



Streamlining Data Pre-processing and Analysis through the V-FOR-WaTer Web Portal

Marcus Strobl¹, Elnaz Azmi¹, Bischof Balazs², Safa Bouguezzi¹, Alexander Dolich², Sibylle K. Hassler^{2,3}, Mirko Mälicke², Ashish Manoj J², Jörg Meyer¹, Achim Streit¹, and Erwin Zehe²

¹Scientific Computing Center (SCC), Karlsruhe Institute of Technology (KIT), Eggenstein-Leopoldshafen, Germany
(marcus.strobl@kit.edu)

²Institute for Water and Environment (IWU), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

³The Institute of Meteorology and Climate Research Atmospheric Trace Gases and Remote Sensing (IMKASF), Karlsruhe Institute of Technology (KIT), Eggenstein-Leopoldshafen, Germany

The rapid growth of environmental data and the complexity of data pre-processing tasks poses significant challenges to environmental scientists. Repetitive and error-prone manual data preparation methods not only consume valuable research time but also introduce potential data quality issues. Also, individually pre-processed datasets are hardly reproducible. The V-FOR-WaTer virtual research environment (VRE) addresses these challenges as a powerful tool that seamlessly integrates data access, data pre-processing, and data exploration capabilities.

V-FOR-WaTer has an automated data pre-processing workflow to improve data preparation by eliminating the need for manual data cleaning, standardization, harmonization, and formatting. This approach significantly reduces the risk of human error while freeing up researchers to focus on their actual research questions without being hampered by data preparation. The pre-processing tools integrated in the virtual research environment are designed to handle a wide range of data formats, ensuring consistent and reliable data preparation across diverse disciplines. This empowers researchers to seamlessly integrate data from various sources in a standardized manner.

The web portal's user-centric design facilitates data exploration and selection through map operations and filtering options, empowering researchers to efficiently identify and focus on relevant data for their analyses. The scalability and extensibility of the V-FOR-WaTer web portal ensures that it can accommodate the ever-growing volume of environmental data and adapt to the evolving research landscape. Its ability to integrate user-developed tools reflects the dynamic nature of environmental research and ensures that the virtual research environment stays up-to-date with the latest advancements in data processing. The comprehensive features and user-friendly interface position it as a valuable tool for environmental scientists, fostering collaboration, streamlining data analysis, and accelerating the advancement of knowledge in the field of hydrology.