

Understanding and interpreting seasonal climate forecasts

Data sources

The seasonal climate forecasts are summarised from an ensemble of bias-corrected forecasts (25 members) generated by the ECMWF's long-range forecasting system SEAS5 and accessed through the Copernicus Climate Data Store. ERA5 data was used for bias correction to local conditions.

Forecasts of broad seasonal climate averages, not daily weather

Seasonal climate forecasts are model predictions of how the weather will evolve over the next 3-6 months. Climate systems are chaotic and so these predictions, which look far out into the future, cannot be used to predict the day-to-day weather. Instead, they are used to look at whether the next 3-6 months will, on average, show broad differences to normal conditions. We adopt the common practice of summarising predictions in terms of whether they fall into one of three categories (for temperature and precipitation):

- Cooler/drier than normal
- Normal
- Warmer/wetter than normal

The temperature/precipitation values which separate these three classes represent the *terciles* of weather data over a historic period (1993-2019). i.e. historic seasonal forecast data for each variable are ordered and then split into three equal parts, each containing a third of the observations. In Vansjø, the following boundaries define the terciles, and were used when categorising future predictions:

Season	Class boundary	Seasonal mean temperature (°C)	Seasonal precipitation sum (mm)
Early summer	Below normal – Normal	14.3	209
	Normal – Above normal	15.1	265
Late summer	Below normal – Normal	11.6	219
	Normal – Above normal	12.4	294
Early winter	Below normal – Normal	-1.2	124
	Normal – Above normal	0.5	173
Late winter	Below normal – Normal	1.2	185
	Normal – Above normal	2.8	243

Forecast uncertainty

A seasonal forecast has no value without an indication of how much the predictions can be trusted. We consider two main sources of measures of forecast quality.

- 1) Prediction spread across terciles/tercile probability:** Seasonal climate forecasts are *probabilistic*. A model is run from a range of plausible initial conditions, resulting in a set (*ensemble*) of predictions (*ensemble members*). These represent uncertainty in the model, are all equally likely, and allow the probability of different events to be estimated. For example, the percentage of members which fall within a given tercile (below, normal or above the seasonal average) is interpreted as the *probability* of that tercile.

The probability of a tercile was discretized into four categories, ranging from very low to high, as follows:

Qualitative probability of the tercile	% of members which predict the tercile
Very low	Less than 35%
Low	35% - 50%
Medium	50% - 65%
High	65% or more

Example 1: 95% of forecast members for next summer suggest it's going to be warmer than average. Conclude: there is high probability of it being warmer than average.

Example 2: 10% of forecast members suggest it's going to be cooler than average, 40% that it will be normal, and 50% that it will be warmer than average. Conclude: there is medium probability of it being warmer than average. We can also say there is high probability of it *not* being cooler than average.

- 2) **Historic performance:** How well historic forecasts matched observations. A strong signal in a forecast (i.e. low spread between members) is useless or even dangerously misleading if the forecast system has no track record of successfully predicting such events in the past. Here, historic performance was assessed using:
 - 1) The **ROCSS** (Relative Operating Characteristic skill score). This is calculated for each weather variable, season and tercile. It measures how well the forecast discriminates between binary events (here, whether the forecast correctly predicts observations occurring within the tercile). ROCSS values range from 1 (a perfect forecast) to -1 (a perfectly bad forecast). Zero indicates no skill compared to a random prediction. We chose to define forecasts as having historic skill when the ROCSS was significantly positive (i.e. the probability of obtaining this ROCSS was less than 5% under the assumption of no forecast skill).
 - 2) **Tercile plots.** The shading in each square represents the likelihood of each tercile for each year (i.e. the proportion of forecast members which predict that tercile). Darker shades depict higher likelihood. White dots indicate the terciles which were actually observed (using the ERA5 dataset). Where white dots coincide with dark (high likelihood) terciles, this indicates good model predictions for that year. ROCSS values are reported next to the plots, and significant values (95% confident interval) are marked with an asterisk.

Summarising forecast quality

Both forecast spread across terciles and historic performance must be taken into account when judging the quality and strength of the forecast and deciding whether action should be taken based on it. We combine forecast spread and historic skill to give an overall indication of the confidence that can be placed in each tercile of the forecast, as follows:

Probability of the tercile	Historic skill	Confidence that tercile will happen	Confidence that tercile won't happen
Very low (less than 35%)	Some (significant ROCSS)	Very low	High
Low (35% - 50%)		Low	Medium
Medium (50% - 65%)		Medium	Low
High (65% or more)		High	Very low
Any value	None (insignificant ROCSS)	None	None

Forecasts are issued by NIVA as part of the ERA4CS-funded WATExR project

