Interactive comment on "Stream temperature prediction in ungauged basins: review of recent approaches and description of a new physically-based analytical model" *by* A. Gallice et al.

Anonymous Referee #1

Received and published: 27 May 2015

This manuscript presents a stream temperature model based on an analytical solution of a simplified energy balance equation. In spite of this claim, the end result is a regression model, mostly based on climatic and physiographic data. The model is applied successfully to estimate monthly river temperatures at the outlet of a number of drainage basins in Switzerland.

This paper is of interest to all water temperature modellers and should eventually be published in my view. However, there are a number of major revisions required. Gen-

erally, the claim that the model is a net improvement over other regressions models is a bit of an overstatement. I see it more as a regression model for which the selection of predictors is based on known relationships between water temperature, climate, solar radiation and local physiography/hydrology. The authors claim (line 210) that the selection of predictor variables relies on the experience of the authors in statistical models. This is not always necessarily the case. Some authors have used objective or semi-objective methods for this selection. In the submitted manuscript, the selection of proxy variables for the net radiative flux is also based on experience and known relationships. Parcimony was only used for model variants, but when the author's model is compared to the regression approach, parcimony is evacuated from the comparison.

Specific comments: Line 78. Las sentence appears to be a strong statement, perhaps not fully substantiated. Line 185: "for each group, they computed the characteristics of the thermal regime by river class. . ." Line 210: As stated earlier, not all statistical models are built using a subjective selection of predictors. For instance, multivariate analysis has been used in some cases. Line 236: Linear functions of available data are being used. Why linear? Is there proof of linearity in all cases? Line 315: "All sources of the network are supposed to have the same discharge". What are the potential repercussions of this simplification, especially in small drainage basins? Line 404: Meteorological stations were selected based on the fact that they are located at 20 km or less from the outlet. What is the impact of this selection when Lc is larger than 20 km? Line 421: Speaking of subjective selection : setting at least two of the coefficients to zero. At least? Were there more? Why two? Line 530: Number of terms in the model was fixed to six to prevent over-parameterization. Given that the final model is a multiple regression? Why not use a stepwise approach? Line 595 (and elsewhere): ". . . data are. . ." Line 620: Threshold of 30°C. Although this is a high value, might it not be feasible that some rivers occasionally reach this temperature in shallow, open areas? Line 669: Subjective selection of 5 of the 26 catchments. Why not do a bootstrap on the stations excluded? Line 713: Why not show AIC separately for calibration and validation? Lines 1015-1020 read more like a discussion than a

conclusion.