Modeling the contribution of e-flows to sustainable agriculture, food security and livelihoods in South Africa’s Limpopo basin

Eike Luedeling, Cory Whitney

This document outlines a collection of holistic modeling techniques aimed at describing the link between river flows and livelihoods. The model simulations we present here are generated using environmental flows (e-flows), which are defined as water provided within a river or wetland to maintain aquatic ecosystems which in turn support dependent livelihoods. These can also be thought of as the ‘ecological water demand’ needed to protect freshwater-dependent ecosystems. The overall objective of the model we outline here is to simulate how river flow, in particular e-flow, impacts smallholder agriculture. The objective is to illustrate the linkage between sustainable e-flows in the rivers, and the water-requirements of sustainable agriculture. The simulation offers insights into the role that river flows play in the ability of agriculture to be sustainable, and the consequent risks to agriculture when river flows are either optimal or when they become marginal. The simulations show a strong positive link between e-flows and livelihoods and provide justification for understanding e-flows as more than just a mechanism for positive biodiversity outcomes.

We generate a holistic model to simulate the contribution of e-flows to sustainable agriculture, food security and livelihoods. Spatially, we do this for only a small portion of the basin as a test-case. We apply holistic modeling approaches to generate conceptual impact pathways and quantitative models to forecast decision outcomes (see Do, Luedeling, and Whitney 2020; Lanzanova et al. 2019; Cory Whitney et al. 2018). This includes collaborative model development (C. Whitney, Shepherd, and Luedeling 2018) to assess farming futures given e-flow forecasts under different management options. To build these simulations we use functions from the *decisionSupport* (Luedeling et al. 2021), *dplyr* (Wickham, François, et al. 2021), *nasapower* (Sparks 2021), *patchwork* (Pedersen 2020), *tidyverse* (Wickham 2021) and E*vapotranspiration* (Guo, Westra, and Peterson 2020) libraries in the R programming language (R Core Team 2021).

## The model

Decision-makers often wish to have a quantitative basis for their decisions. However, ‘hard data’ is often missing or unattainable for many important variables, which can paralyze the decision-making processes or lead decision-makers to conclude that large research efforts are needed before a decision can be made. That is, many variables decision makers must consider cannot be precisely quantified, at least not without unreasonable effort. The major objective of (prescriptive) decision analysis is to support decision-making processes where decision makers are faced with this problem. Following the principles of Decision Analysis can allow us to make forecasts of decision outcomes without precise numbers, as long as probability distributions describing the possible values for all variables can be estimated.

The *decisionSupport* package implements this as a Monte Carlo simulation, which generates a large number of plausible system outcomes, based on random numbers for each input variable that are drawn from user-specified probability distributions. This approach is useful for determining whether a clearly preferable course of action can be delineated based on the present state of knowledge without the need for further information. If the distribution of predicted system outcomes does not imply a clearly preferable decision option, variables identified as carrying decision-relevant uncertainty can then be targeted by decision-supporting research.

We apply the *mcSimulation* function from the *decisionSupport* package to conduct decision analysis (Luedeling et al. 2021). We use this to conduct a Monte Carlo analysis with repeated model runs based on probability distributions for all uncertain variables. We built a custom data table and model to fit the particulars of the different flow scenarios. We worked with project partners to generate a conceptual model of the social effects of altered river flows on the sustainability of livelihoods in the Limpopo Basin (Figure 1). To support the model building process, we design an input table to store the *estimate* values (Annex). The distributions are described by 90% confidence intervals, which are specified by lower (5% quantile) and upper (95% quantile) bounds, as well as the shape of the distribution (Annex).

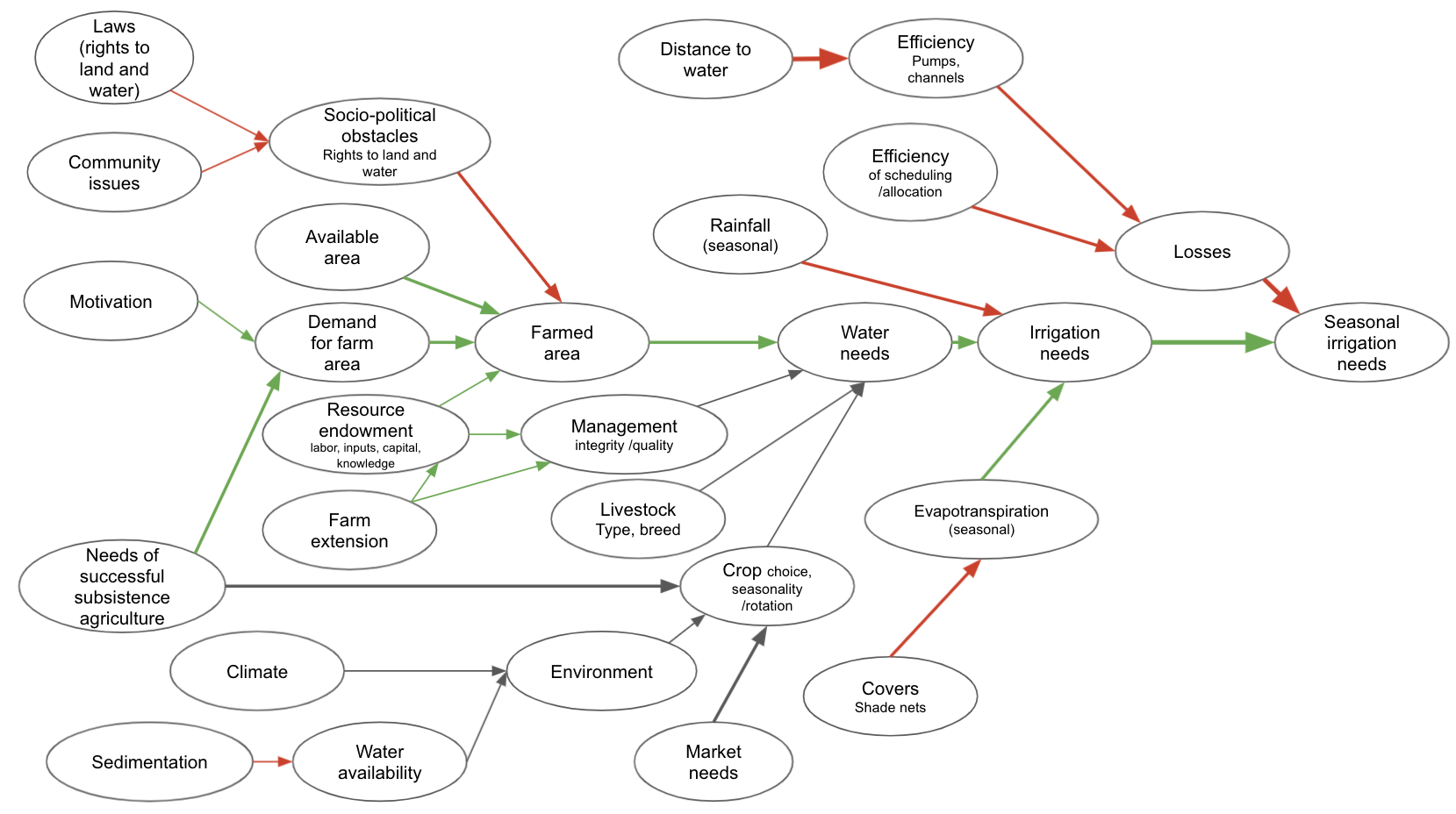


Figure 1. Conceptual model of the social effects of altered river flows on the sustainability of livelihoods in the Limpopo Basin

We used the conceptual model (Figure 1) to generate an R function that takes in the variables provided in the input table and produces a model output. In the following we use of various *decisionSupport* functions, which use the *tidyverse* libraries (Wickham et al. 2019) including *ggplot2* (Wickham, Chang, et al. 2021), *plyr* (Wickham 2020) and *dplyr* (Wickham, François, et al. 2021) among others in the [R programming language](https://www.r-project.org/) (R Core Team 2021). We generated the model as a function using the *decisionSupport* functions *vv()* to produce time series with variation from a pre-defined mean and coefficient of variation and *chance\_event*() to simulate whether events occur for our intervention comparison.

### Scenarios

The model function defines 3 scenarios:

* Scenario 1 - no e-flows: This is a scenario without e-flows. Farmers extract water according to their irrigation needs. Extractions are only limited by the minimum water level that allows operating the pumps.
* Scenario 2 - restricted extraction (only ecological system protection is considered in e-flows): This is an e-flow scenario, in which e-flows are interpreted in a purely ecological sense. Whenever e-flows aren’t achieved, water extraction is curtailed. There are no measures to add water to the river in such events. We simulate Scenario 1 with our own functions and some from the *nasapower* (Sparks 2021) and *Evapotranspiration* (Guo, Westra, and Peterson 2020) packages. Here farmers are considered an obstacle for the e-flows and are (e-flows do not include livelihoods requirement)
* Scenario 3 - dam releases (ecosystems and livelihoods are included in the e-flows): This is an e-flow scenario, in which e-flows are interpreted as encompassing the ecological as well as the smallholder irrigation requirement. In case e-flows aren’t naturally met, water is released from upstream dams to ensure e-flows. Extraction by smallholder farmers is restricted only by the ability to operate the pumps. We assume here that the e-flow plus the smallholder farmer needs are available. In this scenario farmers are part of the e-flows and livelihood requirements (including agriculture) are included in the e-flows.

We performed a Monte Carlo simulation by generating distributions of all variables in the input table as well as the specified model outputs. We ran the model with 10,000 random draws in a defined function (Annex 1, Limpopo model).

## Results

The results refer to the scenarios we formulated in the model (Annex). We use the results of the simulations of Scenario 1 for comparisons. We programmed this scenario so that farmers extract water according to their irrigation needs and this is only limited by the minimum water level that allows them to operate their water pumps. We use this scenario for testing the model simulations of the other scenarios. We do this by subtracting the farm level water gap resulting from this scenario from the water gap for each of the other scenarios.

#### Plot distributions

Model results show the difference in farm level crop water gap (i.e. how much water is missing from the smallholder farmers needs) between Scenario 1, no e-flows and e-flow Scenario 2, restricted extraction. We simulated this scenario so that whenever e-flows aren’t achieved, water extraction is curtailed and there are no measures to add water to the river.

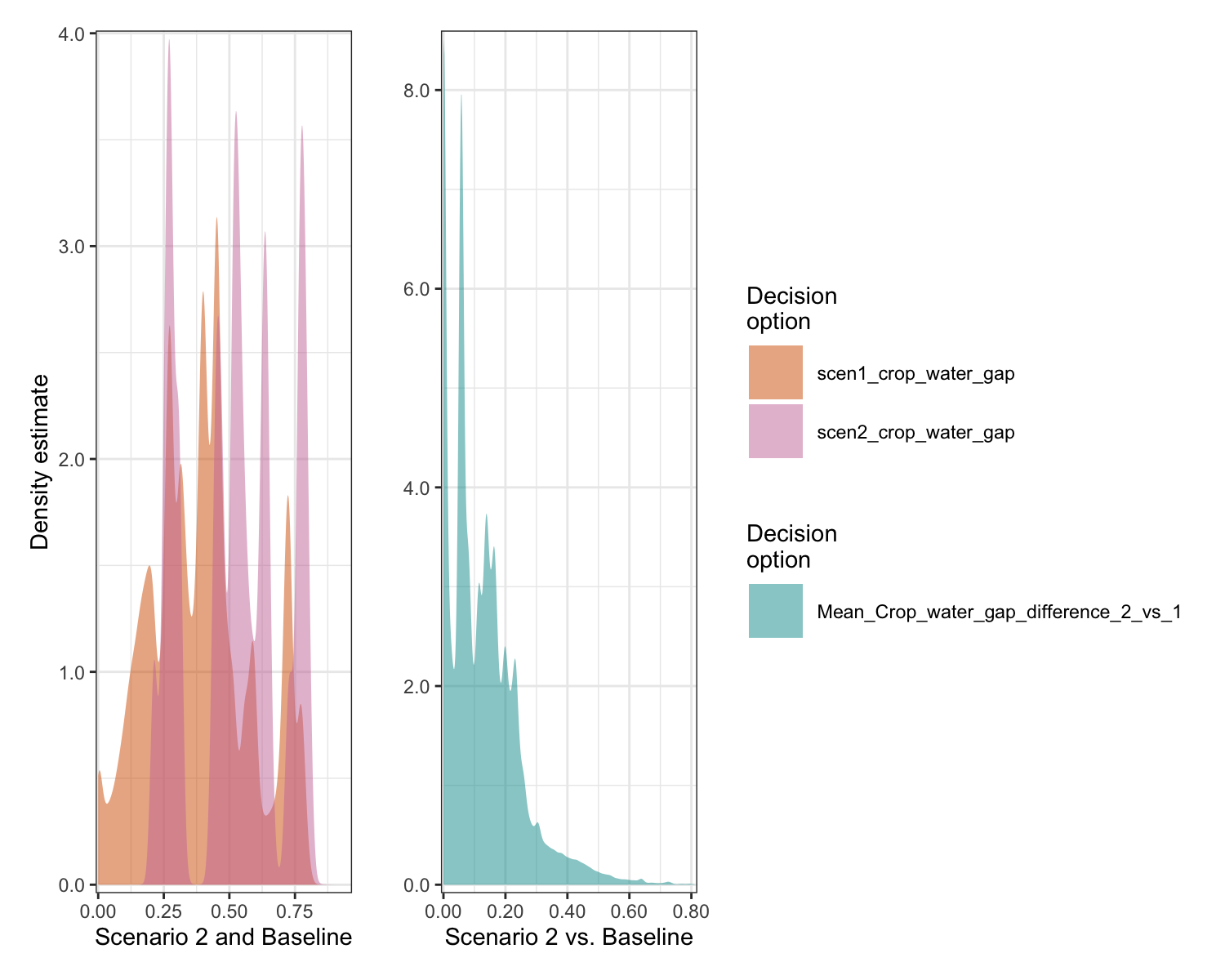


Figure 2. Distributions of differences in the simulation results between the Scenario 1 with no e-flows and Scenario 2 wherein farmers are restricted from extracting water below e-flow levels.

The crop water gap differences in the simulation results between the Scenario 1 with no e-flows and Scenario 2 wherein farmers are restricted from extracting water below e-flow levels are shown in Figure 2. The results indicate that Scenario 2 can lead to greater crop water gaps and have a negative impact on the livelihoods of smallholder farmers. In short, an e-flow scenario that simply curtails extraction but doesn’t alleviate sub-e-flow flows, is likely to cause considerable irrigation water shortages.

Figure 4 shows the difference in farm level crop water gap (i.e. how much water is missing from the smallholder farmers needs) between Scenario 1, no e-flows and Scenario 3, dam releases. We simulate this scenario so that when e-flows aren’t naturally met, water is released from upstream dams to ensure e-flows. Smallholder farmers can access water so long as their pumps are working.

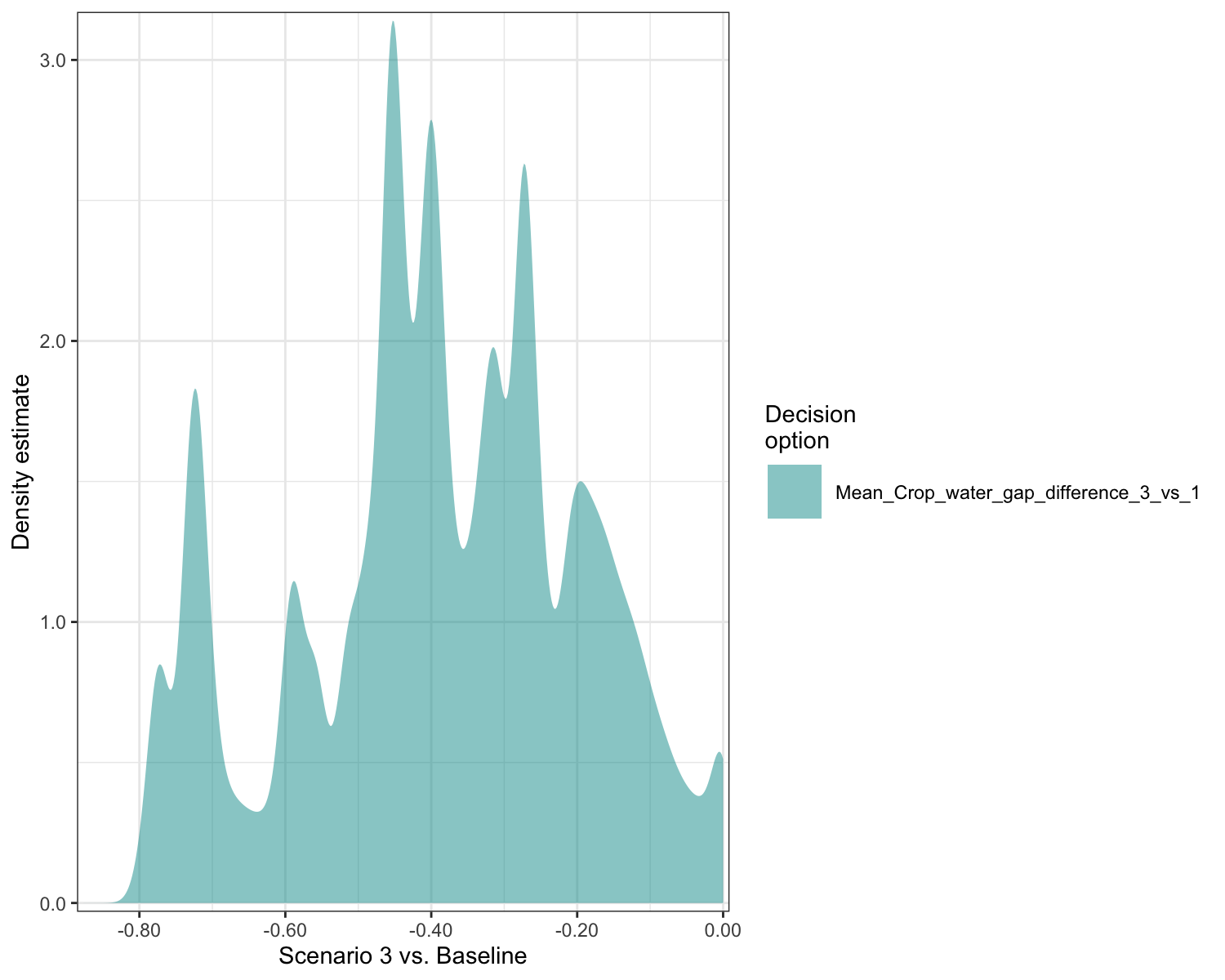


Figure . Monte Carlo simulation results between Scenario 1 with no e-flows and Scenario 3 wherein water is released from upstream dams to maintain e-flows

The differences in the simulation results between Scenario 1 with no e-flows and Scenario 3 wherein water is released from upstream dams to maintain e-flows are shown in Figure 3. The results demonstrate the effectiveness of e-flows for supporting smallholder farmers and their livelihoods by effectively eliminating the crop water gap. The results indicate that an e-flow scenario that includes dam releases to ensure e-flows can benefit farmers by effectively precluding irrigation water shortfalls.

#### Flow analysis

We calculated a downstream differenceoutput from in our simulation function (Annex) to show the difference in downstream river flow over the 12 months of the simulation. In this part of the simulation whenever the present flow is below the eflow requirement, water is released from an upstream dam to ensure that the eflows are met.

Here we plot the distribution of downstream flow difference over the entire simulated year. Our model simulation

Figure 4 shows the flows for scenario 2 and for scenario 3 vs. the no-e-flow Scenario 1.

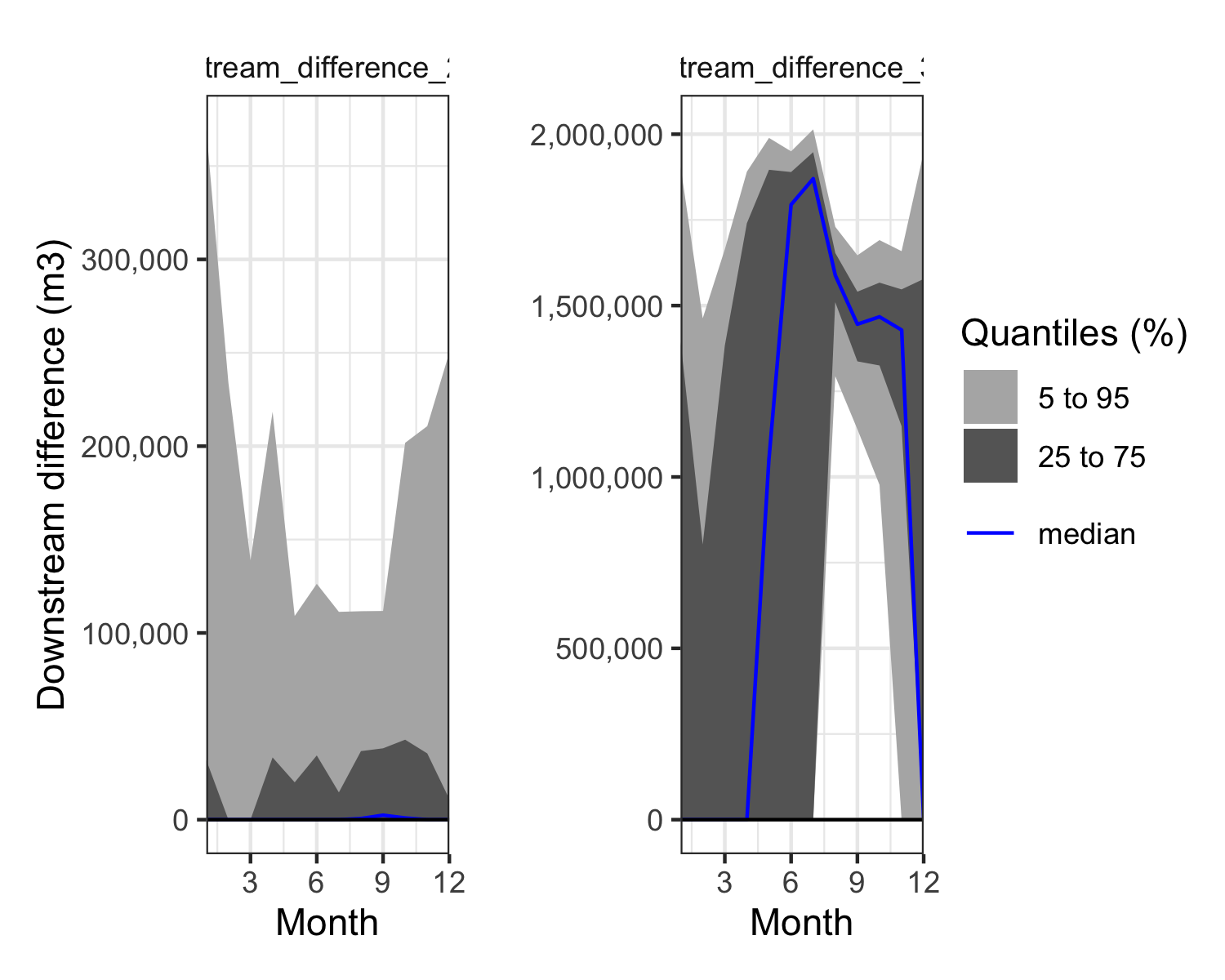


Figure 4. Flows for scenario 2 (left) and for scenario 3 (right) vs. the no-e-flow Scenario 1 across the 12 months of the simulation.

And these are the dam releases required to maintain scenario 3 (Figure 5).

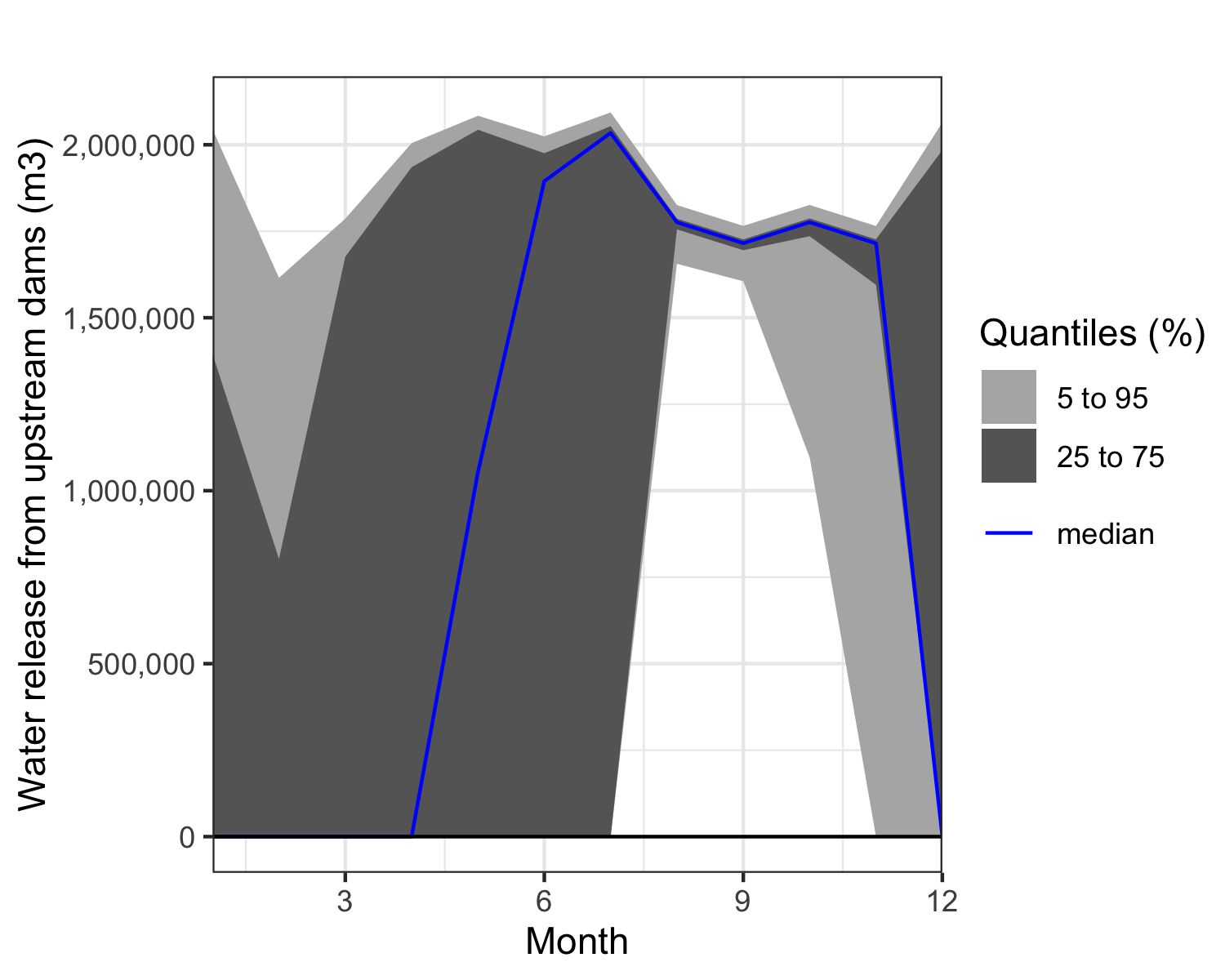


Figure 5. dam releases required to maintain scenario 3

Here similar plots of Crop\_water\_gap\_difference to show the crop water gap over time (across 12 months of the simulation).

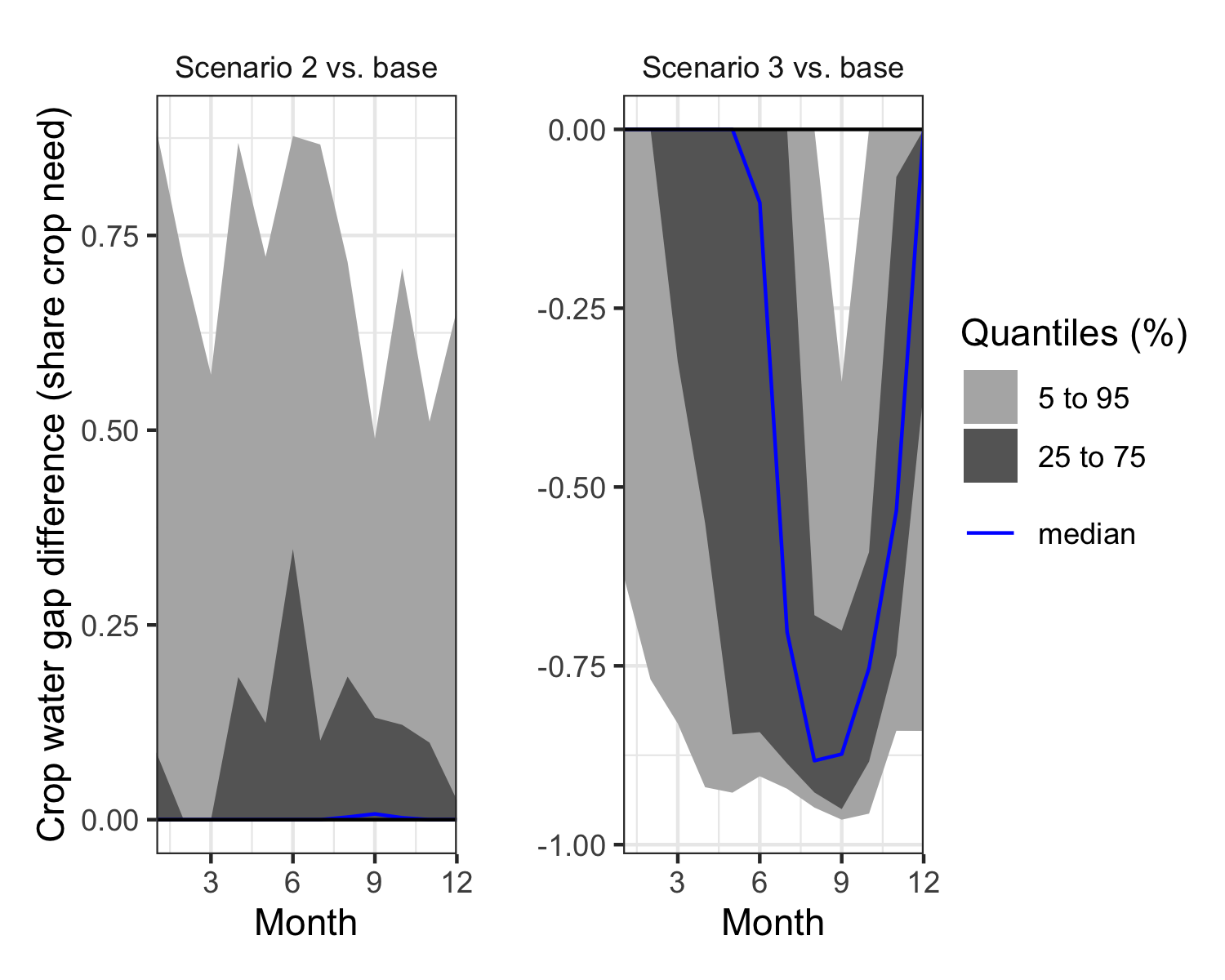


Figure 6. Crop water gap difference indicating the crop water gap over the 12 months of the simulation.

#### Model sensitivity analysis

We apply a post-hoc analysis in the form of a regression model and variable importance analysis to the model outputs. We do this by calculating *Variable Importance in the Projection (VIP)* scores and coefficients of a Projection to Latent Structures (PLS) regression model (Wold, Sjöström, and Eriksson 2001). The VIP scores help us to estimate the importance of each variable in the projection used in the PLS model. We use the VIP as a parameter for calculating the cumulative measure of the influence of individual variableson our model.

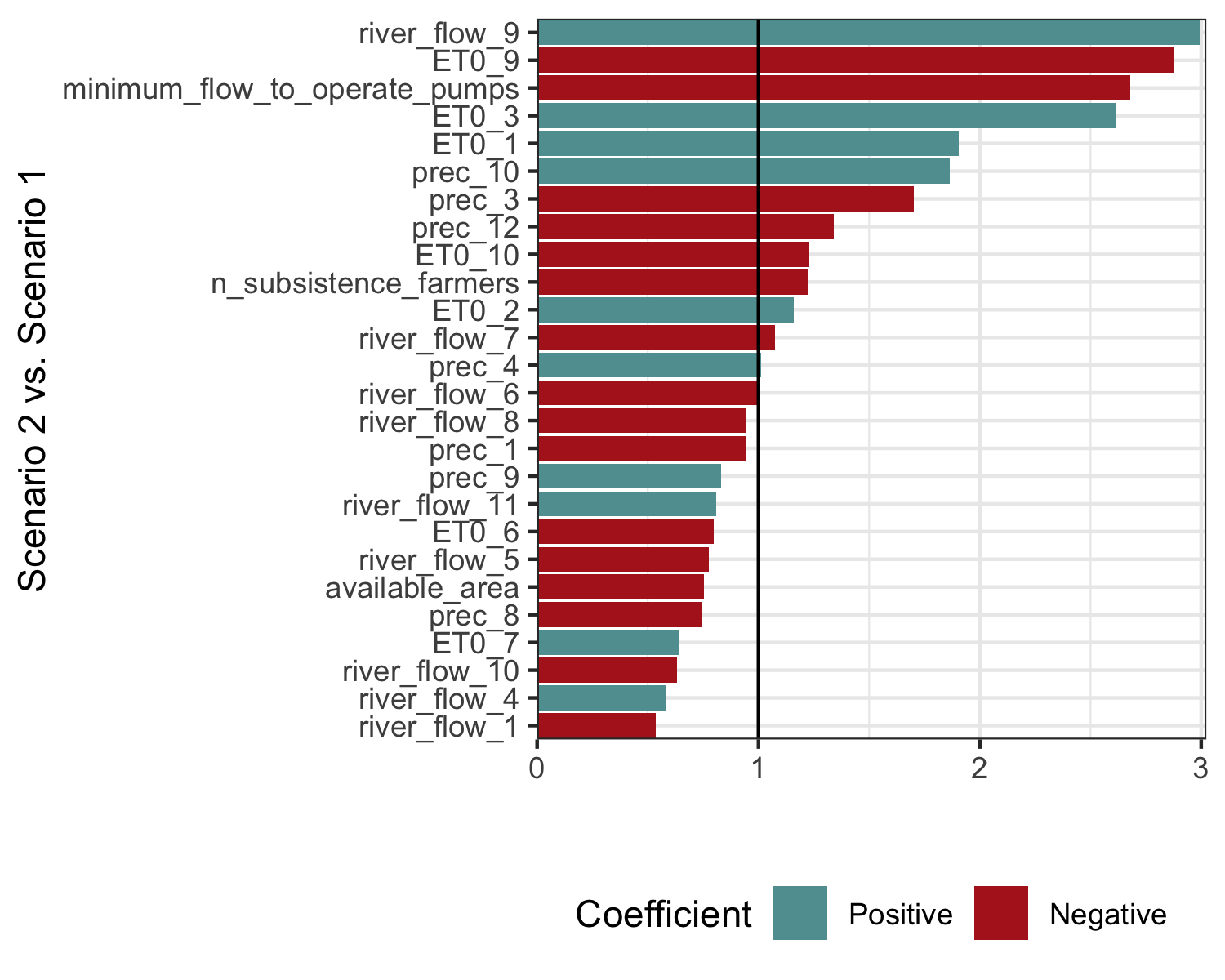


Figure 7. *Variable Importance in the Projection (VIP)* score and coefficients of a *Projection to Latent Structures (PLS)* regression model to assess the mean crop water gap in Scenario 1 and 2. The length of the bars is equal to VIP, the vertical line at ‘1’ on the x-axis indicates the cut-off for VIP used for variable selection. The overall plot only shows those variables with a VIP > 0.5. The colors of the bars represent the positive or negative coefficient of the given input variable with the output variable.

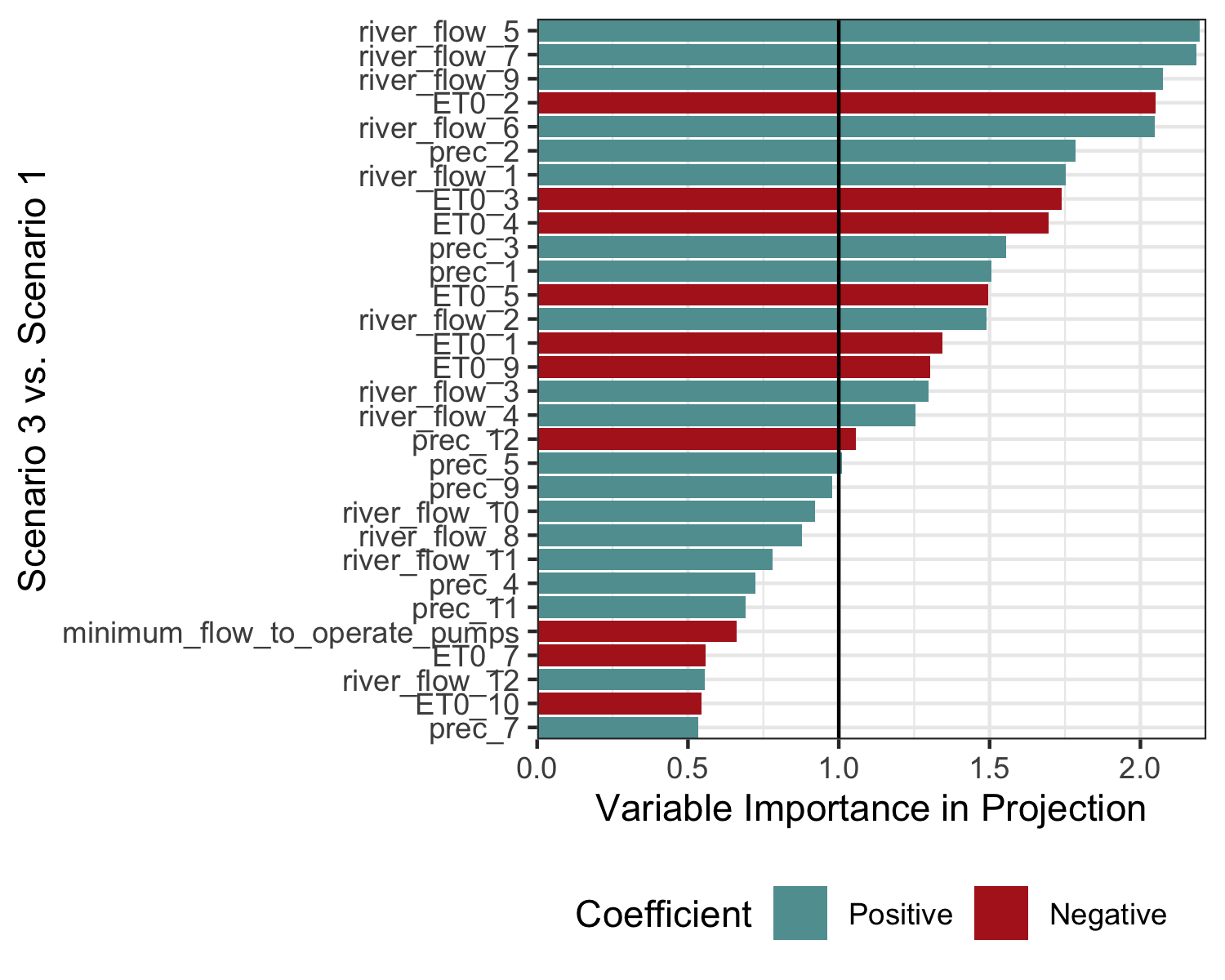


Figure 8. *VIP* score and coefficients of a *PLS* regression model to assess the mean crop water gap in Scenario 1 vs. 3. The length of the bars is equal to VIP, the vertical line at ‘1’ on the x-axis indicates the common value used for selecting variables. The overall plot only shows those variables with a VIP > 0.5. The colors of the bars represent the positive or negative coefficient of the given input variable with the output variable.

The VIP scores show us the importance of each variable in the projection used in our PLS models for the mean crop water gap in scenario 2 and 3 vs.  Scenario 1 (Figure 7; Figure 8).

### Discussion

We assume that the Scenario 3 water is available in reservoirs

The upstream effects of Scenario 3 are not considered here

* Could be a disaster for a whole region if all farmers demand Scenario 3

### Addendum

The objective of the procedures used in the *decisionSupport* package is to make it easier for analysts to produce decision-relevant information that adequately reflects the imperfect nature of the information we usually have. Adding probabilistic elements to a simulation adds substantial value to an analysis. Mostly, it avoids making spurious assumptions, replacing uncertainty with ‘best bets’ and producing results that do not reflect the knowledge limitations that make decision-making so challenging. More information on all this is contained in the [decisionSupport manual](https://cran.r-project.org/web/packages/decisionSupport/decisionSupport.pdf), especially under *welfareDecisionAnalysi*s.

### References

Do, Hoa, Eike Luedeling, and Cory Whitney. 2020. “Decision Analysis of Agroforestry Options Reveals Adoption Risks for Resource-Poor Farmers.” *Agronomy for Sustainable Development* 40 (3): 20. <https://doi.org/10.1007/s13593-020-00624-5>.

Galindo-Prieto, Beatriz, Lennart Eriksson, and Johan Trygg. 2014. “Variable Influence on Projection (VIP) for Orthogonal Projections to Latent Structures (OPLS): Variable Influence on Projection for OPLS.” *Journal of Chemometrics* 28 (8): 623–32. <https://doi.org/10.1002/cem.2627>.

Guo, Danlu, Seth Westra, and Tim Peterson. 2020. *Evapotranspiration: Modelling Actual, Potential and Reference Crop Evapotranspiration*. <https://CRAN.R-project.org/package=Evapotranspiration>.

Lanzanova, Denis, Cory Whitney, Keith Shepherd, and Eike Luedeling. 2019. “Improving Development Efficiency Through Decision Analysis: Reservoir Protection in Burkina Faso.” *Environmental Modelling & Software* 115 (May): 164–75. <https://doi.org/10.1016/j.envsoft.2019.01.016>.

Luedeling, Eike, Lutz Goehring, Katja Schiffers, Cory Whitney, and Eduardo Fernandez. 2021. *decisionSupport: Quantitative Support of Decision Making Under Uncertainty*. <http://www.worldagroforestry.org/>.

Luedeling, Eike, and Keith Shepherd. 2016. “Decision-Focused Agricultural Research.” *Solutions* 7 (5): 46–54. https://doi.org/<https://www.thesolutionsjournal.com/article/decision-focused-agricultural-research/>.

Pedersen, Thomas Lin. 2020. *Patchwork: The Composer of Plots*. <https://CRAN.R-project.org/package=patchwork>.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Sparks, Adam H. 2021. *Nasapower: NASA POWER API Client*. <https://docs.ropensci.org/nasapower/>.

Whitney, Cory, D. Lanzanova, C. Muchiri, K. Shepherd, T. Rosenstock, M. Krawinkel, J. R. S. Tabuti, and E. Luedeling. 2018. “Probabilistic Decision Tools for Determining Impacts of Agricultural Development Policy on Household Nutrition.” *Earth’s Future* 6 (3): 359–72. <https://doi.org/10.1002/2017EF000765/full>.

Whitney, C., K. Shepherd, and E. Luedeling. 2018. “Decision Analysis Methods Guide; Agricultural Policy for Nutrition.” *World Agroforestry (ICRAF)* Working Paper series (275): 40.

Wickham, Hadley. 2020. *Plyr: Tools for Splitting, Applying and Combining Data*. <https://CRAN.R-project.org/package=plyr>.

———. 2021. *Tidyverse: Easily Install and Load the Tidyverse*. <https://CRAN.R-project.org/package=tidyverse>.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.

Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. 2021. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. <https://CRAN.R-project.org/package=ggplot2>.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.

Wold, Svante, Michael Sjöström, and Lennart Eriksson. 2001. “PLS-Regression: A Basic Tool of Chemometrics.” *Chemometrics and Intelligent Laboratory Systems*, PLS Methods, 58 (2): 109–30. <https://doi.org/10.1016/S0169-7439(01)00155-1>.