Review of HESSD-11-6441-2014

The manuscript 'HESSD-11-6441-2014' details a field and modelling study that aims to explain the energy processes driving observed instantaneous longitudinal negative water temperature gradients in a semi-forested stream reach downstream of an open area. In contrast to previous studies, the authors conclude that advection due to groundwater discharge and hyporheic exchange are not needed to produce negative temperature gradients. Instead, negative temperature gradients can be explained by water parcel travel times along the reach and the associated advection of relatively cooler water from upstream to downstream.

This study addresses an important stream temperature management topic and employs a good use of field and modelling approaches. The primary novelty of this study is that the authors put an emphasis on the role of longitudinal advection to explain instantaneous negative temperature gradients between an upstream open site and a downstream forested site. As Dr Westhoff's review has outlined, previous diagnostic water temperature model studies have included longitudinal advection, although that has not been the emphasis of those studies because longitudinal advection alone could not explain downstream cooling. Although the conclusions of this study are reasonable, revisions are needed on what results are presented and how those results are presented to support the conclusions. My primary comments concern evaluation and uncertainty of the water temperature and net radiation models, providing more convincing evidence on the role (or lack thereof) of advection associated with groundwater and hyporheic exchange, and improving the figures so that readers can better interpret the results.

I would also like to note that I am in agreement with Dr Westhoff's review. Therefore, I have tried to not duplicate any of his comments here.

General comments

Water temperature model

The water temperature model is critical to this study in order to establish that longitudinal advection explains the negative temperature gradients observed at the site. The water temperature model relies on field measurements and other models and estimates (flow routing, net radiation, turbulent energy exchanges) that are all associated with errors and uncertainties, yet the manuscript does not address the issue of model uncertainty at all. The only presented evaluation of the model is from page 6456, lines 2–3, 'predictions of downstream water temperature change were typically good' and Figure 6. Examining Figure 6, there appears to be periods when the model over- and under-predicts water temperature by 1 to 2 °C at the downstream locations, mostly during the clear sky days. This error is of similar magnitude to the observed negative temperature gradient signal that this study is trying to explain. Are these errors due to uncertainties in the net radiation model (see below), not including groundwater or hyporheic advection (see below), discharge errors, or uncertainties in the flow routing model (how uncertain are the travel times?). In order to have confidence in the conclusions of this study, a more robust evaluation of the model is needed. Some report of the error statistics (e.g. RMSE) for the model would be helpful. One suggestion could be evaluating the downstream predictions for all time periods during the study week (not just the four times per day examined here) and plotting the model residuals against time. This would be valuable in determining how prevalent the prediction errors are and whether they are systematic (and associated with misrepresentation of a certain process) or noise.

Net radiation model

How uncertain are the estimates of modelled net radiation? Was the net radiation model evaluated against observed net radiation at the site? What threshold value was applied to the hemispherical images to convert them to binary images within Gap Light Analyzer? How was this threshold selected? I have some concern over the selected threshold, considering Figure 4a has some noticeable riparian vegetation in the top left-hand corner, but was classified as having 0.0% canopy density. In addition, how representative are the hemispherical images of lateral variations in canopy cover structure since photographs were only taken in the centre of the stream? Do canopy, terrain, and bank shading vary laterally across the stream?

I agree with Dr Westhoff, in that I do not understand why a smoothing procedure was used on the canopy density and the energy fluxes at the stream surface. The water temperature model can be coded to include the spatiotemporal heterogeneity, and the smoothing procedure creates undesirable artifacts such as the negative canopy density for the first few downstream metres (Figure 5a) and interpolating between distinct riparian vegetation conditions.

Groundwater/hyporheic exchange

The authors position the findings from this study, that longitudinal advection drives observed negative temperature gradients, as an alternative explanation to cooling caused by groundwater discharge and hyporheic exchange, citing the work by Brown et al. (1971) and Story et al. (2003). I am surprised that only minimal efforts to characterize hyporheic exchange and groundwater discharge were made to reject this competing explanation. The authors conducted differential streamflow gauging and cite previous research by Malcolm et al. (2005). However, differential streamflow gauging on its own is known to have limitations on characterizing groundwater and surface water interactions (e.g. Payn et al. 2009), particularly for water temperature modelling (Leach and Moore, 2011). The authors cite Malcolm et al. (2005) to justify that groundwater discharge and hyporheic exchange is minimal in the study reach. However, upon a quick review of that article, it was difficult to confirm whether the study reach here is influenced by hyporheic exchange, since some of the sites which appear to be located in the study reach of this paper (numbered 13-15 in the Malcolm et al. (2005) study) did appear to have distinct surface water and hyporheic water qualities (particularly for DO). Also, I cannot find mention of hydraulic gradient measurements or downwelling patterns in Malcolm et al. (2005) as mentioned on page 6445, lines 18–19. In fairness, I did give Malcolm et al. (2005) only a cursory read; therefore, if more convincing evidence is provided in that paper I would recommend that it be specified and elaborated on in this manuscript. Of course, the water temperature modelling provides a means to evaluate whether groundwater and hyporheic energy exchange processes influence the thermal regime; however, since there may be some errors in the modelling (see comment above - although it is difficult to tell from the the limited model evaluation), it would be important to elaborate on these topics in the discussion.

Figures and data visualization

I think the choice of what data is displayed in figures and the visualization approaches used inhibit full interpretation of the study results. In particular, the 3D plots are aesthetically pleasing; however, I feel that they fail to communicate the rich dataset and modelling results produced by this study. In Figures 3a, 3c-d, and 7a-c, it is difficult to read the absolute temperatures from the figures, and for Figures 3a and 7a, most of the rising and falling diurnal periods of the signal are hidden behind the diurnal maximums. For Figures 5b-d, it is difficult to tell when the net energy fluxes are above or below zero. I suggest more 2D heat plots (aka image or raster plots, as used in Figure 3b) for showing model output when interpolation could be warranted or using time series line plots when showing observed data.

A really useful plot, particularly to support Section 4.2, would be a simple time series line plot of the water temperature at AWS_{open} and AWS_{FDS}. You should also include the instantaneous difference between the two sites. This will allow the reader to see more clearly (than is provided in Figures 3a-d) the diurnal patterns, the lag in daily maximum temperature, and difference between these two sites.

I assume that Figures 3a-d are presenting the measured water temperatures? If so, what kind of interpolation approach was used to generate these plots? Is interpolation of these data warranted?

For Figures 7a-c, it mentions the black lines represent the water parcels. Because of the 3D plot, it is very difficult to determine the travel time for one parcel to travel from 0 m to 1000 m, and my best guesses from the figures suggest anywhere between 4 to 6 hours although the manuscript reports that travel times were on average 7.5 h. Am I misreading these figures?

Specific comments

Title: I would consider revising the title. I appreciate the succinctness; however, it gives the impression that the study will provide a generalized explanation for the drivers of negative temperature gradients in (multiple) forested reaches. Instead, the study reports on a specific case study where groundwater discharge and hyporheic exchange are assumed to have no impact on the thermal regime. Therefore, at best the findings of this study are limited to reaches that meet these conditions.

Page 6442, line 14 and page 6446, line 10: '$>$ 200 hemispherical photographs'; exactly how many hemispherical photographs were taken?

Page 6444, lines 8–10: What is 'point-scale' defined as here? Story et al. (2003) and Leach and Moore (2011) were conducted over reach lengths of about 250 m and 1500 m, respectively. Are these considered point-scale studies?

Page 6445, line 25: Is 9.5 m the channel width or the wetted width?

Page 6446, line 19: You have field data from October 2011 to July 2013, perhaps you can use these data to calculate instantaneous longitudinal temperature gradients for the whole record and highlight the distribution of gradients and show the frequency and magnitude of negative gradients. This would put the detailed week long study into broader context. This is just a suggestion.

Page 6447, section 3.2.1: How much lateral variability was there at installation locations of the temperature loggers?

Page 6447, section 3.2.2: How many discharge and stream surface width surveys were conducted during the study week?

Page 6447, lines 15–16: Is the error in the discharge measurements asssumed to be $\pm 10\%$ or was this error quantified for these measurements by using replicated measurements or some other approach? Does this uncertainty impact your flow routing and water temperature model results?

Page 6448, section 3.3: How was the bed heat conduction flux estimated? What field data were used?

Page 6453, line 2: How and why is the discharge scaled by catchment area?

Section 4 'Results': Please consider using more specific quantitative language when describing the results. There is considerable usage of terms such as 'very low', 'lower', and 'high'.

Page 6456, lines 4–5: The flow routing model suggests a mean average travel time of 7.5 h. How much did this vary during the study period?

Page 6460, line 12: This modelling approach requires considerable field data collection and parametrization to run the model. I question whether it is a realistic tool to be used for areas where observational datasets are unavailable.

Technical corrections

Page 6443, line 23: Perhaps replace 'decreases in temperature' with 'negative instantaneous differences in temperature'.

Page 6444, line 4: Replace 'Storey' with 'Story'.

Page 6445, line 21: 'Dominated predominantly' is redundant.

Page 6452, line 18: What does the '900' refer to in 'Δ_{900}'?

Page 6466, Figure 1: The AWS labels are incorrect. Also, please add the forest cover to the plot.

Page 6471, Figure 6: Add letters to the plots for reference. Also, perhaps use different symbols or different line types for the four time periods examined, since it is difficult to tell the colours apart when printing in greyscale.

References

Brown, G. W., Swank, G. W., and Rothacher, J. 1971. Water temperatures in the Streamboat Drainage. Technical report, USDA Forest Service Research Paper PNW-199.

Leach, J. A. and Moore, R. D. 2011. Stream temperature dynamics in two hydrogeomorphically distinct reaches. *Hydrological Processes*, 25(5):679–690.

Malcolm, I. A., Soulsby, C., Youngson, A. F., and Hannah, D. M. 2005. Catchment scale controls on groundwater-surfacewater interactions in the hyporheic zone: implications for salmon embryo survival. *River Research and Applications*, 21:977–989.

Payn, R. A., Gooseff, M. N., McGlynn, B. L., Bencala, K. E., and Wondzell, S. M. 2009. Channel water balance and exchange with subsurface flow along a mountain headwater stream in Montana, United States. *Water Resources Research*, 45(11):W11427.

Story, A., Moore, R. D., and Macdonald, J. S. 2003. Stream temperatures in two shaded reaches below cutblocks and logging roads: downstream cooling linked to subsurface hydrology. *Canadian Journal of Forest Research*, 33(8):1383–1396.