**Dataset underlying the publication “Flood characteristics drive river-scale macroplastic deposition”**

Louise J. Schreyers1, Rahel Hauk1, Nicholas Wallerstein1, Adriaan J. Teuling1, Remko Uijlenhoet1,2, Martine van der Ploeg1 and Tim H.M. van Emmerik1

1Wageningen University, Hydrology and Environmental Hydraulics Group, Wageningen University, 6708 PB, Wageningen, The Netherlands.

2Delft University of Technology, Department of Water Management, Delft University of Technology, 2628 CN, Delft, The Netherlands.

Contacts: [l.schreyers@gmail.com](mailto:l.schreyers@gmail.com) and [tim.vanemmerik@wur.nl](mailto:tim.vanemmerik@wur.nl).

The publication will be made openly available at the following link: 10.1021/acs.est.5c02969.

**Introductory information**

This dataset presents the data used in the publication “Flood characteristics drive river-scale macroplastic deposition”. This dataset is on macroplastic and other macrolitter sampled on floodplains along the IJssel and the Meuse rivers in the Netherlands. We sampled 23 locations between 25 January 2024 - 14 February 2024, following a winter flood event along the IJssel river. We include the observations from a sampling campaign conducted along the Meuse river, which sampled 25 locations between 22 July and 4 August 2021 (Hauk et al., 2023). In addition, we include observations conducted by the Schone Rivieren program, for lower flood magnitude and low-flood conditions, along the IJssel and the Meuse rivers. For the Meuse river, these include seven sampling campaigns, from fall 2018 to fall 2022. For the IJssel river, five sampling campaigns are included, from fall 2020 to fall 2023.

The dataset is provided both in an csv format (10 files) and in Excel format (10 datasheets). Four main types of datasheets are presented:

1. **Raw data.** The datasheet “RawDataIJsselWinter2024” presents all the raw, unprocessed data for the IJssel winter 2024 flood event.
2. **Input/Training data.** The dataset consists of four datasheets containing the training data used to model macroplastic concentrations along the floodplains of the Meuse and IJssel rivers. Two of these datasheets correspond to higher-magnitude flood events: “InputDataMeuseSummer2021” and “InputDataIJsselWinter2024”. These events have return periods of 111 years and 2.9 years, respectively. The remaining two datasheets, “InputDataMeuseAllOtherEvents” and “InputDataIJsselAllOtherEvents”, cover lower-magnitude floods and non-flood conditions. Each datasheet provides observed macroplastic quantities (expressed as item counts, total mass, item concentrations, and mass concentrations) together with explanatory variables. These include characteristics of the floodplain and river at sampled locations, as well as the distance to potential macroplastic sources.
3. **Model results**. Three datasheets present the model results. One datasheet (“ModelResultsCoefficients”) presents all the model formulations, with the coefficients per variables, as well as R2 and AIC scores per model. This table is the digital version of Table S2 in the publication. In addition, two datasheets present the detailed results, with modeled macroplastic concentrations per each 100 m of floodplain along the rivers IJssel and Meuse (respectively, “ModelResultIJsselWitner2024” and “ModelResultMeuseSummer2021”). We present the modeled macroplastic concentrations using the two best performing models – namely, ‘1j’ for the Meuse summer 2021 flood event and ‘10c’ for the IJssel winter 2024 flood event. We do not include the detailed results for other flood events, as they were not extensively used in the publication due to low performance (R2 < 0.5).
4. **Discharge records**. Two datasheets present river discharge data, obtained from Rijkswaterstaat (the Dutch Ministry for Water and Infrastructure). One datasheet (“SintPieterDischargeMeuse”) presents discharge historical records for the Meuse river, at Sint-Pieter gauging station from January 1991 to May 2024. The remaining datasheet (“OlstDischargeMeuse”) presents the discharge record for the IJssel river, at Olst gauging station from January 2017 to May 2024. The discharge records were resampled at 2-hours resolution. Higher-resolution temporal discharge records can directly be downloaded at: https://waterinfo.rws.nl/.

**Additional information**

1. **Raw data. T**he datasheet “RawDataIJsselWinter2024” contains raw data collected during the IJssel winter 2024 flood event. Each row generally represents a single observed macrolitter item, along with contextual information about the environment in which it was found (see *Feature type* column) and the item’s characteristics (see *OSPAR ID code* column). The table below describes each column in the dataset. Note that in the publication associated to this dataset, we did not analyze all elements related to this data, such as the macrolitter type and in which type of elements they were found. This is left for further analysis.

**Table 1.** Description of columns in datasheet “RawDataIJsselWinter2024”.

|  |  |
| --- | --- |
| **Column name** | **Description** |
| Schone Rivieren code | Unique identifier for Schone Rivieren location |
| Observation type | Describes the type of observation recorded.   * Transect: Regular measurement, usually a 2 m wide strip taken perpendicular from the waterline towards the highest floodline. * Water\_line: Measurement taken along the current waterline. * Extra or other: used interchangeably for additional measurements outside transects, e.g., debris piles or trees showing signs of inundation. |
| Distance transect to SR point | Distance from transect [meters] to the main Schone Rivieren coordinates |
| Position transect | Relative orientation of the transect with respect to the Schone Rivieren coordinates |
| Longitude | Longitude coordinate of Schone Rivieren location in decimal degrees |
| Latitude | Latitude coordinate of Schone Rivieren location in decimal degrees |
| Transect ID | Unique ID for the transect. Transects were typically 2 m wide (parallel to river course) and varied in length depending on floodplain extent. |
| Section ID | Unique ID for each section within a transect. Sections were 2 m wide and varied in length: 3 m when covering mixed features (see *Feature type* column), and up to tens of meters when coverage was uniform. |
| Survey type | Type of survey recorded. Certain survey belong to regular transects, other were extra measurements when interesting accumulation zones were found by the surveyors, but were located outside of the regular sampled transects. |
| Comments | Notes for individual observation |
| Comments entire section | Notes for the full section |
| Feature type | Type of feature (i.e. landcover) recorded (e.g.: vegetation elements, debris pile, bare grass). |
| Distance start section [m] | Start distance along the section [meters] from the current waterline |
| Distance end section [m] | End distance along the section [meters] from the current waterline |
| Width section [m] | Width of the section [meters] |
| Area covered by grass [m²] | Area covered by grass in the section [m²] |
| Area covered by trees [m²] | Area covered by trees in the section [m²] |
| Inundation signs | Evidence of flooding/inundation in the trees |
| Highest tree inundation mark [m] | Maximum height of water mark on trees [m] in the section (from ground) |
| Lowest tree inundation mark [m] | Minimum height of water mark on trees [m] in the section (from ground) |
| Area covered by shrubs [m²] | Area covered by shrubs in the section [m²] |
| Highest shrub height [m] | Maximum height of shrubs [m] in the section |
| Minimum shrub height [m] | Minimum height of shrubs [m] in the section |
| Area covered by bushes [m²] | Area covered by bushes in the section [m²] |
| Maximum bush height [m] | Maximum bush height [m] in the section |
| Minimum bush height [m] | Minimum bush height [m] in the section |
| Area covered by debris pile [m²] | Area covered by debris pile in the section [m²] |
| Maximum height debris pile [cm] | Maximum height of the debris pile [cm] in the section |
| Minimum height debris pile [cm] | Minimum height of the debris pile [cm] in the section |
| Length extra measurement [m] | Length of extra measurement [m], if any |
| Width extra measurement [m] | Width of extra measurement [m], if any |
| Height extra measurement [cm] | Height of extra measurement [cm], if any |
| Distance along section where item was found [m] | For sections of length > 3 m, the distance from the start of the transect (i.e. waterline) at which the items were found was noted. |
| Count macrolitter [#] | Number of macrolitter items observed [#] (always one, because each row refers to an individual macrolitter item) |
| Comments processing | Notes on data processing |
| Total feature area [m²] | Total area of feature [m²] in section |
| Actual sampled feature area [m²] | Area of the feature actually sampled within the section [m²]. This value may differ from the *Feature area [m²]* column. In cases where macrolitter quantities were visibly very high, surveyors intentionally sampled a smaller area to reduce sampling time. This was done especially for debris piles. |
| Actual sampled height [cm] | Height of the feature actually sampled within the section [cm]. This column typically applies only to debris piles, where a vertical dimension was relevant. |
| Adjusted macrollitter count [#] | Adjusted number of macrolitter items [#]. This value can differ from that in *Count macrolitter [#]* column. When the actual sampled area differed from the total feature area in the section, an extrapolation factor was applied to reflect the adjusted macrolitter count. |
| OSPAR ID code | Unique ID for OSPAR classification for macrolitter. More details on the OSPAR classification can be found at van Emmerik et al., 2020. |
| Mean mass for OSPAR category [g] | Average mass of items in the corresponding OSPAR category [g]. Macrolitter items in this sampling campaign were not weighed directly; instead, mean masses were derived from de Lange et al., 2023. |
| Median mass for OSPAR category [g] | Median mass of items in the corresponding OSPAR category [g]. Macrolitter items in this sampling campaign were not weighed directly; instead, mean masses were derived from de Lange et al., 2023. |
| Min mass for OSPAR category [g] | Minimum mass of items in the corresponding OSPAR category [g]. Macrolitter items in this sampling campaign were not weighed directly; instead, mean masses were derived from de Lange et al., 2023. |
| Max mass for OSPAR category [g] | Maximum mass of items in the corresponding OSPAR category [g]. Macrolitter items in this sampling campaign were not weighed directly; instead, mean masses were derived from de Lange et al., 2023. |
| Mean width for OSPAR category [cm] | Average width of items in the corresponding OSPAR category [g]. Macrolitter item size in this sampling campaign were not measured directly; instead, mean sizes were derived from de Lange et al., 2023. |
| Median width for OSPAR category [cm] | Median width of items in the corresponding OSPAR category [g]. Macrolitter item size in this sampling campaign were not measured directly; instead, mean sizes were derived from de Lange et al., 2023. |
| Min width for OSPAR category [cm] | Minimum width of items in the corresponding OSPAR category [g]. Macrolitter item size in this sampling campaign were not measured directly; instead, mean sizes were derived from de Lange et al., 2023. |
| Max width for OSPAR category [cm] | Maximum width of items in the corresponding OSPAR category [g]. Macrolitter item size in this sampling campaign were not measured directly; instead, mean sizes were derived from de Lange et al., 2023. |
| Mean length for OSPAR category [cm] | Average length of items in the corresponding OSPAR category [g]. Macrolitter item size in this sampling campaign were not measured directly; instead, mean sizes were derived from de Lange et al., 2023. |
| Median length for OSPAR category [cm] | Median length of items in the corresponding OSPAR category [g]. Macrolitter item size in this sampling campaign were not measured directly; instead, mean sizes were derived from de Lange et al., 2023. |
| Min length for OSPAR category [cm] | Minimum length of items in the corresponding OSPAR category [g]. Macrolitter item size in this sampling campaign were not measured directly; instead, mean sizes were derived from de Lange et al., 2023. |
| Max length for OSPAR category [cm] | Minimum length of items in the corresponding OSPAR category [g]. Macrolitter item size in this sampling campaign were not measured directly; instead, mean sizes were derived from de Lange et al., 2023. |
| Largest dimension (length or width) of item [cm] | Largest dimension (length or width) of item [cm] |
| Material | Type of material of the item (plastic, metal, glass, etc.) |

The raw dataset for the Meuse summer 2021 flood event can be found at Hauk et al., 2023b. The raw datasets for all other events are owned by the Schone Rivieren program and are therefore made available here only in processed form.

1. **Input/training data.** The two datasheets (“InputDataMeuseSummer2021” and “InputDataIJsselWinter2024”) contain rows of data, each representing a sampling location covering up to 100 m of river length. If sampling at the same location extended beyond 100 m, an additional row was added. As a result, a single Schone Rivieren code may appear in multiple rows, meaning that each row does not necessarily correspond to a unique code. This occurs because multiple sampling transects were usually conducted per Schone Rivieren location, often dictated by time constraints and accessibility limitations. Consequently, some transects may belong to one 100 m river section, while others belong to another. Additional details about the sampling locations are provided in columns D:I (IJssel) and D:L (Meuse), including the transects’ positions relative to the original Schone Rivieren coordinates and their distances in meters. Schone Rivieren codes are unique identifiers used by the Schone Rivieren program (English: Clean Rivers). Although the macroplastic data for the IJssel (winter 2024) and the Meuse (summer 2021) were collected directly by the authors rather than through the Schone Rivieren program, we use these codes to facilitate comparison. This is because Schone Rivieren routinely monitors the riverbanks and floodplains of Dutch waterways.

In addition, these two datasheets present ten variables corresponding to key characteristics of riverbanks and floodplains, as well as potential macroplastic sources (columns: J:S and L:U, for the IJssel and Meuse, respectively). We refer to the section “Predictive Modeling of Macroplastic Concentrations” of the publication for a detailed description of these variables, and their sources.

Finally, these two datasheets presented the total macroplastic (expressed in total item count and total mass), area considered and macroplastic concentrations (expressed in item and mass concentrations) for each sampling location.

1. **Model results***.* In “ModelResultIJsselWinter2024” and “ModelResultMeuseSummer2021” we present the model for two formulations applied to the IJssel (winter 2024) and Meuse (summer 2021) flood events. These correspond to model formulations ‘1j’ and ‘10c’ in the datasheet “ModelResultsCoefficients”, where the specific coefficients used for each input variables can be found. The input variables used to model macroplastic concentrations are listed in columns C–L. In both formulations, however, the variables in columns K–L were ultimately not included in these models, as they did not improve performance. These variables are nonetheless provided in the datasheets to allow replication of other model formulations (see “ModelResultsCoefficients”).

Note that macroplastic concentrations were modeled for each 100 m section of river length along the two river systems, and for both left and right floodplains (column B). Column A indicated either “no\_obs” if no macroplastic sampling was done (so no observational values were available) or the Schone Rivieren code, when sampling was conducted.

The datasheet “ModelResultsCoefficients” presents all the model formulations, with the coefficients per variable, as well as R2 and AIC scores per model. The type of transformation is indicated, “linear” corresponding to linear terms, and “negative exponential” to negative decay terms. For example, for model 1a, the coefficient for distance (4.8) corresponds to an exponential decay term exp(−distance).

**Sharing and access information**

No restrictions are placed on the use of this data. License: CC-BY. This license allows users to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for commercial use.

**References**

De Lange, S., Mellink, Y., Vriend, P., Tasseron, P., & van Emmerik, T. (2022). Replication Dataset for "Sample size requirements for riverbank macrolitter characterization" (Version 1) [Data set]. 4TU.ResearchData. <https://doi.org/10.4121/19188131.V1>

Hauk, R., van Emmerik, T. H. M. (T., M.J. (Martine) Van Der Ploeg, de Winter, W., Boonstra, M., Löhr, A. J., & A.J. (Adriaan) Teuling. (2023). Data underlying the publication: Macroplastic deposition and flushing in the Meuse river following the July 2021 European floods (Version 1) [Data set]. 4TU.ResearchData. <https://doi.org/10.4121/21360531.V1>

Van Emmerik, T., Roebroek, C., De Winter, W., Vriend, P., Boonstra, M., & Hougee, M. (2020). Riverbank macrolitter in the Dutch Rhine–Meuse delta. *Environmental Research Letters*, *15*(10), 104087.