Response to the editor and the referee — second revision

Recent thermohaline shifts in the Adriatic Sea and their potential consequences

Elena Terzić and Ivica Vilibić
Ocean Science, egusphere-2026-3080

October 15, 2026

Dear Dr Hoppema,

Thank you for the decision and for the referee's careful second reading. We have taken up every comment in the annotated manuscript. Of the 50 annotations, five express approval, one (R-47) is a stray single character, and 44 ask for a change; all 44 have been incorporated, and each is answered individually below. None of the changes made in answer to the referee affects a result or a conclusion. They clarify wording the referee found unclear or jargon-like, define the acronyms used in Table 1 and the text, add the Argo acknowledgement and data statement, put the Zenodo links into the text, and take up two presentational suggestions: an introductory paragraph and a methods figure for the two detection criteria (Sect. 2 and the new Fig. 3, so that the former Figs. 3–6 are now Figs. 4–7), and the removal of the 5 m summer bin from the difference row of Fig. 1, which the text already identified as a sampling artefact.

Two small changes were made on our own initiative while working through the comments. The data availability statement said that the 2020–2025 part of the record is supplied by Argo alone; the Croatian monitoring casts to 2023 also contribute, and the sentence now says so. And, following the referee's remark that a dash before a number can be read as a minus sign, we rephrased four further places where that could happen. Cross-references were added so that the main-text figures, the appendix figures and the tables are each first cited in numerical order.

While revising we also re-checked the text against the analysis code and the data. All analyses were re-run on a re-checked version of the dataset. Several values changed slightly as a result; they are marked in the tracked version, and no conclusion changed. In Eq. (2) the numerator and denominator of the Turner angle had been printed the wrong way round; the printed form is corrected, the computation was right. The checks also led to the corrections below.

1. **Upper-layer temperature ratio (Sects. 2.4 and 3.3, Abstract, Conclusions).** The first revision stated that above 200 m the southern Adriatic warms more slowly than the Mediterranean median (layer ratios 0.5 and 0.7) and called this robust. It is not. The layer values are medians over depth bins: at 5–10 m the southern Adriatic warms at or above the Mediterranean median rate, and the low values come from 30–100 m. Most of that difference reflects where the casts were taken: casts over water deeper than 800 m make up about 40 % of a monthly mean until 2014 and about 75 % from 2015, and at 30–100 m the open basin is on average 0.4–0.7 °C colder than the shelf in the same month. With only open-water casts used in every sub-basin, the layer ratios are 0.79 and 0.86 and not distinguishable from one. The text now states that there is no thermal amplification above 200 m and that slower warming cannot be established; the sentences that defended the slower warming (the test without the shallowest bins, the ranking among sub-basins and the quadratic fit) are removed. The two ratios of Table 2 that lie below one are given to two significant figures (0.55 and 0.72), and the caption carries a note. From 250 m down the temperature ratios do not depend on the cast population (1.5 and 6.0 on open-water casts, against 1.5 and 6.2). For salinity the text now says that the ratio at 75–200 m (1.3) is not distinguishable from one and that the near-surface ratio is 3.4 on open-water casts, against 4.3.

2. **Change-point test (Sects. 2.5 and 3.5, Fig. 4).** The level shift in surface saline lakes is now tested on the casts themselves, with a logistic model of a step, the season and the platform (Sect. 2.5), instead of the Pettitt test on the annual series. In an annual series every

year has one vote whatever its cast count, and the test cannot separate the arrival of Argo from the change, since floats detect a surface saline lake about twice as often as a ship cast in the same year and season. With season and platform allowed for, the step falls after 2014 in all casts and in the Argo profiles alone, and after 2016 in the ship casts alone, where the years 2006–2016 cannot be told apart; in all casts it takes warm-season occurrence from 4.9 to 38.4 % ($p = 0.02$). Fig. 4A marks the step, and the Abstract and Conclusions say that the shift falls after 2014 in ship and float records alike. The regime shift is therefore confirmed, after 2014 as in the first revision. Casts that cannot be evaluated by the detection, the single-level underway records and 100 SeaDataNet casts without a usable temperature, are now left out of the cast set of Sect. 2.6; the SeaDataNet detection rate of Sect. 3.5 is 0.56 % (previously 0.58 %), and the resolution-matched set behind Fig. 4 is unchanged.

3. **Double salinity maxima (Sects. 2.7 and 3.4).** A salinity maximum is a level saltier than the levels above and below it, and only maxima that stand at least 0.010 above the lowest salinity on the way to a higher maximum are kept. Archived salinities carry three decimals, so on the flat top of a salty layer two levels a few metres apart often hold exactly the same value, with a dent of a few thousandths between them. While re-deriving the numbers behind R-13 we found that the peak-finding routine treats such levels as two maxima and, because neither is higher than the other, credits each with the full height of the layer; the two-per-cast cap then kept these twins and dropped the genuine second maximum, in about a fifth of the casts with two maxima. Neighbouring maxima are now counted as one unless the salinity dips by at least 0.010 between them, the same threshold that separates a maximum from a ripple, so no parameter is added. Fig. A3 shows that the two results of Sect. 3.4 hold with this rule, without it, with a minimum separation of 20 m instead, and after smoothing the profiles; its new panel (c) shows the occurrence of two maxima under each setting, the one quantity on which the paper draws no conclusion. The conclusions are unchanged: the share of maxima in the 100–200 m band still falls from about 41 to 16 %, the upper maximum still lightens four to five times as much as the lower one, and the occurrence of two maxima still shows no consistent change. The numbers that describe the two maxima did change, in every case in the direction the paper already reported: the within-cast distance grows from a median 122 to 210 m (previously 92 to 113 m), the lower maximum sinks from 205 to 333 m (previously 201 to 325 m), the upper one rises from 82 to 64 m (previously 100 to 78 m), the share of lower maxima at 400 m or deeper rises from 16.4 to 43.4 % (previously 16.1 to 41.9 %), and the salinity contrast of the pair is 0.030 and 0.026 in the two periods (previously 0.021 in both). Sect. 2.7 also now says that the casts reaching 700 m with both variables date from 1972–1998 and 2010–2025: in SeaDataNet the deep casts of 1999–2009 are bathythermographs or two CTD surveys with temperature flagged bad, and the IOR database holds only 16 such casts, in 2000–2009. The 21 shipborne casts of 2010–2019 in the IOR database that reach 700 m, with the 26 MSFD monitoring casts of 2019–2023 the only ship casts of the float years, place 23 % of their maxima in the 100–200 m band (20 % for the MSFD casts), against 15 % for the floats of the same years and 41 % for the ship casts of 1950–1999; adding the IOR and MSFD casts to the two populations moves the pooled shares of Sect. 3.4 by at most 1.6 points, the medians by at most 2 m and the confidence limits by at most 10 m, and changes no conclusion. Fig. 5, Fig. A3 and the corresponding numbers in Sects. 2.7, 3.4, 3.7 and the Conclusions have been updated; Fig. 3b is unaffected.

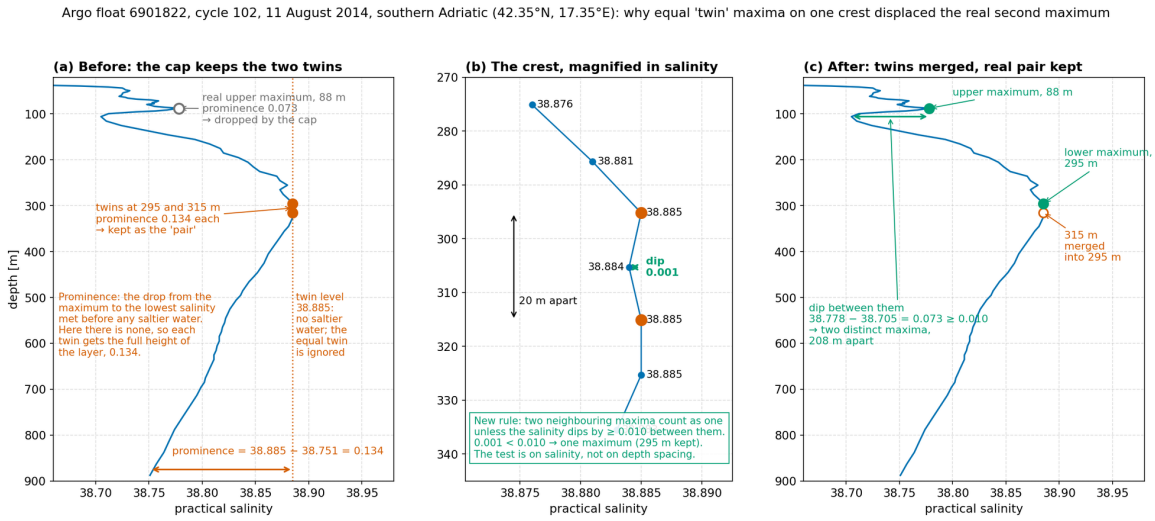


Figure L1. The twin-peak case on a measured profile (Argo float 6901822, cycle 102, 11 August 2014). (a) Before: the two equal maxima of one crest, at 295 and 315 m, each received the full prominence of the layer, so the two-per-cast cap kept them as the pair and dropped the real upper maximum at 88 m. (b) The crest magnified: the dip between the twins is 0.001. (c) After: the twins count as one and the pair is the real one, 88 and 295 m.

4. **Argo-only analyses (Sects. 2.8, 2.9, 3.6, 3.7 and 4.2).** The Argo set behind these sections has been extended from September to December 2025 with the profiles of the January 2026 retrieval that the climatologies use: 757 profiles from 18 floats (previously 694 from 17). Table 4, Table 5, Fig. 7 and the numbers of Sect. 3.7 follow; the conclusions of Sect. 3.7 are unchanged, with the rise in vigour below the core now 28.0 to 39.0 % at 400–600 m and 10.7 to 37.4 % at 600–1000 m. The window of the steric trends, January 2014 to June 2024, holds 116 monthly values, not 121; Table 3 and Fig. A2 are unchanged. Sect. 3.6 now gives the net trend of each longer window, −1.2, −1.4 and −2.2 mm yr^{-1} on the whole record from April 2010, and on January 2012 and January 2014, to December 2025, none distinguishable from zero, and Sect. 4.2 and the Conclusions say that the negative sign of the net is a property of the 2014–2024 decade. The comparison with the long climatological baseline at the end of Sect. 3.6 was recomputed on the climatologies without the float copies (+7.4, −7.0 and +0.4 mm yr^{-1} over 21 years; previously +7.6, −6.8 and +0.8 over 19.8 years).

5. **Smaller corrections, with no change in any conclusion.** Sect. 2.3 described the uncertainty of the climatological differences as a resampling bootstrap over cruise-days; the intervals of Figs. 1–2 are computed analytically, with all casts of one source taken on one day counted as a single independent observation and, for potential density, by Monte Carlo propagation, and the text now says so. The vigour of the salt-finger regime is the fraction of finger-regime levels with $R_\rho < 2$, not of all levels (Sect. 2.9, Table 5, Fig. 7). The deep warming of Sect. 3.1 is 0.5–0.9 °C below 200 m (the 0.4–0.6 of the first revision was carried over from the 2000–2019 analysis), and warming is significant at 83 of the 88 depth–season combinations (94 %; the 97 % of the first revision counted two significant negative differences). In Sect. 3.2 the shares of significant combinations count warming and salinification only (79 and 58 %, previously 88 and 71 % of either sign), and the section names the station sampled several times a day in 1999–2004 that supplies 54 % of the modern northern casts. The eight-level floor of Sect. 2.7 sets aside 479 of 5069 casts (9.4 %), not 469 of 5761. The caption of Fig. 2 quoted cast counts that mixed the two periods, and the caption of Fig. 1 now describes the count panel correctly. Sect. 3.7 adds two checks: on annual means only the increase in vigour and the fall of N^2 in the deep layer remain significant, and with gradients taken over 25 dbar the vigour still rises below the core and not within it; the section now rests the contrast between the core and the

layers below it on vigour, and Sect. 4.2 qualifies the agreement with Amorim et al. (2024) accordingly.

Panel (d) of Fig. 7 now shows the vigour of the salt-finger regime by layer and year instead of the mean extent before and from 2017. The extent is close to saturation in the deep layer and changes little in the core, so that panel repeated the least informative row of Table 5, and a two-period mean hides the course of the change: the deep layer rises steadily from 2013 to a peak in 2021–2022 and relaxes in 2025, while the 400–600 m layer was already vigorous in 2013–2014. Sect. 3.7 now states which contrasts depend on the split year (the 400–600 m one does, the deep-layer one does not, and the core shows no change for any split) and quotes the annual values. Table 5 and its tests are unchanged.

A few sentences in Sects. 2.3, 2.4 and 2.7 were also reworded in plainer language. All changes are marked in the version with tracked changes against the first revision, provided together with the clean manuscript.

With best regards,
Elena Terzić and Ivica Vilibić

Response to the referee

We are grateful for a close and constructive second reading. Comments are numbered R-01 to R-50 in the order in which they appear in the annotated PDF, with the page and line of the first revision. Because a figure has been added in Sect. 2, Figs. 3–6 of the first revision are now Figs. 4–7; figure numbers below refer to the revised manuscript, with the former number in brackets where it differs.

R-01 (p. 1, ll. 14–18, Abstract)

good new summary with improvements to the previous version. I like the new extended version about data and methods.

Response. Thank you.

R-02 (p. 2, l. 26)

good introduction to the relevant processes

Response. Thank you.

R-03 (p. 3, l. 62) — "Argo profiles come from Euro-Argo ERIC"

It sounds a bit odd that the data were provided by the ERIC. Anyway Argo would like that users acknowledge the data . . . could you please add the statement to the manuscript at the end

Response. Reworded and added. Sect. 2.1 now says that the Argo profiles were retrieved through `argopy` from the Argo Global Data Assembly Centre. The statement the referee quotes is in the Acknowledgements, where it replaces Euro-Argo ERIC in the list of data providers, and the data availability statement points to the International Argo Program.

R-04 (p. 3, l. 55) — Sect. 2, "Data and methods"

there is a lot of detail in the following subsections and maybe it could help the readers if some introductory statements could be made . . . often a schematic graphic says more than a lot of words to explain the methods for the SSLs and the double maxima.

Response. Both suggestions are taken up. Sect. 2 now opens with a paragraph stating what each group of analyses needs from the data, and a new Fig. 3 shows the two detection criteria on measured profiles: (a) an Argo profile containing a surface saline lake, with the criteria of Sect. 2.6 annotated; (b) a cast carrying two salinity maxima, with the quantities of Sect. 2.7 marked.

R-05 (p. 3, l. 67) — "analysis code by Terzić and Vilibić (2026b)"

can you please add the link to the zenodo here

Response. Added, here and in the data availability statement (R-39). The record cited so far holds the Mediterranean-wide processing pipeline only, so a third record with the analysis code specific to this paper (Terzić and Vilibić, 2026d, 10.5281/zenodo.23016822) has been deposited and is cited alongside it.

R-06 (p. 4, l. 119) — "To make the amplification claim quantitative"

maybe rephrase since such a claim has not yet been made at this point in the manuscript. Or just delete that part of the sentence

Response. Rephrased. The paragraph now opens with "Adriatic rates of change are compared with those of the rest of the Mediterranean."

R-07 (p. 5, l. 121) — "the only difference between the two sides of the comparison is the ocean rather than the method"

It is not quite clear what is meant here

Response. Replaced by "so that the Adriatic and Mediterranean rates are directly comparable".

R-08 and R-09 (p. 5, ll. 125–126) — dashes around "5–50, … 700–1200 m"

please delete, looks a bit like a minus sign

Response. Both dashes are removed; the layers are given in parentheses. Four further places where a dash stood next to a number were rephrased in the same way.

R-10 (p. 5, l. 129) — "The window is 2000–2025, matching"

My suggestion to rephrase: The time window 2000–2025 is matching

Response. Adopted: "The time window, 2000–2025, matches the climatologies".

R-11 (p. 5, ll. 134–135) — "Because it searches the years itself, the result is not biased by our picking a year by eye."

sounds a bit like jargon, please rephrase

Response. The sentence no longer exists: the change-point test has been replaced (cover letter, item 2), and Sect. 2.5 now describes the model in plain terms.

R-12 (p. 6, l. 161) — "Maxima are found by peak prominence in salinity"

sounds odd

Response. Rewritten: "A maximum is a level saltier than the levels above and below it. Most such levels are small ripples, so each is graded by its prominence: the drop in salinity from the maximum to the lowest value that must be crossed to reach a higher maximum in the same profile (Fig. 3b). A maximum is retained when its prominence is at least 0.010, ten times the precision of the archived values."

R-13 (p. 6, ll. 170–171) — the two sentences on the cap of two maxima per cast

Unclear sentence

Response. Rewritten: "15 expendable-CTD casts of one cruise in May 1995 return 162–266 each and would, without the cap, supply 59 % of all 1950–1999 detections among casts reaching 700 m. The cap falls on both periods alike: 38.8 % of those casts in 1950–1999 and 39.2 % in 2000–2025 carry more than two maxima." The values are those of the corrected detection (cover letter, item 3).

R-14 (p. 7, l. 189) — "The anomaly form is required rather than merely preferred"

unclear

Response. Rephrased: "The integrand must be the anomaly Θ' rather than Θ itself …"

R-15 (p. 7, l. 211) — "Fingers transport appreciably"

transport what? salt? heat?

Response. Salt; the object is now given here and in the caption of Fig. 6 [5], and the clause of Sect. 3.7 in which the verb also stood alone has been removed.

R-16 (p. 7, l. 212) — "A column"

water column?

Response. Yes; "A water column can therefore be entirely finger-favourable . . . ", and likewise earlier in the same section.

R-17 (p. 8, l. 217) — "They are evaluated in three layers, because the layers behave differently and averaging across them cancels their changes out."

It is not quite clear what is meant here

Response. Rewritten: "Both quantities are evaluated separately in three depth layers, because their changes differ between layers and partly cancel when averaged over the whole column."

R-18 (p. 8, l. 221) — "and carries no result here"

????

Response. Deleted.

R-19 (p. 8, ll. 237–238) — "We therefore do not interpret the uppermost summer bin"

maybe it would be more pedagogic to remove the difference bin in the plot and just explain in the text and figure caption why this has been done

Response. Done. The 5 m summer bin (difference and confidence band, all three variables, main panels and insets) is removed from the difference row of Fig. 1, and Sect. 3.1 and the caption say that it is omitted and why. Sect. 3.1 also states that the 10 and 20 m summer differences (-0.44 and $-0.17\,^{\circ}\mathrm{C}$) are not interpreted either, and gives the uneven spread of the summer casts within the season as a second reason.

R-20 (p. 8, l. 240) — "Potential density has the weaker signal"

than what?

Response. ". . . a weaker signal than either temperature or salinity".

R-21 (p. 9, l. 248) — "The northern shelf is where the clustering correction is largest"

maybe add a speculation why that is the case.

Response. Added, and on checking the cast index the reason turned out to be specific: one station at 13.7°E, 45.7°N was sampled several times a day in 1999–2004 and supplies 54 % of the 2000–2025 northern casts, "so that consecutive casts share one synoptic state (Sect. 2.3)".

R-22 (p. 9, ll. 260–261) — "The platform change still acts inside that window . . . "

still sounds a bit too much as jargon. Consider rephrasing

Response. The sentence has been removed: Sect. 3.3 was rewritten (cover letter, item 1) and now states in plain terms how the change from ship casts to Argo floats affects the upper-layer ratios.

R-23 (p. 9, l. 266) — "Ranked against all 28 sub-basins on this pipeline"

what is meant here? what are the 28 sub-basins? I assume it refers to the nomenclature of Bruno but has not been explained before

Response. The referee's assumption is right. Sect. 2.4 now introduces the 28 Mediterranean sub-basins used by Terzić and Vilibić (2026c), labelled after Manca et al. (2004), three of which are Adriatic, and gives the Mediterranean reference as the median over the other 25. The ranking sentence itself was removed in the rewriting of Sect. 3.3.

R-24 (p. 9, l. 271) — "deep-specific"

does not sound like the right word

Response. Removed. The sentence now says that the southern Adriatic is "warming faster than it only below about 300 m", "it" being the Mediterranean median.

R-25 (p. 10, l. 290) — "This band is where the detections themselves concentrate and is not a definition of the core"

unclear

Response. Rewritten: "This band is not a definition of the core but simply the interval in which the historical detections cluster; it lies immediately above the classical 200–400 m range."

R-26 (p. 10, ll. 300–301) — "These pooled medians describe the two populations and are not the separation within a cast . . . "

very unclear

Response. Rewritten: "The difference between these pooled medians need not equal the median within-cast distance: the pooled medians describe where each maximum typically lies, the within-cast distance how far apart the two are in a typical cast."

R-27 (p. 11, l. 314) — "12.14 % for MSFD CTD"

why are you treating the MSFD CTDs separately. They are not done by a different CTD?

Response. They are ordinary CTD profiles, listed separately because they come from a separate archive (the Croatian national monitoring programme, which is not part of SeaDataNet), as in Table 1. The text now says that the sources are archives rather than instrument types and that they cover different years, so that the spread of detection rates mixes period and vertical resolution; the degradation experiment that follows isolates the latter.

R-28 (p. 11, l. 313) — "IOR"

what does IOR stands for

Response. Institute of Oceanography and Fisheries; defined at first use in Sect. 2.1, together with RBI, CTD and MSFD.

R-29 (p. 12, l. 363) — "from a single platform"

platform type? you probably did not mean a single instrument

Response. Correct; now "from Argo floats alone".

R-30 (p. 12, l. 375) — "separate cleanly"

differ? diverge?

Response. Now "The double-diffusive diagnostics behave differently above and below the base of the classical LIW core".

R-31 (p. 13, ll. 378–379) — "Only stratification weakens, median N^2 falling by 16 %."

odd fragment of a sentence

Response. Rewritten: "The only significant change within the core is a weakening of the stratification, the median N^2 falling by 16 %."

R-32 (p. 13, ll. 389–390) — "and the core it has moved out of is the layer in which nothing changed"

odd wording

Response. Rewritten, and since then qualified by the numbers of Sect. 3.4: "Fingering intensified in the layers into which a growing share of the deeper maxima has moved, and did not change in the core, where their share fell from 37 to 23 %."

R-33 (p. 14, l. 428), R-34 (p. 14, ll. 443–444) and R-35 (p. 15, l. 473, Limitations)

good / good / good summary

Response. Thank you.

R-36 to R-38 (p. 16, ll. 490–498, Conclusions)

maybe you use bold font in the manuscript to highlight your conclusions for the reader

Response. Adopted: the three findings and the two further results are set in bold.

R-39 (p. 17, l. 514, data availability)

maybe add the link to the zenodo directly here and not only in the references

Response. Added; the three Zenodo DOIs (data, pipeline code, code specific to this paper) appear in the data availability statement.

R-40 (Table 1)

there are some acronyms in the table which need explaining

Response. The caption now lists RBI, IOR, MSFD, CTD, Argo and the season codes. The source label "Croatian MSFD / JADMON CTD" now reads "Croatian MSFD monitoring CTD", since the casts in that row are the MSFD monitoring casts of 2019–2023.

R-41 (Table 2 caption) — "from an identical pipeline, binning, quality control and estimator on both sides"

is a bit too much jargon

Response. Now "computed with the same method on both sides (Sect. 2.4)".

R-42 (Table 3 caption) — "and are wide because the effective sample size is a fraction of the number of months"

unclear

Response. Rewritten: "Because successive monthly values are correlated, the number of independent values is much smaller than the number of months, and the intervals are correspondingly wide (Sect. 2.8)."

R-43 (Table 5 caption) — "split before and from 2017"

sounds odd. I assume you mean one time interval ending in 2017 and another starting in 2017 and continuing until present

Response. Exactly so; now "divided into those taken before 2017 (257 profiles) and those taken from 2017 onward (500 profiles)", also in Sect. 2.9 and the caption of Fig. 7 [6].

R-44 (Table 5 caption) — "Nothing changes significantly inside the core."

odd sentence. Sounds like interpretation and might not be needed in a table caption

Response. Deleted.

R-45 (Fig. 1 caption) — "and observational support"

????

Response. The fourth column is now described as "the sampling behind them": the positions of the casts by source for each period (top two rows) and the number of casts per depth bin by source and season (bottom row).

R-46 (Fig. 1 caption) — "that is the band against which significance should be judged"

interpretation as this are not necessary in a figure caption

Response. Deleted.

R-47 (p. 26)

n

Response. This annotation is the single character "n", which we took to be accidental. If a comment was intended there, we would be glad to address it.

R-48 (Fig. 6 [5] caption) — "This is a schematic, not a data plot."

please note, . . .

Response. Now "Note that the figure is a schematic and not a data plot."

R-49 (Fig. A1 caption) and R-50 (Fig. A3 caption)

this is odd in a figure caption. maybe you need a short few sentences before the additional figures which motivates the questions and analyses in the appendix / see above comment on the necessity of a few introductory sentences in the appendix.

Response. Appendix A now opens with a short paragraph saying what each appendix figure supports, and the captions of Figs. A1 and A3 begin with a description of the figure instead of an argument.