Referee 1:
We thank the reviewer for the positive assessment of our work.

We will include his editorial comments in the revised version.

The reviewer requests clarification regarding the slope (*m*) of breakthrough curves defined as the exponent of the power law decrease with time *t* of the concentration *c* such as $c \sim t^{-m}$. Unfortunately, it is not well established what controls it. Willmann et al (2008) found some correlation between the degree of connectivity and the slope. The more connected the field, the smaller the slope. In this context, fractured media represent the lowest bound (*m*=1.5), which is controlled by diffusion into immobile regions. A slope of 2.5 may therefore represent a heterogeneous but poorly connected hydraulic conductivity field, where late time arrival is controlled by slow advection. We have expanded the discussion in the revised version as follows:

> *The power-law distribution is consistent with observed breakthrough curves in Heterogeneous Porous Media, which often display long tails that appear linear in log(c) versus log(t) (Gouze et al., 2008; Haggerty et al., 2004; Li et al., 2011; Silva et al., 2009; Willmann et al., 2008). This tailing is well modeled by a power law, such that the breakthrough concentration c evolves as $c \sim t^{-m}$. Haggerty et al. (2000) showed that the slope m relates to the exponent of the power-law distribution of the MRMT rates (equation (3)). m is generally found to be in the interval [1.5;2.5] but litte is known about its relationship to the geological heterogeneity. Willmann et al (2008) found some correlation between the degree of connectivity and the slope. The more connected the field, the smaller the slope. In this context, fracture/matrix exchanges in fractured media represent the lowest bound (m=1.5), which is controlled by diffusion into immobile regions (Haggerty and Gorelick, 1995). On the contrary, a slope m of 2.5 may represent a heterogeneous but poorly connected hydraulic conductivity field, where late time arrival is controlled by slow advection.*

Referee 2:
We thank the reviewer for the positive assessment of our work.

We will include his editorial comments in the revised version.

The referee raises three essential issues:

1 - How do you choose MRMT parameters in the sensitivity analysis?

The reviewer is right in that sensitivity around a point in the parameter space is insufficient to characterize the behavior of the solution. We did a lot more runs than presented in the paper, but limited the cases presented in the manuscript to facilitate readability. To maintain it, we will not discuss additional parameter sets, but we will mention that the analysis has also been made for other combinations of parameters.

Extreme values have been investigated to get the possible range of behaviors. For slopes, we adopted the range observed in nature. The porosity ratio does not have an upper bound. In fact, ideally, the mobile porosity could be zero. We adopted 150 as a large upper value. Larger upper values would not affect results and might cause numerical difficulties. The same can be said for t_N/t_1. We have added the following justification:

For MRMT, extreme values have been investigated to get the possible range of behaviors. For slopes m, we adopted the range observed in nature as discussed in section 2.1 with m varying between m=1.5 (typical fracture/matrix case) and m=2.5. The Single Rate Mass Transfer (SRMT) is also shown for comparison. The porosity ratio β does not have an upper bound. In fact, ideally, the mobile porosity could be zero. We adopted β=150 as a large upper value. Larger upper values would not affect results and might cause numerical difficulties. The same can be said for t_N/t_1 for which we took $t_N/t_1=10^3$ as the upper bound. The analysis presented hereafter has been made for different combination of paramters within these bounds. As all models lead to consistent conclusions, we only present the most characteristic results.

2 - Is it possible to fit MRMT on gamma rather than on conservative breakthrough curves?

This is a very interesting question. In fact, it is what motivated our work. If one could do it, then it would be possible to find an effective model that indeed honors spreading and mixing. What we find is that, in general, this cannot be achieved as there is already a tradeoff to fit gamma between the amplitude of the deviation and the timing of it.

3- The authors claim here (page 12990, lines 19-20?) that dispersivity is always negligible with respect to mass exchange effects. Can the authors provide a reference to support this statement? Is this general? And if so, why not completely disregarding dispersivity?

We do not quite understand this comment as we never make such comment for dispersivity in general. The referee may refer to dispersivity in the mobile zone of MRMT models, in which case, we agree. We might as well neglect mobile zone dispersivity. We will quote earlier work on MRMT that show the dominance of exchanges on residence time distribution and thus effective dispersivity of MRMT models.

However, throughout the paper, when we discuss dispersivity, we refer to either dispersivity of the HPM, which drives the generation of gradients in concentration and thus mixing, or effective dispersivity of the MRMT model, which is controlled by mobile-immobile exchanges. In either case, dispersivity is not negligible with respect to mass exchange effects.

We have expanded the discussion and clarify these three issues in the revised version of the paper as follows:

In summary, in HPM, dispersivity comes primarily form the velocity structure, which drives the generation of gradients in concentration and thus mixing. Instead, in MRMT, effective dispersivity is controlled by mobile-immobile exchanges and delays the actual mixing between the immobile and mobile solute concentrations.