Review for Mundo, I.A., Masiokas, M.H., Villalba, R., Morales, M.S., Neukom, R., Le Quesne, C., Urrutia, R.B. and Lara A. Multi-century tree-ring based reconstruction of the Neuquén River streamflow, northern Patagonia, Argentina. Clim. Past Discuss., 7, 3541-3575, 2011

Reviewers' comments in black
Authors' responses and comments in blue

C. Woodhouse's comments

This paper takes advantage of the excellent network of moisture-sensitive tree-ring chronologies in Chile and Argentina. A number of the authors of this paper have built on and greatly expanded the work of earlier dendrochronologists to develop a set of chronologies from long-lived trees that are being used to extend records of past climate and hydrology back in time. This paper features the latest in a series of reconstructed streamflow records which are providing potentially useful insight for water resource management in the region. The focus of this paper is the reconstruction of the Neuquén River, a major waterway in Argentina with headwaters in the Andes. This reconstruction is the longest flow record for South America, to date, extending to AD 1346.
The authors make use of a suite of statistical methods to decompose the large set of tree-ring chronologies, enhance the common climate signal, and extend the reconstruction back in time as far as possible, give the signal strength of the trees. The reconstruction is then analyzed to investigate drought, spectral characteristics, and relationships of flow to large scale circulation. The paper is well written and methods are clearly documented. Results are reported in a straightforward way, supported with tables and graphics. The disappointing part of this paper is the very minimal discussion of results, beyond the description of the reconstruction, and periods of drought.

Following C. Woodhouse comments, in the new version we present an extended and more comprehensive discussion of the results based on the issues raised by the reviewer.

How are the ranking 20th century droughts and pluvial related to atmospheric/oceanic circulation (only 1998 is mentioned)?

In our previous version, information on the 20th century droughts was centered on the 1998 event. In the discussion of the revised manuscript we added information on the atmospheric/oceanic anomalies related to other events of the 20th century such as the emblematic droughts of 1924 and 1968 as well the 1914 pluvial.

What is the significance of the 6-7 yr and 17.6-yr periodicities; to what might they be related?

Although information on both periodicities was briefly mentioned in the previous Discussion section, in the new version we included the following paragraph dealing specifically with periodicities in the reconstruction:

"Oscillations on the order of 17 years have also been documented in temperature and precipitation reconstructions for northern Patagonia (Villalba et al., 1996; 1998). In the Neuquén River streamflow reconstruction, this periodicity is more evident from 1600 to around 1890 (Fig. 6). This pattern was already detected in the Puelo River reconstruction (Lara et al., 2008) and it is also a common pattern to several proxy records across the Pacific basin, suggesting that interdecadal variations in this region were probably more conspicuous prior to the 20th century (Villalba et al. 2001). The shortest mode of variability in the reconstruction (3.8 - 3.4 years) may correspond to the El Niño–Southern Oscillation recurrence cycle."

How do the SAM, AAO, and ENSO interact in this region to influence water supply? How is this manifested in this and other reconstructions?

Although significant relationships between large-scale circulation indices and the Neuquén river streamflow have been reported, there are not previous studies dealing specifically with interactions between SAM and ENSO on the climate of the region. To add information on this particular issue, we have included the following paragraphs in the Discussion section:

"Previous studies have reported significant relationships between ENSO and mean annual discharges of the Neuquén River (Scarpati et al. 2001; Seoane et al. 2005). However, the comparison of the prewhitened versions of the ENSO indices and the October-June Neuquén River flow (both observed and reconstructed records) shows positive but not statistically significant correlations between these variables. This might be related to a weakened influence of ENSO over the Andean portion of the Neuquén basin (Garreaud et al. 2009), and the transitional location of the study area between the ENSO-dominated region of Central Chile to the north and the SAM-dominated region across Patagonia to the south."

"Recent studies of *Araucaria araucana* fire regimes in northern Patagonia show that large fires are associated with positive SAM departures concurrent with La Niña events in the tropical Pacific (Mundo 2011). Dry springs were also related to the combination of positive anomalies in SAM and La Niña conditions in the Pacific Ocean. Fogt and Bromwich (2006) and Fogt et al. (2011) have shown that the magnitude of ENSO teleconnections over South America intensifies when positive SAM anomalies occur in phase with La Niña events. Although there was a slight tendency of low streamflow during La Niña years and pluvials coincident with El Niño events along the 20th century, no clear relationships were recorded between SAM and ENSO interactions and the reconstructed Neuquén River discharge."

How does this understanding provide information useful for water resource management?

The following paragraph was added to the Discussion section:

"The streamflow reconstruction described in this paper might be used as input in water system models to assess the reliability of water supply systems under a broader range of conditions than those afforded by the gage record alone. Collaborative work with water managers should be encouraged to make streamflow reconstructions compatible with current planning and management tools. Several issues that might be investigated are (1) the fidelity of the tree-ring data in reflecting the severe drought conditions in particular years (e.g. 1924, 1968, 1998); (2) the uncertainty in the reconstructions and their incorporation into water management strategies and modeling; and (3) the feasibility of reconstructing other hydroclimatic metrics that are also critical to water managers in the region."

Specific comments

P. 3547, lines 20-21. While using the common interval, 1800-1950, to extract tree-ring chronology PCs does utilized the period of highest replication, it also assumes that grouping of chronologies over this time period is representative of the full record.

We select the common interval 1800-1950 of highest replication to identify the dominant patterns of tree growth. We note that using longer common intervals produce some awkward grouping patterns that reflect the low replication in the early part of some chronologies rather than the common signal between records. The use of longer periods produce less reliable grouping pattern as they include low sample replication during earlier centuries in some chronologies.

P. 3458, line 7. Define the ARSTAN chronology

The concept of "ARSTAN" chronology was added in the methodology section.

P. 3553, line 3. It is not clear why the beta weights are given just for Model 2. Why not the others? It would be interesting to see if the weights changed between the different models (at least among the shared chronologies).

Beta weights are presented only for model 2 because it represents the longest reconstructed period with the highest percentage of total explained variance in the streamflow.

P. 3553, Frequency, intensity, and duration of events: this analysis only uses n-year running means, which only addresses the intensity component of drought. With the relatively low explained variance, the most robust information is duration and frequency. This section needs more analysis to characterize these features of drought.

Following the reviewer's comments, additional information on the frequency and duration of drought events was incorporated in the Result and Discussion sections:

"Analyses of the frequency of single- and multiple-year droughts show an expected even decline from a peak in numbers of single year droughts to fewer, increasingly longer droughts (Fig. 4). The longest dry period was centered in the 1888-1897 period. In terms of frequency, the mean drought interval (MDI) ranges between 15 years for single events to 175 years for 8-year droughts.
The extreme dry years (< 10th percentile) are not evenly distributed through the complete reconstruction period and were slightly more frequent during the 19th and 20th centuries rather than in the previous four centuries (Fig 5). Of particular note is the clustering of extreme dry event years 1810-1820s, 1860s, 1910s and at the end of 1960s with three consecutive years of extreme drought in 1818–1820, 1863-1865 and 1967-1969. Extremes are less frequent and more evenly distributed in the 15th, 16th, 17th and 18th centuries."

P. 3555, lines 3-4. The similarity is arguable, but if the patterns are described more generally in terms of zonal patterns, this might be OK.

Agreed. This issue was clarified in the text.

P. 3556, line 5. I think you mean 43.1 to 48.2% (the range for the three reconstruction models)?

Corrected.

P. 3556. The lowest reconstructed flow year is 1968, but it is the lowest value in the gage record. What is the rank of 1968 in the gage record?

The flow registered in 1968 is the sixth lowest in the gauge record. This information was added in the discussion.

P. 3556, line 23-24 (also on P 3558), what is the relationship between the Antarctic Oscillation mentioned here and the SAM?

The Southern Annular Mode (SAM) is also referred to as the Antarctic Oscillation (AAO). This was clarified in the methodology.

P. 3558, line 8. It is not clear why QBO is mentioned here, as the periodicity is not apparent in the reconstruction.

Following the reviewer's advice, the reference to QBO was deleted.

P. 3558, line 14. No analyses were shown that indicate that droughts lasted 25 years (a running average does not establish duration).

Agreed, corrected in the new version. We added a histogram that groups droughts according to their duration in years (new figure 4).

Figure 3c: The uncertainty based on RMSE in the early part of the record looks narrower than from 1700-1900, but the variance explained in the earlier period is less. Is this correct?

As it was noted by the Reviewer, the three nested reconstructions have slight differences in the RMSE and not totally consistent. In order to clarify this issue, in the revised version we used the ±2SE for establishing the uncertainty band. The ±2SE uncertainty band is also widely used in dendrochronological reconstructions. Although there are small differences in the uncertainty band over time in the new version, the last period (1346-1486) shows the highest uncertainties.

Figure 4a: Is this needed?

The Figure 4 was removed and a drought duration histogram was included instead (see above).

Figure 7a: Please add a legend or indicate which series is black and which is red.

Corrected. New Figure 7 caption:

"Fig. 7. (a) Relationship between the October-June Neuquén River streamflow (black) and annual SAM index (red) for the period 1887-2000. SAM index has been inverted to facilitate the comparison with the streamflow data. (b) 21-year central moving correlations between the streamflow and the annual SAM index. Dashed lines indicate the 95% significance level. Note the change in strength of this relationship over time.''

GTP Pederson's comments

General Comments:
The manuscript entitled "Multi-century tree-ring reconstruction of Neuquén River streamflow, northern Patagonia, Argentina" by Mundo et al. provides a high-quality, long-term dataset useful for water management planning and investigating drivers of trends and variability in a hydrologically important Argentinean river system. The paper builds substantially on previous work within the region, providing the longest reconstruction of streamflow in South America, and adding an impressive network of new chronologies while utilizing collections extending back to Holmes and LaMarche's early work. The authors employ a chronology screening criteria and a nested PCA reconstruction methodology with the intention of enhancing the common climate/flow signal, and extending the reconstructions as far back as possible. The methodological approach is sound, and the interpretation of results is supported by the data – though the authors might consider adding a more quantitative ranking of drought event severity, magnitude, and duration (see Biondi et al. 2002, 2005, 2008). Relationships between streamflow variability and atmospheric circulation are explored using indices and geopotential heights. The documented association between Neuquén River flow and the Southern Annular Mode is striking, and speaks to the fidelity of the climatic signal (precipitation) captured by the

reconstruction. Overall, the paper is well written and methodologically robust with high-quality figures. Due to the importance of this river system for regional water resources and ecosystems I expect this paper will be of broad interest within the scientific and resource management communities, and is suitable for publication within Climate of the Past with only minor revisions.
Though the manuscript is generally well written, the discussion of results presented within the paper could be substantially improved before official publication. The discussion seemed to lack a strong flow and focus, which would likely become more apparent with the potential addition of a summary figure (or two), and perhaps an additional minor analysis. For example, the start of the discussion related your new streamflow reconstruction to the early work of Holmes and LaMarche along with other regional reconstructions of streamflow and precipitation. I'm not sure if any of this data is available (it's not listed on NOAA paleoclimate), but it would be helpful for the reader if there were a stacked or compilation figure of the discussed reconstructions. If space is limiting, figures 4 and 5 could be moved to the supplemental.

The reconstruction of the annual Neuquén River streamflow by Holmes et al. (1979) is not totally independent from our reconstruction. The seven chronologies (five *Araucaria* and two *Austrocedrus* records) used by Holmes et al. (1979) to develop their reconstruction were also included in the set of predictors in our study. Additionally, the early work of Holmes et al. (1979) is based on a different calibration season: January-December (Holmes et al. 1979) vs. October-June (our reconstruction). In consequence, we decided not to include a figure comparing both reconstructions. Similarly, more recent developed streamflow reconstructions in northern Patagonia, such as the Puelo River by Lara et al. (2008) and the Maule River by Urrutia et al. (2011), are partially based on some of the chronologies used in our study. Consequently, most of the flow reconstructions in the Andes of northern Patagonia are not entirely independent and similarities between them may reflect common predictors in the reconstructions.

Also, a more quantitative ranking of drought/pluvial magnitude, intensity, and duration (e.g. Biondi et al. 2002, 2005, 2008) may better contextualize the importance of a reconstructed or observed event, and aide in the discussion of the hydrological importance of the event. This may help contextualize a discussion on the relevance of the SAM (or ENSO) in forcing major drought or pluvial events.

Following both reviewers' advices, additional analyses of duration and frequency of droughts were included in the new version of our manuscript. Please, see the new figures 4 and 5 which show a histogram of drought frequency by drought length and the duration of droughts using a shading-coded percentile classification of the reconstructed flow by year, respectively.

Also, how are the SAM, ENSO, AAO all potentially related to the dominant frequencies of the reconstruction that are shown with SSA and the Blackman-Tuckey spectral analyses? Does their influence on streamflow modulate through time? Is the same behavior shown in other regional reconstructions (e.g. the comparison figure)? And, can we infer more about the regional expression of the climate forcings?

No significant, decadal-scale cycles were identified in the Power Spectrum (Blackman-Tukey) analysis of the surface pressure-based SAM record of Marshall (2003), probably due to the shortness of this series. The longer (1887-2005) reconstructed SAM series from Visbeck (2009) showed a cycle of ~46 years which was also detected in the singular spectrum analysis -SSA- of the Neuquén river reconstruction (see Fig. 6). However, a coherency analysis between the Visbeck SAM series and the Neuquén river reconstruction shows just a significant peak at the 2.5 yrs. Therefore, it is difficult to identify statistically significant similarities in the periodicities of SAM index and the Neuquén River streamflow. Still weaker relationships were recorded between the Neuquén River reconstruction and different ENSO indexes. In most cases, there are not significantly correlations between ENSO indices and the Neuquén streamflow.

Focusing on improving the summary figures, statistics, and discussion may greatly improve the overall flow and focus of the paper, which may help convey the major points to a water management audience.

As requested by both reviewers, a paragraph indicating the potential use of the information here presented by water resource managers in the northern Patagonian Andes was included in the Discussion section. Please see above.

References
2008 Biondi, F., T.J. Kozubowski, A.K. Panorska, and L. Saito. A new stochastic model of episode peak and duration for eco-hydro-climatic applications. Ecological Modelling 211: 383–395.
2005 Biondi, F., T.J. Kozubowski, and A.K. Panorska. A new model for quantifying climate episodes. International Journal of Climatology 25(9): 1253–1264.
2002 Biondi, F., T.J. Kozubowski, and A.K. Panorska. Stochastic modeling of regime shifts. Climate Research 23: 23–30.

Specific Comments:
Abstract and throughout. Comparison of the 20th century condition to conditions of the past millennium is made. The major reconstruction presented here is nearly 800 years long. Are you referring to longer, unpublished results you have that include the last 1000+ years? Consider contextualizing the results within the past ~800 years instead of 1000, unless citations to ongoing or unpublished work are made.

Considering that our reconstructions spans for 654 years, we agree with the Reviewer that is not correct to refer our work to the past millennium. We have modified the abstract and the text accordingly to reflect this point.

Pg. 3545 ln. 13-16. Consider citing the method used to classify/quantify the duration, magnitude, and intensity of drought/pluvial events, or, add a follow-up sentence describing how a drought event is defined.

The following paragraph was included in the Methods section:
"A drought was simply defined as a year or set of consecutive years with streamflows below the long-term mean. The frequency of droughts of 1-20 yrs in length was calculated together with the mean drought interval (MDI, the number of years between two consecutive droughts of the same length). We also analyzed the distribution of extreme dry years (< 10th percentile) over the reconstructed period."

Pg. 3546 ln. 12-14. Can this sentence be clarified? It's unclear exactly what is meant by "The average monthly streamflow was evaluated with tree-ring data to determine the most appropriate months or seasons to develop the streamflow reconstruction."

Following the reviewer's comment, the sentence was rephrased as follows:
"Correlations between all possible monthly streamflow averages and tree-ring data were evaluated in order to determine the most appropriate season to develop the streamflow reconstruction."

Pg. 3548 ln. 7-8. What specifically are ARSTAN chronologies? Are you referring to the standardization program or the version of chronology used here? Specifically, did you use the standard (std), residual (resid) or the arstan (ars, residual chronology with pooled autocorrelation added back in) version of the chronology? What's plotted in Figure S1 looks like either the std or ars chronology but this should be clarified along with a simple justification (e.g. retention of low-frequency climate information) for using either the std or ars versions of the chronologies.

In the revised version, we clearly stated that the ARSTAN version of the chronologies was used in the reconstruction. In addition, the definition of the ARSTAN chronology (residual chronology with pooled autocorrelation added back in) was added in the methodology.

Pg. 3550. The defined method for ranking intensity, duration, and magnitude characteristic of a drought is insufficient. The rankings only appear based on what would classically be defined as the intensity of the drought. See the above Biondi et al. references for a formal statistical test and ranking procedure related to these three parameters of a drought and consider applying this method in your final analysis.

As noted above, discussion of results regarding ranking intensity, duration and frequency were included in the revised version of this manuscript. Please see the previous responses to reviewers' comments and the new figures 4 and 5 in the revised manuscript.

Also, you have adjusted the mean and the variance of your nested reconstructions to match the "best" reconstruction. The "best" reconstruction, however, has slightly different statistical properties (particularly variance) than the instrumental record. Do you see any problems with making direct comparisons to droughts in the reconstructed record to droughts in the instrumental record? Some caveats should be placed on the discussion comparisons made between droughts in the reconstructed vs. observed streamflow record.

We normalized the mean and variance of each nested series to those of the most replicated period following D'Arrigo et al. (2011). We are aware of the slightly different statistical properties (particularly variance) in the instrumental *vs.* the reconstructed streamflow series. An important caveat in the reconstruction is the reduced capability to reconstruct extreme wet events, particularly those preceded by a long-term period of dry events. Persistence is a common feature in tree-ring series that prevents a rapid recovery of growth after a period of adverse weather conditions. This limitation is clearly acknowledged in the revised version as indicated in the following paragraphs: "Pluvial events of the late 20th century rank among the wettest in the reconstruction, but in some cases the intensity of these events was underestimated by the regression model (see e.g. the year 1914). Therefore, our reconstruction should be considered conservative in the representation of extreme wet events. This limitation is a common feature in hydroclimatic reconstructions from tree rings and it has previously been identified by several authors (Maxwell et al., 2011, Masiokas et al., 2012)"

Pg. 3551 Ln. 13-15. The method of pre-whitening then correlating two autocorrelated time series together is one approach to estimating accurate p-values. Prewhitening largely destroys low-frequency information in both time series, which could have the effect of artificially inflating of deflating the actual r value. In the future, one way around this potential problem is to simply calculate the p-value using the effective degrees of freedom(N) in the time series with autocorrelation. For example, see www.atmos.umd.edu/~ekalnay/notes4a.pdf.

Thanks to the Reviewer for this useful piece of advice. However, pre-whitening of the time series prior to correlation is also a robust approach to account for the serial persistence seen in the original time series (e.g. Dawdy and Matalas, 1964).

Pg. 3552 Ln. 7. Table 1 should be cited here, correct?

Corrected, please see the revised version of the manuscript.

Pg. 3553. This section on drought ranking needs more work (see above). Pointing out the major intervals discussed here on the reconstruction in figure 3 might also be helpful.

As previously explained, the discussion of this section was improved and two new figures (4 and 5) were incorporated in the manuscript.

Pg. 3554. Consider listing correlation coefficients and p-values in the text alongside mentioning their significance. Any expectations of what may have altered correlation strength between the SAM and streamflow through time? Do you get more or less the same results if non-prewhitened time series are used? Consider discussing/exploring potential modulating roles of ENSO or the AAO.

Agreed. The correlation coefficients and p-values were included in section 3.5.
A brief discussion on the changes in correlations between SAM and the Neuquén River streamflow over time is included in the new version:

"Changes in the relationships between the reconstructed Neuquén River streamflow and the SAM over time might be associated with the persistent trend in the SAM toward its positive phase since late 1950's. Observational analyses of sea-level pressure on the Southern Hemisphere have shown a statistically significant increase in the difference in zonal mean sea level pressure (MSLP) between 40ºS (increasing) and 65ºS (decreasing), with the trend being most pronounced since the mid-1960s (Marshall, 2003). Changes in the 1960s reflect the marked shift southward of the subtropical band in the Southern Hemisphere (see e.g., van Loon et al. 1993; Hurrell and van Loon 1994). These changes in atmospheric circulation may have introduced changes in the strength of the relationship between SAM and the Neuquén River streamflow."

Pg. 3555. Composite maps during years of insignificant correlation with the SAM could be used here to explore if atmospheric structure resembling other known major modes of circulation (e.g. ENSO, AAO) are important when the SAM is not. If this is a transitional basin between major atmospheric modes of circulation (as you mention in the following discussion), you may be able to elicit both when and where other circulation patterns have a large influence on the Neuquén River flow.

We agree with the Reviewer that this is an important issue to explore. However, it is beyond the scope of this contribution. The main purpose of this study is to develop a multi-century reconstruction of the Neuquén River streamflow based on a well replicated tree-ring network in Central Chile and Northern Patagonia. Composite maps to elucidate the influence of SAM versus other climate forcings on the Neuquén River variability will be included in a future contributions specifically dealing with climate influences on the Neuquén River streamflow.

Pg 3556 Ln. 23-27. The discussion of ENSO and the AAO here seems to come up unexpectedly since their potential role in driving drought and streamflow variability is downplayed in the results section. I like this discussion, but it also justifies spending a bit more effort investigating the potential influence of these circulation modes.

Lines 23-27 in page 3556 emphasize the influence of SAM, and its interaction with ENSO, on the severe drought of 1998. In addition, a detailed discussion dealing with SAM influences on the Neuquén River streamflow is included afterward. The sequence of topics in the Discussion section follows the same order of the topics presented in the Results section.

Pg. 3558 Ln. 7-9. The discussion on the QBO seems to appear unnecessarily here. Within the paper its potential influence wasn't investigated so it may be advisable to delete the sentence.

As previously indicated, this sentence was deleted.

Pg. 3558 Paragraph 3. The discussion on the influence of the SAM and the AAO should be more clearly articulated. As the paragraph is currently structured it hops from one index to the next without clearly explaining the relevant influence of either. It may be helpful to diagram an idealistic schema of the circulation patterns associated high and low streamflow and precipitation across the region and then focus the discussion within the text. For example, I provided a similar diagram for the Northern U.S. Rocky Mountains in a recent paper, which was extremely helpful for structuring the text. See: Pederson, Gregory T., Stephen T. Gray, Toby Ault, Wendy Marsh, Daniel B. Fagre, Andrew G. Bunn, Connie A. Woodhouse, Lisa J. Graumlich, 2011. Climatic Controls on the Snowmelt Hydrology of the Northern Rocky Mountains. J. Climate, 24, 1666–1687. doi: 10.1175/2010JCLI3729.1

We agree with the Reviewer that the proposed topic is very interesting, but we also believe that is beyond the scope of our present contribution. In the new version we added, in the discussion section, information on the SAM influence on the Neuquén River, on the variations in the relationships between SAM and the streamflow over time and on the interactions between SAM with ENSO. We find very interesting the diagram presented in Pederson et al. (2011). However, to develop an idealistic schema relating atmospheric circulation patterns with the Neuquén River streamflow will require a significant (some patterns of atmospheric circulation in South America still are under discussion) and a long-term effort. As previously indicated, a study to elucidate the influence of atmospheric circulation features on the Neuquén River streamflow will be conducted in the near future.

Table 2. Explain in the table caption why certain reconstructions are shown in bold typeface. It won't be clear to readers who skim the methods section that these are the reconstructions you retained and used.

Corrected.

Figure 1. Nice maps and an impressive array of chronologies.

Figure 3. Why are the calibration and residual panels smaller and offset to the right of the rest of the figure? Enlarge these to match the rest of the graphic, or make the graphic a 2-column graphic with the calibration interval on the left. If more space is needed, the CE and RE statistics for the split calibration and verification intervals could be dropped from the graphic.

The calibration and residual panels were presented smaller and offset of the figure to illustrate that they cover a period of time different from that cover by the reconstruction. In the new version, it was enlarged to match the size of the figure.

Also, double-check your code for the mean and the variance scaling of the reconstruction and reconstruction error. The error bars appear to get smaller over the earliest part of the reconstruction when they should be increasing substantially.

Please see the reply to a similar comment by reviewer #1.

Figure 4 and 5. Consider moving to the supplemental if more space is needed to add a nice summary figure or two that supports the discussion. Nice figures though.

Please see comments above.

Figure 7a. The instrumental SAM and streamflow records need a key or to have their colors defined in the legend.

Corrected. The caption of Figure 7 was modified to identify both records.