Responses to Reviewer 1

General Comments

This paper proposes a new forecast technique for water supply forecasting in Australia designed to allow for more frequent updates as well as to improve forecast accuracy and to make the forecasts more timely. Currently operational forecasts are released 7 days into a forecast target period. This proposed technique would include the entire forecast period in the future. Overall, the paper was logically organized and easy to follow. The verification metrics used were applicable to the problem and applied in a logical manner. The results supported the conclusions in the paper well.

I am curious why the authors did not investigate incorporation of weather prediction into their algorithm. This would seem to be a major area ripe for improvement as the forecast skill at the weather timescale (0-2 weeks into the future) is much larger than at climate time scales and is ever increasing as computer power continues to grow.

Response: The reviewer makes a valid point about the potential for weather forecasts to improve the skill of streamflow forecasts up to two weeks into the future. It is an active area of research within our research group however it is not within the scope of this study. This study considers the forecasting of three-month seasonal streamflow totals using statistical methods where most of the skill comes from initial catchment conditions and limited predictability stems from climate.

It was striking that there was no discussion on how forecast utilization. I recommend including information on how the current forecasts are utilized and how the new forecasts might improve application.

Response: We agree. We are able to include more specific examples about how streamflow forecasts are used by water managers in Australia. We have prepared several paragraphs on 1. Overall forecast use 2. More specific uses, and 3. Advantage of earlier forecasts. These paragraphs are included near the beginning of the introduction (P1 L29 – P2 L13).

Specific comments

[page 3; lines 9-11] How is the undesirability of forecasts beyond 1 one month consistent with the premise of this paper of improving 3 month forecasts?

Response: We explain that the primary source of skill is initial catchment conditions and catchments can have limited memory or short response times (P3 L24-25). It is reasonable to expect that a three-month-total forecast issued today will have much higher skill than a three-month-total forecast issued four weeks ago.

[page 3; line 13] What is N equal to?

Response: "N-days" no longer appears in the manuscript. We have replaced all instances with more precise descriptions.

[page 4; line 23] is daily updated subsurface temperatures really necessary for this method? Does it change that much from one day to the next?

Response: The subsurface ocean temperatures will not change much from day to day and, strictly speaking, it would not be essential to have daily updates to use subsurface temperatures in our approach. We believe the comment about not having daily subsurface temperatures was distracting and so it has been removed.

Responses to Reviewer 2

General Comments

The paper covers the important topic of forecast release time of monthly to seasonal forecasts. Many available observation products used for statistical forecasting, such as mean monthly SST of the last month, are generally available several days after the beginning of the forecast period. Additional time is needed for data control, the generation of the seasonal forecasts and the development of key messages and other communication products. In the Australian forecast system presented here the forecasts are generally issued with a lag of 7 days. New data products based on daily SST from NOAA are tested for a timelier release of the forecast. The verification of the results is straightforward and the methods used for verification, PIT and CRPSS, are suitable for this topic. The results show only a small degradation of average skill for forecasts with 7-days lead time compared to the current method. This is a good message as timelier release of forecast increases its potential value in decision making. Another advantage of the presented method is that the predictands are not limited to calendar months any more. In theory when using daily data a forecast of the next 28 days could be released every week. The paper is of great importance, well written and easy to understand. It fits nicely the topic of the special issue and should be foreseen for publication.

The only thing I am struggling with a little bit is that the current version of the paper is between a technical study and a scientific paper. Additional references of methods used for statistical forecasting and the differences compared to the BJP and the advantages of the BJP should be presented in the introduction.

Response: The reviewer makes a useful point and we agree it would be beneficial to include more information about BJP and how it fits in the frame of statistical seasonal streamflow forecasting methods in Australia. We have added a number of references to the introduction and indicated why BJP was selected for operational use (P3 L32 – P4 L2).

An additional interesting verification metric would be the sharpness of the predictive uncertainty of the different lead times compared to the original method (not only in the combined measure CRPSS). As the training data set of the available daily SST data product (starting 1982) is smaller than the training data set of the monthly SST data (starting 1854 but also depending on the length of the stream- flow observation record) it is expected that the parameter and the total uncertainty of the BJP predictions are larger for the predictors with a smaller observation period.

Response: Sharpness is a useful measure and we agree it could be useful to have a distinct analysis of sharpness in our paper. We have investigated how forecast sharpness changes as forecast lead time is increased. We measure sharpness as the width of forecast intervals relative to lead time 0 forecasts. The results are now included and discussed in the paper, and the methods updated accordingly (Figure 11 and P7 L13-17 and P9 L23–32)

The point about forecast sharpness and length of record is very valid. In our study, we set up models as consistently as possible, using a fixed period, to enable comparisons between forecast locations and seasons. In practice, the weather Bureau may use longer periods of streamflow data to establish the models at individual forecast locations and thus achieve improved forecast sharpness. Whilst an important point from an operational point of view, we haven't discussed this matter in the paper.

Specific Comments:

p 4, l 5: Explain shortly the main characteristics of the different runoff regimes for non-Australian river experts. Probably add the regions of the different runoff regimes to Figure 1.

Response: Figure 1 has been updated to include climate zones to which runoff regimes are closely related. The text has been updated to reflect the change (P4 L14-15).

p 4, l 6: Add range of catchment areas considered: ". . . ranging from 102 to 36 230 km2 . . ."

Response: Easily done (P4 L16).

p 4, l 10: Length of streamflow observation records? This is important to get an impression of the number of data points used for parameter estimation of the BJP (see general comment).

Response: Thanks for pointing out the omission. The data period coincides with the verification period. We have added this information to section 2.2 on streamflow data (P4 L18).

p 4, l 15: Make clear that the predictands are still the three-month totals starting at the beginning of each month.

Response: Easily done (P4 L24-25).

p 4, l 25 – l 30: Is there a relationship between subsurface ocean temperatures and SOI lagged by two / three months with the predictant? These predictors could still be used in the system with lead times up to 28 days.

Response: Yes, however we cannot guarantee nor quantify the usefulness of this approach. We believe the comment about not having frequently updated SOI was distracting and so it has been removed.

p 8, l 8: Add % to the CRPSS values. In many other applications a maximum CRPSS of 1 (100% in your case) is used. This could be a little bit confusing.

Response: Thanks for pointing out the omission. We have corrected the methods section to point out that the CRPS skill scores should be interpreted as percentages (P7)

Responses to Reviewer 3

Summary

This paper proposes to increase the lead time at which operational forecasts of three month streamflow volumes can be issued by the Bureau of Meteorology in Australia. Currently, the data collection and processing cause the forecasts to be sometimes issued in the second week of the forecast period. The development of sub-seasonal forecasts in the region prompts the need for timelier forecasts. In this paper, the authors first present a forecasting method that would allow forecasts to be issued several days prior to the forecast period. Then, they investigate the relationship between forecast lead time and forecast quality. The authors show that being able to issue forecasts several days prior to the forecast period can be achieved without any loss in forecast reliability. Nevertheless, accuracy decreases as the lead time increases. The optimal trade-off between forecast accuracy and forecast lead time must then be decided on. Forecasts with lead times shorter than 7 days may offer an acceptable loss in accuracy as compared to the gain in anticipation time.

General comments

The topic is an interesting one and allows to look at forecast quality within an operational context, with timing constraints The paper is very well written and is easy to follow. The proposed figures are also easy to read and well illustrate the paper. A catchment set is used to obtain relatively general results and the forecast evaluation criteria are relevant. The detailed comments are only minor suggestions and questions I had when reading the manuscript.

Detailed comments

Throughout the paper, "N-days lead time" and "N-month lead time" were confusing at first, probably due to the "N". Could you maybe reformulate or add a small sentence in the first occurrences to define the terms?

Response: "N-days" no longer appears in the manuscript. We have replaced all instances with more precise descriptions.

In Section 2.3, have you compared the performance of the current forecasting system with that of the proposed system? Do you have an idea of their difference in skill, if the forecasts are both run at lead time 0?

Response: A very interesting question and one we pondered how to address when designing our study. Ultimately, we decided there are too many confounding factors to include a direct comparison in our paper.

However, in Australia, most of the forecast skill is acquired through the initial catchment condition predictors. Indeed, many models have no climate predictors. Therefore we expect the skill of the two systems to be relatively on par at 0 lead time.

Page 8, Lines 8-9: Is there a specific reason why you chose -5 and 5 as thresholds?

Response: These thresholds are essentially rules-of-thumb based on years of experience and some bootstrap experiments in previous studies.

Page 8, Lines 12-13: Could you maybe add a short sentence to further explain this? For instance, is the number of cases below -5 sufficient to assert this?

Response: The BJP modelling approach should theoretically produce a reliable climatology-like forecast in the absence of predictor-predictand relationships. In practice, under our stringent leave-five-years-out cross-validation, instances of skill negative skill can occur for various reasons related to short data records, low flows, a high density of zero flows, limited catchment memory etc. We have updated the text accordingly (P8 L31–33). We have also included more information in the methodology section about BJP's propensity to produce climatology-like forecasts in the absence of a forecasting relationship (P6 L11–13).

Page 9, Lines 26-27: This sentence is, I think, very important in this paper, as it links forecast quality with operational expectations and requirements. I am just curious: have the authors further investigated the topic? How would you discuss / How have you discussed the acceptable loss in quality with operational managers? Is there a way to weigh the loss in skill more pragmatically, and relate it to the gain in anticipation time?

Response: The study was motivated by the demands of water managers to have timelier forecasts and so the weather Bureau will continue to engage with operational managers to determine the best outcome for them. We have revised the introduction to include more information about the usage of the forecasts in various agencies and the potential benefits of timelier forecasts (P1 L29 – P2 L13).

Figure 4 and 5: It is interesting to put in parallel these figures with Figure 1 and Table 1. But how are the catchments ordered? I could not find it in the text, nor when looking at Table 1.

Response: The order of catchments in Table 1 and Figures 4 and 5 was not consistent – thanks for detecting. We now order the catchments clockwise from northeast Australia to southwest Western Australia.

Figure 5: If I understand correctly, the objective of this figure is to show that the patterns between 0-day lead time and 7-days lead time are quite similar. Have you thought of showing the skill score of the 7-days lead time computed with the 0-day lead time as reference to highlight the differences between Figures 4 and 5? I understand that Figures 6 and 7 go in this direction but Page 8 Lines 3-6 could be illustrated in that way as well.

Response: It is certainly feasible to calculate the skill scores as suggested, to show the relative difference in error between the 0-day and 7-day lead time forecasts. We have included the suggested figure and updated the text accordingly (Figure 6; P8 L20-24). However, as rightly pointed out, the differences in skill are illustrated by subsequent figures.

Optimising Seasonal Streamflow Forecast Lead Time for Operational Decision Making in Australia

Andrew Schepen[1], Tongtiegang Zhao[2], Q.J Wang[2], Senlin Zhou[3] and Paul Feikema[3]

[1]CSIRO Land and Water, Dutton Park, 4102, Australia
5   [2]CSIRO Land and Water, Clayton, 3168, Australia
[3]Bureau of Meteorology, Melbourne, 3001, Australia

Correspondence to: Andrew Schepen (Andrew.Schepen@csiro.au)

Abstract. Statistical seasonal forecasts of three-month streamflow totals are released in Australia by the Bureau of
10   Meteorology and updated on a monthly basis. The forecasts are often released in the second week of the forecast period, due
to the onerous forecast production process. The current service relies on models built using data for complete calendar months,
meaning the forecast production process cannot begin until the first day of the forecast period. Somehow, the Bureau needs to
transition to a service that provides forecasts before the beginning of the forecast period; timelier forecast release will become
critical as sub-seasonal (monthly) forecasts are developed. Increasing the forecast lead time to one month ahead is not
15   considered a viable option for Australian catchments that typically lack any predictability associated with snowmelt. The
Bureau's forecasts are built around Bayesian joint probability models that have antecedent streamflow, rainfall and climate
indices as predictors. In this study, we adapt the modelling approach so that forecasts have ~~N~~ any number of days lead time.
Daily streamflow and sea surface temperatures are used to develop predictors based on 28-days sliding windows. Forecasts
are produced for 23 forecast locations with 0–14 and 21-days lead time. The forecasts are assessed in terms of CRPS skill
20   score and reliability metrics. CRPS skill scores, on average, reduce monotonically with increase in days of lead time, although
both positive and negative differences are observed. Considering only skilful forecast locations, CRPS skill scores at 7-days
lead time are reduced on average by 4 percentage points with differences largely contained within +5 to -15 percentage points.
A flexible ~~N days lead time~~ forecasting system that allows for any number of days of lead time could benefit Australian
seasonal streamflow forecast users by allowing more time for forecasts to be disseminated, comprehended and made use of
25   prior to the commencement of a forecast season. The system would allow for forecasts to be updated if necessary.

1 Introduction

The Australian Bureau of Meteorology (the Bureau) operates a statistical seasonal streamflow forecasting service to assist
water management agencies make informed decisions about water management strategies in the season ahead. The forecasts
provide probability distributions of the total volume of streamflow over the next three-month period. ~~Forecasts are~~ The forecasts
30   are used by a variety of groups, including federal, state and local governments and their agencies, such as water management

1

authorities; agriculture and water management sectors; private businesses, general public and local communities. The forecasts help reduce the uncertainty in seasonal flow volumes to be expected, and therefore provide users with more certainty in decision making. For example, water authorities use forecasts to assist decisions on water allocation and water restrictions, manage river operations, schedule environmental watering, develop water transfer strategy and provide water order advice.

5   They also use forecasts to guide future storage levels, help decide on releases, produce water allocation outlooks to inform water markets, and manage risks at construction sites along rivers. State government agencies use forecasts to make decisions about environmental monitoring of streams, schedule irrigation, assess flood potential, and determine agency resourcing and public messaging.

The forecasts are currently updated once per month. However, the forecasts are frequently not released until the second week

10   of the forecast season because the forecast production process is overly time-consuming. Feedback from a user survey showed many users would prefer the forecasts to be issued earlier in the month, which would fit in better with their reporting schedules and operational timelines. The need to have forecasts issued earlier becomes more pronounced when issuing a one-month forecast rather than a three-month forecast.

The first step in the forecast process is to gather predictor data and perform quality control. Once the predictor data are assured

15   to be of good quality, forecasts and forecast products are generated. The forecasts subsequently need to be checked for modelling errors and inconsistences. Once all the forecast products are ready, key messages and other communication products are developed. Once the communication strategy is in place, the forecasts are released. The Bureau now issues forecasts for over 200 locations and so, understandably, the forecast production process is a time-consuming one.

Currently, the statistical forecasting models rely on predictor data observed up to the day prior to the first day of the forecast

20   period. For example, a forecast for the austral spring (September–October–November) requires observed data up to (and including) the 31st of August. The forecasts effectively have no lead time and so the forecast lead time can be considered to be 0 ~~months or 0~~ days. Suppose predictor data were available immediately on the first day of the forecast period; then forecasts could be generated on the first day of the forecast period. Even in this case, the forecast release would occur several days into the forecast season, after forecast products and communication messages were created. In reality, by the time the forecasts are

25   released, the forecasts typically have a lag time of at least 7 days.

The lag time in forecast release diminishes the value of the forecasts. Ideally, the forecasts would be in the hands of decision makers well ahead of the forecast season. Furthermore, the lag time in forecast release is likely to become more prominent in the short term because the Bureau plans to release sub-seasonal (i.e. monthly) streamflow forecasts by 2017. It is therefore important to investigate ways to issue the forecasts earlier and to quantitatively analyse how forecast accuracy and reliability

30   are affected.

The Bureau's current forecasting system is built around the Bayesian joint probability (BJP) modelling approach (Wang and Robertson, 2011;Wang et al., 2009). Although other statistical methods have been previously investigated for seasonal streamflow forecasting in Australia (e.g. Westra et al., 2008;Piechota et al., 2001, 1998;Chiew and Siriwardena, 2005), the BJP approach was initially adopted for operations owing to its wide potential applicability to perennial and ephemeral

catchments, in cases of pervasive missing data and/or pervasive zero flows and in situations where multiple sources of predictability were identified. BJP models make use of two types of predictors: predictors representing initial catchment conditions and predictors representing the climate state. The predictors are selected following a rigorous predictor selection process (Robertson and Wang, 2012). Initial catchment condition predictors, which act as a soil-moisture proxy, include

5   streamflow totals over the preceding 1, 2 or 3 months. Predictors representing the climate conditions are monthly climate indices lagged up to 3 months (Schepen et al., 2012;Kirono et al., 2010) and include indices representing the El Niño Southern Oscillation, variability in the Indian Ocean and variability in the southern polar circulation.

Predictors in the BJP modelling approach vary by season and by catchment. Therefore in operational forecast production, it is necessary to prepare a wide range of predictor data from a variety of data sources. Since the Water Act (2007) was introduced

10   in Australia, the Bureau of Meteorology is authorised to collect streamflow data from gauge owners. Streamflow data is therefore easily obtained for most gauges; however, the data comes in a multitude of formats and needs to be processed by data managers prior to being made available to forecasters. Usually, daily streamflow data for each of the forecast locations is available with a few days lag. Climate data is sourced from the Bureau of Meteorology and the United States National Oceanic and Atmospheric Administration (NOAA). NOAA ERSST (Huang et al., 2015) monthly sea surface temperature (SST) grids

15   are used for calculation of climate indices such as Nino3.4 and the IOD mode index. The ERSST data set for the previous month is normally available by the 4th day of the month. Other indices, such as the Antarctic Oscillation Index (AAO) (Mo, 2000) representing the Southern Annular Mode can take up to a week to become available. The Bureau of Meteorology updates Southern Oscillation Index values relatively swiftly, normally within a couple of days.

As just described, up to 7 days delay in the forecast production process can be attributed to ~~delays~~delay in acquiring predictor

20   data. Much of the problem lies in the reliance on data for complete calendar months. An immediate and obvious resolution to the problem is to produce forecasts with 1-month lead time. ~~N~~Forecasts with 1-month lead time ~~forecasts~~ are certainly feasible in a statistical framework and are operationally useful in many parts of the world where snow-melt is a major source of seasonal streamflow predictability, e.g. the western United States (e.g. Pagano et al., 2004) and the Three Gorges system in China (Xu et al., 2007)~~.~~. However, increasing the forecast lead-time for Australian catchments beyond one month is undesirable because

25   the primary source of skill is initial catchment conditions and catchments have limited memory or short response times.

An alternative approach to get forecasts to users in good time is to make use of daily streamflow and climate data to generate predictors and forecasts with ~~N~~ any number of days lead time. Although it could be presumed that skill will tend to be reduced as lead time increases, it is not known to what degree skill will be impacted. ~~An~~The optimal ~~N-days~~forecast lead time ~~forecast is~~(in days) corresponds not only to the most skilful forecast that can be released prior to the beginning of the forecast period,

30   ~~given consideration~~but needs to allow for forecast reliability, forecast preparation and communication time, and ~~allowing~~giving enough time for users to ingest streamflow forecasts into their models ahead of the forecast season.

In this study, we develop BJP models to produce ~~N-~~forecasts with up to 21 days lead time ~~forecasts~~ for 23 catchments across Australia and seek to demonstrate that it is possible to release operational forecasts ahead of the commencement of the forecast target season. We investigate the availability of daily climate data and establish necessary modifications to the predictor

choices and length of record used to establish the models. Cross-validation forecasting experiments are conducted to evaluate the quality of forecasts ~~based on the N-days lead time approach.~~. The skilfulness of forecasts are compared for 0–14 and 21 days lead time. The results give an estimate of how forecast skill changes with increasing forecast lead time. Subsequently, the optimal ~~N-days~~ lead time ~~forecast~~ is determined, based on considerations of forecast skill, data availability and forecast preparation time.

The remainder of the paper is organised as follows. Section 2 describes the study forecast catchments, streamflow and climate data. Section 3 details the study methods, including information about the BJP modelling approach and forecast verification. Section 4 presents the results. Section 5 discusses the results. Section 6 concludes the paper.

~~.~~

2 Catchments and Data

2.1 Catchments

Twenty-three forecast locations are selected for this study including forecast locations in all states: Queensland, New South Wales, Victoria, Tasmania, South Australia, Western Australia; plus the Northern Territory. Flows at the forecast locations are a mixture of perennial, ephemeral and intermittent flows. Table 1 summarises information about the forecast locations, including long catchment name, short ID, state (in the jurisdictional sense), upstream catchment area and approximate centroid latitude and longitude. The catchments range in size from 102–36,230 km^2. Furthermore, the forecast locations are plotted on a map of Australia in Figure 1~~.~~, which also indicates whether the catchment is in a temperate, subtropical or tropical climate zone.

2.2 Streamflow data

Daily streamflow data for the period 1982–2012 is sourced from the Bureau of Meteorology for all 23 forecast locations. The Bureau sources the data from the respective data owners and performs quality control. Daily streamflow data records often contain missing values, for example, due to the failure of gauging equipment. The Bureau partially infills missing data records. Records for all forecast locations are infilled by linear interpolation if the number of missing days is equal to or less than 3 days. For some forecast locations, records are infilled by linear regression with nearby forecast locations if the number of missing days is equal to or less than 14 days.

In this study, the daily streamflow data are aggregated to 28 day totals for use as predictors and three-month totals (always beginning on the first day of the month) for use as predictands. Many of the data records still contain missing data after the infilling process, however the proportion of missing data is never greater than 10%.

2.3 Climate Data

The Bureau's current BJP seasonal streamflow forecasting models rely on predictors including surface and subsurface ocean temperatures, the Southern Oscillation Index (SOI) and the Antarctic Oscillation Index (AAO). The predictors are identified through a rigorous predictor selection process (Robertson and Wang, 2012). In this study, we adopt the climate predictors used

5 by the Bureau in their operational models with some changes. The biggest change is that we restrict the set of climate predictors to sea surface temperature (SST) predictors. ~~The subsurface ocean temperatures and SOI are not updated on a daily basis and it is therefore not possible to obtain up-to-date predictor values for earlier release forecasts. Because the SOI and subsurface temperatures reflect ENSO, we substitute~~, substituting an ENSO SST index wherever the climate predictor in the Bureau's operational model is subsurface ocean temperatures or SOI. SST climate predictors representing ENSO and the Indian Ocean

10 state are likely to be stable across a period of several weeks and are known to have strong lagged relationships with Australian rainfall up to three months ahead (Schepen et al., 2012). Therefore, the predictors are likely to remain valid predictors as forecast lead time is increased by a few weeks. On the other hand, it is less certain that the AAO, representing the Southern Annular Mode, will remain a valid climate predictor as lead time is increased since it has a weak lagging relationship with Australian seasonal rainfall (Schepen et al., 2012). In cases where AAO is a climate predictor in the Bureau's operational

15 model, no climate predictor is used in this study.

Daily SST data is obtained from the NOAA 1/4° daily Optimum Interpolation Sea Surface Temperature (daily OISST) (Reynolds et al., 2007). Full years of data are available from 1982 onwards. The daily OISST is constructed by combining data from satellites, ships, buoys, probes and ocean-dwelling robots. Interpolation fills in missing gaps and ensures a complete historical record with no missing data. Compared to monthly ERSST records, which go back to 1854, the daily OISST record

20 is relatively short but it is necessary for generating timelier forecasts. To obtain SST climate predictors we calculate 28-day average sea surface temperatures relative to 1982-2010 climatology. Area averaging is used to obtain the set of climate predictors: Nino3.4, Nino3, Nino4, the El Nino Modoki Index (Ashok et al., 2007), the Indian Ocean Dipole (Saji et al., 1999) and the Indonesian Index (Verdon and Franks, 2005). Figure 2 demonstrates that the monthly Nino3.4 values calculated from daily OISST and NOAA's monthly ERSSTv4 product are almost identical for the period 1982-2011 with $R^2 = 0.98$.

25 **3 Methods**

3.1 BJP forecasting models

Forecasting models are set up using the Bayesian joint probability (BJP) modelling approach (Wang and Robertson 2011; Wang et al. 2009). Separate forecasting models are established for each forecast location (23 locations), season (12 three-month seasons from the start of each month) and for 0–14 and 21 days lead time (16 lead times).

30 The BJP forecasting models make use of two types of predictors: predictors representing initial catchment conditions and predictors representing the climate state. For all models, the initial catchment condition predictor is fixed to the previous 28

days total flow volume. This predictor will not be optimum for all forecast locations; however, it is a pragmatic choice that is likely to represent the overall catchment wetness reasonably well. Additionally, it gives a consistent model set up without needing to undertake initial catchment condition predictor selection. SST climate predictors are adapted from the Bureau's current operational forecasting models as described in section 2.3. The same climate predictors are applied at all lead times to ensure a consistent model set up.

The full mathematical formulation of the BJP modelling approach is presented in Wang et al. (2009) and Wang and Robertson (2011). Here, we note some key features of the BJP modelling approach.

The BJP models are able to effectively handle missing and non-concurrent records. The BJP models are based upon the multivariate normal distribution after allowing for data transformation using either the log-sinh ~~(Wang et al. 2012) or Yeo-Johnson (Yeo and Johnson 2000) transformations.~~(Wang et al., 2012) or Yeo-Johnson (Yeo and Johnson, 2000) transformations. If a set of variables follow a multivariate normal distribution then a subset of those variables also follows a multivariate normal distribution. Thus the many instances of missing data in streamflow records, as described in section 2.2, are easily handled. Several of the forecast locations experience ephemeral and intermittent flows, which can result in a probability mass for zero flows. This problem is handled in the BJP modelling approach by treating zero flows as censored data, meaning that the observations of zero flow are treated as being of unknown precise value equal to or below zero. Uncertainty in the model parameters due to the short data records is handled by inferring parameters using Markov Chain Monte Carlo methods (MCMC). Probabilistic (ensemble) forecasts are produced using conditional multivariate normal distribution equations. When predictor-predictand relationships are weak, the BJP modelling approach is designed to produce reliable forecasts that approximate climatology (i.e. frequency distribution of historically observed streamflow).

Some aspects of the BJP implementation used in this study vary compared with those published in Wang et al. (2009) and Wang and Robertson (2011). We use fixed transformation parameters together with simplified parameter reparameterisations for more straightforward numerical inference~~.~~ (Zhao et al., 2016) . The changes reflect our experience in achieving more robust and efficient BJP modelling.

3.2 Verification

In this study, we assess the performance of forecasts for the period from JFM 1982 to DJF 2011-12. A separate BJP model is established for each season, forecast location, and lead time. Across different years, the model parameter inference and forecast process is cross-validated using leave-five-years-out cross-validation. ~~In other words, for~~For each historical forecast event to be tested, the data points for the year to be forecast plus the subsequent four years are left out. The leave-five-years-out procedure is designed to account for strong persistence in streamflows, potentially over many years.

Forecasts from the BJP modelling approach are probabilistic. The continuous ranked probability score (CRPS; Matheson and Winkler, 1976) is used to assess full forecast probability distributions, therefore involving forecast ensemble spread as well as

forecast accuracy. ~~For convenience we define CRPS as the average CRPS across forecast time periods t ($t = 1, 2, ..., T$),~~ The CRPS for a given forecast and observation is defined as,

$$CRPS = \frac{1}{T}\sum_{t=1}^{T}\int \left[F(y^t) - H(y^t - d_y^t)\right]^2 dy^t \quad CRPS = \int \left[F(y) - H(y - y_{obs})\right]^2 dy$$

(1)

~~Where~~ where ~~y^t~~ y is the forecast variable ~~for time period, t,~~ ~~d_y^t~~ y_{obs} is the ~~corresponding predictand~~ observed value, $F(.)$ is the forecast cumulative distribution function (CDF) $H(.)$ is the Heaviside step function, which equals 0 if ~~$y^t < d_y^t$~~ $y < y_{obs}$ and equals 1 otherwise. ~~To assess forecast value, the CRPS score is converted~~ Model forecasts are compared to ~~a skill score, to measure the relative improvement of the CRPS over CRPS$_{ref}$, the CRPS for leave-five-years-out climatology~~ reference forecasts by calculating skill scores:

$$CRPS_{SkillScore} = \frac{CRPS_{ref} - CRPS}{CRPS_{ref}} \quad CRPS\ skill\ score = \frac{\overline{CRPS_{ref}} - \overline{CRPS}}{\overline{CRPS_{ref}}} \times 100 \quad (\%)$$

(2)

where the overbar indicates averaging across a set of events. A climatology forecast is used as the reference in this study, although one may choose to use other reference forecasts. The CRPS skill score is positively oriented (whereas CRPS is negatively oriented). ~~A~~ As a percentage, a maximum score of 100 is indicative of perfect forecasts. A score of 0 indicates no overall improvement compared to the reference forecast. A negative score indicates poor quality forecasts in the sense that a naïve climatology forecast is more skilful. CRPS skill scores are calculated for each catchment and season.

Reliability refers to the statistical agreement of forecast probabilities with observed relative frequencies of events, which can be checked using probability integral transforms (PITs). At each lead time, the PIT values are calculated from the BJP-generated ensemble forecasts for every forecast event, forecast location and season, and pooled in the analysis. The PIT represents the non-exceedance probability of observed streamflow obtained from the CDF of ensemble forecast. If the ensemble spread is appropriate and the forecasts are free of bias then observations will be contained within the forecast ensemble spread. Reliable forecasts are evidenced by PIT values that follow a uniform distribution between 0 and 1.

Sharpness refers to the concentration of the ensemble members and is a property of the forecast only. Forecast sharpness is desirable provided the forecast is reliable. A common measure of sharpness is the width or, more pertinently, the relative width of a forecast quantile range. Comparisons between forecasts can be made by comparing average widths or using a sharpness boxplot (Gneiting et al., 2007). In our study, we compare the quantile range widths of forecasts with lead time of 1–21 days with the quantile range widths of forecasts with 0-days lead time. Averages are taken for each catchment and season.

4 Results

4.1 Reliability

We first assess the reliability of the forecasts for the different lead times. Histograms of PIT values are constructed for each lead time from 0–14 days (Figure 3). For each lead time, the PIT histograms follow a similar distribution: approximately

5 uniform but not perfectly so. There are a couple of reasons why the PIT histograms are not perfectly uniform. First, it is very difficult to model the shape of the forecast distributions perfectly, particularly under cross-validation. Second, a prior distribution over the data transformation parameters is influential in preventing very skewed distributions when there is limited data or many instances of zero flows. Thus, the model attempts to reliably reflect uncertainty, but cannot tightly fit the shape of the data in all cases.

10 On-the-whole, the PIT values suggest that the forecasts capture the range of observations sufficiently well and there is no strong evidence of bias. Therefore, the spread of the ensembles is deemed to be appropriate. However, note that the shape of the PIT histograms suggests that an alternative distribution, for example a fat-tailed distribution, may be able to fit the tails of high-flow events better. The PIT histograms exhibit a similar shape for all lead times, meaning that the forecasts are similarly reliable across all lead times.

15 ### 4.2 Accuracy ~~Skill~~

We now assess the accuracy ~~skill~~ of the forecasts for the different lead times. To assess forecast accuracy we calculate CRPS for each forecast location, season and lead time. It is possible to decompose the CRPS into components reflecting accuracy and reliability. Results in the previous section demonstrate that the forecasts are similarly reliable across all times. Therefore

20 differences in the CRPS will mainly reflect differences in forecast accuracy. CRPS scores are in the original units of the data and therefore it is difficult to compare across forecast locations and seasons. CRPS skill scores, as defined in Equation (2), however, measure the ~~relative~~percentage reduction in error relative to a reference forecast, and therefore it is easier to compare relative skill levels across forecast locations, seasons and lead times. Hereafter, we consider only CRPS skill scores with modelled leave-five-years-out climatological distributions as reference forecasts unless otherwise specified.

25 CRPS skill scores for forecasts with 0-days lead time are plotted in Figure 4. Forecast locations form rows and target seasons form columns. The CRPS skill scores vary highly between forecast locations and seasons, due to variations in catchment memory and climate predictability. The sources of predictability are not the main concern here though. Here, the aim is to analyse the changes in forecasting skill with increasing lead time. As an example, forecasts with 7-days lead time are plotted in Figure 5. By visual comparison, skill is overall reduced, however there are individual cases where skill increases. Instances

30 of positive CRPS skill score differences are conceivable given the small sample size (30 years) and associated uncertainty in the skill score. The patterns of skill across forecast locations and seasons are largely consistent at 0- and 7-days lead time. Overall, increasing the forecast lead time to 7 days does not ~~significantly decrease forecasting skill.~~ appear to significantly

8

decrease forecasting skill. To investigate further, CRPS skill scores for 7-day lead time forecasts are calculated with the 0-day lead time forecasts as the reference rather than climatology (Figure 6). The median CRPS skill score (considering all catchments and seasons) is -1.8. Catchment by catchment, the median CRPS skill score across all seasons ranges from -5.1 to 0.8, indicating that CRPS scores typically increase by less than about 5%, although larger increases in errors are observed in individual catchments and seasons.

Figure 67 illustrates for lead times of 0-, 7- and 14-days, the proportion of cases (catchments and seasons) where CRPS skill score thresholds are exceeded. CRPS skill scores between -5 and 5 are considered unskilful, below -5 is considered negative skill, and above 5 is considered skilful. The proportion of cases where certain CRPS skill score thresholds are exceeded decreases as the lead time increases. A CRPS skill score of -5 is exceeded in approximately 95%, 94% and 92% of cases for 0-, 7- and 14-days lead time respectively, indicating only a small proportion of cases exhibit skill worse than climatology. The small difference in exceedance probabilities with increasing lead time is consistent with the knowledge that the BJP modelling approach should produce forecasts approximating climatology in the absence of any real forecasting skill. Under stringent leave-five-years-out cross-validation, instances of skill negative skill can occur for various reasons related to aridity, poor catchment memory, extreme events and data problems. A CRPS skill score of 5 is exceeded in approximately 66%, 59% and 57% of cases for 0-, 7- and 14-days lead time respectively. The larger difference in exceedance probabilities suggests that the number of days lead time is important for skilful forecast locations and catchments.

To evaluate the change in forecast skill for each 1 day increase in lead time, we calculate the difference between CRPS skill scores for 1–14 and 21-day lead times with CRPS skill scores at lead time 0. For each lead time, the skill score differences for all forecast locations and seasons are pooled. Simplified boxplots of the CRPS skill score differences for each lead time are plotted (Figure 78). In each plot, the boxes represent the [0.25,0.75],] and [0.10, 0.90] and [0.05, 0.95] quantile ranges with the median marked by the line crossing the box and the mean by the dot.. Outliers are ignored. The mean and median of the boxplots indicateindicates a clear monotonic trend, demonstrating that, on average, skill decreases as every 1 day forecast lead time increases. However, for individual forecast locations and seasons, the pattern of changes in skill is more variable. In some cases, the CRPS skill score can decrease by up to 18 percentage points (pp) as lead time increases up to 14 days. On the contrary, CRPS skill scores can increase by up to 10 pp as lead time increases up to 14 days. The trend of decreasing CRPS skill scores continues to 21-days lead time.

From Figure 4 and 5, it is evident that many forecast locations and seasons have low skill for reasons that may include aridity, poor catchment memory, extreme events and data problems.. We now partition the catchment and seasons into two groups based on CRPS skill scores at 0-days lead time. If the CRPS skill score for a catchment and season at 0-days lead time is > 5, the case is assigned to the skilful group. Otherwise, the case is assigned to the unskilful group. Figure 8 plots the change in forecast skill for each 1 day increase in lead time, considering only skilful cases. The reduction in skill is more marked for skilful cases compared with all cases. Figure 9 plots the change in forecast skill for each 1 day increase in lead time, considering only skilful cases. The reduction in skill is more marked for skilful cases compared with all cases. Figure 10 plots the change in forecast skill for each 1 day increase in lead time, considering only unskilful cases. The average reduction in CRPS skill

scores is near zero, with variation mainly within ±5 pp. The results in Figure 89 (considering only skilful cases) are arguably the most important for analysing reductions in skill because the Bureau issues climatology forecasts when hindcast skill is poor. From Figure 89, at 7-days lead time the mean reduction in CRPS skill scores is about 4 pp, whereas at 14-days lead time it is about 6 pp. At 21-days the reduction in CRPS skill scores averages 7.5 pp with few cases associated with skill

5   improvements.

4.3 Sharpness

Forecast sharpness is a property of the forecasts only and relates to the narrowness of the forecast ensemble spread. That is, a forecast with a narrower spread is sharper than a forecast with a wider spread. We assess how forecast sharpness changes as forecast lead time is increased. For each forecast event, the [0.1,0.9] quantile range of 1–21 day lead time forecasts is divided

10 by the [0.1,0.9] quantile range for the 0-day lead time forecast. An average is then calculated for each catchment and season, which is termed the average relative width (ARW). Simplified boxplots of the ARW for each lead time are plotted (Figure 11). In each plot, the boxes represent the [0.25,0.75] and [0.05, 0.95] quantile ranges of ARW with the median marked by the line crossing the box. Sharpness is seen to decrease gradually as forecast lead time is increased. At 7-days lead time, the forecasts are typically only marginally wider than at 0-days lead time, with the median ARW approximately 1.02. However, the forecasts

15 can be considerably wider in some cases, with the values of ARW up to 1.2 quite possible in some catchments and seasons. At 21-days lead time the median ARW increases to 1.08.

5 Discussion

The results demonstrate that forecasts are similarly reliable for all lead times from 0-14 days. Hence, from the perspective of

20 reliability, any forecast lead time is similarly optimal. Forecasts released with 7- or 14-days lead time will be similarly reliable to forecasts released with 0 days lead time. Therefore, considerations for optimal forecast lead time do not need to be based on reliability.

The results demonstrate that mean skill, averaged across all forecast locations and seasons, decreases monotonically for lead times from 0-14 days and the trend continues to 21-days lead time. However, the reduction in skill ought to be distinguished

25 for skilful catchments, since the skill for difficult to model catchments with poor skill remains largely unchanged and the Bureau's forecasts are forced to a climatological distribution when skill is very low or negative. From the perspective of accuracy skill, the monotonicity of decreasing skill scores with lead time suggests that the optimum forecast lead time is the shortest lead time that allows for other factors including the number of days required for forecast preparation, dissemination, comprehension and application.

30 As discussed in the introduction, the Bureau currently issues forecasts at least 7 days into the forecast season. Thus 7 days can be regarded as the current time needed for forecast preparation and dissemination. Much of the delay in preparing the current

operational forecasts is attributable to delays in obtaining various climate data sets and quality streamflow data. The approach taken in this study of relying solely on daily SST data for climate predictors significantly reduces the burden in preparing climate indices. Preliminary daily OISST data sets are available within 1 day, although they are subject to revision for up to 14 days. Thus the necessary climate indices can theoretically be ready for inclusion in forecasting models within 1-2 days. As

5   discussed in the introduction, quality-controlled streamflow data can take up to 3 days to enter the Bureau's forecasting database. It is therefore expected that the minimum predictor data preparation time is 3 days.

Forecast and communication strategy production is a process that takes 1-2 days. Thus for the Bureau to consider releasing forecasts prior to the beginning of the target season, it would be a safe choice to prepare forecasts with 7-day lead time. Compared with forecasts with 0-days lead time, the mean reduction in CRPS skill scores in skilful cases is approximately 4

10   pp (Figure ~~8~~9), which is likely to be tolerated by forecast users in exchange for earlier forecast release. However, it is to be reasonably expected that CRPS skill scores will reduce by up to 10–15 pp in some instances (Figure ~~8~~9). The significant reduction in skill scores at 21-days lead time (Figure ~~8~~9) highlights the importance of short-lead time forecasts for Australian catchments and confirms that simply switching to a 1-month ahead forecast system is undesirable.

A better, more flexible operational forecasting system for Australia could be built upon a flexible ~~N-days lead time forecasting~~

15   ~~strategy. An N-days lead time forecasting~~strategy that allows for any number of lead times (in days). Such a system allows for multiple forecast runs prior to forecast release. To reduce forecast preparation time, while striving for optimal skill, a communication strategy for the forecasts can be developed based on a preliminary forecast run with, for example, 7-days lead time. The final forecast release could subsequently be based on a shorter lead time, e.g. 4-days. Furthermore, it sometimes happens that unexpected heavy rains fall within the period of forecast generation and forecast release. If significant events

20   occur that change the hydrological outlooks dramatically, having the option ~~for N-days lead time~~ to reissue forecasts ~~allows for reissue of forecasts~~could benefit users.

The discussion thus far has considered only the forecast preparation and dissemination time. As identified previously, optimal forecast lead time for operational decision making depends on other factors including the time needed for comprehension and application. That is, forecast users need time to understand the forecasts and the likely impact on their operations. The process

25   of understanding may include further sophisticated modelling using streamflow forecasts as inputs. For water managers, lead times of a few days may be sufficient to assimilate the forecast information. For other operators, longer lead-time forecasts may be preferred. In fact, the optimal forecast lead time for operational decision making across a range of industries is likely to vary. It therefore remains a research questions whether water forecasting services need to evolve to cater for the needs of different (sophisticated) water forecast users.

30   **6 Conclusion**

Currently the Bureau of Meteorology releases seasonal streamflow forecasts approximately 7 days into the forecast target season. In this study we develop seasonal streamflow forecasting models with 0–14 and 21 days lead time to demonstrate that

11

it is possible to release skilful operational forecasts ahead of the commencement of the forecast period. Forecasts were produced for 23 of the Bureau of Meteorology's seasonal streamflow forecast locations, using the Bayesian joint probability modelling approach. The forecasting models were constructed similarly to the Bureau's official models, using climate predictors and initial catchment condition predictors. Climate predictors were adapted to be based on SST predictors so that daily SST data sets could be adopted.

The skill and reliability of the 0-14 day lead time forecasts were assessed. Reliability was found to be similar for all forecast lead times. Average skill reduces monotonically for each 1 day that forecast lead time increases. For forecasts with 7-days lead time, the mean reduction in CRPS skill scores is small, approximately 4 percentage points, although skill score differences within a range of +5 to -15 are possible. For forecasts with 14-days lead time, the mean reduction in CRPS skill scores is approximately 6 percentage points. In correspondence with decreasing forecast skill, forecast sharpness reduces slightly as forecast lead time is increased. The reductions in skill are very likely to be tolerated by forecast users in exchange for forecasts released ahead of the commencement of the forecast target season. Particularly as the Bureau moves towards a monthly forecasting service, timelier forecast release is going to become critical.

References

Ashok, K., Behera, S. K., Rao, S. A., Weng, H., and Yamagata, T.: El Niño Modoki and its possible teleconnection, Journal of Geophysical Research: Oceans (1978–2012), 112, 2007.

Chiew, F. H., and Siriwardena, L. W.: Probabilistic seasonal streamflow forecasting methods, 29th Hydrology and Water Resources Symposium: Water Capital, 20-23 February 2005, Rydges Lakeside, Canberra, 2005, 208,

Gneiting, T., Balabdaoui, F., and Raftery, A. E.: Probabilistic forecasts, calibration and sharpness, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69, 243-268, 2007.

Huang, B., Stone, P., Sokolov, A., and Kamenkovich, I.: Coauthors, 2015: Extended reconstructed Sea surface temperature Version 4 (ERSSTv4). Part I: upgrades and intercomparisons, J. Climate, 28, 911-930, 2015.

Kirono, D. G., Chiew, F. H., and Kent, D. M.: Identification of best predictors for forecasting seasonal rainfall and runoff in Australia, Hydrological processes, 24, 1237-1247, 2010.

Matheson, J. E., and Winkler, R. L.: Scoring rules for continuous probability distributions, Management science, 22, 1087-1096, 1976.

Mo, K. C.: Relationships between Low-Frequency Variability in the Southern Hemisphere and Sea Surface Temperature Anomalies, Journal of Climate, 13, 3599-3610, doi:10.1175/1520-0442(2000)013<3599:RBLFVI>2.0.CO;2, 2000.

Pagano, T., Garen, D., and Sorooshian, S.: Evaluation of Official Western U.S. Seasonal Water Supply Outlooks, 1922–2002, Journal of Hydrometeorology, 5, 896-909, doi:10.1175/1525-7541(2004)005<0896:EOOWUS>2.0.CO;2, 2004.

Piechota, T. C., Chiew, F. H., Dracup, J. A., and McMahon, T. A.: Seasonal streamflow forecasting in eastern Australia and the El Niño–Southern Oscillation, Water Resources Research, 34, 3035-3044, 1998.

Piechota, T. C., Chiew, F. H., Dracup, J. A., and McMahon, T. A.: Development of exceedance probability streamflow forecast, Journal of Hydrologic Engineering, 6, 20-28, 2001.

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G.: Daily high-resolution-blended analyses for sea surface temperature, Journal of Climate, 20, 5473-5496, 2007.

Robertson, D. E., and Wang, Q.: A Bayesian approach to predictor selection for seasonal streamflow forecasting, Journal of Hydrometeorology, 13, 155-171, 2012.

5 Saji, N., Goswami, B. N., Vinayachandran, P., and Yamagata, T.: A dipole mode in the tropical Indian Ocean, Nature, 401, 360-363, 1999.

Schepen, A., Wang, Q. J., and Robertson, D.: Evidence for Using Lagged Climate Indices to Forecast Australian Seasonal Rainfall, Journal of Climate, 25, 1230-1246, 10.1175/JCLI-D-11-00156.1, 2012.

Verdon, D. C., and Franks, S. W.: Indian Ocean sea surface temperature variability and winter rainfall: Eastern Australia, Water resources research, 41, 2005.

10 Wang, Q., Robertson, D., and Chiew, F.: A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites, Water Resources Research, 45, 2009.

Wang, Q., and Robertson, D.: Multisite probabilistic forecasting of seasonal flows for streams with zero value occurrences, Water Resources Research, 47, 2011.

Wang, Q. J., Shrestha, D. L., Robertson, D. E., and Pokhrel, P.: A log-sinh transformation for data normalization and variance stabilization, Water Resources Research, 48, n/a-n/a, 10.1029/2011WR010973, 2012.

15 Westra, S., Sharma, A., Brown, C., and Lall, U.: Multivariate streamflow forecasting using independent component analysis, Water Resources Research, 44, 2008.

Xu, K., Brown, C., Kwon, H.-H., Lall, U., Zhang, J., Hayashi, S., and Chen, Z.: Climate teleconnections to Yangtze river seasonal streamflow at the Three Gorges Dam, China, International Journal of Climatology, 27, 771-780, 10.1002/joc.1437, 2007.

20 Yeo, I. K., and Johnson, R. A.: A new family of power transformations to improve normality or symmetry, Biometrika, 87, 954-959, 2000.

Zhao, T., Schepen, A., and Wang, Q.: Ensemble forecasting of sub-seasonal to seasonal streamflow by a Bayesian joint probability modelling approach, Journal of Hydrology, 2016.
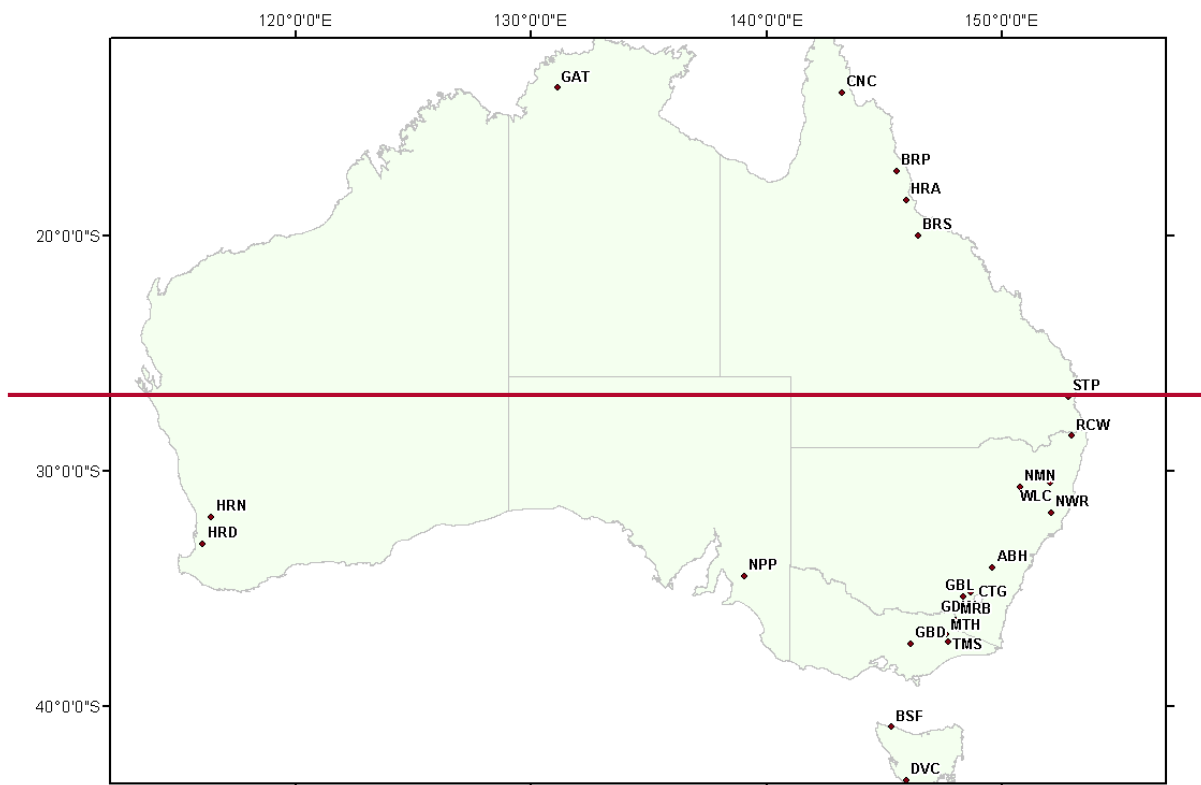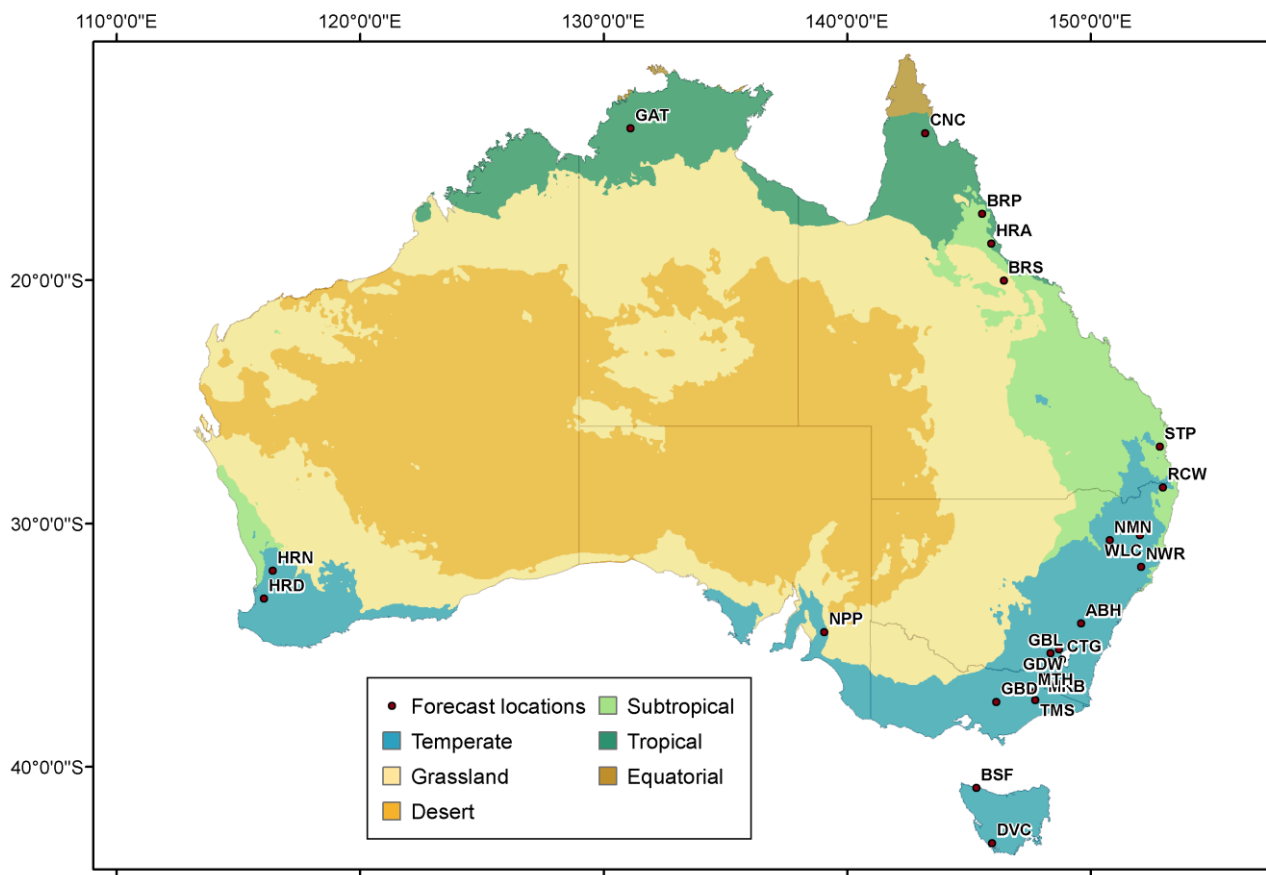
Figure 1: The 23 forecast locations ~~across Australia for which timelier forecasts are produced~~ and their climate zones. **The locations are current Bureau of Meteorology seasonal streamflow forecast locations.**
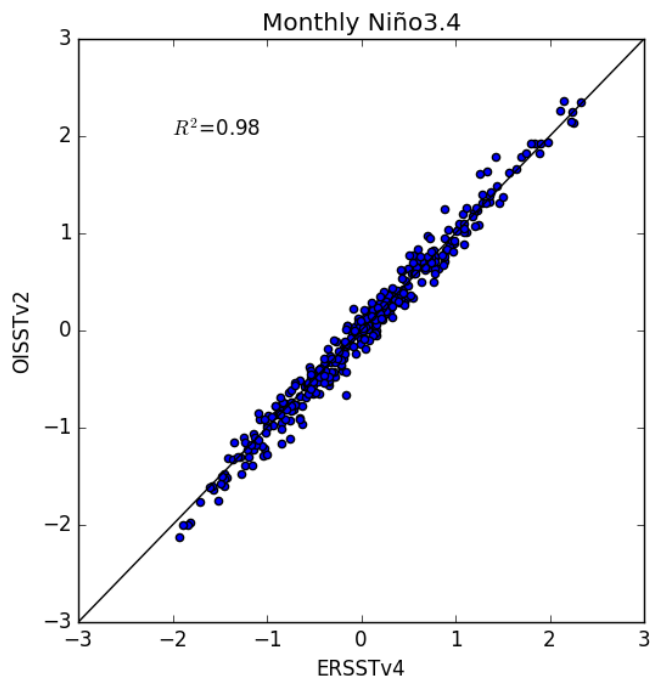
Figure 2: Scatterplot demonstrating the strong relationship between monthly Niño3.4 anomalies calculated from monthly ERSST v4 data and daily OISST v2 data.
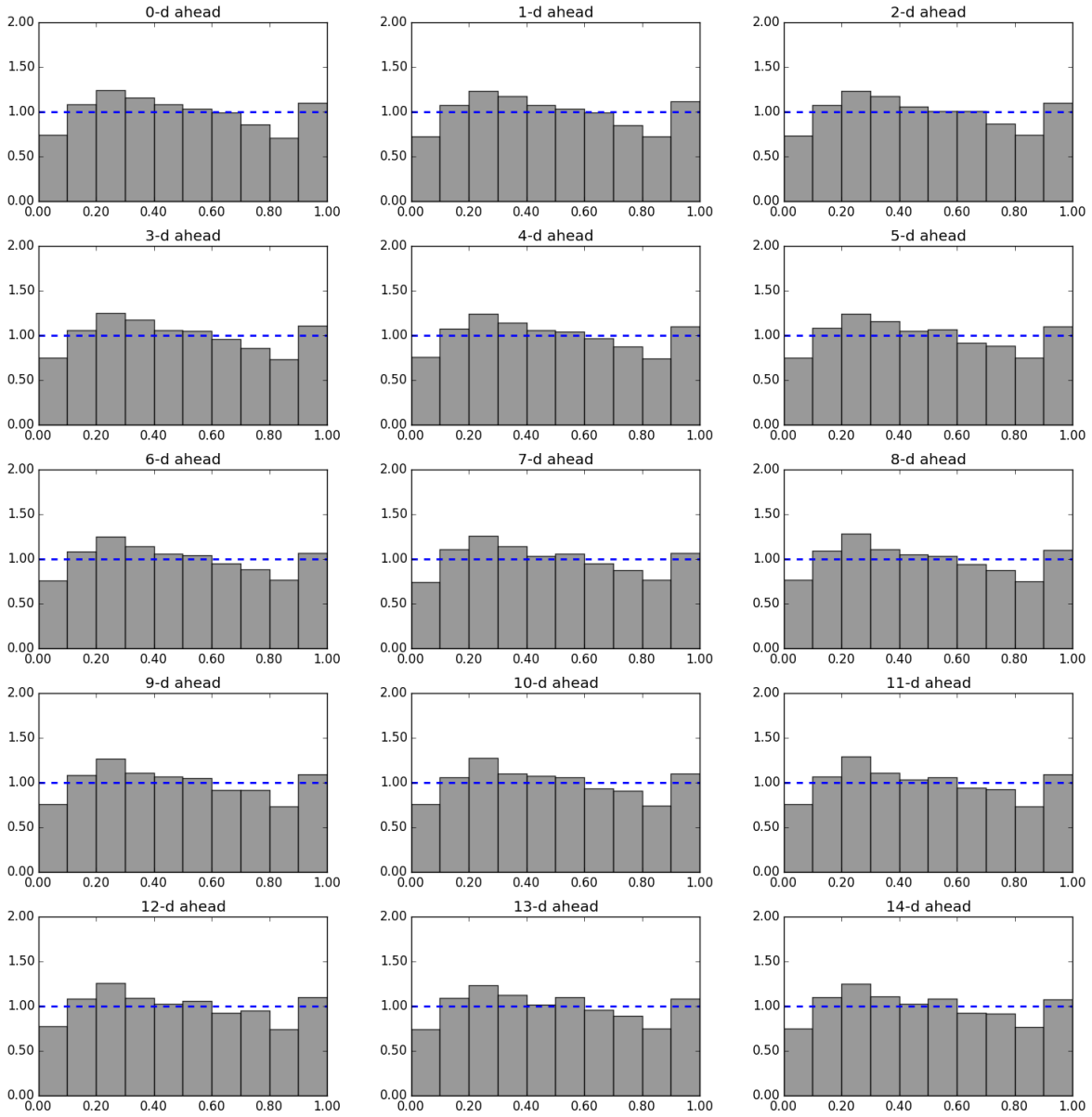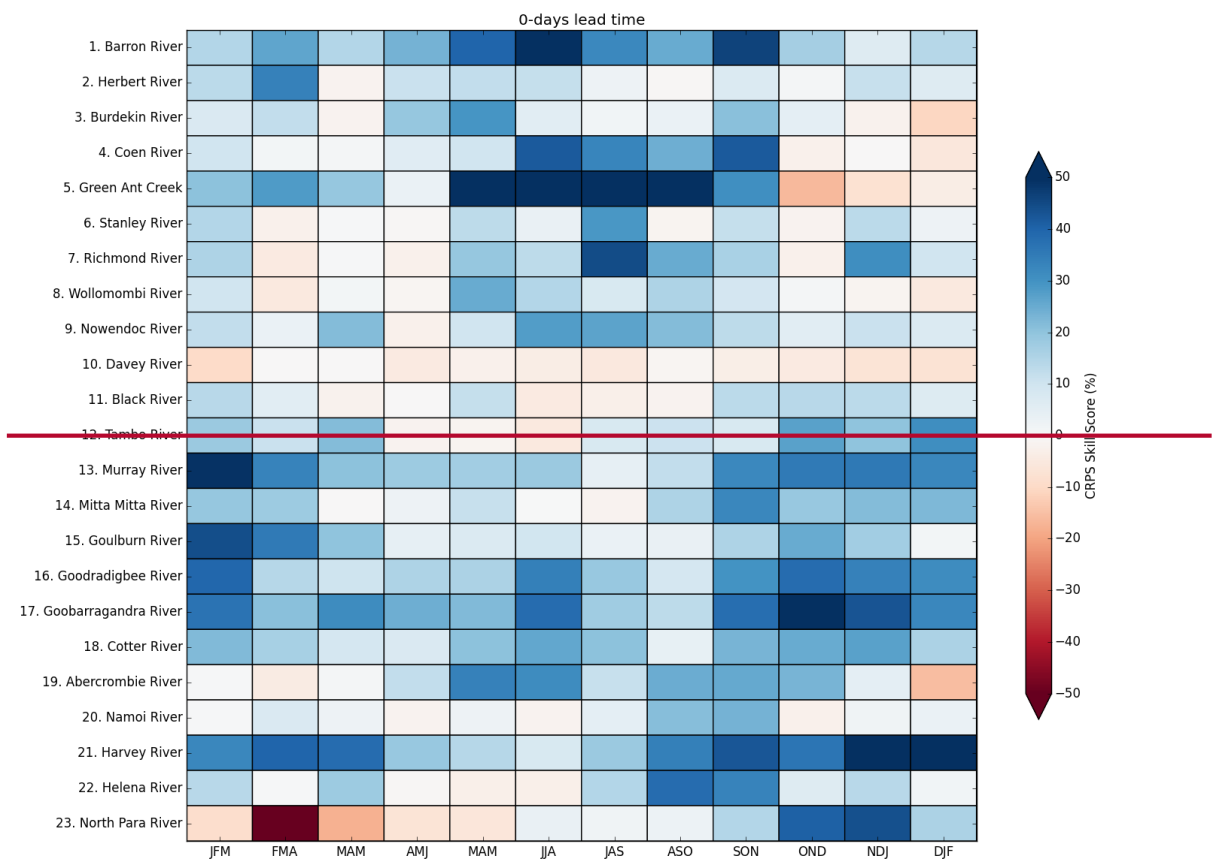
5

16

Figure 3: Figure 3: PIT histograms of seasonal streamflow forecasts for lead times ranging from 0–14 days (the x-axis of the plot represents PIT value and the y-axis normalised frequency). The blue dotted line marks the expected frequency.
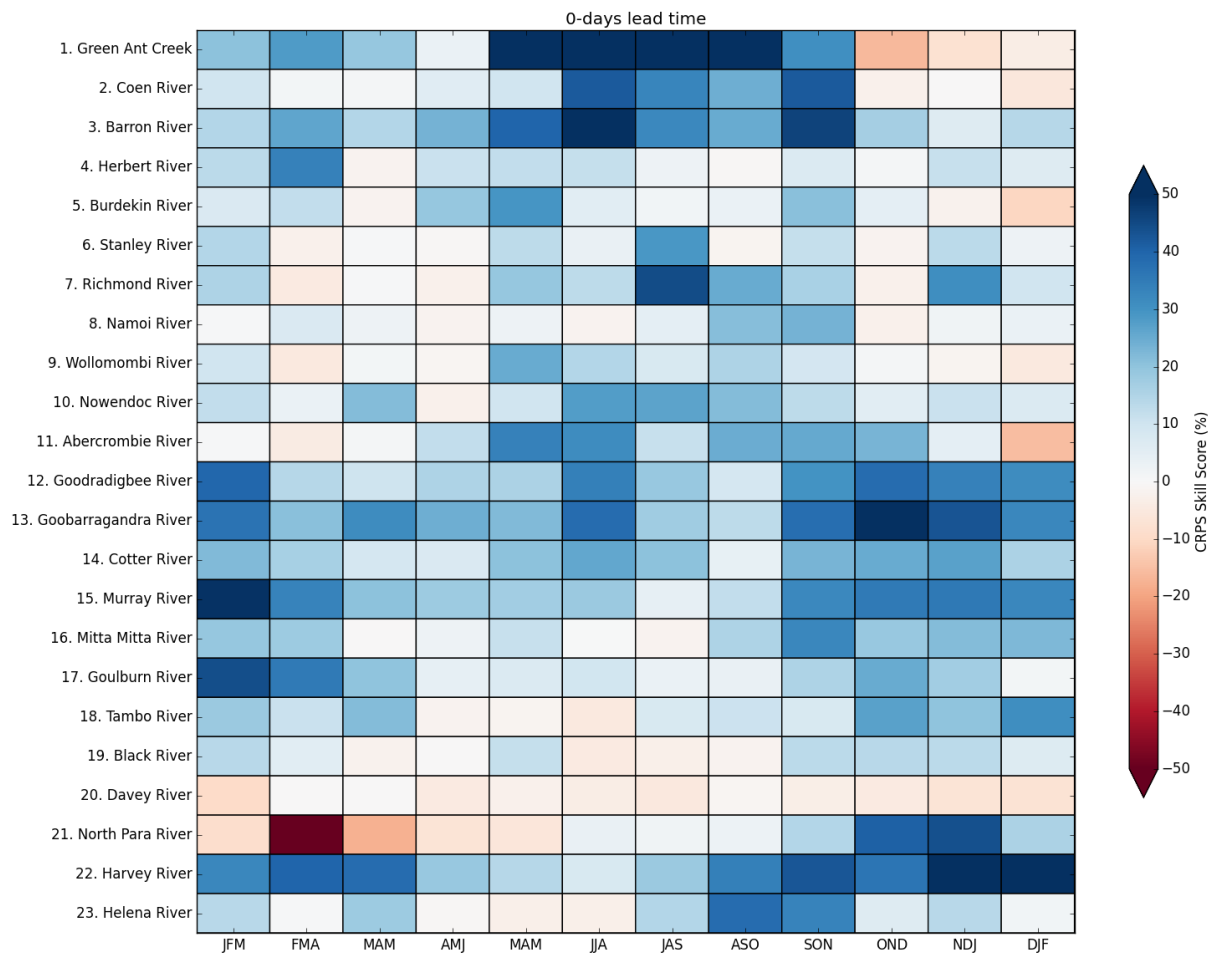
0-days lead time

**Figure 4: CRPS skill scores for each catchment and target season at 0-days lead time. Skill scores are relative to climatology. Leave-five-years-out cross validation is applied for the period 1982–2011.**
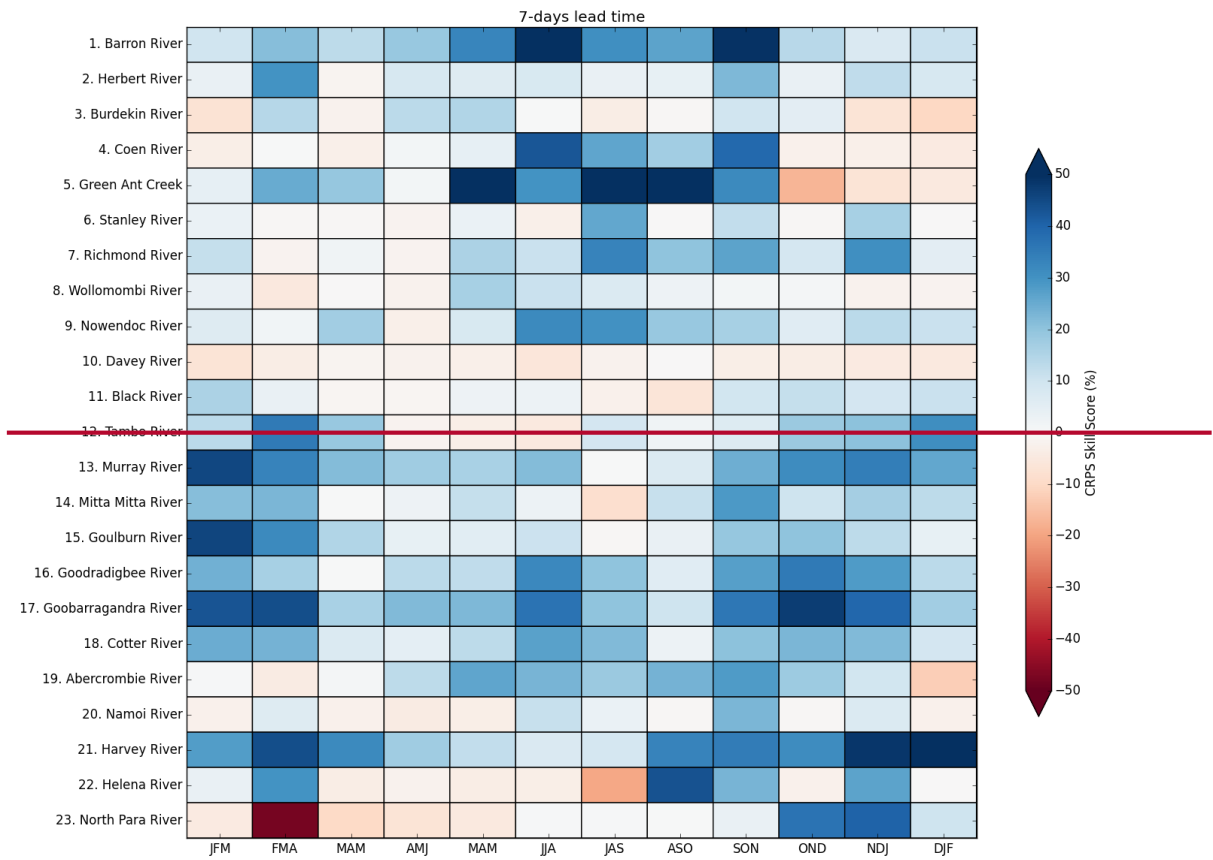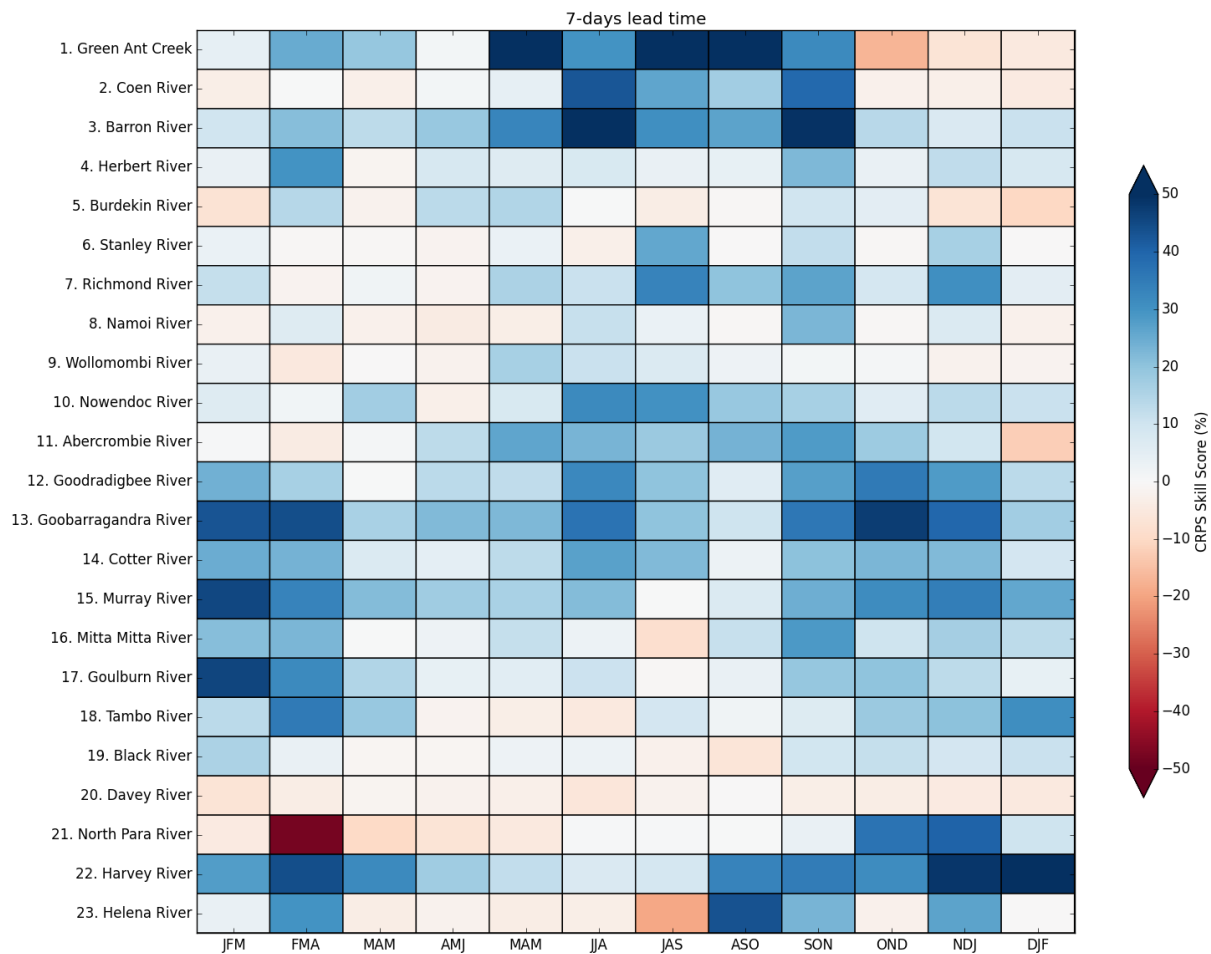
7-days lead time

Figure 5: ~~As~~CRPS skill scores **for** ~~Figure 4, but for~~ each catchment and target season at 7-days lead time. Skill scores are relative to climatology. Leave-five-years-out cross validation is applied for the period 1982–2011.
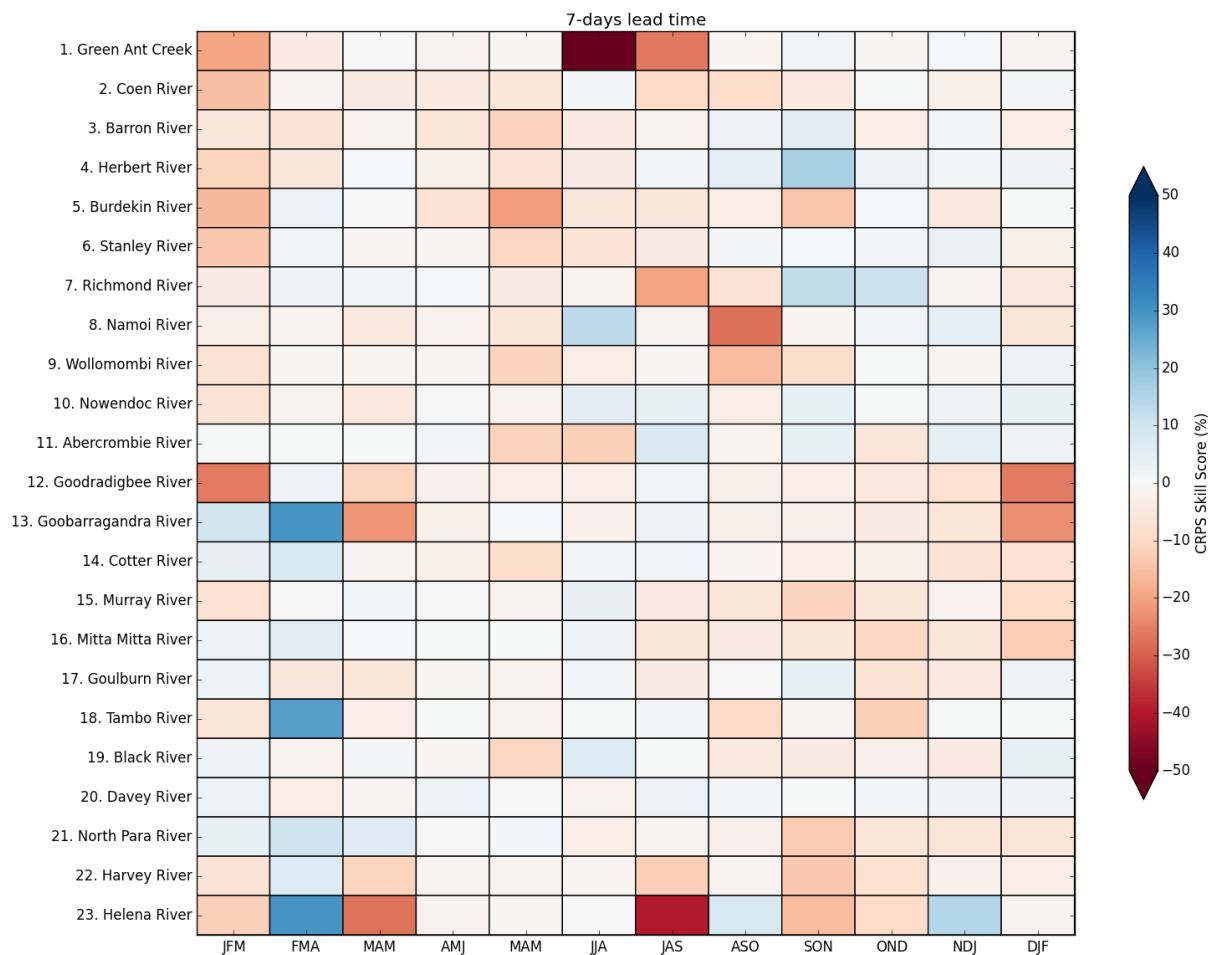
Figure 6: CRPS skill scores for each catchment and target season at 7-days lead time. Skill scores are relative to 0-day lead time forecasts.

5
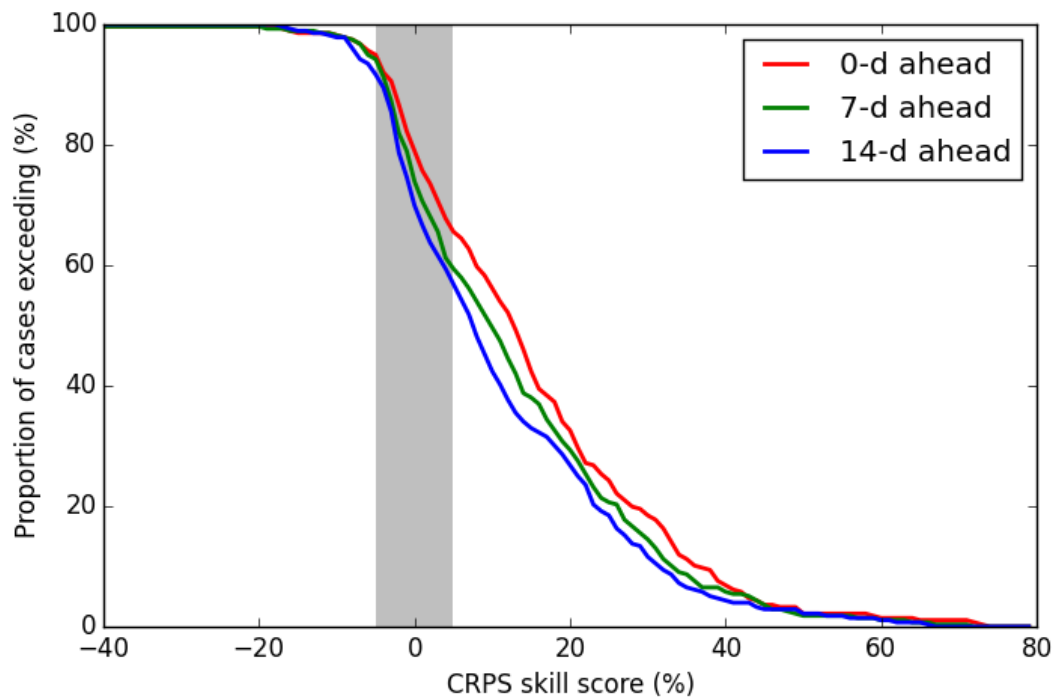
Figure 67: Proportion of cases exceeding CRPS skill scores thresholds for lead times of 0-, 7- and 14-day lead times. For each lead time, the CRPS skill scores for each forecast location and season have been pooled. The grey shaded region indicates neutral skill (±5).

5

All cases

All cases
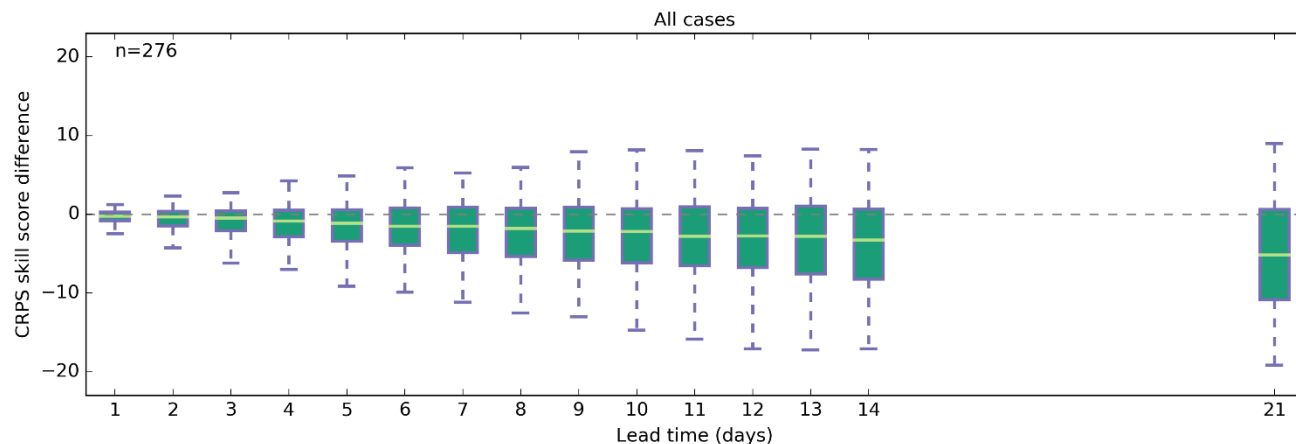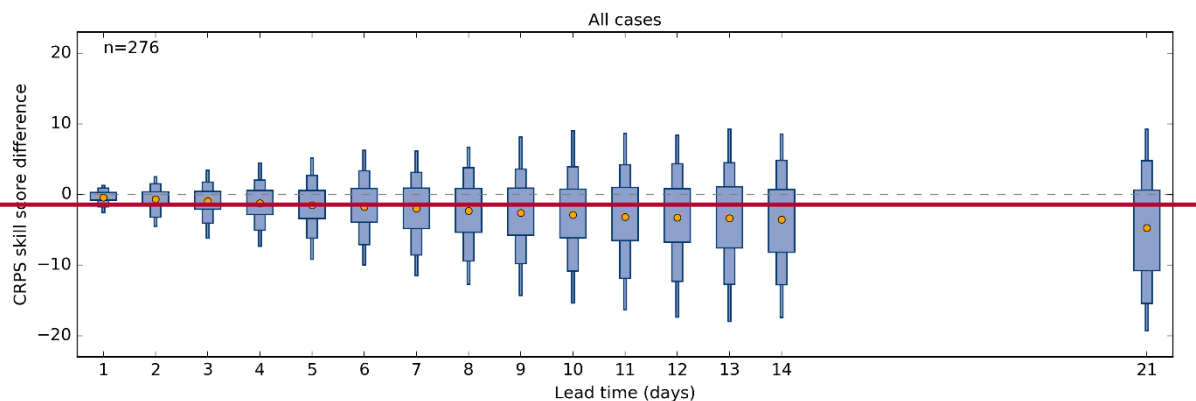
Figure 78: Box-plots of ~~the distribution of~~ CRPS skill score ~~difference~~differences between forecasts at 1–21 days lead time ~~N~~ and lead time 0. The boxes capture the median and the [0.25,0.75~~], [0.10,0.90~~] and [0.05,0.95] quantile ranges. ~~Within each [0.25,0.75] box, the dot is the average and the horizontal line is the median.~~ For each lead time, the CRPS skill score differences for every forecast location and season have been pooled. n is the number of cases per boxplot (23 locations x 12 seasons = 276 cases).
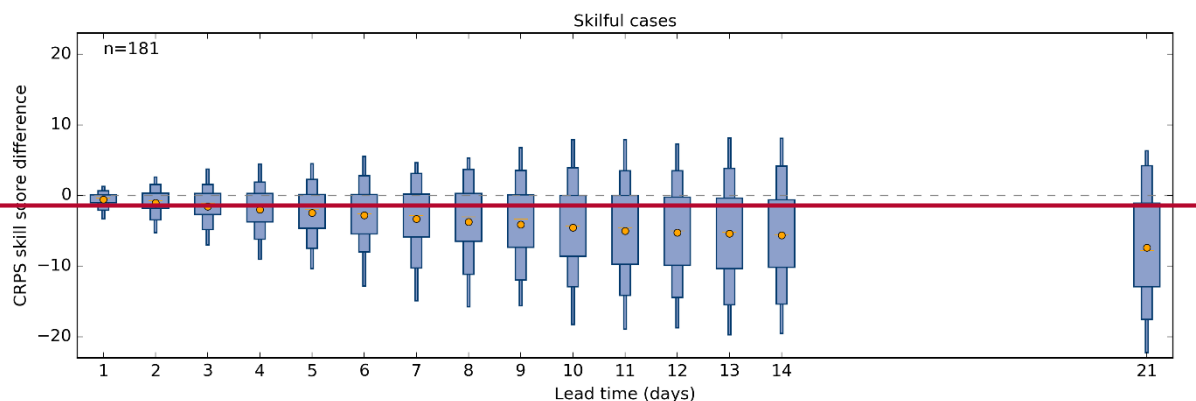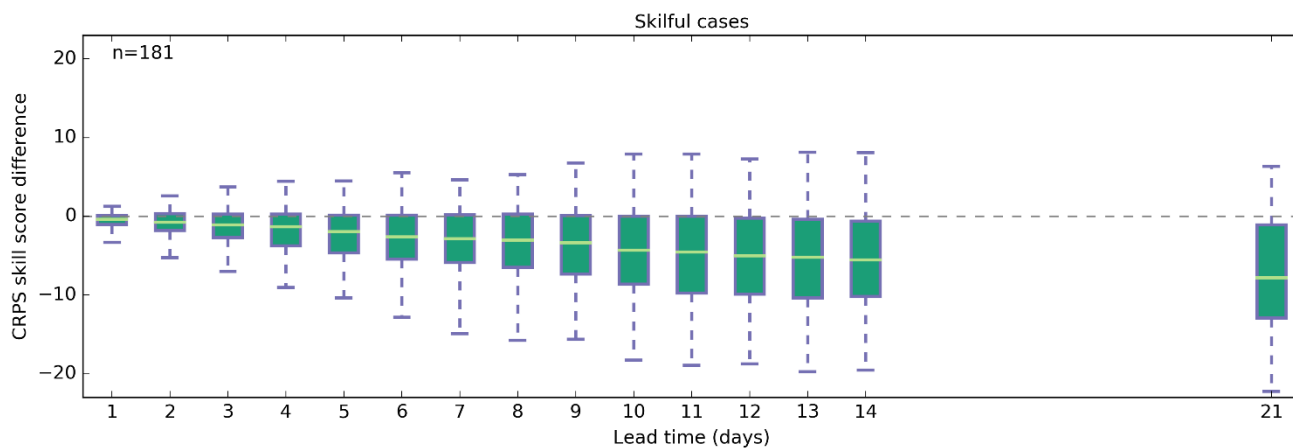

Skilful cases

Figure 89: As for Figure 78, except considering only cases where the 0-day lead time forecast is skilful (CRPS skill score > 5)
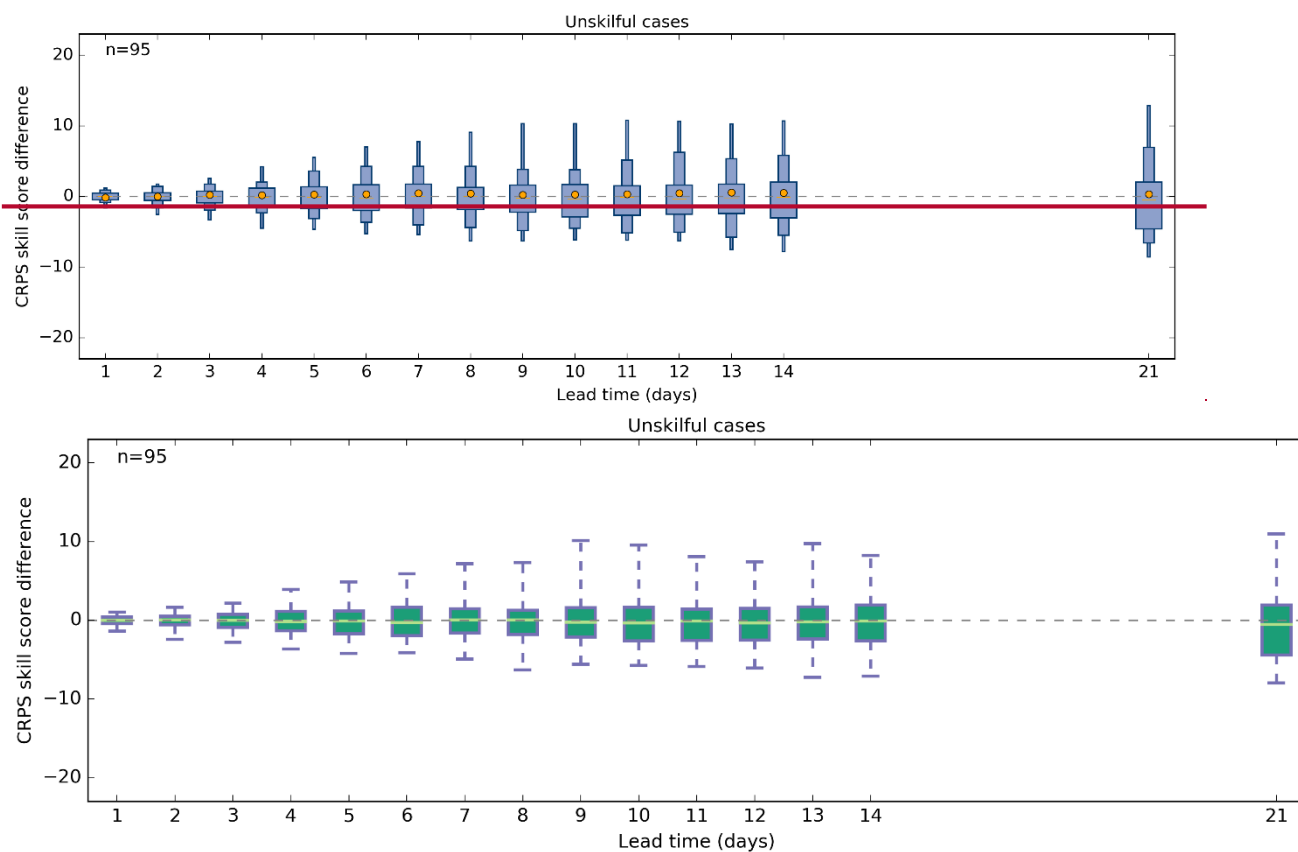

5

Figure 910: As for Figure 78, except considering only cases where the 0-day lead time forecast is not skilful (CRPS skill score <= 5)
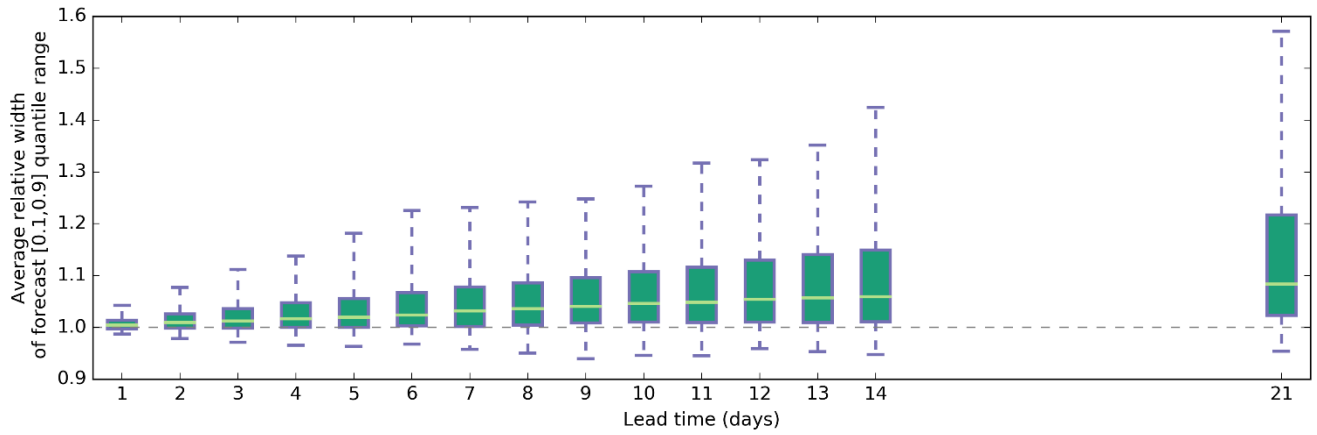
25

Figure 11: Box-plots of the average relative width of the [0.1, 0.9] forecast quantile range of forecasts at 1–21 days lead time. The relativity is to forecasts at lead time 0. The average relative width is calculated for each forecast location and season. Cases where a forecast quantile range is 0 are omitted.

5

Table 1: Catchment information for the 23 forecast locations

Catchment No.	Short ID	Long name	State	Area (km2)	Lon	Lat
G8140161~~473~~	~~DVC~~GAT	~~Davey River above D/S Crossing Rv~~Green Ant Creek at Tipperary	~~TAS~~NT	~~698~~416	~~145.95~~131.1	-~~43.14~~13.7
922101B~~14213~~	~~BSF~~CNC	~~Black~~Coen River ~~at South Forest~~above Coen Racecourse	~~TAS~~QLD	~~319~~170	~~145.3~~143.2	-~~40.87~~13.9
110003A	BRP	Barron River above Picnic Crossing	QLD	239	145.~~5~~45	-17.~~2~~63
116006B	HRA	Herbert River above Abergowrie	QLD	7486	145.~~9~~29	-18.~~4~~95
120002	BRS	Burdekin River above Sellheim	QLD	36230	146.~~4~~34	-20.~~0~~1
143303A	STP	Stanley River above Peachester	QLD	102	152.~~8~~48	-26.~~8~~48
203005	RCW	Richmond River above Wiangaree	QLD	712	~~152.97~~153	-28.~~5~~15
419005~~206014~~	~~WLC~~NMN	~~Wollomombi~~Namoi River above ~~Coninside~~North Cuerindi	~~NSW~~QLD	~~3772~~532	~~152.03~~150.8	-30.4~~8~~7
206014~~208005~~	~~NWR~~WLC	~~Nowendoc~~Wollomombi River above ~~Rocks Crossing~~Coninside	NSW	~~1893~~377	152.~~08~~	-~~31.78~~30.5
208005~~223202~~	~~TMS~~NWR	~~Tambo~~Nowendoc River above ~~Swifts Creek~~Rocks Crossing	~~VIC~~NSW	~~899~~1893	~~147.72~~152.1	-~~37.26~~31.8
412066~~401012~~	~~MRB~~ABH	~~Murray~~Abercrombie River above ~~Biggara~~Hadley No.2	NSW	~~1257~~1631	~~148.05~~149.6	-~~36.32~~34.1
410024~~401203~~	~~MTH~~GDW	~~Mitta Mitta~~Goodradigbee River above ~~Hinnomunjie~~Wee Jasper	~~VIC~~NSW	~~1518~~990	~~147.61~~148.7	-~~36.95~~35.2

410057405219	GBDGBL	GoulburnGoobarragandra River above DohertysLacmalac	VICNSW	700668	146.13148.4	-37.3335.3
410730410024	GDWCTG	GoodradigbeeCotter River above Wee Jasper (Kashmir)Gingera	NSW	990130	148.698	-35.176
401012410057	GBLMRB	GoobarragandraMurray River above LacmalacBiggara	NSW	6681257	148.351	-35.3336.3
401203410730	CTGMTH	CotterMitta Mitta River above GingeraHinnomunjie	NSWVIC	1301518	148.82147.6	-35.5937
405219412066	ABHGBD	AbercrombieGoulburn River above Hadley No.2Dohertys	NSWVIC	1631700	149.6146.1	-34.1137.3
223202419005	NMNTMS	NamoiTambo River above North CuerindiSwifts Creek	QLDVIC	2532899	150.78147.7	-30.6837.3
14213613002	HRDBSF	HarveyBlack River above Dingo Roadat South Forest	WATAS	148319.1	116.04145.3	-33.0940.9
473616013	HRNDVC	HelenaDavey River at Ngangaguringuringabove D/S Crossing Rv	WATAS	316698	116.4146	-31.9443.1
A5050517922101B	CNCNPP	CoenNorth Para River above Coen Racecourseat Penrice	QLDSA	170121	143.2139.1	-13.9434.5
613002A5050517	NPPHRD	North ParaHarvey River at Penriceabove Dingo Road	SAWA	121148	139.06116	-34.4633.1
616013G8140161	GATHRN	Green Ant CreekHelena River at TipperaryNgangaguringuring	NTWA	416316	131.1116.4	-13.7431.9