26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Automating the extension of time series inputs for the Source GBCCL model

Xue Hou [a], Tony Sheedy [b]

[a] Department of Energy, Environment and Climate Action, Melbourne, Victoria, Australia, [b] WaterRes Pty Ltd, Carnegie, Victoria, Australia
Email: xue.hou@deeca.vic.gov.au

Abstract: Maintaining the operational relevance of large-scale hydrological models requires regular extension of time series inputs, such as climate, demand, and inflow data. Traditionally, these tasks have relied on manual, time-intensive processes, limiting efficiency and consistency. A Python-based automation tool was developed to streamline data acquisition, processing, and input extension for the GBCCL Source model. Its recent application, extending inputs to 30 June 2024, significantly reduced manual effort while improving workflow reliability and auditability. By enabling faster, repeatable, and more accurate updates, the tool enhances the operational utility of hydrological models and supports more timely decision-making in water resource management. Broader application across Victoria's Source models could establish a scalable, auditable framework that underpins continuous model improvement and long-term planning.

Keywords: *Automation, data extension, time series inputs, Source model, Python*

1. INTRODUCTION

Surface water models are essential tools for understanding Victoria's changing surface water resources, to share insights and support evidence-based decision-making for water resource planning, policy, investments and operations. In Victoria, the Goulburn, Broken, Campaspe, Coliban and Loddon (GBCCL) Source model is one of the larger and more complex systems. It includes over a century of daily hydrological time series data and detailed water sharing arrangements across rural and urban systems.

Maintaining the model's relevance requires regular updates to a wide range of input datasets, including climate, water demands, inflows, gauged observations, and water management data such as seasonal determinations and accounting rules. Traditionally, this task involved extensive manual effort and was susceptible to data handling errors, making the process inefficient and difficult to repeat consistently.

To address this, DEECA developed a Python-based automation tool within a Jupyter notebook environment to streamline the extension of time series inputs which previously covered the time period from January 1891 to June 2020. The tool was first used to extend the GBCCL Source model inputs to June 2023, and then recently further developed and was used to extend the inputs through to 30 June 2024 (DEECA and WaterRes, 2025), demonstrating its ongoing operational value.

This work builds on decades of modelling development in the GBCCL region, from the monthly time step Goulburn Simulation Model (GSM) which used the REALM software through to the GBCCL model in the Source platform. The automation work reflects a broader shift towards transparent, efficient, and scalable modelling practices. It helps ensure the model remains responsive to evolving climatic, policy, and operational conditions. The approach also provides flexibility to support broader applications across other Victorian Source models into the future.

2. METHODOLOGY AND RESULTS

The DEECA automation tool is implemented in Python and incorporates multiple key processes to efficiently extend the time series inputs for the GBCCL Source model. The workflow begins by downloading essential climate data from the SILO (Scientific Information for Land Owners) database, streamflow data from the Water Measurement Information System (WMIS) system, and historical seasonal determinations from Northern Victorian Resource Manager (NVRM). In combination with historical records and data provided by stakeholders (gauged flows, demands, historical storage data, etc.), they form the meteorological and hydrological inputs of this model.

Following data acquisition, the automated time series processing screens for outliers, treats noise for storage and reach mass balances, and infills missing data to maintain continuous and quality-controlled datasets.

© MSSANZ and The Authors, CC-BY

Historical demands are derived using a combination of observed time series flow data, observed annual metered data, simulated irrigation demands and defined patterns. 1993/94 demand time series are created using a combination of irrigation demand models and regression approaches. The inflow derivation methods for tributary inflows are gauged flow data infilled with rainfall runoff model flows, while ungauged area inflows are derived by mass balance with regressions applied where relevant to gauged flow data.

The tool generates automated comparison plots that visualise updated versus previous time series, along with cumulative curve comparisons and exceedance curve comparisons. Additionally, one of the scripts can perform annual data update comparison and general version comparison, which can identify missing values, compare file differences, provide detailed comparison summary, and undertake directory-wide processing.

The recent enhancement of the tool has enabled successful extension of the GBCCL model inputs through to 30 June 2024 (DEECA and WaterRes, 2025), significantly reducing manual processing time and minimising human error. The automated plots (example - Figure 1) include a comparison of the previous update with an overlap and model extension period using the latest data for calculations. The plots provide clear visual confirmation of input continuity and highlight data changes over the extended period. Together, these outcomes illustrate the tool's scalability and robustness, supporting efficient and repeatable updates for complex catchment models like Source GBCCL.

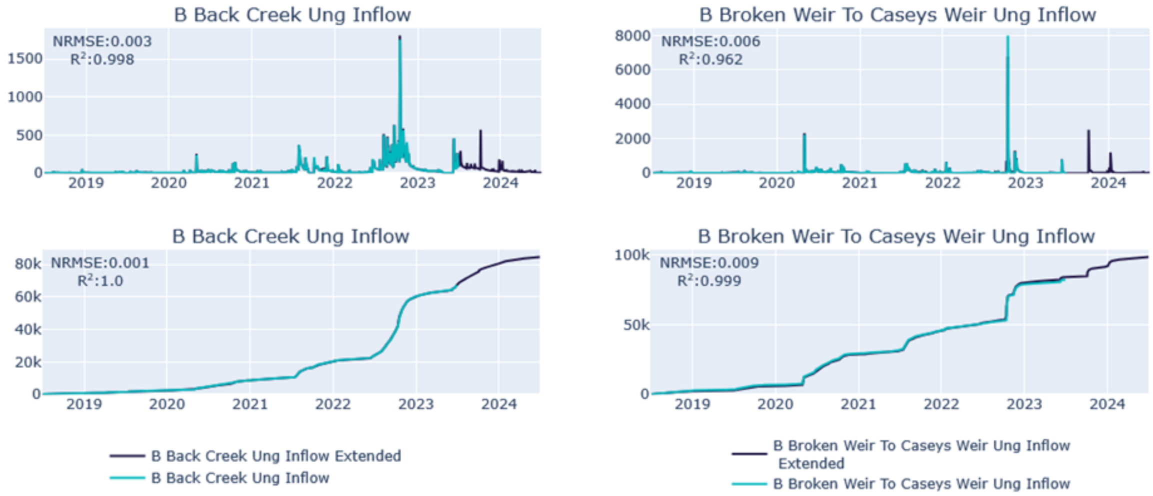

Figure 1 Time series and cumulative comparison plots of inflows for selected Broken River basin sites

3. CONCLUSIONS

A Python-based automation tool has been developed and refined to efficiently extend time series inputs for the Source GBCCL model. The tool has been applied successfully to update inputs through to 30 June 2024. In doing so, the tool has demonstrated practical value in reducing manual workload and improving the timeliness of model updates for large, complex systems.

The automation framework ensures consistent formatting and structure across inputs, enhances reproducibility through fully scripted and version-controlled processes, and supports more transparent and auditable data handling practices. While the current implementation is tailored to the GBCCL model, the underlying logic and structure have the potential to be adapted to other Source models with similar configurations.

This work contributes to the broader goal of modernising water resource modelling workflows by reducing reliance on manual processing and improving the reliability of model input preparation. Ongoing enhancements will focus on incorporating additional automated checks, improving usability to further accommodate Python beginners, and exploring opportunities for broader application across Victoria's Source models and beyond.

ACKNOWLEDGEMENTS

We acknowledge the Murray–Darling Basin Authority (MDBA) for co-funding through the Integrated River Modelling Uplift. (IRMU) Program for supporting improvements to this tool and thank for Surface Water Assessment and Modelling (SWAM) team members for their generous time and guidance, particularly during the early stages of idea development and tool testing.

REFERENCES

DEECA and WaterRes (2025) Inputs to the GBCCL Source Model to 30 June 2024: Python User Manual

© MSSANZ and The Authors, CC-BY