Response to anonymous Referee #1.

We thank the referee for the time spent on our manuscript and for these comments. We have realised that several aspects of our methodology were not so clearly explained and we believe that these comments will greatly improve the clarity of our manuscript. We believe that most of the points raised by the referee can be addressed by clarifying our method and we hereafter provide more details on how we will address these comments.

1) The distinction between "meltwater production" and "runoff" is unclear throughout the paper. This is especially the case as the authors investigate both runoff as a contribution to sea level changes and the "emergence of runoff conditions necessary for hydrofracturing". I find these two outcomes of 'runoff' a bit contradictory because if meltwater runs off into the ocean and contributes to sea level rise, then it can't induce hydrofracture events, which result from the pressure of ponded meltwater (i.e. Bell et al 2018). I think the authors should be careful to clarify this terminology throughout the paper. For example, in L27 the authors state: "The exact warming level needed to trigger important production of runoff on a given ice shelf depends on the amount of snowfall and on the snow/firn temperature and density (Donat-Magnin et al 2021, van Wessem et al., 2023)" However, these studies specifically look at meltwater production, which is the important component for hydrofracture on ice shelves, not runoff.

⇒ We agree that it was misleading and we will better define the quantities in the revised manuscript.

Runoff into the ocean is a negative contribution to the surface mass balance. It is produced if surface melt and/or rain rates are high enough (i) to bring the temperature of underlying snow and firn layers to the freezing point and (ii) to percolate and saturate the pore space in the snow and firn layers, which is sometimes referred to as firn air depletion (Pfeffer et al., 1991; Kuipers Munneke et al., 2014; Alley et al., 2018; Donat-Magnin et al., 2021). Surface melt and/or rainfall beyond the pore space saturation do not necessarily lead to runoff into the ocean. Liquid water in excess can alternatively be transported horizontally and stay on the ice shelf and form ponds. In some circumstances, these processes can trigger ice-shelf collapse through hydrofracturing (Bell et al., 2018; Robel and Banwell, 2019; Lai et al., 2020).

The MAR model does not represent the horizontal transport of meltwater within the firn or at the surface of ice shelves, but assumes that the meltwater in excess for individual snow/firn column is removed from the system, which is why it is usually referred to as "runoff" in this model (e.g., Agosta et al., 2019; Kittel et al. 2021). We will reword "runoff" as "liquid water in excess" when this is relevant throughout the manuscript.

2) Methodology

There are many parts of the methodology that are not well explained and remain unclear. This makes it challenging to assess the robustness of the methodology and therefore the results. I suggest a graphical figure detailing the methodology.

⇒ This is a good suggestion; we will make a graphical summary of our methodology in the revised version.

Below are some specific sentences or sections I did not understand or believe require additional detail:

a) L109: "To extend surface variables to a given local warming or cooling level, we always start from

20 different years (i.e. different values of Tref), then we average the 20 extended values." What does this mean? Does this essentially create a smoothed reconstruction?

⇒ We will better explain this part in the revised manuscript. To reconstruct the annual melt rate (or SMB minus runoff) at a given location and on a given year, an option is to start from the annual melt rate in another simulation and to correct this value to account for the air temperature difference (ΔT) between the two simulations. Given internal climate variability and uncertainty in the exponential fits, there is no reason to only use 1986 from simulation A to reconstruct 1986 in simulation B. So we reconstruct 1986 in simulation B twenty times, for years between 1976 and 1995 (and different ΔT values for every reconstruction) in simulation A, then we take the average of these 20 reconstructions.

We process similarly when we extend a simulation back in time: for example, 1950 (or any other year before 1980) is reconstructed 20 times from individual years in the corresponding MAR simulation that covers 1980–1999, and then we take the average.

This removes a part of the internal climate variability in the melt or SMB pattern of a given year (e.g., the influence of a strong synoptic event at a given year) but keeps the internal climate variability of air temperature in the reconstructed timeseries.

b) L114: "The a and b parameters are obtained through a least-mean-square-fitting of an exponential curve for SMB minus runoff on the one hand and the surface melt rate of the other hand." A supplementary figure off this exponential curve would be very helpful.

⇒ This will be added.

c) L115: "The fit is done on the original model grid as regridding does not preserve exponential relationships." I find this statement to be concerning. My understanding is that the exponential relationship between the two variables may weaken in the regridding but should remain? Further, if the exponential relationship parameters are fit on the original grid, and this relationship is not preserved in the regridding, is it then appropriate to apply this fit on the regridded data?

⇒ Yes, in principle, regridding weakens rather than removes the exponential relationship, but we will remove this statement because the simulations were interpolated conservatively from a 35 km grid to a 4 km grid, so the values on the interpolated grid are almost the same as on the original grid.

d) I am unsure how to interpret the r parameter (Eq. 3, 4, L121-124). Is this the percent of excess meltwater production that is converted to runoff (as opposed to that which ponds or refreezes)? If so, I expect that this value might be different on the grounded ice sheet vs ice shelves due to due to higher slopes on the grounded ice sheet.

⇒ This will be clarified with a better definition of runoff and liquid water in excess in a given firn column. The r parameter accounts for the fact the meltwater first needs to deplete the firn air content before being available for potential ponding (this is well defined in Donat-Magnin et al. 2021 and van Wessem et al., 2023). If there is a lot of snowfall, all meltwater ends up refreezing in the firn. If snowfall is low, it is easier to saturate the firn with meltwater (i.e., deplete the firn air content), after which any additional meltwater is potentially available for ponding. Therefore, if the melt to snowfall ratio exceeds r, it just means that there is locally a potential for ponding and hydro-fracturing.

We do not simulate the horizontal transport of meltwater. Nevertheless, when we explore the hydrofracturing potential at the end of the paper, we average quantities at the scale of an ice shelf so that our estimates are less affected by how meltwater has moved at the surface. Furthermore, we account for the sloping grounded basin upstream of the ice shelf by assuming that the meltwater in excess over the grounded ice flows towards the ice shelf.

e) Section 2.2.4: It is unclear how you reconstruct a cooler scenario from a warmer one. Do you use SSP5 to reconstruct SSP1? Or use warmer years as a reference time and reconstruct back in time? Also, what is the purpose of this and how will this be useful?

⇒ As written in section 2.2.4, "we evaluate the reconstructions of both SSP1-2.6 and SSP2-4.5 from SSP5-8.5". We will indicate the purpose of this in section 2.2.4, which is currently only shown in sections 3 and 4. This basically allows to have a large ensemble of SMB/melt projections based on a limited number of regional climate model (RCM) simulations. RCM simulations are numerically expensive and require 6-hourly CMIP model outputs which are not always available. So based on a single RCM simulation under SSP5-8.5, it is possible to emulate any colder scenario over the same period.

f) L157: "Similarly as in the previous subsection, each reconstructed year is the average of 20 reconstructions from a reference ranging from 10 years before to 9 years after the reconstructed year." I don't fully understand this sentence and have had to read it several times. It is perhaps related to the point mentioned in a) above?

⇒ Yes, see previous response. This will be clarified through the graphical summary.

g) Section 2.2.5: A graphic or schematic outlining the workflow here would be extremely helpful because I don't understand how the emulation is done. Additionally, Figure 3 is not very intuitive for me and should be better explained as I am unsure how to interpret it.

⇒ Yes, we will summarise the workflow in a graphical summary.

We do not really understand what to improve in the description of Fig. 3. The caption is quite explanatory, and this way of presenting results is also found in other papers (e.g., Figs. 4 & 7 in Barthel et al., 2020). Each diagonal direction represents one CMIP model, in black for the original MAR simulation forced by this CMIP model, and in colour for the emulation from another MAR simulation forced by a different CMIP model. The emulations from a given CMIP model are represented by small circles of the same colour and are linked together with a line to give a better overview of the general behaviour.

h) Figure 4: In panels e and f, the reconstructed runoff anomaly is too low, despite the melt anomaly (panel c and d) being fairly accurate for ice shelves and too high over the grounded ice sheet. Does this suggest some mis-parameterization in your method? Perhaps the wrong r value?

⇒ No, the reconstructed melting runoff anomalies are both slightly too high in Fig. 4, as indicated by the positive bias values indicated in brackets for panels c, d, e, f. If the referee was thinking about Fig. 2 and not Fig. 4, there is indeed a positive bias in the melt anomaly and a negative bias in the runoff (i.e., excess of liquid water beyond firn saturation). Decreasing the r value would clearly reduce the runoff bias in this figure, but not in others. This indeed shows that the parameterisation is not perfect. We will add a comment on this.

i) L205-208: For these seven simulations… apply a ramping transition between the two methods from 2101 to 2120." I don't follow what is being done here and again, I think a figure or something would help the reader understand the methodology.

⇒ Yes, this will be clarified in the graphical summary. There are two ways to emulate a simulation beyond 2100: extending from 2081-2100 (method described in section 2.2.3) and reconstructing from another simulation that covers 2101-2200 (here IPSL-CM6-LR--SSP5-8.5; method described in section 2.2.4 for other scenarios of IPSL-CM6-LR or 2.2.5 for other models). The extension from 2081-2100 allows some continuity with the existing 1980-2100 simulation but its fidelity decreases with time. Therefore, we do a linear combination of the 2 reconstructions from 2101 to 2120, with a weight of the extension from 2081-2100 that decreases linearly from 1 to 0 over the 20 years.

j) L280: The choice of a 100kg/m2/yr runoff threshold for triggering hydrofracture seems extreme and is not well-defended in the text. This is 50-67% less than the average meltwater production estimated prior to the collapse of Larsen B (200-300 according to the text). How was this threshold "empirically" chosen? Was it just based on Larsen A/B?

⇒ We can just tell that Larsen B was over the threshold when it collapsed, not just at the threshold given that mechanical conditions also need to be satisfied. Therefore, we took a threshold value smaller than the estimated production of meltwater beyond firn saturation at the time of collapse. Here we emphasize that this is not the meltwater production, this is the excess of meltwater beyond firn saturation, so any kg m^{-2} yr^{-1} is potentially available for hydrofracturing. In comparison, van Wessem et al. (2023) consider that melt ponds form when the melt over accumulation ratio exceeds r=0.7, i.e. a firn saturated with meltwater, which is equivalent to having a threshold just above zero (we will mention this in the revised manuscript). There is of course some sensitivity to this threshold, as written in our Discussion section:

"Last but not least, our estimation of the dates when runoff production becomes prone to hydrofracturing was based on a runoff threshold of 100 kgm^{-2} yr^{-1} over the ice shelf. This was motivated by the estimates of 200-300 kg m^{-2} yr^{-1} over Larsen B prior to its collapse, suggesting the need for a smaller or equal threshold. All the results presented in this paper are based on a threshold of 100 kg m^{-2} yr^{-1}, which is an empirical choice. Decreasing the threshold to 50 kg m^{-2} yr^{-1} shifts the dates by less than 10 years in the past for half of the ice shelves, but makes pre-industrial conditions favorable for 15% more ice shelves. Increasing the threshold to 150 kg m^{-2} yr^{-1} shifts the dates by less than 20 years in the future for half of the ice shelves. For some ice shelves, there can nonetheless be several decades of differences, again indicating that these dates are more indications than real projections."

k) L291: How do you define "likely" or "very likely"?

⇒ We use these terms as in the IPCC reports where they indicate the assessed likelihood of an outcome or a result: virtually certain 99-100% probability, <u>very likely</u> 90-100% probability, <u>likely</u> 66-100% probability. We agree that this needs to be better defined. Currently it is only defined in the caption of Tab. 4 in terms of percentiles "(likely range, i.e., 17–83th percentile) [very likely range, i.e., 5–95th percentile]". This will be clarified earlier in the revised manuscript.

3) Figure 1

L148: It seems a bit of a stretch to say that the extension of the RCM simulation is suitable for 25 years over the grounded ice sheet…. Really, there is just one year anomalously high SMB year at ~2125. Otherwise, the original MAR simulation and the reconstruction have opposite trends, even during this first 25 years.

In general, Figure 1 is concerning for me. It seems that this reconstruction method cannot be applied in a warming climate. However, the authors do apply some sort of reconstruction to obtain the results in Figures 6-12. How were these reconstructions obtained when Figure 1 demonstrates issues for applying this method in a warming climate? How can we trust the results presented here in light of Figure 1?

⇒ Our intention is not to assess the detail of the interannual variability over these 25 years, but to describe the overall bias that could affect an ice-sheet model in the long term. We nonetheless understand the need for a more objective statement and we will mention that the reconstructed SMB has a relative bias smaller that 12% on average over 2101-2125 for the grounded ice sheet and over 2101-2150 for the ice shelves (for r=0.5 and r=0.6).

We will also better emphasize in the text that <u>this method is not suitable for extensions to a much warmer climate</u>. In the final reconstructions used in section 3, we only use this method over a 20-year ramp down transition (2101–2120) as explained previously, and to extend the 1980–1999 period to colder conditions (1850–1979) which is less problematic. The main reconstruction of 2101–2200 for all models and scenarios is based on the MAR–IPSL-CM6A-LR simulation that covers 2101–2200.

4) Relation to previous studies

In general, this manuscript is lacking some references to and context within recent literature. For example, how do the results in section 3.3 add to or fit within the context of previous ice-shelf potential instability studies (i.e. van Wessem et al., 2023; Dunmire et al., 2024; Alley et al., 2018; Lai et al., 2020)).

Additionally, some reverences to previous AIS SMB studies are missing (e.g., Gorte et al., 2020, Noel et al 2023).

⇒ That's a fair point, we will add more comparisons to similar types of predictions.

Van Wessem et al. (2023) assume r=0.7 which was calculated by Pfeffer (1991) for a particular snow density and temperature. As Donat-Magnin et al. (2021), we find that slightly smaller values better match with the snow firn properties simulated by MAR in Antarctica. Van Wessem et al. (2023) then used air temperatures from the CMIP6 ensemble to extrapolate melt and accumulation based on a similar fitting method as the one used in our article (which is also similar to Donat-Magnin et al., 2021).

Dunmire et al. (2024) was published one and a half months after our submission. It has some similarities with our study, so we will definitely include a comparison to their results in the revised manuscript. They emulate firn air content which is exactly the same as the liquid water in excess that we emulate to estimate conditions prone to hydrofracturing, and they also use air temperature from the CMIP6 models, but the detail of their emulation and the training database are very different from our study.

We believe that most original aspects of our approach compared to these studies are (i) the time coverage back to 1850 and until 2200 while other studies cover ~1980-2100, (ii) the weighted ensemble members to have a more realistic representation of the equilibrium climate sensitivity than the raw CMIP6 data, (iii) the relatively large number of MAR simulations used to assess and calibrate our simple emulator.

In terms of results, our findings are in line with previous results: the Getz, Ross and Ronne-Filchner ice shelves do not reach melt conditions prone to hydrofracturing before 2100 whatever the emission scenario, while the ice shelves in the Peninsula and East Antarctica easily reach melting conditions prone to hydrofracturing before 2100 (Alley et al., 2018; Donat-Magnin 2021; van Wessem et al. 2023; Dunmire et al., 2024). As van Wessem et al. (2023), we estimate that Shackleton, Amery, Roi Baudoin and Larsen C ice shelves are prone to hydrofracturing in all scenarios before 2100.

Regarding the other articles mentioned by the reviewer:

- Lai et al. (2020) is already mentioned and provides the description of the mechanical conditions that can lead to hydrofracturing in the presence of meltwater beyond firn saturation.

- Noel et al. (2023) was published just one month before our submission so we did not have time to include it but we will do it in the revised manuscript. The aim of their statistical approach is to provide SMB and melt rates at a higher horizontal resolution than the original regional climate model. We will suggest that future work combines their approach and our approach to emulate large ensemble of SMB and melt rate projections.

- Gorte et al. (2020) provide an interesting evaluation of SMB in the CMIP5 and CMIP6 models, but assume that SMB can be approximated as precipitation minus evaporation/sublimation, i.e., they assume that runoff is negligible, which is not a good approximation beyond 2100 in the warmest scenarios.

5) Finally, the motivation for this work, and specifically how this method could be used in the context of ISMIP7, should be elaborated.

⇒ We will better explain how all this work can be useful to ISMIP7 or similar ice sheet projections. This reconstruction method is already used to provide the surface mass balance (SMB) in the ice-sheet multi-model projections of the PROTECT European project (https://doi.org/10.5194/egusphere-egu24-17095). In addition, our reconstructed melting beyond firn saturation could also be used to feed calving parameterisations or to impose dates of collapse as was done in ISMIP6.

The idea is that the CMIP models usually don't have a good snow/melt physics (some like IPSL don't represent the fate of meltwater in the firn), so that it is preferable to use regional climate models that were developed for polar regions (e.g., MAR, RACMO). However, the weakness of these regional models is the associated requirement for additional skills, additional inputs (e.g., 6-hourly 3-dimensional fields) and processing/computing time. Thanks to the emulation, we were able to provide the SMB for multiple CMIP models and scenarios beyond 2100 (which are difficult to process through a regional climate model given that the 6-hourly CMIP outputs were not saved after 2100 except for IPSL-CM6A-LR--SSP5-8.5 on our request). We were also able to reconstruct the SMB for CMIP models that were never processed through any regional climate models, which increases the diversity of SMB projections. Furthermore, our method can also be used to reconstruct SMB back to 1850 a period for which we currently have no regional climate simulations.

Minor comments

L7: "After correcting the distribution of equilibrium climate sensitivity of 16 climate models…" From just the abstract, it is unclear what this means.

⇒ We believe that the concept of "equilibrium climate sensitivity" is sufficiently known in the climate community to appear in the abstract, but we will define it in the text.

L7: "… we find a likely contribution of surface mass balance to sea level rise of 0.4 to 2.2 cm from 1900 to 2010…". It does not make sense that the contribution of SMB to SLR would be positive for this period so I'm assuming this is with respect to a reference period? Same for the SLR contribution ranges in the following lines?

⇒ Thank you, the entire sentence was wrong and will be replaced with: "we find a likely contribution of surface mass balance to sea level rise of **-2.2 to -0.4** cm from 1900 to 2010, and -3.4 to -0.1 cm from **2000** to 2099 under the SSP1-2.6 scenario, versus -4.4 to -1.4 cm under SSP2-4.5 and -7.8 to -4.0 cm under SSP5-8.5".

L25: "Hydrofracturing may strongly enhance the contribution of upstream glaciers to sea level rise." This is a bit misleading. The papers cited (among other work) indicate that the removal/collapse of ice-shelves (perhaps due to hydrofracture events) causes a speed-up of upstream glaciers, not just the hydrofracture event itself.

⇒ We will replace "hydrofracturing" with "ice shelf collapse resulting from hydrofracturing" in this sentence.

L41: "Because of these difficulties, only… which is generally insufficient to sample the CMIP model diversity." The "- when produced –" in this sentence threw me off a bit and I had to read it a few times to understand what was being said.

⇒ Ok, we will remove this part of the sentence.

L43: "… correct unrealistic Equilibrium Climate Sensitivity…" A brief explanation for this concept would be helpful here.

⇒ Yes, we will add "The ECS is the long-term increase in global mean surface temperature after a doubling of CO_2 concentrations. It can be used to characterise the sensitivity of individual climate models to changes in the radiative forcing."

L45: "Over the years, Antarctic Ice Sheet modellers have often scaled their best estimates of present-day accumulation to temperature anomalies from the CMIP models…". This sentence fragment is unclear to me.

⇒ We will replace with "Over the years, Antarctic Ice Sheet modellers have often scaled their best estimates of present-day accumulation to temperature anomalies from the CMIP models (e.g., based on the Clausius-Clapeyron relationship as in Gregory and Huybrechts, 2006)".

L67: "The surface mass balance and melting… Donat-Magnin et al (2020) and Kittel et al (2021)." I think a brief explanation of the results of these papers would be helpful here. I am left wondering: And how does MAR do in comparison to observational products?

⇒ In the revised manuscript, we will summarise the main results of the melt and SMB evaluations presented in these articles:
"Agosta et al. (2019) used firn-core SMB estimates to evaluate a MAR configuration covering the entire ice sheet: their SMB spatial pattern was well captured and the mean bias was 4%. Donat-Magnin et al. (2020) compared their MAR configuration of the Amundsen Sea sector to automatic weather stations, airborne-radar and firn-core SMB, melt days from satellite microwave, and melt rates from satellite scatterometer. They obtained good results for near-surface temperatures (mean overestimation of 0.1°C), near-surface wind speeds (mean underestimation of 0.42 ms^{-1} and SMB (local biases lower than 20%). The mean surface melt rate over the Amundsen Sea region was underestimated by 18% but the interannual variability was well captured for both melt rate and the annual number of melting days. The aforementioned MAR simulations were forced by atmospheric reanalyses and Kittel et al. (2021) showed that the Antarctic SMB simulated by MAR forced by climate models was close to MAR forced by the ERA5 reanalysis over the recent decades".

L101: Assuming that all precipitation is entirely made of snow is a big assumption to make, especially for projections that extend to 2200 in high-emission scenarios. The impact of this assumption should at least be discussed somewhere in the paper.

⇒ This will be added.

Donat-Magnin et al. (2021) find that mean rainfall in 2081-2100 under the RCP8.5 scenario is still one to two orders of magnitude smaller than snowfall over ice shelves in the Amundsen Sea. The only regional climate simulation from MAR until 2200 (the one forced by IPSL-CM6A-LR--SSP5-8.5) does simulate the actual rainfall/snowfall distribution and still has a positive surface mass balance over ice shelves in 2200, which indicates that snowfall is still dominant over rainfall. Furthermore, as discussed in Appendix B of Donat-Magnin et al. (2021), a greater rainfall rate than melting rate is required to saturate the firn (because melting both removes snow and depletes its air content, while rainfall just depletes the air content). The emergence of rainfall nonetheless contributes to decrease the fidelity of our method towards the end of the 22^{nd} century.

L132-134: Should this really be interpreted as a 'mass loss rate' if Figure 1 shows anomaly values with respect to a reference period? The SMB for the reference period is positive (although the specific reference period value from MAR should be mentioned somewhere in the paper). Even though the line in Figure 1b decreases throughout the timeseries, mass loss doesn't occur until it reaches the negative magnitude of the reference period. For example, if the reference period SMB for ice-shelves ~500 Gt/yr, then surface mass loss doesn't really occur until approximately 2120 (when the time series reaches -500 Gt/yr).

⇒ Yes, this will be reformulated. It is clearly shown in Fig. 8 that the current SMB over ice shelves is ~500 Gt/yr, so an anomaly of 2200 Gt/yr is just a negative SMB of 1700 Gt/yr (and an actual mass loss depends on basal melt, calving and the ice flux across the grounding line).

L138: "… although this is still an improvement compared to the original IPSL-CM6A-LR outputs". I think it would be very interesting to have these original ESM timeseries plotted in Figure 1 as well.

⇒ We agree, this will be added.

L123: "… covering the aforementioned range, i.e. 0.5 to 0.9." This range is different from that mentioned before (0.6-0.85, L98).

⇒ We will reformulate. The 0.5–0.9 range covers the 0.60–0.85 range discussed in previous studies.

Section 2.2.3: Somewhere in this section it should be clarified that Equation 4 is used to do this reconstruction.

⇒ This will be clarified as suggested.

L179-181: "The realistic SMB reconstructions derived from MAR-ACCESS1.3 are mostly compensations between overestimated melt and overestimated accumulation". Why do you say this?

⇒ Because the melt rates emulated from MAR-ACCESS1.3 are largely overestimated, as seen by the large green pentagon in Fig. 3c, which means that the realistic SMB is due to underestimated accumulation. This is to argue that emulations from MAR-ACCESS1.3 are outliers despite realistic emulated SMB over the grounded ice sheet.

L210: "… with a 20-year transition." What does this mean?

⇒ This will be clarified, see previous responses.

L238: Figure 8 is mentioned before Figure 7

⇒ This will be corrected in the revised manuscript.

L253: What does "weaker SMB" mean?

⇒ This will be replaced with "lower SMB".

L265: "Spatially, a net surface mass loss arises for several ice shelves…" I find this to be a bit misleading since Figure 7 shows SMB anomalies from a reference period, not absolute SMB. Is there actually a net surface mass loss or is it just lower than the reference period?

⇒ This will be reformulated. The anomaly is compared to the present-day SMB in Fig. 8.

L294 – It should be mentioned that George VI ice shelf has compressive stresses which do not promote hydrofracture occurrence (Labarbera et al., 2011)

⇒ We will mention this.

L 311: What is the A1B scenario?

⇒ This is a scenario that was used in CMIP3 and IPCC-AR4 which is cited in this sentence.

Technical corrections:

⇒ Thank you, these will be corrected.

L33: progresses à progress
L62: Citation needed for the pore close-off density.
L180: “ooverestimated”
L185: “emulation” à emulations
L234: “scenario” à “scenarios”
L 254: End parenthesis after “(Fig. 9.”
L262: “to the exception” à “with the exception”
Figure 8 caption: “same as Fig. ?? is also shown”