Use of machine learning and deep learning to predict particulate 137Cs concentrations in a nuclearized river

Hugo Lepage, Valerie Nicoulaud Gouin, Kathleen Pele, Patrick Boyer

Use of machine learning and deep learning to predict particulate ^{137}Cs concentrations in a nuclearized river

Lepage Hugo[1,*], Nicoulaud-Gouin Valérie [1], Pele Kathleen [1], Boyer Patrick [1]

[1] Institut de Radioprotection et de Sûreté Nucléaire (IRSN), PSE-ENV/SRTE/LRTA, F-13115, Saint-Paul-lez-Durance, France

* Corresponding author: hugo.lepage@irsn.fr

Abstract

Cesium-137, discharged by nuclear installations under normal operations and deposited in watersheds following atmospheric testing and accidents (i.e. Chernobyl, Fukushima…), has been studied for decades. Thus, modelling of ^{137}Cs concentration in rivers have been developed based on geochemical approaches and equilibrium assumptions (solid/liquid ratio) as this radionuclide has moved into rivers and oceans due to soil erosion. Recently a new approach is possible to model these concentrations with the popularization of data-driven models based on data acquired in the environment by monitoring networks. In this study, the concentrations of particulate cesium-137 measured near the mouth of the Rhône River (France), a highly nuclearized river, are simulated using two data-driven models, a Hierarchical Attention-Based Recurrent Highway Networks (HRHN) and a Random Forest Regressor (RF). The data-driven predictions were done using only hydrological data (water discharge and suspended solid fluxes) and industrial input of ^{137}Cs. Although the data-driven models provided a better prediction than a recent empirical model, the best prediction (R^2 = 0.71) was obtained with HRHN, a model that considers the temporal aspect of the monitoring data. The most important predictors were the hydrological data at the monitoring station and of the tributary that generate the most sediment flux (Durance River). In fact, the concentration of ^{137}Cs in the perimeter of this study was more related to hydrology than to nuclear release, as there were few events with high ^{137}Cs concentrations (concomitant nuclear release and low water discharge). However, the HRHN approach, which is more complex to implement than RF, can predict the concentrations of such events correctly despite their low representation of these events. The results of this study demonstrate the usefulness of data-driven models to assist monitoring programs by filling in gaps or helping to understand observed concentrations.

Use of machine learning and deep learning to predict particulate ^{137}Cs concentrations in a nuclearized river

Keywords: radioactivity, suspended sediment, flux, neural network, modeling

1. Introduction

Monitoring the pollution in rivers has been an accepted necessity for many years as aquatic ecosystems are unfortunately the receptors of emissions from many sectors such as agriculture, industry or urbanization (Horowitz, 2009; Syvitski et al., 2005; WFD, 2000). For many reasons (logistical, economic, societal), it is sometimes difficult to set up a sustainable monitoring system (Dethier et al., 2020). Therefore it makes sense to model the concentrations and fluxes of pollutants (Yang and Wang, 2010) as modelling appear to be an alternative solution to improve water quality monitoring cost. Thus, many works have taken place during the last decades to set up empirical models or models based on geochemical processes to predict the concentration of pollutants such as heavy metals, mercury, nitrates or radionuclides (Braga et al., 2010; Hu et al., 2007; Zheleznyak et al., 2022; Zhu et al., 2018). However, these models are often based on restrictive assumptions and require parameters that are not always available (Cho et al., 2016; Ciffroy and Benedetti, 2018; Desai et al., 2013; Hilko et al., 2012; Lu et al., 2019). With the recent advances on computer calculations and artificial intelligence applications in many fields (Abiodun et al., 2018), the use of data-driven models bring a new tool to model the concentrations and the behavior of theses pollutants (Lu et al., 2019; Yaseen, 2021; Ye et al., 2020). The number of published works on heavy metals simulation using machine learning models greatly increased from less than 5 publications by year before 2011 to 49 publications in 2020 (Yaseen, 2021). Models such as Artificial Neural Network (ANN), classification and regression tree, linear regression or support vector machine models are now used to predict the concentration of pollutants in the environment.

However, their use to model radionuclides concentrations in water and suspended sediment (SS) of rivers remain sparse and new. In a recent review, Dragović (2022) highlights the potential use of ANN in environmental radioactivity but the studies regarding the hydrosphere are still sparse and mainly focus on water, not on sediment. For example, Korobitsyn et al. (2008) demonstrated that ANN could be used to understand the high concentrations of dissolved Strontium-90 observed in the Techa River (Russia). Kulahci et al. (2006) used an ANN construction to predict the total alpha and total beta radioactivity as a function of pH, total hardness, electrical conductivity and depth. Seasonal and spatial variability and classification of radionuclide concentrations might also be done using neural network such as self-organizing maps (Skwarzec et al., 2009; Tutu et al., 2005) and help improve monitoring programs. While other machine learning tools such as random forest (with predictive or causal techniques) had been recently used to analyze cesium-137 (^{137}Cs) contamination of terrestrial plants (Shuryak, 2023, 2022), no similar work was conducted on liquid or particulate concentration in rivers. Due to its anthropogenic presence in many environmental compartments, this radionuclide and its dynamics has been studied for decades, especially in waterways (Antonelli et al., 2008; He and Walling, 1996; Konoplev et al., 2020; Lepage et al., 2014). Thus, during the last decades, empirical modelling of ^{137}Cs concentration in rivers have been developed, based on assumption like the solid/liquid ratio or the size of the particles (Ciffroy et al., 2001; Konoplev et al., 2020; Tomczak et al., 2019). However, to our knowledge, the dynamic in rivers of this radionuclides was not yet subjected to a machine learning analysis. To pursue this effort of using artificial intelligence in radioecology, a machine learning and a deep learning approaches were used and compared to a semi-empirical model to estimate the ^{137}Cs particulate concentrations in a nuclearized river.

2. Material and method

2.1. Study area

This study was conducted near the outlet of the Rhône River basin (≈ 95 000km²) within the Rhône Sediment Observatory (Lepage et al., 2021; Fabien Thollet et al., 2021) (Fig. 1). This river is the largest

source of freshwater to the Mediterranean sea and deliver 6.6 Mt yr^{-1} of sediment each year (Poulier et al., 2019). The Rhône watershed is also characterized by a large climatic and geological heterogeneity (Olivier et al., 2022) that leads to strong variations of annual SS fluxes (from 1.4 Mt to 18.0 Mt) (Delile et al., 2020). Moreover this river is one of the most nuclearized river over the world with many nuclear facilities (Eyrolle et al., 2020) such as Nuclear Power Plant (NPP), fuel cycle facilities and research centers (Fig. 1). The basin has also been affected by the fallouts from the atmospheric weapons tests and the accident of Chernobyl (Roussel-Debel et al., 2007). Total inventory of ^{137}Cs was estimated to be 9.4 10^{14} Bq in 1986. This work focuses on the ^{137}Cs concentrations in SS measured at the SORA monitoring station from 2010 to 2019. Located at Arles (Eyrolle et al., 2010), this station is the reference station to evaluate SS and radionuclides concentrations and fluxes near the outlet of the Rhône River in the Mediterranean Sea (Fig. 1).

2.2. Suspended sediment sampling

The sampling method used to characterize the ^{137}Cs concentration in SS of the Rhône River is described in Eyrolle et al. (2020). Briefly, for baseflow period (water discharge lower than 3000 $m^3 s^{-1}$), 13.5 L of water are automatically sampled and filtered at 0.5 µm every 80 min and integrated for a month. For flood period, the previous sampling strategy is stopped, and a specific flood event automatic sampling is settled using up to five acetate cartridges (0.5 µm) with 5 L collected every 60 min until the clogging of the cartridge.

2.3. Radionuclide analyses and concentrations (endogenous variable)

Measurement of ^{137}Cs is described in Eyrolle et al. (2020). The SS samples were slowly (approx. 2 weeks) evaporated (80 °C) to dryness, ashed and put into tightly closed plastic boxes (17 mL or 60 mL) for gamma-ray spectrometry measurements (20–60 g) using low-background and high-resolution High Purity Germanium detectors. Measurements are regularly checked for background noise and drift using multi-gamma calibration sources in accordances with national standards. The detectors are also used in national and international intercomparisons. Results are expressed in Bq kg^{-1} (d.w.) and each

sample was measured for 3 days to achieve detection limits of 0.5 Bq kg^{-1} (d.w.). Measured activities are decay-corrected to the date of sampling. The activity uncertainty (k=2) was estimated as the combination of calibration uncertainties, counting statistics, and summing and self-absorption correction uncertainties.

For the studied period (2010-01-05 to 2019-06-04), 269 SS including 112 during baseflow and 157 during flood events were measured (Fig. 2– supplementary information). From 2010 to 2019, the average particulate concentration was 7.8 ± 6.1 Bq kg^{-1} for ^{137}Cs. Higher mean concentration was observed during baseflow (11.4 ± 7.4 Bq kg^{-1}) than during flood events (5.3 ± 3.1 Bq kg^{-1}).

The annual flux of particulate ^{137}Cs is calculated by multiplying its particulate concentration by the flux of SS (water discharge multiplied by SS concentration) for the sampling periods and summed by year. In 2015, while water discharge and SS concentration were measured, most of the SS sampling was not conducted due to logistical issues. To estimate the flux during this period, the year has been cut according to the hydrology to distinguish floods and baseflows regarding the flood threshold (3000 m^3 s^{-1}).

2.4. Machine learning methodology

2.4.1. Input variables (exogenous variable)

Two different types of exogenous variable were used: routine radioactive liquid releases and hydrological information. Note that with this approach, we do not apply any transit time between the monitored stations (Fig1) and Arles, and we do not consider physicochemical variables such as particle size or organic content as such variables were not available for the whole studied period. All the data used are available in the supplementary information (SI).

2.4.1.1. Nuclear liquid effluents

As explained previously, nuclear facilities are allowed to release radioactive effluent directly into the Rhône River. Such release must respect concentration thresholds and must be carried out under

normal hydrological conditions (baseflow), excluding low-level water and flood. For ^{137}Cs, the main source of liquid effluent in the Rhône River is the reprocessing center of Marcoule (Fig. 1) which represent most of the annual liquid emission of the ^{137}Cs. In fact, the mean annual releases from 2010 to 2016 was 20.3 GBq y^{-1} of ^{137}Cs for the center of Marcoule while the sum of the mean annual releases of the four NPPs was 0.1 GBq y^{-1} of ^{137}Cs. Quantity of ^{137}Cs (MBq) released was estimated for each measurement at the monitoring station without applying a transit time. The duration of the releases (in hours) and the total flux of SS (t) that transited in the river during the release periods was also used as exogeneous data.

2.4.1.2. Water discharge and suspended sediment concentration

In addition to the SORA station, water discharge (Q in $m^3 s^{-1}$) and SS concentration (SSC in mg L^{-1}) were also acquired at Jons, which is the reference station to evaluate concentrations and fluxes from the Upper Rhône River, and at the outlet of the main tributaries: Saône, Isère, Ardèche and Durance rivers (Fig. 1). Acquisition of Q and SSC are described in (Lepage et al., 2021) and was conducted within the Rhône Sediment Observatory. Hourly water discharges (Q_H) were calculated by conversion of water level measurements through stage-discharge rating curves, otherwise through numerical modelling (F. Thollet et al., 2021). During the studied period, the mean water discharge was 1520 ± 818 $m^3 s^{-1}$, 566 ± 317 $m^3 s^{-1}$, 65 ± 131 $m^3 s^{-1}$, 93 ± 125 $m^3 s^{-1}$, 312 ± 150 $m^3 s^{-1}$, 399 ± 420 $m^3 s^{-1}$, respectively at Arles, Jons, Ardèche, Durance, Isère and Saône stations (Fig. 3). The SSC at most stations are derived from in-situ turbidity measurements conducted every 10 minutes (Le Bescond et al., 2018). The SSC is then calculated through the site-specific turbidity-SS rating curve (Navratil et al., 2011), which is determined on each site for a wide range of concentrations. At Arles, the SSC is measured by filtration of water sample collected by an automatic water sampler by the MOOSE network (Mediterranean Ocean Observing System for the Environment) (Raimbault et al., 2014). During the studied period, the mean SSC was 51 ± 160 mg L^{-1}, 19 ± 46 mg L^{-1}, 9 ± 18 mg L^{-1}, 114 ± 606 mg L^{-1}, 82 ± 455 mg L^{-1}, 14 ± 17 mg L^{-1}, respectively for Arles, Jons, Ardèche, Durance, Isère and Saône rivers (Fig. 3). While Q_H was used as input variable (referred to as Q_Name-of-the-station in $m^3 s^{-1}$), SSC was not used in this form but after

125 calculating the flux of SS (t h^{-1}) by multiplying SSC by Q. The SS flux is referred to as FSS_Name-of-the-
126 station.

127 2.4.2. Models

128 2.4.2.1. Random Forest - RF

129 The random forest (RF) is a supervised learning algorithm consisting of a set of decision trees (Breiman,
130 2001). It uses the "bagging" method whose general idea is that a combination of learning models
131 increases the overall result. The regressor algorithm is used to process the regressions. The RF
132 algorithm randomly selects observations and features to construct multiple decision trees and then
133 averages the results. Bootstrap samples are used when building trees. Several hyperparameters are of
134 interest, including:

135 • The n_estimators hyperparameter which is the number of trees the algorithm built before
136 taking the maximum vote or taking the averages of the predictions. In general, a higher
137 number of trees increases performance and makes the predictions more stable, but it also
138 slows down the computation. 10 values were chosen among a uniform distribution in [80,
139 200].
140 • The max_depth hyperparameter is the maximum depth of the tree: 10 values were chosen
141 regularly distributed in the interval [5, 30].
142 • The max_features hyperparameter which is the maximum number of features that the random
143 forest considers splitting a node. It is chosen equal to the number of features which are in our
144 case the columns of water discharges at the different stations.
145 • The min_sample_leaf hyperparameter which determines the minimum number of leaves
146 needed to split an internal node. Ten values were chosen among a uniform distribution in [1,
147 10].

- The min_sample_split hyperparameter which determines the minimum number of samples required to split an internal node. Ten values were chosen among a uniform distribution in [1, 30].

A randomized search on hyper parameters has been performed, optimized by 5 cross-validated searches over parameter settings.

2.4.2.2. Hierarchical Attention-Based Recurrent Highway Networks - HRHN

The proposed deep learning method aims at predicting the future T of the timeseries of ^{137}Cs (endogenous variable) from a set of past information t at T-1 of the presented exogenous variable and the associated ^{137}Cs evolution. The architecture chosen here is the HRHN (Hierarchical Attention-Based Recurrent Highway Networks) an encoder-decoder neural network (Tao et al., 2018). The principle of this architecture is the use of two neural networks. The first one encodes the input data (exogenous variable at time [t,T-1]) into a latent representation. A hierarchical attention layer weights the importance of the different elements of this representation. Then, the second network must decode this information to predict the future T of the time series of ^{137}Cs concentration in relation to the history of concentration ^{137}Cs at time t to T-1. The decoder and encoder may have different layers. Here the encoder is composed of convolutional layers (Lecun et al., 2015) and Recurrent Highway Network (RHN) (Zilly et al., 2017) layers. The decoder is composed of RHN layers. The purpose of the recurrent layers is to capture the temporal dynamics of the series. The convolutional layer aims to detect patterns in the series (seasonality, peak). Max pooling (Aggarwal, 2018) is also performed between successive convolutional layers, which can reduce the size of feature maps to avoid overfitting and improve efficiency. For more information on the nature of the layers we refer the reader to (Goodfellow et al., 2016).

This neural network architecture is associated with several hyperparameters linked to the nature of these layers. The optimization is carried out with the Hyperband method (Li et al., 2018). The aim of the optimization algorithm is to minimize the error on both the training and the test base to propose

the best performing model with the most interesting parameters and hyper-parameters. The detail of the grid of possible combinations for hyperparameters is presented. We have decided to keep the number of convolution layers given in (Tao et al., 2018) (i.e. 3), however their size and the associated max-pooling will be determined in the following interval for each:

- CNN window size (nbr_filter_cnn) in [3,5,7]. The convolution window size determines the size of the region over which convolution is applied at each time step. It is important to choose an appropriate window size to capture the relevant temporal patterns in the data.
- Number of filters (dim_filter_cnn) in [8,16,32,64,128,256]. The number of filters determines how many different patterns the network can learn. The higher the number of filters, the more complex the network can be, but this can also make training more difficult.
- Max pooling size (dim_max_pooling) in [2,3,4,5]. The pooling window size determines the region of the input that will be aggregated into a single output element. In general, a larger pooling window size reduces the spatial resolution of the output but can also improve the robustness of the network to minor variations in the input. In contrast, a smaller pooling window size retains more detail of the input but may also make the network more sensitive to noise or minor variations.

As in Tao et al. (2018), we assume the RHN has same structure in the encoder and the decoder:

- Hidden layers (layer_RHN) in [1,2,3,4,5]. The hidden layers allow the neural network to model non-linear relationships between inputs and outputs. Each hidden layer in a deep neural network computes a non-linear transformation of the previous layer's outputs, allowing the network to learn increasingly abstract and complex features as information is propagated through the network.
- Dimension of hidden state (dim_RHN) in [8,16,32,64,128,256]. The dimension of the hidden state determines the size of the hidden state vectors that are calculated at each time step of the model. A higher dimension of the hidden state can allow the model to capture more

complex and subtle information in the data, but it can also make the model slower to train and require more training data.

The intervals chosen for these different hyperparameters are based on the following references (Chollet, 2018; Goodfellow et al., 2016) and documentation available on Tensorflow. The algorithm Hyperband has been customized to include optimization of the number of time steps. It implies that for each combination of selected hyperparameters an update of the data size is performed. This hyperparameter linked to the data history allows us to determine the quantity of past information most relevant to predict the future evolution of the endogenous variable. The sequence length (here monthly) is selected between 9 and 20 past data.

2.4.3. Splitting methods and cross-validation

Two processes were used to split the dataset:

- The dataset was chronologically separated as a time series as it is required for the HRHN approach. The train dataset ranges between 2010 to end of 2014 with 198 observations and the test dataset ranges between end of 2015 to 2019 with 72 observations (Table 1). For RF, this modelling will be referred as RF_TS. All the mean with their range of the data used by the models are displayed in the Table 1.
- For RF, an additional five cross-validation procedure was performed on the whole data. The split was performed randomly on the 269 observations. This modelling will be referred as RF_CV. With this method, several subsets of training and test data were used, thus considering the full range of exogenous and endogenous variables.

Table 1: mean (min - max) values of the data used by the data-driven models after chronological splitting.

Mean (min – max)	**Train dataset**	**Test dataset**
Number of data	198	71
water discharge at Arles ($m^3 s^{-1}$)	2646 (595 - 5420)	2586 (579 - 6857)
water discharge of Ardèche River ($m^3 s^{-1}$)	202 (5 - 2167)	209 (6 - 2502)

water discharge of Durance River ($m^3 s^{-1}$)	201 (8 - 1010)	270 (10 - 1715)
water discharge of Isère River ($m^3 s^{-1}$)	415 (97 - 1130)	375 (118 - 1083)
water discharge at Jons ($m^3 s^{-1}$)	880 (187 - 2143)	891 (214 - 2932)
water discharge of Saône River ($m^3 s^{-1}$)	792 (56 - 1713)	853 (38 - 2124)
suspended sediment flux at Arles ($t h^{-1}$)	3119 (16 - 48947)	8547 (8 - 105011)
suspended sediment flux of Ardèche River ($t h^{-1}$)	53 (0,05 - 1640)	68 (0,04 - 2066)
suspended sediment flux of Durance River ($t h^{-1}$)	655 (0,1 - 19807)	7874 (0,1 - 113703)
suspended sediment flux of Isère River ($t h^{-1}$)	389 (3 - 6715)	1351 (4 - 20508)
suspended sediment flux at Jons ($t h^{-1}$)	293 (1 - 4324)	434 (1 - 4714)
suspended sediment flux of Saône River ($t h^{-1}$)	96 (0,5 - 622)	195 (1 - 1593)
^{137}Cs release by nuclear industries (MBq)	479 (0 - 7586)	671 (0 - 3428)
Duration of the nuclear release (h)	18 (0 - 211)	16 (0 - 81)
Total suspended sediment flux during nuclear release (t)	4320 (0 - 103254)	4397 (0 - 32921)
^{137}Cs concentration ($Bq kg^{-1}$)	7,8 (2,1 - 24,7)	7,9 (2,2 - 43,9)

2.4.4. Prediction indicators

The proposed models were evaluated using the Root Mean Squared Error (RMSE) following Eq1 and the coefficient of determination (R^2) following Eq2 as it is commonly used in such studies:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left([^{137}Cs]\, i_{measured} - [^{137}Cs] i_{modeled}\right)^2} \quad \text{Eq.1}$$

$$R^2 = 1 - \left(\sum_{i=1}^{N}\left([^{137}Cs] i_{measured} - [^{137}Cs] i_{modeled}\right)^2 / \sum_{i=1}^{N}\left([^{137}Cs] i_{measured} - [^{137}Cs]_{mean}\right)^2\right) \quad \text{Eq. 2}$$

With:

- N is the number of data
- $[^{137}Cs] i_{measured}$ and $[^{137}Cs] i_{modeled}$ respectively the ith occurrence of the measured and modeled concentrations of particulate ^{137}Cs
- $[^{137}Cs]_{mean}$ the mean of the $[^{137}Cs]_{measured}$.

These performance indicators are classically used in model evaluation. The R^2 measures the degree of variability of the dependent variable that can be explained by the model. The R^2 is a good measure of how well the model fits the dependent variables.

2.4.5. Sensitivity of the input variables

The importance of a feature is computed as the (normalized) total reduction of the criterion brought by that feature. It is also known as the Gini importance (Breiman, 2001). It calculates each feature importance as the sum over the number of splits (across all trees) that include the feature, proportionally to the number of samples it splits.

Regarding the HRHN approach, sensitivity analysis in neural networks with a complex architecture is still a subject in full development notably because of the large number of hyperparameters and parameters of these models. It (Doshi-Velez and Kim, 2017). We propose here the Permutation feature importance, a simplistic approach to try to bring elements of answer on the importance of the various exogenous variables. Initially used for random forests, it is applicable to any model (Wei et al., 2015). This method has the advantage of not requiring a re-training phase for the model or long simulations, which can be costly in terms of computing time for HRHN. The concept is straightforward: we measure the importance of a feature by calculating the increase in the model's prediction error after permuting the feature. A feature is "important" if shuffling its values increases the model error, because in this case the model relied on the feature for the prediction. A feature is "unimportant" if shuffling its values leaves the model error unchanged, because in this case the model ignored the feature for the prediction. This method is applied to the test set rather than the training set to assess the importance of the variables in the ability of the model to generalize to the unknown data. This is because the model is trained on the training set and its coefficients are adjusted to minimize the error on the training data. If we use the training set to assess the importance of variables, this may lead to an overestimation of the importance of some variables, as the model has been optimized to minimize the error on these specific data. In contrast, the validation set is used to assess the ability of the model to generalize to

new data. By calculating the importance of the variables on the validation set, we can assess the importance of the variables in the ability of the model to generalize and predict new data. The results obtained correspond to the average of 100 simulations of Permutation feature importance for each variable.

2.5. Empirical modelling of ^{137}Cs

In the Rhône River, recent work was conducted to model the solid/liquid fractionation of ^{137}Cs as a function of the hydro-sedimentary conditions (water discharge, suspended sediment concentration and particle size), the radioactive liquid discharges and the mean ^{137}Cs concentrations of the superficial soils of the watershed (Tomczak et al., 2021). With this approach, hourly ^{137}Cs concentrations in suspended sediments were modeled in the Rhône River at the SORA station using the following equations:

$$\left[^{137}Cs\right]_{SS} = \left(Kd_{\delta} \cdot \left[^{137}Cs\right]_{w} \cdot [SS]_{R} + \left[^{137}Cs\right]_{SOIL} \cdot [SS]_{NR}\right) / [SS] \quad \text{Eq. 3}$$

With:

- Kd_{δ} (l.kg^{-1}), the equilibrium solid/liquid ratio of the superficial reactive layer of particles of thickness δ (µm).
- $\left[^{137}Cs\right]_{SOIL}$(Bq.kg^{-1}), the mean ^{137}Cs concentration of the superficial soils of the catchment.
- $[SS]_R$ (mg.l^{-1}), the reactive fraction of suspended sediments given as a function of d_{50} (µm), the median of their granulometric distribution:

  $[SS]_R = [SS] \cdot (1 - (1 - 2 \cdot \delta / d_{50})^3)$ if $d_{50} > 2 \cdot \delta$ and $[SS]_R = [SS]$ if $d_{50} \leq 2 \cdot \delta$ Eq. 4
- $[SS]_{NR}$ (mg.l^{-1}), the non-reactive fraction of suspended sediments: $[SS]_{NR} = [SS] - [SS]_R$
- $\left[^{137}Cs\right]_{w}$ (Bq.l^{-1}), the dissolved concentration of ^{137}Cs:

  $$\left[^{137}Cs\right]_{w} = \left[^{137}Cs\right]_{E} / \left(1 + Kd_{\delta} \cdot ([SS]_{c} + [SS]_{R})\right) \quad \text{Eq. 5}$$
- $[SS]_c$ (mg.l^{-1}), the colloidal fraction of suspended sediments.

- $\left[^{137}Cs\right]_E = \frac{q_{release}}{Q} + \left[^{137}Cs\right]_{SOIL} \cdot [SS]_R$ (Bq.l^{-1}), the total exchangeable fraction of the ^{137}Cs, where Q (m^3.s^{-1}) is the waterflow of the river and $q_{release}$ (Bq.s^{-1}) is the flux of ^{137}Cs releases by the routine discharges.

To compare this model approach with the measures and the other approaches, the calculated hourly series is integrated over the sampling period of average activity was calculated for each SS sampling period.

3. Results

3.1. Parameters

The optimized parameters used for the Random Forest and the HRHN approaches are detailed in the Table 2. Before making a prediction, the HRHN model receives as input a batch of data corresponding to a small history. According to the optimization, the hyperparameter of the table indicates a sequence length of data history equal to 16: the first 16 values from the input time series and the target time series (^{137}Cs concentration) are therefore used to predict the future ^{137}Cs concentration (i.e. the 17th value of the target series). This 17th value will then be included in the next history of the target variable in the next prediction. The process is repeated until the last time in the series is reached.

Table 2: Values of the hyperparameters of random forest and HRHN

Random Forest		HRHN	
Parameter	Value	Parameter	Value
n_estimators	146	dim_RHN	8
max_depth	13	layer_RHN	2
min_sample_leaf	1	dim_max_pooling	[3,4,2]
min_sample_split	3	dim_filter_cnn	[7,5,5]
		nbr_filter_cnn	[8,32,16]

		sequence_length *of data history*	16

295

296 ## 3.2. Modeling

297 The modeling of particulate ^{137}Cs in the SS in the Rhône River at Arles is presented in the Figure 4. For
298 the different approaches, the modelled concentrations fall within the range of the measurements with
299 the exception of the empirical approach that overestimates the concentration in the early dataset (Fig.
300 4A).The approaches correctly follow the trends with lower concentrations during floods and higher
301 concentrations during baseflow and low-level water but the RF and empirical approaches tend to
302 underestimate the concentration during low-level water (Fig. 4B) while the HRHN correctly estimate
303 the concentration for the whole range of hydrological data. The difference between measured and
304 modelled values is lower for the HRHN and RF_CV approaches with values of RMSE lower than 4 Bq
305 kg^{-1} (Table 3) and R2 higher than 0.5.

306 *Table 3: Prediction indicators for the different approaches.*

	Empirical	RF_TS	RF_CV	HRHN
RMSE Train (Bq kg^{-1})	5.74	0.96	3.54 ± 2.7	0.69
RMSE Test (Bq kg^{-1})		6.52		3.59
R2 Train	0.10	0.96	0.59 ± 0.31	0.98
R2 Test		0.41		0.71

307 ## 3.3. Sensitivity

308 The sensitivity analysis on the input variables shows contrasted results between the HRHN approach
309 and the two others (Fig. 5). For HRHN, SS fluxes of the Durance River and at Arles had a sensitivity
310 higher than 25% while their sensitivity was lower than 5% for RF_TS. For RF_CV, FSS Durance only was

311 negligible. For both RF approaches, the water discharge at Arles (Q_Arles) and from the Saône River
312 (Q_Saone) had a sensitivity higher than 20%. For RF_TS, the water discharge from the upper Rhône
313 River at Jons (Q_Jons) had a sensitivity higher than 10%. Finally, nuclear release had a sensitivity lower
314 than 5% for the different approaches.

315 ### 3.4. Fluxes

316 For the periods with measurement of ^{137}Cs in SS, fluxes of ^{137}Cs were calculated for the three modelling
317 approaches and compared to the measured flux (Fig. 6A). For HRHN, the data history (the sixteen first
318 values of train and test sub datasets – Table 1) represented around 35% of the annual flux of 2010 and
319 2016 so the estimated fluxes are at two third recalculated by this approach in 2010 and 2016. Overall,
320 the different approaches had annual fluxes close to the measured, with slight overestimation in 2010,
321 2013 2016 and 2017. The average absolute differences for the test sub-datasets are respectively 36.0
322 ± 25.0%, 46.7 ± 22.4%, 21.7 ± 19.3% and 30.6 ± 24.7% for HRHN (excluding 2016 due to the data history
323 used), RF_TS, RF_CV and empirical approaches. Finally, flux of 2015 was reconstructed by the three
324 approaches and the result is similar with a flux of ^{137}Cs estimated between 20.0 to 24.5 GBq (Fig. 6B).

325 ## 4. Discussion

326 The two data-driven approaches (HRHN and RF) achieved good prediction scores using only
327 hydrological data (water discharge et SPM fluxes) and industrial input of ^{137}Cs. These results show that
328 the use of the main Rhône River tributaries hydrology alone is sufficient to obtain good results, without
329 any knowledge of the actual input of ^{137}Cs through erosion from the catchment (i.e. soil affected by
330 the global fallout and the Chernobyl accident). Furthermore, all these input variables do not seem to
331 be necessary as shown in the sensitivity analysis. Indeed, industrial releases represent only 5% of the
332 sensitivity analysis (Fig. 5). These results show that near the mouth of the Rhône River, it is rather the
333 hydrology (and the ^{137}Cs concentration in eroded soils) that governs the concentrations of particulate
334 ^{137}Cs than the industrial releases. However, a bias is possible as there are very few events with high
335 concentrations of ^{137}Cs (only 4 on 269 values over 25 Bq kg^{-1} – Fig. 2), events related to the combination

of nuclear release with low water discharge (Q_{events} = 717 ± 98 m^3 s^{-1} vs Q_{mean} = 2586 m^3 s^{-1}) resulting in a low dilution of the concentration released. So, the weight of the releases could be higher if the representation of these events was larger in the dataset and the score improved for all the models as they underestimate the concentration during low-level water (Fig. 4B). This confirms the interest of regularly updating the parameterization of the models with larger datasets. Also, the sensitivity analysis shows that some tributaries have more weight than others: the Durance tributary for HRHN and the Saône tributary for RF. The weight of the Durance River is logical as this tributary is characterized by intense floods which generate very important sediment flux (Bodereau et al., 2022; Delile et al., 2020; Poulier et al., 2019) for a relatively low water discharge compared to the other tributaries (Fig. 7A). This result confirms the necessity to not only use the water discharge but also the flux of SS as the relation between the two parameters is not linear (Fig. 7B). For the RF approaches, the negligibility of the Durance tributary and the weight of the Saône tributary illustrate the difficulty in the choice of input variables when they are correlated (Gregorutti et al., 2017) which is the case of the hydrological data of this study. For these approaches, the weight of the hydrology at the monitoring station (Arles) is high and might also explain the low weight of the Durance tributary due to the proximity of this tributary (Fig. 1). Therefore, the models could be improved by removing the insignificant variables that indicate redundancy of information in the algorithm.

Regarding the score between the different models, the data-driven approaches have a better prediction of ^{137}Cs than the empirical approach and among the data-driven approaches, the most complex one (HRHN) displayed better prediction. However, these machine learning and deep learning approaches proposed in this article are very different. First, Random Forest (RF) is a machine learning method that will make a prediction based on a set of regression trees, each having learned its tree structure from a sample of the training base. Therefore, its operation will be sensitive to the training frame and to a small number of extreme values (peaks in the data). This sensitivity explains the better scores obtained after the cross validation splitting method (RF_CV) than with temporal splitting (RF_TS) as the CV splitting considered a larger number of extreme events. In fact, the FSS_Arles and

FSS_Durance variables used for RF_TS training are not representative of the full dataset (Table 1), which explains the low sensitivity scores for these variables with this approach. Thus, a number of nodes and leaves will not be present, making it difficult to predict particular cases. However, RF is distinguished by the very intuitive understanding of its algorithm and will find the average trends. Furthermore, its low number of hyperparameters makes the algorithm quickly applicable, so it can be considered to increase its learning base in order to increase its search area. Score of the random forest could also be improved by considering a causal techniques as the algorithm causal forest (Shuryak, 2023). Second, the HRHN approach aims to understand the dynamics of the target variable: understanding past events in order to predict future behavior. This understanding is achieved through the different layers of HRHN, by detecting patterns in the time series (seasonality, peak...) through convolutional layers and by trying to capture the temporal dynamics through recurrent layers. During training, the network will parameterize itself (weighting) to stimulate the most important information in the network to estimate the future prediction. The model will therefore a priori have a greater robustness in the most extreme cases, which explain why the results are better than RF_TS despite the non-representativeness of the FSS_Arles and FSS_Durance variables in the training set (Table 1). However, its architecture is associated with a large number of hyperparameters and parameters which makes its optimization phase expensive in terms of computation time. Therefore, the interpretability of the model is made complex. Although basic sensitivity analysis methods can be implemented to provide a measure of the relative importance of input characteristics, it will be necessary to develop more suitable algorithms to capture the full richness of the model. It should be noted here that the impact of the history of the target variable has not been studied. Even more, prediction with HRHN requires knowledge of a time history which limits its application to a framework where past data is available. Finally, regarding the particulate ^{137}Cs annual fluxes, the different approaches can be used to predict it as the differences are generally low, especially in 2015 (Fig. 7). At the outlet of the Rhône River, the flux of ^{137}Cs is mostly related to the flux of SS than the concentration of ^{137}Cs. In fact, the

^{137}Cs concentration varies by one order of magnitude while the FSS_Arles varies by five orders of magnitude (Table 1).

To conclude, while the data-driven models can estimate concentrations in the perimeter of this study (i.e. normal condition of nuclear releases) without the need of the ^{137}Cs released by nuclear industries (as lower than 5% of the sensitivity analysis), the empirical model needs this information which is difficult to obtain. This limitation shows that an AI approach seems to be easier to apply on this catchment to predict particulate ^{137}Cs in normal condition and that this could be a useful approach for interpreting and supplementing routine monitoring datasets. In case of accidental releases (e.g. concentration several orders of magnitude above those in this study), the AI models will struggle to predict the right concentrations if their training dataset does not incorporate such situations and empirical approaches would then be more useful. However, in this example, the AI approach would reveal an anomaly between the prediction and the measurement, and therefore allow further work to be undertaken to understand the anomaly. Moreover, the AI approach might be used to estimate the quantity of ^{137}Cs released using the ^{137}Cs concentration in SS as input data. Going further, whatever the sediment-bound contaminant of interest, AI seems to be a good tool for predicting concentrations or sources as long as any of this information is available.

5. Conclusion

Concentrations of ^{137}Cs is suspended sediment near the mouth of the Rhône River were estimated using three data-driven models and an empirical approach. Thanks to the long-term monitoring of this radionuclide, the data-driven models were able to accurately predict its concentration with better scores than the empirical approach. These results demonstrate that the use of a basic data-driven model (random forest) allow to obtain a better prediction of this radionuclide with a few diversities of input variables: water discharge and suspended sediment flux, and ^{137}Cs release from nuclear facilities. In order to improve the prediction, a more complex model was used (HRHN) with an architecture that permit to understand the temporality of the ^{137}Cs concentration. In summary, these methods offer

different means of prediction with greater or lesser advantages. It will then be up to the user to choose the most suitable model according to the objectives and the data available.

Acknowledgement

This study was conducted within the Rhône Sediment Observatory, a multi-partner research program funded through the Plan Rhône by the European Regional Development Fund (ERDF), Agence de l'eau RMC, CNR, EDF, and three regional councils (Auvergne-Rhône-Alpes, Région Sud - PACA, and Occitanie). The post-doctoral position of K. Pele was funded under the ANR TRAJECTOIRE project (ANR-19-CE3- 0009, 2020–2024). The authors would like to thank the OSR staff for suspended particulate matter sampling and analytical measurement, especially A. Gruat, F. Thollet, M. Lagouy, P. Paulat, D. Mourrier, F. Giner, F. Eyrolle, C. Antonelli, G. Dur, J. Labille and S. Igguy. We would particularly like to thank the Environmental Control and Impact Assessment Laboratory (LCEI) of the Marcoule center for transmitting information on nuclear releases.

Author contributions.

HL conceptualized the study; HL curated the data; VNG, KP and PB performed the modelling; HL acquired the funds; all authors developed the methodology; HL performed the visualization; all authors contributed to the original draft of the paper.

5.1. Bibliographie

Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Mohamed, N.A.E., Arshad, H., 2018. State-of-the-art in artificial neural network applications: A survey. Heliyon 4, e00938. https://doi.org/10.1016/j.heliyon.2018.e00938

Aggarwal, C.C., 2018. Neural Networks and Deep Learning. Springer 10, 978.

Antonelli, C., Eyrolle, F., Rolland, B., Provansal, M., Sabatier, F., 2008. Suspended sediment and 137Cs fluxes during the exceptional December 2003 flood in the Rhone River, southeast France.

Geomorphology 95, 350–360. https://doi.org/10.1016/j.geomorph.2007.06.007

Bodereau, N., Delaval, A., Lepage, H., Eyrolle, F., Raimbault, P., Copard, Y., 2022. Hydrological classification by clustering approach of time-integrated samples at the outlet of the Rhône River: Application to Δ14C-POC. Water Res. 220, 118652. https://doi.org/10.1016/j.watres.2022.118652

Braga, M.C.B., Birkett, J.W., Shaw, G., Lester, J.N., 2010. Modelling the long-term fate of mercury in a Lowland Tidal River. II. Calibration and comparison of two models with field data. Arch. Environ. Contam. Toxicol. 58, 383–393. https://doi.org/10.1007/s00244-009-9378-8

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Cho, E., Arhonditsis, G.B., Khim, J., Chung, S., Heo, T.Y., 2016. Modeling metal-sediment interaction processes: Parameter sensitivity assessment and uncertainty analysis. Environ. Model. Softw. 80, 159–174. https://doi.org/10.1016/j.envsoft.2016.02.026

Chollet, F., 2018. Deep learning with Python. Simon and Schuster.

Ciffroy, P., Benedetti, M., 2018. A comprehensive probabilistic approach for integrating natural variability and parametric uncertainty in the prediction of trace metals speciation in surface waters. Environ. Pollut. 242, 1087–1097. https://doi.org/10.1016/j.envpol.2018.07.064

Ciffroy, P., Garnier, J.M., Khanh Pham, M., 2001. Kinetics of the adsorption and desorption of radionuclides of Co, Mn, Cs, Fe, Ag and Cd in freshwater systems: Experimental and modelling approaches. J. Environ. Radioact. 55, 71–91. https://doi.org/10.1016/S0265-931X(01)00026-1

Delile, H., Masson, M., Miège, C., Le Coz, J., Poulier, G., Le Bescond, C., Radakovitch, O., Coquery, M., 2020. Hydro-climatic drivers of land-based organic and inorganic particulate micropollutant fluxes: The regime of the largest river water inflow of the Mediterranean Sea. Water Res. 185, 116067. https://doi.org/10.1016/j.watres.2020.116067

458 Desai, H.K., Christian, R.A., Banerjee, J., Patra, A.K., 2013. A fuzzy approach for modelling radionuclide
459 in lake system. J. Environ. Radioact. 124, 37–43. https://doi.org/10.1016/j.jenvrad.2013.03.010

460 Dethier, E.N., Renshaw, C.E., Magilligan, F.J., 2020. Toward Improved Accuracy of Remote Sensing
461 Approaches for Quantifying Suspended Sediment: Implications for Suspended-Sediment
462 Monitoring. J. Geophys. Res. Earth Surf. 125. https://doi.org/10.1029/2019JF005033

463 Doshi-Velez, F., Kim, B., 2017. Towards A Rigorous Science of Interpretable Machine Learning.

464 Dragović, S., 2022. Artificial neural network modeling in environmental radioactivity studies – A review.
465 Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2022.157526

466 Eyrolle, F., Antonelli, C., Raimbault, P., Boullier, V., Arnaud, M., 2010. SORA: a high frequency flux
467 monitoring station at the lower Rhône River, in: Proceedings of the 39th CIESM Congress, Venice,
468 Italy. pp. 10–14.

469 Eyrolle, F., Lepage, H., Antonelli, C., Morereau, A., Cossonnet, C., Boyer, P., Gurriaran, R., 2020.
470 Radionuclides in waters and suspended sediments in the Rhone River (France) - Current contents,
471 anthropic pressures and trajectories. Sci. Total Environ. 723, 137873.
472 https://doi.org/10.1016/j.scitotenv.2020.137873

473 Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

474 Gregorutti, B., Michel, B., Saint-Pierre, P., 2017. Correlation and variable importance in random forests.
475 Stat. Comput. 27, 659–678. https://doi.org/10.1007/s11222-016-9646-1

476 He, Q., Walling, D.E., 1996. Interpreting particle size effects in the adsorption of 137Cs and
477 unsupported 210Pb by mineral soils and sediments. J. Environ. Radioact. 30, 117–137.
478 https://doi.org/10.1016/0265-931X(96)89275-7

479 Hilko, O.S., Kundas, S.P., Gishkeluk, I.A., 2012. Radionuclides migration modelling using artificial neural
480 networks and parallel computing. Eur. water 39, 3–13.

Horowitz, A.J., 2009. Monitoring suspended sediments and associated chemical constituents in urban environments: Lessons from the city of Atlanta, Georgia, USA water quality monitoring program. J. Soils Sediments 9, 342–363. https://doi.org/10.1007/s11368-009-0092-y

Hu, X., McIsaac, G.F., David, M.B., Louwers, C.A.L., 2007. Modeling Riverine Nitrate Export from an East-Central Illinois Watershed Using SWAT. J. Environ. Qual. 36, 996–1005. https://doi.org/10.2134/jeq2006.0228

Ikenoue, T., Shimadera, H., Kondo, A., 2020. Impact of soil erosion potential uncertainties on numerical simulations of the environmental fate of radiocesium in the Abukuma River basin. J. Environ. Radioact. 225, 106452. https://doi.org/10.1016/j.jenvrad.2020.106452

Konoplev, A., Kanivets, V., Laptev, G., Voitsekhovich, O., Zhukova, O., Germenchuk, M., 2020. Long-Term Dynamics of the Chernobyl-Derived Radionuclides in Rivers and Lakes. Behav. Radionuclides Environ. II 323–348. https://doi.org/10.1007/978-981-15-3568-0_7

Korobitsyn, B.A., Chukanov, V.N., Yakshina, N. V., 2008. Artificial neural net modeling of the radioactive contamination of the Techa River. At. Energy 105, 138–144. https://doi.org/10.1007/s10512-008-9077-y

Kulahci, F., Özer, A.B., Doğru, M., 2006. Prediction of the radioactivity in Hazar Lake (Sivrice, Turkey) by artificial neural networks. J. Radioanal. Nucl. Chem. 269, 63–68. https://doi.org/10.1007/s10967-006-0230-6

Le Bescond, C., Thollet, F., Poulier, G., Gairoard, S., Lepage, H., Branger, F., Jamet, L., Raidelet, N., Radakovitch, O., Dabrin, A., Coquery, M., Le Coz, J., 2018. From water fluxes to suspended particulate matter and associated contaminant fluxes: Management of hydro-sedimentary stations on the Rhône River. Houille Blanche 63–70. https://doi.org/10.1051/lhb/2018033

Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444.

Lepage, H., Evrard, O., Onda, Y., Patin, J., Chartin, C., Lefèvre, I., Bonté, P., Ayrault, S., 2014.

505 Environmental mobility of110mAg: Lessons learnt from Fukushima accident (Japan) and potential
506 use for tracking the dispersion of contamination within coastal catchments. J. Environ. Radioact.
507 130, 44–55. https://doi.org/10.1016/j.jenvrad.2013.12.011

508 Lepage, H., Gruat, A., Thollet, F., Le Coz, J., Coquery, M., Masson, M., Dabrin, A., Radakovitch, O.,
509 Eyrolle, F., Labille, J., Ambrosi, J.-P., Delanghe, D., Raimbault, P., 2021. Concentrations and fluxes
510 of suspended particulate matters and associated contaminants in the Rhône River from Lake
511 Geneva to the Mediterranean Sea. Earth Syst. Sci. Data 7, 2369–2384.
512 https://doi.org/10.5194/essd-2021-350

513 Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., Talwalkar, A., 2018. Hyperband: A novel bandit-
514 based approach to hyperparameter optimization. J. Mach. Learn. Res. 18, 1–52.

515 Lu, H., Li, H., Liu, T., Fan, Y., Yuan, Y., Xie, M., Qian, X., 2019. Simulating heavy metal concentrations in
516 an aquatic environment using artificial intelligence models and physicochemical indexes. Sci.
517 Total Environ. 694. https://doi.org/10.1016/j.scitotenv.2019.133591

518 Navratil, O., Esteves, M., Legout, C., Gratiot, N., Nemery, J., Willmore, S., Grangeon, T., 2011. Global
519 uncertainty analysis of suspended sediment monitoring using turbidimeter in a small
520 mountainous river catchment. J. Hydrol. 398, 246–259.
521 https://doi.org/10.1016/j.jhydrol.2010.12.025

522 Olivier, J., Carrel, G., Lamouroux, N., Dole-Olivier, M.-J., Malard, F., Bravard, J., Piégay, H., Castella, E.,
523 Barthélemy, C., 2022. The Rhône River Basin, in: Rivers of Europe. Elsevier, pp. 393–453.
524 https://doi.org/10.1016/B978-0-08-102612-0.00011-0

525 Poulier, G., Launay, M., Le Bescond, C., Thollet, F., Coquery, M., Le Coz, J., 2019. Combining flux
526 monitoring and data reconstruction to establish annual budgets of suspended particulate matter,
527 mercury and PCB in the Rhône River from Lake Geneva to the Mediterranean Sea. Sci. Total
528 Environ. 658, 457–473. https://doi.org/10.1016/j.scitotenv.2018.12.075

529 Raimbault, P., Lagadec, V., Garcia, N., 2014. Water sample analyses - MOOSE - Rhone river. SEDOO
530 OMP. https://doi.org/10.6096/MISTRALS-MOOSE.767

531 Roussel-Debel, S., Renaud, P., Métivier, J.M., 2007. 137Cs in French soils: Deposition patterns and 15-
532 year evolution. Sci. Total Environ. 374, 388–398. https://doi.org/10.1016/j.scitotenv.2006.12.037

533 Shuryak, I., 2023. Analysis of causal effects of 137 Cs deposition on 137 Cs concentrations in trees after
534 the Fukushima accident using machine learning. J. Environ. Radioact. 264, 107205.
535 https://doi.org/10.1016/j.jenvrad.2023.107205

536 Shuryak, I., 2022. Machine learning analysis of 137Cs contamination of terrestrial plants after the
537 Fukushima accident using the random forest algorithm. J. Environ. Radioact. 241, 106772.
538 https://doi.org/10.1016/j.jenvrad.2021.106772

539 Skwarzec, B., Kabat, K., Astel, A., 2009. Seasonal and spatial variability of 210Po, 238U and 239+240Pu
540 levels in the river catchment area assessed by application of neural-network based classification.
541 J. Environ. Radioact. 100, 167–175. https://doi.org/10.1016/j.jenvrad.2008.11.007

542 Syvitski, J.P.M., Vörösmarty, C.J., Kettner, A.J., Green, P., 2005. Impact of humans on the flux of
543 terrestrial sediment to the global coastal ocean. Science (80-.). 308, 376–380.
544 https://doi.org/10.1126/science.1109454

545 Tao, Y., Ma, L., Zhang, W., Liu, J., Liu, W., Du, Q., 2018. Hierarchical Attention-Based Recurrent Highway
546 Networks for Time Series Prediction.

547 Thollet, F., Le Bescond, C., Lagouy, M., Gruat, A., Grisot, G., Le Coz, J., Coquery, M., Lepage, H.,
548 Gairoard, S., Gattacceca, J.C., Ambrosi, J.-P., Radakovitch, O., Dur, G., Richard, L., Giner, F., Eyrolle,
549 F., Angot, H., Mourier, D., Bonnefoy, A., Dugué, V., Launay, M., Troudet, L., Labille, J., Kieffer, L.,
550 2021. Observatoire des sédiments du Rhône. https://doi.org/10.17180/OBS.OSR

551 Thollet, Fabien, Rousseau, C., Camenen, B., Boubkraoui, S., Branger, F., Lauters, F., Némery, J., 2021.
552 Long term high frequency sediment observatory in an alpine catchment: The Arc-Isère rivers,

553 France. Hydrol. Process. 35, 1–6. https://doi.org/10.1002/hyp.14044

554 Tomczak, W., Boyer, P., Eyrolle, F., Radakovitch, O., Krimissa, M., Lepage, H., Amielh, M., Anselmet, F.,
555 2021. Modelling of solid / liquid fractionation of trace metals for suspended sediments according
556 to the hydro-sedimentary conditions of rivers - Application to 137Cs in the Rhône River (France).
557 Environ. Model. Softw. 145, 105211. https://doi.org/10.1016/j.envsoft.2021.105211

558 Tomczak, W., Boyer, P., Krimissa, M., Radakovitch, O., 2019. Kd distributions in freshwater systems as
559 a function of material type, mass-volume ratio, dissolved organic carbon and pH. Appl.
560 Geochemistry 105, 68–77. https://doi.org/10.1016/j.apgeochem.2019.04.003

561 Tutu, H., Cukrowska, E.M., Dohnal, V., Havel, J., 2005. Application of artificial neural networks for
562 classification of uranium distribution in the Central Rand goldfield, South Africa. Environ. Model.
563 Assess. 10, 143–152. https://doi.org/10.1007/s10666-005-0214-x

564 Wei, P., Lu, Z., Song, J., 2015. Variable importance analysis: A comprehensive review. Reliab. Eng. Syst.
565 Saf. 142, 399–432. https://doi.org/10.1016/j.ress.2015.05.018

566 WFD, 2000. Directive of the European Parliament and of the Council 2000/60/EC. Establishing a
567 Framework for Community Action in the Field of Water Policy. Off. J. Eur. Parliam. L327, 1–82.

568 Yang, Y.S., Wang, L., 2010. A review of modelling tools for implementation of the EU water framework
569 directive in handling diffuse water pollution. Water Resour. Manag. 24, 1819–1843.
570 https://doi.org/10.1007/s11269-009-9526-y

571 Yaseen, Z.M., 2021. An insight into machine learning models era in simulating soil, water bodies and
572 adsorption heavy metals: Review, challenges and solutions. Chemosphere 277, 130126.
573 https://doi.org/10.1016/j.chemosphere.2021.130126

574 Ye, Z., Yang, J., Zhong, N., Tu, X., Jia, J., Wang, J., 2020. Tackling environmental challenges in pollution
575 controls using artificial intelligence: A review. Sci. Total Environ. 699, 134279.
576 https://doi.org/10.1016/j.scitotenv.2019.134279

577 Zheleznyak, M., Kivva, S., Pylypenko, O., Sorokin, M., 2022. Modeling of Behavior of Fukushima-Derived
578 Radionuclides in Freshwater Systems. Behav. Radionuclides Environ. III 199–252.
579 https://doi.org/10.1007/978-981-16-6799-2_11

580 Zhu, S., Zhang, Z., Žagar, D., 2018. Mercury transport and fate models in aquatic systems: A review and
581 synthesis. Sci. Total Environ. 639, 538–549. https://doi.org/10.1016/j.scitotenv.2018.04.397

582 Zilly, J.G., Srivastava, R.K., Koutnik, J., Schmidhuber, J., 2017. Recurrent highway networks. 34th Int.
583 Conf. Mach. Learn. ICML 2017 8, 6346–6357.
584 https://doi.org/https://doi.org/10.48550/arXiv.1607.03474

585

Figure 1: Location of the monitoring stations and the nuclear facilities within the French part of the Rhône River downstream Lake Geneva and inventory of ^{137}Cs estimated at the date of may 1986 (Roussel-Debel et al., 2007).

Figure 2: Concentration of particulate ^{137}Cs measured at Arles during the studied period.

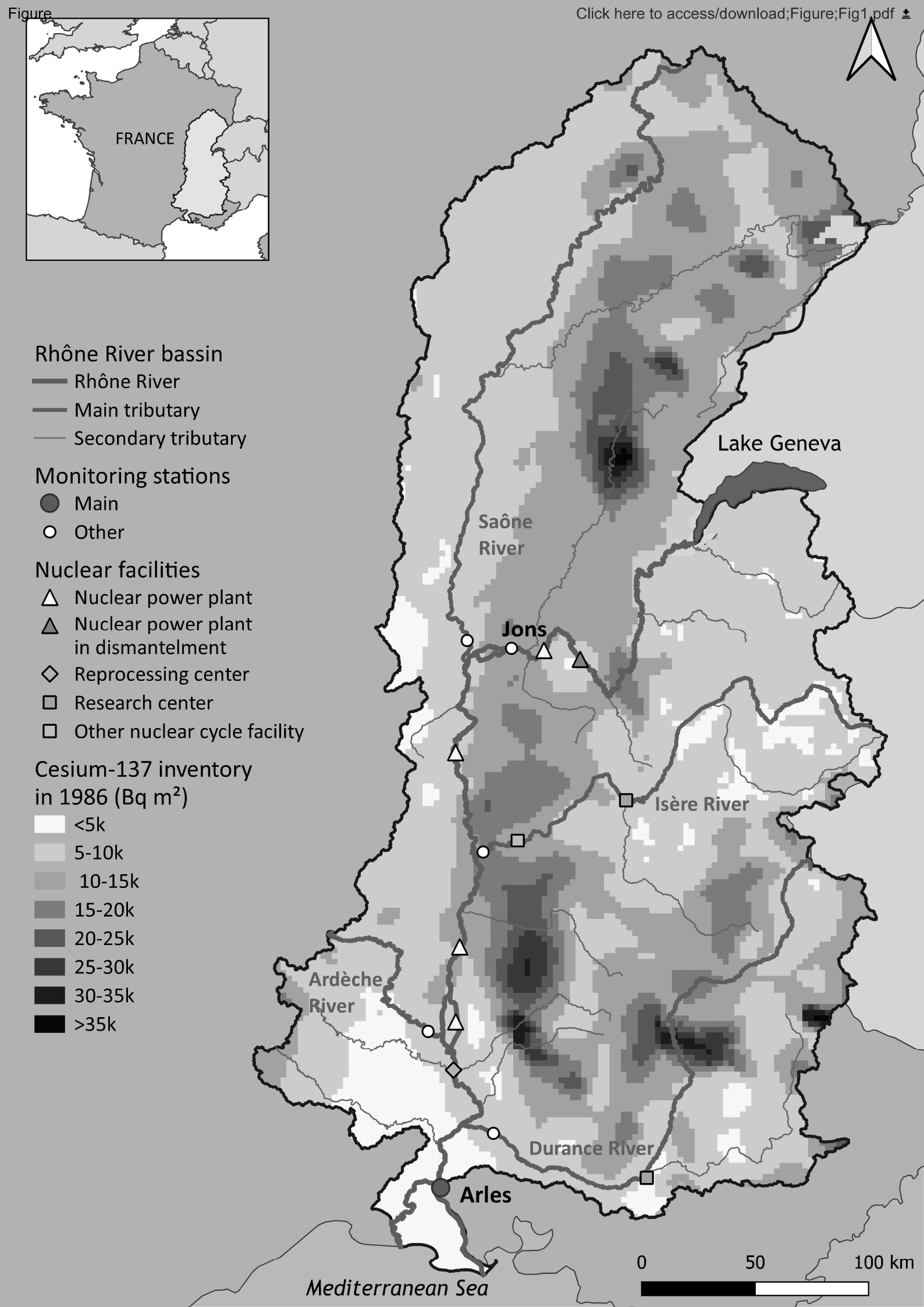
Figure 3: (A) Mean hourly water discharge and (B) suspended sediment concentration at the monitored stations during the studied period.

Figure 4: (A) measured (black) and modeled (grey) concentration of particulate ^{137}Cs and (B) relation between measure and modelling on the test sub-datasets. For the HRHN approach, the 16th firsts values of the train (2010 – 2014) and the test (2015 – 2019) datasets were used as historic in the input dataset and not reported on the B graphic.

Figure 5: Sensitivity (%) of the input variables for the different approaches.

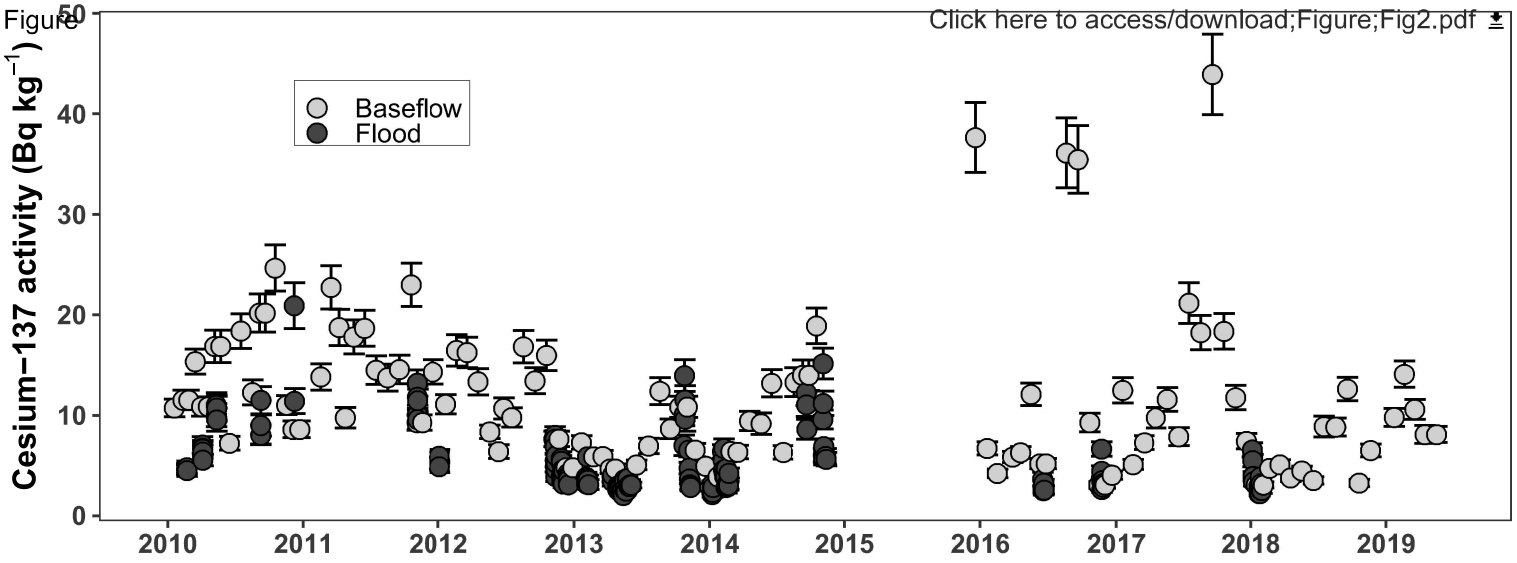
Figure 6: (A) measured and modelled annual flux of ^{137}Cs including only periods with suspended sediment sampling and (B) estimation of the total flux in 2015 with completion of the periods without sampling by the four models.

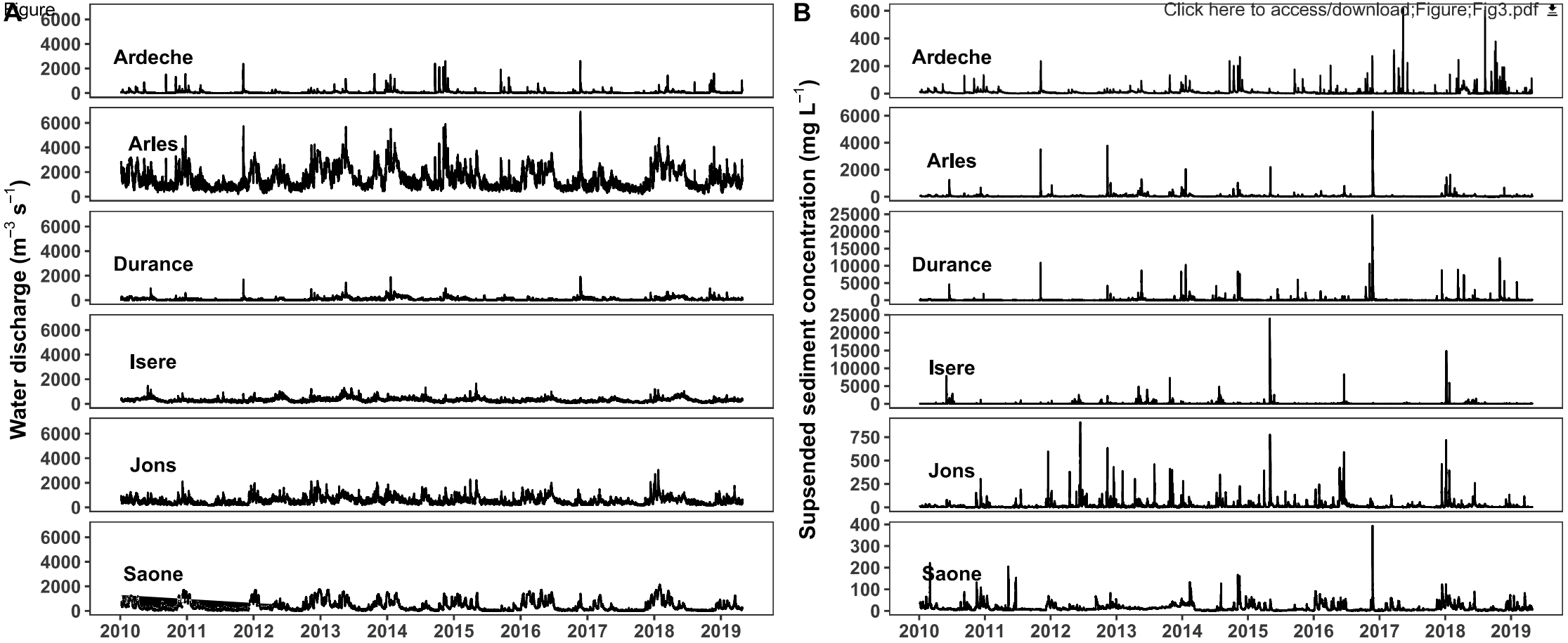
Figure 7: (A) mean annual flux of water and suspended sediment for the 5 stations and (B) their relation between the water discharge and the suspended sediment flux.

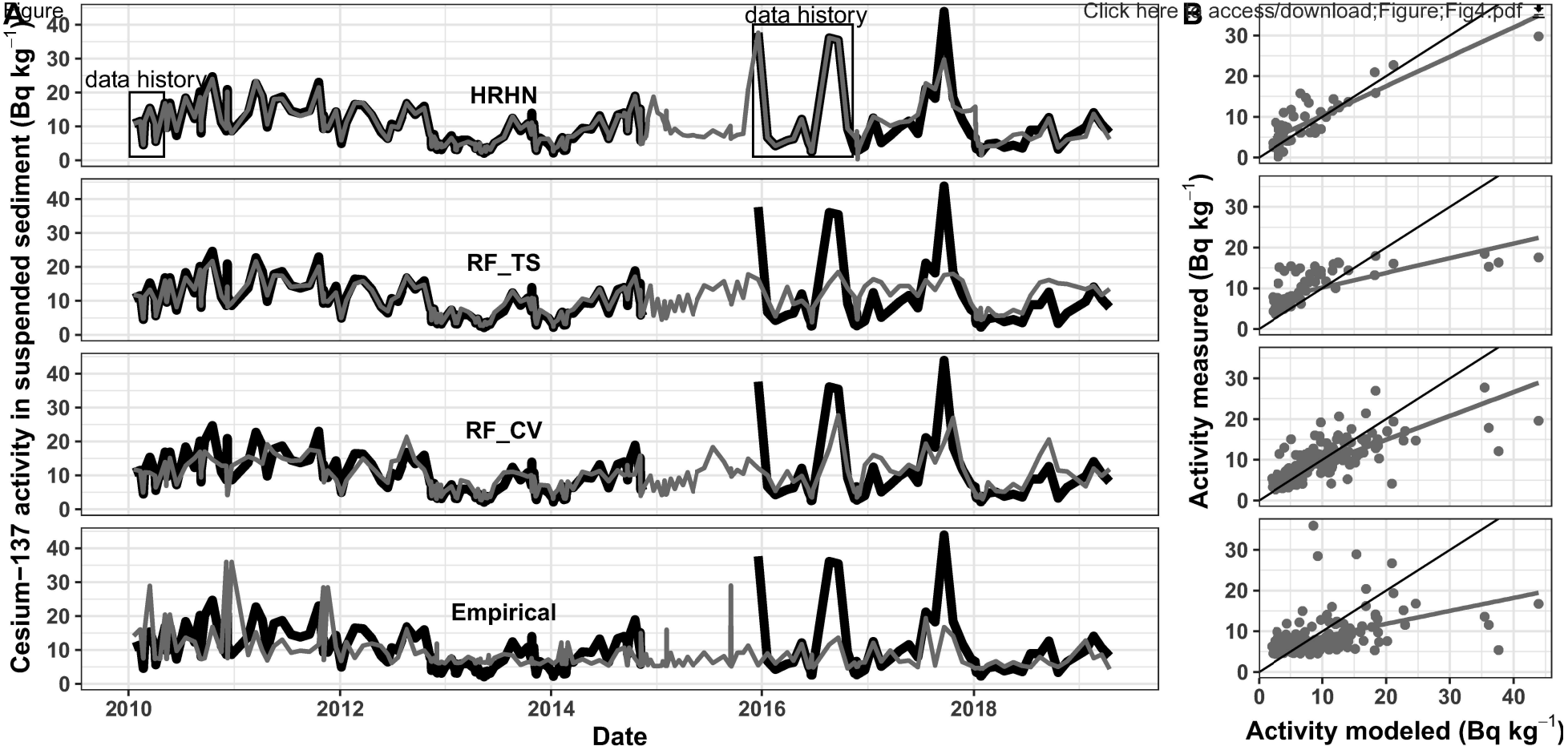


Figure

Click here to access/download;Figure;Fig2.pdf

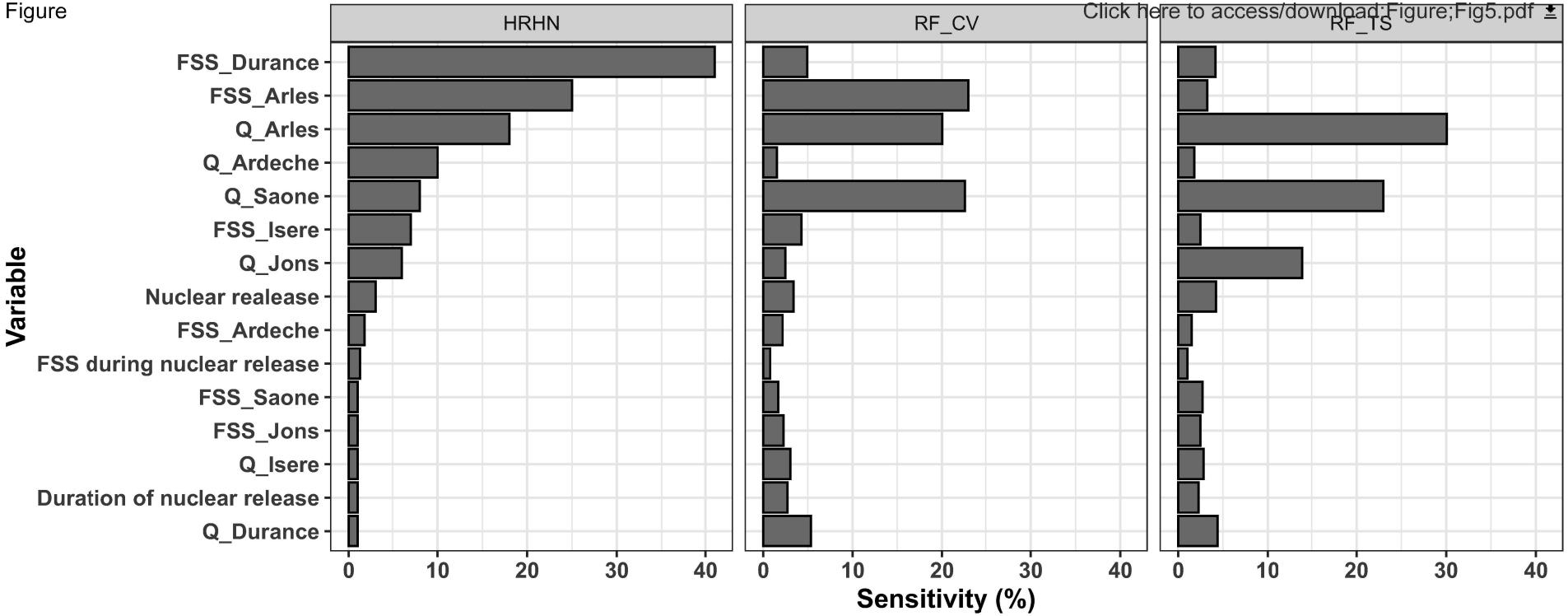






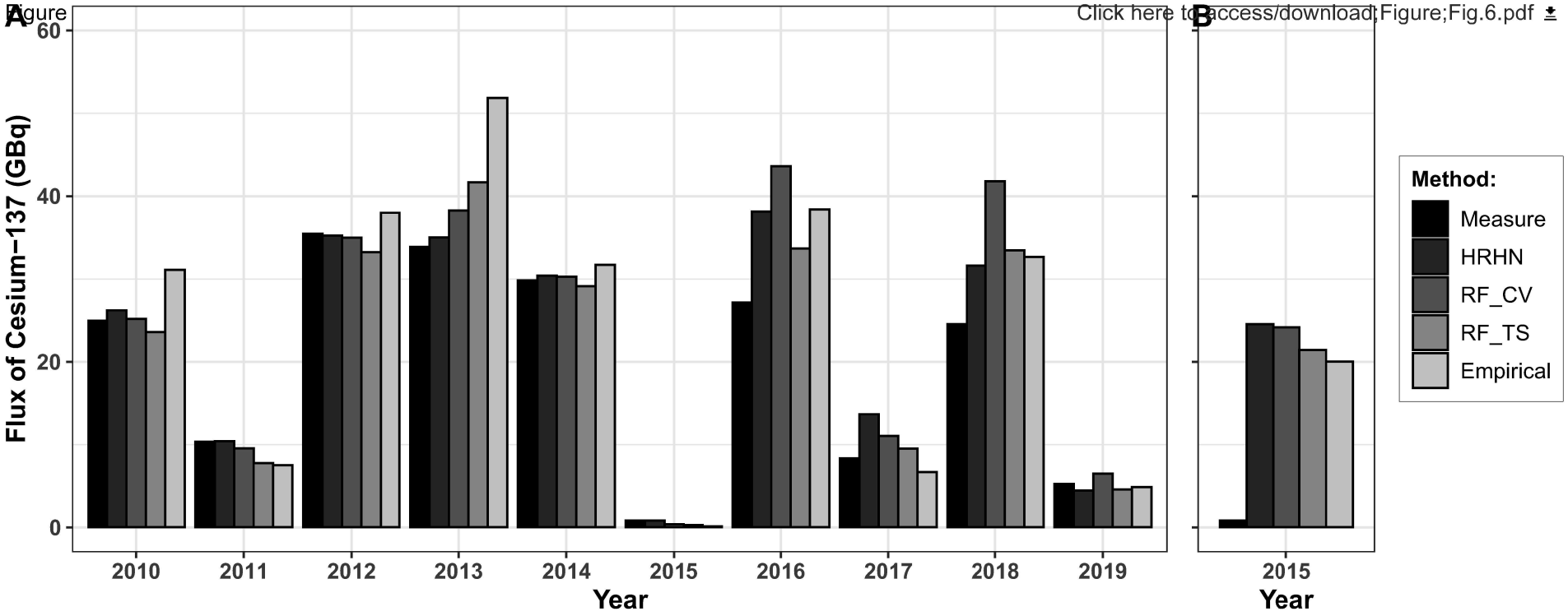
Figure

Click here to access/download;Figure;Fig5.pdf



Figure

Click here to access/download;Figure;Fig.6.pdf



Figure

Click here to access/download;Figure;Fig7.jpg