GMD egusphere-2026-4305 – Author responses to referees

Dear Dr Polcher,

We thank you for your thorough and challenging review. There is no need to apologise for being demanding: we agree that an operational system whose data will be widely used should be held to the highest scientific standards.

We agree in particular that the assumptions and limitations of the system must be stated explicitly, and the main change to the manuscript will be a new subsection documenting them. We will also clarify the coupling between ecLand and CaMa-Flood, extend the discussion of the global discharge evaluation, provide more context for the regional cases, and broaden the references to related modelling systems.

Our planned responses to each comment follow below.

Major comments

M1 – "Physically consistent" and the pristine-continent assumption (L46–48)

Referee comment. *With the sentence starting in line 46 ("The comprehensive and Physically consistent …") two promises are made which are not realised in the paper. To demonstrate that the system is physically consistent the various assumptions made have to be spelled out so that the reader can determine which relations can be expected between the various variables made available. Implicitly throughout this paper the assumption is made that the continental water cycle is pristine, i.e. without human intervention. If that assumption would be spelled out, the authors would not claim in the second part of that sentence that a "realistic representation of freshwater fluxes from river" is achieved. Many studies have shown that the major driver for river discharge changes over the last century in Europe has been human intervention (Vicente-Serrano et al. 2019 or Collignan et al. 2025). The same can be assumed for other mid-latitude continents.*

Author response. We agree. The current operational system does not represent human-induced changes to the hydrological cycle such as water use (irrigation, domestic or industrial abstraction) and reservoir operation. The CaMa-Flood reservoir operation scheme is not active in the operational configuration. By "physically consistent" we meant that the land and river components share the same atmospheric forcing and land state. Runoff is transferred between grids without loss or gain of mass. We did not intend to claim a complete representation of the continental water cycle, and we will make this distinction explicit. We also agree that the second part of the sentence is not yet realised in the operational system: simulated discharge is not yet passed to the ocean model, which still uses climatological river inflow (L748–749), and in the one-way land-to-river configuration, river and floodplain water neither evaporates nor feeds back on land–atmosphere exchanges. For the same reason, rivers do not yet contribute to the land surface energy exchanges referred to in the first part of the sentence. We will revise the sentence and related statements accordingly.

Planned changes to the manuscript:

- Abstract (L19), Introduction (L46–48, L55–59) and Conclusions (L760, L789, L796): define "physically consistent" as described above, and remove the claims of "realistic" freshwater fluxes to the ocean and atmosphere, which are not yet represented operationally.

- New Sect. 3.3.6 “Assumptions and limitations”: state that simulated discharge is naturalised flow, with no reservoir regulation or human water use, citing Vicente-Serrano et al. (2019) and Collignan et al. (2025).

M2 – Placing the system within the state of the art

Referee comment. *Placing the ECMWF operational system within the state of the art achieved by the community is simply a matter of scientific honesty. In the references provided at the end of the paper, less than 30% are not from the co-authors or their institutions ! Other Earth system models which strive to provide a “physically consistent treatment of the water cycle” exist but are not referenced : JULES, SURFEX, CLM, WRF-HMS, ORCHIDEE, …. They do not all aim at the medium range weather forecast but that does not make them less ambitious in achieving physical consistency and realism. To maintain our intellectual rigour we have to keep in mind that we are standing on the shoulders of giants !*

Author response. We acknowledge the imbalance in the reference list and agree that the work of the wider community should be better recognised in the manuscript. We will expand the introduction with a concise overview of how river routing and floodplain dynamics are represented in other land surface and Earth system models used for weather and climate prediction and position the ECMWF system relative to these.

Planned changes to the manuscript:

- Introduction (around L58): add a paragraph on how river routing and floodplain dynamics are represented in other land surface and Earth system models, covering:
  - JULES (Best et al., 2011 and Clark et al., 2011);
  - ISBA-CTRIP/SURFEX (Decharme et al., 2019);
  - CLM/CTSM with MOSART (Lawrence et al., 2019);
  - ORCHIDEE (Guimberteau et al., 2012; Schrapffer et al., 2023);
  - WRF-Hydro and WRF-HMS (Gochis et al., 2015 and Wagner et al., 2016);
  - COSMO-CLM-ParFlow (Shrestha et al., 2014);
  - MIROC with CaMa-Flood (Nitta et al., 2020).

  We will also cite the relevant GEWEX/WCRP activities (as summarised in Stephens et al., 2023; see m1).

M3 – Awareness for data users and machine-learning applications

Referee comment. *Because of the position of ECMWF within the community, it is unavoidable that the data produced by this system will be evaluated in different studies and used to train machine learning models. Who will remember that in this system the assumption of pristine continental surfaces has been made if it is not spelled out here ? What will ML models do if they ingest data from this system and actual observations of the hydrological system ? The user community has to be made aware of the risk of incompatibilities.*

Author response. We thank the referee for raising this point. We agree that, as the referee notes, the river outputs of the IFS are likely to be used widely, including to train and evaluate data-driven models. Users therefore need to know what the simulated variables represent and where they are expected to differ from observations.

In the current system, simulated discharge is naturalised flow: reservoir operation, water abstraction and irrigation are not represented (see M1). In basins that are strongly regulated or heavily used,

simulated discharge will therefore differ systematically from observed discharge in magnitude, seasonality and variability. A data-driven model trained on a mix of these simulations and observations could learn inconsistent relationships between, for example, precipitation, soil moisture and discharge. Similarly, users should not expect the water budget to close exactly in outputs from configurations with data assimilation, because the analysis increments are not conserved (see M4).

We will address this both in the paper and in the user documentation of the river products.

Planned changes to the manuscript:

- New Sect. 3.3.6 (see M1): add a short "guidance for data users" paragraph, noting the risk of combining these data with observed, human-influenced discharge when training data-driven models.
- Conclusions and Code and data availability: repeat the main points of this guidance
- Introduction (L55–59): revise the statements presenting Earth system model output as physically consistent training data, so that they refer to the stated assumptions and limitations.

M4 – Conservation laws and coupling terms (Sects. 3.1–3.2), lakes and energy

Referee comment. *In the description of ecLand (section 3.1) and CaMa-Flood (3.2) nowhere are the chosen conservation laws defined and which terms achieve the coupling. Not in Bousetta et al. 2021 either. ecLand output is known to neglects the water convergence in the lake component (river and groundwater inflows minus water outflow) and thus breaking the catchment level water conservation (Polcher et al. 2026). The energy convergence through lateral water advection is also neglected (Can be significant for some lakes !) At least this assumption is consistent between ecLand and CaMa-Flood ! It is essential to write these assumptions down so that the reader can evaluate which conservation principles can be assumed when using the data. I know all these principles get broken during the data assimilation phase (section 4.6.2) but respecting them in a free running configuration of the system is part of what I would call a "physically consistent treatment of the water cycle".*

Author response. We agree that the conservation assumptions and the coupling terms must be stated explicitly. In the revised manuscript, we will summarise which conservation properties hold in each configuration and include a table of the variables exchanged (see M6).

For resolved lakes, ecLand uses the FLake model with a constant global lake bathymetry (Choulga et al., 2019) and no prognostic water storage (Balsamo et al., 2012). The precipitation falling onto lake is immediately transferred as total runoff (L363–364). Lateral inflow and outflow and a full lake water balance are therefore not yet part of the ecLand lake representation. Lake evaporation can be however withdrawn from the inundated water area. Over resolved lakes, since lakes are not dynamically evolving, the evaporation flux draws from an unlimited water reservoir, therefore the water closure can be achieved only accounting for these explicitly simplifications at the catchment level. Lateral heat advection is also not represented in the ecLand lake model, and CaMa-Flood does not yet simulate water temperature in our configuration. Ongoing works to add water temperature are pursued within the CaMa-Flood community (see Tokuda et al., 2019). In CaMa-Flood, lakes are part of the river and floodplain network: they receive lateral inflow and release outflow, but in the operational configuration no vertical exchanges with the atmosphere is currently active.

In the operational configuration (2-way atmosphere–land, 1-way land–river), river and floodplain water does not feedback on the land: it neither evaporates nor infiltrates into the soil. In free-running mode, the water budget of each land column is closed for the soil, snow and vegetation tiles, but not over resolved lakes, for the reason given above. Closure at catchment level between the two components is therefore not guaranteed. Differences between the land–sea masks of the two components also contribute (see M5). As the referee notes, column closure does not hold when data assimilation is active: soil moisture increments add or remove water without a compensating flux (Sect. 4.6.2). We will state this explicitly.

The required level of budget closure depends on the modelling objectives. For operational medium-range weather and flood forecasting, the current system is designed to conserve water within the river routing model and to consistently represent the processes most relevant to forecast skill. Strict closure of the entire coupled water budget, including small coastal catchments and land–sea mask mismatches, is not necessarily a requirement for these applications. However, it becomes more important at longer time scales, for example, in seasonal forecasts and long climate simulations, and for the coupling of river discharge to the ocean (Sect. 4.6.4). We will state these limitations explicitly and include improved overall budget closure among the future developments.

Planned changes to the manuscript:

- Sect. 3.3.1: add a table listing, for each configuration (operational, 2-way, 1-way), the variables exchanged, their direction, units and coupling interval.
- Sect. 3.3.1 (L363–364): state the lake water and energy assumptions, citing Polcher et al. (2026).
- New Sect. 3.3.6 (see M1):
  - summarise which conservation properties hold, and that they apply to free-running configurations only, since data assimilation increments are not conserved;
  - state the level of budget closure required by each type of application;
  - list the known limitations (lakes, land–sea mask mismatches; see M5);
  - identify their improvement as future work for seasonal and climate applications and for the coupling to the ocean.

M5 – Mass-conservative regridding, land/sea masks and floodplain evaporation feedback

Why would a "A mass-conservative re-gridding" between ecLand and CaMa-Flood be important if the water mass conservation equations are different or inconsistent in both models ? Just a few cases to illustrate the criticality of this point :

a) *ecLand has a lake water continuity equation without the horizontal divergence term while CaMa-Flood has that term but not the vertical fluxes, how is this made physically consistent ? The interpolation chosen does not change anything to the problem as long as the correct fluxes and state variables are not exchanged.*
b) *ecLand and CaMa-Flood run on different grids and thus do not have the same land/sea mask. How are catchments defined (The basic hydrological closed system on which water conservation can be assume.) if some points of CaMa-Flood are ocean for IFS ? This issue also needs to be addressed in section 4.6.4.*
c) *CaMaFlood computes "water level and inundation extent" how does that inform the continental evaporation in ecLand or ensures that the soil moisture below is filled to saturation first ? Some models have implemented solutions to this (Guimberteau et al. 2012, Decharme et al. 2019, Schrapffer et al. 2023, ...)*

If such points are not clearly defined it is not obvious that the system represents a physically consistent water cycle.

Author response

a) Mass-conservative regridding is necessary but not sufficient: it ensures that the runoff volume generated by ecLand reaches CaMa-Flood without loss or gain, but we agree that it does not resolve the differences between the lake representations of the two components. These differences stem from the design of the ecLand lake scheme, which preceded the river routing. The integration of CaMa-Flood and the availability of inundated areas can make use of the energy exchange between lakes and the atmosphere (Balsamo et al., 2012). This paper documents the integration of river routing into the IFS. Revising the lake scheme would be a separate development, which we will mention among possible future developments. We will clarify this in the revised manuscript (see M4).

b) The area found only in the CaMa-Flood domain (1.76 million km², 1.1%) is mainly mismatches along the coastlines, especially in areas with many small islands. ecLand produces no soil runoff there. Conversely, 0.42 million km² (0.28%) is land/lake in the IFS but outside the CaMa-Flood domain, which mainly consists of estuaries such as the Sea of Azov and Lake Maracaibo. Precipitation falling on the resolved lakes is passed to CaMa-Flood as surface runoff, but lake evaporation is not removed, so the net water balance of the lakes is currently not represented. Inflow from upstream is routed through these cells as ordinary river/floodplain storage, with no explicit lake dynamics. This affects the volume and timing of discharge from rivers draining these lakes, such as the St Lawrence.

   Runoff generated outside the CaMa-Flood domain is discarded, because it falls outside the runoff-to-catchment input matrix. This mainly affects small coastal catchments and has a limited effect on forecasts for most inland rivers, but they matter for basins with large lakes, for coastal applications and for long-term budget closure (see M4). We will describe them in Sect. 3.3.1 and extend Sect. 4.6.4 accordingly.

c) In the operational configuration, inundation does not feedback to ecLand: water level and inundation extent are diagnostic outputs. In the experimental 2-way configuration, the flooded fraction is treated as open water that evaporates and withdraws water from river storage, so water is conserved between the two components (L367–370). Infiltration of floodplain water into the soil, i.e. filling the soil to saturation before water remains on the surface, is not yet represented. Approaches implemented in ORCHIDEE (Guimberteau et al., 2012; Schrapffer et al., 2023) and ISBA-CTRIP (Decharme et al., 2019) provide useful references for future developments of the 2-way configuration.

Planned changes to the manuscript:

- Sect. 3.3.1 (L364–366): clarify that mass-conservative regridding ensures that runoff volume is transferred without loss or gain, but does not reconcile the lake representations of ecLand and CaMa-Flood (see M4).
- Sect. 3.3.1: describe how the land–sea masks of the two components differ and how this is handled: catchment definition, the large lakes treated as sea by the IFS, coastal IFS land outside the CaMa-Flood domain, and Antarctica. Include the overlap statistics.
- Sect. 3.3.1 (L367–371): state that, in the operational configuration, water level and inundation extent are diagnostic outputs with no feedback to ecLand. In the 2-way configuration, evaporation from the flooded fraction is withdrawn from river storage, and infiltration of floodplain water into the soil is not yet represented.

- Sect. 4.6.4: discuss the effect of the land–sea mask mismatches on freshwater input to the ocean and on budget closure.
- New Sect. 3.3.6 (see M1): list the land–sea mask mismatches and the lake representation among the known limitations relevant to data users.
- Conclusions: mention the lake water balance and floodplain–soil interaction among the possible areas for future development, citing Guimberteau et al. (2012), Decharme et al. (2019) and Schrapffer et al. (2023).

M6 – Figure 3: name the fluxes and state variables

Referee comment. *Figure 3 is difficult to follow without further information. I guess the arrows between ecLand and the atmosphere is some sort of implementation of the PILPS interface. What do the arrows from the atmosphere to the rivers represent ? Is this rainfall over lakes and floodplains ? If so, where can this water evaporate and how is consistency ensured ? Some groups have properly describe these coupling issues for their models : Shrestha et al. 2014 for COSMO-CLM-ParFlow, Wagner et al. 2016 for WRF-HMS, Hallouin et al. 2022 for JULES, ...Your graphics are nicer than those in the cited papers, but you do not name the fluxes or state variables which are exchanged. That information is not in the legend of the figure or the manuscript.*

Author response. We agree that the exchanged fluxes and state variables should be named. Fig. 3 was intended as a guide to the coupling configurations described in Sect. 3.3.1, with a focus on the integration of CaMa-Flood, and we will make this purpose explicit. The arrows between ecLand and the atmosphere represent the standard IFS land–atmosphere interface, which is unchanged by this work and is described elsewhere (ECMWF, 2024; Boussetta et al., 2021). We will add a reference to it in the caption.

The arrow from the atmosphere to the rivers is misleading, and it will be removed.

In the operational configuration, evaporation from river and floodplain surfaces is not represented, so this water is routed downstream. In the experimental 2-way configuration, evaporation from the flooded fraction is withdrawn from the river and floodplain storage (see M5c). The land-to-river arrows represent surface and subsurface runoff, and in the 2-way configuration, the river-to-land arrow represents the flooded fraction and the infiltration of floodplain water into the soil.

Planned changes to the manuscript:

- Fig. 3: label the river-related arrows with the exchanged fluxes and state variables, consistent with the coupling table in Sect. 3.3.1 (see M4).
- Fig. 3 caption (L387–393): state that the figure focuses on the river coupling, refer to the documentation of the land–atmosphere interface, and correct the description of the dashed arrow.
- Sect. 3.3.1: cite Shrestha et al. (2014), Wagner et al. (2016) and Hallouin et al. (2022) as examples of documented coupling interfaces.

M7 – Global evaluation of river discharge (Fig. 6)

Referee comment. *The global evaluation of river discharge is sketchy. Figure 6 shows the flow direction of the major rivers as imposed on CaMa-Flood by the MERIT Digital Elevation Model. The insert is much more interesting. If I read correctly, 90% of the GRDC stations have a KGE below 0.6. A KGE of 0.41 means that the model predictions are not better than climatology and thus that systematic biases exit. This is the case for more than 70% of all stations ! Why does the model*

perform so poorly ? This should be discussed and linked to the assumptions made or propagation of errors within the system. An essential information for all users of the output from this system !

Author response. We thank the referee for this careful reading. The threshold in the inset is the mean-flow benchmark KGE = 1 − √2 ≈ −0.41 (Knoben et al., 2019): stations above −0.41 perform better than the observed mean flow. In the preprint, a line break separates the minus sign from the number in the caption (L556), so it reads as 0.41. We apologise for the confusion and will correct the caption and state the benchmark explicitly in the text.

Firstly, since in the meantime the 2025 WMO State of Global Water Resources Report, issued this year, has also become available, our analysis in the revised manuscript will focus on the reference simulation provided for this 2025 report.

Of the 1492 GRDC stations available in GloFAS, we used the 898 stations classified as having no major reservoir or lake influence in the GloFAS (version 3.1) station-quality metadata. Of these, 708 stations had at least 10 years of observations in 1991–2020 and were included in the verification.

The station-quality classification is not based on fixed thresholds of reservoir or lake impact indices. For each station, we jointly inspected the observed discharge time series, the upstream reservoirs and lakes from the GRanD v1.3 (Lehner et al., 2011) and GLWD v1 (Lehner and Döll, 2004) databases, and derived indices of the estimated reservoir and lake impact on observed and simulated discharge. A station was considered affected when the observed hydrograph showed variability that could not be explained by natural hydrological behaviour (e.g. abrupt step changes, damped peaks or regular release patterns) and this was consistent with the presence of upstream reservoirs or lakes. This is therefore an expert-based, partly subjective decision. Other human influences, such as urbanisation, irrigation and water abstraction, were largely not considered because they are difficult to quantify, so some of the selected stations may still be affected by them. In the revised manuscript, we will add a short paragraph to Section 4.1 describing the station selection, including the number of stations at each step, the expert-based screening for reservoir and lake influence, and its limitations.

Other human influences, such as urbanisation, irrigation and water abstraction, were largely not considered because they are difficult to quantify. Some of the remaining stations may therefore still be affected, which should be kept in mind when interpreting biases.

With the 2025 reference simulation, 80% of the 708 stations show skill relative to the mean-flow benchmark (KGE > −0.41), with a median KGE of 0.42, while 74% have a KGE below 0.6.

Nevertheless, as the referee notes, a large fraction of stations have a KGE below 0.6, and we agree that the causes must be discussed. The reference simulation is not calibrated, and its skill is limited by several factors that we will discuss in the revised manuscript:

- precipitation errors in the ERA5 forcing, particularly in data-sparse regions;
- the partitioning between surface and subsurface runoff in ecLand;
- the absence of a groundwater reservoir in the system: all runoff, including subsurface runoff, is routed directly to the river channel, which can lead to flows that respond too quickly to rainfall and to underestimated baseflow in dry periods;
- river parameters that are derived from global relationships or set to default values (channel depth, riverbed roughness; L441–443);
- remaining human influences, such as irrigation and abstraction.

A low KGE does not necessarily indicate a systematic bias; it can also result from timing or variability errors. Decomposing the KGE into its components (correlation, bias ratio and variability ratio) will

distinguish errors in timing from errors in water volume and variability, and give a first indication of their likely causes.

Planned changes to the manuscript:

- Verification results will be based on the reference simulation for the 2025 WMO State of Global Water Resources report.
- Fig. 6 caption (L554–557): keep the −0.41 benchmark on one line, and state the station selection criterion and the number of stations.
- Fig. 6: give more prominence to the KGE evaluation, as a separate figure or a larger panel showing the KGE components.
- Sect. 4.1:
  - describe the station selection, and the quality checking procedure;
  - state that the reference simulation is not calibrated and was run without the groundwater reservoir;
  - report the share of stations with skill relative to the mean-flow benchmark and the median KGE;
  - show the KGE and its components (correlation, bias ratio, variability ratio) at each station, mapped and summarised by region or climate zone;
  - discuss the likely causes of low skill listed above and their link to the system assumptions (see M1).

M8 – Scientific level of the regional applications (Sects. 4.2–4.4)

Referee comment. *The 3 regional applications shown (Australia, Zambia and Pakistan) are perfect for an ECMWF brochure but do they reach the level of a scientific publication ? The base line climatology of the model on which these extreme events are predicted is not evaluated. The quality of the reproduced hydrological events is not quantified (is the extent and timing of expansion and retreat of the flooded are of Pakistan correct ?). There is no reference either to a study which analyses this. A trained eye notes that in the Australian and Zambian cases, the simulated discharge peaks (or water levels) are one to two days earlier than the observations. Is this an error in forecasted precipitation, is this linked to the spin-up of the river model, is the infiltration/runoff partition in ecLand correct, are there parameters of the aquifer to be adjusted ? Some working hypothesis and improvement path should be presented to the future user of this operation forecasting system.*

Author response. We thank the referee for this comment. We agree that, as written, Sects. 4.2–4.4 could be read as an evaluation of model performance, which was not our intention. These sections illustrate the range of applications now enabled by the coupling of river routing in the IFS (e.g. discharge and water-level forecasts, assessment of flood inundation extent), and the cases were chosen to show the breadth of capabilities across different hydro-climatic settings, not to quantify forecast skill. This is a first technical implementation, and several parameterisations are still being developed (e.g. runoff generation at high resolution, Sect. 4.6.3). In addition, the CaMa-Flood river parameters, such as channel depth, bank height and riverbed roughness, are currently derived from global relationships or set to default values and have not yet been tuned; their refinement, using altimetry and in situ observations where available, is part of the planned development of the system. A systematic evaluation of the hydrological climatology and forecast skill of the operational configuration would therefore be premature here; it will be reported in future work. The baseline climatology of the long-term simulation is, however, evaluated globally in the extended Sect. 4.1 (see M7).

Nevertheless, we agree that the illustrative cases benefit from quantitative context. For the Luangwa case, we have compared the simulated water levels at the three resolutions with the 29 DAHITI satellite-altimetry observations available for 2023 (Schwatke et al., 2015). Satellite altimetry is itself uncertain, particularly over medium-sized rivers such as the Luangwa, where the radar signal can be affected by the surrounding land and the river banks. DAHITI provides an error estimate for each observation (e.g. ±0.15 m for the peak on 19 March 2023; L614–615). All three simulations reproduce the temporal variations well (Pearson $r = 0.95$, $p < 0.001$), but underestimate the observed water levels, with mean absolute errors of 2.8 m (6 arcmin), 3.0 m (1 arcmin) and 3.7 m (3 arcmin). Given the high correlations, a large part of this error is likely a systematic offset, much larger than the uncertainty of the altimetry, which may reflect differences in vertical reference between the model and the altimetry product and the comparison of a point observation with a grid-cell value, as well as model errors. With only 29 observations from a single year, these results are a preliminary consistency check rather than a validation, and the sampling interval does not allow timing errors of one to two days to be detected.

Regarding the earlier simulated peaks, river model spin-up is an unlikely cause, because the river analysis that initialises the forecasts is a continuous simulation. Our working hypotheses are errors in the timing of the forecast precipitation, river channel parameters (riverbed roughness uses a default value and channel depth is derived from a global relationship; L441–443), the absence of a groundwater reservoir, so that subsurface runoff reaches the channel without delay, and the partitioning between surface and subsurface runoff in ecLand. We will discuss these in the revised manuscript, together with the planned work to test them.

Planned changes to the manuscript:

- Sect. 4 (opening paragraph): group Sects. 4.2–4.4 under “Illustrative applications”, keeping the GRDC evaluation (Sect. 4.1) as a separate evaluation of the long-term simulation. State that the cases demonstrate capability rather than assess performance, and that a systematic evaluation of the operational system is out of scope for this paper.
- Sect. 4.2 and Fig. 7 (Burdekin): report the peak magnitude and timing errors at Sellheim for the analysis and the forecasts, with the KGE over 2005–2024 for context, and add the rainfall forcing (see RC3).
- Sect. 4.3 and Fig. 8 (Luangwa): add the comparison with the DAHITI altimetry observations (correlation, mean absolute error, mean error at each resolution), with its limitations: altimetry uncertainty, vertical reference, point vs grid-cell comparison, and the small sample.
- Sect. 4.4 and Fig. 9 (Pakistan): compare simulated inundation extent with Sentinel-1-derived flood maps for the 2022 event at coincident satellite acquisition dates, using hit rate, false alarm ratio and critical success index (CSI). We will also compare the observed and simulated temporal evolution of total inundated area during the expansion and recession of the flood.
- Sects. 4.2–4.3: Add a paragraph discussing the earlier simulated peaks in the Burdekin case and, as noted by the referee, in the Luangwa case, with working hypotheses (e.g., absence of a groundwater reservoir, timing of the precipitation forcing, river roughness and channel geometry, runoff partitioning in ecLand) and planned work to address them.
- L511–514, L585, L612–613, L651 and L789: temper the statements that the illustrative cases do not support (“skillful hydrological forecasts”, “realistic”, “matching both the timing and magnitude”, “accurately representing the flood extent”, “robust foundation”).

- Sect. 3.4 (L441–443) and Conclusions: state that the CaMa-Flood river parameters have not yet been tuned, and mention their refinement using altimetry and in situ observations among possible future developments.

M9 – CH_4 fluxes and floodplain extent (Fig. 10)

Referee comment. *The CH4 fluxes displayed for May in figure 10 are quite different. Which one is closer to truth ? How correct are the floodplain extents predicted by the model and does this give arguments to assume that the output of the coupled system is more realistic ? It could also be that the inter-annual variability afforded by the coupling is so large that even a systematic error is acceptable because of the gain brought by the weather/climate feedbacks on the fluxes is valuable ? This should be documented.*

Author response. We thank the reviewer for this comment and agree that the fluxes in Fig. 10b and 10d differ. Part of this difference is expected. Fig. 10b shows a single day (15 May 2021) using a year-specific wetland fraction informed by GIEMS and CaMa-Flood. Fig. 10d shows the model-based monthly climatology used operationally since Cycle 49r1. Because it is a climatology, Fig. 10d is smoother in space and does not reflect the conditions of any particular year or day. Similar differences between year-specific and climatological inundation have been reported previously (Marthews et al., 2022; McNorton, 2016; McNorton et al., 2016; Parker et al., 2022).

Whether either estimate is closer to the "truth" cannot be determined directly. Wetland extent is defined differently across products, and the relationship between inundation and CH_4 flux is complex. Even with a common definition, suitable validation data do not exist as in-situ inundation observations have large representation errors at model grid scale. For the same reason, the modelled floodplain extents cannot be fully validated. We have therefore evaluated the fluxes indirectly, by comparing simulated atmospheric CH_4 concentrations with TCCON observations [see Fig. S1 below, also included in the Supplement]. The simulation using the sum of the GIEMS climatology and the CaMa-Flood 2019 wetland extent produces the best fit to observations. Currently the wetland model only accounts for the inter-annual variability of temperature. However, previous studies found better performance with year-specific inundation than with a climatology (McNorton, 2016). This is currently being explored using real time satellite observations of wetland extent coupled with CaMa-Flood inundation.

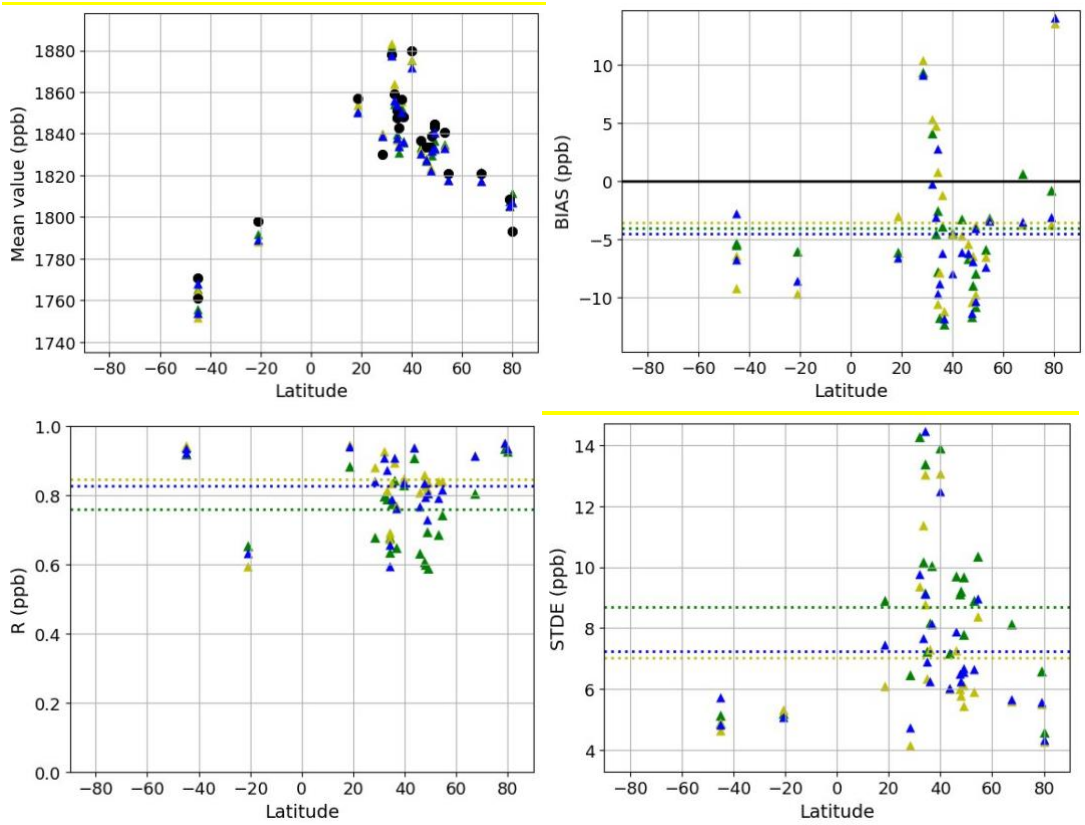


Figure S1: Evaluation of impact of wetland extent on atmospheric total column mean dry molar fraction XCH_4 [ppb] simulated by the IFS at TCCON sites (http://www.tccon.org) with annual mean values (top left panel), annual bias (top right panel), Pearson correlation coefficient (R, lower left panel) and random error (STDE, lower right panel) in 2018. Blue triangles correspond to the simulation using the maximum of the GIEMS climatology and CaMa-Flood 2019 wetland extent; Yellow triangles correspond to the simulation using the sum of the GIEMS climatology and CaMa-Flood 2019 wetland extent (used currently in the operational CAMS CH4 forecast/analysis); Green correspond to the simulations using a climatology of CH_4 wetland fluxes based on the LPJ-WHyMe dataset (Spahni et al., 2011) and used in the CAMS GHG re-analysis (Agustí-Panareda et al., 2023). TCCON observations are plotted with black circles. The dashed lines show the mean value across all TCCON sites for each IFS simulation. The simulation using the sum of the GIEMS climatology and the CaMa-Flood 2019 wetland extent produces the best fit to observations.

Planned changes to the manuscript:

- Fig. 10 caption: describe each panel precisely:

  (a) the GIEMS May climatology (1992–2020);

  (c) the May 2019 flooded fraction from the IFS river reference simulation;

  (b) the online flux for 15 May 2021;

  (d) the LPJ-WHyMe climatology (Spahni et al., 2011) used in CAMS before Cycle 49r1.

- Sect. 4.5 (L678–689): rewrite the second half of the section to cover five points:
  - explain the change from the offline LPJ-WHyMe climatology to fluxes computed online within ecLand since Cycle 49r1, and how the wetland fraction is constructed;

- explain why the fluxes in Fig. 10b and 10d differ (a single day vs a multi-year monthly mean, and different representations of wetland extent), with reference to previous studies;
- explain why neither the floodplain extent nor the fluxes can be validated directly, and summarise the indirect evaluation against TCCON column-averaged CH_4, in which the combination of GIEMS and the CaMa-Flood flooded fraction performed best;
- describe the main advance (fluxes that respond to the simulated weather at every timestep, including over the forecast range) and the main limitation (climatological wetland extent, so no interannual or sub-monthly variability in inundation);
- outline the planned use of real-time satellite inundation, the IFS river analysis and the IFS river forecasts.

- Supplement: add Fig. S1 with the TCCON evaluation.

M10 – Runoff generation at different resolutions; hillslope and aquifer processes (Sect. 4.6.3)

Referee comment. *Section 4.6.3 : The runoff generation at different spatial resolutions is only part of the issue. At higher resolution the recharge of soil moisture from the shallow aquifers is critical for the spatial distribution of surface fluxes, as hypothesized by Fan et al. 2019. It has now been demonstrated that a km-scale resolution this contribution of hill slope process to evaporation is a missing process in most land surface models (Polcher at al. 2026). This goes back to the above issue of how to define the coupling between the vertical energy, water and carbon fluxes treated in ecLand with the lateral water fluxes predicted by CaMa-Flood.*

Author response. We agree, and we thank the referee for raising this point. It reflects a broader challenge for land surface modelling at kilometre scale. Most land surface schemes, including ecLand, were designed for grid spacings of 100–200 km, where hillslope-scale processes, such as lateral subsurface flow and groundwater redistribution from ridges to valleys, are unresolved and can at best be parameterised statistically. At the kilometre-scale resolutions now used in the IFS and in DestinE (~4.5 km; Sect. 4.6.3), these processes are increasingly resolved by the grid, yet they are still represented by parameterisations designed for much coarser scales or not represented at all. The scale dependence of runoff generation shown in Sect. 4.6.3, where the sub-grid orography controlling the saturated fraction behaves differently across resolutions, is one symptom of this mismatch. As the referee notes, the missing contribution of shallow groundwater to soil moisture and evaporation (Fan et al., 2019; Polcher et al., 2026) is another.

Addressing this requires revising the coupling between the vertical energy, water and carbon fluxes in ecLand and the lateral water fluxes, as discussed in M4, M5 and m7. Representing lateral groundwater redistribution is beyond the scope of this paper, which documents the integration of river routing as the first lateral process in the IFS land surface. We will, however, discuss this limitation explicitly and mention scale-aware process representations, including lateral groundwater redistribution, among possible future developments. Besides, there are ongoing developments to include specifically a water table depth and a ground water recharge mechanism, as these are identified as shortcomings in the timing of runoff.

Planned changes to the manuscript:

- Sect. 4.6.3: add a paragraph explaining that land surface parameterisations designed for coarse grids are now applied at kilometre scale, where hillslope processes become resolved. Present the scale dependence of runoff generation as one example, and the missing

groundwater contribution to soil moisture and evaporation as another, citing Fan et al. (2019) and Polcher et al. (2026).
- Conclusions (future developments): mention scale-aware land surface process representations and lateral groundwater redistribution among possible future developments.

Minor comments

m1 – L39: GEWEX/WCRP

Referee comment. *Line 39 : Such developments are also supported by GEWEX/WCRP (see Stephens et al 2023).*

Author response. Agreed.

Planned changes to the manuscript:

- Sect. 1 (L39): cite Stephens et al. (2023) at L39.

m2 – L164: move Sect. 2.3 to an appendix

Referee comment. *Line 164 : Section 2.3 could move to an annex. It is expected from any professional Earth system model development team to have quality control procedures in place for the code and the model output.*

Author response. We thank the referee for this suggestion and agree that quality control procedures are expected of any model development team. Sect. 2.3 serves a different purpose: it shows what integrating a new component into an operational system involves, beyond developing it offline. River routing has to pass the same technical, scientific and computational acceptance steps as any other IFS change, including reproducibility tests, forecast scorecards and strict timing constraints. These requirements directly shaped the design choices described in Sect. 3, such as the single code base across configurations and the basin-based parallelisation. We believe this is useful for groups planning to bring hydrological components into operational prediction systems, and RC1 also noted that it gives insight into the complexity of such an undertaking. We would therefore prefer to keep Sect. 2.3 in the main text, but we will shorten it and make its link to the river routing integration explicit.

Planned changes to the manuscript:

- Sect. 2.3: shorten and make its link to the river routing integration explicit.

m3 – L224: 1D site runs and river evaluation

Referee comment. *Line 224 : How can the 1D simulations be used to evaluate the atmosphere-to-land-to-river ? These cannot be point scale simulations but have to be at the catchment level. Can you please clarify.*

Author response. We thank the referee for pointing this out. The 1D site runs are not used to evaluate river routing. As shown in Fig. 1, CaMa-Flood is linked only to the 2D area/global runs, and the 1D site runs evaluate the land surface only. The description at L224, “1-way atmosphere-to-land-to-rivers coupling”, was meant to refer to the type of coupling configuration (Sect. 3.3.1), but we agree it suggests that rivers are included, and we will correct it. River routing is evaluated in the 2D

runs, over regional domains covering whole catchments (e.g. the Luangwa case, Sect. 4.3) or globally (Sect. 4.1).

Planned changes to the manuscript:

- Sect. 2.3.4 (L224–225)**:** describe the 1D site runs as 1-way atmosphere-to-land coupling, without river routing.

m4 – Fig. 1: evolution of IVER and QUAVER

Referee comment. *Figure 1 and explanatory text : It is not explained in the paper how IVER and QUAVAR have evolved with the addition of the lateral water transports in the ESM. Surface hydrology is arguably one of the best observed branch of the water cycle and thus offers excellent validation opportunities as outlied later in the paper.*

Author response. We agree that this should be explained, and we share the referee's view that river observations offer an excellent opportunity for validation. The integration of river variables into IVER and QUAVER is currently under way. In the meantime, river discharge is evaluated against in situ observations outside these systems, as in Sect. 4.1. Once the integration is complete, river discharge will be verified routinely alongside other surface variables. We will mention this in the revised manuscript.

Planned changes to the manuscript:

- Sect. 2.3.4 (L244–252): add a sentence noting that the integration of river variables into IVER and QUAVER is under way.

m5 – L323: reservoirs and human water use

Referee comment. *Line 323 : CaMa-Flood includes a reservoir operation scheme. Does eLand also include a water usage model (irrigation, domestic, industrial) to estimate the water consumed (i.e. evaporated) ? Is there or is there not an assumption of pristine continents in the system ? Please clarify if this option is activated or not.*

Author response. We thank the referee for asking us to clarify this. Neither the CaMa-Flood reservoir operation scheme (Hanazaki et al., 2022) nor the simple flood protection (levee) scheme (Zhao et al., 2025) is active in the operational configuration. ecLand does not represent human water use, such as irrigation or domestic and industrial abstraction. The system therefore assumes natural conditions, and simulated discharge represents naturalised flow. We will state this explicitly (see M1).

Planned changes to the manuscript:

- Sect. 3.2 (L322–324): state that the reservoir operation and flood protection schemes are available in CaMa-Flood but not active in the operational configuration.
- New Sect. 3.3.6 (see M1): state that the system assumes natural conditions, with no reservoir regulation or human water use.

m6 – L334: river initial conditions and spin-up

Referee comment. *Line 334 : What are "river initial conditions" how can they be separated from groundwater initial conditions in particular for near surface aquifers. How long do these slower reservoirs take to spin-up from a old or warm start depending on the size of the catchment and its*

hydrogeology ? This spin-up temporal scale is essential to determine the inertia of the various water reservoir and thus the importance of the initial conditions for this forecasting system.

Author response. We thank the referee for this question. River initial conditions are the CaMa-Flood prognostic states listed at L336–337: river and floodplain storage and discharge, and river depth. CaMa-Flood includes an optional groundwater reservoir, but it was not activated in the operational configuration or in any of the simulations shown in this paper. Surface and subsurface runoff from ecLand are therefore routed directly to the river channel. ecLand has no explicit groundwater store, so river and groundwater initial conditions do not need to be separated. The slower water reservoirs are the deep soil layers of ecLand, initialised by the land data assimilation in the operational configuration, and the river, floodplain and lake storage in CaMa-Flood. We will correct the description of the operational configuration in Sect. 3.3.1 accordingly.

The river states are initialised from the IFS river reference simulation (Sect. 3.6). From a cold start, our experiments show that river and floodplain storages take up to 2 years to spin up, however large lakes in the CaMa-Flood domain can take up to 10 years to equilibrate. To reduce that spin-up time, we have implemented a long-term mean initial conditions (based on ERA5), that will reduce the spin-up time for these large lakes to 2 years. We will clarify this in Sect. 3.3.2.

Planned changes to the manuscript:

- Sect. 3.3.1 (L359–363): correct the description of the operational configuration: the groundwater reservoir is available but not active, and surface and subsurface runoff are routed directly to the river channel.
- Sect. 3.3.2: list the states included in the river initial conditions, state that the groundwater reservoir is not active, and report the spin-up time scales from a cold start.

m7 – L350: "coupling land to river"

Referee comment. *Line 350 : Does "coupling land to river" make sense when rivers are an integral part of the land ? It has been shown that redistribution of water by groundwater and rivers shape the landscape and drives evapotranspiration fluxes (Fan et al. 2019, Polcher et al. 2026). So why limit the coupling to rivers ?*

Author response. We thank the referee for this comment and agree that rivers, floodplains and groundwater are integral components of the land hydrological system, and that the lateral redistribution of water influences soil moisture and evapotranspiration (Fan et al., 2019; Polcher et al., 2026). Our wording ("coupling from land to rivers") was imprecise and may have suggested a conceptual separation that we did not intend. What is coupled here are two model components: the vertical, column-based hydrology of ecLand, which generates surface and subsurface runoff, and the lateral routing scheme CaMa-Flood. Since ecLand does not represent any lateral water transfer, river routing is the first lateral process introduced into the IFS land surface, and the one-way configuration described here is a first step.

In this context, it is useful to distinguish exchanges with the atmosphere from feedbacks on the land. Evaporation from floodplain water is an exchange between the river storage and the atmosphere. In the experimental 2-way configuration, it is withdrawn from the CaMa-Flood storage (see M5c). It would become a process of the land surface itself only once ecLand includes a storage for transient surface water. The direct feedback of rivers on the land is the infiltration of floodplain water into the soil, which is not yet represented. Both this and lateral groundwater redistribution (see M10) are possible future developments.

Planned changes to the manuscript:

- Sect. 3.3 (L349–353): revise the wording to describe the coupling between the vertical land hydrology of ecLand and the lateral routing of CaMa-Flood. Distinguish evaporation from floodplain water (an exchange with the atmosphere) from infiltration (a feedback on the land), and state that infiltration and lateral groundwater redistribution are not yet represented.

m8 – L367: flooded fraction vs lake fraction

Referee comment. *Line 367 : Why is only the flooded fraction passed to ecLand and not the lake fraction? It is a minimum to ensure consistent assumptions on the water mass conservation. Later in this paragraph the authors understand the limit of their inconsistent assumptions (as "infiltration of the flooded water" is need.) between both models. Posing these assumption up-front would make the sentence starting in line 370 clearer.*

Author response. We thank the referee for this comment. In the experimental 2-way configuration, the flooded fraction is passed as an additive input to ecLand while the lake fraction is already passed as input, since lakes are already represented in ecLand by a static lake fraction and the FLake scheme (Balsamo et al., 2012). Passing the CaMa-Flood flooded fraction does not incur in double counting as this is added to the existing lake cover fraction. We agree, however, that this does not ensure consistent water mass conservation for lakes: as explained in M4, lake evaporation in ecLand is not withdrawn from the routed water, and a full lake water balance is not yet represented. Achieving this would require revising the lake scheme, a separate development that we will mention among future developments (see M5a). We also agree that these assumptions should be stated before the description of the 2-way coupling, and we will reorder the paragraph accordingly.

Planned changes to the manuscript:

- Sect. 3.3.1 (L367–371): reorder the paragraph so that the assumptions are stated before the description of the 2-way coupling.

On behalf of all co-authors, Cinzia Mazzetti

References

Agustí-Panareda, A., Barré, J., Massart, S., et al.: Technical note: The CAMS greenhouse gas reanalysis from 2003 to 2020, Atmos. Chem. Phys., 23, 3829–3859, https://doi.org/10.5194/acp-23-3829-2023, 2023.

Balsamo, G., Salgado, R., Dutra, E., Boussetta, S., Stockdale, T., and Potes, M.: On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model, Tellus A, 64, 15829, https://doi.org/10.3402/tellusa.v64i0.15829, 2012.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, Geosci. Model Dev., 4, 677–699, https://doi.org/10.5194/gmd-4-677-2011, 2011.

Boussetta, S., Balsamo, G., Arduini, G., Dutra, E., McNorton, J., Choulga, M., Agustí-Panareda, A., Beljaars, A., Wedi, N., Muñoz-Sabater, J., de Rosnay, P., Sandu, I., Hadade, I., Carver, G., Mazzetti, C., Prudhomme, C., Yamazaki, D., and Zsoter, E.: ECLand: The ECMWF land surface modelling system, Atmosphere, 12, 723, https://doi.org/10.3390/atmos12060723, 2021.

Choulga, M., Kourzeneva, E., Balsamo, G., Boussetta, S., and Wedi, N.: Upgraded global mapping information for earth system modelling: an application to surface water depth at the ECMWF, Hydrol. Earth Syst. Sci., 23, 4051–4076, https://doi.org/10.5194/hess-23-4051-2019, 2019.

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics, Geosci. Model Dev., 4, 701–722, https://doi.org/10.5194/gmd-4-701-2011, 2011.

Collignan, J., Polcher, J., Bastin, S., and Quintana-Seguí, P.: Identifying and quantifying the impact of climatic and non-climatic drivers on river discharge in Europe, Water Resour. Res., 61, e2024WR038220, https://doi.org/10.1029/2024WR038220, 2025.

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J.-P., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., and Voldoire, A.: Recent changes in the ISBA-CTRIP land surface system for use in the CNRM-CM6 climate model and in global off-line hydrological applications, J. Adv. Model. Earth Syst., 11, 1207–1252, https://doi.org/10.1029/2018MS001545, 2019.

ECMWF: IFS Documentation CY49R1 – Part IV: Physical Processes, ECMWF, Reading, UK, https://doi.org/10.21957/c731ee1102, 2024.

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., et al.: Hillslope hydrology in global change research and Earth system modeling, Water Resour. Res., 55, 1737–1772, https://doi.org/10.1029/2018WR023903, 2019.

Gochis, D. J., Yu, W., and Yates, D. N.: The WRF-Hydro model technical description and user's guide, version 3.0, NCAR Technical Document, 123 pp., NCAR, Boulder, CO, USA, https://ral.ucar.edu/sites/default/files/public/WRF_Hydro_User_Guide_v3.0_CLEAN.pdf, 2015.

Guimberteau, M., Drapeau, G., Ronchail, J., Sultan, B., Polcher, J., et al.: Discharge simulation in the sub-basins of the Amazon using ORCHIDEE forced by new datasets, Hydrol. Earth Syst. Sci., 16, 911–935, https://doi.org/10.5194/hess-16-911-2012, 2012.

Hallouin, T., Ellis, R. J., Clark, D. B., Dadson, S. J., Hughes, A. G., Lawrence, B. N., Lister, G. M. S., and Polcher, J.: UniFHy v0.1.1: a community modelling framework for the terrestrial water cycle in Python, Geosci. Model Dev., 15, 9177–9196, https://doi.org/10.5194/gmd-15-9177-2022, 2022.

Hanazaki, R., Yamazaki, D., and Yoshimura, K.: Development of a reservoir flood control scheme for global flood models, J. Adv. Model. Earth Syst., 14, e2021MS002944, https://doi.org/10.1029/2021MS002944, 2022.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrol. Earth Syst. Sci., 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., et al.: The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing

uncertainty, J. Adv. Model. Earth Syst., 11, 4245–4287, https://doi.org/10.1029/2018MS001583, 2019.

Lehner, B. and 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.

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., and 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.

Marthews, T. R., Dadson, S. J., Clark, D. B., Blyth, E. M., Hayman, G. D., Yamazaki, D., Becher, O. R. E., Martínez-de la Torre, A., Prigent, C., and Jiménez, C.: Inundation prediction in tropical wetlands from JULES-CaMa-Flood global land surface simulations, Hydrol. Earth Syst. Sci., 26, 3151–3175, https://doi.org/10.5194/hess-26-3151-2022, 2022.

McNorton, J. R.: Analysis of recent atmospheric methane trends using models and observations, PhD thesis, University of Leeds, UK, https://etheses.whiterose.ac.uk/id/eprint/13294/, 2016.

McNorton, J., Gloor, E., Wilson, C., Hayman, G. D., Gedney, N., Comyn-Platt, E., Marthews, T., Parker, R. J., Boesch, H., and Chipperfield, M. P.: Role of regional wetland emissions in atmospheric methane variability, Geophys. Res. Lett., 43, 11433–11444, https://doi.org/10.1002/2016GL070649, 2016.

Nitta, T., Arakawa, T., Hatono, M., Takeshima, A., and Yoshimura, K.: Development of Integrated Land Simulator, Prog. Earth Planet. Sci., 7, 68, https://doi.org/10.1186/s40645-020-00383-7, 2020.

Parker, R. J., Wilson, C., Comyn-Platt, E., Hayman, G., Marthews, T. R., Bloom, A. A., Lunt, M. F., Gedney, N., Dadson, S. J., McNorton, J., Humpage, N., Boesch, H., Chipperfield, M. P., Palmer, P. I., and Yamazaki, D.: Evaluation of wetland CH4 in the Joint UK Land Environment Simulator (JULES) land surface model using satellite observations, Biogeosciences, 19, 5779–5805, https://doi.org/10.5194/bg-19-5779-2022, 2022.

Polcher, J., et al.: A framework to evaluate and identify development requirements for land-surface models at km-scale resolution: Application to a semi-arid and mountainous region, Q. J. Roy. Meteor. Soc., https://doi.org/10.1002/qj.70229, 2026.

Schrapffer, A., Polcher, J., Sörensson, A., and Fita, L.: Introducing a new floodplain scheme in ORCHIDEE (version 7885): validation and evaluation over the Pantanal wetlands, Geosci. Model Dev., 16, 5755–5782, https://doi.org/10.5194/gmd-16-5755-2023, 2023.

Schwatke, C., Dettmering, D., Bosch, W., and Seitz, F.: DAHITI – an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry, Hydrol. Earth Syst. Sci., 19, 4345–4364, https://doi.org/10.5194/hess-19-4345-2015, 2015.

Shrestha, P., Sulis, M., Masbou, M., Kollet, S., and Simmer, C.: A scale-consistent terrestrial systems modeling platform based on COSMO, CLM, and ParFlow, Mon. Weather Rev., 142, 3466–3483, https://doi.org/10.1175/MWR-D-14-00029.1, 2014.

Spahni, R., Wania, R., Neef, L., van Weele, M., Pison, I., Bousquet, P., Frankenberg, C., Foster, P. N., Joos, F., Prentice, I. C., and van Velthoven, P.: Constraining global methane emissions and uptake by ecosystems, Biogeosciences, 8, 1643–1665, https://doi.org/10.5194/bg-8-1643-2011, 2011.

Stephens, G., Polcher, J., Zeng, X., van Oevelen, P., Poveda, G., Bosilovich, M., Ahn, M.-H., et al.: The first 30 years of GEWEX, B. Am. Meteorol. Soc., 104, E126–E157, https://doi.org/10.1175/BAMS-D-22-0061.1, 2023.

Tokuda, D., Kim, H., Yamazaki, D., and Oki, T.: Development of a global river water temperature model considering fluvial dynamics and seasonal freeze-thaw cycle, Water Resour. Res., 55, 1366–1383, https://doi.org/10.1029/2018WR023083, 2019.

Vicente-Serrano, S. M., Peña-Gallardo, M., Hannaford, J., Murphy, C., Lorenzo-Lacruz, J., Dominguez-Castro, F., López-Moreno, J. I., Beguería, S., Noguera, I., Harrigan, S., and Vidal, J.-P.: Climate, irrigation, and land cover change explain streamflow trends in countries bordering the Northeast Atlantic, Geophys. Res. Lett., 46, 10821–10833, https://doi.org/10.1029/2019GL084084, 2019.

Wagner, S., Fersch, B., Yuan, F., Yu, Z., and Kunstmann, H.: Fully coupled atmospheric-hydrological modeling at regional and long-term scales: Development, application, and analysis of WRF-HMS, Water Resour. Res., 52, 3187–3211, https://doi.org/10.1002/2015WR018185, 2016.

Zhao, G., Yamazaki, D., Tanaka, Y., Zhou, X., Li, S., Hu, Y., Hirabayashi, Y., Neal, J., and Bates, P.: Developing a levee module for global flood modeling with a reach-level parameterization approach, Water Resour. Res., 61, e2024WR039790, https://doi.org/10.1029/2024WR039790, 2025.

Use of AI tools: in preparing these responses, the authors used Claude (Anthropic) to check that all referee comments had been addressed, to identify inconsistencies between the responses and the preprint, and to suggest edits to the wording. All responses and proposed changes were reviewed, revised where needed and approved by the authors, who take full responsibility for their content.