



Accessing environmental time series data in R from Sensor Observation Services with ease

Daniel Nüst¹, Eike H. Jürrens², Benedikt Gräler², and Simon Jirka²

¹University of Münster, Institute for Geoinformatics, Opening Reproducible Research (o2r), Münster, Germany
(daniel.nuest@uni-muenster.de)

²North Initiative for Geospatial Open Source Software GmbH, Germany

Time series data of in-situ measurements is the key to many environmental studies. The first challenge in any analysis typically arises when the data needs to be imported into the analysis framework. Standardisation is one way to lower this burden. Unfortunately, relevant interoperability standards might be challenging for non-IT experts as long as they are not dealt with behind the scenes of a client application. One standard to provide access to environmental time series data is the Sensor Observation Service (SOS,) specification published by the Open Geospatial Consortium (OGC). SOS instances are currently used in a broad range of applications such as hydrology, air quality monitoring, and ocean sciences. Data sets provided via an SOS interface can be found around the globe from Europe to New Zealand.

The R package `sos4R` (Nüst et al., 2011) is an extension package for the R environment for statistical computing and visualization (`ggplot2`), which has been demonstrated as a powerful tool for conducting and communicating geospatial research (cf. Pebesma et al., 2012;). `sos4R` comprises a client that can connect to an SOS server. The user can use it to query data from SOS instances using simple R function calls. It provides a convenience layer for R users to integrate observation data from data access servers compliant with the SOS standard without any knowledge about the underlying technical standards. To further improve the usability for non-SOS experts, a recent update to `sos4R` includes a set of wrapper functions, which remove complexity and technical language specific to OGC specifications. This update also features specific consideration of the OGC SOS 2.0 Hydrology Profile and thereby opens up a new scientific domain.

In our presentation we illustrate use cases and examples building upon `sos4R` easing the access of time series data in an R and Shiny () context. We demonstrate how the abstraction provided in the client library makes sensor observation data for accessible and further show how `sos4R` allows the seamless integration of distributed observations data, i.e., across organisational boundaries, into transparent and reproducible data analysis workflows.

References

Nüst D., Stasch C., Pebesma E. (2011) Connecting R to the Sensor Web. In: Geertman S., Reinhardt W., Toppen F. (eds) *Advancing Geoinformation Science for a Changing World*. Lecture Notes in

Geoinformation and Cartography, Springer.

Pebesma, E., Nüst, D., & Bivand, R. (2012). The R software environment in reproducible geoscientific research. *Eos, Transactions American Geophysical Union*, 93(16), 163–163.