

Answers to the reviewer

Review of the manuscript egusphere-2025-3407:

Evaluation of coupled and uncoupled ocean-ice-atmosphere simulations using icon_2024.07 and NEMOv4.2.0 for the EURO-CORDEX domain

by Vera Maurer, Wibke Düsterhöft-Wriggers, Rebekka Beddig, Janna Meyer, Claudia Hinrichs, Ha Thi Minh Ho-Hagemann, Joanna Staneva, Birte-Marie Ehlers, and Frank Janssen

The manuscript describes the evaluation of a new atmosphere-ocean coupled model and the two respective atmospheric and oceanic stand-alone models on the ERA5/ORAS5 period, on the EURO-CORDEX domain. The setup of the coupled model is well described. The results are often compared to observations, showing that they are reasonably realistic, and that the coupled model can later be used to perform historic and scenario simulations. The evaluation is more detailed on the oceanic part, which is very interesting per se, but I think it could be a bit enlarged on the atmospheric part.

I find the manuscript well written. The most serious restriction I have about the study concerns the period, which is often inconsistent with the 1979-2021 one mentioned on the first line of the abstract, and the 1979-2020 one on line 216, without any explanation. Indeed the other periods are:

- (1) 1979-2020 on the overview of the experiment (Line 216)
- (2) 1983-2020 for SST from ROAM-NBS except for marine heat waves (MHW from 1989 on)
- (3) 1981-2020 for SST from NEMO-NBS except for MHW (I understand 1981 is chosen for the comparison to Copernicus data, but the 1st years could be shown at least on fig. A4)
- (4) 1989-2021 for MHW (1989 coinciding with observations)
- (5) 1979-2020 for atmospheric variables
- (6) 1979-2020 in fig. 7 with also the skin temperature

If (2) is an effect of ocean spin-up (and even 1984 seems out of range in fig. 3 and A4), then the atmospheric variables should not be shown before.

(4): There is no reason to add 2021 for the MHW only, without showing this year for the other diagnostics.

So to my opinion, one option is to explain why the first years of ROAM-NBS and NEMO-NBS are not shown, and then only show the same period for the atmospheric ROAM-NBS; the second option is to announce a 1983-2020 evaluation simulation.

With this minor revision, I believe the article will be ready for publication.

Answer:

Thank you very much for this thorough review and the helpful comments.

The evaluation of the atmosphere over land is only shortly done as the respective atmosphere-only simulations with ICON-CLM are submitted to CORDEX likewise and will be evaluated elsewhere. It is shown here that the atmospheric variables over land are, on average, not strongly affected by the coupling. For the evaluation of the atmospheric part over the ocean area, unfortunately not many high-resolution reference data are available. The seasonal mean precipitation bias against GPM (which also contains data over the ocean) and of longwave radiation against CERES were added to Fig. A1 (also to consider the specific comments below). However, the focus of the evaluation is clearly lying on the ocean part, as stated in the second last sentence of the introduction: "As the ocean component is of additional benefit compared to most other CMIP6-CORDEX simulations, a particular focus is put on the evaluation of NEMO-NBS and the ocean part of ROAM-NBS."

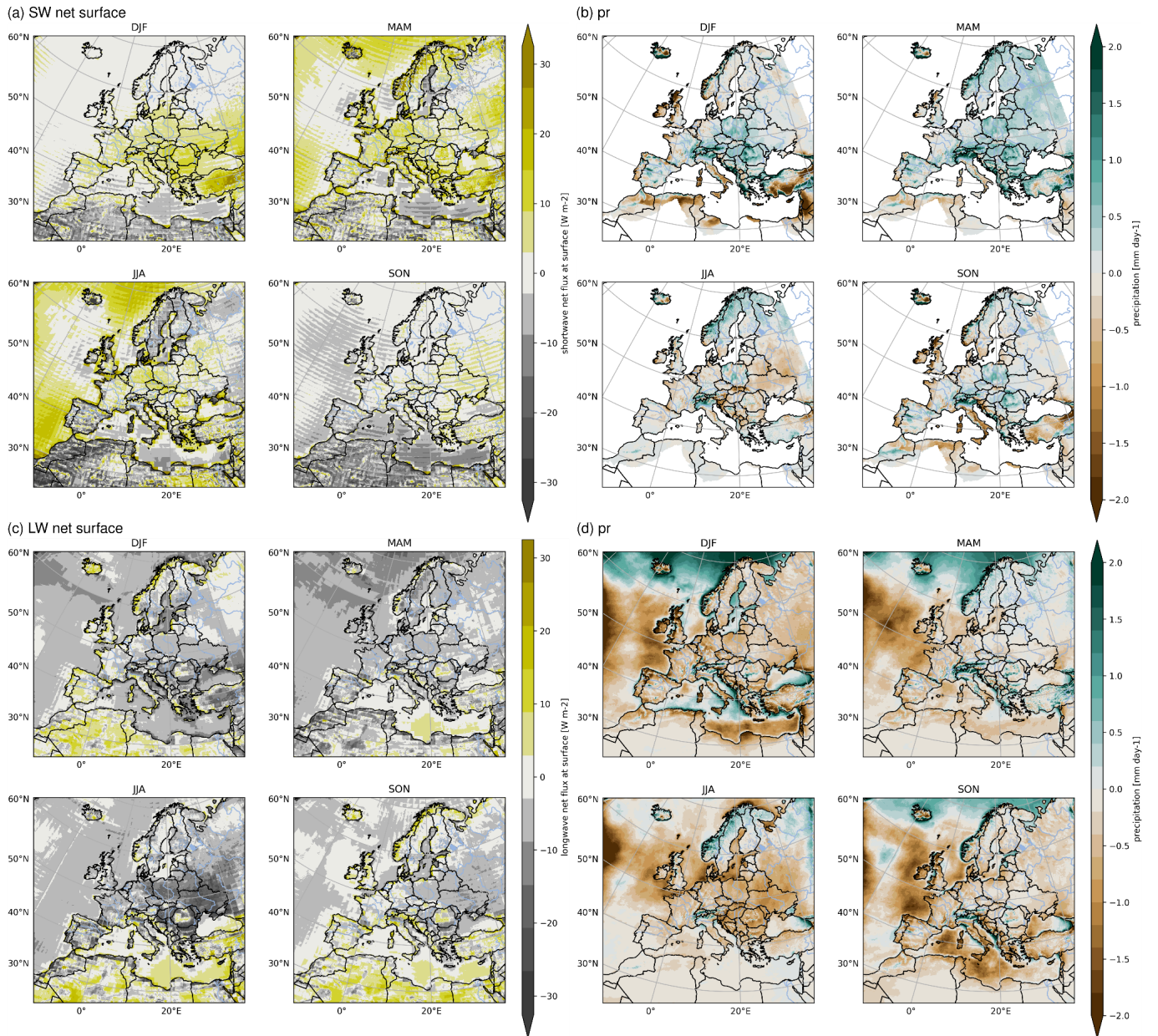


Figure A1. Seasonal mean biases of surface shortwave net radiation for ROAM-NBS against CERES (a), of precipitation against E-OBS, of surface longwave net radiation against CERES (c), and of precipitation against IMERG (d), all averaged over 2001–2020.

On the evaluation period:

Indeed, the evaluation period of 1979–2020 is not optimal for the ocean part as none of the reference data sets for the ocean is available for the whole period. However, we wanted to stick to the whole period, as it is the minimum period required for CORDEX, for which we will also deliver the data.

A more elaborate explanation was added in Sect. 2.4 “Overview of experiments”:

“[...] we are overall evaluating and comparing three simulations for the years 1979–2020, which is the minimum period required for CORDEX. However, especially for the ocean part, many reference data are available for shorter time periods only: SST and sea-ice data are available from September 1981, salinity

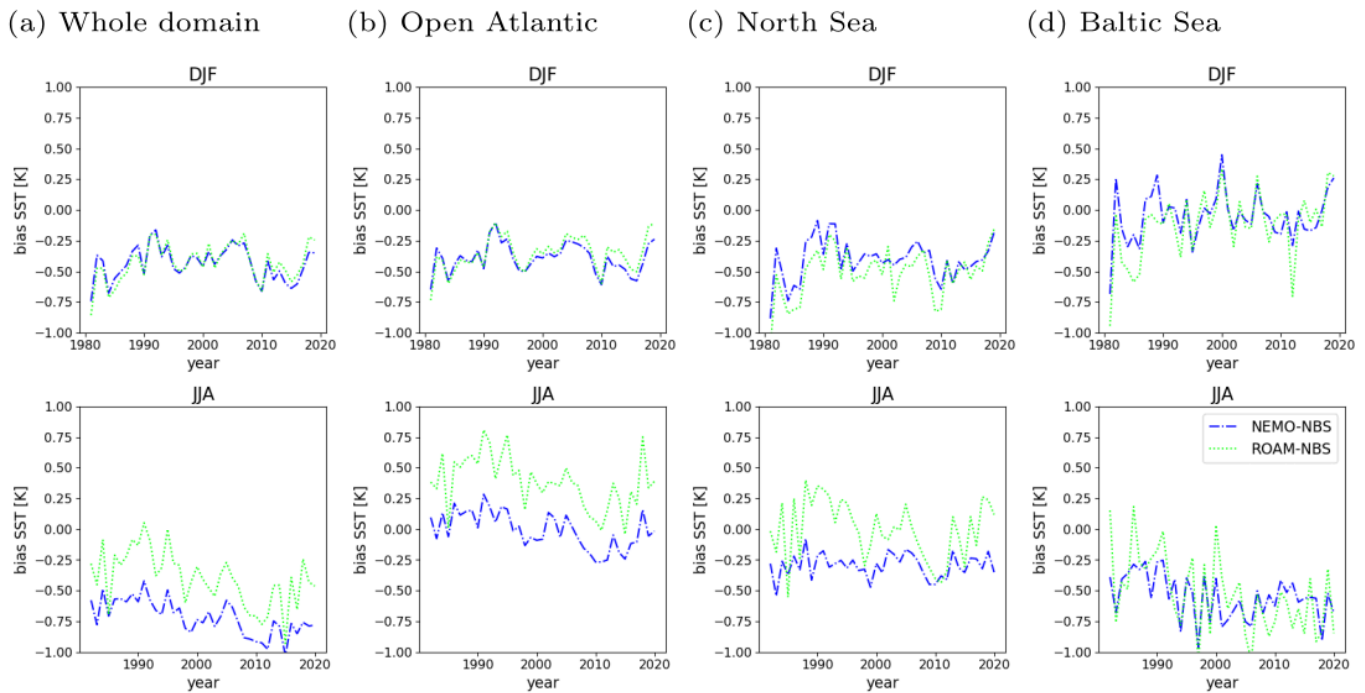
from 1993 and station data are very sparse before 1993. Therefore, the evaluated time periods had to be adapted in these cases. An overview is given in Tab. 2. For evaluations in which statistics from hourly data were calculated, shorter time periods were selected, partly due to limited data availability, partly to reduce the computational costs.”

Table 2. Overview of datasets used for evaluation; references are given in the respective sections; the full years were used if not denoted otherwise.

Dataset	Evaluated variables	Used in Sect.	Evaluated time period
Copernicus ESA SST CCI and C3S reprocessed SST analyses	SST, sea ice	3.1, 3.3.1, A2	1981–2020
ERA5	SST	3.2.1	1979–2020
E-OBS	tas, tasmin, tasmax, precipitation	3.2.1, A1	1979–2020 (2001–2020 in Fig. A1)
meteorological stations	hourly wind speed (10 m)	3.2.3	2011–2020
FINO1 wind measurement	hourly wind speed (100 m)	3.2.3	2004–2010
Copernicus Baltic Sea- In Situ Near Real Time Observations	ocean temperature and salinity (profiles)	3.3.2, 3.3.3, A2	1979–2020
Copernicus Multi Observation Global Ocean Sea Surface Salinity and Sea Surface Density	surface salinity	3.3.3	December 1993–November 2020
GESLAv3.0 observational data	sea surface heights (SSH)	3.3.4, 4.2, A2	2015–2019 (or selected events)
Baltic thalweg level 4 dataset	temperature and salinity (thalweg)	4.1	November 2014, February and March 2015
Copernicus Baltic Sea Physics Reanalysis	SST	4.3	1989–2020
CERES	surface radiation	A1	2001–2020
IMERG	precipitation	A1	2001–2020
Copernicus Atlantic- European North West Shelf- Ocean Physics Reanalysis	SST	A3	1989–2020

The missing years of ROAM-NBS in Fig. 3 (1981–1982) were due to a post-processing error for SST and not related to a masking out of a spinup effect. The figures 2,3, 10, 12 and A4 were corrected accordingly and the years 1979–1980 were added to the absolute time series in Fig. A4.

Below is the updated version of Fig. 3. We now also consequently masked out points with sea ice in the observations, since the SSTs in the Copernicus reanalysis are artificially set to -1.8°C in the regions covered by sea ice.



Please find below some more specific remarks, suggestions and corrections:
(answers in bold grey)

- abstract: it should be precised that it is an ERA5/ORAS5 simulation; **done**

- L12: in the abstract and other parts in the manuscript, the authors say that in the coupled model, the SST bias leads to biases in the ocean-atmosphere fluxes: I don't agree with that. A coupled model develops it's own state of equilibrium, SST and atmosphere-ocean surface are linked.

The formulation was modified: "Differences of fluxes and precipitation over the ocean between the coupled and uncoupled simulation are largely related to SST differences." See also the answers to the resp. comments below.

- L28: the delay behind global simulations is also true for regional downscaling of the atmosphere, because they are also forced by global simulations;

Exactly this was meant by the "downscaling chain". To make it clearer, the formulation was modified to: "Since the standalone ocean models are ideally forced by the output of the regional atmospheric models, the ocean simulations can only be delivered with a considerable delay compared to the global climate simulations due to the downscaling chain."

- L57: in the manuscript one can find either SI3 or SI3: it should be normalized;
done, it was changed to SI3 in the coupling section

- The new CORDEX Task Force on Regional Climate Projections (<https://cordex.org/strategic-activities/taskforces/task-force-on-regional-ocean-climate-projections/>) could be mentioned in the Introduction. → **added in line 27 and 28 within the introduction**

- L98: add Northern to “its adjacent seas” (there is no Mediterranean or Black Sea); **done**
- L137: I think that the tuning mentioned in L262-263 should be presented here
The values and the switch are given now in the model description section; however, we preferred to keep the main part of the explanation in Sect. 3.2.1 together with the discussion of the results, which is better understandable then
- L138: “For ROAM-NBS as well as for UDAG” : add “and ICON-CLM” → **done**
- L138: change icon-2024.07 for icon_2024.07 (as in the title and L56)
changed to icon-2024.07 everywhere to make it consistent with the registered model name on <https://github.com/WCRP-CORDEX/cordex-cmip6-cv> ; the hyphen instead of the underscore is also used on [icon-model.org](https://github.com/WCRP-CORDEX/cordex-cmip6-cv)
- L145: two ((; **done**
- L147: the lateral resolution is 2nm? → **changed to: the horizontal resolution is 2 nm**
- L148: could the authors precise more the vertical distribution of the 50 layers? Which is the depth of the first ones?
A more detailed description of the vertical distribution being dominantly sigma levels with a hyperbolic tangent transition following Madec et al. 1996, between top and bottom was added in section 2.2. The upper 16 levels are < 1.0m resolution within the whole domain.
- L149: 2 “chosen” → **corrected**
- L151: the reference is Madec et al. and not Gurvan et al. (to be corrected in the bibliography as well); **done**
- L182: Craig et al. is the reference for OASIS3-MCT_3.0: the manual of OASIS3-MCT_5.0 by Valcke et al. 2021 could be added; → **done**
- L224: is it the same ocean restart for ROAM-NBS?
**A sentence was added in the manuscript:
 “The restart field for 1 September 1978 from the spin-up simulation with NEMO-NBS was then used to start ROAM-NBS.”**
- L245 and followings: replace sea surface temperature by SST; → **done**
- L247: precise the Copernicus data period; → **done**
- L253: the authors chose the same seasons as in atmospheric studies, but for information the seasons usually used for ocean variables are JFM, AMJ, JAS, OND;
The chosen “standard” seasons DJF etc allow a better comparability with the atmospheric part.
- L263: cf. L137 above → **see above**
- L267-268: this parametrisation should be discussed in §2.2 or 2.4;

A section on the turbulence parametrisation including values for eddy-diffusivity and eddy-viscosity is added to 2.2.

- L289: don't put "surface air temperature": it is confusing with surface temperature above; → **ok**

- Figure 6: the "S" of seasonal is missing; → **corrected**

- §3.2.2: I think LongWave fluxes should be shown as well; and compare to the observations, not only ROAM-NBS to ICON-CLM, which can be biased as well;

Seasonal LW biases against CERES (2001-2020) and precip bias against GPM were added to the appendix (Fig. A3), the text was adapted accordingly; it is a problem to find good measurements (also note that both CERES sfc radiation and GPM precip are derived products and not direct measurements of the respective quantities) over the ocean, especially at appropriate resolutions.

- Fig.7 and in the text: I don't understand how the authors choose now the skin temperature: do they have it as an output of NEMO4 (and from 1979)? The SST comparison to ERA5 SST (which is the one imposed to ICON-CLM) is enough, as it was compared to Copernicus before.

Thank you for the hint, the nomenclature here was indeed a bit confusing. The reason is that we tried to stick to cmor variable names (tas, tasmin, tasmax, skt, lhfl, ...) where applicable. In this part of the evaluation, skt refers to ICON output (T_G = surface temperature) which is (over the ice-free ocean) identical to the SST, as we are not using a dedicated skin temperature parameterization over the ocean as e.g. IFS (it is available in ICON now, but not yet in the version 2024.07 which we are using). It is the same quantity as shown in Fig. 4. We now replaced all occurrences of skin temperature with surface temperature (and skt with Tsfc).

- L316: I don't like the formulation "the skt difference determines the sign of the flux..."; as I said before they are linked in the coupled model; besides they qualify their statements on L325;

The main aim of 3.2.2 is to show that there is a strong relationship between the SST biases (which are not identical, but still very similar in the coupled and uncoupled ocean parts, which was discussed in the first parts of Sect. 3) and the flux differences. As the SST in ICON-CLM is prescribed by ERA5, it is assumed to be more realistic than in ROAM-NBS and therefore, the SST difference between both is sometimes called "bias". The interpretation is that the precipitation and flux differences between the coupled and the uncoupled ICON-CLM can be largely explained by the SST differences. The resp. text passages were re-formulated and an introduction was added to Sect 3.2.2, which explains why the evaluation is done in that way.

- L389: "which" will be shown; **done**

- L393: a space is missing between "evaluation," and "the bias"; **done**

- L413-414: indeed the sea surface salinity is highly linked to the E-P-R flux; concerning the runoff, there are options in NEMO which for example propagate the runoff through the vertical, and also to enhance the vertical mixing at the river mouths: the choice here could be precised in § 2.2 or 2.4, or it could be discussed here;

An evaluation for sea surface salinity was done for both NEMO-NBS and ROAM-NBS and is only shown for ROAM-NBS due to minimal differences in the bias, therefore the E-P-R flux was not identified as the main reason for the bias. A one year test run using the ehype runoff data instead of the presented mix of observational and WaterGap runoff data showed promising results of an

approximately 1 psu smaller salinity biases along German coasts. Within the current NEMO-NBS and ROAM-NBS setups, the runoff is only applied in the upper layer and no enhanced treatment available in NEMO is applied. A sentence discussing these options was added in L413-414.

- L425: the restart of the simulation can sometimes also explain some biases in the deeper layers; Thank you for this comment, that is of course correct, especially within the enclosed basins in the Baltic Sea. A small addition referencing the initial data was added in L425.

- L433: add fig.8 for Cuxhaven;

A reference to Fig. 8 was added after mentioning the station Cuxhaven.

- L435: add “nearly” at “a higher correlation at all station” because it’s not the case for all; done

- L444: I’m not a specialist, but is “wind surge” appropriate here when the authors show only the SSH? Thank you for this attentive remark. Indeed the term wind surge is not quite appropriate here and was changed to storm surge throughout the complete document. In the SSH evaluation section mainly the detided SSH results with additionally removed mean sea level are presented in the figures as well as the table and therefore the storm surge is evaluated as explained in lines 431-433. The explanation is updated to a more detailed version and the text in this section is slightly adjusted for more clarity.

- Fig. 13: replace “blue” by “left” and “green” by “right”;

Thank you for the remark, this was included and the text in brackets changed.

- Fig. 14: what are the isolines for?

The isolines display discrete values of salinity and temperature for easier comparison of the stratification within the Baltic. In the caption of Fig. 14 and Fig. 15 a descriptive sentence was added.

- L465: it would be of interest to explain what is the “normal” situation of the inflow in the Baltic;

A section on the “normal” inflow situation in the Baltic was added in 4.1, including three new literature citations and a short definition of Major Baltic Inflow events.

- Fig. 16 c: add ICON-CLM?

ERA5 is included instead of ICON-CLM as this is used for the forcing of NEMO-NBS (both Fig. 16 c show then consistently observations, results/forcing from NEMO-NBS and results from ROAM-NBS)

- L517: I guess Lighthouse Kiel is also Leuchtturm Kiel of fig.8: choose the same name;

The station name was changed to Leuchtturm Kiel for consistency.

- Fig. 17: Orange stars, not green; it would be interesting to show the MHW from the beginning of the simulations, even if there are no observations;

Corrected, thank you for pointing this out. MHW were chosen not to be shown from 1979 due to missing observational data.

- L523: idem L517, and add Cuxhaven of fig. 8 for UFS German Bight (if I’m right?)

That is the station UFS Deutsche Bucht. We changed the station name here to German. Cuxhaven is a different station.

- L536: the authors must compute the correlations to say that;

The computed correlation based on linear regression for both models to observed MHW frequency, intensity and days at station Leuchtturm Kiel has been added to the text.

- Chapter 4: at the end of this chapter it would be nice to add a conclusion; →added:

“Overall, the evaluation of variability and extreme events shows that both NEMO-NBS and ROAM-NBS can generally reproduce but underestimate the Major Baltic inflow event, that they are able to represent storm surge events, and capture MHWs.”

- L582: the imposed runoff comes from observations, so do the authors think they might be too strong? cf. Remark for L413-414, and also there could be a discussion about coupling the runoff with a hydrological model: it is the best solution for future scenario simulations. Besides the E-P budget is also of much importance in the surface salinity;

Yes, since our test run using a different runoff data set (ehype) results in an approximately 1psu smaller salinity bias along German coasts than our evaluation runs using the mixed observational and model data set provided by BfG, we think that the runoff combined with prescribing it in only the upper cell is too strong, also cf. comment above (L413-414).

Yes, for the coupled historicals and scenarios, the online coupling with HD will be used. A comment was added in the text.

- L593: indeed the spin-up period is very important for the ocean, but also for the coupled model to reach its equilibrium, and as I said before concerning the period of the study, the authors must explain if they think that the 1st years of the ROAM-NBS simulation is in a spin-up phase.

After correcting Fig. 2 and Fig A4, it is much clearer that the coupled model does not show an additional spinup phase. Especially the absolute time series show this.

- L607: the authors don't show the computed correlations;

The computed correlation based on linear regression for both models to observed MHW frequency, intensity and days at station Leuchtturm Kiel has been added to the text in the MHW chapter.

- L633: add Fig. Before 6a → **done**

- Fig. A5: replace “left, mid, right” by “a, b, c” → **done**

- Fig. A7: replace “blue” by “left” and “green” by “right” → **done**