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 forecasting methods. We are very willing to enhance the introduction to include more information and references in this regard.

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 are prepared to investigate how forecast sharpness changes as forecast lead time is increased. We would measure sharpness as width of forecast intervals relative to reference forecasts (e.g. climatology or lead 0). If the results are conclusive and informative, we will be happy to include them in the paper.

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. This point will be added to the discussion.

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 will be updated to include climate zones to which runoff regimes are closely related.

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

Response: Easily done

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 will add this information to section 2.2 on streamflow data.

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

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 our comments of subsurface and atmospheric predictors are distracting and can be safely 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: Easily done