Response to Referee Bethanna Jackson

REFEREE: Although this manuscript is much improved compared to an earlier submission to this journal, it still is not clearly demonstrating a contribution to new ideas /methods; I would like to see in a revision, significant effort on further positioning it versus other models and methods to demonstrate a uniqueness of GLOBAL-FATE versus other models and software available.

AUTHORS: We thank Dr. Jackson for this comment, which is somehow in agreement with a concern raised by Alberto Bellin. In the new version of the manuscript, we put particular emphasis on this aspect, from the abstract to the discussion. We understand GLOBAL-FATE as an advanced in the state-of-the-art of global contaminant modelling, offering in a single tool features that are scattered in several models, but not available in a single application. Therefore, we do not consider GLOBAL-FATE as offering novel scientific advances in the way we parametrize the processes at play, but it would undoubtedly make global contaminant modelling accessible to a much wider community of scientist and policymakers. This will ultimately help the progress of large-scale contaminant modelling offering an open source and flexible platform to test new parameterizations (hypothesis), and also allowing policymakers to plan global or continental strategic actions.

We have modified the manuscript to make this point much clearer, including statement in several places:

- Line 11-Abstract : “GLOBAL-FATE is the first open-source, multiplatform, user-friendly, and modular contaminant fate model operating at the global scale linking human consumption of pharmaceutical-like compounds with their concentration in the river network.”

- Line 23-Abstract: “GLOBAL-FATE will be a valuable tool for the scientific community and the policymaking arena, and could be used to test the effectiveness of large scale management strategies related to pharmaceutical consumption control and wastewater treatment implementation and upgrading.”

- Line 62: “GLOBAL-FATE has been designed to overcome these constraints, offering the first contaminant fate model operating at the global river network, including lakes and reservoirs, which is at the same time open-source, multiplatform, user-friendly, and modular. This will make global contaminant calculations accessible to a much wider community of scientists and practitioners, opening the door for including pharmaceutical pollution into influential assessments of climate change impacts (e.g., the Inter Sectoral Impact Model Intercomparison project) and global policy instruments like the UN Sustainable Development Goals agenda. GLOBAL-FATE calculates the steady-state concentration of a user-defined down-the-drain contaminant through the global river network, including lakes and reservoirs. GLOBAL-FATE is offered as an open-source, GIS-based model programmed in the C language, allowing researchers to select the input information (water routing, hydrology, population, etc.) and the spatial resolution at which the model has to perform. So forth, the model can include new or different hydrological datasets and other input information, and hence it is not fundamentally restricted to a single modelling resolution, hydrological, or socio-economic scenario. The model simulates the propagation of down-the-drain contaminants along the river network, and the constituent decreases at a rate proportional to its concentration in the aquatic media. GLOBAL-FATE is also computationally efficient, can be run in Windows or Linux machines, and can take advantage of parallel computing in multi-processor computers or clusters. It can also be run as a user-friendly plug-in in QGIS, and the modular structure of its code allows switching different functions of the model on and off."

Line 83: "GLOBAL-FATE is a physically-based model for simulating constituent inputs to the river network and their routing along the river network at the global scale. Our approach shares key assumptions and modelling mechanisms with other large scale pharmaceutical models for the river network (i.e., Keller et al. 2006; Pistocchi 2014; Grill et al., 2019), including the use of per capita mass emissions of the contaminant of interest, simplified parameterization of losses due to human metabolism and removal in wastewater treatment plants, and dilution and first order attenuation dynamics upon discharge into natural waters. However, GLOBAL-FATE is the first model natively operating at the global scale including all those mechanisms, including explicit routing and attenuation in lakes and reservoirs."

- Line 418: "GLOBAL-FATE is an open-source, multiplatform, and modular contaminant fate model that links human consumption of pharmaceutical-like compounds with their concentration in the river network. GLOBAL-FATE is also computationally efficient, and can solve the whole global streamflow generation and contaminant routing in less than five minutes in a customary PC. It provides practical guidelines (through readme files and example datasets) to assist non-specialist users in computer programming. At the same time, it has a fully commented code that experienced users can easily customize and further develop to adapt to their needs. The model is also available as a user-friendly QGIS plug-in. Through simple menus, an inexperienced user can conduct simulations and produce basic outputs on the QGIS canvas. This will make global contaminant calculations accessible to a much wider community of scientists and practitioners."

I am in strong agreement with the other referee, that the current main argument does not fully hold - it is not appropriate to consider that GLOBAL-FATE is not associated to a spatial resolution. It may RUN at any spatial resolution, but that is very different to it being methodologically appropriate at any spatial resolution. Not only detailed physics, but even very simplified physically based approximations to processes/integration of rates of change understanding break down once space and/or time steps become too large. Please add very strong warnings about upper limits.

We cannot but agree with Dr. Jackson at this point, which was also raised by Dr. Bellin. We acknowledge that we somehow oversold the scale-free feature of GLOBAL-FATE, because although it is a potential advantage over other available models, it also leaves the door open for gross misuses of the model. To avoid this, we worked in two directions: first, we substantially expanded the section where we assess the limitations of GLOBAL-FATE as implemented in the example, including a new figure showing detailed results for a single watershed; and second, we included a clear warning in the discussion about the use of GLOBAL-FATE at low resolutions.

For the first point, we included the following text (lines 363-375):

"The concentration maps in Figs. 3 and 4 do not show pixels with less than 100 mm year-1 of runoff, which correspond to arid regions. We decided to discard concentration values in these areas because the quality of the runoff product we used is very poor below this threshold (Fekete et al., 2002), so any result would be unreliable. In addition, we also identified unrealistic, huge diclofenac concentrations in large urban areas due to unrealistic representation of river reaches and water infrastructure at our working resolution in these areas (sewage infrastructure in large urban areas is not accounted for in our model). To overcome this limitation, no diclofenac concentration is reported for cells accumulating contaminant mass for less than three upstream cells i.e., in Eq. (5). The two filters described above exemplify how the interpretation of GLOBAL-FATE outcomes depends on the available input datasets, both in terms of quality and resolution. Considering that working resolution and input datasets are user-dependent in GLOBAL-FATE, the criteria to assess model results quality and reliability are case dependent, and the filters suggested here may not be convenient in all circumstances. In any case, users must be aware that the simplified representation of complex processes like water and contaminant routing along natural and engineered systems currently coded in GLOBAL-FATE implies serious limitations on the spatial scale at which the model delivers meaningful results (see Section 4 for a comprehensive discussion on this issue)."

And also in lines 433-458:

"In our example, the combination of the working spatial scale (1/16 of a degree, ~7 km), the complexity of fine-scale interactions between engineered systems and the river network (e.g., the exact location of effluent discharges, extensive sewage networks, poor representation of small streams), and the input data available translates into several model inadequacies that pose limits on the interpretability of the results. We already mentioned that the quality of the runoff map precludes the interpretation of any results for regions where runoff is below 100 mm year^{-1}, and that the calculated concentration are unreliable for watersheds smaller than ~150 km^2 (this roughly relates to river reaches of ~20 km) due to inexact effluent discharge locations in small streams and the absence of data on sewage networks in large urban areas, that would route the contaminant load downstream towards larger rivers resulting in higher dilution and lower contaminant concentration. These limitations were easily spotted as they resulted in very unrealistic high diclofenac concentrations scattered throughout the global network, which attracted our immediate attention. However, other assumptions of the modelling approach do not leave such a conspicuous mark in the model output. For instance, consumption data is homogeneous at the country level, while variability inside large countries may be substantial (urban vs. rural regions, for instance). Also, we have averaged information on intensity of treatment also at the country scale, when this may change even at very local scales. This implies that the model results are not necessarily unbiased beyond the threshold mentioned earlier (150 km^2), because all uncertainties and biases propagating from model inputs and assumptions must also have a reflection in the spatial dimension at varying scales. For instance, the comparison between observed and modelled diclofenac concentration along the main axis of the Rhine river (Fig. 7) shows that the model was able to spot a concentration increase at 300 km upstream the river mouth (in the sense that the model predicts an increase that goes beyond 100 ng L^{-1}, the basic threshold we were interested in). However, in the same basin close to the river mouth (∼50 km) the model could not mimic an increase in concentration beyond 100 ng L^{-1}. Our opinion is that GLOBAL-FATE, as implemented in the example, should be used to answer questions which are general in nature. For instance, "contaminant concentration downstream large urban areas in Central Europe frequently exceeds 100 ng L^{-1}", and related statements concerning remediation measures. We advise against the use of GLOBAL-FATE as implemented in the example to support statements concerning particular places at or near the working resolution (for instance "the remediation measures seem insufficient to lower concentrations below 100 ng L^{-1} downstream from Cologne")."

The second point is also addressed in the introduction, lines 426-433:

“One of the features of GLOBAL-FATE is that it is not fundamentally associated with a spatial resolution or extent. Users can define the working spatial resolution and extent just adapting the resolution of the raster inputs and the region of interest (for instance, a single continent or subcontinent). Although this is an obvious advantage over other large-scale contaminant models, it also harbors the significant risk that users may assume that the model delivers meaningful results at any working spatial scale. We strongly advise against the uncritical use of GLOBAL-FATE, particularly when working at coarse working resolutions or with highly spatially aggregated input data. We do not want to suggest a spatial resolution threshold from which results from GLOBAL-FATE could be considered as reliable, because the criteria to assess model results quality and reliability are case dependent, and guidelines suggested in a given situation may not be convenient in all circumstances.”

and lines 468-480:

“Nonetheless, we discussed the limitations of GLOBAL-FATE as applied in our example (~7 km pixel resolution), but even exercises using models working at much finer resolution in smaller areas (e.g., China at 0.5 km resolution, Grill et al. 2018) found substantial uncertainties related to unaccounted variability regarding input variables and poor representation of small streams. Therefore, we strongly suggest to carefully assess model performance irrespective of the working resolution, and to pay special attention to the spatial scales at which answers are required and its compatibility with the aggregation of input information and the representation of the river network. Finally, GLOBAL-FATE includes very simplified physically based approximations for attenuation in the river network, which a priori are mathematically robust to changes in the spatial resolution, but that assumes homogeneous properties along calculation units (river reaches) such as water velocity and mixing. Although those assumptions do not hold even at very local scales (tens of meters), empirical research on river ecology suggests that this approach is reasonable for rivers reaches up to ~10 km (Marcé et al., 2018). Beyond this, substantial heterogeneity of the river network is overlooked, with potential effects on the contaminant mass balance (Darracq and Destouni, 2007).”

Also, we totally agree that physically-based formulations may lose its physical meaning when working at resolutions very far from the ones used to conceptualize the underlying model. Therefore, we included the following lines in the paper, lines 474-480:

“Finally, GLOBAL-FATE includes very simplified physically based approximations for attenuation in the river network, which a pripori are mathamatically robust to changes in the spatial resolution, but that assumes homogenous propierties along calculation units (river reaches) such as water velocity and mixing. Although those assumptions do not hold even at very local scales (tens of meters), empirial reearch on river ecology suggests that this approach is reasonable for rivers reaches up to ~10 km (Marcé et al., 2018). Beyond this, substantial hetereogenity of the river network is overlooked, with potential effects on the contaminant mass balance (Darracq and Destouni, 2007).”

So what new thing/contribution is being brought? Is it a science contribution, a software contribution that allows others as well as you to take the science further, which is still a contribution, or both?
Please, see the response to the first comment that already addressed this concern.

I also asked a colleague without a specific understanding of contaminant transport, but with a strong computational modelling background, to do a usability review, which I provide below: as per the comments on science and code, note its not damning but not yet convinced if its great worth

We thank the reviewer for this detailed technical assessment, and we want to apologize because we upload the wrong version of the plug-in code into GitHub, which was the ultimate reason of the problems encountered when trying to use it, as detailed below.

A technical review of GLOBAL-FATE: A GIS-based model for assessing contaminants fate in the global river network by Carme Font et al. (2019)

Main executable
I was able to get the GLOBAL-FATE executable running on my PC and tested it with the sample data provided. It took approximately 20 minutes to run, but my laptop processor (Intel Core i5-2435M CPU @ 2.40GHz) is not as powerful as the one mentioned in the paper, so this running time seems about right compared with the five minutes given in the paper.

It would be useful to provide Windows binaries in the repository rather than users having to compile them themselves. As I was not familiar with this process this took me quite some time. If you would prefer not to supply the binaries, then some clearer instructions would be helpful, especially with regards to installing Cygwin.

Regarding the installation of the model, we are not allowed to load executable files in Github, but we have included clear indication that executables will be sent to any user under request, both in the main body of the paper (under the section Code availability) , and also in the instructions at the GitHub site.

While there are some comments within the C code, these could be improved and added to, to allow users such as myself to gain a better understanding of what the code is doing.

We have included a lot more comments in the code. It would be cumbersome to detail here everything we added, but we invite the referee to consult the source code files to have an idea of the commenting level of the new version, which we think it is high.

QGIS plugin

I also tried to use the GLOBAL-FATE QGIS plugin. It would be helpful if you mention that the plugin is only compatible with QGIS 2 and not QGIS 3 as I first installed QGIS 3.6 in order to try this plugin but was told that I had to use QGIS 2 when I tried to install the plugin (I used QGIS 2.18). It would also be useful to provide some brief instructions on how to install the QGIS plugin for non-expert users such as myself.

The first version of the plugin was made for QGIS 2, now we have been working on a new version for QGIS 3. Therefore, the final version of the plug-in will work on QGIS 3. For the installation of the plugin, we updated the file in Github explaining all the steps to follow:
https://github.com/icra/GLOBALFATE/blob/master/QGIS%20plug-in/INSTALL.txt

When running the plugin I came across the error that the directory 'C:\tmp' had not been created (IOError: [Errno 2] No such file or directory: 'C:/tmp/dir.txt'). This error could be mitigated by either creating the directory for the user if it does not exist, or by asking the user for a temporary directory as one of the inputs. I simply created the directory as a workaround to this problem.

We apologize because we also encountered this error during plug-in development, and it was already solved. However, for some unknown reason, the version of the plug-in we uploaded to Github still included the bug, which we see precluded a proper assessment of the tool. As the referee suggested, the workaround consisted in creating creating the directory directly from the python script, so the user wouldn't have to bother on it.

The input parameters are split into two dialog boxes. Could these be combined into one dialog box as this could be more intuitive to the user?

The reason of two dialog boxes is that you can switch off some modules of the code (for instance, the hydrological calculations), for instance to run different scenarios. This is why we decided to split the input in two boxes. The second box only appears if hydrological calculations are performed. We think that this is better than having a single window, because it may confound users that will input those files even in cases when they are not necessary.

When the GLOBAL-FATE plugin started executing, another dialog box popped up immediately giving the elapsed time. I was unable to determine why the GLOBAL-FATE code was not executing, so I was unable to run the plugin successfully. Also, I was unsure where the data would be saved to, or if maps showing the data would just load within QGIS.

We apologize again because we also encountered this error during plug-in development, and it was already solved. However, for some unknown reason, the version of the plug-in we uploaded to Github also included this bug, which we see precluded a proper assessment of the tool. The point of this error was the presence of an absolute address that precludes the execution of the model in different computers. This bug is now solved, with a box in the main menu of the plugin, where the user can specify the directory in which to save the results. We apologize the referee could not assess the tool because this error. Just for the record, the main result, i.e. the map of contaminants concentrations, is automatically loaded in the canvas.

Response to Referee Alberto Bellin

REFEREE: This manuscript presents a generalization of the FATE model for global applications. The model inherits all the simplifying assumptions and limitations of FATE and focuses on providing a GIS platform suitable for global applications.

AUTHORS: We appreciate the thoughtful revision by Dr. Alberto Bellin, and want to make some precisions about our modelling strategy and additional clarifications that we think addresses the reviewer concerns. First, we want to make clear that GLOBAL-FATE is not a revision or upgrade of the PhATE model, as the first comment by the reviewer seems to suggest. Although some assumptions and approaches are shared between GLOBAL-FATE and PhATE, and in fact with a number of other contaminant models, the development of our model has been totally independent of the former or any other contaminant model (except of course for the inspiration and guidance collected from all past work on large scale modelling we found in the literature).

General comments
The modeling part is simplistic, as in FATE, and boils down to the application of the following first order decay equation, providing the contribution of a cell to the annual load observed at the reference cell: $L = L_0 \exp[-k\tau]$, where L is the contaminant annual load [g/year] at the reference cell, L_0 is proportional to the population of the contributing cell and τ is the residence time from the contributing cell to the reference one. The loads of all the cells contributing to the reference one are added such as to obtain the total load, which is then divided by the annual water discharge at the same cell. This approach does not capture important mechanisms, such as seasonality in the releases, temperature and hydrology, which may cause significant fluctuations of the contaminant concentration. This is somewhat acknowledged by the authors in the discussion.

We totally agree that our modelling approach is simple, but we do disagree with the adjective “simplistic”. We decided to work with such a simple model structure because we understand that working at global scales precludes any attempt to parameterize a complex model including a lot of processes, such as relationship with temperature or seasonality of releases. This is particularly true if the available information collected in the field is scarce, as it is the case. A limitation to the level of detail captured in GLOBAL-FATE is the need of consistent and complete datasets with global coverage and the variety of sources and unknown sampling methodologies that make difficult to use the data as reference datasets.

It is a piece of fundamental knowledge that any attenuation process in the river network depends on temperature one way or the other, but there are simply not enough data in the literature to parameterize this dependence at large scales. And the seasonality of releases, as modelled by GLOBAL-FATE, would need global, gridded information on seasonality of population. To the best of our knowledge, such a data product does not exist yet. All in all, our choices concerning model structure were not the result of a naive approach to the problem (and thus “simplistic”), but of a careful consideration of pros and cons considering the kind of questions we are anticipating to answer with a model like this and the information available to parameterize a model at such large scales.

There is no doubt that hydrology exerts a prominent role in defining pollutant concentrations in the river network, including seasonal variations. In fact, we already make this point clear in section 4 in the manuscript, where we discuss limitations of GLOBAL-FATE. However, working at annual, average streamflow conditions does not preclude the model being useful to answer many relevant questions concerning pollution in the river network. It precludes, for instance, to answer questions about the impacts of extreme events, although there is the possibility to run the model for different synoptic situations. We do not deny that seasonal and short-term hydrological variability is relevant for contaminant transport, we simply do not intent to answer questions related to this variability with GLOBAL-FATE. Needless to say, the decision to work with annual averaged streamflow assuming steady-state was also related to the complexity and computing needs that come with a dynamical hydrological model working at daily or hourly time steps.

However, we understand the referee’s concerns, and have included a number of warnings and considerations in the manuscript. The main modifications are listed here:

- We made clear that the degree of simplification in GLOBAL-FATE is similar to other large-scale contaminant models, such as for instance the most recent development in Grill, G., Li, J., Khan, U., Zhong, Y., Lehner, B., Nicell, J., Ariwi, J.: Estimating the eco-toxicological risk of estrogens in China's rivers using a high-resolution contaminant fate model, Water Research, 145, 707-720, https://doi.org/10.1016/j.watres.2018.08.053, 2018.

- We expanded the discussion on the limitations of our modelling approach due to the rough temporal resolution for hydrology. Particularly, we have in line 466 these sentences:

“GLOBAL-FATE is a steady state model, and although synoptic conditions like low o high flows or climate change scenarios can be modelled, it cannot dynamically simulate extreme events or seasonality. This should be considered when formulating research or management questions for which hydrological seasonal or subseasonal variability is relevant.”

Recently, we published a more comprehensive model (Diamantini et al., 2019), which under suitable assumptions can be reduced to the approach presented in the manuscript, but that is more general and allows to take into account the above processes. This previous published work includes also the effect of lakes that the authors claim they introduced for the first time. I think that the work we did is relevant to this contribution since it represents a generalization of the proposed approach.

The model introduced by Diamantini et al. (2019) is a very fine work, and we agree that constitutes a basic antecedent that we have incorporated in our discussion about the limitations of our approach. It nicely incorporates time-varying forcing functions like population and hydrology. And, indeed, it links lakes and reservoirs with the river network. However, the fact that the authors simulated a small-medium watershed (12.000 km^2) to exemplify the applicability of their model already shows which are the spatial scales for which this model has been conceptualized. In our opinion and making clear that we do not diminish in any way Diamantini and co-workers' approach, that we like very much, it is not fair comparing the complexity in terms of model structure of models devised to work at such diverging spatial and temporal scales. On one hand, we posit that applying Diamantini's model at the global scale using the same approach as in the original paper would imply a gigantic (and most probably unsuccessful) effort related to parameterization of the model and computing resources. On the other hand, we acknowledge that applying GLOBAL-FATE to answer the questions posed by Diamantini and co-workers in the Adige basin would be inappropriate. We tried to incorporate this reasoning in the new version of the manuscript in several places, but we think this new sentence catches the point quite conveniently (line 443):

"We acknowledge that this restriction limits the usability of the current version of GLOBAL-FATE to answer questions that require precise information at a scale of ~20 km of river network. In such cases, models operating at local (i.e., single watershed scale) or regional (i.e., country level) scales may be a better option (e.g., Diamantini et al., 2019)."

I appreciated the disclaimer the authors introduced in the conclusions, where they warned users against the application of the model at what they call the "very local" scale. However this scale is not adequately defined, though by mentioning the watershed scale as an example of scale at which the model cannot be applied and the following suggestion of not using the model "below the country level" provides some, but still ambiguous, guidelines. This notwithstanding, the disclaimer poses strong limitations to the analyses that can be done and a more comprehensive discussion about the limits of applications is needed, in my view, to avoid misuses of the proposed model. Considering that the model cannot provide valuable information at important scales, such as the watershed scale and downstream large urban areas (see sentence beginning at line 9 of page 15), where the impacts are evaluated, I am wondering what type of indications the model can actually provide, besides suggesting the reduction of drug consumption, a recommendation that can be done by considering the total consumption based on census information. In other words, my concern is that hydrological processes may not be so relevant for the type of questions that the model can actually answer, considering the level of simplification introduced, thereby making this model not clearly preferable to alternative approaches, such as simple regressions or machine learning, for example. A discussion supporting the utility of the model is needed here.

We thank Dr. Bellin for his insightful comments on the spatial scales at which GLOBAL- FATE deliver meaningful and usable results, and the implications for the overall value of our model. We agree that we were not particularly brilliant at this respect, as we introduced some ambiguity and vagueness that did not help to convey the message. In fact, this point also puzzled the Associate Editor, who also wondered what was the main contribution of our approach.

In the new version, we have tried to locate our model in the landscape of contaminant models in a more explicit way, making a strong case for the step forward GLOBAL-FATE constitutes. First, making clear which is the technical novelty GLOBAL-FATE offers (lines 55-76):

"Recently, other approaches specifically designed for very large scales have used a Geographical Information System (GIS) framework to solve the routing of chemicals along the river network (Pistocchi et al., 2012; Dumont et al., 2015; Grill et al., 2016; Rice and Westerhoff 2017). Most of these models use a much simpler model parameterization, in order to make continental and global calculations accessible. However, some of them assume that chemicals do not decay when travelling through the river network, and simply rely on dilution factors once pollutants enter in the river network. Further, they work at a fixed spatial scale which is either very rough to adequately represent the river network (e.g., 0.5 degrees), or too detailed to be practical for global calculations due to computational requirements (e.g., 500 m, Grill et al., 2018).

GLOBAL-FATE has been designed to overcome these constraints, offering the first contaminant fate model operating at the global river network, including lakes and reservoirs, which is at the same time open-source, multiplatform, user-friendly, and modular. This will make global contaminant calculations accessible to a much wider community of scientists and practitioners, opening the door for including pharmaceutical pollution into influential assessments of climate change impacts (e.g., the Inter Sectoral Impact Model Intercomparison project) and global policy instruments like the UN Sustainable Development Goals agenda. GLOBAL-FATE calculates the steady-state concentration of a user-defined down-the-drain contaminant through the global river network, including lakes and reservoirs. GLOBAL-FATE is offered as an open-source, GIS-based model programmed in the C language, allowing researchers to select the input information (water routing, hydrology, population, etc.) and the spatial resolution at which the model has to perform. So forth, the model can include new or different hydrological datasets and other input information, and hence it is not fundamentally restricted to a single modelling resolution, hydrological, or socio-economic scenario. The model simulates the propagation of down-the-drain contaminants along the river network, and the constituent decreases at a rate proportional to its concentration in the aquatic media. GLOBAL-FATE is also computationally efficient, can be run in Windows or Linux machines, and can take advantage of parallel computing in multi-processor computers or clusters. It can also be run as a user-friendly plug-in in QGIS, and the modular structure of its code allows switching different functions of the model on and off. ”

Also, from line 461 on, we state that:

“In any case, GLOBAL-FATE can be used to test the effectiveness of large scale management strategies related to pharmaceutical consumption control and wastewater treatment implementation and upgrading, in order to deliver influential assessments of climate change impacts on pharmaceutical consumption and river network ecosystem health (e.g., the Inter Sectoral Impact Model Intercomparison project), and also for informing global policy instruments like the UN Sustainable Development Goals agenda. This is already common practice in other sectors using large scale, coarse resolution models such as impacts of climate change on marine life (Lotze et al., 2019), on lake physics (Woolway and Merchant 2019), on soil moisture (Samaniego et al., 2018), or on economic losses due to river flooding (Dottori et al., 2018), to cite just a few recent examples.”

The authors remark that GLOBAL-FATE is not associated to a spatial resolution, or extent, and consider this as the "main strength" of the proposed approach. I disagree with this conclusion. The size of the cell has an impact on the way the river systems are represented and a coarse gridding may produce inaccurate estimates of the residence time. For instance, the raster of 1/16 degree used in the example of application is already too coarse and does not guarantee a good reproduction of the river system in densely populated areas, such as in Europe for example. On the other hand, this gridding may be ok in large rivers with low population density, but as a consequence with low impact. An upper limit should be indicated here and a warning to avoid improper applications with large cells should be issued.

We acknowledge that we somehow oversold the scale-free feature of GLOBAL-FATE, because although it is a potential advantage over other available models, it also leaves the door open for gross misuses of the model. To avoid this, we worked in two directions: first, we substantially expanded the section where we assess the limitations of GLOBAL-FATE as implemented in the example, including a new figure showing detailed results for a single watershed; and second, we included a clear warning in the discussion about the use of GLOBAL-FATE at low resolutions.

For the first point, we included the following text (lines 363-375):

“The concentration maps in Figs. 3 and 4 do not show pixels with less than 100 mm year-1 of runoff, which correspond to arid regions. We decided to discard concentration values in these areas because the quality of the runoff product we used is very poor below this threshold (Fekete et al., 2002), so any result would be unreliable. In addition, we also identified unrealistic, huge diclofenac concentrations in large urban areas due to unrealistic representation of river reaches and water infrastructure at our working resolution in these areas (sewage infrastructure in large urban areas is not accounted for in our model). To overcome this limitation, no diclofenac concentration is reported for cells accumulating contaminant mass for less than three upstream cells i.e., in Eq. (5). The two filters described above exemplify how the interpretation of GLOBAL-FATE outcomes depends on the available input datasets, both in terms of quality and resolution. Considering that working resolution and input datasets are user-dependent in GLOBAL-FATE, the criteria to assess model results quality and reliability are case dependent, and the filters suggested here may not be convenient in all circumstances. In any case, users must be aware that the simplified representation of complex processes like water and contaminant routing along natural and engineered systems currently coded in GLOBAL-FATE implies serious limitations on the spatial scale at which the model delivers meaningful results (see Section 4 for a comprehensive discussion on this issue).”

And also in lines 433-458:

“In our example, the combination of the working spatial scale (1/16 of a degree, ~7 km), the complexity of fine-scale interactions between engineered systems and the river network (e.g., the exact location of effluent discharges, extensive sewage networks, poor representation of small streams), and the input data available translates into several model inadequacies that pose limits on the interpretability of the results. We already mentioned that the quality of the runoff map precludes the interpretation of any results for regions where runoff is below 100 mm year^{-1}, and that the calculated concentration are unreliable for watersheds smaller than ~150 km^2 (this roughly relates to river reaches of ~20 km) due to inexact effluent discharge locations in small streams and the absence of data on sewage networks in large urban areas, that would route the contaminant load downstream towards larger rivers resulting in higher dilution and lower contaminant concentration. These limitations were easily spotted as they resulted in very unrealistic high diclofenac concentrations scattered throughout the global network, which attracted our immediate attention. However, other assumptions of the modelling approach do not leave such a conspicuous mark in the model output. For instance, consumption data is homogeneous at the country level, while variability inside large countries may be substantial (urban vs. rural regions, for instance). Also, we have averaged information on intensity of treatment also at the country scale, when this may change even at very local scales. This implies that the model results are not necessarily unbiased beyond the threshold mentioned earlier (150 km^2), because all uncertainties and biases propagating from model inputs and assumptions must also have a reflection in the spatial dimension at varying scales. For instance, the comparison between observed and modelled diclofenac concentration along the main axis of the Rhine river (Fig. 7) shows that the model was able to spot a concentration increase at 300 km upstream the river mouth (in the sense that the model predicts an increase that goes beyond 100 ng L^{-1}, the basic threshold we were interested in). However, in the same basin close to the river mouth (∼50 km) the model could not mimic an increase in concentration beyond 100 ng L^{-1}. Our opinion is that GLOBAL-FATE, as implemented in the example, should be used to answer questions which are general in nature. For instance, "contaminant concentration downstream large urban areas in Central Europe frequently exceeds 100 ng L^{-1}", and related statements concerning remediation measures. We advise against the use of GLOBAL-FATE as implemented in the example to support statements concerning particular places at or near the working resolution (for instance "the remediation measures seem insufficient to lower concentrations below 100 ng L^{-1} downstream from Cologne")."

The second point is also addressed in the introduction, lines 426-433:

"One of the features of GLOBAL-FATE is that it is not fundamentally associated with a spatial resolution or extent. Users can define the working spatial resolution and extent just adapting the resolution of the raster inputs and the region of interest (for instance, a single continent or subcontinent). Although this is an obvious advantage over other large-scale contaminant models, it also harbors the significant risk that users may assume that the model delivers meaningful results at any working spatial scale. We strongly advise against the uncritical use of GLOBAL-FATE, particularly when working at coarse working resolutions or with highly spatially aggregated input data. We do not want to suggest a spatial resolution threshold from which results from GLOBAL-FATE could be considered as reliable, because the criteria to assess model results quality and reliability are case dependent, and guidelines suggested in a given situation may not be convenient in all circumstances."

and lines 468-480:

"Nonetheless, we discussed the limitations of GLOBAL-FATE as applied in our example (~7 km pixel resolution), but even exercises using models working at much finer resolution in smaller areas (e.g., China at 0.5 km resolution, Grill et al. 2018) found substantial uncertainties related to unaccounted variability regarding input variables and poor representation of small streams. Therefore, we strongly suggest to carefully assess model performance irrespective of the working resolution, and to pay special attention to the spatial scales at which answers are required and its compatibility with the aggregation of input information and the representation of the river network. Finally, GLOBAL-FATE includes very simplified physically based approximations for attenuation in the river network, which a priori are mathematically robust to changes in the spatial resolution, but that assumes homogeneous properties along calculation units (river reaches) such as water velocity and mixing. Although those assumptions do not hold even at very local scales (tens of meters), empirical research on river ecology suggests that this approach is reasonable for rivers reaches up to ~10 km (Marcé et al., 2018). Beyond this, substantial heterogeneity of the river network is overlooked, with potential effects on the contaminant mass balance (Darracq and Destouni, 2007)."

Detailed comment

I am wondering how the value that the NS assumes after log-transforming the data compares with that obtained without the transformation. In Figure 5 the points are rather disperse and this may be due to the attenuating effect of errors when the log-transform is applied.

Concerning the comment about the log scale used for comparing observed and modelled values, we had no other option considering that the magnitude of the errors was proportional to the modelled value. This effect in a modelling exercise spanning 3 orders of magnitude forced us to use the log scale for a proper calibration of the tool.

Response to Executive editor Astrid Kerkweg

COMMENT: In my role as Executive editor of GMD, I would like to bring to your attention our Editorial version 1.1: http://www.geosci-model-dev.net/8/3487/2015/gmd-8-3487-2015.html. This highlights some requirements of papers published in GMD, which is also available on the GMD website in the 'Manuscript Types' section: http://www.geoscientific-model-development.net/submission/manuscript_types.html. In particular, please note that for your paper, the following requirements have not been met in the Discussions paper:

15• "The main paper must give the model name and version number (or other unique identifier) in the title."
• "If the model development relates to a single model then the model name and the version number must be included in the title of the paper. If the main intention of an article is to make a general (i.e. model independent) statement about the usefulness of a new development, but the usefulness is shown with the help of one specific model, the model name and version number must be stated in the title. The title could have a form such as, "Title outlining amazing generic advance: a case study with Model XXX (version Y)"."

AUTHORS: We now refer to the model as GLOBAL-FATE (version 1.0.0) in the paper and the Github repository.

• "All papers must include a section, at the end of the paper, entitled 'Code availability'. Here, either instructions for obtaining the code, or the reasons why the code is not available should be clearly stated. It is preferred for the code to be uploaded as a supplement or to be made available at a data repository with an associated DOI (digital object identifier) for the exact model version described in the paper. Alternatively, for established models, there may be an existing means of accessing the code through a particular system. In this case, there must exist a means of permanently accessing the precise model version described in the paper. In some cases, authors may prefer to put models on their own website, or to act as a point of contact for obtaining the code. Given the impermanence of websites and email addresses, this is not encouraged, and authors should consider improving the availability with a more permanent arrangement. After the paper is accepted the model archive should be updated to include a link to the GMD paper." Note, that the exact code version described in this article should be permanently accessible. Thus please consider to make the exact version, your article refers to, available via a permanent archive providing a DOI (e.g. Zenodo). Additionally, please add a version number identifying this version to the title of your article upon submission of the revised manuscript.

AUTHORS: We are storing the exact version of the code in a Github repository. We are ready to link this to Zenodo, but we are still waiting for permission of the owner of the ICRA Github repository to make an explicit link of the GLOBAL-FATE repository to Zenodo. After this step, which is going to be completed by early October 2019, we will be able to provide a DOI for the GLOBAL-FATE (version 1.0.0) repository.

GLOBAL-FATE (version 1.0.0): A GIS-based model for assessing contaminants fate in the global river network

Carme Font[1,2], Francesco Bregoli[1,*], Vicenç Acuña[1,2], Sergi Sabater[1,3], Rafael Marcé[1,2]

[1] Catalan Institute for Water Research (ICRA), Emili Grahit 101, 17003 Girona, Spain
[2] University of Girona, Girona, Spain
[3] Institute of Aquatic Ecology, University of Girona, Campus Montilivi, 17071 Girona, Spain
[*] Current address: Water Science and Engineering Department, IHE Delft Institute for Water Education, Westvest 7, 2611 AX, Delf, The Netherlands

Correspondence to: Rafael Marcé (rmarce@icra.cat)

Abstract. GLOBAL-FATE is the first open-source, multi-platform, user-friendly, and modular contaminant fate model operating at the global scale linking human consumption of pharmaceutical-like compounds with their concentration in the river network. GLOBAL-FATE simulates human consumption and excretion of pharmaceuticals, the attenuation of the contaminant load in wastewater treatment plants, as well the attenuation of the contaminant load in river reaches, lakes, and reservoirs, as a first order decay depending on residence time. We provide a comprehensive description of model equations and the overall structure of the model, with spacial attention to input/output datasets. GLOBAL-FATE is written in C, can be compiled in any platform, and uses inputs in standard GIS format. Additionally, the model can be run inside QGIS as a plug-in. The model has no built-in working resolution, which depends on the intended use and the availability of appropriate model inputs and observed data. We exemplify the application of GLOBAL-FATE solving the global concentration of diclofenac in the river network. A comparison with a dataset of diclofenac concentration observations in rivers suggest that GLOBAL-FATE can be successfully applied in real case modelling exercises. The model is particularly sensitive to the generation of contaminant loads by human pharmaceutical consumption, and to the processes governing contaminant attenuation in the river network. GLOBAL-FATE will be a valuable tool for the scientific community and the policymaking arena, and could be used to test the effectiveness of large scale management strategies related to pharmaceutical consumption control and wastewater treatment implementation and upgrading.

1 Introduction

The United Nations 2030 Agenda for Sustainable Development identifies 17 master goals, amongst which is the availability and sustainable management of water and sanitation. This agenda establishes as a goal the improvement of water quality by reducing pollution, eliminating dumping, and minimizing release of hazardous chemicals to the river network by 2030 (UN, 2015). However, a large proportion of surface water networks is currently severely affected by sewage inputs from waste water treatment plants (WWTP) or by direct sewage disposal (Richardson et al., 2005; Stewart et al., 2014; Hernández et al., 2015; K'oreje et al., 2016). Sewage disposal inputs organic matter, nutrients, and fecal bacteria to the river network, together with a whole plethora of chemicals related to household human activities. These include micro-plastics and nanomaterials (Besseling et al., 2017), pharmaceuticals (Li et al., 2016), personal care products (Arlos et al., 2014), and even illicit drugs (Postigo et al., 2010). This increment of down-the-drain chemicals reaches to the river network and affect both humans and biodiversity and ecosystem function (Rudd, 1970), so forth posing at risk water security (Goldman and Koduru, 2000; Vörösmarty et al., 2010).

Assessing chemical discharges and their fate in the river network is thus vital to evaluate both the health of aquatic ecosystems and the security of water supplies for human needs. This requires adequate models to consider the spread and dynamics of chemicals at large spatial scales, both for assessing the current water quality status in regions with poor monitoring programs coverage (Strokal et al., 2019), and for planning management and mitigation measures. Existing models approach the fate of contaminants in multimedia (air, water, soil) and using steady-state models working at regional scales such as ChemCAN, HAZCHEM, or EUSES (MacLeod et al., 2011; Gouin et al., 2013; Lindim et al., 2016). Others are process based, operating at the watershed scale, and perform as dynamical in-stream water quality models, such as MIKE11, SWAT, WASP, QUAL2E, or DELWAQ (Liang et al., 2015; Santhi et al., 2005; Di Toro et al., 1983; Brown et al., 1987; Van Wijngaarden, 1999). Another set of models analyse the dynamics of down-the-drain pollutants, considering the linkages between engineered systems (e.g., WWTP) and natural systems (e.g., rivers). This includes PhATE, GREAT-ER, LF2000-WQX, STREAM-EU, iSTREEM, or ePiE (Anderson et al., 2004; Feijtel et al., 1997; Johnson et al. 2007; Boxall et al., 2014; Lindim et al., 2016; Kapo et al., 2016, Oldenkamp et al., 2018), and frequently target on pharmaceutical products though are also suitable to simulate the fate of any compound decreasing following first order decay dynamics (Table 1). Particularly, models in Table 1 are applied for chemicals whose dominant emission source to the environment is via WWTP effluents. Most of these models are highly data demanding and use many adjustable parameters. This makes some of them computationally inefficient; others have non-open source codes, which make their use for global or continental scale calculations cumbersome.

Recently, other approaches specifically designed for very large scales have used a Geographical Information System (GIS) framework to solve the routing of chemicals along the river network (Pistocchi et al., 2012; Dumont et al., 2015; Grill et al., 2016; Rice and Westerhoff 2017). Most of these models use a much simpler model parameterization, in order to make continental and global calculations accessible. However, some of them assume that chemicals do not decay when travelling through the river network, and simply rely on dilution factors once pollutants enter in the river network. Further, they work at a fixed spatial scale which is either very rough to adequately represent the river network (e.g., 0.5 degrees), or too detailed to be practical for global calculations due to computational requirements (e.g., 500 m, Grill et al., 2018).

GLOBAL-FATE has been designed to overcome these constraints, offering the first contaminant fate model operating at the global river network, including lakes and reservoirs, which is at the same time open-source, multi-platform, user-friendly, and modular. This will make global contaminant calculations accessible to a much wider community of scientists and practitioners, opening the door for including pharmaceutical pollution into influential assessments of climate change impacts (e.g., the Inter Sectoral Impact Model Intercomparison project) and global policy instruments like the UN Sustainable Development Goals agenda. GLOBAL-FATE calculates the steady-state concentration of a user-defined down-the-drain contaminant through the global river network, including lakes and reservoirs. GLOBAL-FATE is offered as an open-source, GIS-based model programmed in the C language, allowing researchers to select the input information (water routing, hydrology, population, etc.) and the spatial resolution at which the model has to perform. So forth, the model can include new or different hydrological datasets and other input information, and hence it is not fundamentally restricted to a single modelling resolution, hydrological, or socio-economic scenario. The model simulates the propagation of down-the-drain contaminants along the river network, and the constituent decreases at a rate proportional to its concentration in the aquatic media. GLOBAL-FATE is also computationally efficient, can be run in Windows or Linux machines, and can take advantage of parallel computing in multi-processor computers or clusters. It can also be run as a user-friendly plug-in in QGIS, and the modular structure of its code allows switching different functions of the model on and off. Here we describe the structure, functioning, and strengths and limitations of GLOBAL-FATE. First, we explain the structure and functioning of the model, focusing on the type of input data structure and the formulation of the different hydrological and biogeochemical processes. Then, the application of the model is exemplified solving the worldwide propagation of the pharmaceutical diclofenac throughout the global river network. Finally, we discuss strengths and limitations of GLOBAL-FATE, and point at future developments.

2 Methodology

GLOBAL-FATE is a physically-based model for simulating constituent inputs to the river network and their routing along the river network at the global scale. Our approach shares key assumptions and modelling mechanisms with other large scale pharmaceutical models for the river network (i.e., Keller et al. 2006; Pistocchi 2014; Grill et al., 2019), including the use of per capita mass emissions of the contaminant of interest, simplified parameterization of losses due to human metabolism and removal in wastewater treatment plants, and dilution and first order attenuation dynamics upon discharge into natural waters. However, GLOBAL-FATE is the first model natively operating at the global scale including all those mechanisms, including explicit routing and attenuation in lakes and reservoirs.

The model uses GIS input files to solve for the contaminant fate at every element (cell) of the domain (raster). The model is multi-platform, written in C, and uses multi-core parallel computing via OpenMP. Additionally, GLOBAL-FATE can be used from QGIS (QGIS Development Team, 2018) as a plug-in, so it can be executed in a *push-button* fashion, in order to load all the layers and information the model needs in a user-friendly way, as well as automatically producing basic visualizations of the results. The code (including compilation instructions) is freely available at https://github.com/ICRA/GLOBALFATE, including the QGIS plug-ín. Pre-built executables are available under request.

GLOBAL-FATE simulates the fate of contaminants that behave as human pharmaceuticals (Fig. 1). That is, the model assumes that the origin of the contaminant load is the consumption of a pharmaceutical by population, which can differ in different regions of the World considering population density and per capita consumption. No other origins, such as diffuse sources through spreading of pharmaceutical-rich farm manure on agricultural fields, are currently included. The model assumes an excretion rate by population, and the fraction finally excreted will reach the river network either directly or after some attenuation in wastewater treatment plants (WWTPs). The fraction of load treated can be dependent on the region of the World, while the decay in WWTPs is an input parameter that applies globally. Finally, the contaminant load is routed along the river network, considering that the contaminant will decay following first-order kinetics dependent on water residence time in the river network reaches. GLOBAL-FATE also considers the presence of lakes and reservoirs in the river network and includes a particular solution for water residence time in these systems in order to calculate contaminant decay. The main output is a global map of predicted contaminant load or concentration throughout the river network.

The model workflow (Fig. 2) is based on the input of 9 global maps in the form of raster datasets and the definition of 8 parameters (Table 2). The model has been designed to work with raster data with the geographic coordinate system WGS84 in decimal degrees, but it has not a predefined spatial resolution. The only prerequisite is that all input rasters must have the same resolution and extent. Raster input data files are expected to have ASCII ESRI grid format, which makes GLOBAL-FATE very easy to set up using customary GIS software even for non-experienced users.

2.1 Model workflow

The model is composed of six main functions that are executed successively, being some outputs the inputs to the next function. First, the geographical related functions are run, which deliver streamflow and water residence time along the river network as main outputs. Then, the contaminant related functions calculate the load of contaminant to the river network by population, after discounting for wastewater treatment, and also the routing of the contaminant through the river network considering that the contaminant decays following a first order reaction dynamics. The following is a description of the calculations performed by the six main functions in GLOBAL-FATE (see also Table 3 and the code and example input files at the GitHub repository).

2.1.1 Cells area (function *Area_m2_fun.c*)

This function performs an auxiliary calculation for the flow routing function. In order to route locally generated runoff the area (m^2) of each cell in the domain is necessary. Considering WGS84 as a reference coordinate system, cells height is the length of the arc formed by the angle δ (raster resolution in decimal degrees transformed to radians) and is constant in the whole grid, calculated as:

$$H=\delta R, \tag{1}$$

where R is the authalic Earth radius ($6{,}371{,}007.2$ m). In its turn, the width of each cell depends on latitude:

$$W(y)=R\left(\sin\left(y+\frac{\delta}{2}\right)-\sin\left(y-\frac{\delta}{2}\right)\right) , \tag{2}$$

Here y corresponds to the latitude and comes from:

$$y=y_0+\delta(nr-j), j=1,\ldots,nr , \tag{3}$$

where y_0 is the southern cell latitude and nr is the number of rows in the raster, so *j* is a latitude index. Due to the fact that cells at the same latitude have equal width, both area (m^2) and width are calculated for one meridian:

$$A(y)=H\cdot W(y), \tag{4}$$

2.1.2 Flow routing (function *Flow_accumulation_m2.c*)

Streamflow in each cell of the raster is computed using a runoff accumulation approach. First, for each basin, cells are enumerated from headwaters ($l=0$) to river mouth ($l=L$), following the hierarchical organization of the river network. The J_l is the set of cells indexes at stage $l=0,\ldots,L$ .. This cells classification is obtained as a raster input dataset (Table 3), and is a typical product of many GIS hydrological algorithms (often defined as an area or flow accumulation layer). Second, we define the amount of runoff locally generated in each cell due to the precipitation-evaporation water budget. For this, we use available products of mean annual runoff (m $year^{-1}$) at the global scale in raster format, re-scaled to the same resolution as the rest of the hydrological input rasters. Finally, we route the locally generated runoff along the river network following the hierarchical order defined in the first step. Streamflow in cell j in $J_{l>0}$

is computed as the sum of runoff inputs from the surrounding cells and that generated within the cell. In order to determine the input from the neighboring cells, a raster of flow direction is used. Such raster must be encoded following the D8 method (O'Callaghan et al., 1984). Finally, average annual streamflow in m^3 $year^{-1}$ in each cell is calculated after summing up the multiplication of cells runoff (m $year^{-1}$) by the corresponding cells area (Eq. 4, m^2):

$$q_j = \sum_{i \in N_j} q_i + runoff_j A_j, j \in J_{l>0}, \quad (5)$$

where $N_j \subseteq J_{l-1}$ is the set of indexes of the neighborhood of cell j, such that $i \in N_j$ implies that flow of cell *i* goes to cell *j*, and A_j is the area of cell *j*. Note that for $l=0$, J_0 represents the set of headwater cells, where there is no neighboring inputs, and Eq. 5 simplifies to:

$$q_j = runoff_j A_j, j \in J_{l=0}, \quad (6)$$

2.1.3 Water residence time calculation in river cells (function *RT_rivers_calculator.c*)

Residence time (RT) of water in rivers is a key magnitude for the calculation of contaminant decay in the river network. RT at each cell is calculated as the division of the longitude (m) of the river reach (cell) by water velocity (m s^{-1}), i.e.;

$$RT = \frac{l}{v}, \quad (7)$$

The longitude of the flow path through the cell depends on its direction. We differentiate four cases:

$$l = \begin{cases} H, \text{if the flow has North or South direction} \\ W, \text{if the flow has East or West direction} \\ \sqrt{H^2 + W^2}, \text{if the flow has NE, NW, SE, or SW direction} \end{cases}, \quad (8)$$

H and *W* correspond to cell height and width explained in section 2.2.1. The flow velocity within the cell is calculated using the Manning equation:

$$v = \frac{1}{n} R_h^{2/3} S^{1/2}, \quad (9)$$

$$R_h = \frac{wh}{2h + w}, \quad (10)$$

where n is the Manning coefficient (s m$^{-1/3}$), R_h is the hydraulic radius (m) and S is the local slope (m m^{-1}), obtained as an external input in the form of a slope raster dataset. The Manning coefficient is applied globally, in our case we chose 0.044 s·m$^{-1/3}$ following Schulze et al. (2005). The hydraulic radius is calculated after solving for channel width (w) and depth (h) using the power functions of Leopold and Maddock (1953):

$$w=a\,q^{b},\,h=c\,q^{d} \quad , \tag{11}$$

where a,b,c and d are fitted parameters (in our case we chose $a=7.2, b=0.5, c=0.27, d=0.39$ after Andreadis et al., 2013), and q is river discharge (m^3 year^{-1}).

2.1.4 Water residence time calculation in lakes and reservoirs (function ***RT_lakes_incorporation.c***)

Lake and reservoirs are included in GLOBAL-FATE using available global databases on the location, shape, and volume of lakes and reservoirs. These spatially explicit databases must be converted into a raster with the same resolution and projection as the other hydrological rasters. The general strategy is to store all features of a given lake (volume, residence time) in the outlet cell (i.e., the cell routing the streamflow downstream from the lake), making the rest of cells of the lake as mere pipes of water and constituents to that outlet cell, where all contaminant reactions occur. Since most lakes occupy more than one cell in the network, the indexes of the cells belonging to a lake (raster of lakes location and shape) need to be indicated. Being L_j the set of indexes of the cells belonging to lake j, streamflow to lake as calculated by Eqs. 5 and 6, Q_j , corresponds to the outlet cell, i.e., the cell with maximum flow accumulation:

$$Q_j=max\{q_i, i\in L_j\}, \tag{12}$$

And the RT for the lake is the quotient of its volume, V , and streamflow (m^3 year^{-1}):

$$RT=\frac{V}{Q}, \tag{13}$$

The volume of the water bodies, V (m^3), is introduced as a raster input dataset (Table 3) in which the volume information for a particular lake is stored in the outlet cell. This implies that during RT calculation for lakes and reservoirs the cell corresponding to the lake outlet will store the annual average residence time value for the entire lake, while the rest of cells of the lake will be considered as dummy cells in terms of residence time. In its turn, this implies that during calculation of contaminant transport and reaction throughout the network, only the outlet cell of a lake will be reactive in terms of contaminant decay. Thus, the rest of cells pertaining to that lake will transport water and constituents, but all contaminant decay will take place exclusively in the outlet cell. The final implication is that lakes and reservoirs are treated in GLOBAL-FATE as point-like features, with no spatial heterogeneity. The RT raster for the river network obtained using Eq. 7 is finally updated with the RT for lakes and reservoirs (RT for the entire lake in the outlet cells, and a dummy RT value (-9999) for the rest of lake cells).

2.1.5 Contaminant load to the river network (function *Initial_contaminant_load.c*)

The contaminant load to the river network in GLOBAL-FATE is modelled for a constituent that behaves as a human pharmaceutical. Consequently, load from each cell in the raster domain is modelled as a function of the population present in each cell and several parameters accounting for consumption and excretion by population, and contaminant decay in wastewater treatment plants (WWTPs) before the contaminant mass is loaded into the river network. The contaminant load to the river network (L_0) is thus defined as:

$$L_{0,j} = \gamma\, m_j P_j \left(1 - w_{treat}\, \varepsilon\right), \qquad (14)$$

where j is the cell index, P is the population raster, m is the compound per capita consumption raster (g person^{-1} year^{-1}), usually defined at the country level), and γ is a parameter for the human excretion rate. The second term in the equation expresses the loss of contaminant due to wastewater treatment, and includes the proportion of population that is connected to WWTPs (w_{treat} usually available at the country level), and contaminant removal rate during wastewater treatment (ε), which needs to be calibrated or assigned to bibliographical values. The output of Eq. 14 is the contaminant load (g year^{-1}) discharged by any populated cell; this amount is used as initial values in the contaminant routing function.

2.1.6 Contaminant routing (function *Contaminant_accumulation.c*)

The contaminant routing along the river network assumes that once delivered to the river network the contaminant load decays following a first order reaction kinetics:

$$\frac{dC}{dt} = -kC \quad , \qquad (15)$$

where k is the first-order decay constant (hour^{-1}). After reaction during a given period of time, the remaining load will be defined by the solution of the differential in Eq. 15:

$$C(t)=C_0 e^{-kt}, \tag{16}$$

where time t would correspond in GLOBAL-FATE to the time (hours) that the constituent remains into the cell, i.e., the water residence time (*RT*) previously calculated with Eqs. 7 or 13. However, to solve the routing of the contaminant along the network, we also have to take into account the hierarchical relationship between cells. In computational terms, this function works similarly to the flow routing function, with the difference that we have to implement not only the transport of the contaminant, but also the decay in Eq. 16. In this context, the load of contaminant in a cell *j* considering loading from upstream cells and its own local population and first order decay in the cell is defined by:

$$L_j=\left(\sum_{i\in N_j} L_i+L_{0,j}\right)e^{-k\,RT_j},\ j\in J_{l>0}, N_j\subseteq J_{l-1}, \tag{17}$$

where L_i is the load from upstream cell *i*, $L_{0,j}$ is the load from local sources (Eq. 14) in cell *j*, and RT_j is residence time in cell *j*. From this load we can calculate the resulting contaminant concentration in cell *j* (C_j, g m^{-3}) with:

$$C_j=\frac{L_j}{q_j}, \tag{18}$$

where q_j is streamflow in cell *j*. Considering that we have both transport and a first order decay process, the contaminant routing must be solved respecting the hierarchical arrangement of the river network, that is, all contributing upstream cells must be calculated before a particular cell can be solved.

2.2 Coding general strategy

GLOBAL-FATE has been programmed in C. C is a compiled language, so it implements algorithms and data structures swiftly, facilitating faster computation. Furthermore, the use of loops is not as punishing as in interpreted languages, such as Python, R, Matlab, or Octave, which is relevant in a code that has loop structures to solve the water and contaminant routing. Regarding this, we integrated parallelization routines in the code using OpenMP to expedite calculations during time-consuming loop calculations and raster input/output routines. OpenMP supports multi-platform shared memory multiprocessing programming in C. It works out well for any multi-core machine, while still executable in single core computers. The model has been coded using a modular structure in several independent functions, so it is possible to skip the hydrological calculations if they are not relevant for a given analysis (for instance, different wastewater treatment scenarios can be solved without running the hydrological functions every time), but we also offer the possibility to trigger the whole model chain in a single call. The model has been also designed to take command line arguments when executed, if necessary.

This enables the use of pseudo-parallelization to run different model instances with different input arguments, for instance to perform automatic calibration or sensitivity analyses.

Some readers might be surprised by the fact that we programmed our own flow routing function instead of using a customary flow accumulation algorithm from one of the multiple hydrological GIS packages available. This stems from the fact that the contaminant routing function cannot be solved with a standard flow accumulation algorithm with a "negative weights" raster to solve for contaminant decay. The hierarchical nature of the river network is intimately related to contaminant transport and decay, and the process in non-linearly dependent on the mass present in each cell, so there is no way of defining *a priori* a "weighting" raster to solve contaminant transport with a standard flow accumulation algorithm. This means that we had to code the accumulation and decay of the contaminant so that contaminant mass is calculated in each cell appropriately. It is easy to realize that setting the first order decay constant to zero in our code gives a solution that would be similar to the one delivered by a standard flow accumulation algorithm. We decided to calculate flow routing with our algorithm to avoid using two different codes for flow routing and contaminant transport. Although both algorithms would use the same flow direction raster and thus should produce coherent results even using two different codes, we preferred to ensure a total coherence between the two solutions (water and contaminant). Moreover, the fact that our code is programmed in a compiled language with OpenMP parallelisation for loops makes our flow routing algorithm as efficient as any customary GIS flow accumulation function.

3. Example model application: concentration of diclofenac in the global river network

Here we exemplify the application of GLOBAL-FATE, simulating the concentration of diclofenac in the river network at the global scale. Diclofenac is a non-steroidal anti-inflammatory drug used as an analgesic, anti-inflammatory and antipyretic for humans (Todd and Sorkin, 1988). Diclofenac enters the environment through treated or non-treated wastewater discharges (Pistocchi et al., 2012) and it has been shown affecting aquatic organisms (Nassef et al., 2010). Furthermore, this pharmaceutical was included in the EU watch list of emerging contaminants of the Water Framework Directive by the European Commission (EC) under the Water Framework Directive (WFD), as well as by the US Environment Protection Agency (US EPA), with a proposed maximum acceptable concentration of 100 ng L^{-1} (Acuña et al., 2015).

3.1 The input data sets

All rasters in this example were re-scaled and adjusted to match a resolution of 1/16 deg (δ), with extreme positions $x_0 = -180$ (western cell position) and $y_0 = -56$ (southern cell position), and for extension $nr = 2240$ (number of rows) and $nc = 5760$ (number of columns). We want to stress here that the following collection of datasets is just one possible choice; GLOBAL-FATE is not restricted to work with those datasets or resolutions. Researchers are free to choose the data products that best serve the interest of the research question at place. All the example datasets are available in the GitHub repository, and correspond to the datasets identified in Table 3.

3.1.1 Morphology and Hydrology

Flow direction and area accumulation rasters. We used the Dominant River Tracing (DRT) (Wu et al., 2012), a database designed to perform macro scale hydrologic calculations, to build the global river network. We used the flow direction raster at 1/16 of a degree (approx. 7 km) in http://files.ntsg.umt.edu/data/DRT/ to generate a hierarchical cells order raster using the area accumulation algorithm in ESRI ArcGIS Spatial Analyst.

Runoff raster. As a runoff raster, we used the composite global annual runoff from Fekete et al. (2002), which consists in a raster of annual runoff with values in mm year^{-1}. The original raster was rescaled to the same resolution and extent as the other hydrological raster, disaggregating the runoff raster so that the water mass remained the same after disaggregation.

Slope raster. The slope raster was produced in QGIS from the digital elevation models at approximately 1 km resolution in HydroSHEDS (http://hydrosheds.cr.usgs.gov) and Hydro1k (USGS, https://lta.cr.usgs.gov/HYDRO1K) for regions above 60 N.

Lakes locations and shape raster. To identify the location and shape of lakes and reservoirs we merged the GRanD database for reservoirs (Lehner et al., 2011) with GLWD (Level 1) for lakes (Lehner & Doll 2004). Duplicate lakes were removed before producing the final map.

Lakes volume raster. To produce the volume raster, we first identified the pixel with the largest streamflow for each lake and reservoir, and then we stored the volume information for each lake in that particular pixel. The volume of the 41 World biggest lakes was manually introduced after literature review. For reservoirs, the GRanD database already contains the volume of each system, while for lakes volume is not available for all systems. In those cases, we calculated volume through the morphometric relationships reported in Lewis (2011).

Manning coefficient and channel form parameters. These parameters were set at the values provided in section 2.1.3.

3.1.2 Human population and diclofenac consumption

Population raster. Human population was obtained from the Gridded Population of the World version 4 (GPWv4) (Doxsey-Whitfield et al., 2015). GLOBAL-FATE has been designed to overcome these constraints, offering the first contaminant fate model operating at the global river network, including lakes and reservoirs, which is at the same time open-source, multiplatform, user-friendly, and modular. This will make global contaminant calculations accessible to a much wider community of scientists and practitioners, opening the door for including pharmaceutical pollution into influential assessments of climate change impacts (e.g., the Inter Sectoral Impact Model Intercomparison project) and global policy instruments like the UN Sustainable Development Goals agenda. GLOBAL-FATE calculates the steady-state concentration of a user-defined down-the-drain contaminant through the global river network, including lakes and reservoirs. GLOBAL-FATE is offered as an open-source, GIS-based model programmed in the C language, allowing researchers to select the input information (water routing, hydrology, population, etc.) and the spatial resolution at which the model has to perform. So forth, the model can include new or different hydrological datasets and other input information, and hence it is not fundamentally restricted to a single modelling resolution, hydrological, or socio-economic scenario. The model simulates the propagation of down-the-drain contaminants along the river network, and the constituent decreases at a rate proportional to its concentration in the aquatic media. GLOBAL-FATE is also computationally efficient, can be run in Windows or Linux machines, and can take advantage of parallel computing in multi-processor computers or clusters. It can also be run as a user-friendly plug-in in QGIS, and the modular structure of its code allows switching different functions of the model on and off.

Per capita consumption raster. The per capita consumption of diclofenac was calculated from information provided by the IMS-Health dataset for the period 2011-2013 (Acuña et al. 2015). The IMS-Health dataset includes national consumption of diclofenac for 86 nations (expressed as kilograms of consumed compound per year). Therefore, national consumption for the remaining 145 nations had to be estimated. Although IMS-Health data was only available for 38% of the global nations, these included the most populous and up to 82% of the global population. National per capita consumption for the 86 nations included in the IMS-Health dataset was estimated as the total consumption divided by the national population. The per capita consumption values of nations not included in the IMS-Health dataset were estimated as equal as the adjacent nation consumption (using Adjacent Fields function of ArcMap, ESRI; Acuña et al. (2015)).

Excretion parameter. We considered the oral application because it is the main form of administration and account for about 70% of the worldwide diclofenac sales following IMS-Health data (Zhang et al., 2008). We took $\gamma = 12.5\%$ as mean value for excretion rate (Johnson et al., 2013, $\gamma = 9.5\%$; Heberer and Feldmann, 2005, $\gamma = 10 - 15\%$; while Ternes et al., 1998, $\gamma = 15\%$).

3.1.3 WWTP and river removal

Fraction of sewage treated raster. Data of the fraction of wastewater that is treated per country were provided by the framework of “Environmental Performance Index” (EPI, Hsu et al., 2016) of the Yale University. Data were downloaded from https://epi.envirocenter.yale.edu/epi-downloads, and we produced a raster dataset with values per country.

Fraction of contaminant attenuation in WWTP and first order decay rate in the river network. The percentage of removal of diclofenac in water treatment plants, $\varepsilon = 40\%$, was decided as a tentative value between 21-40% and 69% (ranges from data in Zhang et al. (2008) and Ternes et al. (1998)). For this example, the first order decay rate in the river network was set to $k = 0.0096$ (after Pistocchi et al. 2012).

3.2 Model application and testing

Model predictions were obtained with a run-time of 5 minutes using a Desktop PC with Intel Core i5-4590 CPU 3.30 GHz and 8 GB RAM. The global concentration of diclofenac throughout the river network (Fig. 3) shows large areas of the World with very low concentration of diclofenac (mainly in boreal and tropical latitudes), while densely populated areas, particularly in Europe, Asia, and Africa show very high concentrations, sometimes beyond 100 ng L^{-1}. Thresholds of diclofenac concentrations for lowest observed effect on life concentration (LOEC, 30 ng L^{-1}) (Acuña et al. 2015) and the maximum acceptable limit proposed by the Water Framework Directive EC and the predicted non effect concentrations (PNEC) (both at 100 ng L^{-1}, Grill et al. 2016) are crossed in extensive regions of the World (Fig. 3). Simulated concentrations of diclofenac above 100 ng L^{-1} are detected in isolated areas of North America, several areas in Central America and punctually in South America. In Africa concentrations over the above thresholds occur in the occidental Mediterranean coast (Fig. 3 and 4), Nigeria, and oriental and south-east sides of the continent. Furthermore, punctual areas in European countries show very high diclofenac concentrations, with remarkable prevalence in Belgium, central Europe and Ukraine (Fig. 4). Concentrations over the thresholds are also found in occidental Asia. India and Bangladesh stand out, mainly in the Ganges basin. Several regions of China, Thailand, and Japan also show very high concentrations. Concentrations above 100 ng L^{-1} are also observed in some Indonesian islands, such as the Java Island.

The concentration maps in Figs. 3 and 4 do not show pixels with less than 100 mm $year^{-1}$ of runoff, which correspond to arid regions. We decided to discard concentration values in these areas because the quality of the runoff product we used is very poor below this threshold (Fekete et al., 2002), so any result would be unreliable. In addition, we also identified unrealistic, huge diclofenac concentrations in large urban areas due to unrealistic representation of river reaches and water infrastructure at our working resolution in these areas (sewage infrastructure in large urban areas is not accounted for in our model). To overcome this limitation, no diclofenac concentration is reported for cells accumulating contaminant mass for less than three upstream cells i.e., $l < 3$ in Eq. (5). The two filters described above exemplify how the interpretation of GLOBAL-FATE outcomes depends on the available input datasets, both in terms of quality and resolution. Considering that working resolution and input datasets are user-dependent in GLOBAL-FATE, the criteria to assess model results quality and reliability are case dependent, and the filters suggested here may not be convenient in all circumstances. In any case, users must be aware that the simplified representation of complex processes like water and contaminant routing along natural and engineered systems currently coded in GLOBAL-FATE implies serious limitations on the spatial scale at which the model delivers meaningful results (see Section 4 for a comprehensive discussion on this issue).

Although the aim of this exercise was only to exemplify the application of GLOBAL-FATE in a real case, we assessed the goodness-of-fit of model predictions against observed loads of diclofenac in the river network. We used 405 diclofenac loading (concentration times streamflow) values measured in rivers around the globe compiled by Acuña et al. 2015, and compared this with the modelled value in the corresponding cell after log-transforming the two values (the range of observed and modelled diclofenac loadings shows several orders of magnitude). We used the Nash–Sutcliffe model efficiency coefficient to assess model performance:

$$E=1-\frac{\sum\left(L_{obs}-L_{est}\right)^2}{\sum\left(L_{obs}-\left(\overline{\log L_{obs}}\right)\right)^2} \tag{19}$$

The relationship between observed and simulated diclofenac loads (Fig. 5) shows a Nash–Sutcliffe model efficiency of 0.4, which is reasonable considering that we did not calibrate any parameter of the model. Global models for contaminants always suffer from low to medium performance scores due to the scarce and spatially biased datasets available for model evaluation (Strokal et al., 2019), and frequently they only go beyond E>0.5 after intensive calibration procedures (e.g., Harrison et al, 2019).

3.3 Sensitivity analysis

Model simulations may diverge from observed values due to uncertainty in observations and parametric values, and to deliberate simplifications inherent in all phases of the modelling process. Furthermore, most input datasets come from previous modelling exercises with more assumptions and simplifications that may affect the final result. We carried out a sensitivity analysis in order to investigate the propagation of errors to the output from a selection of inputs (population, pharmaceutical consumption, excretion rate, runoff, decay rate in WWTPs and the river network, lakes volume, Manning coefficient, and the d exponent in Eq. 11). This analysis was performed using a local sensitivity, one-at-a-time procedure, changing one input per simulation around a reference parametric point, defined by the values of the original datasets or the parametric value provided in section 3.1. These inputs were perturbed around this reference point, decreasing and increasing the value from -100% to 100% the original figure in 10% increments. In case of raster datasets, the whole domain was perturbed in a homogeneous way. We assessed the sensitivity of the mean diclofenac load in the river network to those perturbations in the inputs, expressed as percent change form the value in the reference condition.

The results from the sensitivity analysis (Fig. 6) suggest that the output of the model is highly sensitive to two groups of inputs. On one hand, everything related to the generation of contaminant mass by population (population, consumption, and excretion rate) showed the largest overall sensitivity (the sensitivities are the same for these parameters because they are multiplying themselves in the model, Eq. 14). On the other hand, the output was also very sensitive to parameters related to the attenuation of contaminant in the river network: the first order decay rate in the river network, and parameters related to water residence time calculation such as the Manning coefficient and the exponent *d* for water depth. The output showed less sensitivity to the rest of tested inputs. These results suggest that the quality of datasets related to the generation of contaminant from human use must be carefully checked, and that attenuation of the contaminants in rivers and lakes plays an important role on defining their presence in the river network. This last point is very relevant considering that data for first order reaction rates in rivers for many contaminants are scarce or non-existent, and that residence time calculation in the river network still depends on global empirical functions that may have large regional variability. Also, these results suggest that mitigation strategies to reduce the prevalence of pharmaceutical contaminants in the river network should point to increasing the assimilation of the drug by the human body and decisions and campaigns devoted to lower the per capita consumption. This would be much more efficient than increasing WWTP treatment technologies to attenuate the contaminant load before reaching the river network, at least in regions where the prevalence of wastewater treatment is already high.

4. Strengths and limitations of GLOBAL-FATE

GLOBAL-FATE is an open-source, multi-platform, and modular contaminant fate model that links human consumption of pharmaceutical-like compounds with their concentration in the river network. GLOBAL-FATE is also computationally efficient, and can solve the whole global streamflow generation and contaminant routing in less than five minutes in a customary PC. It provides practical guidelines (through readme files and example datasets) to assist non-specialist users in computer programming. At the same time, it has a fully commented code that experienced users can easily customize and further develop to adapt to their needs. The model is also available as a user-friendly QGIS plug-in. Through simple menus, an inexperienced user can conduct simulations and produce basic outputs on the QGIS canvas. This will make global contaminant calculations accessible to a much wider community of scientists and practitioners.

One of the features of GLOBAL-FATE is that it is not fundamentally associated with a spatial resolution or extent. Users can define the working spatial resolution and extent just adapting the resolution of the raster inputs and the region of interest (for instance, a single continent or subcontinent). Although this is an obvious advantage over other large-scale contaminant models, it also harbors the significant risk that users may assume that the model delivers meaningful results at any working spatial scale. We strongly advise against the uncritical use of GLOBAL-FATE, particularly when working at coarse working resolutions or with highly spatially aggregated input data. We do not want to suggest a spatial resolution threshold from which results from GLOBAL-FATE could be considered as reliable, because the criteria to assess model results quality and reliability are case dependent, and guidelines suggested in a given situation may not be convenient in all circumstances. In our example, the combination of the working spatial scale (1/16 of a degree, ~7 km), the complexity of fine-scale interactions between engineered systems and the river network (e.g., the exact location of effluent discharges, extensive sewage networks, poor representation of small streams), and the input data available translates into several model inadequacies that pose limits on the interpretation of the results. We already mentioned that the quality of the runoff map precludes the interpretation of any results for regions where runoff is below 100 mm $year^{-1}$, and that the calculated concentration are unreliable for watersheds smaller than ~150 km^2 (this roughly relates to river reaches of ~20 km) due to inexact effluent discharge locations in small streams and the absence of data on sewage networks in large urban areas, that would route the contaminant load downstream towards larger rivers resulting in higher dilution and lower contaminant concentration. These limitations were easily spotted as they resulted in very unrealistic high diclofenac concentrations scattered throughout the global network, which attracted our immediate attention. However, other assumptions of the modelling approach do not leave such a conspicuous mark in the model output. For instance, consumption data is homogeneous at the country level, while variability inside large countries may be substantial (urban vs. rural regions, for instance). Also, we have averaged information on intensity of treatment also at the country scale, when this may change even at very local scales. This implies that the model results are not necessarily unbiased beyond the threshold mentioned earlier (150 km^2), because all uncertainties and biases propagating from model inputs and assumptions must also have a reflection in the spatial dimension at varying scales. For instance, the comparison between observed and modelled diclofenac concentration along the main axis of the Rhine river (Fig. 7) shows that the model was able to spot a concentration increase at 300 km upstream the river mouth (in the sense that the model predicts an increase that goes beyond 100 ng L^{-1}, the basic threshold we were interested in). However, in the same basin close to the river mouth (∼50 km) the model could not mimic an increase in concentration beyond 100 ng L^{-1}. Our opinion is that GLOBAL-FATE, as implemented in the example, should be used to answer questions which are general in nature. For instance, "contaminant concentration downstream large urban areas in Central Europe frequently exceeds 100 ng L^{-1}", and related statements concerning remediation measures. We advise against the use of GLOBAL-FATE as implemented in the example to support statements concerning particular places at or near the working resolution (for instance "the remediation measures seem insufficient to lower concentrations below 100 ng L^{-1} downstream from Cologne"). We acknowledge that this restriction limits the usability of the current version of GLOBAL-FATE to answer questions that require precise information at a scale of ~20 km of river network. In such cases, models operating at local (i.e., single watershed scale) or regional (i.e., country level) scales may be a better option (e.g., Diamantini et al., 2019). In any case, GLOBAL-FATE can be used to test the effectiveness of large scale management strategies related to pharmaceutical consumption control and wastewater treatment implementation and upgrading, in order to deliver influential assessments of climate change impacts on pharmaceutical consumption and river network ecosystem health (e.g., the Inter Sectoral Impact Model Intercomparison project), and also for informing global policy instruments like the UN Sustainable Development Goals agenda. This is already common practice in other sectors using large scale, coarse resolution models such as impacts of climate change on marine life (Lotze et al., 2019), on lake physics (Woolway and Merchant

2019), on soil moisture (Samaniego et al., 2018), or on economic losses due to river flooding (Dottori et al., 2018), to cite just a few recent examples.

Nonetheless, we discussed the limitations of GLOBAL-FATE as applied in our example (~7 km pixel resolution), but even exercises using models working at much finer resolution in smaller areas (e.g., China at 0.5 km resolution, Grill et al. 2018) found substantial uncertainties related to unaccounted variability regarding input variables and poor representation of small streams. Therefore, we strongly suggest to carefully assess model performance irrespective of the working resolution, and to pay special attention to the spatial scales at which answers are required and its compatibility with the aggregation of input information and the representation of the river network. Finally, GLOBAL-FATE includes very simplified physically based approximations for attenuation in the river network, which a priori are mathematically robust to changes in the spatial resolution, but that assumes homogeneous properties along calculation units (river reaches) such as water velocity and mixing. Although those assumptions do not hold even at very local scales (tens of meters), empirical research on river ecology suggests that this approach is reasonable for rivers reaches up to ~10 km (Marcé et al., 2018). Beyond this, substantial heterogeneity of the river network is overlooked, with potential effects on the contaminant mass balance (Darracq and Destouni, 2007).

GLOBAL-FATE is a steady state model, and although synoptic conditions like low o high flows or climate change scenarios can be modelled, it cannot dynamically simulate extreme events or seasonality. This should be considered when formulating research or management questions for which hydrological seasonal or subseasonal variability is relevant. A keen aspect of GLOBAL-FATE is that researchers are free to use the input information they prefer, it is not limited to particular hydrological products, so synoptic conditions can always be modelled, as far as the steady state assumption is reasonable.

GLOBAL-FATE is the first contaminant model operating at the global scale that fully integrates lakes and reservoirs in the routing of a contaminant along the river network. This is a relevant improvement over other modelling approaches, especially considering the long water residence time of lakes and reservoirs compared to river reaches, which implies a prominent role of lakes and reservoirs on the attenuation of contaminants. However, it should be noted that GLOBAL-FATE models lakes and reservoirs as point-like features, with no spatial heterogeneity. This may fail to capture likely gradients of contaminant concentration in large lakes and reservoirs.

Our analyses showed that GLOBAL-FATE will have a performance in terms of goodness-of-fit similar to other global contaminant fate models. However, as in any other modelling exercise, this will be highly dependent on the quality of input data used and the availability of observed data to adjust parameters that cannot be set at confident values using prior information. In any case, large uncertainties will always be present in global models including simplified lumped representations of very complex processes. We pointed to the main limitations of the model and the most sensitive inputs, but researchers will have to re-assess this in a case-by-case fashion.

As for future developments, we envisage the inclusion of diffuse pollution in the current steady-state framework, which would make GLOBAL-FATE useful for a much wider range of pollutants, such as nutrients or agricultural pesticides, and a more detailed accounting of sewage infrastructure to be able to solve contaminant routing at high resolution in very large urban areas. In any case, GLOBAL-FATE will be a valuable tool for the scientific community and the policymaking arena, and could be used to test the effectiveness of large scale management strategies related to pharmaceutical consumption control and wastewater treatment implementation and upgrading.

Code availability

The GLOBAL-FATE code (including compiling instructions, examples, and the QGIS plug-in) is available in the following URL: https://github.com/ICRA/GLOBALFATE. Prebuilt executables for Windows are available under request.

Author contribution

RM, VA, and SS conceived the model. RM produced a preliminary version of the model code, later modified and expanded by FB, CF, and VA. CF and FB built the diclofenac example and performed model runs. CF and RM prepared the manuscript with contributions from all co-authors.

Acknowledgements

This work has been supported by the European Communities 7th Framework Programme funding under Grant Agreement No.603629-ENV-2013-6.2.1-GLOBAQUA. Authors also acknowledge the support for scientific equipment given by the European Regional Development Fund (FEDER) under the Catalan FEDER Operative Program 2007-2013 and by MINECO according to DA3ª of the Catalan Statute of Autonomy and to PGE-2010.

References

Acuña, V., Ginebreda, A., Mor, J. R., Petrovic, M., Sabater, S., Sumpter, J., Barceló, D.: Balancing the health benefits and environmental risks of pharmaceuticals: Diclofenac as an example, Environ. Int., 85, 327-333, https://doi.org/10.1016/j.envint.2015.09.023, 2015.

Anderson, P. D., D'Aco, V. J., Shanahan, P., Chapra, S. C., Buzby, M. E., Cunningham, V. L., Rader, J. C.:Screening analysis of human pharmaceutical compounds in US surface waters, Environ. Sci. Technol., 38, 838-849, https://doi.org/10.1021/es034430b, 2004.

Andreadis K. M., Schumann G. J. P., Pavelsky T.: A simple global river bankfull width and depth database, Water Resour. Res., 49, 7164–7168, https://doi.org/10.1002/wrcr.20440, 2013.

Archundia, D., Boithias, L., Duwig, C., Morel, M. C., Aviles, G. F., Martins, J. M. F.: Environmental fate and ecotoxicological risk of the antibiotic sulfamethoxazole across the Katari catchment (Bolivian Altiplano): Application of the GREAT-ER model, Sci. Total Environ., 622, 1046-1055, https://doi.org/10.1016/j.scitotenv.2017.12.026, 2018.

Arlos, M. J., Bragg, L. M., Servos, M. R., Parker, W. J.: Simulation of the fate of selected pharmaceuticals and personal care products in a highly impacted reach of a Canadian watershed, Sci. Total Environ., 485, 193-204, https://doi.org/10.1016/j.scitotenv.2014.03.092, 2014.

Besseling, E., Quik, J. T., Sun, M., Koelmans, A. A.: Fate of nano-and microplastic in freshwater systems: A modeling study, Environ. Pollut., 220, 540-548, https://doi.org/10.1016/j.envpol.2016.10.001, 2017.

Boxall, A. B. A., Keller, V. D. J., Straub, J. O., Monteiro, S. C., Fussell, R., Williams, R. J.: Exploiting monitoring data in environmental exposure modelling and risk assessment of pharmaceuticals, Environ. Int., 73, 176-185, https://doi.org/10.1016/j.envint.2014.07.018, 2014.

Brown, L. C., Barnwell, T. O.: The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: documentation and user manual, US Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory, 1987.

Darracq, A., Destouni, G.: Physical versus biogeochemical interpretations of nitrogen and phosphorus attenuation in streams and its dependence on stream characteristics, Global Biogeochem. Cycles, 21, GB3003, https://doi.org/10.1029/2006GB002901, 2007.

Di Toro, D. M., Fitzpatrick, J. J., Thomann, R. V.: Documentation for water quality analysis simulation program (WASP) and model verification program (MVP), 1983.

Diamantini, E., Mallucci, S., Bellin, A.: A parsimonious transport model of emerging contaminants at the river network scale. Hydrol. Earth Syst. Sci., 23(1): 573-593, doi:10.5194/hess-23-573-2019, 2019.

Dottori, F., Szewczyk, W., Ciscar, J.-C., Zhao, F., Alfieri, L., Hirabayashi, Y., Bianchi, A., Mongelli, I., Frieler, K., Betts, R. A.., Feyen, L.: Increased human and economic losses from river flooding with anthropogenic warming, Nature Climate Change, 8, 781-786, https://doi.org/10.1038/s41558-018-0257-z, 2018.

Doxsey-Whitfield E., MacManus K., Adamo S. B., Pistolesi, L., Squires, J., Borkovska, O., Baptista, S. R.: Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4, Pap. Appl. Geogr., 1, 226–234, https://doi.org/10.1080/23754931.2015.1014272, 2015.

Dumont, E., Johnson, A. C., Keller, V. D., Williams, R. J.: Nano silver and nano zinc-oxide in surface waters–Exposure estimation for Europe at high spatial and temporal resolution, Environ. Pollut., 196, 341-349. https://doi.org/10.1016/j.envpol.2014.10.022, 2015.

Fekete, B. M., Vörösmarty, C. J., Grabs, W.: High-resolution fields of global runoff combining observed river discharge and simulated water balances, Glob. Biochem. Cycles, 16, 1042, https://doi.org/10.1029/1999GB001254, 2002.

Feijtel, T., Boeije, G., Matthies, M., Young, A., Morris, G., Gandolfi, C., Hanse, C., Fox, K., Holt, M., Koch, V., Schroder, R., Cassani, G., Schowanek, D., Rosenblom, J., Niessen, H.: Development of a geography-referenced regional exposure assessment tool for European rivers-GREAT-ER contribution to GREAT-ER# 1, Chemosphere, 34, 2351-2373, https://doi.org/10.1016/S0045-6535(97)00048-9, 1997.

Ferrer, D. L., DeLeo, P. C.: Development of an in-stream environmental exposure model for assessing down-the-drain chemicals in Southern Ontario, Water Qual. Res. J., 52, 258-269, https://doi.org/10.2166/wqrj.2017.019, 2017.

Gentzsch, W.: Sun grid engine: Towards creating a compute power grid, in: Proceedings First IEEE/ACM International Symposium on Cluster Computing and the Grid, Brisbane, Queensland, Australia, 35-36, 2001.

Goldman, L. R., Koduru, S.: Chemicals in the environment and developmental toxicity to children: a public health and policy perspective, Environ. Health Perspect., 108, 443-448, https://doi.org/10.1289/ehp.00108s3443, 2000.

Gouin, T., Armitage, J. M., Cousins, I. T., Muir, D. C., Ng, C. A., Reid, L., Tao, S.: Influence of global climate change on chemical fate and bioaccumulation: The role of multimedia models, Environ. Toxicol. Chem., 32, 20-31, https://doi.org/10.1002/etc.2044, 2013.

Grill, G., Khan, U., Lehner, B., Nicell, J., Ariwi, J.: Risk assessment of down-the-drain chemicals at large spatial scales: Model development and application to contaminants originating from urban areas in the Saint Lawrence River Basin, Sci. Total Environ., 541, 825-838, https://doi.org/10.1016/j.scitotenv.2015.09.100, 2016.

Grill, G., Li, J., Khan, U., Zhong, Y., Lehner, B., Nicell, J., Ariwi, J.: Estimating the eco-toxicological risk of estrogens in China's rivers using a high-resolution contaminant fate model, Water Research, 145, 707-720, https://doi.org/10.1016/j.watres.2018.08.053, 2018.

Heberer, T., Feldmann, D.: Contribution of effluents from hospitals and private households to the total loads of diclofenac and carbamazepine in municipal sewage effluents - Modeling versus measurements, J. Hazard. Mater., 122, 211–218, https://doi.org/10.1016/j.jhazmat.2005.03.007, 2005.

Harrison, J.A., Beusen, A. H. W., Fink, G., Tang, T., Strokal, M., Bouwman, A. F., Metson, G. S., Vilmin, L.: Modeling phosphorus in rivers at the global scale: recent successes, remaining challenges, and near-term opportunities, Curr. Opin. Environ. Sustain., 36, 68-77, https://doi.org/10.1016/j.cosust.2018.10.010, 2019.

Hernández, F., Ibáñez, M., Botero-Coy, A. M., Bade, R., Bustos-López, M. C., Rincón, J., ... Bijlsma, L.: LC-QTOF MS screening of more than 1,000 licit and illicit drugs and their metabolites in wastewater and surface waters from the area of Bogotá, Colombia, Anal. Bioanal. Chem., 407, 6405-6416, https://doi.org/10.1007/s00216-015-8796-x, 2015.

Hsu, A., Zomer, A.: Environmental Performance Index, in: Wiley StatsRef: Statistics Reference Online, edited by: Balakrishnan, N., Colton, T., Everitt, B., Piegorsch, W., Ruggeri, F., Teugels, J. L., John Wiley & Sons, New York, USA, 1-5, https://doi.org/10.1002/9781118445112.stat03789.pub2, 2016.

Johnson, A. C., Keller, V., Williams, R. J., Young, A.: A practical demonstration in modelling diclofenac and propranolol river water concentrations using a GIS hydrology model in a rural UK catchment, Environ. Pollut., 146, 155-165, https://doi.org/10.1016/j.envpol.2006.05.037, 2007.

Johnson, A. C., Dumont, E., Williams, R. J., Oldenkamp, R., Cisowska, I., Sumpter, J. P.: Do concentrations of ethinylestradiol, estradiol, and diclofenac in European rivers exceed proposed EU environmental quality standards?, Environ. Sci. Technol., 47, 12297-12304, https://doi.org/10.1021/es4030035, 2013.

Kapo, K. E., DeLeo, P. C., Vamshi, R., Holmes, C. M., Ferrer, D., Dyer, S. D., Wang, X., White-Hull, C.: iSTREEM: An approach for broad-scale in-stream exposure assessment of “down-the-drain” chemicals, Integr. Environ. Assess. Manag., 12, 782-792, https://doi.org/10.1002/ieam.1793, 2016.

Keller, V., Fox, K., Rees, H.G. and Young, A.R.: Estimating population served by sewage treatment works from readily available GIS data. The Science of The Total Environment, 360(1-3), 319-327, 774 https://doi.org/10.1016/j.scitotenv.2005.08.043, 2006

Keller, V. D. J., Lloyd, P., Terry, J. A., Williams, R. J.: Impact of climate change and population growth on a risk assessment for endocrine disruption in fish due to steroid estrogens in England and Wales, Environ. Pollut., 197, 262-268, https://doi.org/10.1016/j.envpol.2014.11.017, 2015.

K'oreje, K. O., Vergeynst, L., Ombaka, D., De Wispelaere, P., Okoth, M., Van Langenhove, H., Demeestere, K.: Occurrence patterns of pharmaceutical residues in wastewater, surface water and groundwater of Nairobi and Kisumu city, Kenya, Chemosphere, 149, 238-244, https://doi.org/10.1016/j.chemosphere.2016.01.095, 2016.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, Front. Ecol. Environ., 9, 494-502, https://doi.org/10.1890/100125, 2011.

Lehner, B., Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, J. Hydrol., 296, 1–22, https://doi.org/10.1016/j.jhydrol.2004.03.028, 2004.

Leopold, L. B., Maddock, T. J.: The Hydraulic Geometry of Stream Channels and Some Physiographic Implications, Geol. Surv. Prof. Pap., 252, 1-57, https://doi.org/10.3133/pp252, 1953.

Lewis, Jr. W.: Global primary production of lakes: 19th Baldi Memorial Lecture, Inl. Waters, 1, 1–28, https://doi.org/10.5268/IW-1.1.384, 2011.

Li, Z., Sobek, A., Radke, M.: Fate of pharmaceuticals and their transformation products in four small European rivers receiving treated wastewater. Environ. Sci. Technol., 50, 5614-5621, https://doi.org/10.1021/acs.est.5b06327, 2016.

Liang, J., Yang, Q., Sun, T., Martin, J. D., Sun, H., Li, L.: MIKE 11 model-based water quality model as a tool for the evaluation of water quality management plans. J. Water Supply Res. T., 64, 708-718, https://doi.org/10.2166/aqua.2015.048, 2015

Lindim, C., Van Gils, J., Cousins, I. T.: A large-scale model for simulating the fate and transport of organic contaminants in river basins, Chemosphere, 144, 803-810, https://doi.org/10.1016/j.chemosphere.2015.09.051, 2016.

Lotze, H. K., Tittensor, D. P., Bryndum-Buchholz, A., Eddy, T. D., Cheung, W. W. L., Galbraith, E. D., Barange, M., Barrier, N., Bianchi, D., Blanchard, J. L., Bopp, L., Büchner, M., Bulman, C. M.., Carozza, D. A.., Christensen, V., Coll, M., Dunne, J. P.., Fulton, E. A.., Jennings, S., Jones, M. C.., Mackinson, S., Maury, O., Niiranen, S., Oliveros-Ramos, R., Roy, T., Fernandes, J. A.., Schewe, J., Shin, Y.-J., Silva, T. A. M.., Steenbeek, J., Stock, C. A.. Verley, P., Volkholz, J., Walker, N. D., Worm, B.: Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change, Proceedings of the National Academy of Sciences, 116 (26), 12907-12912, DOI: 10.1073/pnas.1900194116, 2019.

MacLeod, M., von Waldow, H., Tay, P., Armitage, J. M., Wöhrnschimmel, H., Riley, W. J., McKone, T. E., Hungerbuhler, K.: BETR global–A geographically-explicit global-scale multimedia contaminant fate model, Environ. Pollut., 159, 1442-1445, https://doi.org/10.1016/j.envpol.2011.01.038, 2011.

Marcé, R., von Schiller, D., Aguilera, R., Martí, E., Bernal, S.: Contribution of hydrologic opportunity and biogeochemical reactivity to the variability of nutrient retention in river networks, Global Biogeochemical Cycles, 32, 376–388, https://doi.org/10.1002/2017GB005677, 2018.

Nassef, M., Matsumoto, S., Seki, M., Khalil, F., Kang, I. J., Shimasaki, Y., Oshime, Y., Honjo, T.: Acute effects of triclosan, diclofenac and carbamazepine on feeding performance of Japanese medaka fish (Oryzias latipes), Chemosphere, 80, 1095-1100, https://doi.org/10.1016/j.chemosphere.2010.04.073, 2010.

O'Callaghan, J. F., Mark, D. M.: The extraction of drainage networks from digital elevation data, Comput. Vision Graphics Image Process., 28, 328-344, https://doi.org/10.1016/S0734-189X(84)80011-0, 1984.

Oldenkamp, R., Hoeks, S., Čengić, M., Barbarossa, V., Burns, E., Boxall, A. B., Ragas, A.: A high-resolution spatial model to predict exposure to pharmaceuticals in European surface waters–ePiE, Environ. Sci. Technol., https://doi.org/10.1021/acs.est.8b03862, 2018.

Pistocchi, A., Marinov, D., Pontes, S., Gawlik, B. M.: Continental scale inverse modeling of common organic water contaminants in European rivers, Environ. Pollut., 162, 159-167, https://doi.org/10.1016/j.envpol.2011.10.031, 2012.

Pistocchi, A.: GIS Based Chemical Fate Modeling: Principles and Applications, Wiley, ISBN: 978-1-118-05997-5, 2014.

Postigo, C., de Alda, M. J. L., Barceló, D.: Drugs of abuse and their metabolites in the Ebro River basin: occurrence in sewage and surface water, sewage treatment plants removal efficiency, and collective drug usage estimation, Environ. Int., 36, 75-84, https://doi.org/10.1016/j.envint.2009.10.004, 2010.

QGIS Development Team: QGIS Geographic Information System. Open Source Geospatial Foundation Project, 2018.

Rice, J., Westerhoff, P.: High levels of endocrine pollutants in US streams during low flow due to insufficient wastewater dilution, Nat. Geosci., 10, 587-591, https://doi.org/10.1038/ngeo2984, 2017.

Richardson, B. J., Lam, P. K., Martin, M.: Emerging chemicals of concern: pharmaceuticals and personal care products (PPCPs) in Asia, with particular reference to Southern China, Mar. Pollut. Bull., 50, 913-920, https://doi.org/10.1016/j.marpolbul.2005.06.034, 2005.

Rudd, R. L.: Chemicals in the environment, Calif. Med., 113(5), 27-32, 1970.

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E. F., Marx, A.: Anthropogenic warming exacerbates European soil moisture droughts, Nature Climate Change, 8, 421-426, https://doi.org/10.1038/s41558-018-0138-5, 2018.

Santhi, C., Srinivasan, R., Arnold, J. G., Williams, J.R.: A modeling approach to evaluate the impacts of water quality management plans implemented in a watershed in Texas, Environ. Modell. Softw., 21, 1141-1157, https://doi.org/10.1016/j.envsoft.2005.05.013, 2005.

Schulze, K., Hunger, M., Döll, P.: Simulating river flow velocity on global scale, Adv. Geosci., 5, 133–136, https://doi.org/10.5194/adgeo-5-133-2005, 2005.

Stewart, M., Olsen, G., Hickey, C. W., Ferreira, B., Jelić, A., Petrović, M., Barcelo, D.: A survey of emerging contaminants in the estuarine receiving environment around Auckland, New Zealand, Sci. Total Environ., 468, 202-210, https://doi.org/10.1016/j.scitotenv.2013.08.039, 2014.

Strokal, M., Emiel Spanier, J., Kroeze, C., Koelmans, A. A., Flörke, M., Franssen, W., Hofstra, N., Langan, S., Tang, T., van Vliet, M. T. H., Wada, Y., Wang, M., van Wijnen, J., Williams, R.: Global multi-pollutant modelling of water quality: scientific challenges and future directions. Curr. Opin. Environ. Sustain., 36, 116-125, https://doi.org/10.1016/j.cosust.2018.11.004, 2019.

Ternes, T. A.: Occurrence of drugs in German sewage treatment plants and rivers, Water Res., 32, 3245–3260, https://doi.org/10.1016/S0043-1354(98)00099-2, 1998.

Todd, P. A., Sorkin, E. M.: Diclofenac sodium, Drugs, 35(3), 244-285, https://doi.org/10.2165/00003495-198835030-00004, 1998.

UN General Assembly: Transforming our World: The 2030 Agenda for Sustainable Development, Resolution A/RES/70/1, available at: https://sustainabledevelopment.un.org/post2015/transformingourworld, 2015.

Van Wijngaarden, M.: A two dimensional model for suspended sediment transport in the southern branch of the Rhine–Meuse estuary, The Netherlands, Earth Surf. Process. Landf., 24(13), 1173-1188, https://doi.org/10.1002/(SICI)1096-9837(199912)24:13<1173::AID-ESP25>3.0.CO;2-N, 1999.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R., Davies, P. M.: Global threats to human water security and river biodiversity, Nature, 467, 555-561, https://doi.org/10.1038/nature09440, 2010.

Woolway, R. I., Merchant, C. J.: Worldwide alteration of lake mixing regimes in response to climate change, Nature Geoscience, 12, 271-276, https://doi.org/10.1038/s41561-019-0322-x, 2019.

Wu, H., Kimball, J. S., Li, H, Huang, M., Ruby Leung, L., Adler, R.F.: A new global river network database for macroscale hydrologic modeling, Water Resour. Res., 48, W09701, https://doi.org/10.1029/2012WR012313, 2012.

Zhang, Y., Geißen, S.-U., Gal, C.: Carbamazepine and diclofenac: Removal in wastewater treatment plants and occurrence in water bodies, Chemosphere, 73, 1151–1161, https://doi.org/10.1016/j.chemosphere.2008.07.086, 2008.

Zhang, L., Cao, Y., Hao, X., Zhang, Y., Liu, J.: Application of the GREAT-ER model for environmental risk assessment of nonylphenol and nonylphenol ethoxylates in China, Environ. Sci. Pollut. R., 22, 18531-18540, https://doi.org/10.1007/s11356-015-5352-3, 2015.

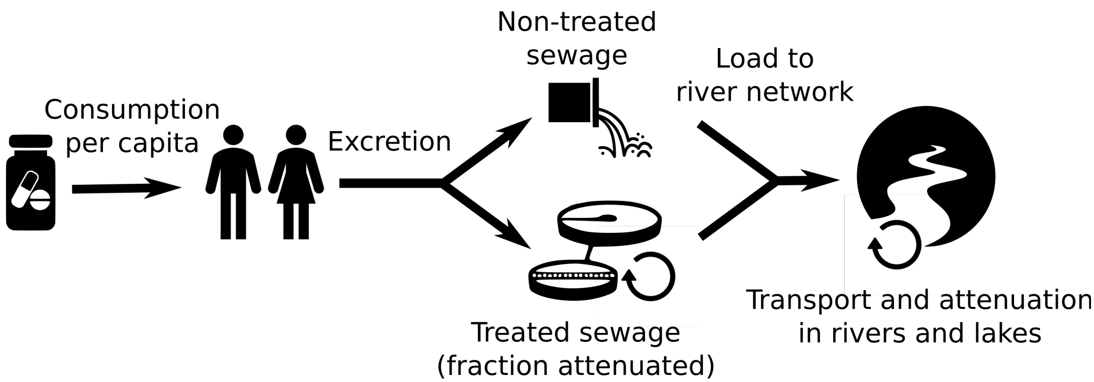

Figure 1. Conceptual diagram of the processes modelled by GLOBAL-FATE

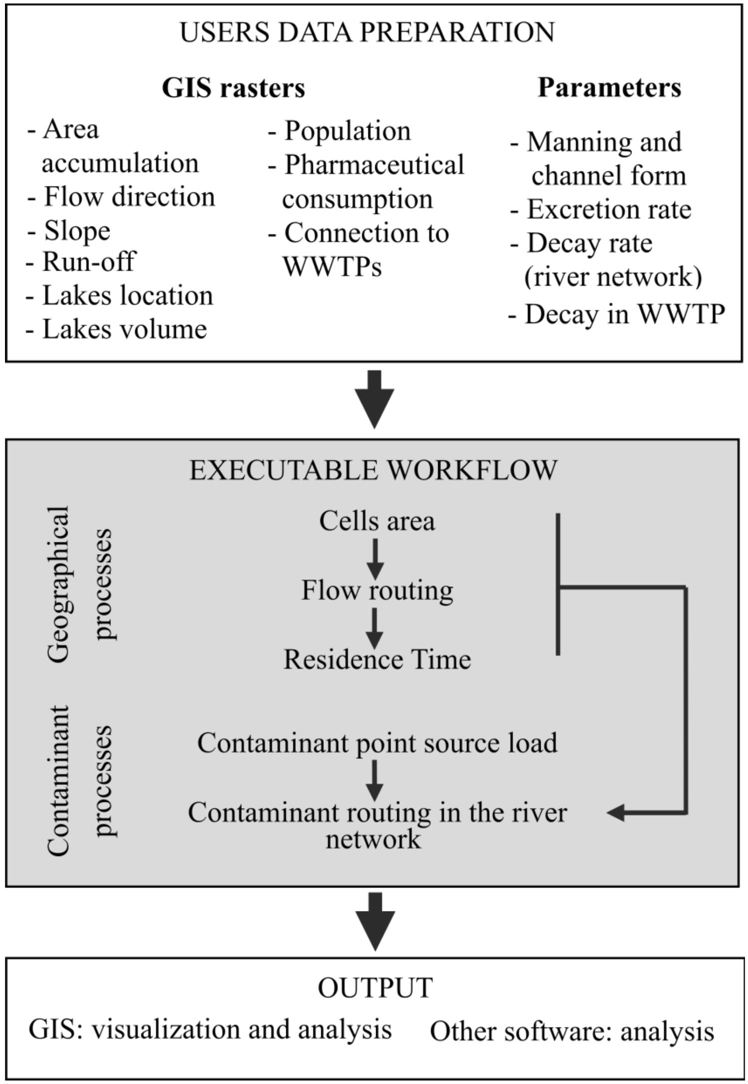

Figure 2. Work flow of GLOBAL-FATE

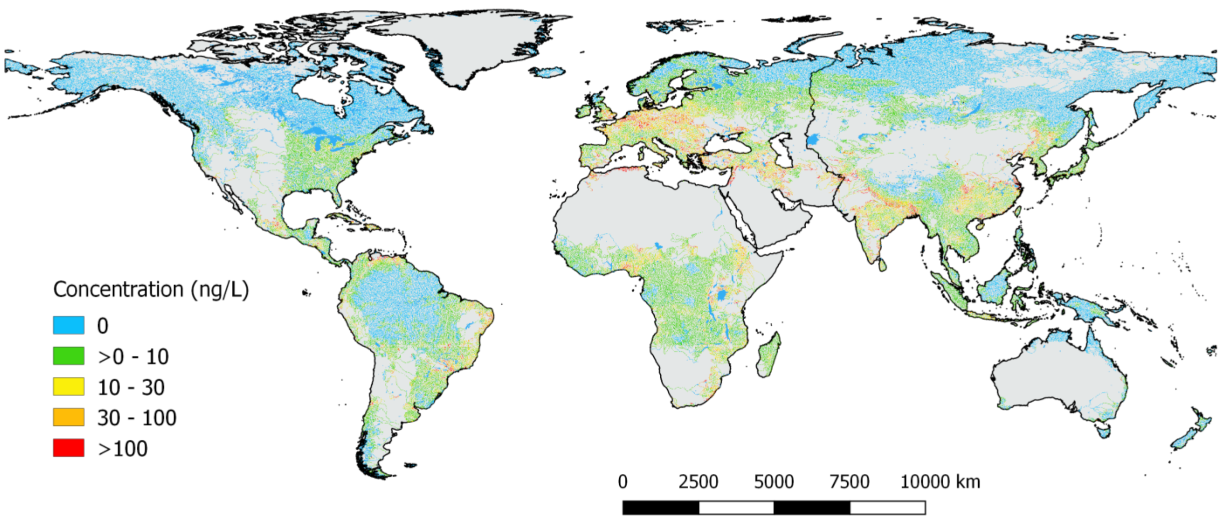

Figure 3. Simulated mean annual diclofenac concentration worldwide.

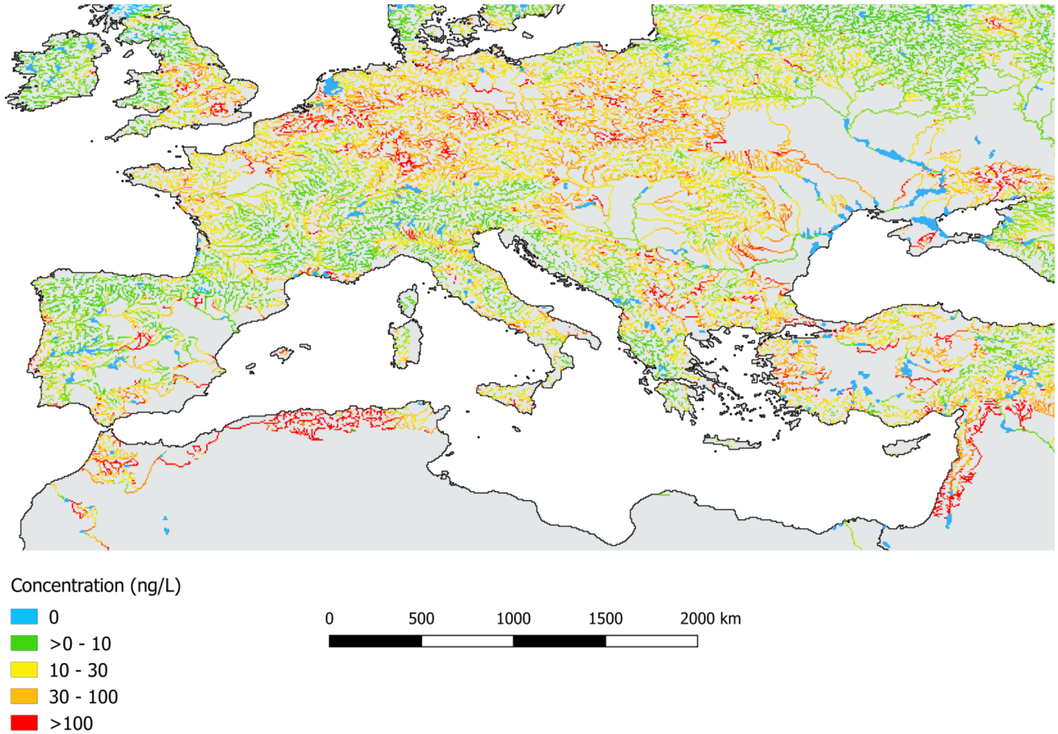

Figure 4. Simulated mean annual diclofenac concentration in Central and Southern Europe and the southern Mediterranean basin.

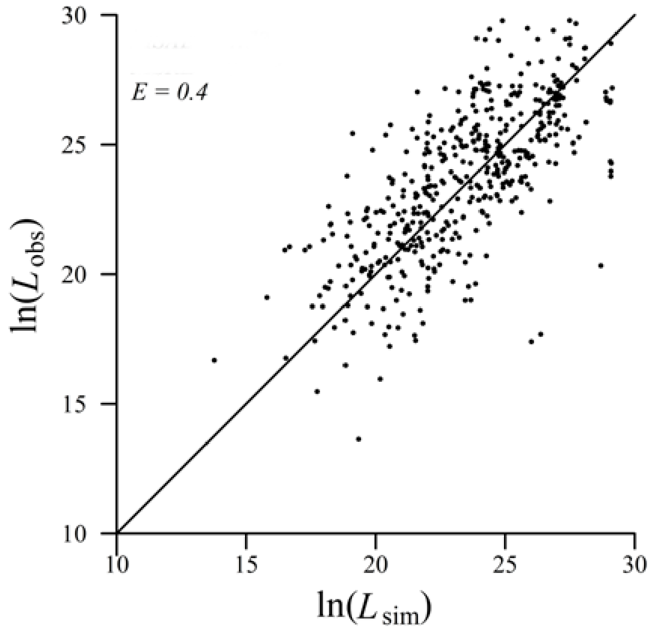

Figure 5. Observed versus simulated load log-values ($\ln(ng/L)$), $N=405$ points. Nash–Sutcliffe model efficiency coefficient (E) is also reported.

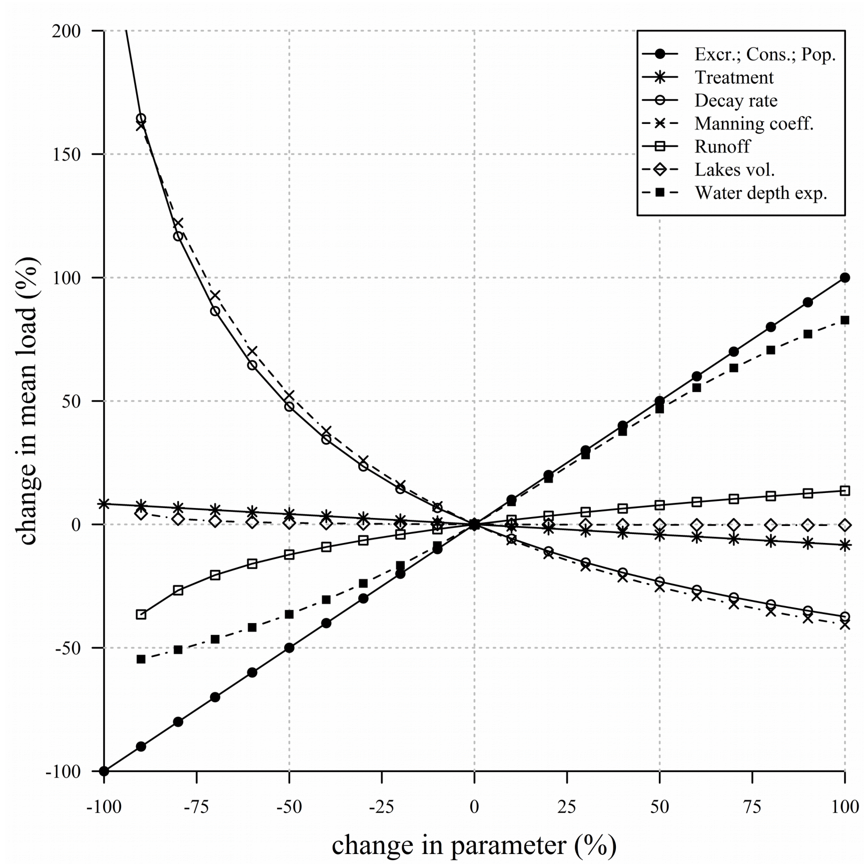

Figure 6. Spider plot of percent changes in the mean load in the river network due to changes in a collection of inputs to GLOBAL-FATE.

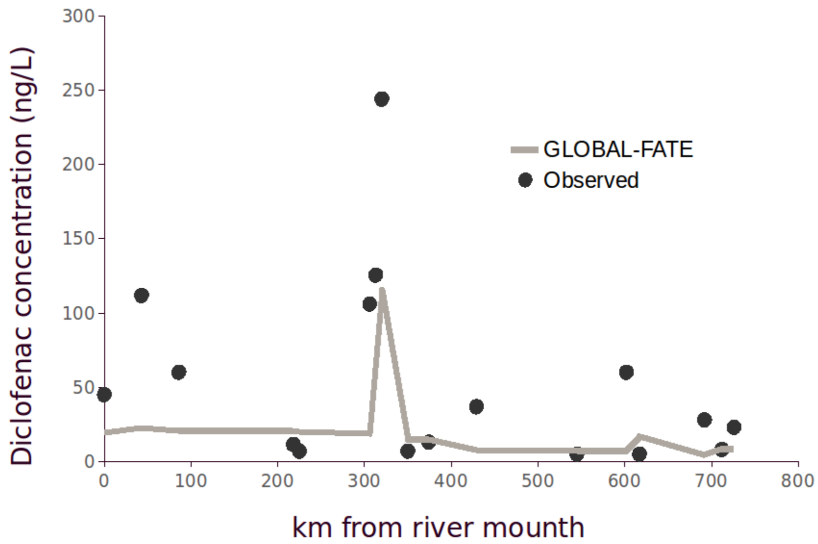

Figure 7. GLOBAL-FATE diclofenac simulation along the Rhine river (Europe), and diclofenac observations from our compiled database.

	PhATE	GREAT-ER	LF2000-WQX	HydroROUT	iSTREEM	ePiE	GLOBAL-FATE
Pollutant modelled	Pharmaceuticals	Diclofenac and propranolol; nonylphenol and nonylphenol ethoxylates; Sulfamethoxazole (antibiotic)	Pharmaceuticals Steroid estrogens	Pharmaceuticals	HHCB and DEET; triclosan and carbamazepine	Pharmaceuticals	Pharmaceuticals
Spatial extent	US	Large river basins	England and Walles	Quebec and Ontario	US	EU	World
Spatial resolution	Discrete segments (~16 km)	Discrete segments	River reach visible at a scale of 1:50,000	Raster of 500 m pixel resolution	River network segment	30 arc seconds (~1 km)	Any raster resolution
Model type	Deterministic	Mixes deterministic with stochastic processes	Mixed deterministic and stochastic model	Deterministic	Deterministic	Deterministic	Deterministic
Sources of pollutants	Point sources, from Publicly owned treatment works (POTW)	Point sources from WWTP	Point sources	Point sources	Point sources from WWTP. Different treatments in WWTP	Point sources from WWTP	Point sources
Model implementation and availability	Microsoft Visual C++ and uses Microsoft Access databases	Implemented as part of a GIS. It is open source software under the GNU Public License	Not available	Not available	Public web application	Written in R	Public code written in C. QGIS plug-in
Transferability to global scale	Limited geographic scope	Restricted to river network dataset and WWTP information availability	Restricted to river network dataset and WWTP information availability	It needs other models (WaterGAP) to estimate runoff	Restricted to river network dataset and WWTP information availability	Restricted to river network dataset and WWTP information availability	Native
References	Anderson et al., 2004	Johnson et al. 2007; Zhang et al., 2015; Archundia et al., 2018	Boxall et al., 2014; Keller et al., 2015	Grill et al., 2016	Kapo et al., 2015; Ferrer and DeLeo, 2017	Oldenkamp et al., 2018	This study

Table 1. Features of a collection of contaminants fate models compared to those of GLOBAL-FATE.

	INPUTS	OUTPUTS
Morphology and Hydrology	• Flow direction • Area accumulation (hierarchic structure) • Runoff (mm year^{-1}) • Lakes location and shape • Lakes volume (m^3) • Slope (m m^{-1}) ○ Manning coefficient (s m$^{-1/3}$) ○ Parameters for channel form (4 of them)	• Cells area (m^2) and width (m) • Streamflow (m^3 year^{-1}) • Residence time in rivers and lakes (hours) • Lake outlet discharge (m^3 year^{-1})
Contaminant	• Population (people per cell) • Contaminant consumption per capita (country level, g person^{-1} year^{-1}) • Population connected to WWTPs (country level, fraction) ○ Decay constant in the river network (hour^{-1}) ○ Human excretion rate (fraction) ○ WWTP attenuation efficiency (fraction)	• Contaminant concentration (g m^{-3})

Table 2. Input and output datasets and parameters for both geographical (morphology and hydrology) and contaminant model processes. Filled bullets represent raster datasets, non-filled bullets stand for parameters.

Process	Description	Inputs	Outputs	C function
Area	Calculates cells area	❖ No direct user inputs, but projection must be WGS84	• Area for each cell in latitude direction* (m^2) • Horizontal cells width for each cell in latitude direction* (m)	Area_m2_fun.c
Flow routing	Calculates streamflow	❖ Raster of flow direction ❖ Raster of area accumulation ❖ Raster of runoff (m $year^{-1}$) ➢ Area (m^2)	• Raster of streamflow* (m^3 $year^{-1}$)	Flow_accumulation_m2.c
Residence Time calculator	Calculates residence time for every cell	❖ Raster of slope (m m^{-1}) ❖ Manning coefficient (s $m^{-1/3}$) ❖ Parameters of channel form (4) ➢ Raster of streamflow (m^3 $year^{-1}$) ➢ Cell height and cell width (m) ✓ Raster of flow direction ✓ Raster of area accumulation	• Raster of flow velocity* (m s^{-1}) • Raster of residence time in rivers (hours)	RT_rivers_calculator.c
Lakes RT incorporation	Incorporates lakes into the RT raster	❖ Raster of lakes location and shape ❖ Raster of lakes volume (m^3)	• Raster of residence time in rivers and lakes* (hours) • Vector of outlet discharge per lake* (m^3 $year^{-1}$)	RT_lakes_incorporation.c
Contaminant load	Calculates consumption by population and attenuation in WTTPs	❖ Population raster (people per pixel) ❖ Raster of pharmaceutical consumption per capita (g $person^{-1}$ $year^{-1}$) ❖ Raster of fraction of sewage treated ❖ Rate of contaminant excretion ❖ Rate of contaminant removal in WWTP	• Raster of contaminant load from human consumption to the river network (g $year^{-1}$)	Initial_contaminant_load.c
Contaminant routing	Calculates contaminant routing in the river network	❖ Exponential decay rate ($hours^{-1}$) ➢ Raster of residence time (hours) ➢ Raster of streamflow (m^3 $year^{-1}$) ✓ Raster of flow direction ✓ Raster of area accumulation	• Raster of contaminant concentration* (g m^{-3}) or load* (g $year^{-1}$) in the river network	Contaminant_accumulation.c

Input flags legend:

❖ Dataset used for the first time

➢ Input coming from previous functions output

✓ Data set used (at least) for the second time

Table 3. Main calculation steps in GLOBAL-FATE, with indication of inputs and outputs used by each process, and the C functions responsible. Outputs with an asterisks can be saved during model execution and accessed afterwards.