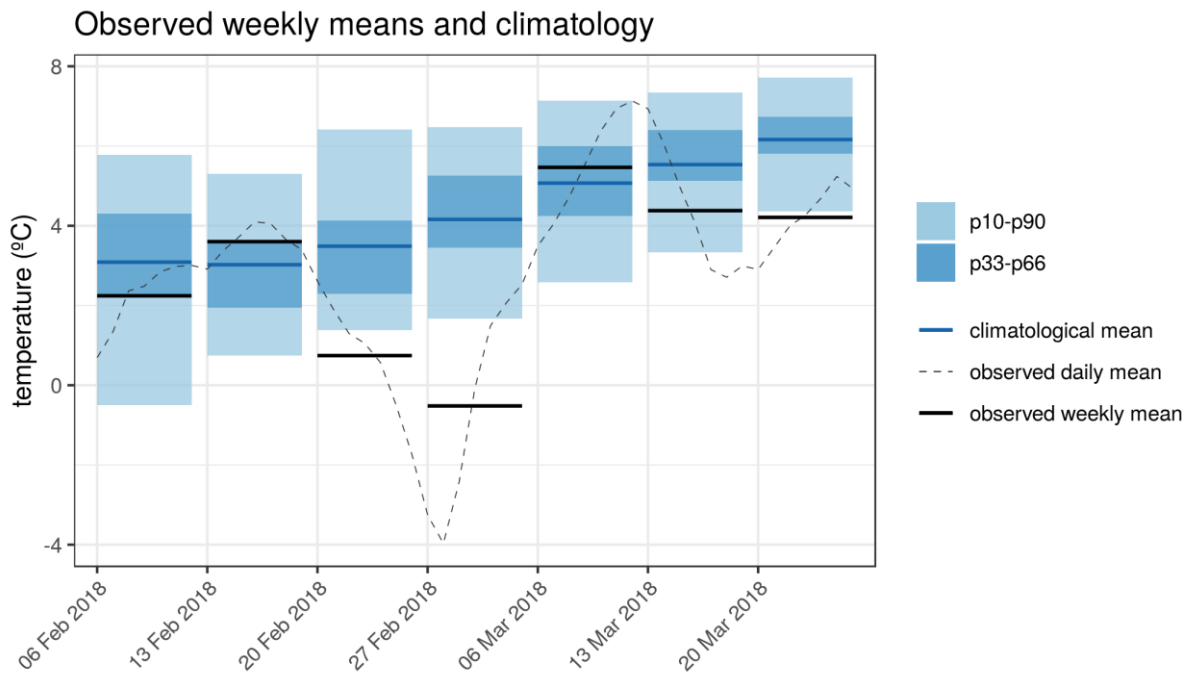
**Figure 63: Standardized anomalies of temperature, precipitation, wind and radiation for the first quarter of 2018, obtained from ERA-Interim.**

Analysis of the anomalies for the week of interest (27 Feb – 5 Mar) shows that precipitation was above normal with high daily and weekly temporal variability (Figure 64). The exceedance of normal conditions continues until the 20<sup>th</sup> March. Temperature evolution was different to the precipitation. Observations showed to be below normal since the 20<sup>th</sup> February and, in particular, for the period 27 Feb – 5 Mar temperature was significantly lower than normal conditions (Figure 65). The daily variation was quite strong with observed temperature differing from climatological mean by about 8°C. These strong anomalies can affect different stakeholders in the energy sector and in particular those that are prone to effects driven by the joint precipitation-temperature patterns.

### Observed weekly means and climatology



**Figure 64: Weekly evolution of the observed precipitation in the region 10°W-30°E and 36°N-65°N during February and March 2018 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis.**



**Figure 65: Weekly evolution of the observed temperature in the region 10°W-30°E and 36°N-65°N during February and March 2018 compared to the climatological distribution. Values obtained from ERA-Interim reanalysis.**

### Available forecasts

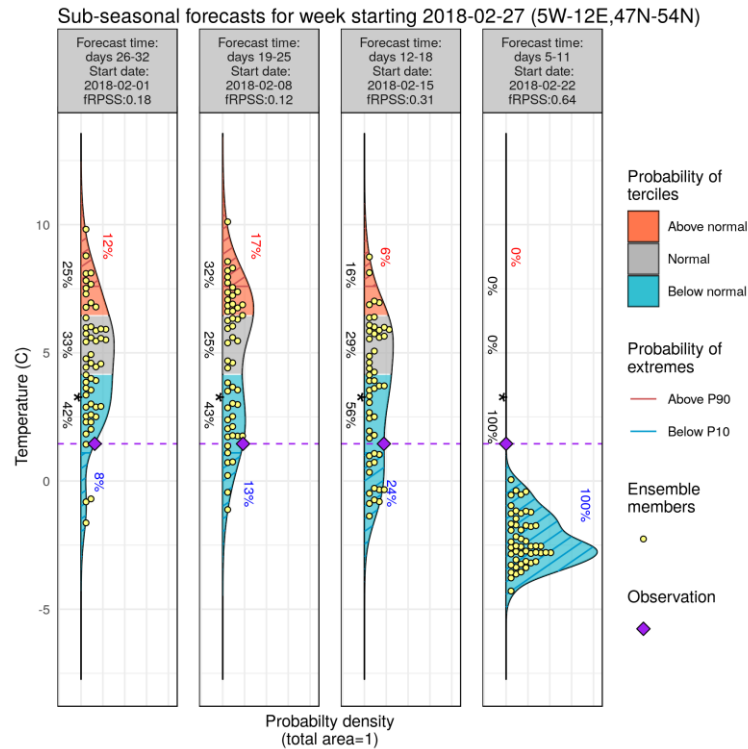
Sub-seasonal forecasts of temperature and electricity demand have been produced for Europe and France respectively using ECMWF SEAS5 (Table 10). Forecasts issued one, two, three and four weeks in advance are presented. The predictions have been bias-adjusted using the calibration variance inflation method with a hindcast covering 20 previous years to the forecast, and with ERA-Interim used as a reference dataset.

Start date	Valid period	Lead time	Variables	Forecast system
01 Feb 2018	27 Feb – 5 Mar 2018	26 to 32 days	- temperature - electricity demand	ENS-ER 1.5 degrees resolution 51 ensemble members
08 Feb 2018	27 Feb – 5 Mar 2018	19 to 25 days		
15 Feb 2018	27 Feb – 5 Mar 2018	12 to 18 days		
22 Feb 2018	27 Feb – 5 Mar 2018	5 to 11 days		

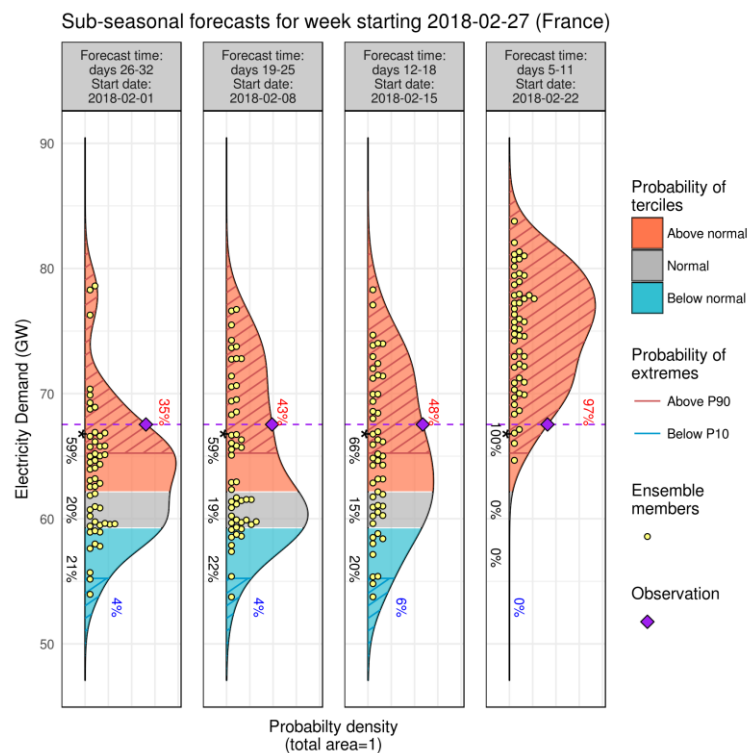
**Table 15: Start dates, lead times, valid period and variables for each of the forecasts presented.**

For each forecast, along with the figures, skill scores have been estimated using the hindcasts. Figure 66 and Figure 67 show the temperature and electricity demand forecasts for the four lead times. The temperature forecasts only show potential of capturing the observation two weeks prior to the event; however results for forecast week 1 were accurate in terms of being below normal conditions but underestimated the temperature value. In terms of electricity demand, and similarly with temperature, the forecasts could not predict the event, only until two weeks prior to it. The signal (described here as probability) was strong that the electricity demand would be above normal conditions.

Looking at skill scores (Annex 3), RPSS and CRPSS (measuring the terciles forecast quality and the forecast distribution quality respectively) are all positive (hence better than climatology) but show acceptable values from days 12-18 prior to the event. Decision making could therefore be made from that time onwards. Brier scores for the extreme conditions (P10 and P90) are also all positive, but only the forecast from days 5-11 prior to the event indicate high values.



**Figure 66: Temperature forecasts for 27 Feb 2018 issued four, three, two and one weeks in advance for the domain (5W-12E, 47N-54N).**



**Figure 67: Electricity demand forecasts for 27 Feb 2018 issued four, three, two and one weeks in advance for France.**



## 6.8. Case study 8 – Spain 2018

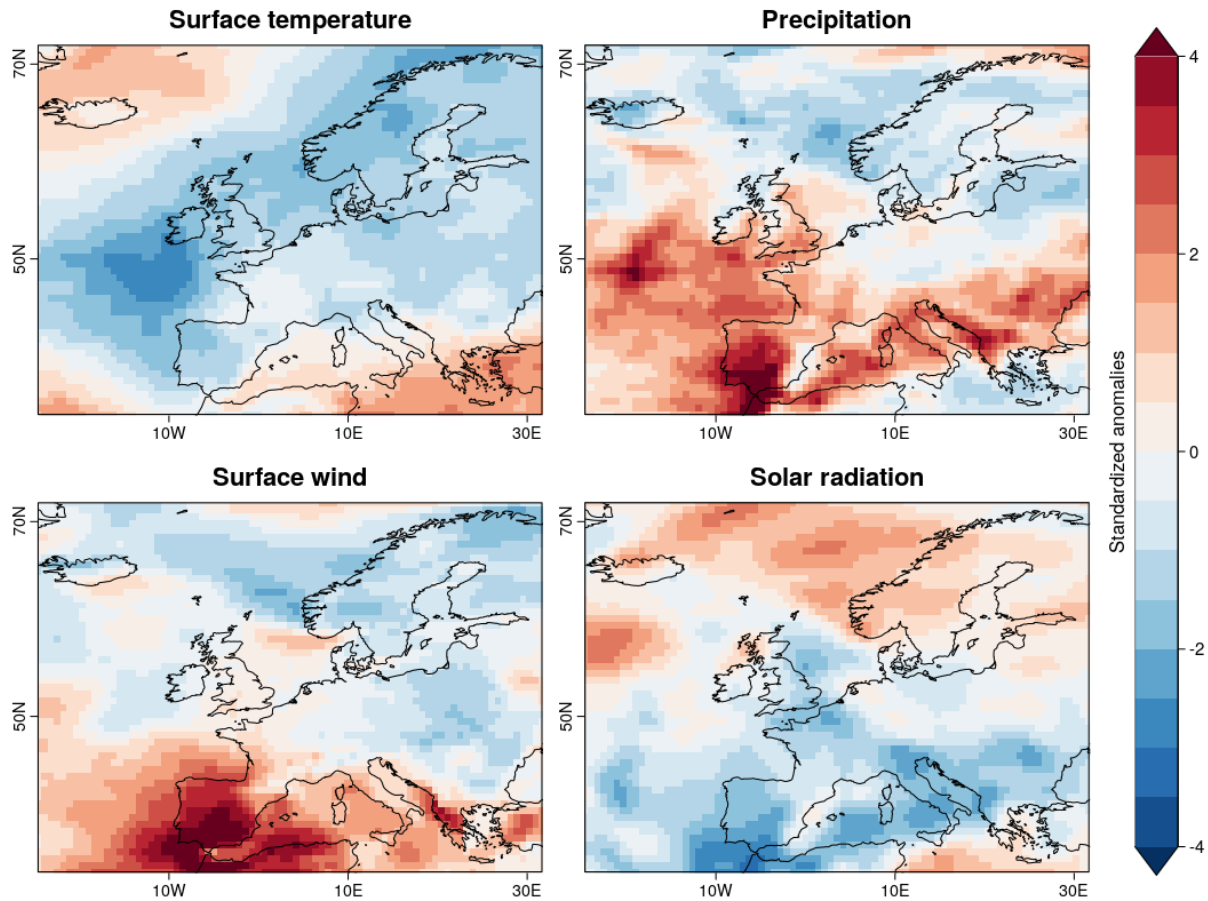
About 44.6% of Spain's power generation came from renewable energy sources in the first quarter of 2018. Wind farms were the top source of electricity power in the quarter, with a share of 26.5%, followed by nuclear power plants (NPPs) with a contribution of 21.6%. In March alone, the share of renewables rose to 56.6%. Wind farms produced 6,937 GWh in the first 28 days of the month, up by 62.7% year-on-year. Thus, the share of wind in March reached 32.9%.

In March 2018, the Spanish wind energy system broke his monthly records			
Region:	Spain	Period:	March 2018
Forecast type:	Seasonal	Main interest:	Wind energy, hydro and solar. Supply vs demand
Forecast variables:	Wind speed, wind capacity factor, Spanish wind power, precipitation, temperature, solar irradiation and solar capacity factor.		

**Table 16: Region, period, forecast type and main interest for case study 8.**

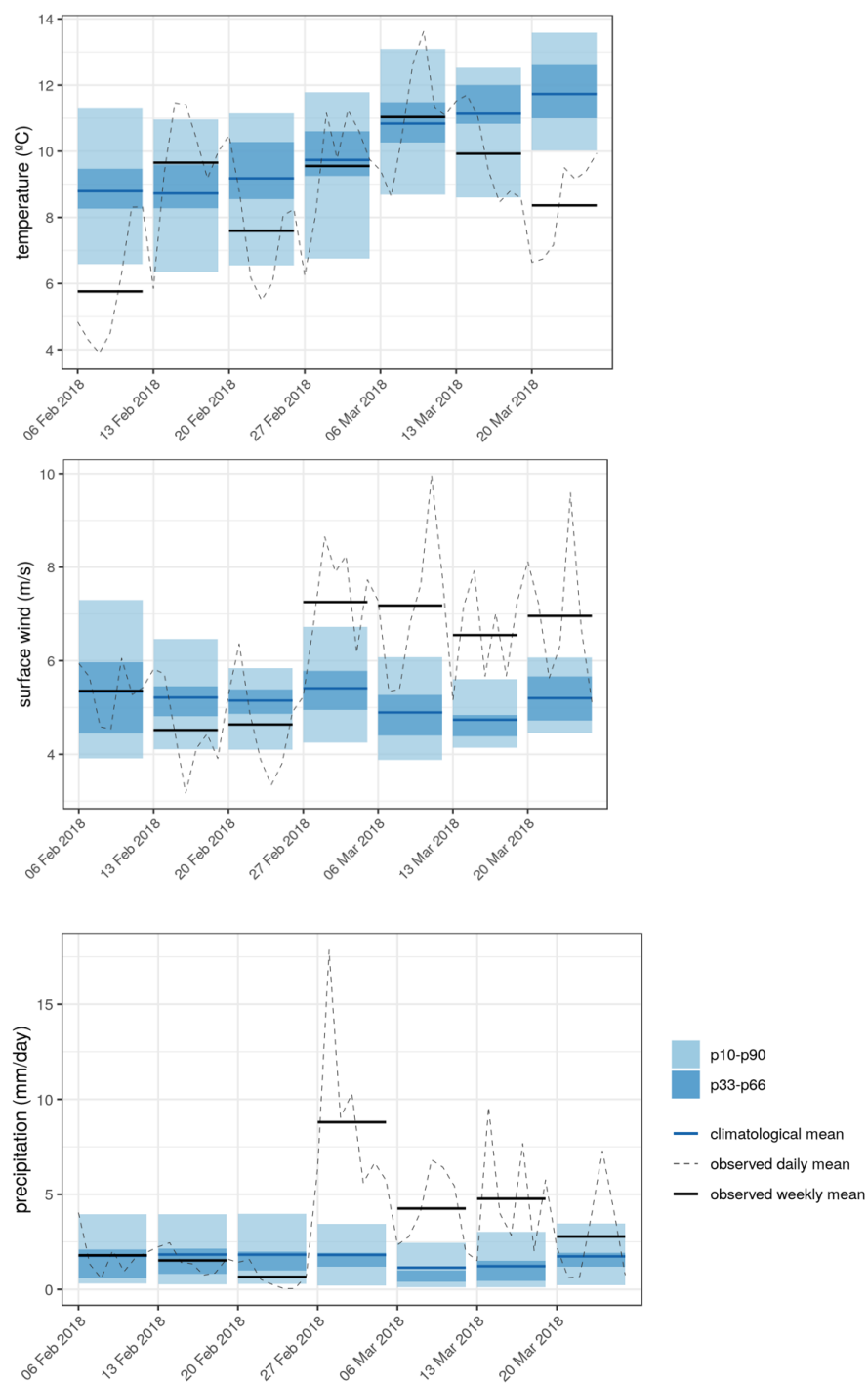
### Event description

March 2018 was characterized by a large area of low pressure in the eastern North Atlantic and western Mediterranean, producing a strong storm activity in western and southern Europe during this period. Consistent with this, most of Europe experienced temperatures slightly below normal (though within ~1 standard deviation of climatology), with strong winds and precipitation (particularly in the west and south (up to 3-4 standard deviations above the climatological norm in Iberia) and reduced insolation (Figure 68).



**Figure 68: Observed and surface meteorological conditions compared to climatology during March 2018.**

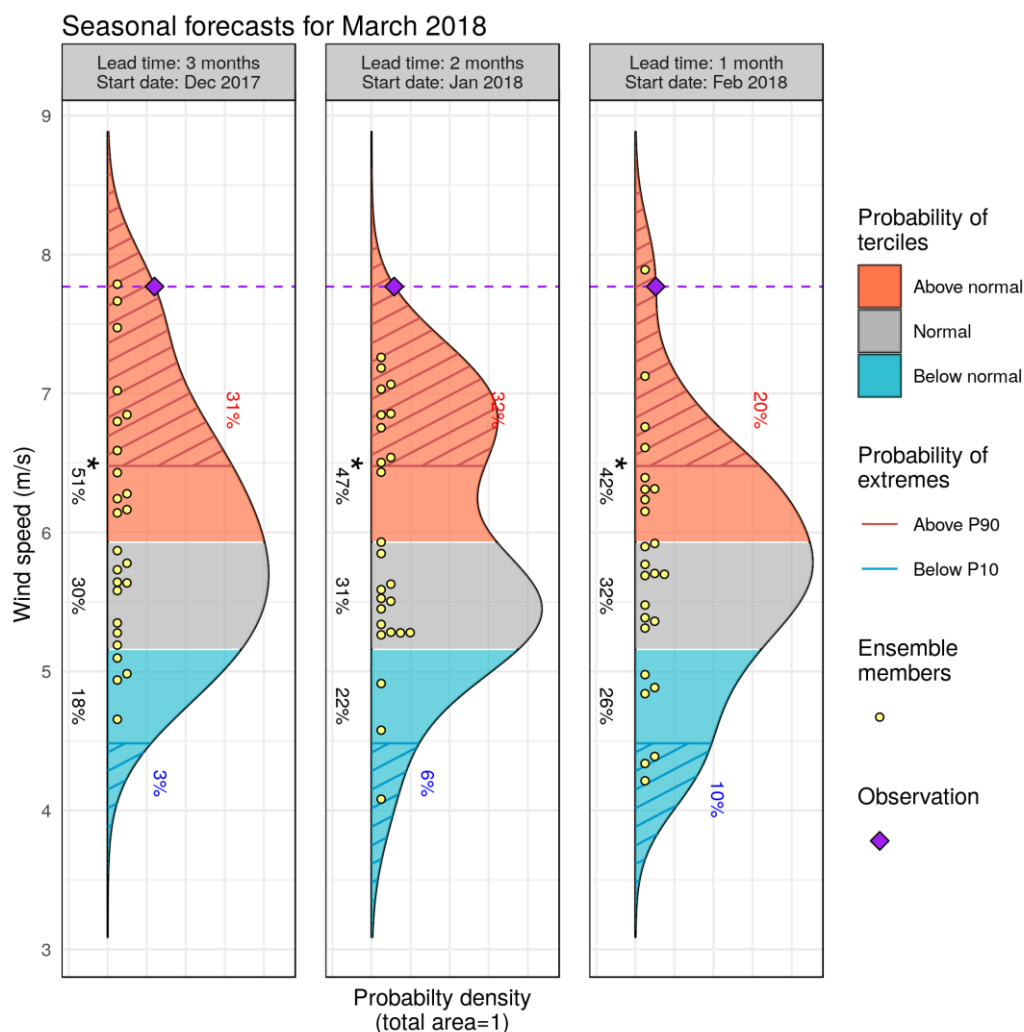
Figure 69 focusses on north-western Iberia: a region with particularly strong weather anomalies in March 2018. This wider view of the season's evolution suggests that the strong winds and precipitation began in the first week of March (the precipitation anomalies were particularly extreme during this week) and continued to exceed the 90th percentile of climatological expectation continuously for three weeks (precipitation) and at least four weeks (wind). In both precipitation and wind, however, there remained significant day-to-day variations, consistent with the episodic nature of synoptic systems affecting the region.



**Figure 69: Observed and climatological surface air temperature (top left), 10m wind speed (top right) and precipitation (bottom) averaged over 8.5°W-3E, 36.5-43.5°N during February and March 2018.**

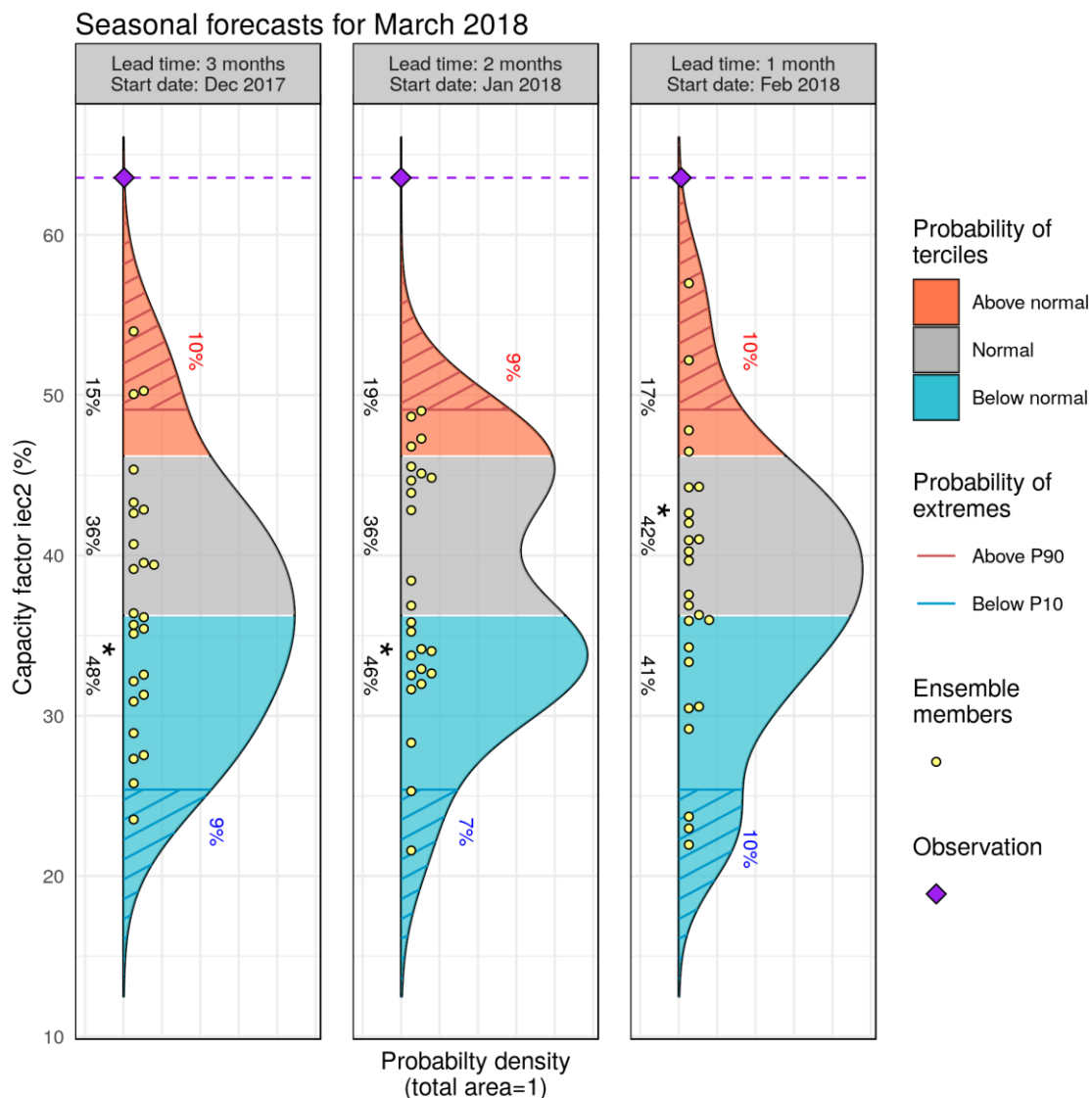
## Available forecasts

Figure 70, Figure 71 and Figure 72 show a sample of sub-seasonal wind and wind-power forecasts that could have been available during this event. For wind-speed, Figure 70 suggests that forecast indicated a weak shift toward above normal values across all lead times (~40-50% probability of the above normal tercile). It is, however, important to note that across these lead times, a formal skill assessment suggests there is little (or even negative) skill in the seasonal forecast for this property on average. Moreover, the forecast's tercile probabilities do not appear to converge on a shift to above-normal winds as lead time reduces (the January launch date suggests a greater probability of above-normal winds than the February launch date). Individual ensemble members typically also suggest much weaker wind speed anomalies than observed, though it is encouraging that one two occasions an ensemble member is found with comparable (or stronger) wind speeds than observed.



**Figure 70: Seasonal forecast for 10m wind speed averaged over 5°W–12°E, 47–54°N for March 2018. From left to right corresponds to forecasts launched from lead times month 3 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast.**

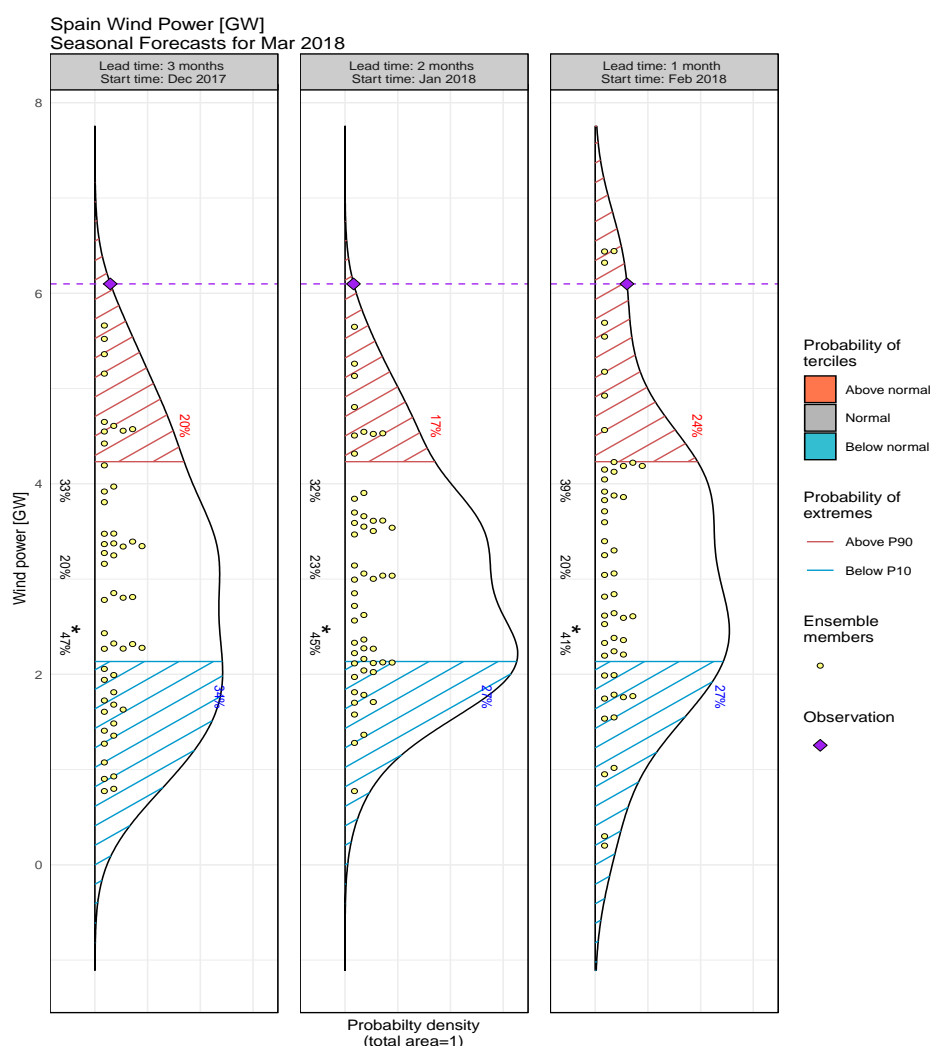
Figure 71 provides a different perspective on the expected wind power by using a standard power-curve to convert wind speeds into estimated capacity factors. Again, a formal skill assessment suggests there is little skill on average for forecasting this quantity; however interestingly, this specific forecast demonstrates a shift toward lower-than-normal capacity factors (40-50% probability of the below-normal tercile). This stands in direct contrast to the shift seen in wind speed and, in this case, no individual member captures the extremely high capacity factors that were observed.



**Figure 71: Seasonal forecast for capacity factors derived from the IEC2 class turbine (Gamesa G87 2MW), averaged over 5 oW - 12 oE, 47-54 oN for March 2018. From left to right corresponds to forecasts launched from lead times month 3 to 1.**

**Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast. An assessment of the skill associated with the forecast is indicated in each header (fRPSS = fair RPSS).**

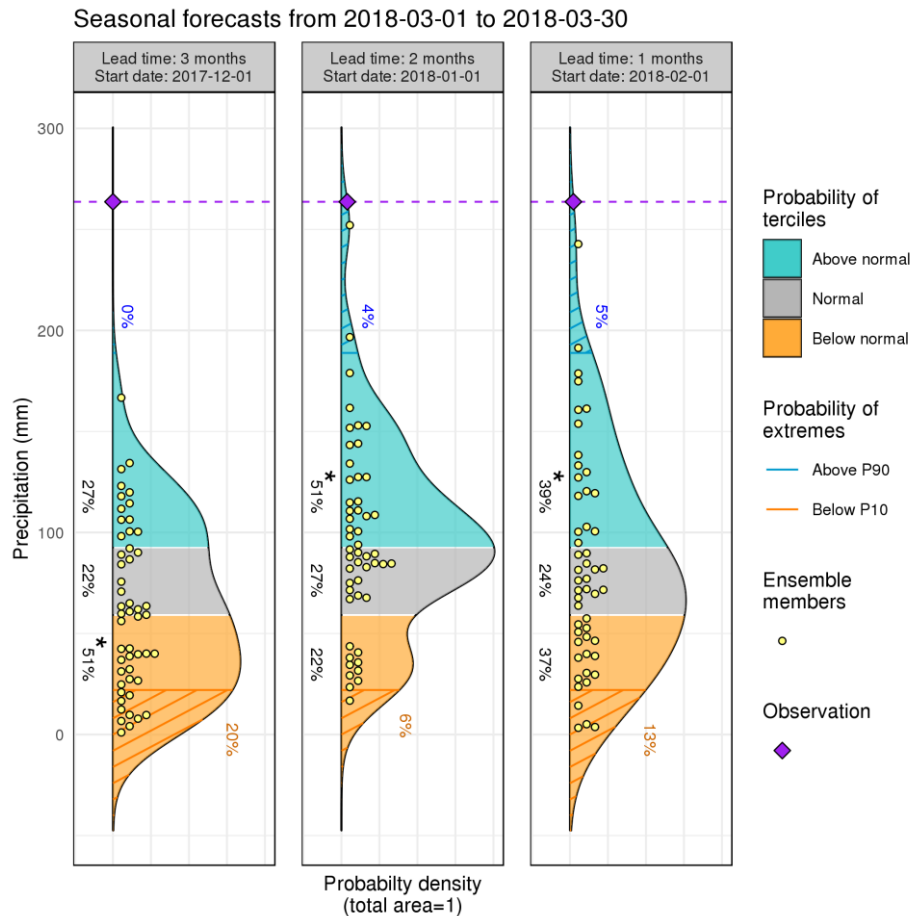
Although this forecast is undoubtedly unsuccessful – either for wind or capacity factor – the qualitatively apparent differences between the output of the two forecasts emphasizes the importance of adequately modelling the transformation of wind into power estimates for useful, skillful predictions. Indeed, taking the analysis a stage further – to simulate the total wind power generation of Spain (Figure 72) – reveals further differences insofar as the PDF is weakly shifted to lower values (similar to Figure 71 and generally in opposition to the observed event), but several individual ensemble members do capture the extremely high wind power output observed (at lead month 1), and the probability of high wind power generation remains high (32-39%) at all lead times.



**Figure 72: Sub-seasonal forecasts for nationally aggregated wind power for Spain for March 2018. From left to right corresponds to forecasts launched from lead times week 4 to 1. Methodology: lead time dependent mean bias correction, applied to both 10m wind speed and national wind power generation (once calculated from the wind), calibrated to ERA5, based on a 17-year hindcast. Note that the ERA5 climatology used to calibrate the forecast leads to unrealistically low wind power generation estimates (in GW).**

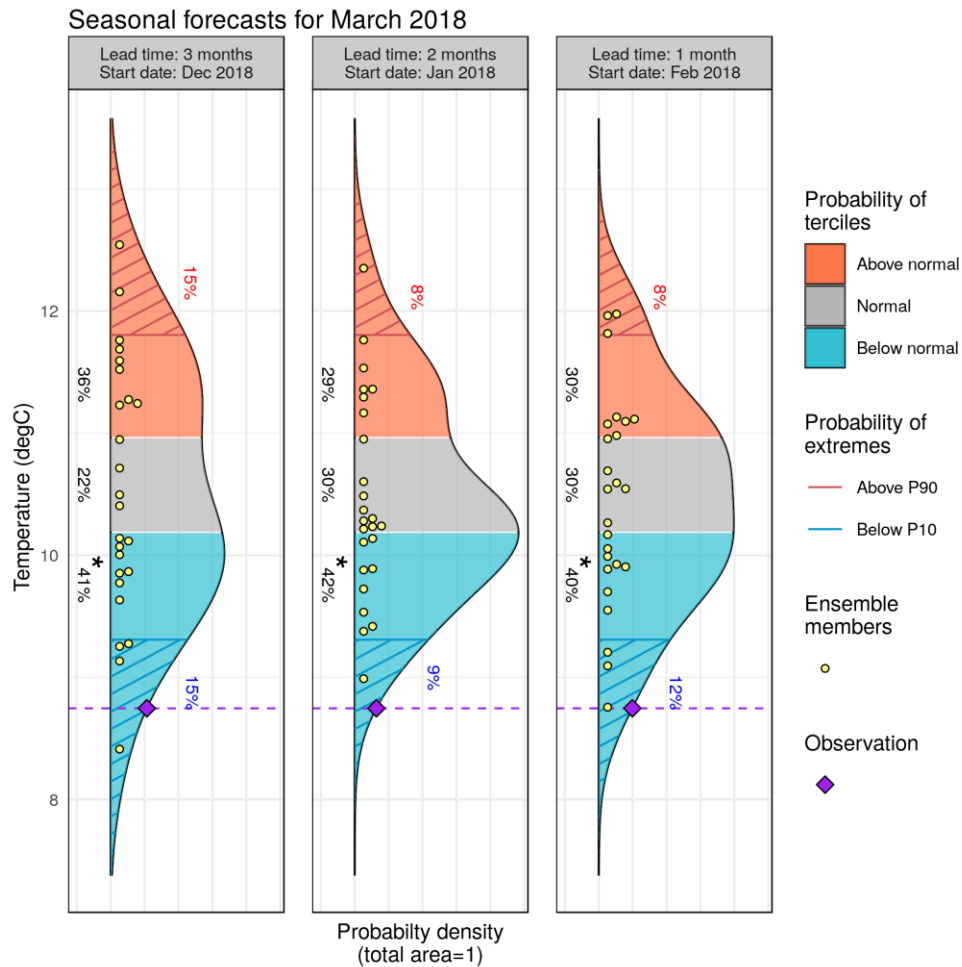


For precipitation for Tajo and Mino regions, Figure 73 suggests that forecast indicated a weak shift toward above normal values across the last two lead times (one and two months in advance). Individual ensemble members also suggest much weaker precipitation anomalies than observed.



**Figure 73: Precipitation forecasts for March 2018 issued three, two and one months in advance for Tajo and Mino regions in Spain.**

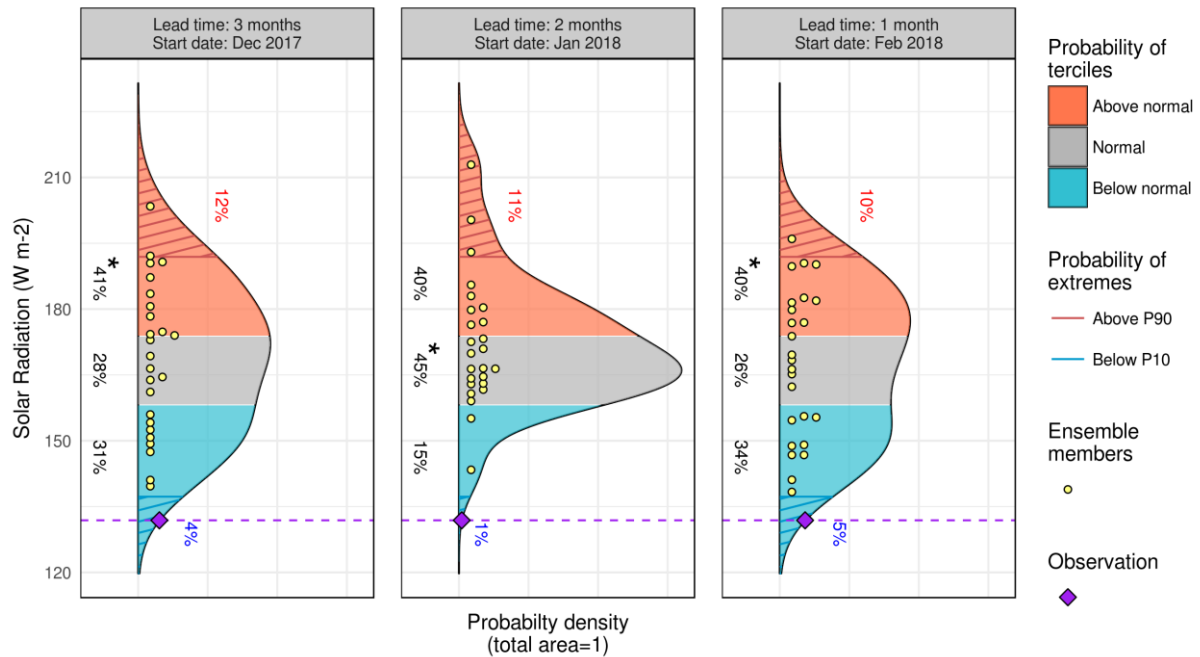
For temperature, Figure 74 shows that the forecast indicates a preference to below normal temperature, with 41, 42 and 40% probability in the below normal tercile.



**Figure 74: Seasonal forecast for temperature averaged over 5°W-12°E, 47-54°N for March 2018. From left to right corresponds to forecasts launched from lead times month 3 to 1. Methodology: variance inflation calibration to ERA-Interim, based on a 20-year hindcast.**

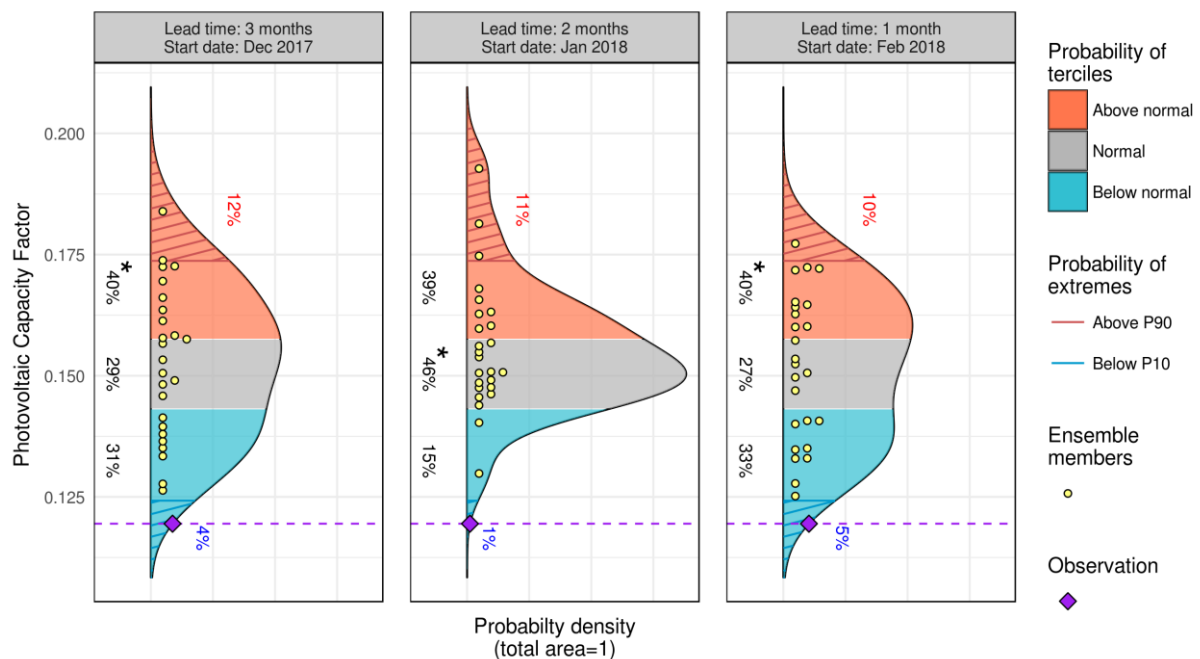
Surface solar radiation and photovoltaic capacity factor forecasts lacked skill with generally a very small probability of capturing the event. Forecasts showed an almost equal probability of the event being above or within the normal conditions. However, the observed values were below normal conditions, and the forecasts were not capable of capturing this at none of the three lead times. Both probabilistic and deterministic skill metrics are not significant with negative or close to zero values for all the lead times. Therefore, for surface solar radiation and photovoltaic capacity factor the skill of the predictions for this case study is not enough to provide added value.

Seasonal forecasts for March 2018 (9W-7W,42N-44N)



**Figure 75: Surface solar radiation forecasts for July 2013 issued three, two and one months in advance for domain (6E-15E, 47N-54N). Methodology: mean inflation calibration to ERA-Interim, based on 38 year hindcast.**

Seasonal forecasts for March 2018 (9W-7W,42N-44N)



**Figure 76: Photovoltaic capacity factor forecasts for July 2013 issued three, two and one months in advance for domain (6E-15E, 47N-54N). Methodology: mean inflation calibration to ERA-Interim, based on 38 year hindcast.**

## 7. Conclusions

The analysis conducted in this report aims to benchmark forecasting skill for a number of climate models, ECVs and indicators and dig deeper in terms of understanding at specific user identified case studies. Here benchmarking is based on a single model approach in which simple methods are used to downscale the forecasted information at the local scale. The analysis is conducted at two spatial scales (the European-wide scale and a local scale) using two ECMWF systems (the ENS-ER and SEAS5 for sub-seasonal and seasonal evaluation respectively).

Here we highlight that both ENS-ER and SEAS5 forecasts have biases that need to be adjusted prior to their evaluation and calculation of the impact indicators. These biases are not similar in terms of magnitude and spatial variability, and hence a different approach had to be applied for the two different systems (e.g. precipitation adjustment over a 3-week moving window and each month for sub-seasonal and seasonal systems respectively). However even when a bias-adjustment methodology is applied remaining biases still exist and their magnitude depends on the variable of interest. Here the objective was not to focus on advanced bias-adjustment methodologies and their inter-comparison, but rather on application of state-of-the-art approaches based on users' experience.

Despite the remaining biases in the meteorological forcing, all ECVs showed skill for the first lead time, however different variables show different level of skill, i.e. wind power forecasts are slightly less skillful than demand. Sub-seasonal forecasts for all ECVs show adequate skill in most European regions in forecast weeks 1 and 2; however the predictability varies both geographically and seasonally. In addition, some skill was observed for some regions in forecast week 3. The aggregation period of interest for seasonal forecasting was monthly averages in comparison to the weekly averages for the sub-seasonal forecasts, yet seasonal forecasts for all ECVs showed strong spatial and temporal variability. The skill in seasonal forecasts is only strongly evident in the 1<sup>st</sup> forecast month, and this decreases rapidly with increased lead time. Some patterns could have been identified in which forecasting skill in general shows more skill in winter than summer (June-August). In addition, the Scandinavia and occasionally central Europe seemed to retain their moderate skill for higher than 2 forecast months ahead.

The eight case studies analysed here gave additional information about the systems (sub-seasonal and seasonal) potential for these extreme events. Although skill is present considering an assessment over a large hindcast period, forecasts for the case studies occasionally failed to capture the magnitude of the event; however in most cases the event was predicted in terms of probability being above/below normal conditions. This investigation over the European-wide domain and specific event assessment sets a benchmark over which further methodologies and systems, beyond those used here, can be tested to assess potential forecasting skill improvements.

## 7.1. Moving forward

Overall, the present study motivates two areas of further research to improve the quality and utility of seasonal and sub-seasonal forecasts for energy users. Firstly, though meteorological skill is clearly available from these forecast systems (particularly for sub-seasonal week 2), it is appropriate to seek improvements in *meteorological* forecast processing to maximize its quality. Promising areas for investigation include a mixture of improved calibration techniques, longer averaging periods, longer hindcast data-sets, multi-model combination, or pattern-based forecasting techniques (Clark et al 2017, Thornton et al 2017, Brayshaw et al 2011) and “windows-of-opportunity” type approaches (Beerli et al, 2017, Cassou 2008, Baldwin and Dunkerton, 2001). Seamless forecasting could be another interesting area of scientific investigation in which seasonal forecasts are improved by the frequent initialization of the sub-seasonal service.

Secondly, the forecast skill *in terms of energy impacts* is dependent on the conversion process from weather to energy. Greater understanding of how the energy-conversion interacts with forecast skill – including steps to quantify uncertainty and improve quality – are clearly needed.

## Bibliography

- Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-de-Pison, F. J., Antonanzas-Torres, F. (2016). Review of photovoltaic power forecasting. *Solar Energy*, 136, 78-111, 10.1016/j.solener.2016.06.069.
- Arnal, L., Cloke, H. L., Stephens, E., Wetterhall, F., Prudhomme, C., Neumann, J., ... Pappenberger, F. (2018). Skilful seasonal forecasts of streamflow over Europe? *Hydrology and Earth System Sciences*, 22, 2057–2072. <https://doi.org/doi.org/10.5194/hess-22-2057-2018>
- Arnal, L., Wood, A. W., Stephens, E., Cloke, H. L., & Pappenberger, F. (2017). An Efficient Approach for Estimating Streamflow Forecast Skill Elasticity. *Journal of Hydrometeorology*, JHM-D-16-0259.1. <https://doi.org/10.1175/JHM-D-16-0259.1>
- Baldwin M.P. and Dunkerton, T.J. (2001). Stratospheric Harbingers of Anomalous Weather Regimes. *Science*, 294, 581-584.
- Bartholomé, E., Belward, A. S., Achard, F., Bartalev, S., Carmona Moreno, C., Eva, H., ... Stibig, H.-J. (2002). *GLC 2000 Global Land Cover mapping for the year 2000*. European Commission, DG Joint Research Centre, EUR 20524 EN, Ispra.
- Bazile, R., Boucher, M. A., Perreault, L., & Leconte, R. (2017). Verification of ECMWF System 4 for seasonal hydrological forecasting in a northern climate. *Hydrology and Earth System Sciences*, 21(11), 5747–5762. <https://doi.org/10.5194/hess-21-5747-2017>
- Beerli, R., Wernli, H., Grams, C.M. (2017). Does the lower stratosphere provide predictability for month-ahead wind electricity generation in Europe? *Quarterly Journal of the Royal Meteorological Society*, 143, 3025–3036.
- Bennett, J. C., Wang, Q. J., Robertson, D. E., Schepen, A., Li, M., & Michael, K. (2017). Assessment of an ensemble seasonal streamflow forecasting system for Australia. *Hydrology and Earth System Sciences*, 21(12), 6007–6030. <https://doi.org/10.5194/hess-21-6007-2017>
- Berg, P., Donnelly, C., & Gustafsson, D. (2018). Near real-time adjusted reanalysis forcing data for hydrology. *Hydrology and Earth System Sciences*, 22, 989–1000. <https://doi.org/https://doi.org/10.5194/hess-22-989-2018>
- Bessec, M., Fouquau, J., (2008). The non-linear link between electricity consumption and temperature in Europe: a threshold panel approach. *Energy Econom.*, 30(5), 2705–21.
- Brayshaw, D. J., Troccoli, A., Fordham, R. and Methven, J. (2011). The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: a case study over the UK. *Renewable Energy*, 36, 2087-2096.
- Bruno Soares, M., Alexander, M., & Dessai, S. (2017). Sectoral use of climate information in Europe: A synoptic overview. *Climate Services*, 1–16. <https://doi.org/10.1016/j.cliser.2017.06.001>



- Bruno Soares, M., & Dessai, S. (2016). Barriers and enablers to the use of seasonal climate forecasts amongst organisations in Europe. *Climatic Change*, 137(1–2), 89–103. <https://doi.org/10.1007/s10584-016-1671-8>
- Cannon, D.J., Brayshaw, D.J., Methven, J., Coker, P.J. and Lenaghan, D. (2015). Using reanalysis data to quantify extreme wind power generation statistics : a 33 year case study in Great Britain. *Renewable Energy*, 75, 767–778.
- Cassou, C. (2008). Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Nature*, 455, 523–527.
- Clark R.T, Bett, P.E., Thornton H.E., Scaife A.A. (2017). Skilful seasonal predictions for the European energy industry, *Environmental Research Letters*, 12, 024002.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Kheer, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, <https://doi.org/10.1002/qj.828>
- De Felice, M., Alessandri, A., Catalano, F. (2015). Seasonal climate forecasts for medium-term electricity demand forecasting. *Appl. Energy*, 137, 435–444.
- De Felice, M., Petitta, M., Ruti, P.M. (2015). Short-term predictability of photovoltaic production over Italy. *Renew. Energy*, 80, 197–204. doi:10.1016/j.renene.2015.02.010.
- Doblas-Reyes F. J., Hagedorn, R., & Palmer, T. (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting – II . Calibration. *Tellus A: Dynamic Meteorology and Oceanography*, 2005(57), 234–252. <https://doi.org/10.3402/tellusa.v57i3.14658>
- Dutra, E., Pozzi, W., Wetterhall, F., Di Giuseppe, F., Magnusson, L., Naumann, G., ... Pappenberger, F. (2014). Global meteorological drought-Part 2: Seasonal forecasts. *Hydrology and Earth System Sciences*, 18(7), 2669–2678. <https://doi.org/10.5194/hess-18-2669-2014>
- Ferro, C. A. T. (2014). Fair scores for ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, 140(683), 1917–1923. <https://doi.org/10.1002/qj.2270>
- Füss R., Mahringer S., Prokopczuk M. (2015). Electricity derivatives pricing with forward-looking information, *Journal of Economic Dynamics and Control*, Volume 58, Pages 34–57, ISSN 0165-1889, <https://doi.org/10.1016/j.jedc.2015.05.016>
- Green, R. (2005). Electricity and Markets. *Oxford Review of Economic Policy*, 21, 67–87.
- Greuell, W., Franssen, W. H. P., Biemans, H., & Hutjes, R. W. A. (2018). Seasonal streamflow forecasts for Europe – Part I: Hindcast verification with pseudo- and real observations. *Hydrology and Earth System Sciences*, 22, 3453–3472. <https://doi.org/10.5194/hess-22-3453-2018>

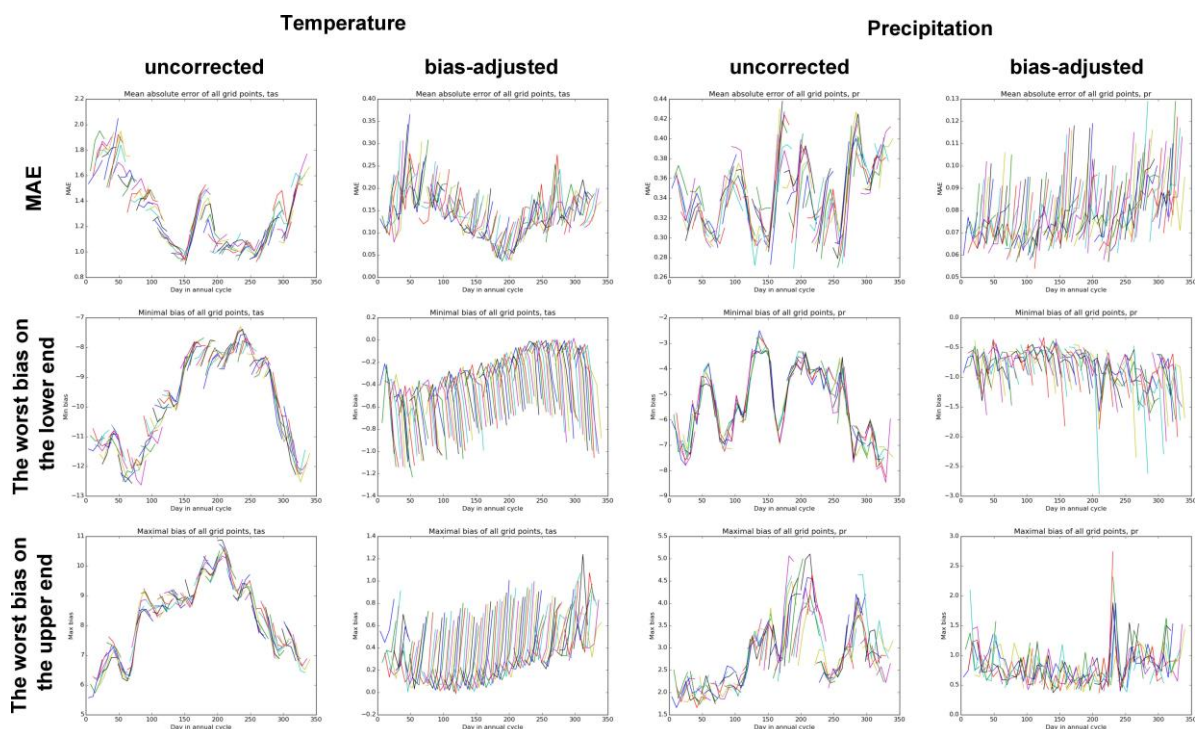
- Hundecha, Y., Arheimer, B., Donnelly, C., & Pechlivanidis, I. (2016). A regional parameter estimation scheme for a pan-European multi-basin model. *Journal of Hydrology: Regional Studies*, 6. <https://doi.org/10.1016/j.ejrh.2016.04.002>
- Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 296(1–4), 1–22. <https://doi.org/10.1016/j.jhydrol.2004.03.028>
- Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., ... Wissler, D. (2011). *Global Reservoir and Dam (GRanD) database - Technical documentation version 1.1*. Bonn.
- Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions AGU*, 89(10), 93–94. <https://doi.org/10.1029/2008EO100001>
- Lynch, K. J. (2016). Subseasonal forecasting for the energy sector. PhD Thesis, University of Reading.
- Manzanas, R., J. M. Gutiérrez, J. Bhend, S. Hemri, F. J. Doblas-Reyes, V. Torralba, E. Penabad, A. Brookshaw (2018). Bias adjustment and ensemble recalibration methods for seasonal forecasting: A comprehensive intercomparison using the C3S dataset, manuscript submitted for publication in Climate Dynamics.
- Meißner, D., Klein, B., & Ionita, M. (2017). Development of a monthly to seasonal forecast framework tailored to inland waterway transport in central Europe. *Hydrology and Earth System Sciences*, 21(12), 6401–6423. <https://doi.org/10.5194/hess-21-6401-2017>
- Mendoza, P. A., Wood, A. W., Clark, E., Rothwell, E., Clark, M. P., Nijssen, B., ... Arnold, J. R. (2017). An intercomparison of approaches for improving operational seasonal streamflow forecasts. *Hydrology and Earth System Sciences*, 21(7), 3915–3935. <https://doi.org/10.5194/hess-21-3915-2017>
- Nachtergaele, F., van Velthuisen, H., Verelst, L., & Wiberg, D. (2012). *Harmonized world soil database version 1.2*. FAO, Rome and IIASA, Laxenburg, Austria.
- Nachtergaele, F., van Velthuisen, H., Verelst, L., & Wiberg, D. (2012). *Harmonized world soil database version 1.2*. FAO, Rome and IIASA, Laxenburg, Austria.
- Najafi, A., Falaghi H., Contreras J., Ramezani M. (2016), Medium-term energy hub management subject to electricity price and wind uncertainty, *Applied Energy*, Volume 168, Pages 418–433, ISSN 0306-2619, <https://doi.org/10.1016/j.apenergy.2016.01.074>
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models. *Journal of Hydrology*, 10, 282–290.
- Pardo, A., Meneu, V., Valor, E. (2002). Temperature and seasonality influences on spanish electricity load. *Energy Econom.*, 24(1), 55–70.
- Portmann, F. T., Siebert, S., & Döll, P. (2010). MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, 24(1), GB1011. <https://doi.org/10.1029/2008GB003435>

- Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., et al. (2014). Skillful long-range prediction of European and North American winters. *Geophysical Research Letters*, 41, 2514–2519.
- Sene, K., Tych, W., and Beven, K. (2018). Exploratory studies into sea-sonal flow forecasting potential for large lakes, *Hydrol. Earth Syst. Sci.*, 22, 127–141, <https://doi.org/10.5194/hess-22-127-2018>
- Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., & Portmann, F. T. (2010). Groundwater use for irrigation - A global inventory. *Hydrology and Earth System Sciences*, 14(10), 1863–1880. <https://doi.org/10.5194/hess-14-1863-2010>
- Stoft, S. (2002). *Power System Economics*. IEEE Press Wiley, Piscataway, NJ.
- Thornton, H. E., Scaife, A. A., Hoskins, B. J. and Brayshaw, D. J. (2017) The relationship between wind power, electricity demand and winter weather patterns in Great Britain. *Environmental Research Letters*, 12, 064017.
- Torrallba, V., Doblas-Reyes, F. J., MacLeod, D., Christel, I., Davis, M., Torralba, V., ... Davis, M. (2017). Seasonal Climate Prediction: A New Source of Information for the Management of Wind Energy Resources. *Journal of Applied Meteorology and Climatology*, 56(5), 1231–1247. <https://doi.org/10.1175/JAMC-D-16-0204.1>
- Valor, E, Meneu, V., Caselles, V. (2001) Daily air temperature and electricity load in Spain. *J Appl Meteorol*, 40(8), 1413–21.
- Wetterhall, F., & Di Giuseppe, F. (2018). The benefit of seamless forecasts for hydrological predictions over Europe. *Hydrology and Earth System Sciences*, 22(6), 3409–3420. <https://doi.org/10.5194/hess-22-3409-2018>
- White, C. J., Carlsen, H., Robertson, A. W., Klein, R. J. T., Lazo, J. K., Kumar, A., ... Zebiak, S. E. (2017). Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorological Applications*, 24(3), 315–325. <https://doi.org/10.1002/met.1654>

## Annex

### Annex 1 – Bias-adjustment of ENS-ER for hydro

Figure A. 1 shows the validation results (1999-2010) produced for the ECMWF ENS-ER forecasts. The analysis is based on the ensemble mean bias and derived the overall mean absolute bias and worst case grid point biases at the lower and upper end for the whole European domain. Results show that no forecast sticks out although there is some seasonal variation in the bias after bias-adjustment. Overall, the bias-adjustment has worked allowing further the hydrological assessment.



**Figure A. 1: Biases in terms of mean absolute error (MAE) of the raw (uncorrected) and post-processed (bias-adjusted) temperature and precipitation sub-seasonal forecasts. The hydroGFD dataset was used as reference, whilst the DBS method was used for bias-adjusting the forecasts.**

## **Annex 2 – European-wide assessment – Additional metrics**

A number of figures have been produced for all ECVs using both ENS-ER and SEAS5, and for all months, lead times and metrics. Below you find a subset of them for selected ECVs, but the reader is directed to two links for ENS-ER and SEAS5 respectively, where more information is available.

### **ENS-ER**

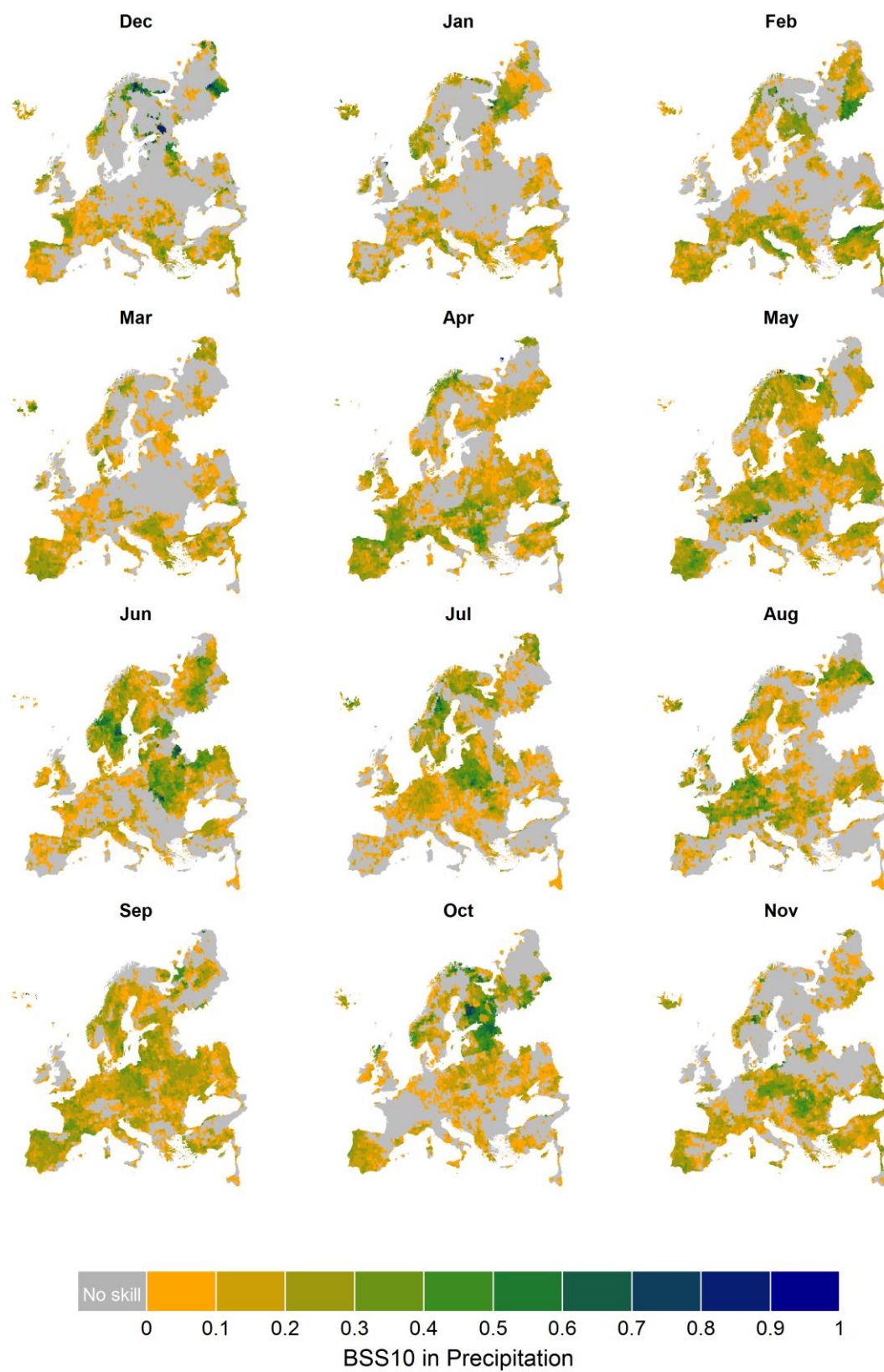
[https://drive.google.com/file/d/1vu59fDQRb9I\\_Pqwh6SI33aqlI8QR77BA/view?usp=sharing](https://drive.google.com/file/d/1vu59fDQRb9I_Pqwh6SI33aqlI8QR77BA/view?usp=sharing)

### **SEAS5**

<https://drive.google.com/file/d/1s9kfQfB6EPWRIJXtn-NmCntY8fBKGqLi/view?usp=sharing>

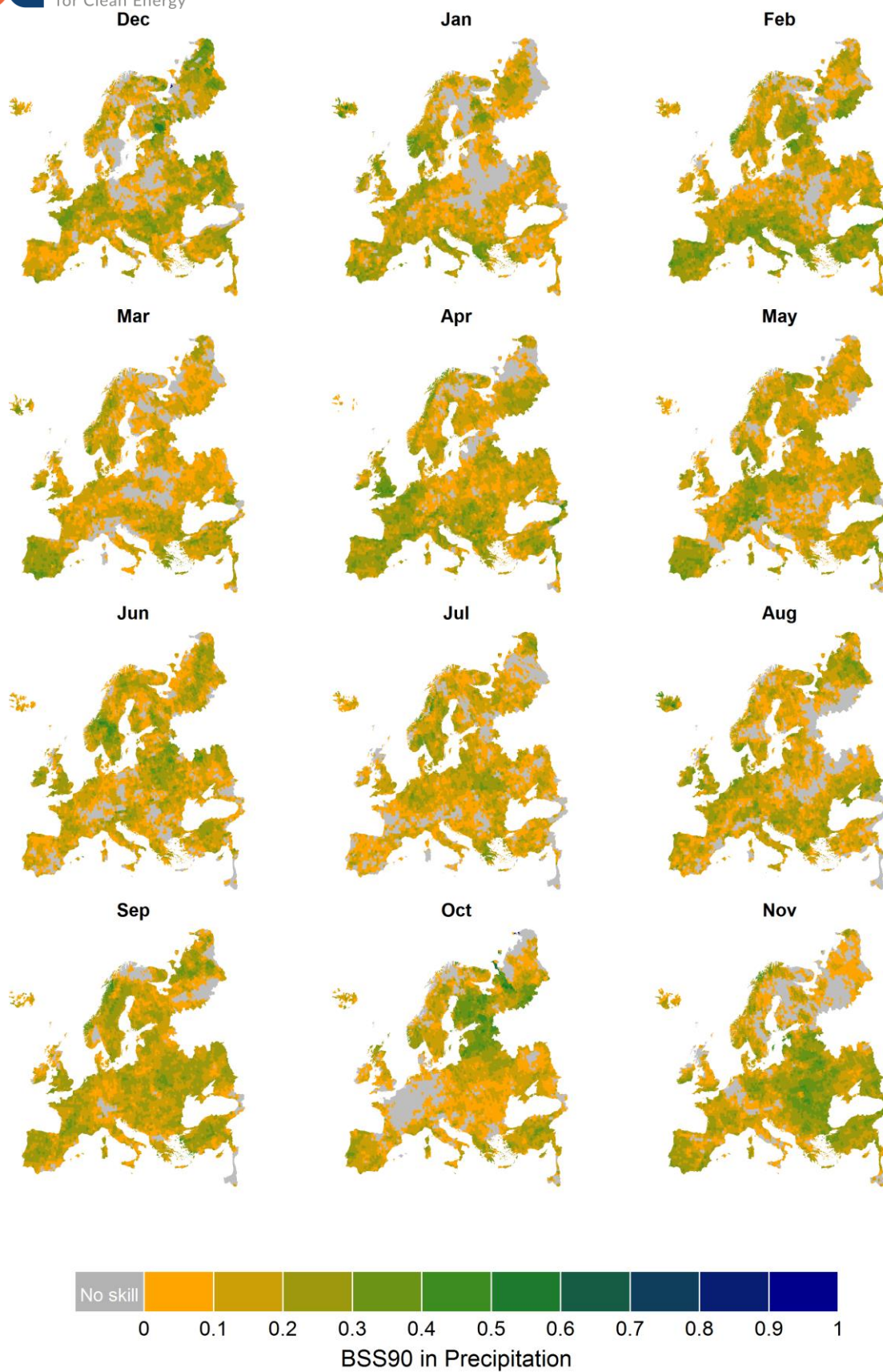


## Annex2.1. Seasonal time-scale

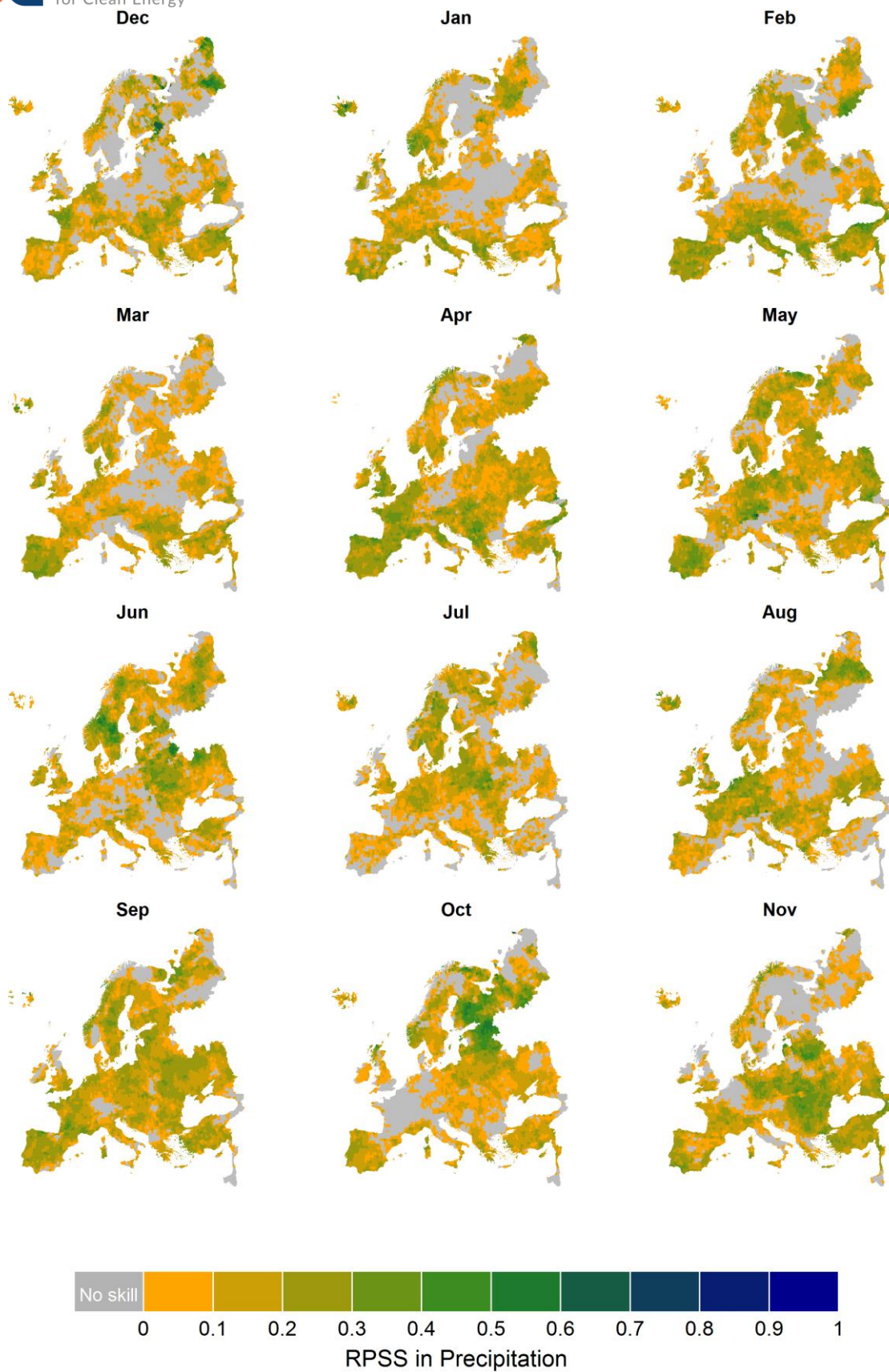


**Figure A. 2: Forecast month 1 monthly spatial variability of the Brier skill score (BSS) for 10<sup>th</sup> percentile.**





**Figure A. 3: Forecast month 1 monthly spatial variability of the Brier skill score (BSS) for 90<sup>th</sup> percentile.**



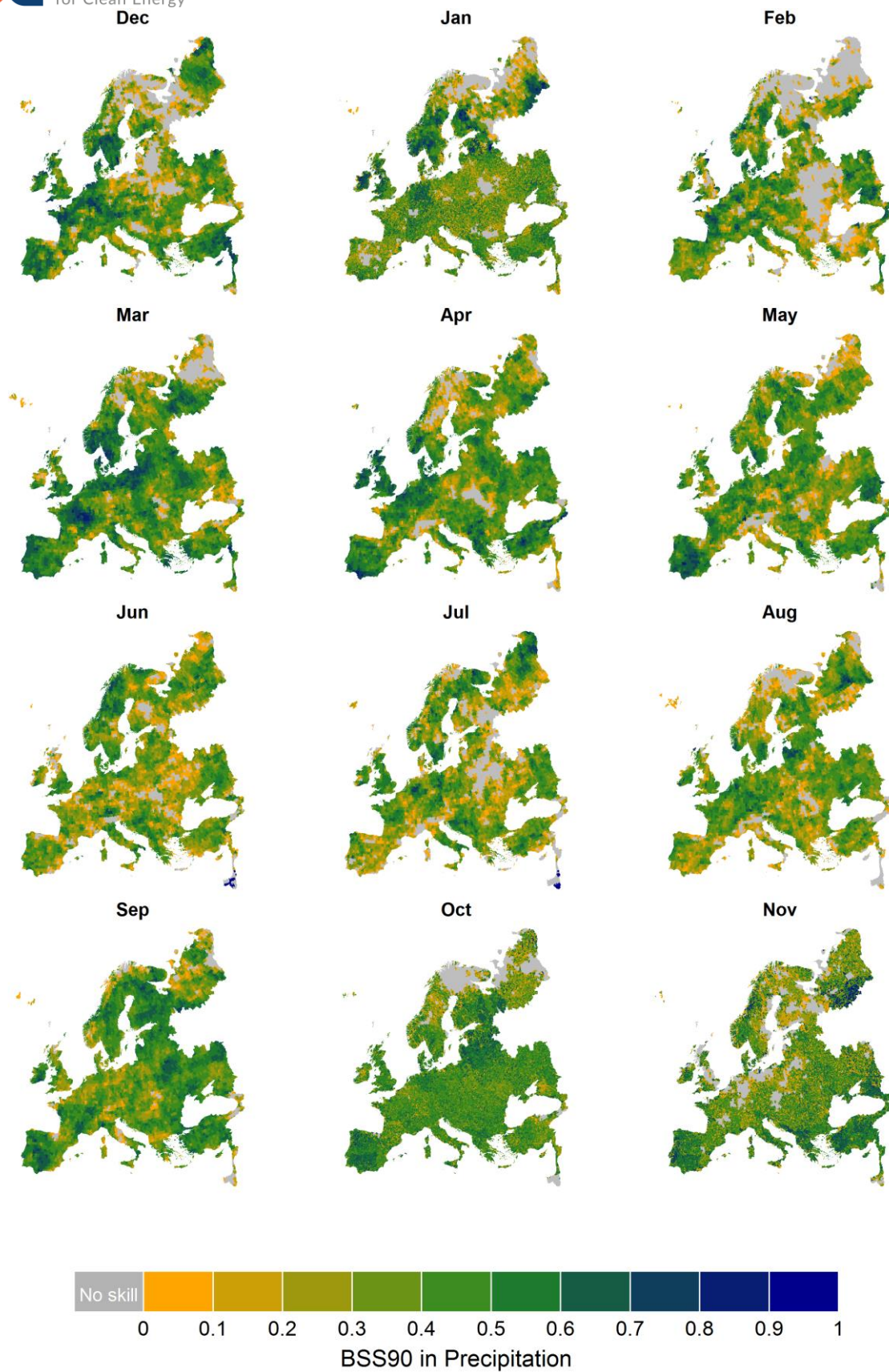
**Figure A. 4: Forecast month 1 monthly spatial variability of the RPS score (RPSS).**



## Annex2.2. Sub-seasonal time-scale



**Figure A. 5: Forecast week 1 monthly spatial variability of the Brier skill score (BSS) for 10<sup>th</sup> percentile.**



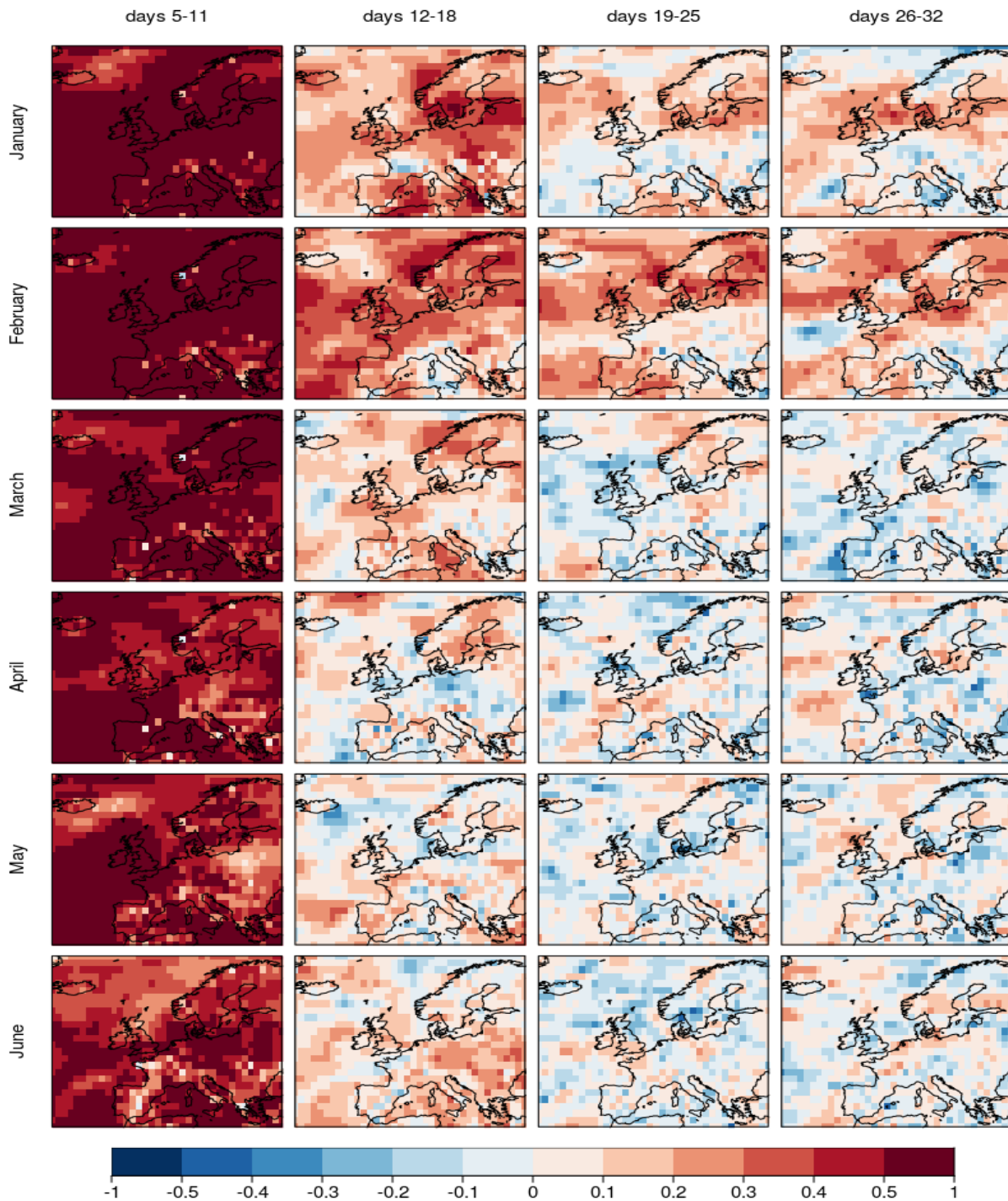
**Figure A. 6: Forecast week 1 monthly spatial variability of the Brier skill score (BSS) for 90<sup>th</sup> percentile.**





**Figure A. 7: Forecast week 1 monthly spatial variability of the RPS score (RPSS).**

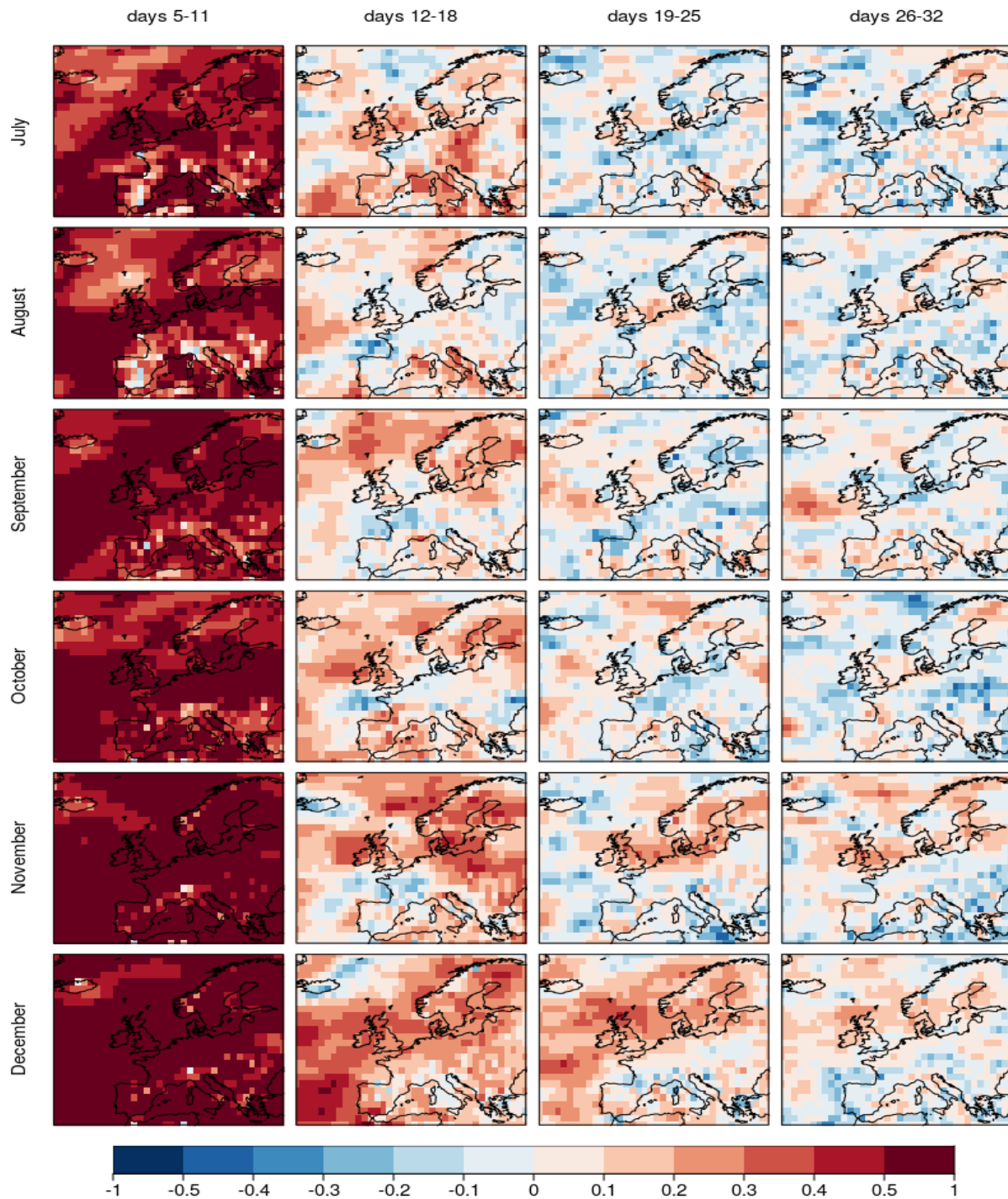
## sfcWind EnsCorr calibration



**Figure A. 8: Ensemble correction (EnsCorr) of 10m wind speed from ECMWF monthly prediction system for January to June and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts were calibrated with the variance inflation method.**

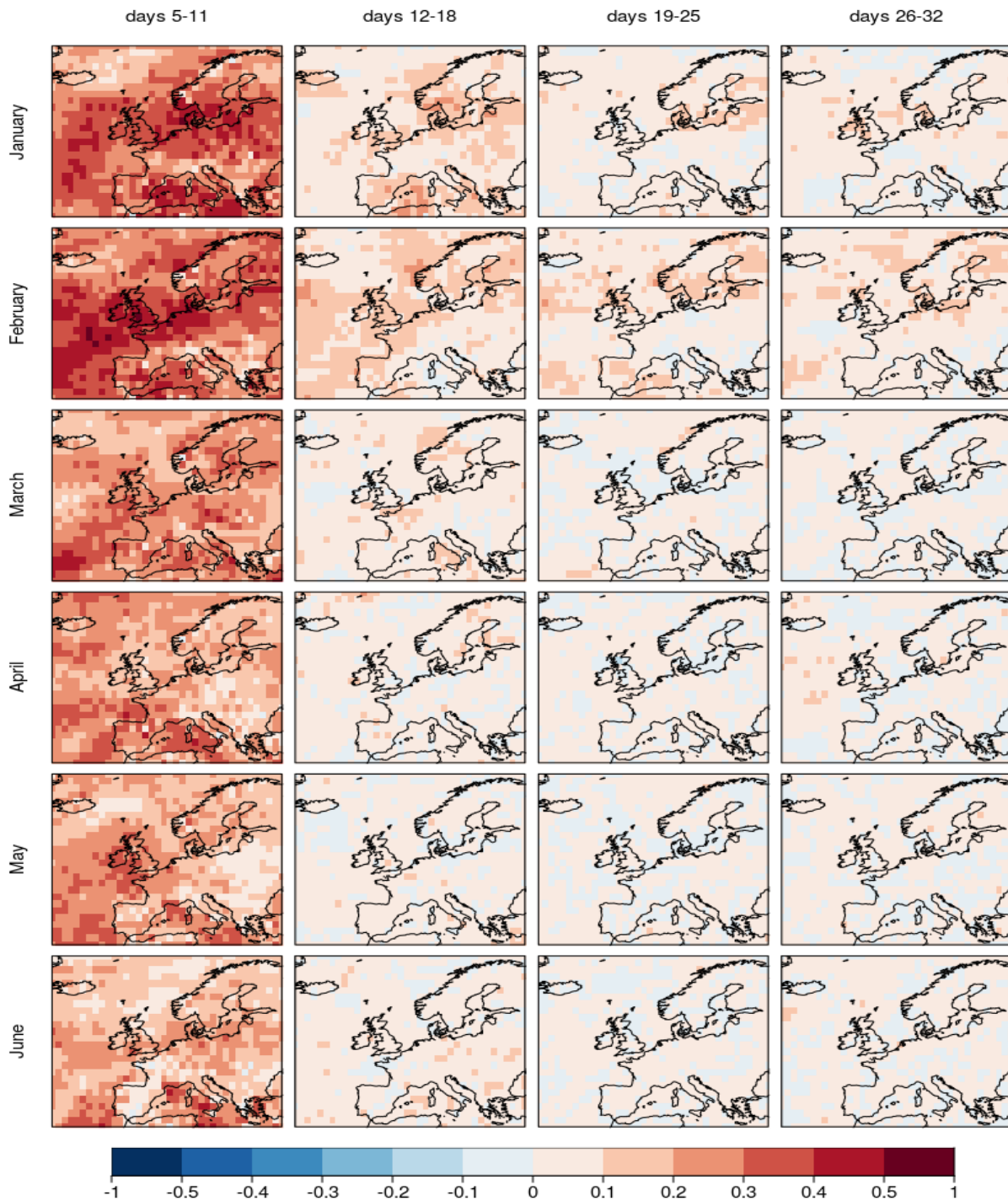


## sfcWind EnsCorr calibration



**Figure A. 9: Ensemble correction (EnsCorr) of 10m wind speed from ECMWF monthly prediction system for July to December and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts were calibrated with the variance inflation method.**

## sfcWind FairRpss calibration



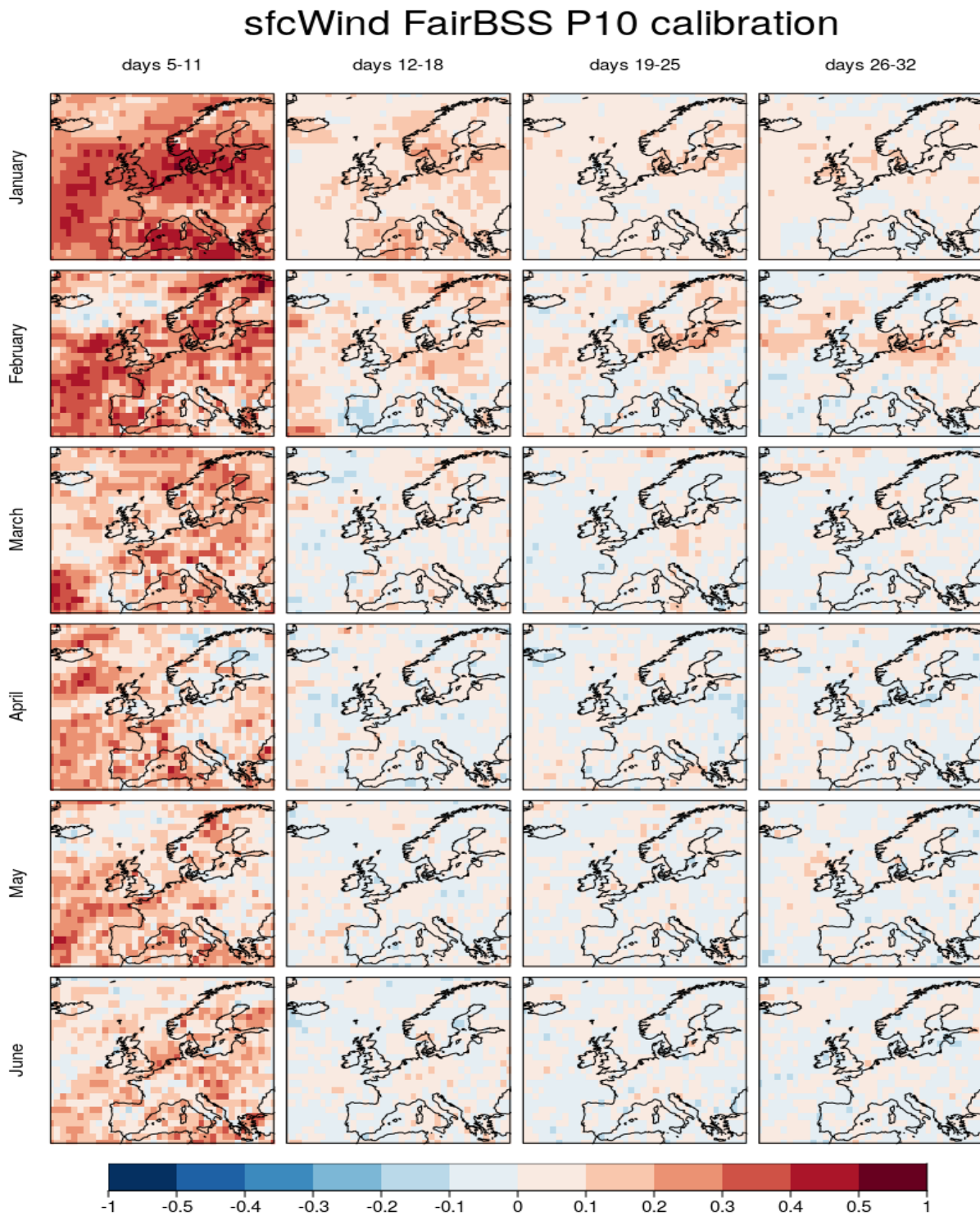
**Figure A. 10: FairRPSS of 10m wind speed from ECMWF monthly prediction system for January to June and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts were calibrated with the variance inflation method.**

## sfcWind FairRpss calibration

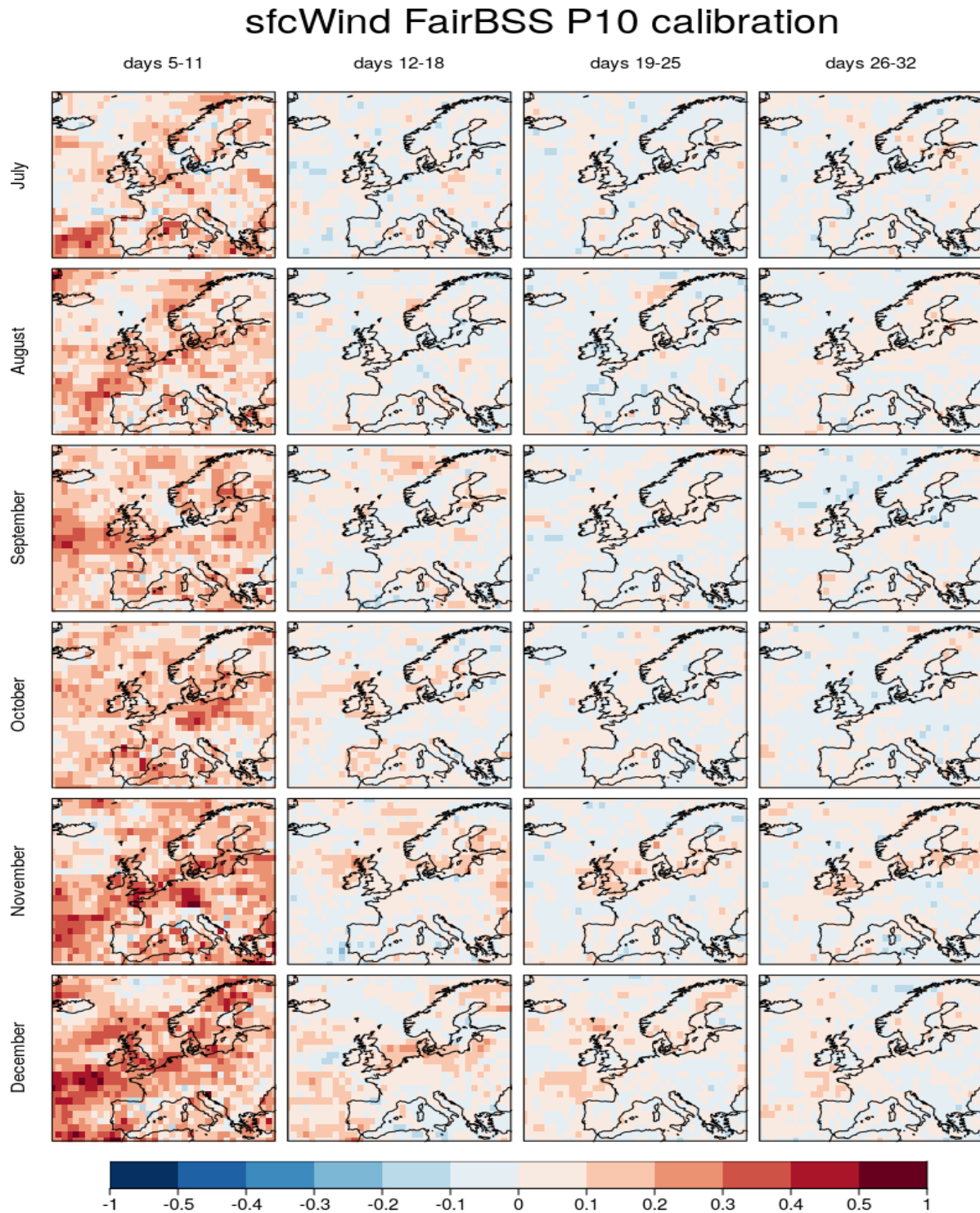


**Figure A. 11: FairRPSS of 10m wind speed from ECMWF monthly prediction system for July to December and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts were calibrated with the variance inflation method.**



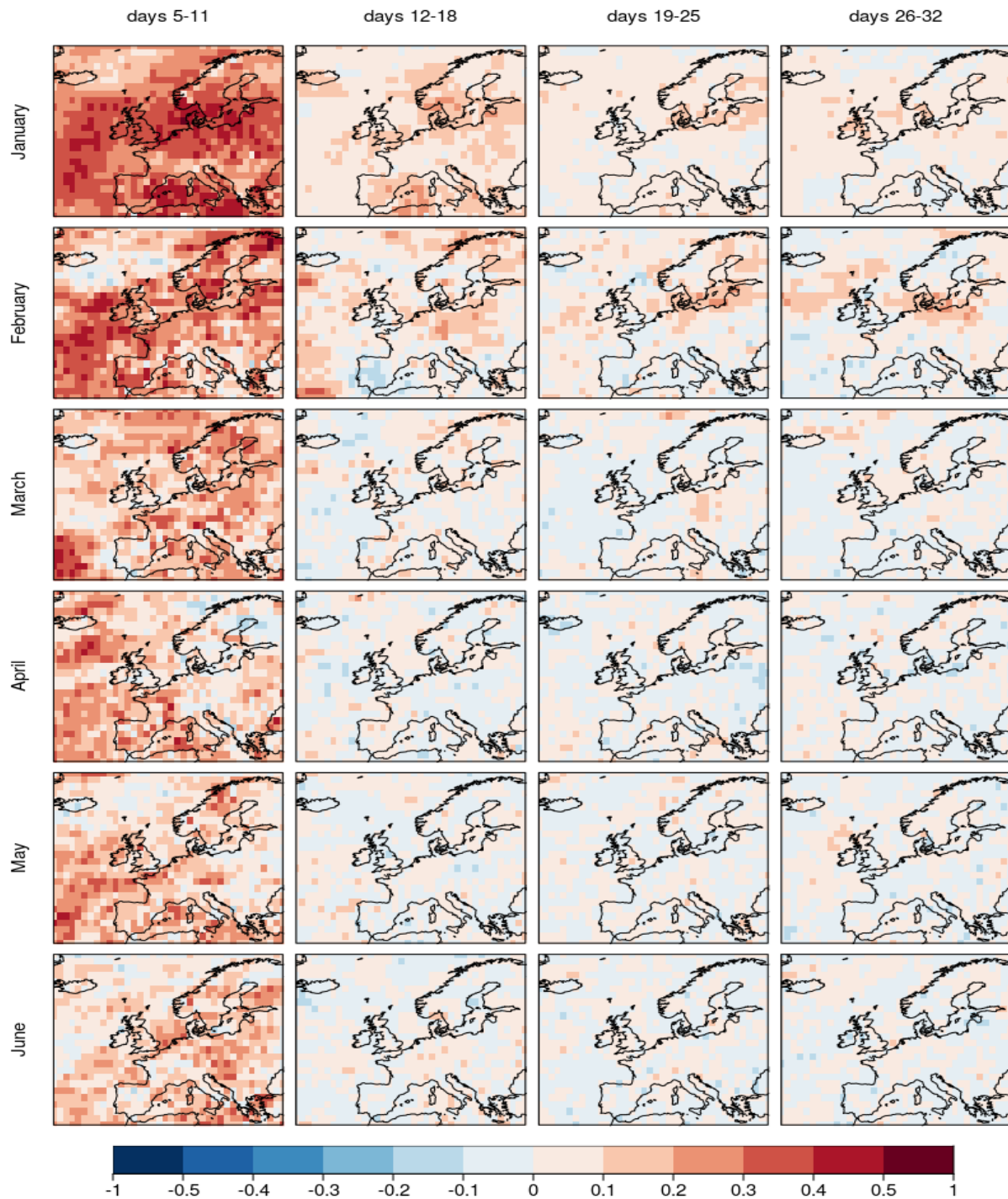


**Figure A. 12: FairBSS for the 10<sup>th</sup> percentile of 10m wind speed from ECMWF monthly prediction system for January to June and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts were calibrated with the variance inflation method.**



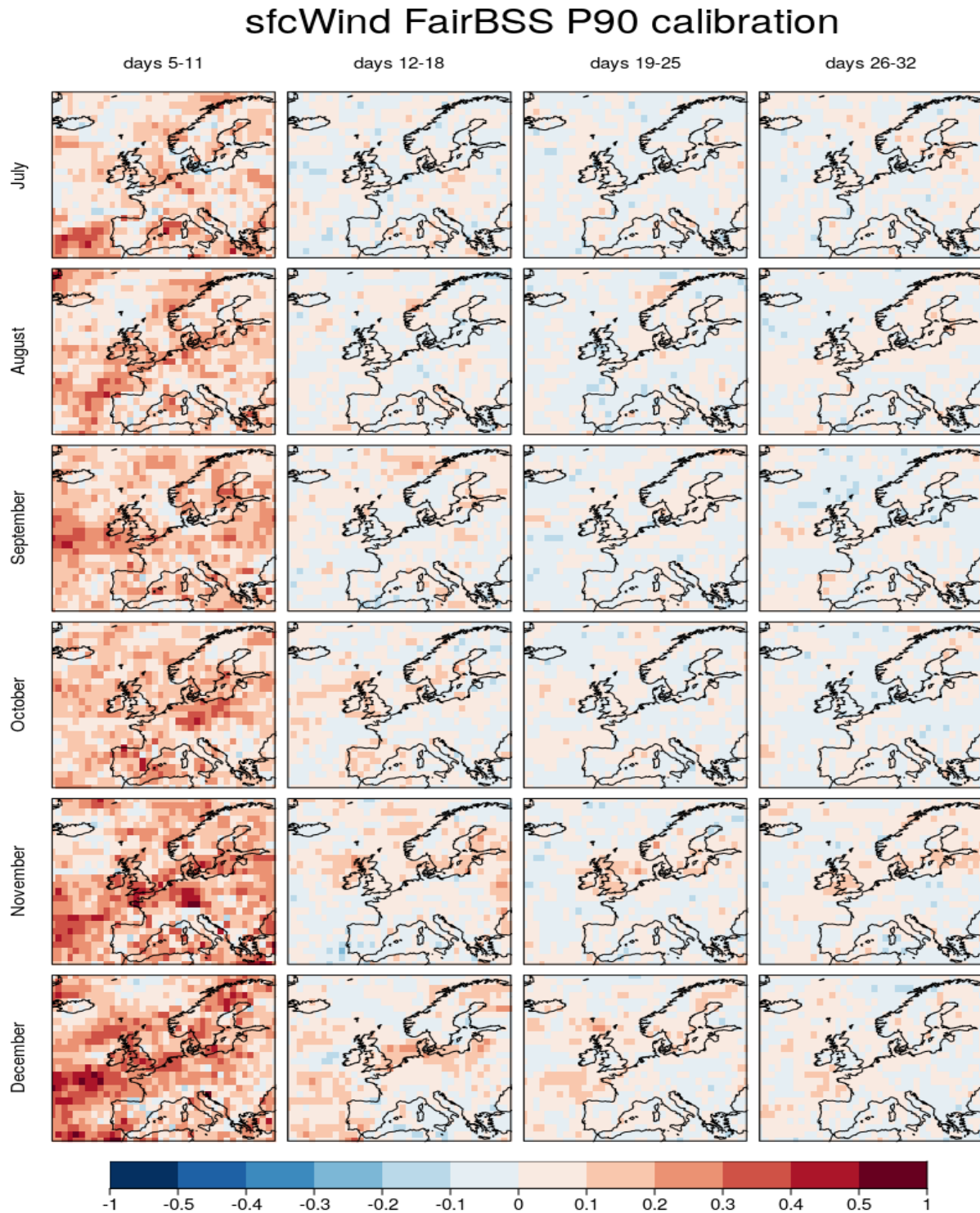
**Figure A. 13: FairBSS for the 10<sup>th</sup> percentile of 10m wind speed from ECMWF monthly prediction system for July to December and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts were calibrated with the variance inflation method.**

## sfcWind FairBSS P90 calibration



**Figure A. 14: FairBSS for the 90<sup>th</sup> percentile of 10m wind speed from ECMWF monthly prediction system for January to June and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts were calibrated with the variance inflation method.**





**Figure A. 15:** FairBSS for the 90<sup>th</sup> percentile of 10m wind speed from ECMWF monthly prediction system for July to December and 4 forecast times: days 5-11, days 12-18, days 19-25, days 26-32. The hindcast period 1996-2015 and reference dataset is ERA-Interim, hindcasts were calibrated with the variance inflation method.

## Annex 3 – Skill assessment of current S2S systems for the user’s selected case studies.

For each forecast, skill scores have been estimated using the hindcasts. This is key in order to understand the past performance of forecasts in this region. According to this skill estimates the forecasts can be trusted or need to be disregarded.

### Case study 1 – France, Germany 2017

	days 26–32	days 19–25	days 12–18	days 5–11
<i>RPSS</i>	0.03	0.02	0.12	0.3
<i>BS P10</i>	0.07	−0.02	0	0.21
<i>BS P90</i>	0.02	0.01	−0.01	0.19
<i>CRPSS</i>	−0.01	−0.01	0.05	0.23
<i>EnsCorr</i>	−0.01	0.05	0.28	0.58

**Table A.1: Skill scores for surface wind forecasts in the area of interest.**

	days 26–32	days 19–25	days 12–18	days 5–11
<i>RPSS</i>	0.13	0.06	0.12	0.37
<i>BS P10</i>	−0.03	−0.04	0.26	0.59
<i>BS P90</i>	−0.12	0.05	−0.15	0.29
<i>CRPSS</i>	0.01	0.01	0.09	0.43
<i>EnsCorr</i>	0.21	0.13	0.46	0.83

**Table A.2: Skill scores for temperature forecasts in the area of interest.**

	Apr	May	Jun
<i>RPSS</i>	0.01	−0.06	−0.05
<i>BS P10</i>	0.1	0.03	−0.03
<i>BS P90</i>	0.02	−0.17	−0.05
<i>CRPSS</i>	−0.01	−0.06	−0.07
<i>EnsCorr</i>	0.06	−0.34	−0.31

**Table A.3: Skill scores for temperature forecasts in the area of interest.**

SKILL	Forecast lead time		
Demand	3-months (April)	2-months (May)	1-month (June)
<b>RPSS</b>	−0.19	−0.06	−0.25
<b>BS P10</b>	−0.34	−0.25	−0.28
<b>BS P90</b>	−0.05	−0.18	−0.20
<b>CRPSS</b>	−0.17	−0.15	−0.18
<b>EnsCorr</b>	−0.21	−0.21	−0.23

**Table A.4: Skill scores for demand forecasts in the area of interest.**

	Apr	May	Jun
<i>RPSS</i>	−0.1	−0.07	−0.03
<i>BS P10</i>	−0.03	0.01	0.02
<i>BS P90</i>	−0.22	−0.08	−0.08
<i>CRPSS</i>	−0.1	−0.07	−0.08
<i>EnsCorr</i>	−0.41	−0.67	−0.08

**Table A. 5: Skill scores for surface wind forecasts in the area of interest.**

	month 3	month 2	month 1
fRPSS	−0.08	−0.10	−0.13
fBSS P10	−0.20	−0.08	−0.03
fBSS P90	0.02	0.00	−0.10
fCRPSS	−0.05	−0.03	−0.08
EnsCorr	−0.16	−0.05	−0.14

**Table A.6: Skill scores of surface solar radiation forecasts for case 2. Values marked with (\*) passed a significance test at 10% level.**

	month 3	month 2	month 1
fRPSS	-0.08	-0.09	-0.13
fBSS P10	-0.20	-0.08	-0.03
fBSS P90	0.00	0.00	-0.10
fCRPSS	-0.05	-0.03	-0.08
EnsCorr	-0.16	-0.05	-0.14

**Table A.7: Skill scores of photovoltaic capacity factor forecasts for case 2. Values marked with (\*) passed a significance test at 10% level.**

	Start Date		
	Forecast month 3	Forecast month 2	Forecast month 1
RPSS	-0.15	-0.07	0.23
BS P10	-0.07	-0.05	0.24
BS P90	-0.05	-0.06	0.21
CRPSS	-0.07	-0.06	0.10

**Table A.8: Skill scores of precipitation forecasts.**

	Start Date		
	Forecast month 3	Forecast month 2	Forecast month 1
RPSS	0.05	0.06	0.35
BS P10	0.08	0.11	0.37
BS P90	0.12	-0.02	0.18
CRPSS	0.06	-0.01	0.29

**Table A.9: Skill scores of inflows Rhine forecasts.**

	Start Date		
	Forecast month 3	Forecast month 2	Forecast month 1
RPSS	0.13	0.15	0.26
BS P10	0.36	0.46	0.70
BS P90	-0.25	0.06	0.49
CRPSS	0.04	0.18	0.43

**Table A.10: Skill scores of inflows Elbe forecasts.**

	Start Date		
	Forecast month 3	Forecast month 2	Forecast month 1
RPSS	0.12	0.21	0.42
BS P10	0.33	0.50	0.35
BS P90	0.04	-0.04	-0.19
CRPSS	0.11	0.16	0.13

**Table A.11: Skill scores of inflows Neckar forecasts.**

### Case study 3 – Spain, 2016

	days 26–32	days 19–25	days 12–18	days 5–11
<i>RPSS</i>	0.01	0.03	0.03	0.05
<i>BS P10</i>	−0.09	0	0.1	−0.12
<i>BS P90</i>	0.01	0.03	−0.07	0.04
<i>CRPSS</i>	−0.03	0	0.02	0.06
<i>EnsCorr</i>	0.17	0.21	0.07	0.33

**Table A.12: Skill scores for wind speed forecasts in the area of interest.**

	days 26–32	days 19–25	days 12–18	days 5–11
<i>RPSS</i>	0.38	0.35	0.21	0.42
<i>BS P10</i>	−0.01	0.08	0.02	0.31
<i>BS P90</i>	0.08	0.18	0.02	0.32
<i>CRPSS</i>	0.27	0.27	0.14	0.34
<i>EnsCorr</i>	0.66	0.66	0.51	0.77

**Table A.13: Skill scores for temperature forecasts in the area of interest.**

### Case study 4 – Sweden, 2015

Start Date			
	Forecast month 3	Forecast month 2	Forecast month 1
<b>RPSS</b>	-0.06	0.10	0.13
<b>BS</b>	-0.09	0.12	0.06
<b>P10</b>			
<b>BS</b>	-0.02	0.08	0.17
<b>P90</b>			
<b>CRPSS</b>	-0.00	0.04	0.10

Table A.14: Skill scores of precipitation forecasts for June for the Umeälven river.

Start Date			
	Forecast month 3	Forecast month 2	Forecast month 1
<b>RPSS</b>	0.21	0.43	0.81
<b>BS</b>	0.23	0.42	0.78
<b>P10</b>			
<b>BS</b>	0.18	0.29	0.91
<b>P90</b>			
<b>CRPSS</b>	0.16	0.39	0.74

Table A.15: Skill scores of snow water equivalent forecasts for July for the Umeälven river.

Start Date			
	Forecast month 3	Forecast month 2	Forecast month 1
<b>RPSS</b>	0.27	0.48	0.67
<b>BS</b>	0.39	0.50	0.71
<b>P10</b>			
<b>BS</b>	0.12	0.51	0.78
<b>P90</b>			
<b>CRPSS</b>	0.23	0.44	0.74

Table A.16: Skill scores of river inflows forecasts for July for the Umeälven river.



### Case study 5 – Romania, 2014

	days 26–32	days 19–25	days 12–18	days 5–11
<i>RPSS</i>	0.08	0.03	0.01	0.2
<i>BS P10</i>	0.08	0.1	0.06	0.03
<i>BS P90</i>	−0.1	−0.07	0.21	−0.01
<i>CRPSS</i>	0.02	0.01	0.05	0.14
<i>EnsCorr</i>	0.09	0.16	0.11	0.42

**Table A.17: Skill scores for temperature forecasts in the area of interest.**

	days 26–32	days 19–25	days 12–18	days 5–11
<i>RPSS</i>	0.13	0.16	0.28	0.49
<i>BS P10</i>	0.06	0.03	0.38	0.58
<i>BS P90</i>	−0.03	−0.04	0.19	0.21
<i>CRPSS</i>	0.09	0.1	0.22	0.47
<i>EnsCorr</i>	0.22	0.38	0.49	0.82

**Table A.18: Skill scores for minimum temperature forecasts in the area of interest.**

### Case study 6 – USA, 2015

	Start Date		
	Oct	Nov	Dec
<b>RPSS</b>	0.35	0.39	0.35
<b>BS P10</b>	-0.07	-0.27	-0.16
<b>BS P90</b>	0.1	0.04	0.07
<b>CRPSS</b>	0.14	0.11	0.14
<b>EnsCorr</b>	0.55	0.54	0.51

**Table A.19: Skill scores for surface wind forecasts in the area of interest.**

	CF IEC1				CF IEC2				CF IEC3		
	Oct	Nov	Dec		Oct	Nov	Dec		Oct	Nov	Dec
RPSS	0.21	0.26	0.2		0.23	0.25	0.24		0.31	0.31	0.25
BS P10	-0.17	-0.25	-0.16		-0.18	-0.23	-0.16		-0.22	-0.34	-0.18
BS P90	0.02	-0.02	0		0.06	0	0.03		0.09	0.09	-0.07
CRPSS	0.11	0.08	0.07		0.11	0.08	0.08		0.12	0.09	0.09
EnsCorr	0.49	0.44	0.4		0.5	0.45	0.42		0.51	0.47	0.44

**Table A.20: Skill scores for CF forecasts in the area of interest.**

### Case study 7 – France, Europe 2018

	days 26–32	days 19–25	days 12–18	days 5–11
<i>RPSS</i>	0.18	0.12	0.31	0.64
<i>BS P10</i>	0.09	0.06	0.19	0.42
<i>BS P90</i>	0.15	0.02	0.09	0.35
<i>CRPSS</i>	0.09	0.08	0.24	0.54
<i>EnsCorr</i>	0.37	0.38	0.6	0.87

**Table A.21: Skill scores for temperature forecasts in the area of interest.**

### Case study 8 – Spain, 2015

	Dec	Jan	Feb
<i>RPSS</i>	-0.06	-0.03	0.05
<i>BS P10</i>	-0.24	-0.06	-0.05
<i>BS P90</i>	-0.03	-0.03	-0.16
<i>CRPSS</i>	-0.17	-0.19	-0.08
<i>EnsCorr</i>	-0.05	-0.45	0.35

**Table A.22: Skill scores for wind speed forecasts in the area of interest.**

	Dec	Jan	Feb
<i>RPSS</i>	-0.16	-0.11	-0.06
<i>BS P10</i>	-0.1	-0.09	-0.08
<i>BS P90</i>	-0.18	-0.18	-0.09
<i>CRPSS</i>	-0.15	-0.14	-0.05
<i>EnsCorr</i>	-0.35	-0.26	0.12

**Table A.23: Skill scores for wind capacity factor (IC2) forecasts in the area of interest.**

	Dec	Jan	Feb
<i>RPSS</i>	-0.07	-0.1	-0.06
<i>BS P10</i>	-0.02	0.1	0.07
<i>BS P90</i>	-0.13	-0.24	-0.22
<i>CRPSS</i>	-0.09	-0.09	-0.04
<i>EnsCorr</i>	-0.18	0.14	0.19

**Table A.24: Skill scores for temperature forecasts in the area of interest.**

<b>Demand – Spain – Mar 2018</b>			
<b>Full calibration period: 2000 – 2016</b>			
	Start: Dec 2017	Start: Jan 2018	Start: Feb 2018
	Lead: 3 months	Lead: 2 months	Lead: 1 month
<b>RPSS</b>	-0.15	-0.34	-0.4
<b>BS P10</b>	-0.04	-0.24	-0.35
<b>BS P90</b>	-0.2	-0.18	-0.4
<b>CRPSS</b>	-0.39	-0.56	-0.82
<b>EnsCorr</b>	-0.33	-0.59	-0.75

**Table A.25: Skill scores for demand forecasts in the area of interest.**

<b>Wind power – Spain – Mar 2018</b>			
<b>Full calibration period: 2000 – 2016</b>			
	Start: Dec 2017	Start: Jan 2018	Start: Feb 2018
	Lead: 3 months	Lead: 2 months	Lead: 1 month
<b>RPSS</b>	-1.96	0.0	-0.07
<b>BS P10</b>	-0.86	-0.21	-0.18
<b>BS P90</b>	-0.86	-0.1	-0.21
<b>CRPSS</b>	-0.77	-0.18	-0.21
<b>EnsCorr</b>	-1	-0.12	0.05

**Table A.26: Skill scores for Spanish wind power forecasts in the area of interest.**

	<b>Start: Dec 2017</b>	<b>Start: Jan 2018</b>	<b>Start: Feb 2018</b>
	Lead: 3 months	Lead: 2 months	Lead: 1 month
<b>RPSS</b>	-0.03	0.09	0.17
<b>BS P10</b>	-0.07	0.01	0.29
<b>BS P90</b>	-0.08	0.05	0.16
<b>CRPSS</b>	-0.09	-0.07	0.18

**Table A.27: Skill scores for precipitation in the area of interest.**

	month 3	month 2	month 1
fRPSS	0.06	0.05	0.05
fBSS P10	-0.09	0.00	0.02
fBSS P90	-0.02	-0.02	-0.11
fCRPSS	-0.02	-0.02	0.01
EnsCorr	-0.14	-0.04	0.05

**Table A.28: Skill scores of surface solar radiation forecasts for case 2. Values marked with (\*) passed a significance test at 10% level.**

	month 3	month 2	month 1
fRPSS	0.05	-0.06	0.05
fBSS P10	-0.09	0.00	0.02
fBSS P90	-0.02	-0.02	-0.11
fCRPSS	-0.02	-0.02	0.01
EnsCorr	-0.14	-0.04	0.05

**Table A.29: Skill scores of photovoltaic capacity factor forecasts for case 2. Values marked with (\*) passed a significance test at 10% level.**



## Annex 4 – Factsheets on the case studies.

The case studies presented in this document support the proof of concept phase of the project and are linked to other activities involving not only climate scientists:

- 1) Understand how potential users would have benefited from S2S forecasts in those contexts. It might be understood as an evaluation from the users' point of view that will be included in Deliverable 2.2.
- 2) Help defining the Decision Support Tool (WP5).

To help the interaction within the interdisciplinary teams that will use the case studies as a tool, a series of factsheets have prepared summarizing the case studies. These factsheets are going to be available on the project website: <https://s2s4e.eu/info/case-studies>

These factsheets are organized in three sections:

- Short summary of the case study including region, period, interest for the energy sector.
- Analysis of the event
- Available forecast



Figure A. 16: Factsheet of case study 6.

S2S4E is an innovative service to improve renewable energy variability management. We produce new research methods exploring the frontiers of weather conditions for future weeks and months and a decision support tool for the renewable industry. Our long-term goal is to make the European energy sector more resilient to climate variability and extreme events



**[www.s2s4e.eu](http://www.s2s4e.eu)**



**Coordinator contact:** Albert Soret  
Barcelona Supercomputing Center (BSC)  
Carrer de Jordi Girona, 29-31  
08034 Barcelona (Spain)  
**[s2s4e@bsc.es](mailto:s2s4e@bsc.es)**



Follow us on Twitter  
and Facebook!

**[@s2s4e](https://twitter.com/s2s4e)**