Abstract

Three recent papers review and analyze large global datasets related to impacts of forest cover on streamflow. Using three different approaches, they all find a strong relationship between forestation, de-forestation and streamflow. However, the results are problematic, the underlying data set is unbalanced, and there are correlations in the data that warrant further investigation as this would influence the results. For example, the area of the catchment is strongly related to the assessment technique and the variability in the response data. For this study, the data for the recent three papers were reviewed, combined, and supplemented with new studies. Subsequently, the data were re-analyzed using generalized additive modelling. The results highlight that there are four interlinked reasons that make the general outcomes from the previous papers problematic: 1) The existence of latent variables in the data that create the appearance of a relationship that really does not exist; 2) The difficulty in fully interpreting the specifics of different studies; 3) The difficulty of integrating data from seemingly similar studies, but with quite different objectives; and 4) The chance of transcription errors influencing the data. Overall this indicates that while valuable data can be extracted from past studies, the above problems need to be considered before results are generalized and extrapolated to continental and global scales.