Author's response to reviews of the manuscript "MFIT 1.0.0: Multiflow inversion of tracer breakthrough curves in fractured and karst aquifers" by Jacques Bodin

jacques.bodin@univ-poitiers.fr

I am grateful to the two referees for their positive and constructive feedback, which has greatly benefited the work presented in this manuscript. I address their specific comments in detail here.

Font legend:

1. *Reviewers' comments are shown in a black italic font.*
2. My responses are shown in a blue normal font. Where not specified, line, table and/or figure numbers refer to the original manuscript.
3. The text that has been modified or added to the revised manuscript is shown in an orange normal font.

Anonymous Referee #1

My major criticism is focused on the code verification, in particular in the BTCs selected for the comparative analysis of simulation results. The five synthetic BTCs generated fail in both the relatively simple curve morphologies and the test duration. Since the four proposed models try to better fit multi-peaks and long-tailed curve shapes, the multimodal curves obtained from real field experiences show more marked/pronounced peaks (very often reaching relatively quite similar tracer concentration, as twin peaks) and the long-tailed ones (even with higher concentrations slowly decreasing along the lower slope ending curve segments) use to be recordered during much more prolonged tests (>100 hours). So, I would recommend incorporates and/or replacing new synthetic BTCs representing more adapted-to-reality morphologies. This will deeply test the code efficiency under more realistic and non-ideal (Fickian) transport dynamics.

The model parameter values in tests 1–5 have been modified following your recommendations. Table 2 and Fig. 2 have been updated.

Table 2. Input parameters for the five verification tests.

Test	Parameters	Values
1 (single flow channel, ADE, instantaneous injection)	Flow rate Q	10 $m^3 h^{-1}$
	Injected mass m	20 g
	Mean transit time T_0	200 h
	Peclet number Pe	2
2 (single flow channel, ADE, exponentially decaying injection)	Mean transit time T_0	70 h
	Peclet number Pe	10
	Initial (maximum) injection concentration C_0	8.0×10^{-3} mg l^{-1}
	Gamma coefficient γ	0.9
3 (single flow channel, SFDM)	Q, m, T_0, Pe	same as Test 1
	Diffusion parameter β	0.04 $h^{-1/2}$
4 (single flow channel, 2RNE)	Q, m, T_0, Pe	same as Test 1
	Length of the flow channel L	1000 m
	Fraction of mobile water ψ	0.7
	Omega coefficient ω	0.1 m^{-1}
5 (two channels, MDM-ADE)	Total system flow rate Q	10 $m^3 h^{-1}$
	Mass flowing through the first channel m_1	12 g
	Mass flowing through the second channel m_2	8 g
	Mean transit time in the first channel T_{01}	170 h
	Mean transit time in the second channel T_{02}	300 h
	Peclet number in the first channel Pe_1	15
	Peclet number in the second channel Pe_2	80

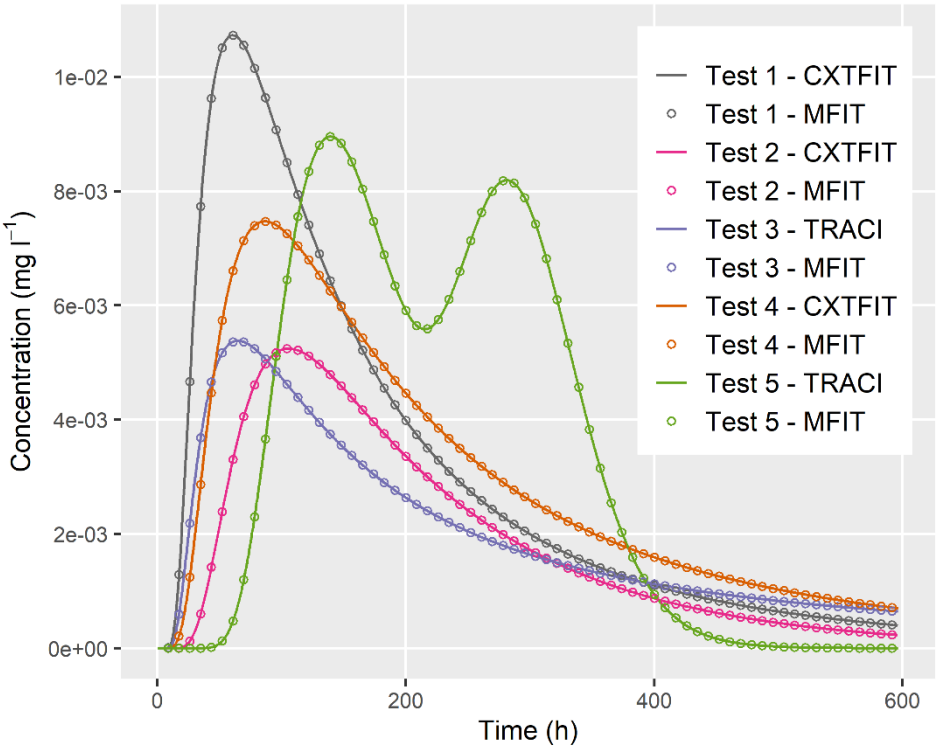

Figure 2. Comparison among MFIT, CXTFIT, and TRACI simulations for test 1 (single flow channel, ADE, instantaneous injection), test 2 (single flow channel, ADE, exponentially decaying injection), test 3 (single flow channel, SFDM), test 4 (single flow channel, 2RNE), and test 5 (two channels, MDM-ADE)

Regarding the modeled BTCs from the HES experimental site, they also display short tracer test duration and local transport dynamics. Some questions arise me, what about longer –multi-kilometers- karst connections and their expected very often long-tailed BTCs? and, what about the degree of flow diversion in anastomosed/ forked karst conduit systems and their associated multi-peak BTCs? I agree with the proposed pathway decomposition in multi-single channel scheme but, how the flow diversion in one or several of them and where (close to the injection point or to the end of the master conduit) may condition the obtained BTC shape?

There is no specific scale attached to the multiflow modeling approach, as depicted in Fig. 1. The lengths of the flow channels can be assumed to be a few meters or several kilometers. The proposed model is applicable to any tracer test, regardless of the distance between the injection site and the monitoring point. Likewise, short-, medium-, or long-duration BTCs can be simulated with the model depending on the values of the parameters that influence the spreading of transit/residence times in the individual flow channels (*Pe*, γ, β, ψ, and ω) and/or by considering multiple flow channels that have a large span of mean transit time (T_0) values. The only unsuitable cases are those for which extensive breakthrough tailing would be due to i) unsteady flow, and/or ii) complex injection signals (e.g., multiple steps), and/or iii) reactive transport processes, which are not considered in the present model version. A noticeable difference between modeling short- or long-distance tracer tests could be the physical interpretation that pertains to the channels of the multiflow model. For short distances, the channels might be considered as model-abstractions of real individual fractures or karst-conduits. For longer distances, this abstraction likely would be physically unrealistic. As discussed in the manuscript (L105–108), in the general case, the channels are not assumed to represent individual fractures or karst conduits but are lumped submodels of the main flow routes employed by the tracer through the fractures/karst conduit network. The multiflow model that will be fitted against a multi-peak BTC generated by a complex flow structure geometry will always be a simplification of the real flow network pattern, which is the essence of modeling. Moreover, the information content of a tracer BTC only does not allow for further structural interpretation. For

instance, the existence of two concentration peaks in a tracer BTC indicates at least one diverging-converging flow structure between the injection site and the monitoring point. However, many more structures may exist in reality; their effects can be masked by mixing at the converging nodes and/or by similar transit times in the different pathways. Similarly, it is impossible to determine the relative location of the diverging and converging nodes between the injection site and the monitoring point. Promising approaches to address this issue are i) the joint inversion of multiple tracer tests with different injection and monitoring locations, which is also referred to as tracer tomography (see, e.g., Borghi et al., Can one identify karst conduit networks geometry and properties from hydraulic and tracer test data? Adv. Water Resour. 2016) and ii) the joint interpretation of tracer test and geophysical data (see, e.g., Shakas et al., Probabilistic inference of fracture-scale flow paths and aperture distribution from hydrogeophysically-monitored tracer tests, J. Hydrol. 2018).

I miss complementary numerical results such as transport parameters and their discussion (i.e. sensitivity analysis) for a deeper comparative analysis of simulation results in section 5. The recovery rate of the injected tracer for the three examined BTCs would be helpful to the reader to have information about how many tracer mass has been lost during the test. This will help to understand the potential role of rock matrix or stagnant zones in the karst circuit by which anomalous transport is reflected as multi-peaks or long-tailed BTC shapes.

A table that lists the fitted parameter values that correspond to Fig. 4 has been prepared: https://doi.org/10.5281/zenodo.3824439. The table also contains i) the injected and recovered mass values for each tracer experiment, ii) the corresponding recovery ratios, iii) the minimum and maximum parameter values considered for the optimization, iv) the composite parameter sensitivities computed with PEST, and v) the post-calibration scaled values of the parameters pairs (Q, m_j) and (C_0, Q_j/Q), which are consistent with the recovered mass values. This post-calibration scaling was necessary because no parameter was fixed prior to the inversion; so a degree of freedom remained in the pairs of parameters due to their balance in the model equations. Owing to its large size (more than 400 parameters for the different models and channel number scenarios), I propose the inclusion of this table as a supplement rather than as an appendix. I also added a discussion about the (consistent) variation in the values of the parameters that influence the spreading of transit/residence times in the individual flow channels according to the different models and/or the number of flow channels.

The optimized parameter values and their composite sensitivities at the end of the optimization process are provided in the Supplement (Table S1). Unsurprisingly, the model parameters that influence the spreading of transit/residence times in the individual flow channels while accounting for different processes (Pe, γ, β, ψ, and ω) are sensitive to the number of channels. For instance, when comparing single- with multiple-channel models, the former requires lower Pe values to compensate for the coarser description of the flow system heterogeneity (recall that the dispersion coefficient integrated in the Peclet number reflects the unresolved variability of the flow velocity below the modeling scale). The same observation holds when comparing single- and double-porosity models with the same number of flow channels, i.e., the Pe values of single-porosity models are lower than the Pe values of double-porosity models because part of the spreading of transit/residence times in the latter case is implicitly captured by solute mass exchanges between the mobile and immobile domains. A noticeable exception is the diffusion parameter β of the SFDM model, whose values are mostly around 1.0×10^{-3} $h^{-1/2}$. This value corresponds to the upper bound of the optimization range set for this parameter, which is based on a matrix porosity of 30 %, a molecular diffusion coefficient of 1.0×10^{-9} m^2 s^{-1}, and a flow-channel aperture of 1.0×10^{-2} m. Beta values larger than 1.0×10^{-3} $h^{-1/2}$ would be physically unrealistic. The fact that the Beta value is limited by its upper bound during the optimization process indicates that the SFDM model is not suitable for describing the HES tracer experiments, as further discussed below. All other parameters have converged to values far from their optimization bounds.

Concerning your last remark, it is not possible, with the specified assumptions of steady-state flow and non-reactive tracer, to establish a link between the incomplete recovery of a tracer and mass-exchanges between flowing and stagnant water regions. These exchanges are assumed to obey physically reversible processes (regarded either as a first-order system or second-order system in the 2RNE model and SFM model, respectively), and therefore, cannot be regarded as potentially responsible for the incomplete mass recovery. For tracer tests performed in steady state conditions that involve non-reactive tracers, an incomplete recovery of the injected mass indicates a diverging flow structure between the injection site and the monitoring point. Unfortunately, no additional information can be drawn about this flow divergence from the recovered tracer data only. This discussion has been added to the revised manuscript.

The mass recovery ratios for these three tracer experiments were 58 %, 79 %, and 60 %, respectively. Note that these recovery data cannot be included in the model because the flow structure assumption that underlies the multiflow approach (Fig. 1) implies that all the mass that enters the system flows out after a certain lapse of time. The same holds for any single- or double-porosity modeling approach based on a 1-D flow assumption. For tracer tests that are performed in steady state conditions and involve non-reactive tracers, an incomplete recovery of the injected mass indicates a diverging flow structure between the injection site and the monitoring point. Unfortunately, no additional information can be obtained about this flow divergence from the tracer data only. Therefore, the total mass in a multiflow model must be consistent with the recovered tracer mass rather than the injected mass.

I recommend adding at the early sections a glossary of acronyms and parameters described throughout the text.

The acronyms and model parameters have been summarized in two separate tables. Owing to their size, I felt that it was more appropriate to place these two tables (A1 and A2) in the Appendix. Two sentences have been added at the end of the Introduction to refer to these tables. Four new references quoted in Table A1 (Diersch, 2014; Langevin et al., 2017; Zheng et al., 2012; Zheng and Bennett, 2002) have been added to the References.

Appendix A: Glossary

Table A1. Acronyms and model abbreviations utilized in the text.

Acronym or model name	Description	Reference
ADE	Advection-dispersion equation	Zheng and Bennett (2002)
BTC	Breakthrough curve	
CATTI	Computer Aided Tracer Test Interpretation: a computer program for tracer BTC fitting	Sauty et al. (1992)
CMA-ES	Covariance Matrix Adaptation – Evolution Strategy: a global optimization algorithm	Hansen and Ostermeier (2001)
CMAES_P	PEST-compatible program that implements the CMA-ES method	Doherty (2019a)
CXTFIT	Computer program for tracer BTC fitting	Toride et al. (1999)
DADE	Dual-advection-dispersion equation	Field and Leij (2012)
FEFLOW	Finite Element FLOW model; a simulation package for flow, heat, and mass transport in groundwater	Diersch (2014)
GUI	Graphical user interface	

HES	Hydrogeological Experimental Site in Poitiers, France	Audouin et al. (2008)
MDM	Multi-Dispersion Model	Maloszewski et al. (1992)
MDMed	Computer program that implements the Multi-Dispersion Model and assumes a non-instantaneous injection (exponentially decaying concentration) at the inlet of the flow system	This article
MDMi	Computer program that implements the Multi-Dispersion Model and assumes an instantaneous injection of tracer at the inlet of the flow system	This article
MDP	Multi-Double Porosity: a combination of multiflow and double-porosity models	This article
MDP_SFDM	Computer program that implements the MDP approach, where the mass exchanges between the mobile and immobile domains are modeled as a second-order (diffusion) process	This article
MDP_2RNE	Computer program that implements the MDP approach, where the mass exchanges between the mobile and immobile domains are modeled as a first-order process	This article
MFIT	MultiFlow Inversion of Tracer breakthrough curves: a GUI for the MDMi, MDMed, MDP_SFDM, MDP_2RNE, and PEST programs.	This article
MIM	Mobile-Immobile Model	Coats and Smith (1964)
MODFLOW	MODular three-dimensional groundwater FLOW model: a computer code developed by the U.S. Geological Survey that numerically solves the groundwater flow equation	Langevin et al. (2017)
MT3DMS	Modular Three-Dimensional MultiSpecies transport model: a numerical code to simulate solute transport in groundwater	Zheng et al. (2012)
OM-MADE	One-dimensional Model for Multiple Advection, Dispersion, and storage in Exchanging zones: a python script to simulate solute transport in multiflow systems with possible mass exchanges between the flow channels	Tinet et al. (2019)
OptSFDM	Computer program for tracer BTC fitting based on the SFDM model	Gharasoo et al. (2019)
OTIS	One-dimensional Transport with Inflow and Storage: a numerical code to simulate solute transport in streams and rivers	Runkel (1998)
PEST	Parameter ESTimation: a collection of computer programs for model-independent parameter estimation and uncertainty analysis	Doherty (2019a)
SCE-UA	Shuffled Complex Evolution method – University of Arizona: a global optimization algorithm	Duan et al. (1992)

SCEUA_P	PEST compatible program that implements the SCE-UA method	Doherty (2019a)
SFDM	Single-Fracture Dispersion Model	Maloszewski and Zuber (1990)
STANMOD	STudio of ANalytical MODels: a collection of computer programs for tracer BTC fitting	van Genuchten et al. (2012)
SVD	Singular value decomposition	Doherty (2015)
TRAC	Computer program for tracer BTC fitting	Gutierrez et al. (2013)
TRACI	Computer program for tracer BTC fitting	Käss (2004)
1-D	One-dimensional	
2RNE	Two-region nonequilibrium equation	Toride et al. (1993)

Table A2. List of model parameters.

Parameter	Description	Unit	Specific model (an empty box means that the parameter is employed in all the models)
b_j	Half-aperture of the *j*-th flow channel	L	MDP-SFDM
C_j	Concentration in the *j*-th flow channel	ML^{-3}	
C_{pj}	Concentration in the immobile domain assigned to the *j*-th channel	ML^{-3}	MDP-SFDM
C_{imj}	Concentration in the immobile domain assigned to the *j*-th channel	ML^{-3}	MDP-2RNE
C_0	Initial (maximum) concentration at the inflow boundary for an exponentially decaying injection concentration	ML^{-3}	MDMed
D_j	Dispersion coefficient in the *j*-th flow channel	L^2T^{-1}	
D_{pj}	Molecular diffusion coefficient in the immobile domain assigned to the *j*-th channel	L^2T^{-1}	MDP-SFDM
L_j	Length of the *j*-th flow channel	L	
m_j	Part of the solute mass flowing through the *j*-th channel	M	MDMi, MDP-SFDM, MDP-2RNE
N	Number of flow channels	-	
N_{max}	Maximum number of flow channels	-	
Pe_j	Peclet number in the *j*-th channel	-	
P	Number of optimized model parameters	-	
PHI	Measurement objective function (sum of the squared weighted differences between the tracer BTCs and the model-fitted curves)	M^2L^{-6}	

Q	Total system flow rate	L^3T^{-1}	
Q_j	Flow rate in the j-th channel	L^3T^{-1}	
t	Time variable	T	
T_{min}	Minimum time value of the user-provided BTC	T	
T_{max}	Maximum time value of the user-provided BTC	T	
T_5	T_5 time, Eq. (18)	T	
T_{5th}	Earliest time at which the concentration values exceed 5 % of the maximum concentration value	T	
T_{95}	T_{95} time, Eq. (19)	T	
T_{95th}	Latest time at which the concentration values exceed 5 % of the maximum concentration value	T	
T_{0j}	Mean transit time in the *j*-th channel	T	
u_j	Advection velocity in the *j*-th flow channel	LT^{-1}	
x_j	Spatial coordinate along the *j*-th flow channel	L	
y_j	Spatial coordinate perpendicular to the *j*-th flow channel	L	
α_j	First-order mass transfer coefficient between the mobile and immobile domains assigned to the *j*-th channel	T^{-1}	MDP-2RNE
β_j	Diffusion parameter in the *j*-th flow channel, Eq. (12)	$T^{-1/2}$	MDP-SFDM
γ_j	Gamma coefficient in the *j*-th flow channel, Eq. (8)	-	MDMed
θ_j	Volumetric water content of the mobile domain assigned to the *j*-th channel	-	MDP-2RNE
θ_{imj}	Volumetric water content of the immobile domain assigned to the *j*-th channel	-	MDP-2RNE
λ_j	Time decay constant that controls the exponentially decaying release of tracer in the *j*-th channel	T^{-1}	MDMed
ξ	Integration variable, Eq. (11)	T	MDP-SFDM
σ_j	Standard deviation of travel times for transport by advection and dispersion in the *j*-th channel	T	
τ	Integration variable, Eq. (15)	L	MDP-2RNE
ψ_j	Fraction of mobile water in the *j*-th channel, Eq. (16)	-	MDP-2RNE
ω_j	Omega coefficient in the *j*-th flow channel, Eq. (17)	L^{-1}	MDP-2RNE

Point-to-point comments:

Page 12: Table 2, test 4 >>> "Partitioning coefficient (ß)" instead of "Fraction of mobile water (ψ)"?

Page 12: Table 2, test 4 >>> "Mass transfer coefficient" instead of "Omega coefficient"?

As mentioned in Table 2 and the text (lines 296–298), test 4 addresses the case of a single-flow channel that is described as a two-region nonequilibrium (2RNE) medium. The 2RNE model is mathematically described by Equations (13) – (17) and the

model parameters are summarized in Table 1. The parameters ψ and ω are part of the 2RNE model and correct in Table 2, test 4. The suggested changes would be inappropriate since the parameter β is the diffusion parameter in the SFDM model, and the "mass transfer coefficient" corresponds more specifically to the parameter α, which is part of ω (refer to line 175 and Eq. 17).

Anonymous Referee #2

Figure 1: should this say what the dashed line represents, is it non flowing water?

The dashed line was supposed to represent possible additional flow channels (greater than 3 and less than N). I removed this line for clarity.

Line 115-116: "A possible reason is the increasing number of fitting parameters, which makes the inverse problem more complicated. The use of modern inversion tools such as PEST enables overcoming this problem, as discussed in section 3" I agree that these methods can efficiently find parameters sets in the situation you outline but I would assume not without the possibility that the parameters best fitting the data are far from unique and more so the greater the number of parameters. For me, this sentence misses a discussion of this important caveat in an otherwise very carefully considered section.

A short discussion about the non-uniqueness (equifinality) issue has been added in this section.

Among the challenges related to the inversion of a multiflow model is the inherent problem of nonuniqueness (or equifinality). A variety of parameter sets can yield nearly identical simulated BTCs because the change in the value of a parameter of a given channel can be compensated by modifying at least one other parameter that pertains to this same channel or the parameters of the other channels. This nonuniqueness causes the inverse problem to be ill-posed in the sense of Hadamard (1902) and requires the use of advanced optimization methods, such as regularization, to make the inverse problem tractable (Tikhonov and Arsenin, 1977; Moore and Doherty, 2006; Zhou et al., 2014).

The uncertainty associated with the inverted parameter set and the methods that can be employed to assess this uncertainty are addressed again at the end of Chapter 3 and at the end of section 5 (application example).

Around line 235: For my understanding, is the optimisation run for a given number of channels and if so should the user seek the minimum number of n that perform well for the measurement objective function and regularisation terms. OK, I see later where this comes in but I'll leave the comment so you can see the issue I had when reading for the first time.

Clarification was necessary; the text has been amended.

The optimization and uncertainty analysis of the model parameters for a given number of flow channels are carried out using PEST routines (Doherty, 2019a, 2019b). The influence of the number of channels on the model fitting performance can be analyzed once a series of calibrations has been performed for a variety of channel numbers, as illustrated below.

Around line 285: A series of utility functions are called here for the uncertainty analysis. I don't think they need further explanation here but a pointer to the relevant documentation/literature on these would aid completeness.

A sentence that specifies appropriate references has been added.

The method is essentially similar to that described by Fang et al. (2019) and relies on the use of the PREDUNC7 and RANDPAR utilities documented in the PEST manual (Doherty, 2019b).

Section 4 code verification – should this also test for the case where n channels is unknown? So for test 5 if n_max was set to 6 would the same results be found as for the current test. Perhaps this goes beyond verification of the transport processes models, which is clearly the aim of this section, but I think checking the multistart would add value if feasible.

A new test has been implemented to assess the multistart method. The related discussion has been placed in a new/separate section (section 5) for the sake of clarity.

5 Assessment of the multistart optimization method

The purpose of this section is to assess the automatic multistart method described in section 3 using a new synthetic test case. A multimodal BTC that corresponds to 3 channels has been simulated using the MDMi program with the parameters listed in Table 3. A "blind" inversion of this BTC has been performed using the automatic multistart method with a maximum number of flow channels N_{max} = 6. The only model parameter that has been fixed prior to the inversion process was the total flow rate Q to simplify the post-comparison of the inverted mass values in each channel with the "true" mass values. Otherwise, a degree of freedom would persist for the pairs of the optimized Q and m_j values, i.e., multiplying or dividing these parameters by the same constant would yield the same BTCs; refer to Eq. (6). The parameters m_j, T_{0j} and Pe_j of the different flow channels were optimized with virtually no upper and lower bound constraints (minimum and maximum allowed parameter values of 1.0×10^{-10} and $1.0 \times 10^{+10}$, respectively). As shown in Fig. 3, the inverted BTCs that correspond to N = 3, 4, 5, and 6 channels overlap perfectly with each other and with the original simulated BTC; and as shown in Table 4, the optimized values for the parameters of the 3-channel model are equal to the "true" parameter values.

Table 3. Model parameters that correspond to the multimodal simulated BTC in Fig. 3.

Parameters	Values
Q	10 $m^3 h^{-1}$
m_1	10 g
m_2	6 g
m_3	4 g
T_{01}	150 h
T_{02}	250 h
T_{03}	350 h
Pe_1	20
Pe_2	50
Pe_3	100

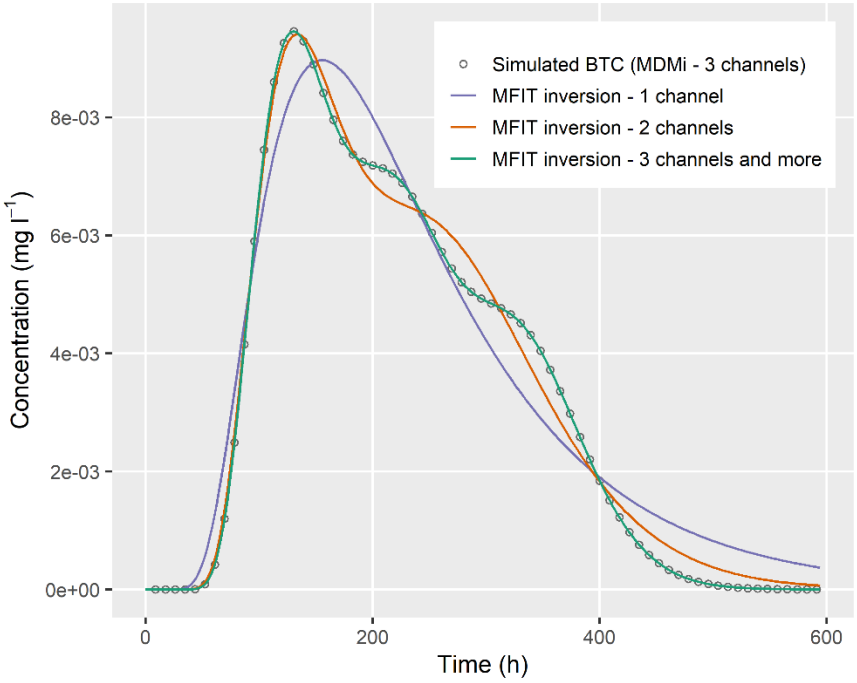

Figure 3. Inversion of the 3-channel-simulated BTC using the automatic multistart method with N_{max} = 6. The inverted BTCs that correspond to N = 3, 4, 5, and 6 channels overlap perfectly with each other and the original simulated BTC.

Table 4. Optimized model parameters that correspond to the inverted BTCs in Fig. 3.

N	1	2	3	4	5	6
m_1 (g)	21.11	10.79	10.00	2.79	2.66	2.66
m_2 (g)	-	9.54	6.00	7.19	7.35	7.35
m_3 (g)	-	-	4.00	6.02	5.99	5.91
m_4 (g)	-	-	-	4.00	2.58	2.62
m_5 (g)	-	-	-	-	1.42	1.45
m_6 (g)	-	-	-	-	-	0.02
T_{01} (h)	239.36	155.82	150.00	126.17	151.55	151.31
T_{02} (h)	-	302.91	250.00	158.91	149.47	149.60
T_{03} (h)	-	-	350.00	250.00	249.97	249.30
T_{04} (h)	-	-	-	350.01	349.48	347.68
T_{05} (h)	-	-	-	-	350.84	351.08
T_{06} (h)	-	-	-	-	-	405.58
Pe_1	6.72	17.52	20.00	24.22	19.80	19.92
Pe_2	-	27.55	50.00	22.18	20.07	20.01
Pe_3	-	-	100.00	49.94	50.04	50.62
Pe_4	-	-	-	100.00	98.45	88.42
Pe_5	-	-	-	-	102.68	120.27
Pe_6	-	-	-	-	-	442.13

Figure 3: In my version the dots and labels overlap, generally this figure could be cleaner, and the scale bar is also quite small. Would it also be possible to highlight the wells used for pumping and injections in the experiments, perhaps with colours or different symbology.

This figure (Fig. 4 in the revised manuscript) has been revised.

Figure 4. Locations of wells at the HES in Poitiers, France. Map data are from Google.

Line 345: Do the parameter bounds come into play in the optimised parameter sets? i.e. do you get parameters optimising to the bounds? Generally, there is not any discussion of the parameters found, we there a reason for this? I think this should be justified.

A table that lists the fitted parameter values that correspond to Fig. 4 has been prepared: https://doi.org/10.5281/zenodo.3824439. The table also contains i) the injected and recovered mass values for each tracer experiment, ii) the corresponding recovery ratios, iii) the minimum and maximum parameter values considered for the optimization, iv) the composite parameter sensitivities computed with PEST, and v) the post-calibration scaled values of the parameters pairs (Q, m_j) and (C_0, Q_j/Q), which are consistent with the recovered mass values. This post-calibration scaling was necessary because no parameter was fixed prior to the inversion; so a degree of freedom remained in the pairs of parameters due to their balance in the model equations. Owing to its large size (more than 400 parameters for the different models and channel number scenarios), I propose the inclusion of this table as a supplement rather than an appendix. I also added a discussion about the (consistent) variation in the values of the parameters that influence the spreading of transit/residence times in the individual flow channels according to the different models and/or number of flow channels. A bound effect on the optimization of parameter β (SFDM model) has been noted and discussed.

The optimized parameter values and their composite sensitivities at the end of the optimization process are provided in the Supplement (Table S1). Unsurprisingly, the model parameters that influence the spreading of transit/residence times in the individual flow channels while accounting for different processes (*Pe*, γ, β, ψ, and ω) are sensitive to the number of channels. For instance, when comparing single- with multiple-channel models, the former requires lower *Pe* values to compensate for the coarser description of the flow system heterogeneity (recall that the dispersion coefficient integrated in the Peclet number reflects the unresolved variability of the flow velocity below the modeling scale). The same observation holds when comparing single- and double-porosity models with the same number of flow channels, i.e., the *Pe* values of single-porosity models are lower than the *Pe* values of double-porosity models because part of the spreading of transit/residence times in the latter case is implicitly captured by solute mass exchanges between the mobile and immobile domains. A noticeable exception is the diffusion parameter β of the SFDM model, whose values are mostly around 1.0×10^{-3} $h^{-1/2}$. This value corresponds to the upper bound of the optimization range set for this parameter, which is based on a matrix porosity of 30 %, a molecular diffusion coefficient of 1.0×10^{-9} m^2 s^{-1}, and a flow-channel aperture of 1.0×10^{-2} m. Beta values larger than 1.0×10^{-3} $h^{-1/2}$ would be physically unrealistic. The fact that the Beta value is limited by its upper bound during the optimization process indicates that the SFDM model is not suitable for describing the HES tracer experiments, as further discussed below. All other parameters have converged to values far from their optimization bounds.

Figure 5: Could the legend be a single legend for all plots. On my version the legend is also quite small making dashed and continuous lines difficult to identify. Worth checking in the final production of the figure for publication.

The figure has been modified to include only one legend. The readability problems are most likely related to the .pdf file that was generated for the review. This figure is a vector graphic and should not pose any production/re-editing issues if the article is accepted for publication in GMD.

Line 390 uncertainty analysis – Could you be more explicit about why the particular model and test case was chosen for the uncertainty analysis. Line 390 uncertainty analysis – Could you be more explicit about why the particular model and test case was chosen for the uncertainty analysis. Line 396: "fairly similar" could you be more precise about how similar was defined. The uncertainty analysis description in the methods is quite brief which means its difficult to fully appreciate the setup here in my opinion. Around line 400 – I feel the discussion of these results is somewhat rushed regarding the uncertainty analysis, I don't feel I fully appreciate the results. Is this conclusion made because only the four and six channel models capture the first peak? MDP-2RNE seems to for the two channels although it's difficult to see if this is really the case.

Since this part of the original manuscript was unclear, I rewrote it and expanded the uncertainty analysis by considering i) two models (MDMi and MDP-2RNE) instead of one model (MDMi in the original manuscript) and ii) different numbers of channels (1, 2, and 3) in each case instead of a single two-channel model in the original manuscript. As outlined in the revised manuscript, the objective is to illustrate the possibilities offered by MFIT for the analysis of uncertainties associated with the inverted parameters. It would be inappropriate to rely on this analysis to compare the performances of the different models since, as discussed in the revised manuscript, the choice of a model and a number of channels (i.e., degree of complexity of the model) depends on the objective pursued by the modeler. In the revised manuscript, the results of Fig. 7 and Fig. 8 are discussed in terms of the equifinality of the inverted model parameters.

The Pareto curves in Fig. 6 (previously Fig. 5) indicate that the *final choice* of a model, if one is to be made, relies on a tradeoff between the desired fitting accuracy and the desired degree of simplification/complexity with respect to the model structure (number of channels and/or number of model parameters). Beyond this subjective (expert) decision, which may depend on the goal of the study, and therefore, will not be discussed further in the present application case, uncertainty remains in the inverted

model parameters as a consequence of the nonuniqueness of the inverse problem. This uncertainty is related to both the equifinality of the model parameters, which is partly due to the multiflow framework structure, and the measurement noise in the tracer BTCs. Figures 7 and 8 illustrate the post-calibration uncertainty analysis capabilities of MFIT, via an assessment of the MDMi and MDP-2RNE model fittings of the M16-M22 tracer BTC with 1, 2, and 3 flow channels. Owing to the balance between the Q and m_j terms in the model equation (Eqs. (6) and (15)), at least one of these parameters must be fixed to assess the uncertainty of the other parameters. Here, the value of Q was set to 25 $m^3 h^{-1}$, which ensures the consistency of the model against the recovered tracer mass that was independently calculated from the experimental data (refer to Table S1). Following the PEST optimization of the different model parameters, 500 calibration-constrained parameter fields were stochastically generated and recalibrated by PEST. Depending on the model (MDMi or MDP-2RNE) and number of flow channels, between 483 and 500 recalibration runs successfully achieved a level of fit that is fairly similar (i.e., within a tolerance of +5 % for the PHI value; refer to section 3) to that associated with the original calibration parameter field. The histograms shown in Figs. 7 and 8 were constructed from these recalibrated parameter sets and illustrate the multitude of parameter combinations that are equally good, for a given number of flow channels, in terms of fitting the M16-M22 tracer BTC. As shown in these figures, the confidence intervals are quite narrow for most parameters but tend to widen as the number of channels increases, which reflects the equifinality of the multiflow modeling approach. Although not shown here, it has been established that the tailed behaviors of the parameters L_j and ω_j in Fig. 8 are due to a partial correlation between these two parameters (refer to Eq. (15)), i.e., fixing the value of one parameter prior to the inversion drastically reduces the uncertainty of the other parameter. As previously discussed, the higher Pe values in Fig. 8 compared to Fig. 7 are due to the fact that the distribution of the transit/residence times with the 2RNE model is primarily controlled by the solute mass exchanges between the mobile and immobile domains.

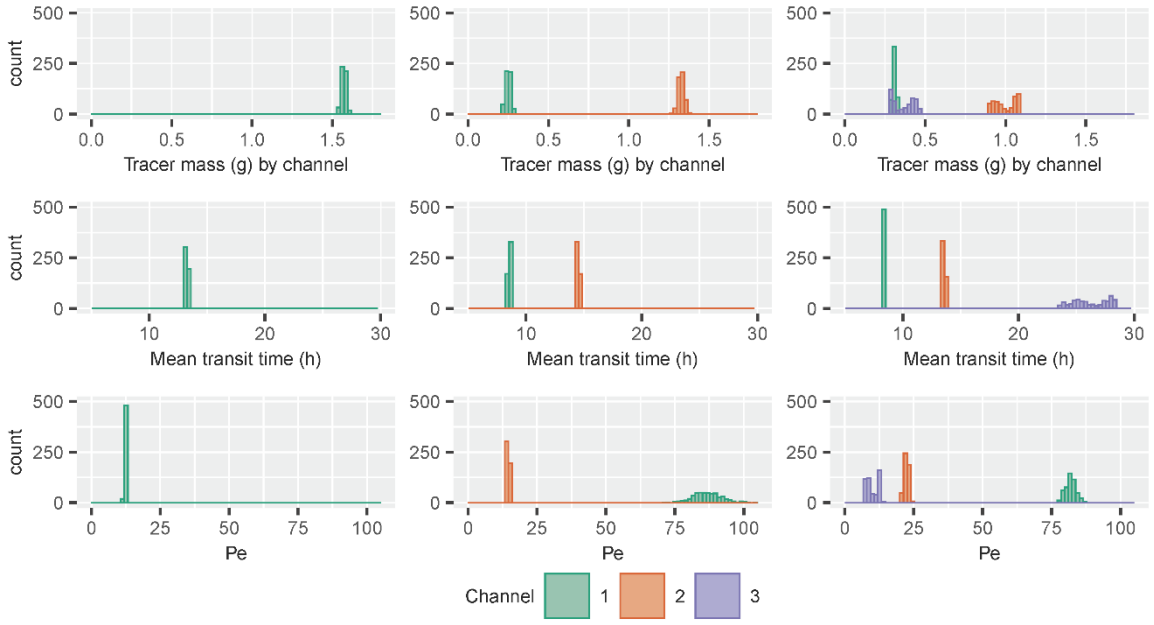

Figure 7. Postcalibration uncertainty of model parameter values for the inversion of the M16-M22 tracer BTC by the MDMi model with 1, 2, and 3 flow channels.

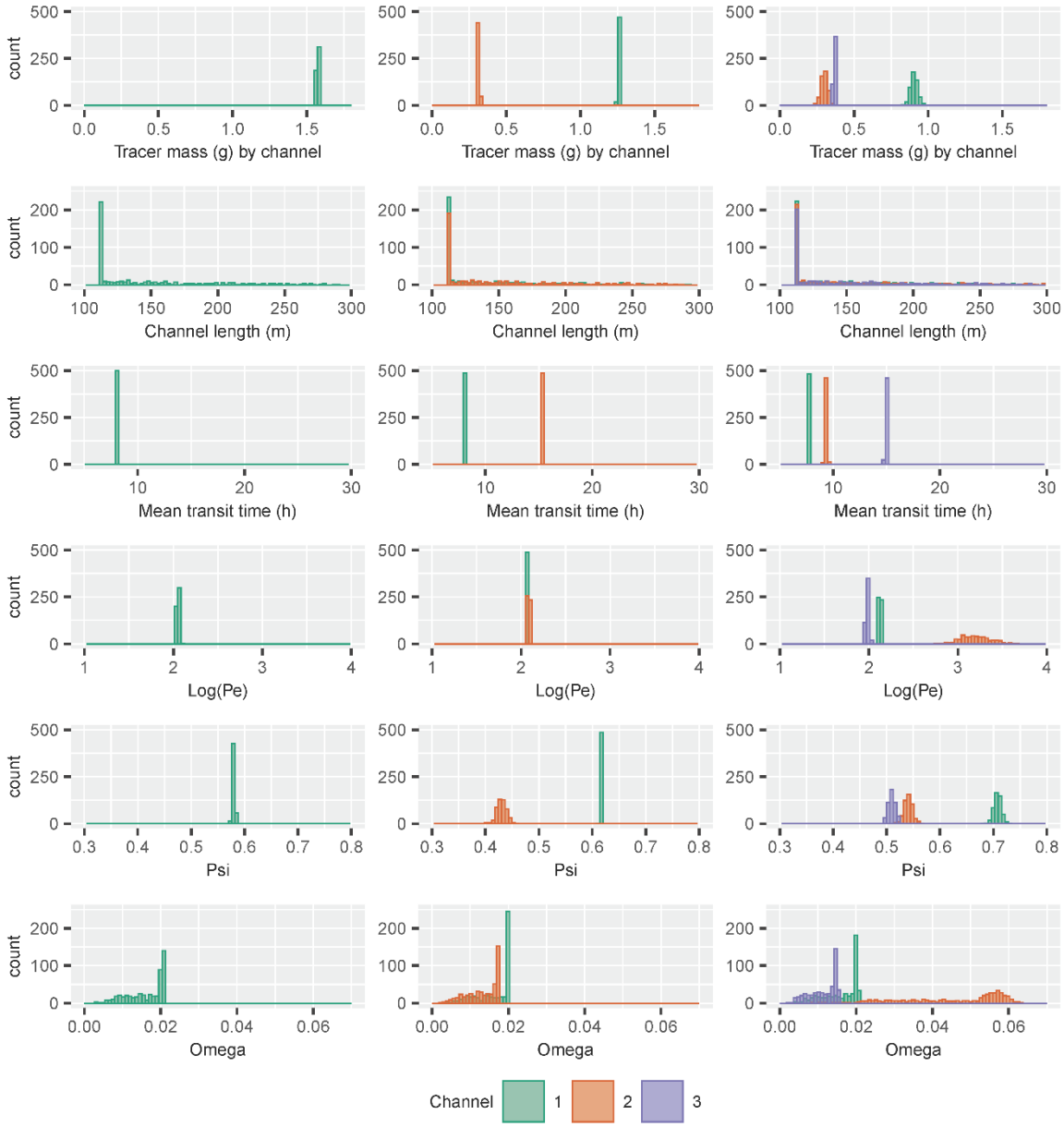

Figure 8. Postcalibration uncertainty of model parameter values for the inversion of the M16-M22 tracer BTC by the MDP-2RNE model with 1, 2, and 3 flow channels. A logarithmic scale has been employed for Pe due to a wider range of values than shown in Fig. 7.

References that have been added to the revised manuscript:

Diersch, H.-J.: FEFLOW: finite element modeling of flow, mass and heat transport in porous and fractured media, Springer-Verlag, Berlin Heidelberg, doi:10.1007/978-3-642-38739-5, 2014.

Fang, Q., Ma, L., Harmel, R. D., Yu, Q., Sima, M. W., Bartling, P. N. S., Malone, R. W., Nolan, B. T. and Doherty, J.: Uncertainty of CERES-Maize calibration under different irrigation strategies using PEST optimization algorithm, Agronomy, 9(5), 241, doi:10.3390/agronomy9050241, 2019.

Hadamard, J.: Sur les problèmes aux dérivées partielles et leur signification physique, Princet. Univ. Bull., 49–52, 1902.

Langevin, C. D., Hughes, J. D., Banta, E. R., Niswonger, R. G., Panday, S. and Provost, A. M.: Documentation for the MODFLOW 6 Groundwater Flow Model, USGS Numbered Series, U.S. Geological Survey, Reston, VA., 2017.

Moore, C. and Doherty, J.: The cost of uniqueness in groundwater model calibration, Adv. Water Resour., 29(4), 605–623, doi:10.1016/j.advwatres.2005.07.003, 2006.

Tikhonov, A. N. and Arsenin, V. Y.: Solutions of ill-posed problems, Winston & Sons, Washington., 1977.

Zheng, C., Hill, M. C., Cao, G. and Ma, R.: MT3DMS: model use, calibration, and validation, Trans. ASABE, 55(4), 1549–1559, doi:10.13031/2013.42263, 2012.

Zheng, C. and Bennett, G. D.: Applied contaminant transport modeling, 2nd ed., Wiley-Interscience, New York, 2002.

Zhou, H., Gómez-Hernández, J. J. and Li, L.: Inverse methods in hydrogeology: Evolution and recent trends, Adv. Water Resour., 63, 22–37, doi:10.1016/j.advwatres.2013.10.014, 2014.