Response to comments by reviewer #5

This paper provides useful insights into the various potential inputs to streamflow by comparing physical and chemical approaches to baseflow separation. It reinforces that both quickflow and baseflow can have multiple sources and that the contributions from these sources can vary significantly over the flow regime. Specific comments mainly relate to apparent inconsistencies and additional clarification:

We thank this reviewer for their considered comments and are encouraged that they found the approach that we took to be useful. We have considered the specific comments and have provided responses and clarifications below.

(1) The baseflow definition in the Introduction assumes that baseflow is only derived from delayed storages in the unsaturated and saturated zone, whilst other storages may theoretically be involved (such as snow melt or return flows from connected lakes etc). This leads to an inconsistency, as the example of draining of floodplain pools is given later in the paragraph.

We agree that the discussion of baseflow should be more consistent throughout the paper. In the literature, there are varying uses of the term. Some studies use baseflow as being synonymous with groundwater inflows (which is certainly an over simplification of the water balance in the rivers as is evident from this and other papers). More commonly, the term baseflow is used to refer to water from the unsaturated and saturated zone, but we agree with the reviewer that all the delayed sources of water should be grouped into the term and we will make this more explicit in our initial definitions in the revised paper.

(2) The literature review of the comparison of chemical and physical baseflow separation methods could really be expanded. A cursory search came up with some additional references that may be relevant– see below.

As with all papers, the number of references used to substantiate individual comments will always be limited (and there are multiple choices for which references to use). It was not our intention to provide an extensive review of the literature (which would be an interesting study in itself) but to provide a few examples of studies where there have been comparisons of physical and chemical baseflow. We will review the studies that we quote in this regard and thank the reviewer for pointing out the Sanford et al. 2012 USGS report of which we were unaware.

(3) The salinity of the return flows from bank storage is assumed to be similar to surface water which may not be the case. Such return flows may involve mobilisation of salt stores in the near-river geological profile. The geological description of the catchment suggests that there are examples of such salt stores in low-lying areas of the floodplain.

This is possible in some catchments but is not likely in the Barwon Catchment. As described in the paper, there are saline water stores on the flood plain (these are the marshes and wetlands described in Section 2). However, there is little evidence of salt stores in the near-river aquifers. Although not mentioned in the paper, the Cl/Br ratios of the groundwater and the river water in the Barwon catchment are approximately those of rainfall, which indicates that neither the groundwater nor the surface water has dissolved halite. This is the case for all catchments in southeast Australia for which there is Cl/Br data. The variations in salinity of groundwater and

surface water in these catchments is driven by evapotranspiration, which is a relatively slow process in the unsaturated zone. Thus there is no mechanism of increasing the overall salinity of water in the bank on the timescales of a few weeks to months over which bank exchange occurs. The bank return water may have mixed with regional groundwater; however, this would result in the chemical mass balance recording the inflow of that groundwater. The Cl/Br ratios were discussed by Cartwright et al. (2013. Applied Geochemistry, 31, 187-198). This is an important point that certainly needs to be clarified and we propose to add a brief explanation to Sections 2 (Local Geology and Hydrogeology) and 5 where the results are discussed.

(4) The description of local geology and hydrogeology, although cursory, suggests that the catchment is hydrogeologically complex, consisting of a combination of sedimentary sequences, volcanics with interbedded sediments, as well as alluvial deposits. There is no summary of the hydraulic connection of these different units to the river and to each other). There is the potential for multiple groundwater systems to be operating at different scales (eg regional, intermediate and local flow systems) that could be contributing baseflow of varying magnitude over different time scales. At a minimum, there needs to be a simplified hydrogeological map of the catchment as well as schematic cross sections that summarise the different scenarios for groundwater surface water interaction in the catchment.

The Barwon Catchment is less complex than perhaps was conveyed in Section 2. The majority of the upper catchment comprises Quaternary alluvial sequences overlying the basalts and these are the units with which the Barwon River interacts. These units are separated from the deeper regional Otway Basin aquifers by low hydraulic conductivity layers and there is little evidence of upward mixing from the deeper aquifers into the shallower units. We can readily provide a few more details of this in Section 2 and modify Fig.1 to provide more of a regional geological and hydrogeological context, including a cross-section. We agree that this would add to the paper.

(5) The chemical mass balance approach is based on a constant groundwater salinity. This is difficult to reconcile considering (i) the hydrogeological complexity and the huge variation in groundwater salinity observed in the catchment and (ii) much of the shallow groundwater being 3500-13000 mg/L (P 5949) and so significantly above the 3200 EC (2100 mg/L) threshold used. There needs to be a discussion about what the 3200 EC figure actually represents – is this an aggregate of all the various groundwater inputs upstream of the gauge, does it reflect a particularly dominant groundwater input in a particular reach, does it reflect a degree of freshening of shallow groundwater near the river, etc.

In practice in this and other catchments, it is difficult to be certain of the net EC of groundwater that actually interacts with the river (as opposed to groundwater that is a few tens of metres away) without a density of monitoring bores that is unrealised in many, if any catchments. As with other studies (e.g. Yu et al., 1999; Gonzales et al., 2009) we have used the EC of the river at baseflow to estimate the EC of the groundwater component (Section 4.3). Since the data is for a single gauge, it does represent an estimate of the average EC of groundwater up catchment of that gauge. This is justified on the following grounds.

1. **It is not possible that the net EC is lower than this otherwise the calculated fluxes are negative at baseflow times.**

2. **It is possible to assign a higher EC to the groundwater component but this would have the effect that rover at low flows would always have a considerable component of surface water (which is unlikely during the prolonged very low flows).**
3. **In most years the highest EC in the river water during the low flows in the summer is similar. This is consistent with the inference that the river is groundwater fed at this time and that the river EC represents the groundwater EC than it is of the groundwater having a higher EC and the river being fed by a similar mixture of groundwater and surface water.**
4. **In terms of the calculations, assigning the groundwater EC to the maximum EC recorded in the river gives the maximum baseflow flux as calculated by Eq. (3). Nonetheless, this is still considerably lower than the baseflow fluxes calculated by the physical methods. Thus, we have been conservative in our approach and in our comparisons.**

Some of this is discussed in Section 4.3, but we can provide additional justification of this point. In particular is important to note that while there is some saline groundwater in the catchment, the average TDS is much less than 13,000 mg/L (EC <20,000 μS/cm); again this will be calcified.

(6) The flow duration curves for the representative years (Figure 5) appear to have some inconsistencies. The 2002 year was chosen to represent the long-term median, but is the one that has the most persistent low flows (<10 ML/d). The 2001 year is the high discharge example, as represented by the greater occurrence of flows >100 ML/d. However, at 100 ML/d the flow duration curve changes abruptly, and low flows are more typical of normal conditions. These features in the low-flow part of the flow duration curves need to be explained. In contrast, the 2006 year appears to be a good representation of a low discharge year.

We are not sure that these are inconsistencies (2002 was close to the long-term median and 2001 was a high discharge year). The behaviour at low flows relates to the rainfall distribution over summer rather than the overall annual rainfall. In a year of high rainfall where much of that rain falls in autumn and winter following a dry summer, the Barwon River will record higher than average discharge for part of the year but there may be considerable periods of low flow over the summer months. We can explain this in the description of the flow duration curves.

(7) The sentence summarising the annual discharge at Winchelsea (p 5950, line 21-22) does not match Table 1. Perhaps all flow/baseflow figures referred to in text need to be checked.

It should be 5.4×10^5 ML. Other figures are correct.

(8) River EC values > 3500 uS/cm have been interpreted to reflect evaporation during stagnant conditions (page), however the discharge of saline groundwater cannot be ruled out.

This is probably true, although it is not clear why the salinity of the groundwater discharging to the river would increase only at this time and not during any other baseflow period. We will explain this briefly in the revised text.

(9) The surface runoff EC is assumed to be the same as the local rainfall. However, the surface runoff may be significantly higher, considering the extent of near-surface salt stores that is apparent in the catchment.

We do discuss this in the paper. Firstly in regard to the hysteresis loops (Section 5.1) we note that the EC values are higher on the rising limbs of the hydrographs due to the flushing of saline water from the marshes and wetlands which results in clockwise EC vs. discharge hysteresis loops. Secondly, in Section 4.3 we discuss using a higher value of surface water EC in the calculations (100 μS/cm). The discussion at the start of section 5.1 is, however, a little brief and could be amplified. In terms of what the paper discusses, raising the assumed EC of the surface water lowers the estimates of baseflow from the chemical mass balance technique and thus our assumptions again produce conservative estimates of the mismatch between the chemical and physical techniques. This is noted at the end of section 4.3, but we can reiterate this important point in the Conclusions.

(10) I found it difficult to differentiate the time series of the various baseflow curves (Figures 2-4). It would also be useful to have time series of the ratios of the physical baseflow estimates to the chemical mass balance baseflow – especially to reinforce some of the points made in the discussion.

Part c on Figs 2-4 are the baseflow ratios derived from the two methods. We will ensure that this is clear in the text and figure captions.

(11) The differences between EC predicted from the physical baseflow methods and the actual EC record for discharge events (Figure 7) need further discussion. There are differences between events – for some the calculated EC values return to high values on the falling limb (such as events 3 and 4) whilst for others the falling limb EC values are relatively low (such as events 1 and 2) – what is the significance of this?

Figure 7 shows the predicted EC vs. discharge if the baseflow estimated by the physical methods reflected groundwater inputs (i.e. it used the baseflow flux). The specific shape of the predicted hysteresis loops depends on a number of factors such as: 1) the calculated baseflow during the period before the discharge event; 2) the maximum discharge during the event; and 3) the rate at which discharge subsides after the peak of the event. Some of the loops are "incomplete" because a second discharge event occurs before the river has returned to low flow conditions. Notwithstanding that the predicted hysteresis loops have slightly different shapes; they are all much steeper than the observed EC vs. discharge loops and that indicates that there must be transient sources of water feeding the river on the rising (saline water flushed from the floodplain) and falling (bank return waters) limbs. This latter point is the most important as far as the aims of this study are concerned, and we will ensure that it is clear.

(12) The physical baseflow estimates have input parameters that are largely empirical or subjective (eg N, BFI-max). The study would benefit from a sensitivity analysis to derive the potential range of baseflow estimates when low and high estimates of these input parameters are applied. These ranges should then be compared with ranges of baseflow derived from a similar sensitivity analysis undertaken for the chemical mass balance approach (by having low and high estimates of surface runoff EC and groundwater EC). This would be a better representation of the level of uncertainty embedded in these two methods.

This is a similar comment to that made by other reviewers. While we agree with the sentiment, in practice it is difficult. For the Eckhardt filter we do vary BFI_{max} and discuss this in section 4.2. The Lyne and Hollick filter has no variable parameters aside from a (which is measured and which exerts very little influence on the results: as discussed by Nathan & McMahon, 1991 and Eckhardt,

2005). The graphical variation technique is also an empirical relationship that when used in other studies seems not to be varied. Varying N in this technique will make a difference to the calculated baseflow estimates. For example increasing 2N to 9 changes baseflow estimates by <5% (the value of 2N is used to define a window in which the turning points of the hydrograph are defined, but changing the size of the window by one or two units does not change many of the values of discharge that are used as the local minima values). Given this, it is difficult to assign an uncertainty to these techniques in the same manner that can be done for chemical mass balance (other studies seem not to have addressed this either). We consider that this is an important point that we will address generally in Section 5.

This reviewer raised a number of important points, the addressing of which will be valuable in clarifying the manuscript. While it was not always explicitly stated, we took a conservative approach in this study. Many of the issues raised above (e.g. the possibility that groundwater was more saline that expected or that surface runoff was more saline than rainfall) increases the mismatch between the chemical and physical estimates of mass balance making it even more likely that there are transient stores of water. This adds weight to our conclusions in this study.