Improving and harmonizing El Niño recharge indices

Takeshi Izumo[1,2] and Maxime Colin[3,4,5]

[1] Institut de Recherche pour le Développement (IRD), EIO laboratory, Université de la Polynésie Française (UPF), Tahiti, French Polynesia

[2] formerly at IRD, Sorbonne Université - CNRS-IRD-MNHN, LOCEAN Laboratory, IPSL, Paris, France

[3]Climate Change Research Centre, University of New South Wales, Sydney, New South Wales, Australia

[4]Laboratoire GEPASUD, University of French Polynesia, Tahiti, French Polynesia

[5]now at Leibniz Centre for Tropical Marine Research (ZMT), Bremen, Germany

Corresponding author: Takeshi Izumo (takeshi.izumo@ird.fr)

Submitted to GRL on February 21st, 2022
Revised , August 2022

Key Points:

- To clarify the debate on the Recharge Oscillator index, we develop an objective approach optimizing equation fit to observations.
- The recharge index must be based on the slow component of sea level/thermocline depth, taken independently from the fast zonal tilt mode: this reconciles usual metrics.
- The optimal index is this independent component taken in the equatorial and southwestern Pacific. It is better suited for operational diagnostics.

Abstract

El Niño Southern Oscillation (ENSO) is the leading mode of interannual climate variability, with large socioeconomical and environmental impacts. The main conceptual model for ENSO, the Recharge Oscillator (RO), considers two independent modes: the fast zonal tilt mode in phase with central-eastern Pacific Temperature (T_e), and the slow recharge mode in phase quadrature. However, usual indices (western or equatorial sea level/thermocline depth h) do not orthogonally isolate the slow recharge mode, leaving it correlated with T_e. Furthermore the optimal index is currently debated. Here, by objectively optimizing the RO equations fit to observations, we develop an improved recharge index. (1) T_e-variability is regressed out, building h_{ind} statistically-independent from T_e. Capturing the pure recharge, h_{ind} reconciles usual indices. (2) The optimum is equatorial plus southwestern Pacific h_{ind_eq+sw} (because of ENSO Ekman pumping meridional asymmetry). Using h_{ind_eq+sw}, the RO becomes more consistent with observations. h_{ind_eq+sw} is more relevant for ENSO operational diagnostics.

Plain Language Summary

El Niño and La Niña events have important impacts globally. A key element for long-lead forecasts is the recharge state of the tropical Pacific Ocean, as captured in the Recharge Oscillator (RO) conceptual model. The RO considers two independent modes of oceanic variability, a fast adjustment process and a slow recharge/discharge process. However, usual recharge indices mix these two modes of variability, and can thus lead to ambiguous operational diagnostics of the actual oceanic recharge state. Here we develop a better recharge index, independent of the fast mode, which reconciles typical indices and allows us to go beyond the current geographical debate on the optimal metrics. We use an objective approach optimizing the RO resemblance to observations to find the optimal index: the independent sea level (or equivalently thermocline depth) averaged over the equatorial and southwestern tropical Pacific. We recommend this simple and unambiguous index for El Niño operational forecasts diagnostics.

1 Introduction

The El Niño Southern Oscillation (ENSO) is the leading mode of climate interannual variability, with large socioeconomic and environmental impacts (e.g. Neelin 1998, Wang and Picaut 2004, Clarke 2008, Timmerman et al. 2018, Jin et al. 2020 reviews). While ENSO predictability skill is now rather good at short leads, it has to be improved at longer leads (e.g. Barnston et al. 2012, 2019). One key element for long-lead seasonal forecasts is the recharge state of the tropical Pacific (be it in terms of Oceanic Heat Content (OHC), thermocline depth h or sea level anomaly (SLA)), as it brings long oceanic memory across ENSO phases. Its role is formalized in the Recharge Oscillator (RO) conceptual model of ENSO (e.g. Wyrtki 1985 ; Jin 1997a,b ; Clarke et al. 2007; Clarke 2010). E.g. during a La Niña, easterlies favor a slow accumulation of OHC in the western and equatorial Pacific. This recharge will progressively favor positive Sea Surface Temperature (SST) anomalies in the central-eastern equatorial Pacific (T_E) and thus El Niño onset through the Bjerkness positive feedback. The El Niño event will in turn lead to a discharge favoring a reversal to La Niña conditions. The RO can thus well explain ENSO cyclic nature (cf. section 3 for RO equations).

Yet there is a debate on the best recharge metric, e.g. western or equatorial Pacific, sea level or thermocline depth. In the RO, the equatorial Pacific basin adjustment is separated into two independent modes: 1) the fast mode associated to a zonal tilt of the thermocline, in phase with zonal equatorial wind stress τ_x and T_E (Fig. 1b); 2) the slow recharge mode in phase quadrature with the fast mode (and thus with τ_x and T_E ;Fig. 1a). Several indices have been developed for this slow recharge mode that all bring predictability skill. Wyrtki (1985) and Jin (1997a,b) originally focussed on the western equatorial Pacific h_w, using equivalently SLA, OHC or thermocline depth for h. Jin's sketch, Meinen and McPhaden (2000), and Burgers et al. (2005) focussed on the mean equatorial band: h_{eq} (e.g. Warm Water Volume WWV based on the 20ºC isotherm depth Z20, or SLA, as usual thermocline depth proxies). h_{eq} has become a widely-used ENSO recharge index. Theoretically, in the RO model, h_{eq} is independent (i.e. uncorrelated at lag 0) from the fast tilt mode (and thus from T_E) within Jin (1997ab) approximations.

This independence/orthogonality property is essential, so that the recharge metric captures solely the *pure* recharge mode and does not mix it with the fast mode. But in observations, h_{eq} (and its proxies) is not independent (Figs 1b and S2f). It is physically ambiguous, mixing the fast adjustement with the slow recharge mode, and thus potentially misleading: h_{eq} is dominated by a short-term Ekman convergence leading to a temporary *fast* "*increase*" with *El Niño westerlies* (thus not a true long-term recharge in the RO sense), two times larger than the long-term *slow discharge* expected from the RO theory. h_{eq} is hence firstly an index of the fast equatorial Kelvin wave rather than of the RO long-term recharge/discharge process (Izumo et al. 2018a; see also Neske and McGregor 2018). Thus h_{eq} is strongly positively correlated to the fast mode (and thus to τ_x and T_E). Hence, Izumo et al. (2018a) recommended the use of a western Pacific index, h_w (in agreement with Ramesh and Murtugudde 2013; Boschat et al. 2013; Graham et al. 2015 ; Lai et al. 2015; Ballester et al. 2016a; Petrova et al. 2017; Jin et al. 2020). Yet, such western Pacific OHC index is also not independent. It is partly negatively correlated to the fast zonal tilt mode (Fig. 1b). h_w could thus also lead to ambiguous diagnostics of the actual oceanic state.

Therefore the fast mode needs to be removed in order to obtain an operationally-useful indicator of the fully-isolated recharge process. Here we will hence develop an improved recharge index h_{ind} independent of the fast mode. We will show that previously-defined indices become closer using h_{ind}. We will develop an objective approach optimizing the skill of the RO differential equations fit to observations for the (T_E,h_{ind}) pair to find the optimal averaging region for h_{ind} : h_{ind_eq+sw}.

2 Data and methods

Here we use classical observations and reanalyses datasets, indices and statistical methods detailed in Suppl. Text S1 (statistics robust thanks to a sufficiently-large number of effective degrees of freedom, ~85 to ~140).The ENSO index T_E, is Niño3.4 relative SST (RSST, i.e. SST minus its 20°N-20°S tropical mean), as recommended by Izumo et al. (2020) and Van Oldenborgh et al. (2021) because atmospheric tropical deep convection interannual anomalies are rather related to RSST than to SST. Theoretically, SST is the variable directly involved in the recharge process for the term F_1*h in the dT_E/dt equation (cf. section 3). Yet, the recharge process is driven by windstress, itself directly driven by atmospheric deep convection and thus by RSST. Therefore, Niño3.4 RSST is better for the term F_2*T_E in the dh/dt equation (Supplementary Text S3). Using SST instead of RSST, or Niño3 instead of Niño3.4, makes T_E slightly less correlated to equatorial Pacific τ_x (i.e. ocean-atmosphere coupling), but leads to very similar results (Supplementary Table S2).

3 Improving the Recharge Oscillator recharge index

3.1 Traditional RO framework revisited

To derive RO dT_E/dt tendency equation (see Jin et al. 2020 review for derivation), some physically-reasonable assumptions are used. 1) τ_x is proportionnal to T_E, i.e. tropical deep convection and related τ_x respond quickly to T_E. 2) The fast oceanic response (i.e. quasi-instantaneous, timescales faster than ~2-3 months) to τ_x leads to a positive Bjerkness feedback term in the dT_E/dt equation (through both the zonal advective and thermocline feedbacks) that is proportional to τ_x and thus to T_E: $R_{BJ\_o}T_E$. 3) Atmospheric fluxes are approximated as a weak Newtonian damping proportional to T_E: $-r_{damp\_o}T_E$. 4) A deepening of the thermocline depth h related to a recharge favors positive T_E on time scales longer than 2-3 months: F_1h (see Supplementary Text S2 on mechanisms). Therefore:

$$\frac{dT_E}{dt} = R_o T_E + F_1 h \qquad (1)$$

the first term representing the net effect of Bjerkness positive feedback and Newtonian damping ($R_o = R_{BJ\_o} - r_{damp\_o}$) and the 2nd one, F_1, the recharge/discharge influence on T_E (following Jin et al. 2020 notations).

Concerning dh/dt equation, in the RO, the slow recharge mode response is the temporal integral of τ_x (e.g. Izumo et al. 2014). Negative T_E associated with easterly anomalies will progressively recharge the equatorial Pacific (see section 3.3 and Suppl. Text S2). This is formalized as a term $-F_{2\_o}T_E$ in the dh/dt equation:

$$\frac{dh}{dt} = -F_{2_o} T_E - \varepsilon_o h \qquad (2)$$

the second term $-\varepsilon_o h$ being formally a Newtonian damping, expected to be weak.

I.e. the tendency equation for the vector $\begin{pmatrix} T_E \\ h \end{pmatrix}$ in matrix form is:

$$\frac{d}{dt}\begin{pmatrix} T_E \\ h \end{pmatrix} = \begin{pmatrix} R_o & F_1 \\ -F_{2_o} & -\varepsilon_o \end{pmatrix}\begin{pmatrix} T_E \\ h \end{pmatrix} \qquad (3)$$

This RO linear equation is simple, but it remains unclear which geographical box to use for h. Furthermore, a standard multivariate linear regression fit minimizing rms error (second method in Burgers et al. 2005) gives significantly-different coefficients for the classical normalized metrics h_w and h_{eq}: $R_o=\varepsilon_o=+0.06\pm0.04$ and -0.18 ± 0.05 month^{-1}, $F_1=F_2=0.15\pm0.04$ and 0.25 ± 0.05 month^{-1} respectively. The fact that R_o and ε_o can change sign so easily for different classical recharge metrics makes it hard to interpret it physically. It evidences some physical inconcistency (Fig 1).

To resolve these geographical debate and physical issues objectively, an empirical way to find the best h is to have no *a priori* on the best region and index. Thus, one may intuitively think of searching for the $h(x,y,t)$ region that statistically optimizes the skill of equation (3). We will actually show later that we need two steps to construct a better index: (1) independence to T_E, (2) geographical optimization of the averaging region. Let us first try the geographical optimization:

$$\frac{d}{dt}\begin{pmatrix} T_E(t) \\ h(x,y,t) \end{pmatrix} = \begin{pmatrix} R_o(x,y) & F_1(x,y) \\ -F_{2_o}(x,y) & -\varepsilon_o(x,y) \end{pmatrix}\begin{pmatrix} T_E(t) \\ h(x,y,t) \end{pmatrix} + \begin{pmatrix} residual_T(x,y,t) \\ residual_h(x,y,t) \end{pmatrix} \qquad (4)$$

In practice we do at each spatial point (x,y) a fit with two multivariate linear regressions, one for equation (1) and one for equation (2), with Pearson correlation skill r_{Te} and r_h respectively. The squareroot of the RV-coefficient (Rho-Vectoriel; Robert and Escouffier 1976), "$r_{Te,h}$", is the equivalent of the Pearson correlation for matrix form, i.e. a measure of the skill for this 2D equation. Where the skill is the highest (residuals variance minimized) should inform us of the best h region to capture the RO processes. The highest skill is in the equatorial and south-western tropical Pacific (Fig. 2i), suggesting that the best region for averaging $h(x,y,t)$ could combine these two regions. However, the regression coefficients contributions have a puzzling spatial distribution. They are highly spatially-correlated with large opposing signs. In the central Equatorial Pacific, T_E and h influences on dT_E/dt would be very large negatively (R_o) and positively (F_1) respectively, which is unphysical. This is actually a statistical artefact because $T_E(t)$ and $h(x,y,t)$ are not statistically independent (i.e. not orthogonal), which makes it difficult to interpret physically the regression coefficients. It therefore suggests that the (T_E,h) basis is not ideal.

3.2 *Towards new RO tendency equations for T_E and independent h_{ind}*

A large component of h responds rapidly to τ_x and T_E. This fast response component of h is correlated to (i.e. linearly dependent of) T_E at timescales longer than ~2-3 months in the equatorial Pacific: Fig. 1e shows the downwelling pattern along the central to eastern equatorial Pacific and the zonal seesaw pattern in the western Pacific, i.e., the fast mode. We suggest to regress out from $h(x,y,t)$ this rapidly-responding component correlated to $T_E(t)$, so as to only focus on the slowly-responding independent component of h, hereafter h_{ind} (Fig. 3b of Izumo et al. 2018a). h_{ind} corresponds to the slow recharge mode related to the slower basin adjustment in disequilibrium with wind stress (e.g. Jin 1997ab, Alory and Delcroix 2002, Masuda et al. 2009, Clarke 2010, Fedorov 2010, Thual et al. 2013, Zhu et al. 2017, Izumo et al. 2018a). Thus, we can formally decompose h:

$$h(x,y,t) = h_{fast\_mode(dependent,\ corr.\ to\ Te)} + h_{slow\_recharge\_mode(independent,\ uncorr.\ to\ Te)}$$

$$h(x,y,t) = K(x,y)\ T_E(t) + h_{ind}(x,y,t) \quad (5)$$

$K(x,y)$ being the regression coefficient of h onto T_E (Fig. 1b).

h_{ind} is the pure recharge component, that allows us to describe the system with an orthonormal basis: (T_E,h_{ind}) (Fig. 1b). An orthonormal basis is more acceptable physically and mathematically.

With this linear transform, equation (4) can then be rewritten as:

$$\frac{d}{dt}\begin{pmatrix} T_E(t) \\ h_{ind}(x,y,t) \end{pmatrix} = \begin{pmatrix} \mathrm{R}(x,y) & F_1(x,y) \\ -F_2(x,y) & -\varepsilon(x,y) \end{pmatrix}\begin{pmatrix} T_E(t) \\ h_{ind}(x,y,t) \end{pmatrix} + \begin{pmatrix} residual_T(x,y,t) \\ residual_{h_{ind}}(x,y,t) \end{pmatrix} \quad (6)$$

with $R = R_o + K\,F_1$, $F_2 = F_{2\_o} + \varepsilon_o\,K + K\,R_o + K^2 F_1$, $\varepsilon = \varepsilon_o + KF_1$ and $residual_{h_{ind}} = residual_h - K\ residual_T$.

So the F_1 term representing the recharge influence on T_E remains the same for h_{ind}, and the physics behind is also the same. F_2 represents the influence of easterly anomalies related to T_E on the h_{ind} recharge. It differs from F_{2_o}.

Fig. 2cdgh shows the coefficients obtained from the multivariate regression onto (T_E,h_{ind}). R and ε are uniformly negligible (~0.00±0.03 month^{-1}). F_1 and F_2 vary spatially similarly, with highest values in the equatorial and southwest Pacific corresponding to the highest skill (Fig. 2k and Supplementary Fig. S1). Hence the best h_{ind} region should combine these two regions. The picture is also clearer and more consistent with RO theory: ε is negligible, consistent with a weak damping (due mainly to oceanic mixing, e.g. Fedorov 2010). r_{Te,h_ind} is clearly larger than $r_{Te,h}$ (Fig. 2i), further confirming that (T_E,h_{ind}) is a better basis of RO phase space.

Here we explain the spatial patterns of Fig. 2's various panels (full explanation in Suppl. Text S2), robust among datasets and periods (cf. Suppl. Fig. S2), and why there are similarities among some of them. F_2 map shows us how $h_{ind}(x,y,t)$ would look like if ENSO windstress anomalies would blow for a long time, e.g. because of long-lasting La Niña conditions. The slow recharge is as expected in the western and central equatorial Pacific, through: 1) downwelling equatorial Rossby waves to the west (Wyrtki 1985, Jin 1997ab), and off-equatorial ones in the southwest; 2) upwelling equatorial Kelvin waves to the east forcing coastal Kelvin waves propagating poleward along the eastern boundary and thus a leakage of negative OHC anomalies towards the poles along the eastern boundary (Wyrtki 1985, Izumo et al. 2018a).

Interestingly the recharge is also in the southwest, because of asymmetric Ekman pumping (poleward shift of the South Pacific Convergence Zone; SPCZ) forcing locally downwelling (in the La Niña case) and thus slow off-equatorial downwelling Rossby waves progressively recharging the southwestern Pacific. Note that the western boundary coastline meridional asymmetry would conversely favor a larger northwest recharge, as shown by sensitivity experiments with the LCS model (Linear Continuously Stratified model; McCreary 1980; configuration of Izumo et al. 2018; Suppl. Fig. S3). F_1 physically represents the slow recharge mode influence on T_E through several mechanisms (Suppl. Text S2).

R_o, ε_o and F_{2_o} spatial patterns can be explained as follow: $R_o \approx \varepsilon_o \approx -KF_1$ and $F_{2\_o} \approx F_2 + K^2 F_1$ (as $\varepsilon \approx 0$ and $R \approx 0$). These relationships explain the paradox with large positive and negative R_o and ε_o values found for classical h_w and h_{eq} respectively: $R_o \approx \varepsilon_o > 0$ ($K<0$) for h_w and $R_o \approx \varepsilon_o < 0$ ($K>0$) for h_{eq}, even if the actual damping ε and net feedback R are weak. So the "damping" term (ε_o) and the "positive feedback" (R_o) would be artificially large. This further confirms that we should use h_{ind} rather than the full h polluted by the fast mode that biases the RO model.

Such transforms from observable variables to new variables that are more relevant physically are often done in physics and geosciences, e.g. to create potential temperature, relative SST, rotating PCs (Takahashi et al. 2011), decomposition into baroclinic modes, spherical harmonics... We just want to clearly isolate the independent influence of the slow recharge mode from the fast tilt mode, and to optimize the RO understanding and metrics (we are not trying to add new physical processes).

Considering h_{ind} instead of h allows us to go beyond the RO metrics debate, by reconciliating indices: they become much closer when considering their independent component. E.g. h_{ind_w} and h_{ind_eq} are much better correlated (r^2 increases from 10% to 69%) with now similar coefficients (Suppl. Figs S4, S5 and Table S2). And Z20 and SLA-based indices become even closer (Suppl. Fig. S6). Trajectories in the (T_E,h_{ind}) phase diagram (Kessler 2002, Dommenget and Al Ansari 2022) become also closer (Fig. 3abdefh; shown here for longer ORAS5 SLA dataset to have a larger density; see Supplementary Figs S7 and S8 for satellite SLA and ORAS5 Z20). The transform is geometrically simply a linear transform of coordinates from the (T_E, h) to the (T_E, h_{ind}) orthogonal coordinate space. Trajectories in the latter space are more circular, closer to idealised RO model circular trajectory.

To conclude, h_{ind} reconciles various recharge metrics such as h_{ind_w} and h_{ind_eq}. Here below we would like to further improve them by objectively finding h_{ind} optimal averaging region.

3.3 Objectively finding the optimal region : Equatorial + South-West (h_{ind_eq+sw})

The optimal region for $h_{ind}(x,y,t)$ will be the one optimizing the skill r_{Te,h_ind} of the RO tendency equation (6) for $\begin{pmatrix} T_E \\ h_{ind} \end{pmatrix}$:

$$\frac{d}{dt}\begin{pmatrix} T_E \\ h_{ind} \end{pmatrix} = \begin{pmatrix} R & F_1 \\ -F_2 & -\varepsilon \end{pmatrix}\begin{pmatrix} T_E \\ h_{ind} \end{pmatrix} \qquad (7)$$

r_{Te,h_ind} (Fig. 2k) is the largest in the equatorial (~5°N-5°S) and southwest (~5°S-15°S) Pacific. The combination of these two boxes should be our optimal region, and r_{Te,h_ind} should be even better thanks to the spatial averaging (reducing noise). To choose objectively the optimal averaging region, we have tested several options. We have firstly tested rectangular boxes, e.g. by averaging $h_{ind}(x,y,t)$ from the usual fixed 120°E western end to a varying eastern edge, for various latitudinal bands. For the classical 5°N-5°S band, the best skill ($r_{Te,h\_ind}=0.61$) is found for an eastern edge around ~90-80°W (Fig. 2l, black line), consistent with classical h_{eq} (5°N-5°S, 120°E-80°W). Another relevant band is the 5°N-15°S, 120ºE to ~150ºW, leading to a similarly high skill (not shown). A "hybrid" choice with two rectangles is even better. Adding to the optimal classical equatorial band (5°N-5°S, 120°E-80°W) a second box in the southwest, along the 5°S-15°S band, from the same 120ºE western edge to a varying eastern edge, further improves the skill to r_{Te,h_ind}=0.69 for an optimal eastern edge around ~170°W (Fig. 2l, red line; we have similarly tested all the other edges of this two-rectangle region, also testing ORAS5 SLA and Z20; Suppl. Figs S9ab and S10). This is our best allround and sufficiently-simple index, hereafter h_{ind_eq+sw} ($h_{ind\_eq+sw}=h_{eq+sw\_ind}$, the regression being linear; interestingly, this choice also conveniently minimizes the correlation between h_{eq+sw} and T_E (Suppl. Fig. S9cd), making h_{eq+sw} closer to h_{ind_eq+sw}). The skill increase from h_{ind_eq} to h_{ind_eq+sw} is statistically significant, more thanks to r_{h_ind} (p=0.01|0.001|0.001) than to r_{Te} (p=0.09|0.01|0.23) for satSLA|ORAS5_SLA|ORAS5_Z20 respectively). And trajectories are even smoother (Fig. 3g). Including the southwest agrees with Izumo et al. (2018a; see also Santoso et al. 2017; Ramesh and Murtugudde 2013). h_{ind_eq+sw} is now our default best recharge index.

Using normalized h_{ind_eq+sw}, we obtain: $F_1 \approx F_2 \approx 0.17 \pm 0.03$, $R \approx \varepsilon \approx 0.00 \pm 0.03$ month^{-1}. These coefficients are robust among datasets and periods (Suppl. Table S1). R and ε are negligible. This reduces the parameter space. And the RO equations system has the form of a harmonic oscillator excited by noise (Burgers et al. 2005), with $d^2X(t)/dt^2 \approx -F_1F_2\,X(t)$, $X(t)$ being T_E or h_{ind}, and angular frequency being the Wyrtki index $W=(F_1F_2)^{1/2} \approx F_1 \approx F_2$ (i.e. reasonable eigenperiod of ~3.1 year). Even if theoretically undamped, the oscillator is actually still damped because $residual_T$ is not a pure red noise but includes non-linear terms neglected in our 1st order linear approximation (Supplementary Table S2).

4 Conclusion

4.1 Summary

Here we have defined a simple Pacific recharge index h_{ind} *independent* of the fast mode, by regressing out T_E-related variability. It unambiguously represents the slow recharge mode, with more physically-consistent RO parameters, conversely to classical recharge indices. h_{ind} harmonizes recharge indices: by taking h_{ind}, averages over the usual western/equatorial Pacific regions based on Z20/OHC/SLA have much more similar time series, equation parameters and phase trajectories, so that all indices converge to a single one. We have also objectively searched for the optimal averaging region to have the most realistic RO tendency equations. The optimal index is h_{ind_eq+sw} averaged over the classical equatorial band (5°N-5°S, 120°E-80°W), extended in the southwest until 15°S (5°S-15°S, 120°E-170°W).

In practice, obtaining h_{ind_eq+sw} is straightforward: 1) average over *'eq+SW'* box SLA and then normalize, 2) remove its dependent part KT_E: $h_{ind\_eq+sw} = h_{eq+sw} - K_{eq+sw}\, T_E$, with regression coefficient $K_{eq+sw} \approx 0.26$ (details in Suppl. Text S2).

4.2 Discussion

For ENSO operational diagnostics, the (T_E, h_{ind_eq+sw}) basis is more relevant to describe the system trajectory than the usual (T_E, h_{eq}) and (T_E, h_w) pairs (Fig. 3). If the latter are used, a situation with anomalous h is ambiguous. Independent h_{ind_eq+sw} better represents actual precursory recharge anomalies. Only with h_{ind} sign (recharged/discharged state) can one get dT_E/dt sign directly (while h sign is not sufficient, as the fast tilt mode signal could blur a weaker long-term build-up; see yellow/blue dots in idealised schematics Fig. 3dh).

To compare h_{ind_eq} and h_{ind_eq+sw} forecasting skills, we have done a preliminary assessment by using the simple multivariate linear regression model combining the recharge index, Indian Ocean Dipole (IOD) index and T_E, all in September-November, to hindcast T_E peak in November-January 14 months later. Adding the Southwest improves the skill for all datasets/periods, more clearly for SLA than for Z20 (Suppl. Table S3; since we have statistically-significant contributions from h_{ind_eq+sw} and IOD, but not from T_E itself, this new recharge index and updated datasets confirm earlier studies of Izumo et al. 2010, 2014, 2016, Dayan et al. 2014, Jourdain et al. 2016).

Results are robust among datasets. Only some subtle 2nd order differences remain between SLA and Z20 (OHC inbetween) for h_{ind}, likely related to different weighting of the first baroclinic modes (Suppl. Figs S2, S6, S8, S10). Note that we are not trying to address "how to best estimate the thermocline depth" (Vijayeta, 2020), but rather "how to best isolate the recharge mode". SLA and Z20 advantages are equivalent with presently-available datasets: Z20 seems better in a perfectly-observed and non-warming ocean (Suppl. Tables S1 and S3), but SLA is observed globally by satellite in near real-time and is more available in climate models outputs (e.g. Coupled Model Intercomparison Project, CMIP).

Here we have neglected possible seasonal cycles of parameters and asymmetries/non-linearities for the sake of simplicity. Knowing ENSO seasonal phase-locking, taking into account such seasonal cycles could be one next step. We could include asymmetrical/non-linear terms (Supplementary Table S2). To have a simple h_{ind} definition, we have also neglected a weak non-linearity in the fast response of h to T_E (Figs 3c, S7c, S8c) likely due to τ_x anomalies longitudinal position, further to the east during El Niño than during La Niña, favoring a larger rapid discharge. This non-linearity would lead to a quadratic term in the dh_{ind}/dt equation.

Our method is based on a statistical optimisation of regression coefficients from observations to fit the RO model, and on an orthogonal basis, which is conceptually satisfying, and practically necessary to determine eigenmodes. But correlations do not mean causality, and this study therefore does not aim to challenge RO physical validity. Although using the physical h variable is necessary to infer causality and derive RO equations, we still argue that having orthogonal co-ordinates thanks to h_{ind} allows for clearer recharge diagnostics, and a properly isolated recharge mode.

Some studies, using usual h indices, have put in question the RO and suggested that the delayed oscillator (DO, Suarez and Schopf 1988, Battisti and Hirst 1989) is more realistic (e.g. Linz et al. 2014, Graham et al. 2015). One reason for this "RO vs DO" debate could be the mis-leading character of usual h indices because of fast tilt mode (dependent) component, which tends to artificially reduce RO skill. It would be interesting to compare the RO to the other ENSO oscillators, using h_{ind_eq+sw} instead, to clarify whether theories, climate models and observations would better agree.

Acknowledgements. MC would like to thank Steven Sherwood for his essential support, supervision, and mentoring, in particular throughout the pandemic, which made this collaboration possible. We would like to thank the two reviewers, notably Dr. Dietmar Dommenget for his critical thinking and interesting discussions, who have pushed us to improve, clarify and strengthen the results, as well as Dr. Matthieu Lengaigne, Dr. Jérôme Vialard, and Pr. Fei-Fei Jin for discussions. T.I. would like to thank UPF for hosting M.C. because of COVID-related closed Australian borders; this latter fact was paradoxically essential for allowing us to make this work mature.

Open Research (Data Availability Statement). All data used is open data: OISSTv2 (https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html), CMAP1 (https://psl.noaa.gov/data/gridded/data.cmap.html) ERA5-ORAS5 (https://www.ecmwf.int), Copernicus SLA (https://doi.org/10.48670/moi-00148). We acknowledge use of NOAA pyferret open source software for analyses and figures.

References

Alory, G., & Delcroix, T. (2002). Interannual sea level changes and associated mass transports in the tropical Pacific from TOPEX/Poseidon data and linear model results (1964–1999). *Journal of Geophysical Research: Oceans*, *107*(C10), 17-1.

An, S. I., Tziperman, E., Okumura, Y. M., & Li, T. (2020). ENSO irregularity and asymmetry. *El Niño Southern Oscillation in a changing climate*, 153-172.

Ballester, J., Bordoni, S., Petrova, D., & Rodó, X. (2015). On the dynamical mechanisms explaining the western Pacific subsurface temperature buildup leading to ENSO events. *Geophysical Research Letters*, *42*(8), 2961-2967.

Ballester, J., Petrova, D., Bordoni, S., Cash, B., García-Díez, M., & Rodó, X. (2016a). Sensitivity of El Niño intensity and timing to preceding subsurface heat magnitude. *Scientific reports*, *6*(1), 1-9.

Ballester, J., Bordoni, S., Petrova, D., & Rodó, X. (2016b). Heat advection processes leading to El Niño events as depicted by an ensemble of ocean assimilation products. *Journal of Geophysical Research: Oceans*, *121*(6), 3710-3729.

Barnston, A. G., Tippett, M. K., L'Heureux, M. L., Li, S., & DeWitt, D. G. (2012). Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing?. *Bulletin of the American Meteorological Society*, *93*(5), 631-651.

Barnston, A. G., Tippett, M. K., Ranganathan, M., & L'Heureux, M. L. (2019). Deterministic skill of ENSO predictions from the North American Multimodel Ensemble. *Climate Dynamics*, *53*(12), 7215-7234.

Battisti, D. S., & Hirst, A. C. (1989). Interannual variability in a tropical atmosphere–ocean model: Influence of the basic state, ocean geometry and nonlinearity. *Journal of Atmospheric Sciences*, *46*(12), 1687-1712.

Boschat, G., Terray, P., & Masson, S. (2013). Extratropical forcing of ENSO. *Geophysical Research Letters*, *40*(8), 1605-1611.

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., & Bladé, I. (1999). The effective number of spatial degrees of freedom of a time-varying field. *Journal of climate*, *12*(7), 1990-2009.

Burgers, G., F.-F. Jin, and G. J. van Oldenborgh (2005), The simplest ENSO recharge oscillator, Geophys. Res. Lett., 32, L13706, doi:10.1029/2005GL022951.

Capotondi, A., Wittenberg, A. T., Newman, M., Di Lorenzo, E., Yu, J. Y., Braconnot, P., ... & Yeh, S. W. (2015). Understanding ENSO diversity. *Bulletin of the American Meteorological Society*, *96*(6), 921-938.

Cibot, C., Maisonnave, E., Terray, L., & Dewitte, B. (2005). Mechanisms of tropical Pacific interannual-to-decadal variability in the ARPEGE/ORCA global coupled model. *Climate dynamics*, *24*(7), 823-842.

Clarke, A. J., Van Gorder, S., & Colantuono, G. (2007). Wind stress curl and ENSO discharge/recharge in the equatorial Pacific. *Journal of Physical Oceanography*, *37*(4), 1077-1091.

Clarke, A. J. (2008). *An introduction to the dynamics of El Niño and the Southern Oscillation*. Elsevier.

Clarke, A. J. (2010). Analytical theory for the quasi-steady and low-frequency equatorial ocean response to wind forcing: The "tilt" and "warm water volume" modes. *Journal of physical oceanography*, *40*(1), 121-137.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., & Bechtold, P. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597.

Dommenget, B. D., & Al-Ansari, M. (2022). Asymmetries in the ENSO phase space. *Climate Dynamics*, 1-20.

Fedorov, A. V. (2010). Ocean response to wind variations, warm water volume, and simple models of ENSO in the low-frequency approximation. *Journal of climate*, *23*(14), 3855-3873.

Gadgil, S., Joseph, P. V., & Joshi, N. V. (1984). Ocean–atmosphere coupling over monsoon regions. *Nature*, *312*(5990), 141-143.

Gasparin, F., & Roemmich, D. (2017). The seasonal march of the equatorial Pacific upper-ocean and its El Niño variability. *Progress in Oceanography*, *156*, 1-16.

Graham, F. S., Brown, J. N., Wittenberg, A. T., & Holbrook, N. J. (2015). Reassessing conceptual models of ENSO. *Journal of Climate*, *28*(23), 9121-9142.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D. and Simmons, A., 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), pp.1999-2049.

Izumo, T. (2005). The equatorial undercurrent, meridional overturning circulation, and their roles in mass and heat exchanges during El Niño events in the tropical Pacific Ocean. *Ocean Dynamics*, *55*(2), 110-123.

Izumo, T., Vialard, J., Dayan, H., Lengaigne, M., & Suresh, I. (2016). A simple estimation of equatorial Pacific response from windstress to untangle Indian Ocean Dipole and Basin influences on El Niño. *Climate dynamics*, *46*(7-8), 2247-2268.

Izumo, T., Lengaigne, M., Vialard, J., Suresh, I., & Planton, Y. (2018a). On the physical interpretation of the lead relation between Warm Water Volume and the El Niño Southern Oscillation. *Climate Dynamics*, *52*(5), 2923-2942.

Izumo, T., Khodri, M., Lengaigne, M., & Suresh, I. (2018b). A subsurface Indian Ocean dipole response to tropical volcanic eruptions. *Geophysical Research Letters*, *45*(17), 9150-9159.

Izumo, T., Vialard, J., Lengaigne, M., & Suresh, I. (2020). Relevance of relative sea surface temperature for tropical rainfall interannual variability. *Geophysical Research Letters*, *47*(3), e2019GL086182.

Jin, F. F. (1997a). An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model. *Journal of the atmospheric sciences*, *54*(7), 811-829.

Jin, F. F. (1997b). An equatorial ocean recharge paradigm for ENSO. Part II: A stripped-down coupled model. *Journal of Atmospheric Sciences*, *54*(7), 830-847.

Jin, F.-F., Lin, L., Timmermann, A., & Zhao, J. (2007). Ensemble- mean dynamics of the ENSO recharge oscillator under state- dependent stochastic forcing. *Geophysical Research Letters*, *34*(3), L03807. https://doi.org/10.1029/2006GL027372

Jin, F. F., Chen, H. C., Zhao, S., Hayashi, M., Karamperidou, C., Stuecker, M. F., ... & Geng, L. (2020). Simple ENSO models. *El Niño Southern Oscillation in a Changing Climate*, 119-151.

Johnson, N. C., & Xie, S.-P. (2010). Changes in the sea surface temperature threshold for tropical convection. *Nature Geoscience*, *3*(12), 842–845.

Johnson, N. C., & Kosaka, Y. (2016). The impact of eastern equatorial Pacific convection on the diversity of boreal winter El Niño teleconnection patterns. *Climate Dynamics*, *47*(12), 3737-3765.
Kessler, W. S. (2002). Is ENSO a cycle or a series of events?. *Geophysical Research Letters*, *29*(23), 40-1.
Khodri, M., Izumo, T., Vialard, J., Janicot, S., Cassou, C., Lengaigne, M., ... & McPhaden, M. J. (2017). Tropical explosive volcanic eruptions can trigger El Niño by cooling tropical Africa. *Nature communications*, *8*(1), 1-13.
Kumar, A., & Hu, Z. Z. (2014). Interannual and interdecadal variability of ocean temperature along the equatorial Pacific in conjunction with ENSO. *Climate dynamics*, *42*(5), 1243-1258.
Lai AWC, Herzog M, Graf HF (2015) Two key parameters for the El Niño continuum: zonal wind anomalies and Western Pacific sub- surface potential temperature. Clim Dyn 45(11–12):3461–3480. https://doi.org/10.1007/s00382-015-2550-0
Lengaigne M, Boulanger J-P, Menkes C, Spencer H (2006) Influence of the seasonal cycle on the termination of El Niño events in a coupled general circulation model. J Clim 19(9):1850–1868. https ://doi.org/10.1175/JCLI3706.1
Linz M, Tziperman E, MacMartin DG (2014) Process-based analysis of climate model ENSO simulations: intermodel consistency and compensating errors. J Geophys Res 119:7396–7409. https://doi. org/10.1002/2013JD021415
Masuda, S., Awaji, T., Toyoda, T., Shikama, Y., & Ishikawa, Y. (2009). Temporal evolution of the equatorial thermocline associated with the 1991–2006 ENSO. *Journal of Geophysical Research: Oceans*, *114*(C3).McGregor, S., N. Ramesh, P. Spence, M. H. England, M. J. McPhaden, and A. Santoso (2013), Meridional movement of wind anomalies during ENSO events and their role in event termination, Geophys. Res. Lett., 40, 749–754, doi:10.1002/grl.50136.
Meinen CS, McPhaden MJ (2000) Observations of warm water volume changes in the equatorial Pacific and their relationship to El Niño and La Niña. J Clim 13:3551–3559
Neelin, J. D., Battisti, D. S., Hirst, A. C., Jin, F. F., Wakata, Y., Yamagata, T., & Zebiak, S. E. (1998). ENSO theory. *Journal of Geophysical Research: Oceans*, *103*(C7), 14261-14290.
Neske S, McGregor S (2018) Understanding the warm water volume precursor of ENSO events and its interdecadal variation. Geophys Res Lett 45:1577–1585. https://doi.org/10.1002/2017GL076439
Okumura, Y. M. (2019). ENSO diversity from an atmospheric perspective. *Current Climate Change Reports*, *5*(3), 245-257.
Palanisamy, H., Cazenave, A., Delcroix, T., & Meyssignac, B. (2015). Spatial trend patterns in the Pacific Ocean sea level during the altimetry era: the contribution of thermocline depth change and internal climate variability. *Ocean Dynamics*, *65*(3), 341-356.
Petrova, D., Koopman, S. J., Ballester, J., & Rodó, X. (2017). Improving the long-lead predictability of El Niño using a novel forecasting scheme based on a dynamic components model. *Climate Dynamics*, *48*(3-4), 1249-1276.
Ramesh N, Murtugudde R (2013) All flavours of El Niño have similar early subsurface origins. Nat Clim Change 3(1):42–46. https://doi. org/10.1038/nclimate1600
Reynolds, R.W., N.A. Rayner, T.M. Smith, D.C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. J. Climate, 15, 1609-1625.

Rebert, J. P., Donguy, J. R., Eldin, G., & Wyrtki, K. (1985). Relations between sea level, thermocline depth, heat content, and dynamic height in the tropical Pacific Ocean. *Journal of Geophysical Research: Oceans*, *90*(C6), 11719-11725.

Escoufier, R., & Robert, P. (1976). A unifying tool for linear multivariante statistical methods: The Rv-coefficient. *Applied Statistics*, *25*(3), 257-265.

Santoso, A., Mcphaden, M. J., & Cai, W. (2017). The defining characteristics of ENSO extremes and the strong 2015/2016 El Niño. Reviews of Geophysics, 55, 1079–1129.

Suarez MJ, Schopf PS (1988) A delayed action oscillator for ENSO. J Atmos Sci 45:3283–3287

Takahashi K, Montecinos A, Goubanova K, Dewitte B (2011) ENSO regimes: reinterpreting the canonical and Modoki El Niño. Geo- phys Res Lett 38:1–5. https://doi.org/10.1029/2011GL047364

Thual S,.Dewitte B, Ayoub N, Thual O (2013) An asymptotic expansion for the recharge–discharge model of ENSO. J Phys Oceanogr 43:1407–1416. https://doi.org/10.1175/JPO-D-12-0161.1

Timmermann, Axel, et al. "El Niño–southern oscillation complexity." *Nature* 559.7715 (2018): 535-545.

Van Oldenborgh, G. J., Hendon, H., Stockdale, T., L'Heureux, M., De Perez, E. C., Singh, R., & Van Aalst, M. (2021). Defining El Niño indices in a warming climate. *Environmental Research Letters*, *16*(4), 044003.

Vijayeta Asha, supervised by D. Dommenget and S. McGregor (2020), ENSO in a changing climate simulations. PhD thesis. Monash University, Australia.

Wang, C., & Picaut, J. (2004). Understanding ENSO physics—A review. *Earth's Climate: The Ocean–Atmosphere Interaction, Geophys. Monogr*, *147*, 21-48.

Williams, I. N., & Patricola, C. M. (2018). Diversity of ENSO events unified by convective threshold sea surface temperature: A nonlinear ENSO index. *Geophysical Research Letters*, *45*(17), 9236-9244.

Wyrtki K (1985) Water displacements in the Pacific and the genesis of El Niño cycles. J Geophys Res 90:7129–7132.

Xie, P., & Arkin, P. A. (1997). Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society*, *78*(11), 2539–2558.

Yokoi, T., Tozuka, T., & Yamagata, T. (2008). Seasonal variation of the Seychelles Dome. *Journal of Climate*, *21*(15), 3740-3754.

Zhu, X., Greatbatch, R. J., & Claus, M. (2017). Interannual variability of tropical Pacific sea level from 1993 to 2014. *Journal of Geophysical Research: Oceans*, *122*(1), 602-616.

H. Zuo, M. A. Balmaseda, S. Tietsche, K. Mogensen, and M. Mayer. The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment*, Ocean Sci., 15, 779–808, https://doi.org/10.5194/os-15-779-2019, 2019.*

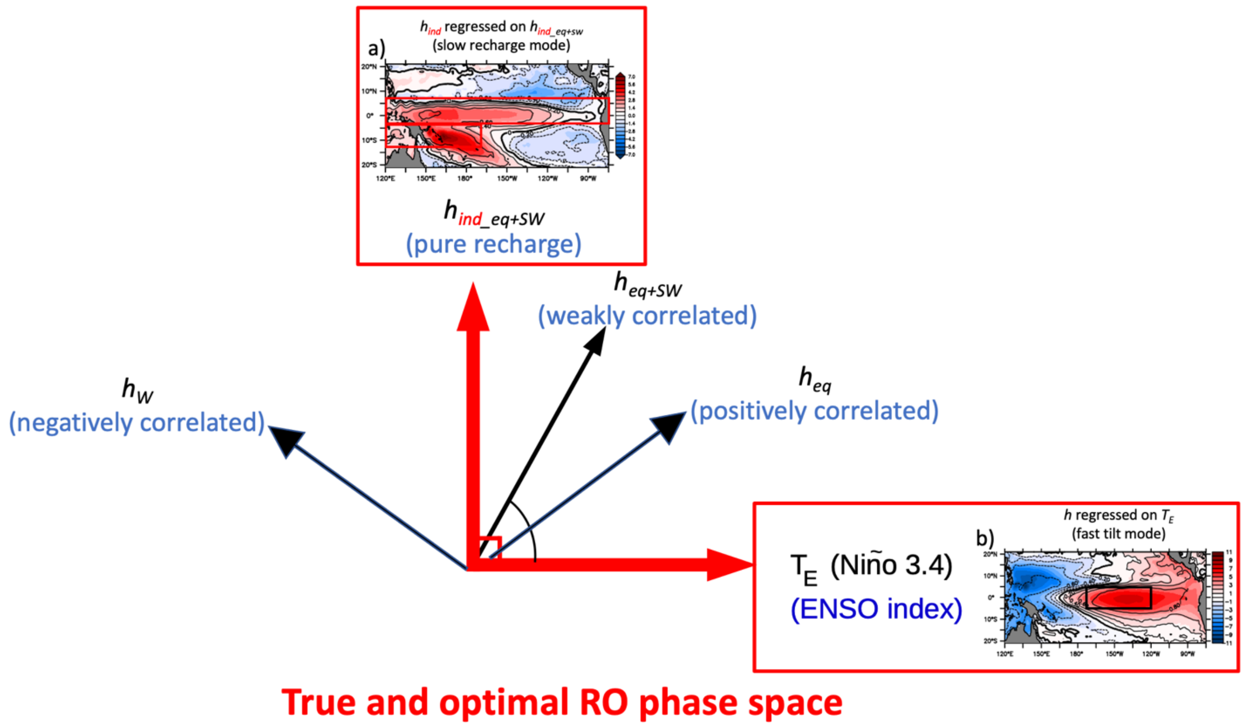

Fig. 1. Schematics of the optimal orthonormal phase space for the Recharge Oscillator (RO). Panels **a** and **b** show respectively the pure recharge mode (h_{ind} regressed onto onto our default best recharge index h_{ind_eq+sw}), independent of (uncorrelated to) T_E, and the fast zonal tilt mode dependent of T_E (h regressed onto T_E, i.e. K; h unit: satellite SLA in cm ; cf. sections S1 and 3 for details). Black arrows illustrate other used indices based on usual h. They are hence partly correlated to T_E.

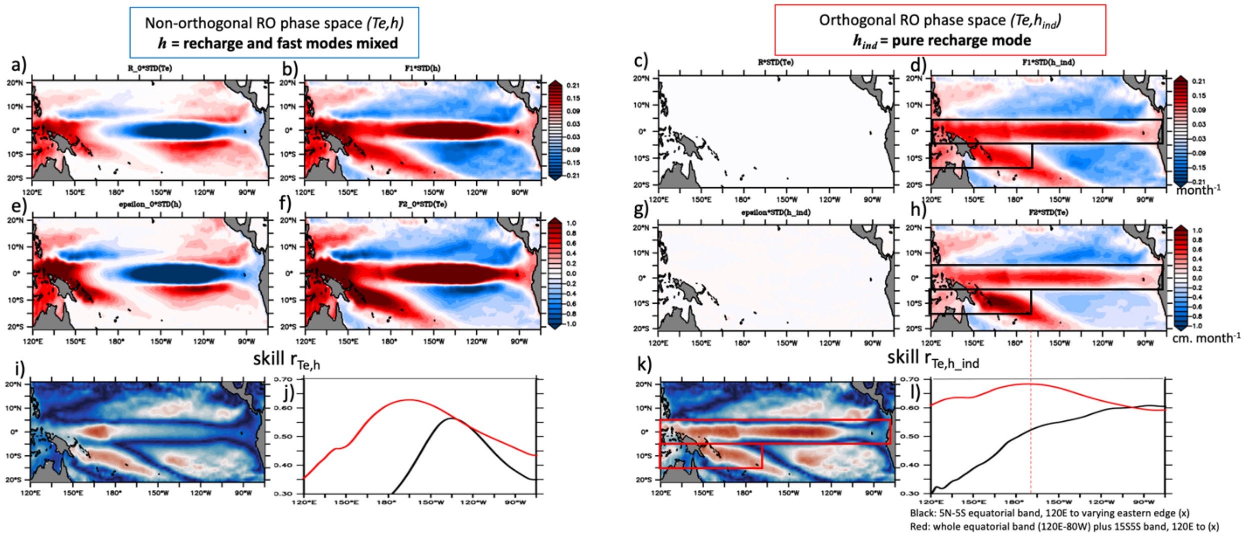

Fig. 2. Searching for the optimal recharge index h_{ind}, representing the actual slow recharge mode independent of the fast mode in RO equations. Usual non-orthogonal basis *(T_E,h)* in the left set of panels, and orthogonal basis *(T_E,h_{ind})* in the right one. Panels **a** and **b** show respectively the coefficients R_o(x,y) and F_1(x,y) for the multivariate regression of $dT_E/dt(t)$ onto $T_E(t)$ and h(x,y,t) (equation 4), multiplied by the STDs of T_E (=1) and of *h(x,y,t)* respectively, to measure their respective contributions to $dT_E/dt(t)$. Panels **e** and **f** are as a and c, but for coefficients for *dh(x,y,t)/dt* regression. Panel **i** shows $r_{Te,h}$ skill using *h(x,y,t)*. Black line in panel **j** shows this skill for h_{ind} averaged over: the equatorial band (5°N-5°S) with its western edge fixed to 120°E and its eastern edge varying, given by the x-axis. Red line is for a two-rectangle region, the classical equatorial box (5°N-5°S, 120°E-80°W) plus a southern 5°S-15°S box with the same 120°E fixed western edge and its eastern edge varying. The right set of panels is the same for the here-developed orthogonal basis *(T_E,h_{ind})*. The boxes overlaid represents the two-rectangle equatorial+SouthWest (*eq+sw*) optimal region finally chosen to average h_{ind}. This defines the suggested improved index: $h_{ind\ eq+sw}$. See text for details.

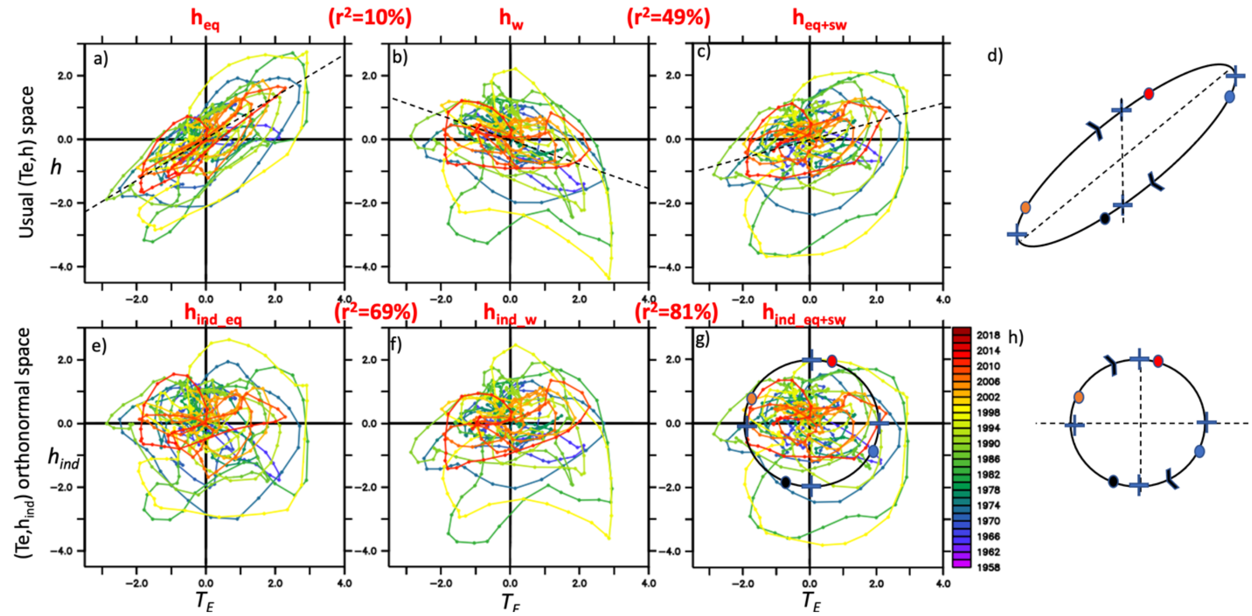

Fig. 3. Observed trajectories of the RO in the usual *(T_E,h)* space (upper panels) **and new orthonormal *(T_E,h_{ind})* space** (lower panels)**, for classical h_{eq} and h_w indices** (1st and 2nd columns)**, and optimal h_{eq+sw}** index (3[rd] column)**.** a) trajectory of the system for the pair of coordinates *(T_E,h_{eq})*, T_E in horizontal axis, *h* in vertical axis, each month being a point, with year indicated in color, for ORAS5 SLA. Other panels are similar, but for other recharge indices. Indices are all normalised. The shared variance (square of correlation) between recharge indices is shown in parenthesis. d) schematized elliptical trajectory if the recharge index is positively correlated to T_E. h) ideal RO circular trajectory, for its independent component. Color dots indicate different positions in the cycle.