Supplement 5: Model Performance

As shown in the Results section of the manuscript, the model performed well over most of the period of study, but failed to reproduce inflow to the TG Halli reservoir in the 1991-1995 period and the groundwater levels (GWL) in the 2006-2010 period. The reasons for these discrepancies have to do with the inability to completely capture heterogeneity and complexity of the system.

For instance, because infiltration excess rainfall is the primary runoff generation mechanism, runoff in the model is primarily generated during short high-intensity events which are heterogeneous in space and localized in time. With high temporal resolution rainfall data (30 minute scales) only available from 2010 onwards, and four operational long-term gauges over the watershed from 1970 onward, capturing the space-time behavior of storms accurately was not possible. Historical daily rainfall data were *probabilistically downscaled* to generate a historical 30 min historical rainfall series, meaning that the range of storm intensities is appropriate, yet this is clearly insufficient to reproduce inflows on short (daily-weekly) timescales. Therefore, the model can only be reasonably compared to long term average inflows into TG Halli reservoir, and it is plausible that spatial variability in rainfall (with high rainfall observed at the four gauges) resulted in overestimates of inflow in the 1991-1995 period.

Furthermore, the observed depths of the groundwater level were highly uncertain. Historical groundwater monitoring well data were found to be completely inaccurate (Fig. S3.1-S3.2), and as a result, simulated groundwater declines could not be compared with monitoring well data. Instead the model was calibrated against the decline inferred from the borewell histories. The 2006-2010 period contained considerable change in the catchment including (a) reductions in groundwater availability and (b) implementation of the watershed development programme. Because farmer decisions are sensitive to groundwater levels and the GWL in the fractured rock aquifer is highly sensitive to changes in water storage, the model was likely unable to fully capture the timing or complex interactions associated with these dynamics.