Review of : „From single Storm to Global Waves: A Global 2.5 km ICON Simulation of Weather and Climate“

The authors analyze a 4-year simulation of the ICOsahedral Non-hydrostatic model with a horizontal grid spacing of 2.5 km. The simulation has been done by porting ICON into GPUs in the ALPS machine as part of the EXCLAIM project. Through different analyses, the authors evaluate the simulation regarding the large-scale and mesoscale features of the simulations by comparing them with observations. The analyses point out that large-scale features are adequately represented (air temperature, precipitation, pattern of tropical cyclones, among others), but biases are observed in meso-scale or regional features. Overall, the manuscript is well written, and the amount of analysis is considerable. However, in the actual state of the manuscript, the little contextualization of the results within the current knowledge of km-scale Earth system community gives the impression that this is an isolated study. Moreover, it is not clear whether the authors aim to give some insights of the simulation or just present the simulations. I think that pointing out what the simulation of a horizontal grid spacing of 2.5 km gives compared to coarser or finer (if it exists) resolution will be a valuable contribution to the community working with km-scale Earth System models. Another criticism of this work is the word “climate simulation”, and I think that this has implications on the technical side. In the following lines, the major concerns about this manuscript are described.

Major comments:

1. The manuscript gives the impression, starting with the title, that the authors will describe climate simulations. I struggled a bit with the word “climate” in this study for two reasons. One is the fact that the concept of climate involves different temporal variability, from synoptic to multidecadal. Second, a climate simulation should involve communication between components of the climate systems. While this is the case for land and atmosphere in this study, the ocean is not, i.e., sea-surface temperatures are prescribed. This means that the global atmosphere-only 4-year simulation presented in this study is not properly a climate simulation, i.e., too short for considering climate and only land-atmospheric interaction. These two points become relevant for the technical part. The throughput of this global atmosphere-only 4-year simulation is 0.25 simulated year per day (SYPD). This means that even in an atmosphere-only configuration, the throughput is much lower than what is expected when conducting climate simulations (ocean-land-atmosphere), which is 1SYPD. Also, this implies that with a dynamical ocean, the throughput will be lower. My suggestion is that the authors refer to the simulations as a 4-year atmosphere-only simulation rather than a climate one. Moreover, if the authors intend to provide climate simulations in the future, it would be adequate to elaborate a bit more on how the authors will increase the throughput.
2. I think that the manuscript will benefit tremendously if the authors describe more of the advances in the global storm-resolving and km-scale Earth System models. The first initiative of conducting global storm-resolving models was done by the Japanese community (Tomita et al., 2005; Miura et al., 2007;

Satoh et al., 2008), developing the Non-hydrostatic ICosahedral Atmospheric Model (NICAM). Now, there are more centers developing global storm-resolving models (DYAMOND; Stevens et al., 2019; Satoh et al., 2020), and a few of them are coupling with the ocean, referred to as km-scale Earth System models (Hoheneggar et al., 2023; Rackow et al., 2025). So, there is a considerable spectrum of simulations using grid-spacing of 10 km or less. In the ICON's universe, the model the authors are using, advances have been done to couple with the ocean (Hohenegger et al. 2023), conducting multidecadal climate simulations (Segura et al., 2025a), and recently, carbon cycle has been incorporated into the model, and a simulation was conducted using a horizontal grid-scaping of 1.25 km (Klocke et al., 2025). Having this in mind is important in order to contextualize the results of the simulation presented in this study. A simple question would be what is the added value of the 2.5 km simulation to the large universe of simulations being produced with coarser or finer (if they exist) horizontal grid spacing?

3. I understand that this work is mostly to introduce the simulation, and future analysis will probably come in the future. However, I found a bit of disproportion in the explanation of the biases. Regarding precipitation and MCS's, the authors give a considerable explanation of how MCS's characteristics are biased compared to observations. Then, the authors go back to the large-scale, indicating that it's probable that most of the precipitation in the large-scale feature of precipitation is related to congestus and shallow clouds. Considering the other features analyzed in the manuscripts, the explanation most of the time is to cite other papers. For instance, the biases in temperature, according to the authors, are related to biases in soil moisture and are probably related to the lack of representation in the lateral flux of water. Nevertheless, there is no analysis of the soil moisture or the partition of energy between latent and sensible heat flux. Or are all the biases related to the lateral flux of water, or could it be related to local precipitation biases? In a similar direction are the tropical cyclones and maximum wind speed. The explanation of the biases lies in the difference between the time step of observations and the simulation. My suggestion is that the authors should reconsider which analyses they want to show. This could be based on what is actually shown in previous studies, selecting the one that provides new insights in the km-scale modeling community.
4. I wonder if the regionalization of the analysis concerning precipitation maximum and diurnal cycle provides insightful results, which does not seem to be the case. At the end, it seems that the good representation of such characteristics is region-dependent. In regional km-scale models, due to the spatial constraint of the boundary conditions, it is expected that precipitation characteristics in the region of simulation should be similar to observations inside the region. However, this is more complex for global atmosphere-only km-scale simulations. Small changes in the pattern of local winds, temperature, or soil moisture can change the precipitation characteristics

locally. My suggestion here is to use a bigger domain to analyze the diurnal cycle of precipitation. For example, land or ocean, tropics and extratropics, similar to what was done in Segura et al. (2022) or Takasuka et al. (2024). This can also be done for the case of precipitation maximum.

Specific comments:

- Line 24: The first ones in developing global storm-resolving models is the Japanese community, who developed NICAM (Tomita et al., 2005; Miura et al., 2007; Satoh et al., 2008).
- Lines 26-27: As I mentioned, global storm-resolving models were already conducted in Japan.
- Line 32: convection-permitting, storm-resolving, or km-scale atmosphere models? Aside from a philosophical/semantical question, try to use one nomenclature across the manuscript. I would prefer storm-resolving or km-scale atmosphere models. Convection-permitting could be arguable due to the type of convection being resolved: shallow, congestus-type, or deep.
- Line 34: I would add Marsham et al. (2013)
- Line 35: What about Holloway et al. (2012) for precipitation extremes
- Line 37: Also Marsham et al. (2013).
- Line 44: I do not think that Hohenegger et al. (2009) showed that lateral groundwater fluxes help in representing soil moisture - precipitation characteristics. The authors evidenced that by explicitly resolving convection, there is a negative feedback between precipitation and soil moisture, i.e., dry soils can get precipitation. This is not observed when convection parameterizations are used. This has been studied using ICON (Lee and Hohenegger 2024), and was pointed out as a possible reason for a good representation of the terrestrial tropical rainbelt in ICON (Segura et al., 2022).
- Lines 49-50: Global km-scale atmosphere simulations were conducted already 20 years ago (Tomita et al., 2005; Miura et al., 2007; Satoh et al., 2008)
- Line 62: Hohenegger et al. (2023) described ICON as a km-scale Earth System Model. Similar for IFS-FESOM/NEMO, which was described in Rackow et al. (2025). Segura et al. (2025a) showed that multidecadal simulations using km-scale Earth System Models (ICON and IFS-FESOM) are feasible.
- Lines 63- 64: Simulations using km-scale Earth System Models were conducted with a horizontal grid spacing of ~ 2.5 km, using ICON (Hohenegger et al., 2023) and IFS-FESOM (Rackow et al., 2025). The integration time is shorter than one year, though.
- Line 70: A better term instead of “weather-to-climate” would be “meso- to large-scale systems” in this study. Climate has several connotations that are not totally addressed by the simulations in this study (interannual to decadal variability or a coupling system).
- Line 95: Is this throughput enough to produce climate simulations?
- Line 227-229: So, the argument is that there is a lack of soil moisture due to a lack of lateral flux of water. My question is how much of this dry bias is related to a precipitation bias. In Figure 3g, there is a bias of 1.8 mm d-1 over North America, in the same region where the warm bias is. Moreover, I do not think that the lateral flux of water can explain the warm bias on a continental scale. Take, for example, the northeast Asia in the winter season (DJF) or the South America continent in JJA or SON. If the lateral flux of water is the culprit, could

the authors elaborate a bit more on how it can explain biases on a continental scale?

- Line 238: Is 7.8 W m-2 referred to 6.8 W m-2 or 25 W m-2?
- Lines 239-240: Is the overestimation in the frequency of shallow clouds the reason for a positive bias in downwelling shortwave radiation at the land-surface? Or is it for the ocean surface? If it is at the ocean surface, I can understand this argument for the subtropical region and eastern sides of the basins, but for the intertropical convergence zone, is that the case?
- Lines 245: What do you mean by surface evaporation or soil evaporation is not well represented? Do you mean the bulk formulas or the forcings as radiation, winds, or soil moisture? How is the ratio in the partition between latent and sensible heat flux? This would indicate that maybe there is a problem with the soil moisture.
- Lines 248-249: In Figures 2d,e, latent heat is slightly overestimated, but in Figures 2f,h, it seems that there is an underestimation. How is it possible? Is it because in the transition season, latent heat flux is overestimated?
- Lines 250-252: Could it be that the soil moisture biases are due to an underestimation of local precipitation?
- Lines 252-253: What does it mean more site-specific?
- Lines 253-254: How do other variables make the result robust? Is it because warm biases is explained by surface heat fluxes?
- Lines 258-259: So, biases in shortwave radiation are not related to shallow clouds, as implied in Section 3.2, but rather related to deep convection.
- Lines 259-260: Well-simulated land rainfall pattern in the tropics has been shown in Segura et al. (2022,2025b) with coarser horizontal grid spacing. So, does it mean that in ICON, rainfall patterns over tropical land are already well simulated, independent of the configuration?
- Lines 266-267: The double band of precipitation has been shown in coupled km-scale (Segura et al., 2022) and atmosphere-only km-scale simulations (Segura et al., 2025b).
- Line 274: Heavier precipitation intensities at hourly time-step than observations have also been observed in IFS-FESOM using horizontal grid spacing of 2.8 and 4.4 km without convective parameterization. Then, is the authors implying that using IMERG is not the best tool to characterize extreme rainfall events?
- Lines 273-300: It’s difficult to get a clear message of the regionalization. Are the authors showing that the IMERG is not adequate for studying the diurnal cycle of precipitation? But according to the regionalization, this is not totally bad. There are, of course, errors in the satellite data, but overall, and in particular in the intensity (Figure 5), it seems good.
- Lines 308-310: Is the underestimation of maximum wind speed related to the frequency of tropical cyclones? Or is it that the tropical cycles in ICON tend to have lower wind speeds than observations?
- Line 310: “Most tropical”? According to Figure S3, there are not many stations in the tropics. And most of them are focused on the Maritime Continent.
- Lines 313-314: It’s very hard to see how the autocorrelation gives a proxy of the structure of the wind-producing storms. Is it possible to elaborate more?
- Lines 315-319: Does it mean that km-scale models need to be evaluated differently?
- Line 326: Did you expect this? Takasuka et al. (2024) also showed this behavior in NICAM.

- Lines 330-331: ICON simulates one peak, but observations have two. The difference in the distribution is due to the track method. This is the explanation for the first peak, but what about the second one?
- Line 333: (Fig. 9), typo
- Line 339: typo
- Line 341: typo. Takasuka et al. (2024) also showed this.
- Lines 344: Takayabu (1994) was the first to use this method.
- Line 351: Were Kiladis et al. (2009) the first ones to show this?
- Line 357: southern hemisphere or Southern Hemisphere?
- Line 364: underestimation of MCS's sizes seems similar in the ocean and on land. It's difficult to see any big differences in the Violin plot.
- Line 371-372: In ICON, MCS's velocities are independent of the size, which is not the case in observations. Is there a hypothesis about this?
- Lines 379-380: I will also point out that the morphology of precipitation between observation and ICON is totally different. In observations, there is a tendency to sustain precipitation in the 20% of the lifetime; the long-lived ones over land are more efficient in doing this, till 40%. On the other hand, simulated MCSs do not do it. They have a rapid decay in terms of precipitation. It seems that long-lived ones are doing it, but during the peak of their growth.
- Lines 388-389: What is the source of bias?
- Lines 394-395: Most of the biases in non-MCS clouds are over land and a little over ocean.
- Line 394: Fig S10 or Fig S8
- Lines 396-401: The authors could refer to Segura and Hohenegger (2024) regarding cumulonimbus, congestus, and shallow clouds and their relationship with precipitation. In their work, shallow clouds only represent 8% of the total amount of precipitation, while congestus accounts for mostly 46%. Is this result going in the same direction? Based on your analysis of the surface energy, shallow clouds are probably small over land, due to more shortwave radiation into the surface. Another question is how much of the precipitation satellite data can observe the shallow or congestus precipitation.
- Lines 432-434: It is difficult to consider a 4-year atmosphere-only km-scale simulation as a climate simulation, through the throughput and the simulated time-scale variability. However, my impression is that this simulation tends to fill the gap between simulating meso-scale and large-scale features of the atmosphere.
- Line 440: Weren't the surface heat fluxes showing the strong biases?
- Lines 447-448: Is this something new?
- Lines 461-462: Could the authors give an estimation of how much comes from shallow clouds? I think most of the precipitation comes from congestus clouds.
- Lines 466-467: Convective Rossby waves are also controlled by column water vapor (Nakamura and Takayabu 2022).
- Lines 469-470: But atmosphere-only simulations can simulate convective coupled equatorial waves (Falko and Rios-Berrios 2021, Takasuka et al., 2024, Ortega et al., 2026)
- Line 477: variability? I did not see any time-scale variability analysis.
- Lines 481-482: It's difficult to see this given the number of projects using km-scale Earth System Models. nextGEMS produced an 8-month simulation (coupled) using IFS-FESOM with a horizontal grid spacing of 2.8 km (Rackow et al., 2025). Using ICON, a two-month simulation was produced with a horizontal grid spacing of 2.5 km (Hohenegger et al., 2023). Recently, ICON's

simulations with a horizontal grid spacing of 1.25 km were conducted, in which the carbon cycle was included (Klocke et al., 2025). The simulation in this study does not show adequate throughput to conduct climate simulations, even if an atmosphere-only configuration is used. Rather, I would say that it's going to fill the gap between the meso- and the large-scale circulation in the atmosphere.

References

- Tomita, H., Miura, H., Iga, S., Nasuno, T., & Satoh, M. (2005). A global cloud-resolving simulation: Preliminary results from an aqua planet experiment. *Geophysical Research Letters*, *32*(8), 1–4. https://doi.org/10.1029/2005GL022459
- Miura, H., Satoh, M., Tomita, H., Noda, A. T., Nasuno, T., & Iga, S. I. (2007). A short-duration global cloud-resolving simulation with a realistic land and sea distribution. *Geophysical Research Letters*, *34*(2), 2–6. https://doi.org/10.1029/2006GL027448
- Satoh, M., Matsuno, T., Tomita, H., Miura, H., Nasuno, T., & Iga, S. (2008). Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. *Journal of Computational Physics*, *227*(7), 3486–3514. https://doi.org/10.1016/j.jcp.2007.02.006
- Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., Düben, P., Judt, F., Khairoutdinov, M., Klocke, D., Kodama, C., Kornblueh, L., Lin, S. J., Neumann, P., Putman, W. M., Röber, N., Shibuya, R., Vanniere, B., Vidale, P. L., … Zhou, L. (2019). DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Progress in Earth and Planetary Science*, *6*(1). https://doi.org/10.1186/s40645-019-0304-z
- Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W. M., & Düben, P. (2019). Global Cloud-Resolving Models. *Current Climate Change Reports*, *5*(3), 172–184. https://doi.org/10.1007/s40641-019-00131-0
- Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., Bao, J., Bastin, S., Behravesh, M., Bergemann, M., Biercamp, J., Bockelmann, H., Brokopf, R., Brüggemann, N., Casaroli, L., Chegini, F., Datseris, G., Esch, M., George, G., … Stevens, B. (2023). ICON-Sapphire : simulating the components of the Earth System and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development*, 779–811. https://doi.org/10.5194/gmd-2022-171
- Rackow, T., Pedruzo-Bagazgoitia, X., Becker, T., Milinski, S., Sandu, I., Aguridan, R., Bechtold, P., Beyer, S., Bidlot, J., Boussetta, S., Deconinck, W., Diamantakis, M., Dueben, P., Dutra, E., Forbes, R., Ghosh, R., Goessling, H. F., Hadade, I., Hegewald, J., … Ziemen, F. (2025). Multi-year simulations at kilometre scale with the Integrated Forecasting System coupled to FESOM2.5 and NEMOv3.4. *Geoscientific Model Development*, *18*(1), 33–69. https://doi.org/10.5194/gmd-18-33-2025
- Segura, H., Pedruzo-Bagazgoitia, X., Weiss, P., Müller, S. K., Rackow, T., Lee, J., Dolores-Tesillos, E., Benedict, I., Aengenheyster, M., Aguridan, R., Arduini, G., Baker, A. J., Bao, J., Bastin, S., Baulenas, E., Becker, T., Beyer, S., Bockelmann, H., Brüggemann, N., … Stevens, B. (2025a). nextGEMS: entering the era of kilometer-scale Earth system modeling. *Geoscientific Model Development*, *18*(20), 7735–7761. https://doi.org/10.5194/gmd-18-7735-2025
- Klocke, D., Frauen, C., Engels, J. F., Alexeev, D., Redler, R., Schnur, R., Haak, H., Kornblueh, L., Brüggemann, N., Chegini, F., Römmer, M., Hoffmann, L., Griessbach, S., Bode, M., Coles, J., Gila, M., Sawyer, W., Calotoiu, A., Budanaz, Y., … Stevens, B. (2025). Computing the Full Earth System at 1km Resolution. *Proceedings of the*

International Conference for High Performance Computing, Networking, Storage, and Analysis, SC 2025, 125–136. https://doi.org/10.1145/3712285.3771789
- Segura, H., Hohenegger, C., Wengel, C., & Stevens, B. (2022). Learning by Doing: Seasonal and Diurnal Features of Tropical Precipitation in a Global-Coupled Storm-Resolving Model. *Geophysical Research Letters*, *49*(24), 1–10. https://doi.org/10.1029/2022GL101796
- Takasuka, D., Kodama, C., Suematsu, T., Ohno, T., Yamada, Y., Seiki, T., Yashiro, H., Nakano, M., Miura, H., Noda, A. T., Nasuno, T., Miyakawa, T., & Masunaga, R. (2024). How Can We Improve the Seamless Representation of Climatological Statistics and Weather Toward Reliable Global K-Scale Climate Simulations? *Journal of Advances in Modeling Earth Systems*, *16*(2). https://doi.org/10.1029/2023MS003701
- Marsham, J. H., Dixon, N. S., Garcia-Carreras, L., Lister, G. M. S., Parker, D. J., Knippertz, P., & Birch, C. E. (2013). The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations. *Geophysical Research Letters*, *40*(9), 1843–1849. https://doi.org/10.1002/grl.50347
- Hohenegger, C., Brockhaus, P., Bretherton, C. S., & Schär, C. (2009). The soil moisture-precipitation feedback in simulations with explicit and parameterized convection. *Journal of Climate*, *22*(19), 5003–5020. https://doi.org/10.1175/2009JCLI2604.1
- Lee, J., Hohenegger Edited by Paul A Dirmeyer, C. I., & Dickinson, R. E. (2024). *Weaker land-atmosphere coupling in global storm-resolving simulation*. https://doi.org/10.1073/pnas
- Holloway, C. E., Woolnough, S. J., & Lister, G. M. S. (2012). Precipitation distributions for explicit versus parametrized convection in a large-domain high-resolution tropical case study. *Quarterly Journal of the Royal Meteorological Society*, *138*(668), 1692–1708. https://doi.org/10.1002/qj.1903
- Segura, H., Bayley, C., Fievét, R., Glöckner, H., Günther, M., Kluft, L., Naumann, A. K., Ortega, S., Praturi, D. S., Rixen, M., Schmidt, H., Winkler, M., Hohenegger, C., & Stevens, B. (2025b). A Single Tropical Rainbelt in Global Storm-Resolving Models: The Role of Surface Heat Fluxes Over the Warm Pool. *Journal of Advances in Modeling Earth Systems*, *17*(7). https://doi.org/10.1029/2024MS004897
- N. Takayabu, Y. (1994). Large-Scale Cloud Disturbances Associated with Equatorial Waves. *Journal of the Meteorological Society of Japan. Ser. II*, *72*(3), 433–449. https://doi.org/10.2151/jmsj1965.72.3_433
- Nakamura, Y., & Takayabu, Y. N. (2022). Convective Couplings with Equatorial Rossby Waves and Equatorial Kelvin Waves. Part I: Coupled Wave Structures. *Journal of the Atmospheric Sciences*, *79*(1), 247–262. https://doi.org/10.1175/JAS-D-21-0080.1
- Segura, H., & Hohenegger, C. (2024). How Do the Tropics Precipitate? Daily Variations in Precipitation and Cloud Distribution. *Journal of the Meteorological Society of Japan*, *102*(5), 525–537. https://doi.org/10.2151/jmsj.2024-028
- Sebastián Ortega, Hans Segura, Victor C. Mayta, et al. Convectively Coupled Equatorial Waves in a Global Storm-Resolving Model. *ESS Open Archive* . February 07, 2026.
- Judt, F., & Rios-Berrios, R. (2021). Resolved Convection Improves the Representation of Equatorial Waves and Tropical Rainfall Variability in a Global Nonhydrostatic Model. *Geophysical Research Letters*, *48*(14), 1–10. https://doi.org/10.1029/2021gl093265