Reply to Anonymous Referee 2

We thank Referee 2 for carefully reviewing our manuscript and providing thoughtful feedback. Below we provide our responses to each point raised. We cite several instances where changes are made to the working manuscript, which will be submitted to HESS (if accepted to move forward) consistent with the review timeline.

Comment: The manuscript is limited to the study case, and as such it is not likely to help readers "gain broader insights in hydrological processes, modelling concepts, and/or improvements of existing modelling tools and methods."

Reply: We agree that the streamflow forecast component represents a location-specific contribution, however we believe this forecast coupled with the human managed allocation system is collectively novel and broadly relevant. While we address the unique set of circumstances posed by the Elqui Valley, Chile, the implications of the framework apply to basins where water rights represent a mechanism to promote equity and efficiency in the use of limited water resources.

The comment is valid, and accordingly we have made this point clearer in the discussion and conclusions. We address this in the working manuscript by describing how water rights driven basins might increase allocation efficiency by implementing forecasts as opposed to climatology based information as part of their decision framework.

Additional novelty lies in the coupling of the Stat-P&S and Stat-PCR forecast models, which provides a May 1st categorical prediction followed by a deterministic prediction on September 1st. The high level of agreement between the models suggests that while determinism is lost extending the lead from September 1st to May 1st, accurate categorical predictions are possible. We are not aware of work which has produced similar forecast skill at a May 1st lead in North Central Chile. This result also holds the potential for broader applications. Coupled forecasts need not be strictly deterministic, and using early categorical forecasts to provide an indication of expected conditions, and reinforcing the prediction with a revised deterministic forecast as more observations of local variables (e.g. precipitation) become available may be useful for a water rights holder.

Our discussion and conclusions section does address this point, but is lacking in terms of linking to a broader application. We address this in the working manuscript by bolstering the existing discussion.

Comment: The manuscript contains no clear hypothesis against which the research can be assessed.

Reply: This is a valid comment and is addressed in the working manuscript both in the introduction, discussion and conclusions by reframing the purpose of the research, namely to test if:

"…skillful streamflow forecasts can be coupled with reservoir allocation decision models to improve allocation efficiency as compared to climatology based decisions."

Additional discussion pertaining to the outcome of the research, as it pertains to addressing the hypothesis, is included in the discussion and conclusions.

Comment: Separating Section 2 into distinct 'data' and 'modelling approach' sections may make the modelling approaches more clear to the reader.

Reply: We are compelled to combine both the data and the modelling approaches in a common section as the data are not shared by each model. The Stat-PCR models are informed by observations of precipitation, soil moisture and sea surface temperatures, while the Stat-P&S model makes use of Niño 1.2 and 3.4 Index values, and the Stat-Dyn uses dynamical model outputs of both precipitation and sea surface temperatures. Rather than consistently refer to a data section when describing the modelling approach, we feel the logical approach is to introduce the data as it corresponds to the appropriate model and lead.

Still, we acknowledge the validity of this comment, noting that data and methods are presented separately in many peer reviewed papers. If the Referee feels strongly about separating the sections, we are happy to do so.

Comment: The Referee notes a lack of description of the reasons for which the modelling approaches are selected. Specifically, the Stat-P&S model, which uses Niño 1.2 and 3.4 Index values seems inappropriate, and leads are not well described.

Reply: We recognize the presentation of forecast leads for each model are not clearly described in the manuscript. To address this concern, we add additional language to the description of each model to clarify the leads considered, and provide supplemental text in the results section further reinforcing the leads and corresponding model skill. Additional information and detail are provided below (and added to the working manuscript.)

There are three distinct streamflow modelling approaches used in this research, aimed at balancing model skill and lead time. All are classified as statistical.

1) The Principal Component Analysis (Stat-PCR) model is meant to provide a deterministic prediction of streamflow using the most skillful and defensible predictors possible for increasing leads. Use of PCR is common in research focused on season-ahead streamflow prediction, and applying the leave-one-out cross-validated methodology adds additional credibility of the approach. Leads extend monthly from June 1st to September 1st. As described in the manuscript, observed data before June does not add to model skill.

2) The use of quantile mapping to correct dynamical model outputs of precipitation and sea surface temperatures (Stat-Dyn) is implemented in the same manner as the Stat-PCR. The main purpose of the Stat-Dyn approach is to increase the lead time of the streamflow predictions beyond what is possible with the Stat-PCR model. For example, the January 1st dynamical model outputs for May-August precipitation and sea surface temperatures are used to produce statistical streamflow predictions with a January 1st lead. Leads extend monthly from January 1st to June 1st.

3) The Phase and Strength model (Stat-P&S) makes use of the persistence of sea surface temperatures by using the Niño 1.2 and 3.4 indices, as opposed to other predictors which are shorter-lived or only become apparent at later leads (e.g. precipitation, soil moisture, pressure). The Stat-P&S approach is only use to provide a categorical prediction of streamflow, and ultimately proves more skillful at a May 1st lead than the Stat-PCR approach. May 1st is the only lead-time fir Stat-P&S.

The coupling of the Stat-P&S and Stat-PCR approach provides a skillful categorical streamflow prediction at May 1st (Stat-P&S), which is solidified by a September 1st deterministic prediction (Stat-PCR). The strength of the coupled model is the high degree of agreement between the two components. For all but two of the 39 years predicted by the Stat-P&S model, the Stat-PCR model provides a deterministic prediction which falls within the same category as the Stat-P&S model.

Comment: The Referee suggests producing predictions of rainfall and subsequently coupling with a physically-based runoff model as a more obvious approach to predicting streamflow.

Reply: This is a valid comment. However, while the method of precipitation runoff routing is well documented, and certainly applicable to the study area, it is perhaps unnecessary considering the correlation between May-August precipitation and October-January streamflow (Pearson's Correlation Coefficient = 0.80) suggests a strong, direct link exists. As such, predicting precipitation to inform a hydrology model is unlikely to add additional (appreciable) skill, while perhaps introducing additional uncertainty. This is further compounded by the relative lack of spatially diverse observational data and the complex topography of the upper basin. Previous research has also found a strong link between precipitation and streamflow within North Central and Central Chile (Waylen and Caviedes 1990; Verbist et al. 2010). Further, a concurrent study (performed by others) utilized the Water Evaluation and Planning (WEAP) model to address the contribution of rainfall runoff to streamflow in the Elqui basin, and found similar skill in predicting October-January streamflow.

Comment: The (Giorgi 1990) reference should be updated.

In the working manuscript, we substitute (Giorgi 1990) for (Fowler and Ekström 2009; Rauscher et al. 2010; Kendon et al. 2014) which each cite the use of general circulation models or regional climate models in seasonal and sub-seasonal precipitation forecasting at or below 20 kilometer resolution, as opposed to 600 kilometers.

Comment: Given that October-January streamflow is heavily influenced by concurrent season snow-melt, snow cover and snow depth should provide predictive strength.

The link between snow-melt and streamflow in the basins of North Central Chile is well documented (Souvignet et al. 2008; Vicuña, Garreaud, and McPhee 2011; Ribeiro et al. 2015). However, snow depth, snow cover, snow water equivalent (SWE) are not well sampled in the Elqui both spatially and temporally. The Dirección General de Aguas (DGA), the body charged with hydrologic and meteorological monitoring for Chile, provides SWE for a single location (La Laguna) for the period 1976-2005, which includes significant data gaps. The correlation between May-August SWE and October-January streamflow (Pearson's Correlation Coefficient = 0.67) is not as strong as the correlation between May-August precipitation and October-January streamflow (see above), and arguably provides the same information to the model. As such, we retain precipitation observations as a predictor in the Stat-PCR model. We agree with the Referee that snow observations would seemingly be an obvious predictor of streamflow, and have explored this thoroughly, however in this case, for the reasons mentioned above, it explains less of the overall variance in streamflow as compared with precipitation. We discuss this explicitly in the manuscript and have further highlighted this point in the working version.

Comment: The allocation model description is unclear. Specifically, the end of year (February) target volume seems too high, and the requirement to carry storage shortfall to the next year implies the reservoir has does not replenish. Is this realistic?

While we agree, for a variety of reasons, that a static target volume is generally a suboptimal operational policy, the target volume (50% of maximum storage) is the operating rule enacted by Puclaro's operators as a response to critically low reservoir levels (< 20 MCM) observed during the recent extended hydrologic drought (2009-2014). It was not the purpose of this research to address the performance of existing reservoir operating policies. Rather, we evaluate the performance of the October-January streamflow and climatological forecasts, translated to per-water right allocation values using the reservoir allocation model, against perfect foresight as a means of assessing the value of the forecast. A concurrent project is aimed at optimizing Puclaro's operational policies.

The purpose of carrying the storage shortfall or surplus from February and using it as a constraint or benefit for the subsequent October-January per-water right allocation value is that it recognizes the storage target as non-binding (can be violated by over or under allocation in the previous year), but consequential, in the allocation model. As such, the reasonable place to impose the effect is in terms of the following year's allocation. Effectively, it represents a mechanism for the reservoir operator to compensate for over or under allocation in the previous year. In addition, reservoir replenishment from snow melt typically does not occur until December, which is three months after the allocation issuance date of September 1st, and such is the reason why a forecast is produced. The use of the deficit or surplus becomes a hedge against the uncertainty of the forecast. Ultimately, we believe the carryover of deficit or surplus is an appropriate way to include the operational goal of the reservoir operators.

Comment: The summary and discussion are uninformative. As such, they should be split into distinct "Discussion" and "Conclusions" sections and significant effort applied to drawing more scientific conclusions from the research.

We agree with the Referee's comment to break the "Summary and Discussion" section into and "Discussion" and "Conclusions" sections, and have done this in the working manuscript. The Discussion section now more thoroughly describes where models are both successful and limited in terms of prediction, and how the limitations (e.g. tradeoff between lead and skill) of the models we construct align with previous research. The Conclusions section establishes the broader insights gained from the research, including the potential for improved water right allocation efficiency achieved by coupling hydroclimate streamflow prediction with a reservoir allocation framework which may benefit both reservoir operators and water rights holders. In addition, the Conclusion presents the coupled Stat-P&S and Stat-PCR models as achieving both increase in forecast lead while maintaining skill, by adjusting the type of forecast provided (categorical to deterministic). The broader insight gained here is that by sacrificing forecast precision, the lead can be skillfully extended. We hypothesize this information to be of potential value to water rights holders who must make decisions (e.g. cropping) prior to the annual setting of the per-water right allocation value.

Comment: The manuscript should be shortened. In depth descriptions of well understood methods and metrics may be removed.

We recognize that the interdisciplinary nature of this manuscript may draw readers who have limited methodological knowledge of hydroclimate prediction and reservoir allocation forecasts. Therefore, we provide explicit detail of both methods and metrics used to construct and evaluate models, respectively. The comment is reasonable, and in the working manuscript we have removed all but necessary discussion of principal component analysis, multiple linear regression, cross-validation, and metrics.