Frontiers in Environmental Science

TYPE Data Report
PUBLISHED 01 July 2025
DOI 10.3389/fenvs.2025.1599188

OPEN ACCESS

EDITED BY
Agnieszka Jedruch,
Polish Academy of Sciences, Poland

REVIEWED BY
Jan Marcin Weslawski,
Polish Academy of Sciences, Poland
Markus Meier,
Leibniz Institute for Baltic Sea Research (LG), Germany

*CORRESPONDENCE
Simon A. Mischel,
mischel@bafg.de

RECEIVED 24 March 2025
ACCEPTED 13 June 2025
PUBLISHED 01 July 2025

CITATION
Mischel SA, Zentis C and Plessow H (2025) The BALTEX dataset of the Global Runoff Data Centre (GRDC): a dataset of river discharge draining into the Baltic Sea.
Front. Environ. Sci. 13:1599188.
doi: 10.3389/fenvs.2025.1599188

The BALTEX dataset of the Global Runoff Data Centre (GRDC): a dataset of river discharge draining into the Baltic Sea

Simon A. Mischel*, Claudia Zentis and Henning Plessow

Global Runoff Data Centre (GRDC), Federal Institute of Hydrology, Koblenz, Germany

KEYWORDS

baltic sea, river discharge, catchment areas, time series, *in situ* stations

1 Introduction

Due to its separation from the Atlantic Ocean, the Baltic Sea is a special marine environment, with a unique brackish water regime (Meier et al., 2022). Its salinity setting is an important driver for the biogeochemical functioning and the ecological status of the Baltic Sea (Kuliński et al., 2022). Global and climate change, however, alter the hydrological cycle including river discharge, which is responsible for most of the freshwater input into the Baltic Sea (Meier et al., 2022).

One of the major studies in the Baltic Sea was the Baltic Sea Experiment (BALTEX) as part of the Global Energy and Water Exchanges project (GEWEX), initiated in the 1980s under the World Climate Research Programme (WCRP). A key objective of GEWEX was to improve the understanding of heat exchange, energy, and water processes within the Earth's climate system at global, regional, and local scales (Raschke et al., 2001; Reckermann et al., 2011). For this purpose, the first river discharge dataset of BALTEX was compiled and made available by the Swedish Meteorological and Hydrological Institute (SMHI) through the BALTEX Hydrological Data Centre (BHDC, BALTEX, 1994). In 2015, the BALTEX database, comprising data from 592 monitoring stations, was transferred to the Global Runoff Data Centre (GRDC) to ensure long-term support and regular updates as part of GRDC's ongoing responsibilities providing river discharge time series, even beyond the active phase of BALTEX (e.g., Phase I, 1993-2002 and Phase II, 2003-2012).

The aim of this data report is to provide an overview about the BALTEX dataset and its extension in order to provide a complete as possible hydrological river discharge database for the Baltic Sea. Similar databases have been collected by Mikulski (1970), who reported the inflow of river water to the Baltic Sea in the period 1951-1960. A more comprehensive database was compiled by Bergström and Carlsson (1994), who provide more than 200 stations with monthly data draining into the Baltic Sea, covering the period of 1950–1990 and around 86% of the total drainage area. They also analyzed the major 10 rivers and divided the data into six major sub-basins of the Baltic Sea. We compare the mean discharges for 30 large rivers in the research area with values presented in the Baltic Sea Model Intercomparison Project (BMIP, Väli et al., 2019).

The data collected and the update provided by GRDC is useful for estimating and analyzing the various components of the water balance in the Baltic Sea, particularly the runoff and discharge contributions from surrounding landmasses. Large-scale hydrological models of river flow to the Baltic Sea exist and the data contributed to advances in research on the regional hydrology of the Baltic Sea, as well as developing different climate change scenarios of the Baltic Sea (Reckermann et al., 2011). With the BALTEX dataset models like E-HYPE (Donnelly et al., 2016) can be further verified and evaluated.

10.3389/fenvs.2025.1599188

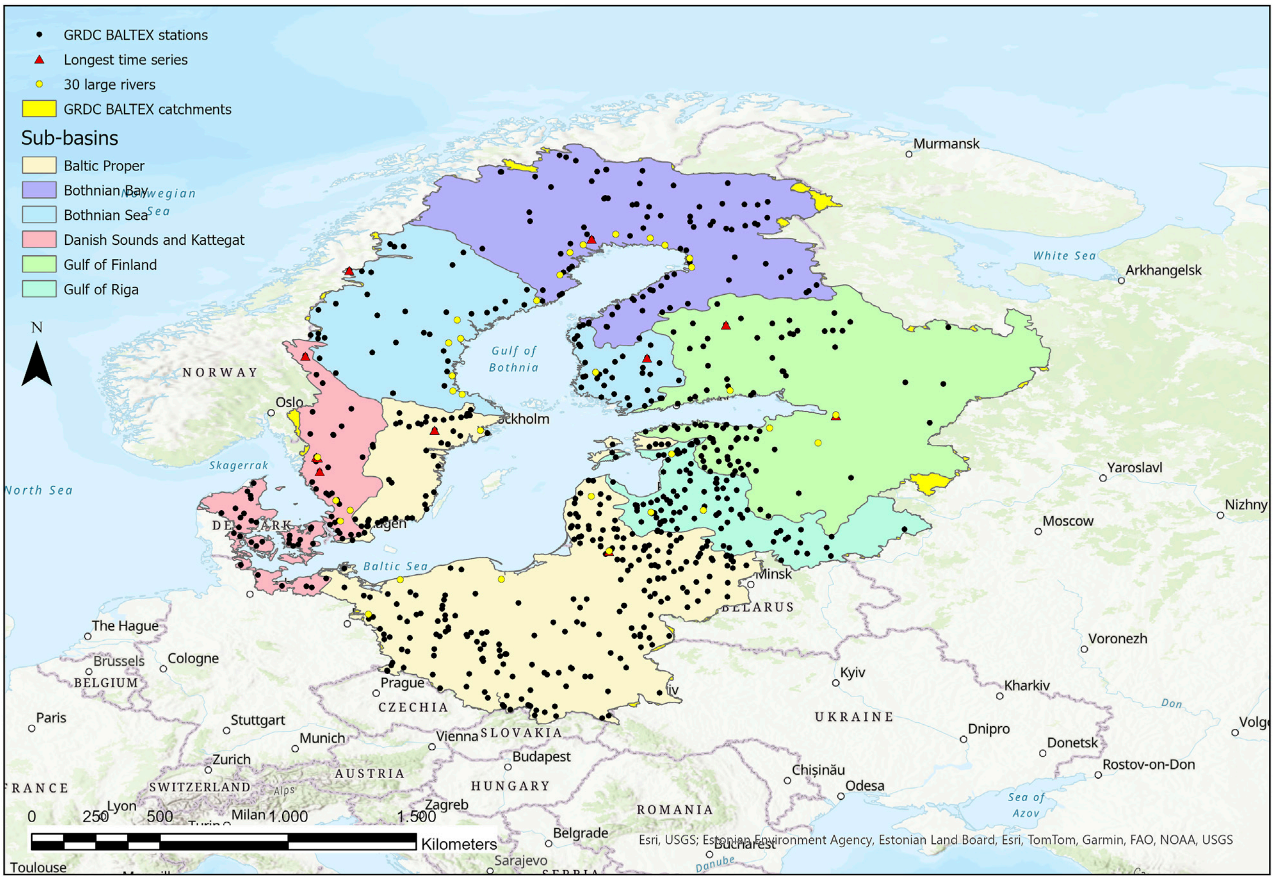

FIGURE 1
Map showing the catchment of the Baltic Sea and its sub-basins (see Table 1 for more details). Black dots mark the stations of the BALTEX dataset of GRDC, the red triangles mark stations with time series before 1900 and yellow dots represent 30 large rivers (see Table 2 for more details).

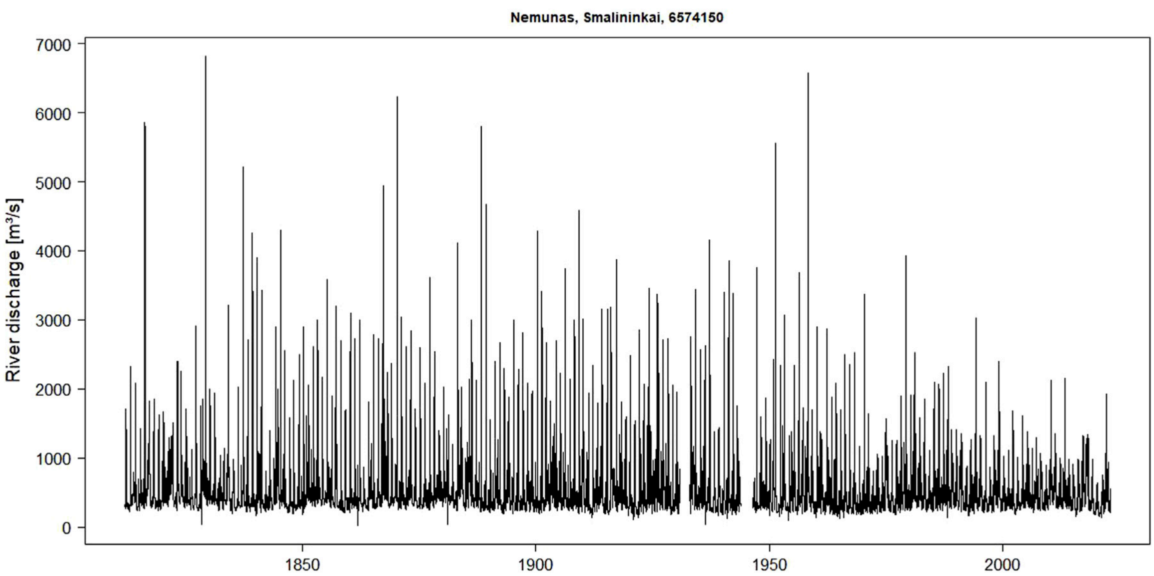

FIGURE 2
Daily time series of the Nemunas River, Smalininkai Station, GRDC No. 6574150, Lithuania. This is the longest available time series of daily data of the BALTEX dataset.

TABLE 1 Area of the sub-basins covered by the BALTEX dataset and average time series length.

Sub-basin	No. of stations	Area [km^2]	Station catchment coverage [%]	Avg. daily time series length [yrs]	Avg. monthly time series length [yrs]	Earliest	Latest
Baltic Proper	287	569,076	85	48	46	1812	2024
Bothnian Bay	95	265,740	89	69	63	1899	2024
Bothnian Sea	85	224,343	84	70	63	1863	2024
Gulf of Finland	96	133,701	90	55	42	1859	2024
Gulf of Riga	66	122,243	62	36	16	1906	2023
Danish Sounds and Kattegat	89	421,916	91	74	62	1807	2024
Total	718	1,737,019	84	59	49	1807	2024

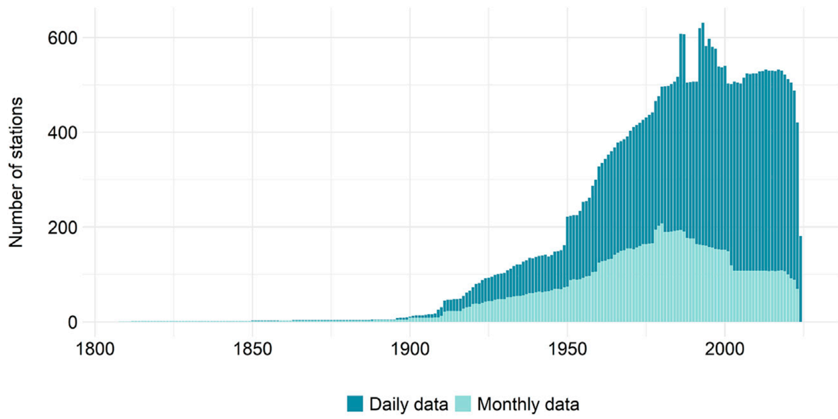

FIGURE 3
The number of stations in the BALTEX dataset, separated into daily and monthly time series.

2 Methods

2.1 The Baltic Sea

The Baltic Sea has a total surface area of 377,400 km^2 with a volume of 21,200 km^3 (Figure 1). The five major sub-basins are the Bothnian Bay, the Bothnian Sea, the Gulf of Finland, the Gulf of Riga and the Baltic Proper. Following the definition of the Baltic Marine Environment Protection Commission (HELCOM, 2014) and Bergström and Carlson (1994), stations draining to the Danish Sound and Kattegat are included due to its significance for the water exchange of the Baltic Sea. Depending on the literature the total drainage area varies between 1,729,000 and 1,739,000 km^2 (Bergström and Carlsson, 1994; Nilsson, 2006). The Baltic Sea basin stretches over 14 countries and different climatic zones, from subarctic climate in northern Finland to the temperate, continental climate in southern Poland (Raschke et al., 2001). During winter, the Baltic Sea freezes to various extend (HELCOM, 2018). At the Baltic drainage basin, river discharge is culminating at the Danish straits, with about 450 km^3 per year, while 50% of the discharge is measured during the melting season only.

2.2 Data collection

The GRDC[1], operating under the auspices of the World Meteorological Organization (WMO), has the aim of supporting and promoting research in the field of global and climate change as well as integrated water resources management. GRDC provides *in situ* river discharge time series, with daily or monthly resolution. To date, the total number of available stations in the database of GRDC has increased to nearly 11,000.

Since the integration of the BALTEX database into GRDC in 2015, the curation and update for the BALTEX dataset is embedded in the regular work of GRDC. Most of data is obtained from the National Meteorological and Hydrological Services (NMHS) of the WMO member states. The main quality-check is the responsibility of the NMHS, who provide the data. GRDC performs an outlier test and a visual quality-check prior to the inclusion into the GRDC database (Färber et al., 2024). If the NMHS provide only daily data, the mean

1 https://grdc.bafg.de/

TABLE 2 Comparison of the mean discharge reported for BMIP and BHDC (Väli et al., 2019) and the BALTEX dataset for 1961-1990. In addition, the mean discharge for the complete available time for each individual station of the BALTEX dataset is provided.

River Name	GRDC Station Name	GRDC Nr	BMIP	BHDC	BALTEX	Complete BALTEX dataset
				mean discharge [m^3/s]		
Angermanalven	Solleftea KRV	6233650	460.1	511.1	483	487.3
Dalalven	Aelvkarleby KRV	6233201			353.4	351.9
Gavlean	Gavle	6233112			20.2	20
Dalalven & Gavlean			440.1	396.7	373.6	371.9
Daugava	Jekabpils	6373304	786.7	637.1	487.3	517.2
Gota Alv	Vargoens KRV	6233315	608.9	577.5	531.8	531
Iijoki	Raasakka (Merikoski)	6854600	189.9	168.1	143.8	173.8
Indalsalven	Bergeforsens KRV	6233401			437	447.2
Ljungan	Skallboele KRV	6233551			123.2	125.3
Indalsalven & Ljungan			569.7	598.1	560.2	572.5
Lielupe	Mezotne	6373400	99.5	127.6	49.7	56.9
Kalixalven	Raektfors	6233850			283.3	294.2
Tornealven	Kukkolankoski Oevre	6232910			387	393.8
Kalixalven & Tornealven			709.5	700.9	670.3	688
Kemijoki	Isohaara	6854700	610.3	620.6	552.9	571.2
Kokemaeenjoki	Harjavalta	6854101	299.9	240.1	231.3	219
Kymijoki	Anjala	6854203	355.7	308.3	300	284
Lagan	Aengabaecks KRV	6233170			62.5	62.4
Ronne A	Forsmollan	6233610			107	11
Nissan	Nissastrom	6233311				41.4
Lagan & Ronne A & Nissan			144.2	164.3	73.2	114.8
Limfjord			74.7	154.6		
Ljusnan	Ljusne Stroemmar KRV	6233221	277.1	242.4	226.9	227
Luga	Tolmachevo	6972400	131.6	107.4	50.3	50.3
Lulealven	Bodens KRV	6233750			482	509.7
Raneaelven						NA
Lulealven & Raneaelven			478.6	549.8	482	509.7
Narva Jogi	Narva Linn	6172350	478.6	363.7	364	397.3
Nemunas	Smalininkai	6574150	753.9	613.6	506.9	527.6
Neva	Novosaratovka	6955430	2421.7	2417	2320.5	2464.9
Norrstrom	Oevre Stockholm	6233410	254.2	176.9	162	174.4
Oder River	Hohensaaten-Finow	6357010	651.6	561.5	555.6	511.7
Oulujoki	Merikoski	6854500	338.6	259.6	259.1	262.9
Parnu Jogi	Oore	6172050	62.9	129.6	49.3	49.3
Parseta						NA
Rega	Trzebiatow	6456107			21.4	19.7
Parseta & Rega			60.6	127.7	21.4	19.7

(Continued on following page)

TABLE 2 (*Continued*) Comparison of the mean discharge reported for BMIP and BHDC (Väli et al., 2019) and the BALTEX dataset for 1961-1990. In addition, the mean discharge for the complete available time for each individual station of the BALTEX dataset is provided.

River Name	GRDC Station Name	GRDC Nr	BMIP	BHDC	BALTEX	Complete BALTEX dataset
				mean discharge [m^3/s]		
Pitealven	Sikfors KRV	6233710	185.7	188.6	158.3	161.2
Pregolia			143.1	156.7		
Skelleftealven	Kvistforsens KRV	6233690	163.7	172.3	151.2	160.9
Umealven	Stornorrfors KRV	6233502			436.9	443.4
Saevarn						NA
Umealven & Saevarn			485.4	462.1	436.9	443.4
Venta	Kuldiga	6373010	133.6	171.2	69.5	68.3
Vistula	Tczew	6458010	1078.6	1128.5	1093.1	1026.5

monthly values are calculated by GRDC. Routinely, the data is requested each year. Where possible, data provision service via application programming interfaces (API) are used. With this approach, the number of stations in the Baltic Sea catchment area of the BALTEX dataset was expanded by 126 stations, receiving regular, up-to-date time series updates. Since not all countries have productive APIs, data provision is partly still delayed between 2 and 5 years, due to the time-consuming verification process. As soon as the data is available to GRDC, the data is provided to the user. An example of a daily time series available for download is presented in Figure 2.

The time series data are available in various formats from the GRDC data portal[2]. The user can request the data for each station individually or by downloading the whole BALTEX dataset using the BALTEX catalogue provided at the GRDC website[3] or using the metadata catalogue provided in the supplementary material (Supplementary Table S1). Together with the time series data, extensive metadata are attached to the files and catchment polygons for each of the stations are available. The BALTEX dataset is published under the GRDC data policy[4] and WMO Resolution 25 (Cg-XIII-1999). Individual data licenses are reported in Supplementary Table S2.

To allocate the GRDC stations to the Baltic Sea catchment, the HELCOM GIS layer[4] available at the HELCOM metadata catalogue, was used, which includes the Baltic Sea's sub-basins drainage areas, aligning with Bergström and Carlsson (1994), resulting in a total of 718 stations.

3 Dataset description

Figure 1 shows the location of all GRDC stations that are included in the BALTEX dataset. Table 1 provides an overview of the number of stations in each sub-basin and their spatial and temporal coverage. There are 718 stations distributed across all sub-basins, which cover a total area of 1,737,019 km^2 and fall within the range of the drainage area of the Baltic Sea mentioned above (see Chapter 2.1). The stations are equally distributed over the entire sub-basins upon availability. The GRDC station catchments cover in total 84% of the HELCOM basin layer. Important to note is that coastal areas are largely excluded and catchment areas are calculated for the location of the station.

One limiting factor for the spatial coverage of the dataset is that GRDC avoids to include stations that are influenced by tides. Therefore, stations directly at the river mouth are excluded, when they contain negative discharge or unreliable discharge values. Another factor is the inability to determine the station catchment area for 34 of the 718 stations, as certain catchments cannot be accurately represented by the elevation models used for catchment delineation. The methodology used to generate the GRDC station catchments is described in Färber et al. (2024).

Table 1 shows that the Baltic Proper has the highest station density, with a total of 287 stations, while in contrast, the Gulf of Riga has only 66 stations. All sub-basins have time series data available until 2024, except for the Gulf of Riga, which has records up to 2023. The oldest time series is in service since 1807, with a total of ten stations in the whole Baltic Sea catchment with data before 1900. The average daily time series length for all sub-basins is 59 years, while the average monthly time series length is 49 years. Therefore, most of the data covers the BALTEX active phase and beyond.

Figure 3 illustrates the availability of mean daily and monthly averages, represented in blue and light blue, respectively. As shown, a few time series reach back to the 19th century. These exceptionally long time series are also geographically well distributed (Figure 1), as indicated in the penultimate column of Table 1. Overall, the number of available time series increases until the mid-20th century, with a noticeable increase in the 1950s, largely due to the incorporation and data provision for many Polish time series during this period. Data availability peaks towards the late 20th century but steadily declines after 2000. The two observed peaks are due to time series from

2 https://grdc.bafg.de/data/data_portal/

3 https://grdc.bafg.de/data/project_datasets/baltex/

4 https://grdc.bafg.de/about/data_policy/

Lithuania and Belarus, covering short periods of 2 and 6 years, respectively.

4 Conclusion

The BALTEX database provides a viable source for river discharge data in the catchment area of the Baltic Sea. In the BALTEX dataset all available data is provided, even small catchments and short time series are included. In total, there are 718 stations with daily and monthly data available, ranging between 1 and 215 years, the catchment size ranges vary from sub-square kilometer up to 218,000 km^2. Monthly time series are separated between original time series and calculated time series. The latter is calculated using daily averages. All of the sub-basins draining into the Baltic Sea are represented with different numbers of total stations. To date, this dataset is the most comprehensive dataset of river discharge data draining into the Baltic Sea. It is freely and unrestricted available.

Comparing the BALTEX dataset with available data is shown in Table 2. The datasets are in good agreement. As the original datasets of BMIP and BHDC are not available, small differences might occur due to a different selection of stations.

Using the values from the BALTEX dataset and the 30 longest rivers allows the calculation of loads when combining the databases from other datasets, for example, from the Global Freshwater Quality Database GEMStat[5,6].

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://grdc.bafg.de/data/project_datasets/baltex/.

Author contributions

SAM: Validation, Data curation, Supervision, Methodology, Conceptualization, Writing – review and editing, Investigation, Resources, Writing – original draft, Visualization. CZ: Visualization, Writing – review and editing, Conceptualization, Writing – original draft. HP: Formal Analysis, Writing – review and editing, Software, Writing – original draft, Data curation.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. GRDC is funded by the Federal Institute of Hydrology, Germany (BfG).

Acknowledgments

GRDC thanks the World Meteorological Organization (WMO) and the Federal Institute of Hydrology, Germany (BfG) for support. The NMHS are acknowledged for providing their valuable data sets.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenvs.2025.1599188/full#supplementary-material

References

BALTEX (1994). *Scientific plan for the Baltic sea experiment (BALTEX)* ed. *E. Raschke* Geesthacht: GKSS Research Centre Geesthacht.

Bergström, S., and Carlsson, B. (1994). River runoff to the Baltic sea: 1950-1990. *Ambio* 23 (Nr. 4-5), 280–287. Available online at: http://www.jstor.org/stable/4314220

Donnelly, C., Anderson, J. C. M., and Arheimer, B. (2016). Using flow signatures and catchment similarities to evaluate the E-HYPE multi-basin model across Europe. *Hyd Sci. J.* 61, 2. doi:10.1080/02626667.2015.1027710

Färber, C., Plessow, H., Mischel, S., Kratzert, F., Addor, N., Shalev, G., et al. (2024). GRDC-Caravan: extending caravan with data from the global runoff data centre. *Earth Syst. Sci. Data Discuss.* doi:10.5194/essd-2024-427

HELCOM (2014). Convention on the protection of the marine environment of He Baltic sea are, 1992 (helsinki convention), HELCOM. Available online at: https://helcom.fi/wp-content/uploads/2019/06/Helsinki-Convention_July-2014.pdf (Accessed March 21, 2025).

HELCOM (2018). State of the Baltic sea – second HELCOM holistic assessment 2011-2016. *Balt. Sea Environ. Proc.* 155.

5 https://metadata.helcom.fi/geonetwork/srv/eng/catalog.search#/metadata/9e132cda-8f69-4f7a-9c66-8b9859e61441

6 GEMStat - The global water quality database

Kuliński, K., Rehder, G., Asmala, E., Bartosova, A., Carstensen, J., Gustafsson, B., et al. (2022). Biogeochemical functioning of the Baltic sea. *Earth Syst. Dynam.* 13, 633–685. doi:10.5194/esd-13-633-2022

Meier, H. E. M., Kniebusch, M., Dieterich, C., Gröger, M., Zorita, E., Elmgren, R., et al. (2022). Climate change in the Baltic sea region: a summary. *Earth Syst. Dynam.* 13, 457–593. doi:10.5194/esd-13-457-2022

Mikulski, Z. (1970). Inflow of river water to the Baltic sea in the period 1951-1960. *Nord. Hydrol.* 4, 216–227. doi:10.2166/nh.1970.0014

Nilsson, S. (2006). "International river basins in the Baltic sea region,", III B. EU BSR INTERREG. Stockholm, Schweden: KTH, Royal Insitute of Technology.

Raschke, E., Meywerk, J., Warrach, K., Andrea, U., Bergström, S., Beyrich, F., et al. (2001). The Baltic sea experiment (BALTEX): a European contribution to the investigation of the energy and water cycle over a large drainage basin. *Bull. Am. Meteorological Soc.* 82, 2389–2413. doi:10.1175/1520-0477(2001)082<2389:tbseba>2.3.co;2

Reckermann, M., Langner, J., Omstedt, A., von Storch, H., Keevallik, S., Schneider, B., et al. (2011). BALTEX – an interdisciplinary research network for the Baltic sea region. *Environ. Res. Lett.* 6, 045205. doi:10.1088/1748-9326/6/4/045205

Väli, G., Meier, H. E. M., Placke, M., and Dieterich, C. (2019). River runoff forcing fr ocean modeling within the Baltic sea model intercomparison project. *Mar. Sci. Rep.* (113).