Response to Reviewer #1.

We would like to thank the reviewer for his constructive comments.

There are some technical issues that need to be addressed. The band-pass filtering of zonal wavenumber k = 1 – 3 for the MJO and k = -1 – -8 for the Rossby wave are inappropriate for model simulations. According to Hayashi (1979), only the part of the eastward power that is incoherent with its equivalent westward power represents true eastward propagating signals. The coherence part represents stationary of standing signals. So using k = 1 – 3 to represent the MJO and k = -1 – -8 to represent the Rossby wave would exaggerate the propagating signals. In observations, the eastwest equivalent signals are weak, so this practice is ok. For model simulations, such east-west equivalent signals are strong, the potential coherence part is great and this practice is problematic. The regression results from Jiang et al (2015, Fig. 3) clearly show the dominant stationary signals in many model simulations. The band-pass filtering method used in this current study would mistakenly extract propagating signals from these simulations when there is none.

In order to address the reviewer's comment, we have carried out additional analyses in order to check the importance of the east-west equivalent signals.

1. The analysis of ITV spectrum (new figure 2) shows the strong westward signal in 5 models among 16 (CanESM2, CNRM-CM5, IPSL-CM5A-MR, MPI-ESM-LR, MRI-CGCM3). These models are excluded from further analysis. In other models the westward power is of the same order than in Reanalysis. Exception is the INM-CM4 models where the westward power is equivalent to eastward one. To verify if the signals are coherent we made the further analysis. Below we present the results for INM-CM4 and two other models for comparison.
2. We recompose the U850 signal in the same frequency intervals as for MJO and ER but for the opposite sign of zonal wave numbers: -1…-3 for MJO (westward propagation) and +1..+8 for Rossby waves (eastward propagation) (Figures A1 andA2). It may be seen that the amplitude of westward analogue of MJO is significantly lower as compare to eastward propagating patterns (except for INM-CM4). For Rossby waves the amplitude of eastward and westward propagating signal is comparable but the timing, spatial localization and speed of propagating signal differ significantly. To confirm quantitatively this suggestion we calculated the correlation between eastward and westward propagating signals (Table 1). The correlation is rather small that allows suggesting that the signals are incoherent.

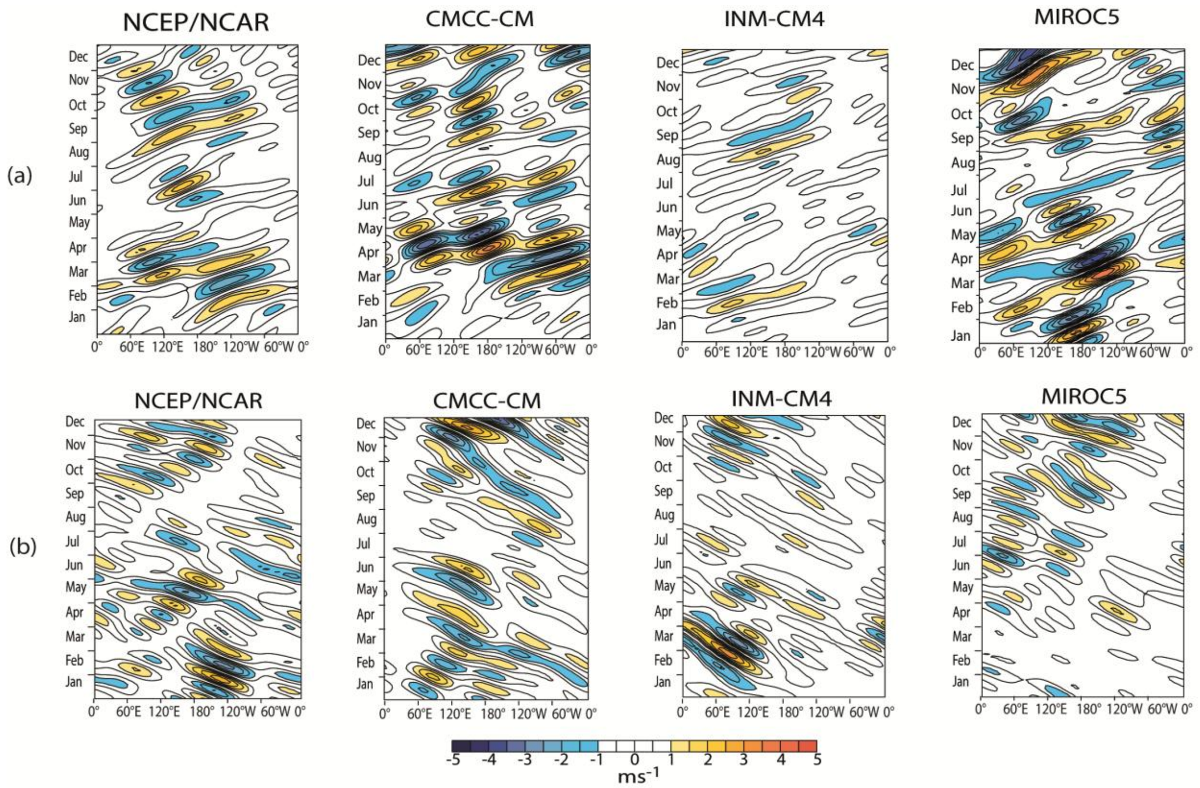

Figure A1: Time-longitude plots of equatorial averaged (5°N–5°S) daily-mean anomalies of MJO filtered U850 from NCEP/NCAR Reanalysis and 3 CMIP5 models (zonal wave numbers: +1...+3 (a), -3...-1 (b)). Contour interval is 0.5 m/s. Negative values ≤–1 m/s are blue shaded, positive values ≥1 m/s are orange shaded.

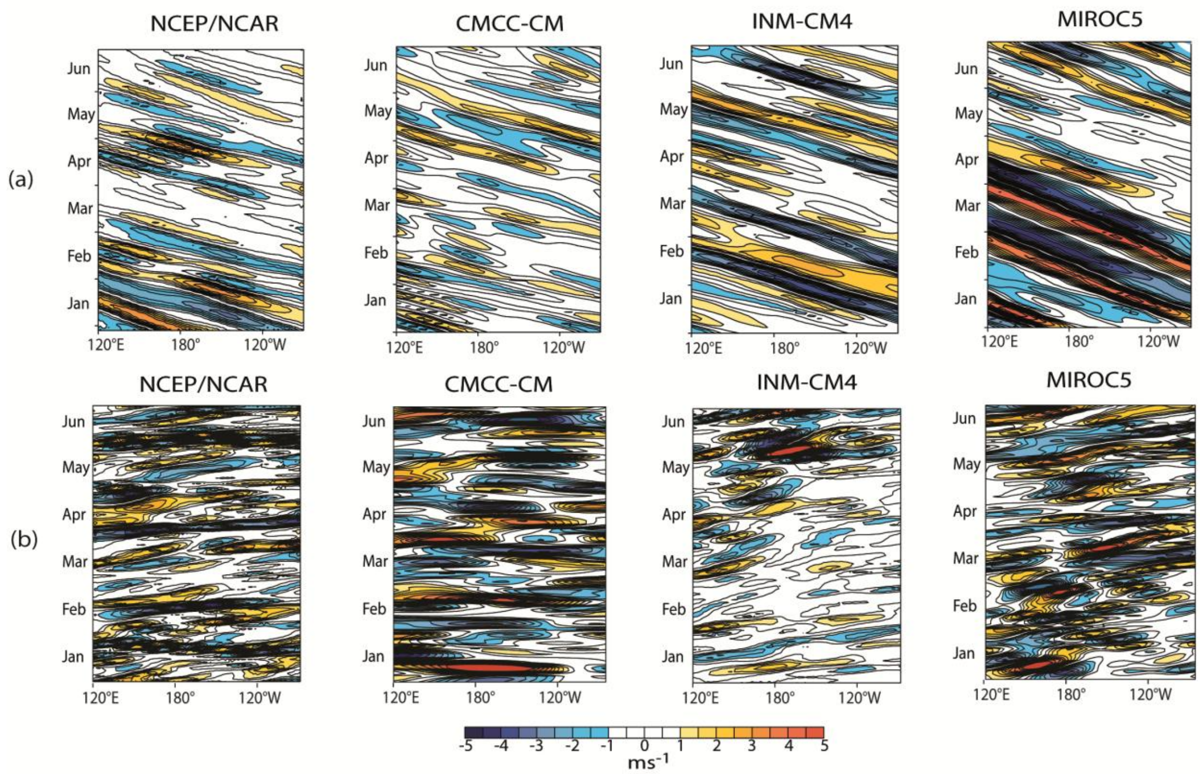

Figure A2: As in Figure A1, but for Rossby waves (zonal wave numbers: -8...-1 (a), +1...+8 (b)).

Table 1: Correlation between MJO and Rossby waves and the signals filtered in the same frequency intervals but opposite sign of zonal wave numbers in CMIP5 models.

	MJO (wave numbers: 1…3, -1…-3)	Rossby waves (wave numbers: +1…+8, -1…-8)
CMCC-CM	0.08	0.23
INM-CM4	0.12	0.3
MIROC5	0.15	0.17

3. We have analyzed the spatial distribution of variance of westward/eastward signal in the frequency interval of MJO (Figure A3). It may be seen that the maximum of variability for MJO are much higher than for its westward counterpart

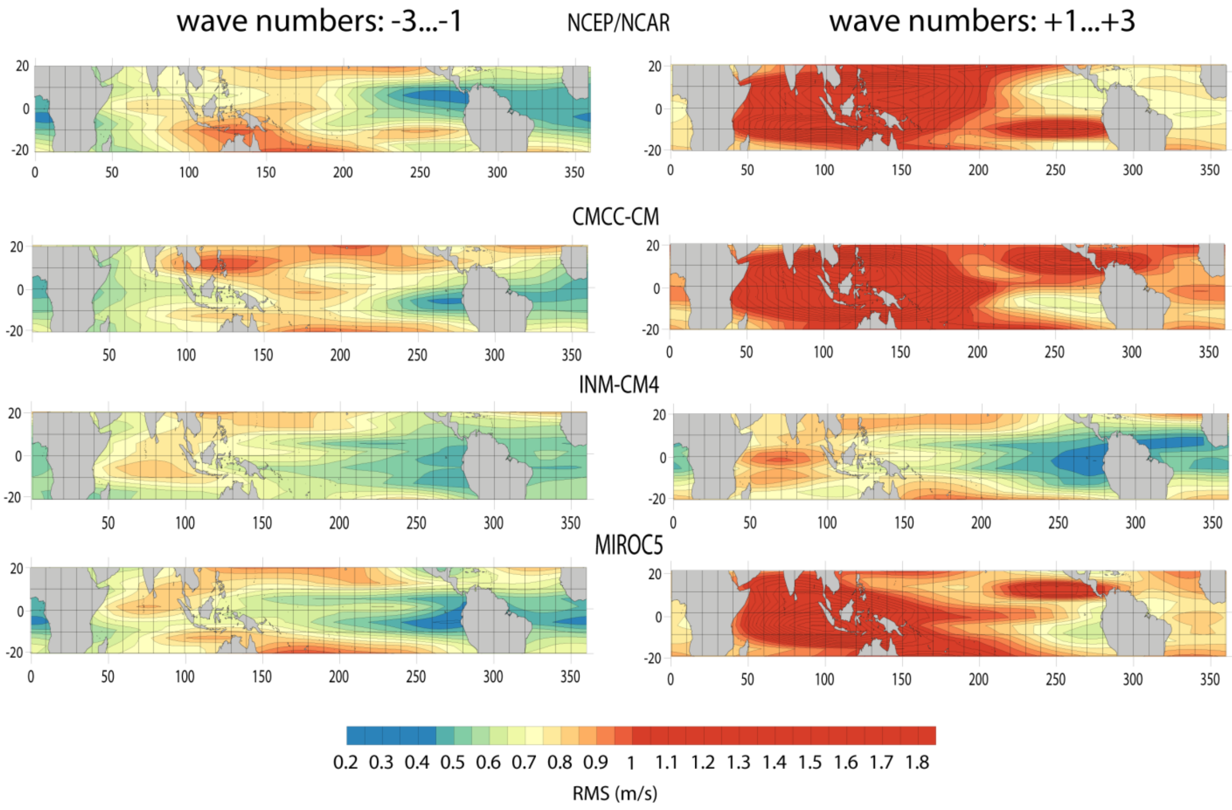

Figure A3: Variance (rms) of U850 filtered in MJO frequency interval (30-60 days) for zonal wave numbers -3...-1 (left column) and +1...+3 (right column).

4. The signal in the frequencies of MJO for zonal wave numbers from -3 up to +3 was recomposed (Figure A4). Figure A4 shows that eastward propagating signal dominates during almost the whole year. The stationary signal can be guessed but its characteristics are comparable to the Reanalysis.

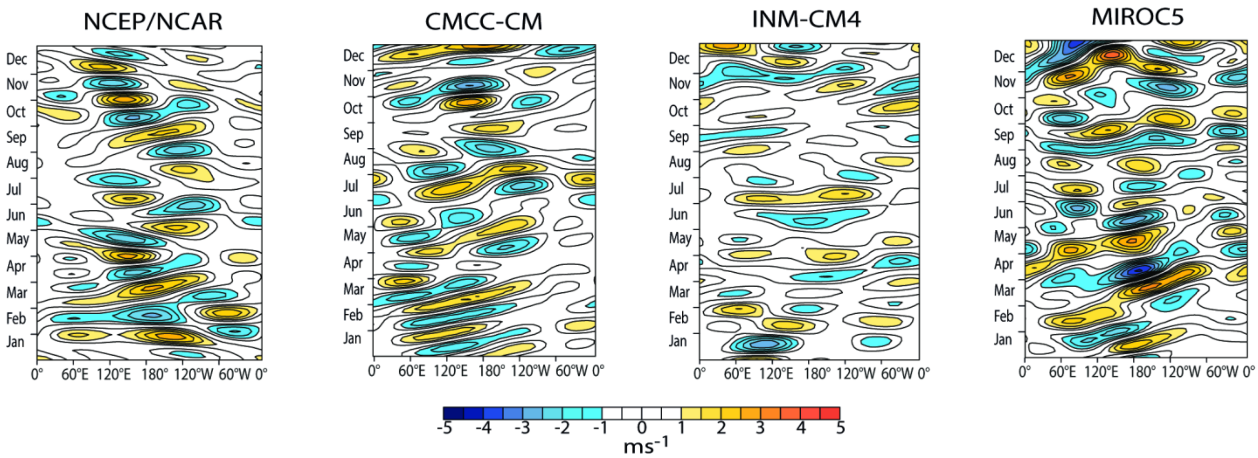

Figure A4: U850 filtered in MJO frequency intervals for zonal wave numbers -3...+3.
Contour interval is 0.5 m/s. Negative values ≤–1 m/s are blue shaded, positive values ≥1 m/s are orange shaded.

The discussion of the coherent signal in the models was added to the revised manuscript.

"Following Hayashi (1979), only the part of the eastward power that is incoherent with its equivalent westward power represents the true eastward propagating signal. Moreover the results of Jiang et al. (2015) emphasize the dominant stationary signals in many model simulations. To verify if the westward counterpart is present in the models, we recomposed the signal in the same frequency intervals that for MJO and Rossby waves but for the opposite sign of zonal wave numbers: -1…-3 for MJO and +1..+8 for Rossby waves. Insignificant correlation between westward and eastward signals confirms that westward and eastward parts are incoherent, validating *a posteriori* our decomposition approach of the model outputs."

Discussions of the results are mostly qualitative and subjective, heavily relying on visual impression. Suggest use quantitative measures to compare models and between models and observations.

Following the reviewer's recommendation we have substantiated our analyses providing metrics of the models' skill in accounting for the ENSO and ITV characteristics. In details:
1) To evaluate quantitatively the simulation of SST distribution associated to the types of ENSO we calculated the spatial correlation between observations and model for SST projected onto the E and C indices (see new Table 2). The models with spatial correlation less than 50% were excluded from further analysis. We also provide the new Figure 1 that summarizes the comparison between observations and models in terms of the spatial structure of ENSO. .
2) For evaluating the ITV characteristics in the models, we now provide the root mean square error (RMSE) of total variance as a function of longitude (Figure 3.cdgh), the RMSE of MJO and Rossby wave seasonal variance in the western and central Pacific respectively (Figure 6). The phase speed values of MJO and ER in the models were compared to the ones of the NCEP/NCAR data (Figures 7 and 8) following the diagnostic of (Hung et al., 2013).

3) We introduced a measure of the predictive score of the ER and MJO with regards to El Niño types, which is used to select the periods over which the statistics is done, recognizing that the seasonal ENSO/ITV relationship has a decadal modulation. This follows the study by Gushchina and Dewitte (2017, submitted to Climate Dynamics). A supplementary material is provided in relation to that.

4) We have introduced the new tables 3 and 4 that summarize the evaluation of the models based on the different diagnostics done in the paper. We acknowledge that the evaluation has a certain degree of subjectivity owing to the difficulty in ranking the importance of the diagnostics between each other.

Significance level of 90% is lower than commonly used 95% in modern literatures.

The results are hardly impacted when we use the 95% significant level. We provide the figures A5 and A6 that illustrates the differences when using 95% instead of 90%.

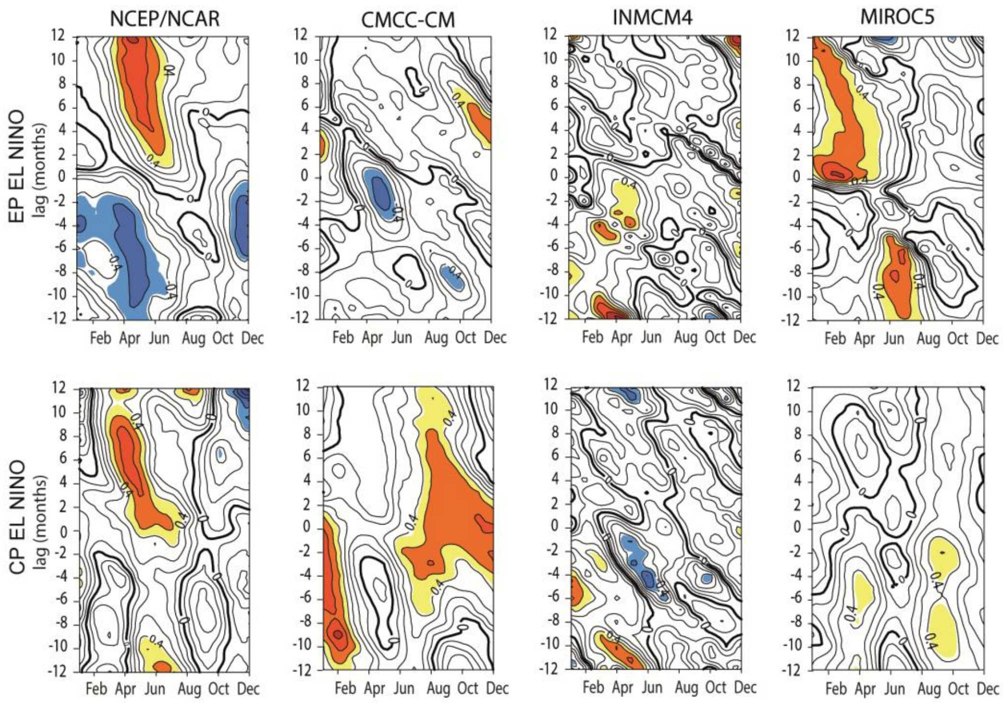

Figure A5: Monthly lagged correlation of E (a-d) and C (e-h) indices as a function of start month with respect to MJO activity index WPacMJOu850 for NCEP/NCAR Reanalysis and 3 CMIP5 models. Contour interval is 0.1. Negative correlation ≤–0.42 is blue shaded, positive correlation ≥0.42 is orange shaded (90% significance level) .

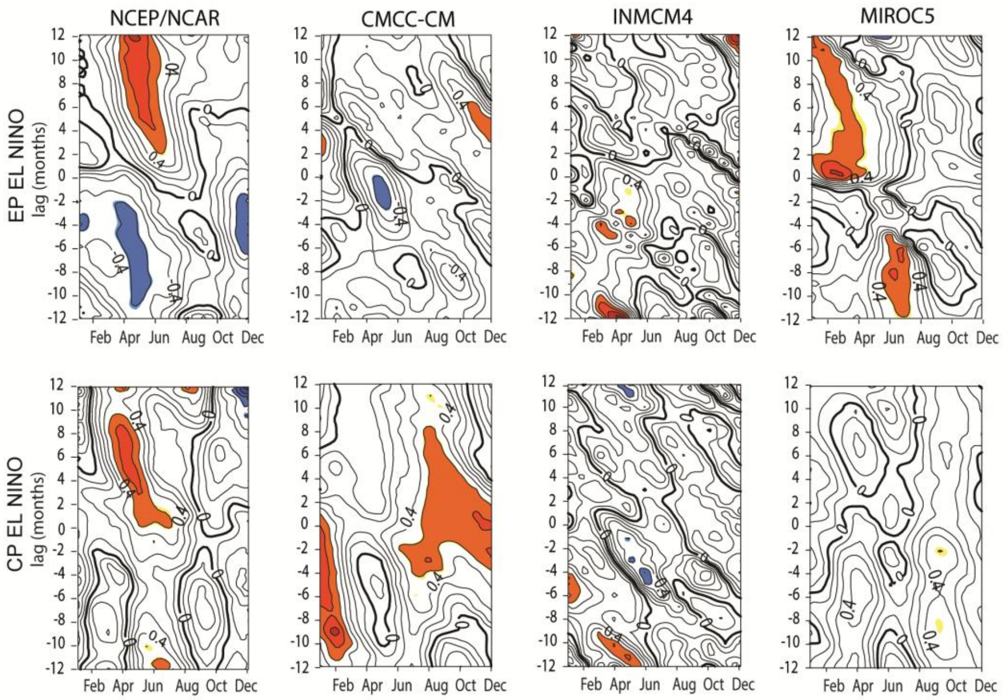

Figure A6: As in Figure A5, but negative correlation ≤–0.49 is blue shaded, positive correlation ≥0.49 is orange shaded (95% significance level).

Using U850 to define the MJO and Rossby wave might be problematic. There are obviously other perturbations in the same frequency band of the Rossby wave (Fig. 3). Why not use precipitation as everyone else did? This would yield results that can be directly compared to others.

In the revised manuscript, we better justify the use of U850 field for deriving the ITV components:
"We use here the U850 field for deriving the various components of the ITV instead of Outgoing Longwave Radiation (OLR) or brightness temperature signals from satellite data noting that the regions in the frequency-wavenumber domains where the spectral energy peaks are similar for OLR and U850, which is also predicted by a simple dynamical model of ITV (Thual et al., 2014). Moreover the use of U850 eases the interpretation of the results since it is the westerly wind anomalies that serve a physical conduit from the ITV to the ENSO dynamics. This approach follows previous relevant studies (McPhaden et al. 2006; Hendon et al. 2007)."

Some missing literature citations should be added: Hendon et al. (2007) for seasonally varying relationship between MJO activity and the ENSO cycle Kessler et al. (1995) for MJO inducing the oceanic Kelvin wave in the Western Pacific Zhang and Gottschalck (2002) for MJO as a precursor of El Nino.

Hendon et al. (2007) was used as a reference in the original manuscript, and the other suggested references were added to the revised version.

Response to Reviewer #2

We appreciate the reviewer's constructive comments.
The following is our point-to-point reply to these comments.

Twenty-three CMIP5 models are investigated for their match with observations in representing aspects of tropical intraseasonal and interannual variability. Despite the title, which emphasises the relationship between interannual and intraseasonal variability, the majority of the paper is first spent analysing which models are best at simulating individual aspects of the variability, namely the two types of ENSO, the MJO, and Equatorial Rossby and Kelvin waves. The results show a large variety of behavior from the models, with very few models showing variability and relationships like observed. This may be of interest to model developers, but I don't think it adds much new insight into the dynamics of the observed variability. Also, I can't see how these results can help pin-point what aspects of the models need to be changed for improvement. I understand this is a difficult task, but is one that needs to be done to help improve the models.

We agree with the reviewer that our paper does not propose or suggest ways to improve the models. It is an evaluation of the realism of the ITV/ENSO relationship, and as such, the paper can be viewed as a preliminary step towards suggesting improvement in the model physics, considering that it proposes a physically-based metrics to evaluate models and classify them into two broad classes, the less and most realistic ones. Such a "classification" could be the basis for identifying differences in some key dynamical aspects of the ITV, like the energy sources of the ITV (i.e. extra-tropical disturbances, tropical instabilities or non-linear interactions of multiple waves), its coupling with SST, its seasonal phase locking etc.
Despite the limitation of not addressing issues on model development, we believe that our results may still fit with the scope of Geoscientific Model Development since it provides "*new methods for assessment of models, including work on developing new metrics for assessing model performance and novel ways of comparing model results with observational data*", as well as proposes "*novel ways of comparing model results with observational data*"

1. The English grammar needs improving to make it easier to read and understand.
For example, there are many instances where the word "the" is inserted incorrectly or missing.

We have thoroughly revisited the text and improved the English grammar

2. Page 3, line 24: Kim and You (2012) missing from reference list.

Added to the reference list

3. Page 5, line 6: "PI" is not defined.

Pre-Industrial – corrected

4. Section 2.2: It is noteworthy that you are using zonal wind data instead of a proxy for clouds and convective rainfall (e.g. outgoing longwave radiation) as used by Wheeler and Kiladis (1999). This means that the variability highlighted by your wavenumberfrequency analysis (Figure 3) is somewhat different to that highlighted in Wheeler and Kiladis (1999). It also means that the variability you show and isolate is not necessarily 'convectively-coupled'. For example, Figure 3 indicates the existence of the global Rossby-Haurwitz waves for low westward-propagating wavenumbers and periods around 5 days. It also means that the convectively-coupled equatorial Rossby (ER) and Kelvin waves are much less clear in Figure 3. This means that your filtered fields will also contain a much greater mix of variability compared to Wheeler and Kiladis. Finally, I note that you use rectangles to define your regions of filtering instead of following the dispersion curves for the equatorial waves. Ideally you should change your fields and filtering to better match the characteristics of the waves. However, I support the use of the western Pacific wind indices later in the paper as this is consistent with the findings of Hendon et al. (2007).

In the revised manuscript, we better justify the use of U850 field for deriving the ITV components:
"We use here the U850 field for deriving the various components of the ITV instead of Outgoing Longwave Radiation (OLR) or brightness temperature signals from satellite data noting that the regions in the frequency-wavenumber domains where the spectral energy peaks are similar for OLR and U850, which is also predicted by a simple dynamical model of ITV (Thual et al., 2014). Moreover the use of U850 eases the interpretation of the results since it is the westerly wind anomalies that serve a physical conduit from the ITV to the ENSO dynamics. This approach follows previous relevant studies (McPhaden et al. 2006; Hendon et al. 2007)."

In the paper, the focus is on two component of ITV – MJO and equatorial Rossby (ER) wave which were shown to be associated to El Niño development (McPhaden et al., 2006, Hendon et al., 2007, Gushchina and Dewitte, 2011, 2012). The Figure B1 provides the wavenumber-frequency spectra for both OLR and U850 and it can be seen that, the domains where the MJO and ER spectral energy peaks are comparable for both fields. The MJO spectral maximum in OLR is shifted to the higher zonal wave numbers as compare to U850 in accordance with the results of previous studies (Zhang, 2005).

For example, Figure 3 indicates the existence of the global Rossby-Haurwitz waves for low westward-propagating wavenumbers and periods around 5 days. It also means that the convectively-coupled equatorial Rossby (ER) and Kelvin waves are much less clear in Figure 3. This means that your filtered fields will also contain a much greater mix of variability compared to Wheeler and Kiladis.

The frequency band for MJO filtering is 30-60 days and for ER – 10-50 days, thus we do not include in our filtered fields the waves with 5 days period. Overall the main difference between OLR and U850 spectra is located at the periods shorter than 10 days, which does not impact our results. We have improved the presentation of figure 3 (new figure 2) so as to better visualize the MJO and ER domains. We applied the same color scale as in Hung et al. (2013) for easing the comparison with their results.

Note that we use rectangles to define our regions of filtering only for MJO as in Wheeler in Kiladis (1999). For Rossby wave following (Wheeler and Kiladis, 1999) we follow the dispersion curves for the equatorial waves with equivalent depth ranging from 8 to 90 m. This is now mentioned in the text of the revised manuscript. “For Rossby waves, the frequency-wavenumber bands is also limited by the dispersion curves corresponding to values of the atmosphere equivalent depth ranging from 8 m to 90 m, which follows (Wheeler and Kiladis, 1999)”

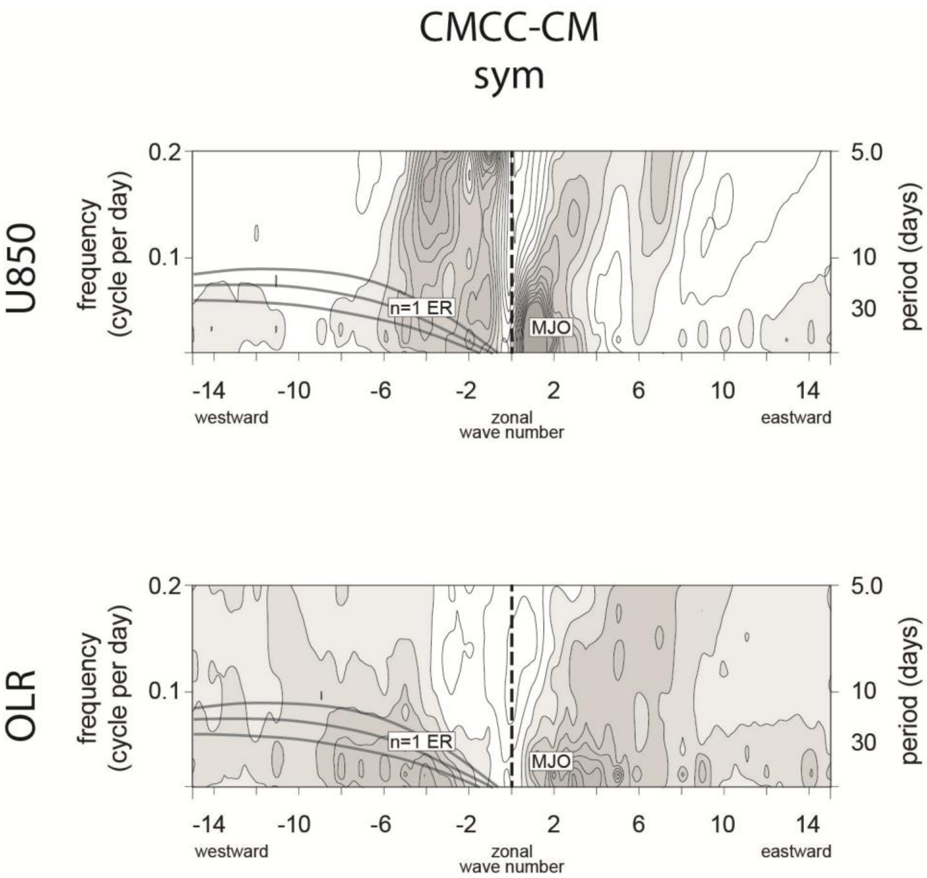

Figure B1: Space–time spectrum of the 15°N–15°S symmetric component of U850 (upper panel) and OLR (bottom panel) divided by the background spectrum for CMCC-CM model

5. How are the values in Table 3 calculated?

We agree with the reviewer that the estimates presented in Table 3 are rather uncertain. We have removed this Table and added to the revised version a figure that is comparable to figure 9 of Hung et al. (2013) and that indicate the main values of expected phase speed (figures 7-8).

6. Page 13, lines 13-15. This is poor style for scientific writing. Please refer to this paper: http://onlinelibrary.wiley.com/doi/10.1029/2010EO450004/full

We have improved the style, which should ease the readability of the revised manuscript.

Response to the Reviewer #3

We thank the reviewer for his/her constructive comments, which has motivated us to improve our manuscript.

General Comments

The authors attempt to relate the ability of CMIP5 coupled models to simulate ENSO with their ability to correctly simulate the seasonal cycle and coupling between atmospheric intraseasonal equatorial waves, namely the MJO and Equatorial Rossby (ER) waves and the ocean. While the observational support for such a relationship in the real world has been well-established by the authors and others, unfortunately most of models studied here appear to only marginally simulate such relationships. The physical relationship between the zonal wind variability and ENSO in the models has not been explored in detail, therefore in my opinion the paper should be revised to include more diagnostics.

We have clarified in the motivations that our main objective is here to evaluate the CMIP5 models in terms of the seasonal ENSO/ITV relationship, which is viewed as a preliminary step for addressing why some models fail more than others to account for this fundamental ENSO property. Compared to the original version of the manuscript, we have also modified our methodology in particular considering the fact that the ENSO/ITV relationship is modulated at decadal timescales, which diagnoses more objectively the dispersion among models. Our main conclusion is indeed that CMIP5 models have limited skill in accounting for the seasonal ENSO/ITV relationship when taking into account ENSO diversity (i.e. the existence of two types of event), which suggests that what produces ENSO diversity (and ENSO itself) in the models may be associated to different forms of the external forcing. We have expanded the discussion in order to emphasize the consistency of our results with recent ENSO studies.

Specific Comments

The paper starts out with an interesting and useful analysis of the behavior of ENSO in the models. However, I did not get a sense from this manuscript of which aspects of the models lead to bad (or better) representation of ENSO. For example, no indication of the oceanic response to wind forcing has been shown. Is it possible that the ocean models might also be an issue? In the end, it seems to me that a paper like this one in a journal such as GMD should lead to some recommendations on how models could be improved. Have the authors checked any of the ocean data (TAO buoys, SODA reanalysis) as in Guschina and DeWitte (2011) to see whether the wind signals they are isolating are actually related to oceanic Kelvin waves? Otherwise trying to relate ITV to ENSO seems speculative. In a better model such as CMCC-CM there should be a realistic relationship between the wind forcing and the ocean response. I encourage the authors to expand on these points in a revision.

First of all, we have modified our diagnostics of the realism of the model in terms of ENSO, which should be clearer (see new figure 1). In particular we base our analysis on the E and C patterns introduced by Takahashi et al. (2011), which compared to the EOF1 and EOF2 mode patterns have a more robust physical interpretation (see Takahashi et al. (2011) for a discussion). These modes encapsulate many aspects of the ENSO processes that are either embedded into the atmospheric and/or oceanic components of the model (since the system is coupled).

Regarding the relationship between high-frequency winds and the intraseasonal equatorial oceanic Kelvin wave in the models, this is difficult to address this issue from the CMIP5 archive owing to the absence of 5 day-mean oceanic fields to derive the Kelvin wave (only monthly mean outputs are available).

The other major issue has to do with the isolation of CCEWs. The authors are using broadly defined filters that are based on OLR or brightness temperature signals from satellite data. Based on the spectra in Fig. 3, there is little basis for using the filter bands they have chosen, which ultimately derive from precipitation signals. Talking about "waves" such as the MJO and ERs in Figs. 7 and 8 is very suspect, since filtering of just red noise will give you similar results. I suggest more diagnostics to establish the existence of zonal wind signals associated with the MJO and ER waves (see below).

To isolate MJO in the zonal wind field we follow previous studies (McPhaden et al., 2006; Hendon et al., 2007) using the same frequency-wavenumber interval for extracting the MJO signal from zonal wind.
In the revised manuscript, we better justify the use of U850 field for deriving the ITV components:

"We use here the U850 field for deriving the various components of the ITV instead of Outgoing Longwave Radiation (OLR) or brightness temperature signals from satellite data noting that the regions in the frequency-wavenumber domains where the spectral energy peaks are similar for OLR and U850, which is also predicted by a simple dynamical model of ITV (Thual et al., 2014). Moreover the use of U850 eases the interpretation of the results since it is the westerly wind anomalies that serve a physical conduit from the ITV to the ENSO dynamics. This approach follows previous relevant studies (McPhaden et al. 2006; Hendon et al. 2007)."

In the paper, the focus is on two component of ITV – MJO and equatorial Rossby (ER) wave which were shown to be associated to El Niño development (McPhaden et al., 2006, Hendon et al., 2007, Gushchina and Dewitte, 2011, 2012). The Figure C1 provides the wavenumber-frequency spectra for both OLR and U850 and it can be seen that, the domains where the MJO and ER spectral energy peaks are comparable for both fields. The MJO spectral maximum in OLR is shifted to the higher zonal wave numbers as compare to U850 in accordance with the results of previous studies (Zhang, 2005).

We have improved the Figure 3 (Figure 2 in revised version) so as to better visualize the MJO and ER domains. We applied the same color scale as in Hung et al. (2013) for easing the comparison with their results. Note that for the selected models in section 3.2 (model name in green in Table 3) the MJO spectral maximum is similar to NCEP/NCAR Reanalysis (new figure 2).

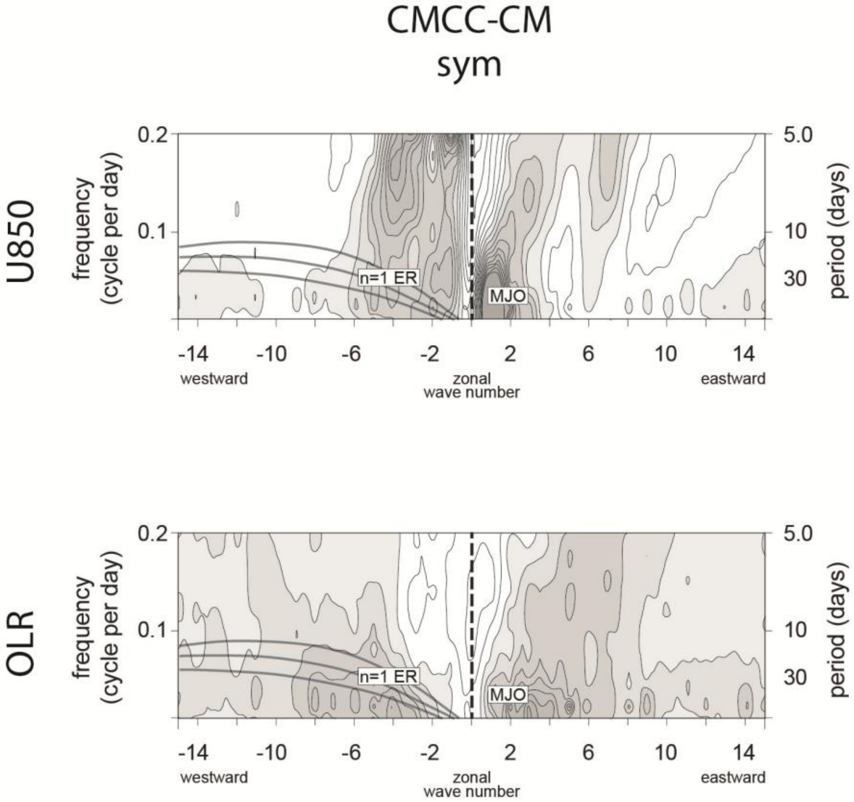

Figure C1: Space–time spectrum of the 15°N–15°S symmetric component of U850 (upper panel) and OLR (bottom panel) divided by the background spectrum for CMCC-CM model

Technical Comments:

Pg. 2, line 21: see also Keen, 1982 Mon. Wea. Rev. pg. 1405.

We have added this reference.

Pg. 3, line 14: “loose” => “lost”

corrected

Pg. 5, line 16: except that Lin et al. and Hung et al. used precipitation not wind, this needs to be pointed out here.

Following the reviewer's recommendation, we now mention explicitly that Hung et al. and Lin et al. use a different approach. In the revised version

"They showed that CMIP5 models exhibit an overall improvement in the simulation of ITV, especially the MJO and several CCEWs as compared to CMIP3 models. The CMIP5 models produce larger total intraseasonal variance of precipitation than the CMIP3 models, including larger variances of MJO, Kelvin, ER, and eastward inertio-gravity (EIG) waves. About one-third of the CMIP5 models generate the spectral peak of MJO precipitation between 30 and 70 days; however, the model MJO period tends to be longer than in the observations and only one of the 20 models is able to simulate a realistic eastward propagation of the precipitation patterns associated to MJO."

Pg. 6, line 3: "the maximum of ITV/ENSO relationship is observed. ' It is not clear what you mean by this. Precisely how are the indices in Table 2 defined? Please provide more detail on this, perhaps by using the "NCEP-NCAR" data as an example, which should be the closest to reality. Also, only 4 models are shown in Table 2 yet other models are analyzed later.

The text was revised accordingly so that this should be now clearer.

In the revised version

"The regions for averaging the MJO and ER running variance correspond to the regions where the maximum of ITV/ENSO relationship is observed in the Reanalysis (Guhschina and Dewitte 2011): western Pacific (120°-180°E; 5°S-5°N) for MJO and central Pacific (140°E-160°W; 5°S-5°N) for Rossby waves. These indices are further referred as MJO and Rossby wave indices."

Pg. 8, Line 8: The spectra in Fig. 3 should be replotted, since it is difficult to see the signals through the dispersion curves. In particular, the MJO peak should show be a wavenumber 1 signal but these are obscured by the dispersion lines. A comparison with Hung et al. for those models that have overlap would be welcome. If I look for example at CanESM2 and CCSM4 spectra of rainfall in Hung et al., it seems that these two models have a good spectral peak for the Kelvin wave in rainfall, but there is no evidence for a corresponding zonal wind peak in Fig. 3. This just illustrates the problem with using wind to define the equatorial wave modes as used here.

The Figure 3 (Figure 2 in the revised manuscript) was replotted using the same color scale than in (Hung et al., 2013), which should ease the comparison. The dispersion curves for Kelvin waves were removed as they obscure the MJO signal and as we do not focus on the analysis of the Kelvin waves. The main focus here is on the ER and MJO components of ITV.

Line 15: "lower" suggest "weaker"

corrected

Pg. 9, line 9: “The maximum: : :” I have no idea what this sentence means to say.

Changed to

“The MJO exhibits a maximum intensity in the summer hemisphere (i.e. in the Northern Hemisphere in July and in the Southern Hemisphere in January) which implies that the MJO variance peaks along the equator in boreal spring when it may act efficiently as an ENSO trigger. Therefore the MJO cross-equatorial seasonal migration is a key feature that needs to be evaluated in the models.”

Pg. 10, line 1: I guess the periods chosen for Fig. 7 are chosen from random, but perhaps the authors looked for good examples from the reanalysis and each model? More detail is needed, including making the obvious but necessary point that the model fields in no way are expected to match the reanalysis or each other. It is not clear where the statements on propagation velocity and intensity come from. These are only one year periods, and it seems that the authors are just making statements by visual comparisons between the plots. The characteristics of the waves in each model could be compared with reanalysis by using diagnostics of the type used by Wheeler et al. 2000 (J. Atmos. Sci. pg. 613) or Hung et al. (their Fig. 9). The characteristics of the “waves” identified here overwhelmingly determined by the filtering, which sets the phase speed in particular. The plots in Fig. 8 are great examples of getting “something from nothing” by filtering: The CCSM4 zonal wind spectrum in Fig.3 shows no signal at all for ER waves, yet Fig. 8c shows lots of westward propagation, which must come primarily from the constraints of the filter. There are lots of other examples of this.

The figures 7 and 8 were removed from the revised manuscript. We added to the revised version the figures that are comparable to Figure 9 of Hung et al. (2013) and that indicate the main values of expected phase speed (new figures 7-8). We also present the distribution of MJO and ER total variance along the equator as in Wheeler et al. (2000) (Figure 3).

CCSM4 has an obvious signal in ER and MJO domain, which is better seen on the modified Figure 2.

To illustrate that models with unrealistic spectra yield an unrealistic filtered signal, we present the figure C2 that corresponds to the CanESM2 model:

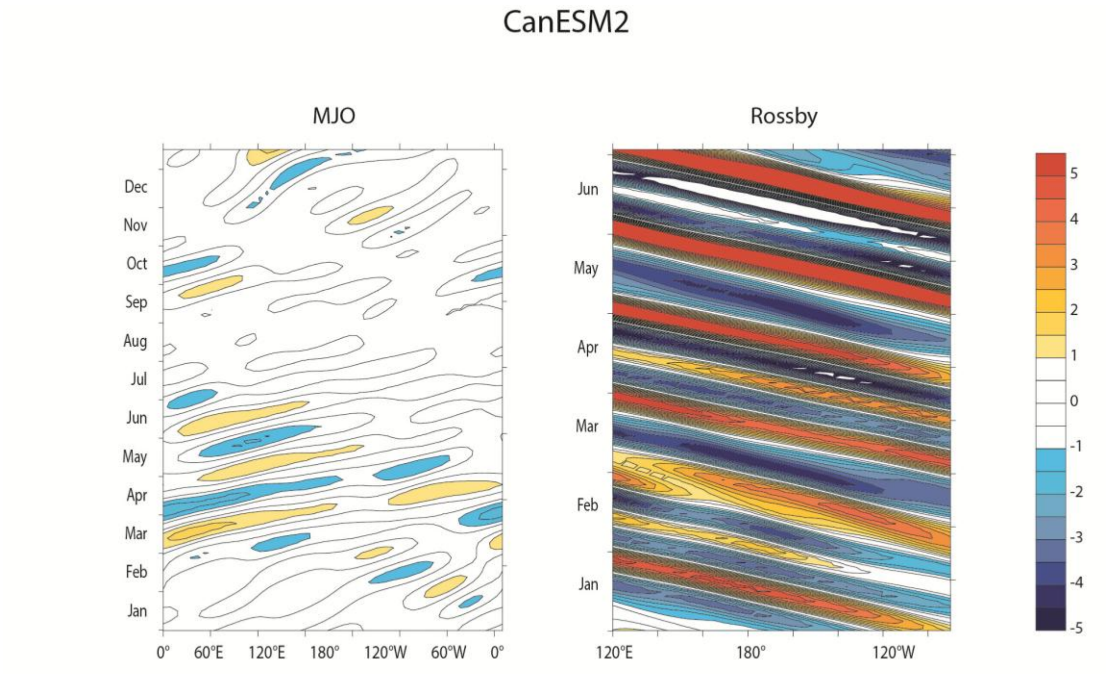

Figure C2: Time-longitude plots of equatorial averaged (5°N–5°S) daily-mean anomalies of MJO and ER filtered U850 from CanESM2 model. Contour interval is 0.5 m/s. Negative values ≤–1 m/s are blue shaded, positive values ≥1 m/s are orange shaded. Zero line is omitted

Pg. 11, top: Much more discussion of what the expected relationship between the zonal wind and ENSO is needed here. Although the authors refer back to Takahashi (2009) and their own previous work, it will not be immediately obvious what Figs. 9a and 10a imply for ENSO forcing. The caption for Fig. 9 does not help. A brief review of the concepts is needed at the start of Section 3.3.1 before Figs. 9 and 10 can be interpreted. Unfortunately, the model results from Figs. 9 and 10 are not very impressive, with little statistical significance indicated. Also, it is difficult to even tell the sign of many of the signals, so I suggest using more color.

Following the reviewer's recommendation, the presentation of figures 9 and 10 (new figures 10 and 11) was improved.
We have also modified our methodology in particular considering the fact that the ENSO/ITV relationship is modulated at decadal timescales following Gushchina and Dewitte (2017, submitted to Climate Dynamics). The ITV/ENSO relationship is analyzed based on 56 years of the historical run.

The text corresponding to the analysis of the ITV/ENSO relationship was significantly modified (Section 3.3).

Pg. 11, line 9: Indian Ocean wind stress could not force ENSO.

In the revised version the forcing from Indian ocean is not more mentioned.

Line 10: 9g => 9h

corrected

Line 21: Here it is difficult to even tell what the sign is in Fig. 10h.

The figures 9-12 (new figures 10 and 11) were modified by adding more color.

Pg. 12, line 2: Puy et al. 2016

corrected

Line 15: I wouldn't say it's "very close", but certainly it's better than the rest. To be honest, since Figs. 10-12 are based on the models' own renditions of ENSO, the huge disparity between them leads to an obvious conclusion: what forces ENSO in most of these models is something different than what forces it in the real world. I think this is the statement you should make more forcefully.

We agree with the reviewer and we have expanded the discussion section so as to emphasize this result. The fact that the models have difficulties in simulating the observed ENSO/ITV relationship suggests that ENSO in the models is influenced by other forms of external forcings not necessarily related with the ITV. This is in fact in line with recent researches that suggest that the role of external forcing on ENSO might be more important than previously thought (Dommenget and Yu, 2017; Takahashi et al., 2017). In particular Takahashi et al. (2017) shows that, based on the experiment with a conceptual non-linear recharge-discharge model, the role of the low-frequency component of the external forcing (interannual timescales) is actually key to trigger El Niño events and that there can be extreme El Niño events without a significant recharge of the heat content. What happens in 2014 when a strong El Niño event was expected after strong WWBs in February-March similar to 1997 (Menkes et al., 2014), was also the indication that external forcing is key for the development of El Niño (Hu and Fedorov, 2016).

References:

Dommenget, D. and Y, Yu, 2017: The effects of remote SST forcings on ENSO dynamics, variability and diversity. Climate Dynamics, in press.

Hu, S. and Fedorov, A.V., 2016: An exceptional easterly wind burst stalling El Niño of 2014. PNAS, doi: 10.1073/pnas.1514182113

Menkes, C. E., M. Lengaigne, J. Vialard, M. Puy, P. Marchesiello, S. Cravatte, and G. Cambon, 2014: About the role of Westerly Wind Events in the possible development of an El Nino in 2014, Geophysical Research Letters, 41(18), 6476-6483.

Takahashi, K., Karamperidou, C. and B. Dewitte, 2017: A simple theoretical model of strong and moderate El Niño regimes. Climate Dynamics. Accepted upon minor revisions

Pg. 14, line 7: “The deficiency of INM-CM4: : :” Little basis for this statement is shown. More detailed diagnostics of the ocean response would bolster claims like this, even by using SST if sea surface height or thermocline depth cannot be obtained.

This part was removed and we now base our analysis on the historical runs taking into account the decadal modulation of ITV/ENSO relationship, which yielded to revise the text.