Reviewer 1

I commend the authors for their work. The manuscript reflects substantial field and modelling efforts and demonstrates a thorough understanding and careful investigation of this important topic. Nonetheless, I have both specific and more general questions and concerns regarding the presented approach, which I believe should be addressed before considering this work for publication in leading journal such as HESS.

We sincerely thank the reviewer for his/her constructive comments on the manuscript. We recognize that many of the comments were highly relevant. Indeed, in several parts of the manuscript, the text was not sufficiently fluid and readable. Moreover, the references were somewhat unbalanced toward the authors' own literature. From a technical point of view, we found the comments on Figure 3 especially helpful. Indeed, the original manuscript contained a misinterpretation of the VCP and HCP signals, a point also raised by Reviewer 2. These suggestions were fully accepted. In any case, we have tried to provide a clear response to all the many technical issues raised by the reviewer, by also integrating or modifying the text of the original manuscript where needed.

L23 – 25 Consider avoiding "microscopic" and "macroscopic" in this context.

In the manuscript, these terms were used to distinguish between the different observation windows of point-scale measurements and the EMI sensor. In any case, they may be somewhat misleading and thus we decided to delete them.

L33 "worldwide", I agree, and references should reflect this. At the moment there seems to be a bias towards the authors' – Italian work, in the first paragraphs but also throughout the manuscript. This also results in a lack of references where the authors possibly did not have suitable works.

We agree. We have included relevant references where needed

L136 As I side note about the ERT, I think ERT is the strong choice for monitoring infiltration processes: autonomous and more stable time-lapse measurement, arguably also higher vertical resolution, and control on the data quality. While the list of advantages and disadvantages is long and case-specific, the wider use of ERT in time-lapse applications is clear and should be kept in mind.
L142 "model-space calibrations are more widely used". The supporting references continue to be from the authors or collaborators, which weakens the introduction and raises some concerns. I would not agree on this, for example.

Yes, we agree with the last two remarks by the reviewer 1. Accordingly, we revised the text as follows:
"Some strategies were developed to calibrate EMI without inversion by using the conductivity models obtained from Electrical Resistivity Tomography, ERT, as inputs for EMI forward simulations (von Habel et al., 2014; Lavoué et al., 2010; von Hebel et al., 2019). ERT is especially suitable for this comparison, as it is a particularly effective approach for monitoring infiltration processes (Binley et al., 2002; Dietrich et al., 2014). It allows autonomous and relatively stable time-lapse measurements, provides greater control over data quality and, depending on the survey configuration and site conditions, may also offer

higher vertical resolution than EMI.In the von Habel et al. (2014) paper, the resulting simulated σ_a is compared to those measured by EMI to calibrate it (see figure 7 in von Habel et al., 2014).This data-space calibration, however, strongly depends on the ERT inversion accuracy, which, like EMI, is an ill-posed problem. Both EMI and ERT inversions are highly sensitive to data quality and the regularization framework. They should incorporate prior constraints like known profile discontinuities. Also, ERT measurements are time and effort consuming at large scales. Moreover, EMI data still needs inversion if the goal is to obtain σ_b distributions. Alternatively, a model-space calibration may be used, where the calibration is carried out on the σ_b distributions coming from preliminary inversion of σ_a series rather than on the original σ_a series themselves (Hendrickx et al., 2002; Coppola et al., 2016; Huang et al., 2016; Dragonetti et al., 2018; 2022; Farzamian et al., 2021; Autovino et al., 2026). In general, the choice of using either a data space or a model space calibration cannot be done a priori, as it depends on the survey configuration, the availability and quality of calibration data, as well as the vertical heterogeneity of the soil profile (Lavoué et al., 2010; Huang et al., 2016)."

L141 "EMI still needs inversion if the goal is to obtain σa distribution." Did you mean σb?

Yes, right.

L165 In Dragonetti et al. (2022), if I understand correctly, the EMI data (σa) or inverted model (σb) were not scaled (e.g., relative to the TDR) but the inverted σb model was converted into moisture content with a laboratory calibration (θ - σb)? This way EMI is converted into moisture content but the - investigated - effect of the footprint is maintained.

Yes, correct. We clarified this in the text:
"It is worth noting that, in their approach, the authors intentionally chose not to scale the original EMI σ_a nor the σ_b derived from the EMI inversion. To the contrary, the inverted σ_b distributions were converted into soil water content using a laboratory-derived θ–σ_b calibration relationship, thus preserving the effect of the EMI measurement footprint as the aim was just to assess how the TDR direct measurements and the EMI observations captured the same infiltration process."

L182 " the hydrological model itself could provide a consistent framework to reconcile the two seemingly different observations from TDR and EMI by appropriately scaling the initial water contents observed by the two techniques and the saturated water contents of their respective water retention curves." Yes, but this would mean having the wrong saturated water contents? And possibly a hydraulic conductivity curve that (although apparently similar) compensate for this? This part is a bit convoluted and I am not sure I can follow it.

We rephrased this section and clarified that:
"EMI and point- scale measurements provide two apparently different representations of the infiltration process. This does not mean that one is more reliable than the other. Simply, they see different water content distributions because of their different observation volumes and spatial sensitivities.
To us, this seems an interesting interpretative key and has important implications for understanding what an EMI sensor and point-scale measurements (TDR, soil sampling, ...) can really observe in a hydrological process, such as the propagation of a wetting front during

infiltration. EMI and point- scale measurements (e.g., TDR and soil sampling) provide two apparently different representations of the infiltration process. This does not mean that one is more reliable than the other. Simply, they see different water content distributions because of their different observation volumes and spatial sensitivities. The key point is that a soil hydrological model may be able to reproduce both representations by simply using the specific initial water-content distributions observed by the two techniques and by scaling the corresponding water retention curves by the ratio between the technique-specific saturated water contents. If this result were confirmed, there would be no need to empirically adjust the EMI-based σ_b or θ time series to match the point-scale measurements at each time and depth. To the contrary, either representation could be converted into the other by scaling the corresponding water retention curve and initial water-content profile."

L234 Consider specifying if you use a GPS or some manual markers to track the continuous Mini Explorer measurements?

Done

L260 – EQ1 I think there is a typo, the denominator should have a minus sign?
To us equation 1 is correct as written. The apparent relative dielectric permittivity is related to the electromagnetic-wave propagation velocity by $K_a = \left(\frac{c}{v}\right)^2$, where c is the velocity of light in vacuum and v is the propagation velocity in the medium.

L279 Consider clarifying how the undisturbed samples are collected, considering the predominant sand content and the relevance of the samples (bulk density and thus moisture contents, saturated moisture content analysis, also laboratory electrical calibration?).

Done. We integrated the text with:
"Due to the strong sandy composition of the soil, the samples were taken at relatively high water content to avoid soil losses from the cylinder during sampling"

L321 Regarding the regularization, yes, I agree on the differences and possible specific advantages of a sharp regularization. However, I also think that the key aspects remain how many measurements are available, as well as their independence and error. Consider balancing these theoretical aspects with more specific details, for example, would it be good to see the inversion misfits associated with the six measurements on each position?

We thank the reviewer for this useful suggestion. We agree that the inversion misfit provides important information on the quality of the inversion. Therefore, we have included in the revised manuscript the data misfit associated with the inversion of the six apparent conductivity measurements, showing that the adopted inversion provides a satisfactory compromise between fitting the data and obtaining a physically plausible conductivity model.We report here the revised text:
"Table 3 reports the Root Mean Square Error (RMSE) between measured and modelled σ_a calculated for all the six coil configurations for different times. For the final time (9 hours), the RMSE values were 4.79 and 1.31 mS m^{-1} for the HCP71 and HCP118 configurations, respectively, with an overall RMSE of 3.84 mS m^{-1} when considering all six configurations.

The RMSE values in the table seems to confirm the behaviour already described above. The highest RMSE were observed for the VCP32 and HCP71 configurations, whereas the lowest values were observed for the HCP118 configuration.
To us, rather than indicating a lower inversion quality, this behaviour reflects the different sensitivity function of the coil configurations. Actually, as already discussed, the VCP32 and HCP71 configurations are mainly sensitive to the shallow and intermediate layers of the soil profile, where the propagation of the wetting front induces strong and rapidly evolving conductivity gradients. These gradients are quite difficult to be reproduced with a 1D layered conductivity model. To the contrary, HCP118 results in a relatively lower misfit because this coil configuration integrates the electrical response over a deeper soil volume, with relatively lower changes of conductivity. To an extent, the increase of the RMSE values associated with the VCP32 with the infiltration evolution likely reflect an increasing complexity of the conductivity profile rather than problems in the inversion procedure. "

In any case, the issue has been already discussed in Deidda et al., 2014, which was already cited in the original paper.

L334 OK the stability of the non-irrigated part, did you also check the water mass balance or other aspects more directly related to the inversion of the dynamic part?

We agree that a water-mass-balance analysis could, in principle, have provided an additional check on the goodness of the inversion. However, a rigorous mass-balance assessment could not be performed as it would have required first a conversion of the EMI-based σb to water contents. More importantly, drainage below the investigated soil layer was not measured during the experiment. Consequently, the amount of water retained within the EMI investigation depth could not be rigorously compared with the applied irrigation volume. For these reasons, we did not attempt a water-mass-balance closure. Instead, the stability of the non-irrigated control area was used to assess the absence of unexpected significant temporal changes, while the dynamic part of the inversion was evaluated through the agreement between measured and modelled σa values

L363 Regarding the negligible lateral fluxes, I don't think horizontal variability is needed to have lateral fluxes, right? For example, the two layers with different Ks could already induce a preferential lateral flow? Consider rephrasing to support your assumption.

We rewrote the section. In principle, we agree that horizontal variability is not really needed to have lateral fluxes. Actually, a contrast in hydraulic conductivity between the two layers in the investigated soil profile may induce some lateral redistribution above their interface. However, it was assumed that the difference between the Ks values of the two layers may affect primarily vertical infiltration across the interface. In the text, we now recognize that local lateral redistribution cannot be completely excluded and that, consequently, the 1D approach is a modelling assumption appropriate to the available observations, rather than as a direct consequence of soil texture.

L378 Regarding the bottom boundary condition, if I understand correctly, you estimated the negative hydraulic head based on the height difference between water table and profile boundary condition? This is ok for the positive heads, but I am not sure this is correct for the negative ones...

In principle, the reviewer is right. Nevertheless, in the specific study under review, the water table was very close to the bottom of the simulate soil profile (about 20-30 cm). In the case of negative pressure head at the bottom boundary (water table deeper than the bottom of the simulated soil profile), we still assumed a hydrostatic equilibrium with the water table. This was justified by the shallow maximum depth to groundwater (about 120-130 cm) and by the value of the α parameter, which would correspond to a characteristic capillary length (α^{-1}) of about 15-20 cm and indicate that the bottom of the simulated profile (100 cm), very close to the water table, was likely affected by the capillary fringe.
This was clarified in the revised text

L405 The comparison between measurements (samples and EMI) and simulations that used the respective initial conditions from samples and EMI (plus SHP and described boundary conditions) makes sense to me. The scaling based on the averages at the different depths is not convincing or straight forward though. It seems a convoluted solution that would address but also mix both the footprint – sensitivity part (target here) and many other aspects, including, for example, an offset (wrong measurements that would need calibration) in one or more of the six EMI datasets, or even calibration (conductivity – moisture content) at different depths.

The comment by the reviewer 1 indicates that the text about the target of the scaling procedure in the manuscript may be a bit misleading in its actual shape. Actually, the scaling was not intended to combine the different measurement footprints of soil sampling and EMI. The aim of the paper is just avoiding confusion between the two technique footprints.
The sample- and EMI-based datasets provide different representations of the same infiltration process because of their different observation volumes and spatial sensitivities. In this sense, the different temporal dynamics observed by each technique, coming from their different spatial sensitivities and observation volumes, is completely preserved.
Only when one looks for converting the soil water-content representation associated with one observation footprint into that associated with the other one representation in the other, one needs to scale the initial water content profiles and the water retention curves. This was clarified better in the text of the revised manuscript:

"It is worth to note here that nowhere this procedure combines the different measurement footprints of soil sampling and EMI. The sample- and EMI-based datasets provide different representations of the same infiltration process because of their different observation volumes and spatial sensitivities. In this sense, the different temporal dynamics observed by each technique, coming from their different spatial sensitivities and observation volumes, is completely preserved. Only when one looks for converting the soil water-content representation associated with one observation footprint into that associated with the other, one needs to scale the initial water content profiles and the water retention curves".

We also clarified that: "In the procedure presented in this paper, the scaling of the initial conditions using a spatially averaged scaling factor that varies with depth deserves further analysis. Differences between the depth-averaged sample- and EMI-based water contents may result not only from their different measurement footprints and spatial sensitivities, but also from uncertainties in the inversion of the EMI data, possible depth dependence of the σ_b–θ calibration relationship, and errors associated with this calibration".

Fig3 I like and agree on showing the apparent conductivity values. The figure is intuitive and summarizes the EMI acquired datasets. My two doubts are:
1) I cannot appreciate significant changes after the first three hours. Possibly only the HCP32 shows an increase from 3h to 5h, but other shallow datasets (VCP) do not support this, right?
2) the deep HCP datasets do not show any particular variation, despite the shallow dynamics, do you suspect some balancing mechanism in the deeper region?
Therefore, I do not particularly agree with your interpretation at L424 on the evident increase in both VCP and HCP.

The same point was also raised by the reviewer 2. Accordingly, this response will be useful to clarify remarks from both the reviewers. We recognize that our original interpretation of the σa evolution was misleading, partly because of a reversed interpretation of the VCP and HCP meaning. We also agree that the relatively modest temporal evolution of the apparent electrical conductivity deserves further explanation. The temporal increase in the values during irrigation was relatively modest, especially for the HCP ρ71 and HCP ρ118. An evident increase in σ_a occurs between t = 0 and t = 3 h in the VCP datasets and in the shallowest HCP dataset (HCP ρ32). A further increase is locally visible mainly in HCP ρ32 between t = 3 and t = 5 h, whereas the other datasets do not show a significant temporal increase. From t = 5 to t = 9 h, the σ_a patterns remain broadly stable. This does not imply that infiltration was limited. In general, the overall observed σ_a patterns may result from a combination of effects related to the integrated response of the EMI sensor and those related to the specific physical system under study. During the first stages of irrigation, the increase in soil water content was mainly restricted to the upper part of the profile, and its electrical response was therefore averaged with that of the underlying soil, which remained practically unchanged. This depth-averaging effect is inherent in EMI measurements and is expected to be more pronounced for the deeper-sensitive HCP configurations, while VCP is relatively more responsive to near-surface changes. In addition, the sandy soil promoted rapid vertical infiltration and redistribution, whereas the relatively low conductivity of the irrigation water likely produced only a moderate electrical contrast between the initial and irrigated soil conditions.

In the manuscript, we have revised the discussion of Figure 3 to account for all the discussion above.

L426 values of lambda between 1 and 5, consider commenting on how robust and general (or case-specific) this choice could be.

We agree that the optimal value of λ cannot be considered universal. The interval between 1 and 5 should not be interpreted as generally applicable, but rather as appropriate for the present dataset, considering the measurement configuration, the expected conductivity contrasts and the noise level. This has been clarified in the revise manuscript

Fig4 The infiltration looks ok, apart from the above concerns related to the absence of significant data changes in the deeper part (>= 50 cm depth) and later times (> 3 h), which the *time-lapse inversion seems to highlight? Any estimates of the misfit for the HCP71 and HCP118 with the final – 9h model?*

We thank the reviewer for this valuable suggestion. We calculated the root mean square error (RMSE) between the measured and forward-modelled apparent electrical conductivities for each coil configuration. For the final model (9 h), the RMSE values were 4.79 mS m^{-1} and 1.31 mS m^{-1} for the HCP71 and HCP118 configurations, respectively, while the overall RMSE considering all six configurations was 3.84 mS m^{-1}.
In any case, it should be noted that the conductivity changes occurring below approximately 50 cm after 3 hours are less evident in the measured apparent conductivities than in the inverted conductivity sections. This is expected because the inversion simultaneously uses all six coil configurations and estimates the conductivity distribution that best explains the complete dataset while satisfying the imposed regularization constraints. Consequently, relatively small conductivity variations distributed over depth may become visible in the inverted σb sections even when they are only weakly expressed in the individual σa measurements.
The configuration-specific RMSE values also provide useful insight into this behaviour. The largest residuals were observed for the VCP32 and HCP71 configurations, whereas HCP118 exhibited one of the lowest RMSE values. We believe that this behaviour reflects the different sensitivity functions of the coil configurations rather than a deterioration of the inversion quality. In fact, VCP32 and HCP71 are mainly sensitive to the shallow and intermediate portions of the soil profile, where the infiltration experiment generated the strongest and most rapidly evolving conductivity gradients associated with the advancing wetting front. Such gradients are inherently more difficult to reproduce with a one-dimensional layered conductivity model under the adopted regularization. Conversely, HCP118 integrates the electrical response over a larger and deeper soil volume, where conductivity changes remained comparatively smoother and less pronounced, resulting in a substantially lower misfit.
Finally, the progressive increase in the VCP32 residuals during the infiltration experiment, rather than a nearly constant error throughout the monitoring period, further suggests that the larger residuals mainly reflect the increasing complexity of the conductivity distribution produced by infiltration rather than instrumental effects or deficiencies of the inversion procedure.

We have added this discussion to the revised manuscript.

L441 "Figure 5 shows the soil-specific σb-θ calibration obtained by the TDR readings, as described in the section 2.2. Building of the site-specific ρb-θ relationship." Did the EMI conversion use this calibration or a laboratory one from the samples at different depths?

This was already said in the original manuscript at lines 443-444.

In any case, in the revised manuscript we now state:

"Figure 6 shows the distributions of θ coming from the conversion of the EMI estimations of σb, based on the site-specific σb-θ relationship shown in figure 5".

This was also clarified in the section 2.2. **Building of the site-specific σb-θ relationship,** where we state:
"All the TDR-based θ and σb data taken during all the infiltration test duration were used to build an independent soil-specific calibration relationship to be used for converting the σb coming from the EMI inversion to as many θ values" (already said in the original manuscript, lines 271-273)"

L460 The spatial part of this paragraph is generally accepted in my opinion, I surely agree with the authors on this. What I think is missing is that it is not just a spatial aspect, but also the specific conductivity model; meaning that talking about different spatial scales and resolutions can be misleading in my opinion, as the EMI sensitivity varies, and could be relatively sensitive to a thin layer or object, as long as it is very conductive. To me, and the authors likely agree, the two things cannot be separated. This, however, hinders a bit the generalization of these comparisons, and the idea of reasoning in terms of spatial frequency, scaling, and low-pass filters. Consider addressing this aspect, possibly moving the paragraph to the discussion.

We fully agree with the reviewer. The comparison between point-scale and EMI-derived observations cannot be explained only in terms of different support volumes. The EMI response also depends on the conductivity distribution within the investigated volume because each conductivity element is weighted by the sensor sensitivity function. We have revised the text to clarify that these two aspects are intrinsically linked and should not be considered independently.

L471 – L483 No EMI here, correct? It is auger-based simulations (with SHP from TIM) compared to the successive auger measurements. If so, consider clarifying that in the caption of figure 7. Figure 8 is already good.

The caption of figure 7 in the original manuscript was "Figure 7: Six of the twelve θ profiles observed by auger sampling (symbols) and simulated by FLOWS (lines) with the SHP measured in the two soil horizons by infiltrometer and reported in table 2. The profiles refer to three different times (1h, 5h and 9h) from the beginning of the infiltration experiment"

We now added " ...with SHP measured by TIM...."

L490 - L497 Trying to summarize the different parts: the EMI-based was calibrated with the TDR measurements at the surface; then, their moisture estimates were scaled with the auger values (averages by depths); and the derived saturated water content (SHP for the simulation) was scaled with this additional factor, while the other the hydraulic parameters were taken and fixed from the tension infiltrometer measurements? It may be me, but this is hard to follow, also on its physical or methodological motivations. What is the physical or methodological meaning of the scaling factor for the saturated water content?

The section was completely rewritten to make the discussion as clearer as possible. However, we would like to clarify here that the EMI-derived water-content estimates were not scaled using the auger measurements. This was emphasised in the text. The revised text now explicates step by step the procedure used to carry out the simulations to reproduce the EMI-based θ profiles. We report here much of the text included in the revised manuscript:

"The simulations were carried out by adopting as the initial condition for each of the twelve profiles, the θ profile estimated by the EMI readings at t = 0. The boundary conditions were again set at the rainfall rate applied by irrigation in the different parts of the transect. No scaling of the EMI-based water contents using the auger-based water contents was done. The EMI-based water content profiles were retained as a self-consistent representation of the infiltration process and were considered as physically plausible as those coming from auger, from which it differs only because of the different observation volume and spatial sensitivity of the technique.
To carry out these EMI-based simulations, the SHP obtained by TIM were scaled by simply reducing the saturated water content by a factor varying in the range 0.7-0.9, to make the SHP internally consistent with the observation scale of the EMI measurements. To identify the value of that factor, three values were used, 0.7, 0.8 and 0.9, and the value 0.8 was found to be the most suitable choice as it led to simulated water contents close to those estimated by EMI, as shown in figure 9. Thus, the scaling factor value was selected empirically. This scaling factor should be regarded as a dataset-specific scale-conversion factor, rather than as a correction implying that either the auger- or EMI-based water-content representation was incorrect. "

Fig9 1) I think that the effect of the scaling factor could be added to all the subplots, it remains readable in my opinion. 2) Would it make sense to add the auger values here for comparison (when and where available)?

Yes, done

In general, EMI estimates and simulations are similar in terms of moisture range, but the profiles – trends differ on their specific features (changes both vertically and in time). I do not agree much on the "very similar representation of the wetting front propagation" (L500).

Yes, clarified in the text, where we stated:

"The simulated θ distributions reproduce quite satisfactorily the measured profiles in terms of moisture range, even if the two profiles show some differences in their trends, which are in any case limited to quite narrow water content bounds. In any case, they provide comparable representations of the wetting front propagation."

Fig11 OK, this is an important figure, the comparison looks good. What I notice though is that the results introduce a significant lateral smoothing, right? Compared for example with auger values in the left panel of figure 8. My guess would be that lateral homogeneity of the SHP and of the EMI scaling can explain this. However, some assumptions seem necessary to define some statistically robust SHP and scaling factors, or even the TDR - EMI calibration itself. Therefore, my doubts are not on the specific choices but on the general contradictions of the approach, for example stressing the vertical smoothing of the inversion, but then introducing general scaling, calibration, and SHP. While I am not an expert, I wonder whether a 2D model with a more classical data assimilation approach would represent a less arbitrary and more *data-based work flow. The authors did not introduce or discuss the existing literature on data assimilation, and related ERT or hydrogeophysics modeling studies. This weakens the motivation and general value of this approach in my opinion.*

We agree that the simulations based on the EMI-derived initial conditions show less lateral variability than the auger-derived distributions. We have made more explicit this in the revised. Related to this, we also shortly introduced the possibility of exploring a 2D data-assimilation framework to evaluate whether the local variability observed by auger sampling can be retained while remaining consistent with the EMI observations.
The added text is:
“The two simulations reproduce a comparable infiltration dynamic at the transect scale, with a similar propagation of the wetting front. However, the latter appears laterally smoother than the point-scale auger observations shown in the left panel of figure 8. For the auger-based simulations (left panel of figure 11), this smoothing is coherent with the spatial assumption of SHP considered horizontally homogeneous within each soil horizon. For the EMI-based simulations (right panel in figure 11) the smoothing also reflects the use of a single scaling factor for initial conditions at each depth for all the profiles along the transect. Overall, these assumptions reduced the impact of the local-scale variability of the SHP in both the simulations and of the local-scale variability of the initial water-content distributions in the EMI-based simulation. In any case, the similar behaviour obtained under common SHP and boundary conditions along the transect suggests that the remaining differences between the initial water content distributions (variable along the transect for auger-based and constant for EMI-based simulations) have only limited effects on the simulated transect scale infiltration pattern.
A spatially distributed SHP dataset (not available in this study) or a 2D data-assimilation framework (Camporese et al., 2011; 2025) would be required to evaluate whether the local variability observed by auger sampling can be retained while remaining consistent with the EMI observations. However, although theoretically feasible with the available EMI, auger, TIM-derived SHP, and irrigation data, a 2D data-assimilation approach would require spatial covariance models for the SHPs, error models for the EMI dataset, as well as an EMI observation operator linking simulated water-content distributions to the measured σ_a. This would represent a substantial methodological development and is beyond the scope of this paper but offers but offers considerable potential for future research. “

Finally, the overlap with the work by Dragonetti et al. (2022) is also concerning. The addition of the EMI scaling part, with the concerns outlined above, does not seem to me a substantial or sufficient methodological difference.

Yes, we agree. In the revised text in the first part of the Conclusions we write:

“The study builds on the approach proposed by Dragonetti et al. (2022) but extends it to a transect-scale infiltration experiment in sandy soil. The methodological contribution is in the use of hydrological modelling as a common process-based framework to compare the time evolution of water content distributions provided by point-scale measurements and EMI. The study focuses on the to the effects of the different measurement support volumes, the depth-dependent scaling factor of the initial conditions and the assumption of horizontally homogeneous SHP within each soil horizon. Specifically, it evaluates whether technique-specific initial conditions and water-retention curves differing mainly in their saturated water contents can reproduce the temporal evolution observed by both techniques, without requiring an empirical rescaling of the EMI-derived water contents at each time step. The main message of the paper is that integrating EMI measurements with hydrological modelling may provide a robust physically based framework for interpreting what an EMI sensor observes during a water infiltration experiment, as well as to estimate soil hydraulic properties noninvasively at the field scale.”