My thanks to Referee # 2 for comments. Responses follow below.

" *The author likes to call distributions with non-zero probability density at time zero as "L-shaped" and argues with modes of the distributions at $t > 0$.* "

It is not a question of "likes to call". The term "L-shaped" is a widely used as a distribution descriptive term. For the purposes of the paper a specific definition is used. That is, a transit time distribution f(t) with origin at zero is defined to be L-shaped if $f'(t) < 0$ for $t \geq 0$. I could not imagine anyone contemplating defining a distribution to be L-shaped on the basis of a non-zero probability density at time zero. I am totally perplexed therefore why I should stand accused by both reviewers of doing just that. After all, $f'(t) < 0$ for $t \geq 0$ is hard to misinterpret.
Of course there are other distributions forms which are also conceptually incorrect. However L-shaped forms (as defined in the paper) are widely used as transit time distributions and focus therefore is only on those forms – as indicated by the title of the paper.

I am entirely happy to remove the thought experiment in the final version of the paper, if accepted. It was presented in the Discussion version in anticipation of the kinds of criticism that would be raised.

The referee's second paragraph raises the important points of parsimony and distribution identifiability. At first glance it all sounds plausible. Assume a simple L-shaped distribution, such as the 1-parameter exponential. This is then applied to data of complex origin and, assuming a satisfactory fit is obtained, it would be reported as "a satisfactory data fit was obtained using an exponential transit time distribution". Similar studies over the years would build up to "numerous studies have shown that the exponential distribution is a suitable model for many transit time distributions", followed by the obligatory extended citation list.

However, knowingly using an incorrect distribution form for convenience carries its own penalty against good hydrological science. That is, using an exponential distribution and doing nothing more ignores the fact that the transit time distribution must in reality have a mode at some $t > 0$. As noted by the referee, distribution identification is essentially impossible with real-world data. That cuts both ways. If the exponential distribution is recognised as never being correct for time zero, then that should raise the question "how far away from time zero might the first mode be located while still leading to an equally good fit to data?". This is a legitimate hydrological question and it does not amount to the impossible task of identification of a specific alternative transit time distribution. For example, a nonparametric histogram might be utilised as alternative. Such a distribution could be constrained to have the correct form (zero probability near zero) and would also provide a degree of sensitivity measure for the transit time distribution application concerned.

The reference in the paper to the possible use of unimodal two-parameter distributions was not in terms of more complex alternatives to the 1-parameter exponential distribution, but as potential alternatives to the 2-parameter gamma distribution with shape parameter < 0. As noted in the paper, the exponential distribution has been shown to be unsuitable as a transit time distribution in a number of catchment studies. And yes, there is no implication of Fickian systems in reality if using inverse Gaussian distributions – just as Kirchner et al (2001) presumably did not consider they were modelling Fickian slope water movement when they used a mixture of inverse Gaussian transit time distributions.

"*As a way out, the author suggests to use the gamma distribution with a shape parameter slightly bigger than one*."

No – in fact for the purposes of using a single parametric distribution where the exponential distribution gives suitable fit to data, the point is made in the paper that there no need to seek a "way out". As mentioned above, the more important issue is the extent to which the data might permit the mode to be displaced from zero, even when the exponential model works well with the data. However, for the purposes of using the exponential distribution as a single parametric distribution which works with the data, there is no need to consider some specific numerical value of a shape parameter marginally greater than 1 because the "L-shaped" and "non-L-shaped" versions will be indistinguishable in this case (unlike gamma distributions with shape parameter considerably less than 1). To that extent the "call" for abandoning L-shaped distributions is moderated for the special case of the exponential distribution.

" *But if your data are not that great, you stick to the simplest distributions thinkable*."

Outside of quantum concepts, the true nature of the physical world exists entirely independently of our ability to measure it. Data quality is therefore no excuse to use a model which we know to be incorrect in some specific aspect, whether the model concerned be simple or complex. Transit time distributions with zero probability density at time zero always exist in nature even if we have no measurements at all. To be sure this is limited knowledge, but it would be unfortunate to ignore it. Of course, with poor data quality any attempt at parameter interpretation will not mean very much.

With respect to the "editorial" comments:

I would not drop the "L-shaped" title because that describes exactly the content of the paper, and just what is meant by L-shaped is exactly defined in the context of the paper, as mentioned earlier.

Point taken with respect to water "stores", though I would prefer something like "water flow system".

It does not make sense to integrate exactly to the river system because that includes particles already at the observation point at time zero and are therefore not part of the transit time distribution. That is why it was necessary for this discussion document to start off with the thought experiment description. With respect to the "rain falling into the river" – if the tracer particles concerned fell into the river anywhere upstream of the conceptual observation point then the particles would require finite time to be carried to the observation point. If the rain falls directly "on" the conceptual recorder then the particles have not been involved in transiting the flow system and are therefore not part of the transit time distribution.

Finally, a response to the referee query "is this new?". In some ways this is a hard question to answer. Perhaps it really is common knowledge that L-shaped transit time distributions cannot exist in reality. If that is the case then it does seem unusual then that, for example, gamma distributions with shape parameters less than 1 should be utilised in transit time modelling. Given that two parameters are required, it would seem preferable to make a thorough search through the many available 2-parameter distributions with zero probability density at time zero to see if the data can be matched as well. To my knowledge there has not been a previous paper drawing attention to the impossibility of L-shaped transit time distributions, nor any call to seek to find how far from time zero the data might permit the first transit time mode to be located.