Forecasting South China Sea Monsoon Onset Using Insight from Theory

Ruth Geen[1]

[1]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK.

Key Points:

- Recent theoretical studies have revealed feedbacks at work during monsoon onset.
- If the spring atmosphere is in a state where these feedbacks are easily triggered, onset is earlier.
- This insight is used to produce a simple forecast of South China Sea Monsoon onset timing.

Corresponding author: Ruth Geen, r.geen@exeter.ac.uk

Abstract

Monsoon onset over the South China Sea occurs in April–May, marking the start of the wet season over East Asia. Nevertheless, skillful prediction of onset timing remains an open challenge. Recently, theoretical studies using idealized models have revealed feedbacks at work during the seasonal transitions of the Hadley cells, and have shown that these are relevant to monsoon onset over Asia. Here I hypothesize that monsoon onset occurs earlier in years when the atmosphere over the South China Sea is already in a state where these feedbacks are more easily triggered. I find that local anomalies in lower-level moist static energy in the preceding Jan–March are well correlated with South China Sea Monsoon onset timing. This relationship remains consistent on decadal timescales, while correlations with other teleconnections vary, and is used to develop a simple forecast model for onset timing that shows skill competitive with that of more complex models.

Plain Language Summary

Arable land in China is estimated to feed 20% of the world's population on 7% of the world's farmlands, with much of this region watered by rainfall from the East Asian Summer Monsoon. Forecasting the arrival of the monsoon, which occurs first over the South China Sea in April–May, is therefore helpful for agricultural planning. However, producing a reliable forecast remains an open challenge. Recently, simple climate models, for example models including seasons but with no land, have been used along with observations to investigate the most basic processes controlling climatological monsoon onset over Asia. In this paper, I suggest that year-to-year variations in monsoon onset timing are controlled by the same processes and, based on these, I suggest climate variables that may help in predicting when the monsoon will begin. By using this approach, I find that the 'Moist Static Energy' (a quantity that combines temperature and humidity) over the South China Sea in Jan–March is strongly correlated with monsoon onset timing. This insight can be used to predict future onset timing competitively compared with more complex forecast models.

1 Introduction

The East Asian Summer Monsoon (EASM) affects up to one third of the world's population (Chang, 2004), and the arable land in China is estimated to feed 20% of the Earth's population on 7% of the world's farmlands (Cui & Shoemaker, 2018). The reliance of agriculture on EASM rainfall means that there is a strong drive to predict when rain will arrive.

In East Asia, the wet season begins with the abrupt reversal of the lower-level zonal wind over the South China Sea (SCS), accompanied by an off-equatorial shift in the tropical rainband over this area, known as the South China Sea Monsoon (SCSM) (Wang & LinHo, 2002). This is followed by the development of an east-west oriented front of precipitation, known as Meiyu in China, which brings intense rainfall to the Yangtze River valley from mid-June to mid-July (Wang & LinHo, 2002; Ding & Wang, 2005). SCSM onset timing is considered as a precursor to the broader-scale onset of the monsoon over East Asia (Lau & Yang, 1997; Wang et al., 2004) and additionally appears correlated to monsoon rainfall intensity over China, with an early onset associated with a drier summer (Huang et al., 2006; He & Zhu, 2015). The monsoon onset over the South China Sea varies interannually from late April to mid June (e.g. Wang et al., 2004, and see Fig. 1), giving over a month's uncertainty. In this paper I propose a novel approach to improving predictability of the SCSM onset, by exploring whether recent theoretical advances in monsoon dynamics can provide insight into controls on SCSM onset timing and inform seasonal forecasting efforts.

SCSM onset timing shows both interannual variability and slower, interdecadal trends (Kajikawa et al., 2012; Kajikawa & Wang, 2012, & Fig. 1). The El Niño Southern Oscillation (ENSO) has been identified as one clear source of interannual variability, with La Niña (El Niño) events linked to early (late) monsoon onset via their influence on the Western North Pacific subtropical high (Zhou & Chan, 2007). However, the strength of the relationship between the SCSM and ENSO varies dramatically on decadal timescales, likely influenced by the the Pacific Decadal Oscillation (PDO) (Chan & Zhou, 2005) and Atlantic Multidecadal Oscillation (AMO) (Fan et al., 2018). In addition to ENSO, SCSM onset variability has been related to a range of factors in the preceding winter and spring: thermal and mechanical forcing over the Tibetan Plateau (G. Wu & Zhang, 1998); temperature contrasts between the South China Sea and Western North Pacific and the land surface temperature to the north (P. Liu et al., 2009); and the cross-equatorial flow over the South China Sea (Lin et al., 2017; Hu et al., 2018).

Physical-Empirical models have been developed to predict SCSM onset timing and intensity based on correlations with SST, sea level pressure and temperature tendency anomalies in the preceding months. These models can show high forecast correlation skill over the time-periods analysed (e.g. $r = 0.72$, Zhu & Li, 2017). Recently, dynamical seasonal forecasting ensembles have also been found to give skillful prediction of the SCSM onset in hindcasts (Fan et al., 2016; Martin et al., 2019). However, the skill of both types of model stems from teleconnections such as ENSO, whose correlation with the SCSM vary decadally.

Here, I try a different approach. First, I identify the processes found to be most important to monsoon onset in idealized modeling studies (e.g. Bordoni & Schneider, 2008; Geen et al., 2018). I then explore whether these insights can help to identify direct, local precursors to the monsoon. My aim is to find common pathways via which multiple teleconnections affect onset timing. In Section 2, I motivate the precursors that I hypothesize to be relevant and detail the datasets used in the paper. In Section 3 I then examine the correlation between SCSM onset and these precursors over different time periods. In Section 4 I use develop a simple forecast of SCSM onset and test its predictive skill. Section 5 concludes.

2 Methods

2.1 Hypothesized Precursors

To select potential mechanistic pathways I apply results from idealized modeling studies that compare the seasonal behavior of the Hadley cell in aquaplanets with Earth's monsoons (Hill, 2019; Geen et al., 2020). One important expectation from this theoretical work is that, if the tropical atmosphere is near convective quasi-equilibrium (CQE) and the influence of extra-tropical eddies on the Hadley cell is weak, then the 0-streamfunction line separating the Hadley cells is colocated with the maximum in subcloud moist static energy (MSE). If this maximum is off the Equator, as is the case during the monsoon season, then the maximum ascent and associated rainfall will lie just equatorward of this (Privé & Plumb, 2007; Nie et al., 2010). MSE is defined

$$h \equiv c_p T + Lq + gz. \tag{1}$$

where c_p is the heat capacity at constant pressure, T is temperature, L is the latent heat of vaporization of water, q is specific humidity, g is the gravitational constant and z is geopotential height.

The connection between the overturning circulation and MSE distribution results in two feedbacks occurring during monsoon onset. First, diabatic heating by the insolation warms the summer hemisphere. In response, the ITCZ shifts into the summer hemisphere and the winter-hemisphere Hadley cell becomes cross-equatorial. This cross-equatorial

cell advects cooler, drier air up the MSE gradient, while diabatic processes increase MSE poleward. As a result, the MSE maximum shifts farther poleward and the cell becomes more cross-equatorial. The result is a positive feedback between the circulation and the thermal forcing, so that the convergence zone jumps abruptly into the summer hemisphere (Bordoni & Schneider, 2008). The second feedback relates to the tropical upper-level easterlies generated by a cross-equatorial Hadley cell. These limit the propagation of eddies to lower latitudes. As a result, the cell becomes primarily thermally-driven, rather than eddy-driven, and responds strongly to changes in the MSE distribution, strengthening the cell, and so further enhancing the easterlies (Schneider & Bordoni, 2008; Geen et al., 2019).

Although these ideas have been developed in an idealized framework, they appear to apply to both the climatology (Bordoni & Schneider, 2008; Geen et al., 2018; Nie et al., 2010; Ma et al., 2019) and interannual variability (Hurley & Boos, 2013) of the Asian monsoons in reanalysis data. In this study, I further hypothesize that in the months prior to monsoon onset, both local and remote influences may cause the atmosphere to be in a state where these feedback cycles will more readily begin, so that onset may then occur earlier in the season. Based on this, I suggest that early SCSM onset will be associated with positive 850-hPa MSE and negative 200-hPa zonal wind speed anomalies in the SCSM region.

2.2 Data and Metrics

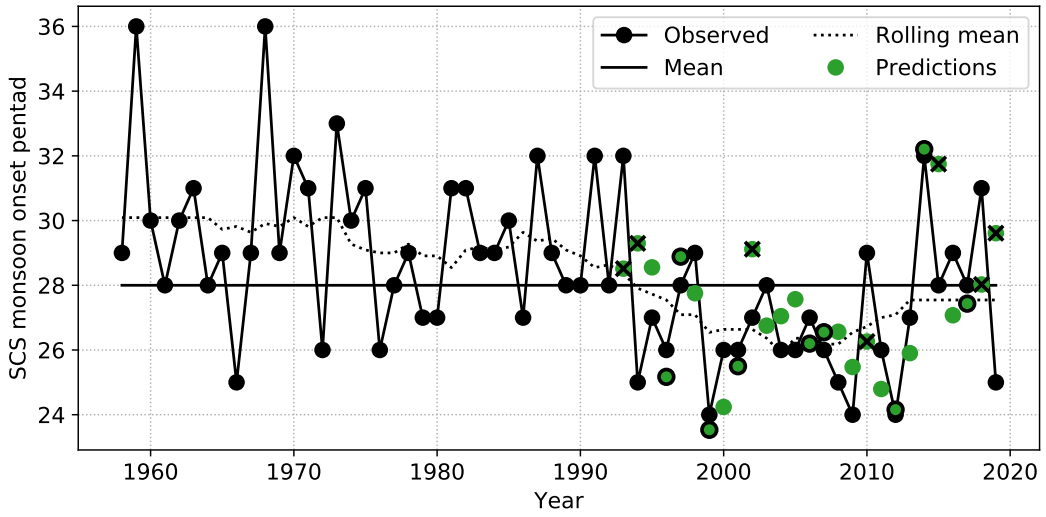

Figure 1. Onset timing of the SCSM in pentads (5-day means; black circles). The dotted line shows the 11-year rolling mean, which I use to distinguish longer term trends from interannual variability. The overall mean onset timing is pentad 28 (solid line). Green circles show predicted onset pentads for each year based on moist static energy over 5–15 °N, 110 – 125 °E. Details of how these forecasts are generated is given in Section 4. Where these are circled in black, they are within 1 pentad of the observed date. Where they are crossed, they are more than 2 pentads from the observed date.

Results are presented for the JRA-55 reanalysis dataset (Japan Meteorological Agency/Japan, 2013; Kobayashi et al., 2015) for years 1958-2019, with SSTs taken from the COBE SST dataset (Japan Meteorological Agency, 2006, Ongoing). Daily-mean 850-hPa zonal wind data were used to establish the SCSM onset pentad using the criteria developed by Wang

et al. (2004). SCSM onset is defined as the first pentad after 25th April (pentad 24) where the average zonal wind speed over 5–15 °N, 110–120°N, U_{SCS}, is westerly, and where U_{SCS} is positive in at least 3 of the 4 subsequent pentads (including the onset pentad) and the accumulative 4-pentad mean of $U_{SCS} > 1$ ms^{-1}. The onset pentads identified in the JRA-55 dataset are shown by the black circles in Fig. 1. These are broadly consistent with dates evaluated in previous studies using the NCEP/NCAR (Kalnay et al., 1996; Wang et al., 2004) or ERA-Interim reanalyses (Uppala et al., 2005; Martin et al., 2019). JRA-55 data are used here due to the long data record and use of 4Dvar data assimilation, but correlations were also checked using the ERA-Interim, NCEP/NCAR and NCEP/DOE-R2 (Kanamitsu et al., 2002) datasets, with similar conclusions obtained overall (not shown).

In this study I aim to explore precursors for interannual variability, but slower decadal trends are also present in the data. When investigating correlations data were therefore detrended with an 11-year rolling mean, which is illustrated for the onset dates by the dashed line in Figure 1. This ensures that the correlations presented relate to interannual variability, rather than to coincident trends in variables due to e.g. global warming or variations in multi-decadal modes. Note that the initial and final 5 years are detrended using the mean of the initial and final 11 years to allow these to be included.

3 Hypothesis Testing

Fig. 2 shows correlations between the hypothesized predictors, averaged from January to March, and the SCSM onset pentad, with both detrended with an 11-year rolling mean. A negative correlation indicates that a positive anomaly is associated with earlier monsoon onset. Looking first over the full reanalysis record, I see correlations that are consistent with the hypothesized relationships. Spring 850-hPa MSE over the South China Sea is negatively correlated with SCSM onset timing (Fig. 2a), while 200-hPa zonal wind over East Asia is weakly positively correlated with monsoon onset (Fig. 2b). MSE anomalies can be expected to relate to anomalies in SST and MSE advection. Over the full record, SCSM onset is correlated with SST over the Philippines to the east (Fig. 2c), but the correlation is weaker than that with MSE. This suggests that the MSE pattern is partially, but not completely, related to local SST anomalies. Breaking down MSE into its contributions from internal, latent, and potential energy, the majority of the correlation was found to come from the latent heat (not shown).

To investigate whether these relationships are consistent throughout the reanalysis record, I also divided the data into two 31-year sections, an early period spanning 1958-1988 (Figs. 2d–f) and a later period spanning 1989-2019 (Figs. 2f–i). A statistically significant negative correlation between Jan-March 850-hPa MSE is present over the South China Sea for both of these time periods. In contrast, for upper-level zonal wind, it becomes clear that the correlation is dominated by the later period, with the earlier period showing no statistically significant relationship over East Asia.

It has been noted that the ENSO-SCSM monsoon relationship appears to have strengthened from the late 1970s onwards (Wang et al., 2009) and that this provides a strong source of predictability for SCSM onset (Martin et al., 2019). The correlations over the later period are consistent with the ENSO teleconnection influence exerting a strong influence on SCSM onset. The MSE correlation pattern shows a clear East-West asymmetry across the Pacific Basin (Fig. 2g), while the zonal wind correlation resembles the upper branch of the Walker cell (Fig. 2h) and the SST correlation shows a clear ENSO pattern (Fig. 2i). The correlation with 200-hPa zonal wind seen in Fig. 2b therefore appears to be a side effect of the relationship with ENSO and the Walker circulation, rather than relating to the Hadley circulation dynamics seen in aquaplanets that were reviewed in Section 2.1. Overall I conclude that upper-level zonal wind does not provide a useful predictor for SCSM onset.

In contrast, in the earlier period there is no clear connection between ENSO and SCSM onset. Instead, the strongest correlation is found to be with MSE over Australia (Fig. 2d). This correlation is not captured by looking purely at SSTs (Fig. 2f), and is predominantly due to the latent heat (not shown). Specifically, in this period, higher MSE over Australia from Jan-March was associated with later onset of the SCSM monsoon. The dipole around the Equator in Fig. 2d suggests that meridional thermal gradients, and their influence on the Hadley circulation, were more important precursors for monsoon onset in this period, compared with the role of zonal thermal gradients and the Walker circulation in the later period. This is supported by Fig. 3, which shows the correlation of onset with the local Hadley and Walker cells, as defined by Schwendike et al. (2014). In the earlier period, Walker circulation anomalies show little correlation with SCSM onset timing, but a weakening of the Northern hemisphere Hadley cell in Jan–March is associated with earlier monsoon onset. In the later period, a strong correlation can be seen between both the Walker and Hadley circulations and SCSM onset, reflecting the influence of ENSO on the circulation.

Fig. 2g shows running correlations with a 31-year mean, indicating how the teleconnections vary in time. The blue and orange lines show MSE averaged over the boxes shown in panels (a) and (d), while the green line shows the SST averaged over the Niño 4 region (panel i). For SCS MSE, a statistically significant negative correlation is found to be present over almost the entire record, although the strength of the correlation does vary in time. The correlation of onset with Australian MSE is in fact stronger than that with SCS MSE, but begins to drop off after 1983, while the relationship with Niño 4 SST strengthens at this point. SCS MSE approximately follows the correlation with both connections, but is consistent across the record. Previous studies (Zhu & Li, 2017) identified different precursors to those presented here; I find that rolling correlations of these also show strong interdecadal variations (Fig. S1).

4 A Simple Forecast

The correlations in Fig. 2 suggest that SCS MSE could provide a useful predictor of SCSM onset. To assess the predictive skill, I test how well data from previous years can be used to predict the onset timing for the next year. I produce predictions from 1993 onwards, which encompasses periods used in previous studies, enabling comparison.

To mimic a plausible operational forecast, an expanding window method is applied to the un-detrended data. The prediction for a given year is estimated by using least squares regression to fit a linear model between the observed onset dates and SCS MSE from all previous years. Each year the window used for generating the forecast expands as new observations are incorporated. As discussed in Section 3, the mean onset date varies in time, and Figs. 2 and 3 indicate that the mechanisms generating the local MSE anomaly vary interdecadally. To account for non-stationary statistics in the forecasting model, a simple option is to apply exponential smoothing (e.g., Brown & Meyer, 1961; Agnew, 1982; Johnstone et al., 1982; Young & Ord, 1985). In order to fit a linear model to generate an onset pentad estimate, $\hat{y}$, from MSE, h, with coefficients β, i.e.,

$$\hat{y} = \beta_1 h + \beta_0, \tag{2}$$

this technique weights the cost-function for the least squares regression using a 'discount factor', λ as:

$$\Sigma_{s=1}^{t} \lambda^{t-s} (y_s - \hat{y}_s)^2. \tag{3}$$

In the above, y_s and $\hat{y}_s$ are the observed and estimated onset pentad in year s and t is the year in which the model is fitted. For $0 < \lambda < 1$, this fit therefore prioritizes minimizing the error for more recent data, while information from historic data is still retained. For $\lambda = 1$ this recovers an unweighted least squares regression.

Predictions for a discount factor of 0.5 are shown by the green circles in Fig. 1. Over the full projected period, I find a Pearson's r of 0.54 (p=0.003), and a root mean square error (RMSE) of 2.05 pentads. The circles and crosses on Fig. 1 indicate where the predicted pentad lies within 1 pentad, or farther than 2 pentads from the observed timing, giving a qualitative picture of the 'direct hit' vs. 'miss' rates. In the 27 years predicted, 7 predictions are more than 2 pentads from what was observed, while 11 are within 1–2 pentads, and 9 are within 1 pentad. For comparison, Fig. S2 shows how the skill score varies with the choice of discount factor, and Fig S3 show the forecasts produced when no discount factor is applied.

The model skill compares well with previous efforts with more complex models. Zhu and Li (2017) applied 3 predictors to model the SCSM onset dates in the NCEP/DOE-R2 dataset, achieving a correlation over their test period, 2005–2014, of 0.72 (RMSE 2.08 pentads). Martin et al. (2019) found that the Met Office GloSea5 ensemble (MacLachlan & Coauthors, 2015; Williams & Coauthors, 2015) could predict SCSM onset with a correlation of 0.5 over a study period from 1993–2015. Over these periods the predictions show correlations with observed onset dates of 0.81 (p=0.004; RMSE 1.35 pentads) and 0.64 (p=0.001; RMSE 1.85 pentads) respectively. Repeating the predictions using only MSE averaged in Jan–Feb or Jan, I find the correlations also remain high when only earlier data are used. For the 2005–2014 and 1993-2015 periods, Jan–Feb averaged MSE gives correlations of 0.77 (RMSE 1.50 pentads) and 0.55 (RMSE 1.95 pentads) respectively, while Jan-mean MSE gives a correlations of 0.66 (RMSE 1.96 pentads) and 0.46 (RMSE 2.19 pentads). Using only a single precursor, this simple model is able to give an initial estimate of SCSM onset timing with roughly 3 to 4 months lead time (given that onset occurs in May on average), and shows correlations competitive with previous models.

5 Discussion

Based on theoretical insights into controls on the meridional overturning circulation that have been developed in aquaplanets, I set out with two hypothesized predictors for SCSM onset: lower-level MSE and upper-level zonal wind in the preceding Jan-March. While the latter does not correlate well with SCSM onset timing, I find that MSE is a useful predictor of interannual variability in SCSM onset. Although the strength of individual teleconnections to the South China Sea varies in time, looking at local MSE allows us to take a step farther along the mechanistic chain from a remote forcing to a local impact on the monsoon onset. Multiple processes can produce local MSE anomalies, but these anomalies are consistently correlated with SCSM onset timing across the JRA-55 record. The correlation strength was also tested in other reanalysis datasets (not shown). In ERA-Interim similar correlation patterns and strengths are seen to those shown in Fig. 2. Correlations in NCEP/NCAR and NCEP/DOE-R2 were also found to follow similar patterns, although these are weaker in magnitude. These findings are thus robust across datasets.

Generating a simple linear regression model based on this single predictor, I produced predictions of onset timing from 1993 onwards via an expanding-window approach. I find predictive skill comparable with more complex models from as early as February. I conclude that local MSE in the months preceding monsoon onset is a useful source of predictability, and would merit further exploration both for use in Physical-Empirical forecast models and for guiding development of dynamical forecasting ensembles.

Despite this favorable comparison in skill, onset in several years was poorly predicted. Subseasonal factors such as intraseasonal oscillations (Zhou & Chan, 2005; R. Wu, 2010) and tropical cyclones (Mao & Wu, 2008; B. Liu & Zhu, 2020) have been found to trigger onset in individual years, limiting seasonal predictability. Although MSE anomalies do appear to precondition the region for early or late onset, I note that it is also highly

important to account for sub-seasonal systems that may cause deviation from the seasonal forecast.

This theory-motivated approach has successfully identified a predictor whose correlation with onset persists across the record, and highlights the direct benefits of better understanding the controls on the large-scale circulation. I hope to build on this success in future by exploring whether similar insights can help understand the spread of projections of the East Asian Summer Monsoon under climate change. However, the second proposed predictor, 200-hPa zonal wind, did not correlate well with monsoon onset. This could relate to a lack of memory in the system for upper-level wind anomalies, or could indicate issues with applying ideas from highly idealized models to Earth. The theoretical foundation for understanding the climatological monsoon still shows some key gaps, in particular the role of zonal asymmetries and transient weather systems in the seasonality of the Hadley cells (Geen et al., 2020). The results presented here suggest that bridging these gaps may provide further opportunities for improved seasonal forecasts.

Acknowledgments

The work was supported by the UK-China Research and Innovation Partnership Fund, through the Met Office Climate Science for Service Partnership (CSSP) China, as part of the Newton Fund. Datasets for this research are available in these in-text data citation references: Japan Meteorological Agency/Japan (2013); Japan Meteorological Agency (Ongoing).

References

Agnew, R. A. (1982). Econometric forecasting via discounted least squares. *Naval Research Logistics Quarterly*, *29*(2), 291–302.

Bordoni, S., & Schneider, T. (2008). Monsoons as eddy-mediated regime transitions of the tropical overturning circulation. *Nature Geoscience*, *1*(8), 515–519. doi: 10.1038/ngeo248

Brown, R. G., & Meyer, R. F. (1961). The Fundamental Theorem of Exponential Smoothing. *Operations Research*, *9*(5), 673–685. doi: 10.1287/opre.9.5.673

Chan, J. C. L., & Zhou, W. (2005). PDO, ENSO and the early summer monsoon rainfall over south China. *Geophysical Research Letters*, *32*(8). doi: 10.1029/2004GL022015

Chang, C.-P. (2004). *East Asian Monsoon* (Vol. 2). World Scientific. doi: 10.1142/5482

Cui, K., & Shoemaker, S. P. (2018). A look at food security in China. *npj Science of Food*, *2*(4). doi: 10.1038/s41538-018-0012-x

Ding, Q., & Wang, B. (2005). Circumglobal teleconnection in the Northern Hemisphere summer. *Journal of Climate*, *18*(17), 3483–3505. doi: 10.1175/JCLI3473.1

Fan, Y., Fan, K., & Tian, B. (2016). Has the Prediction of the South China Sea Summer Monsoon Improved Since the Late 1970s? *Journal of Meteorological Research*, *30*(6), 833–852. doi: 10.1007/s13351-016-6052-8

Fan, Y., Fan, K., Xu, Z., & Li, S. (2018). ENSO-South China Sea Summer Monsoon Interaction Modulated by the Atlantic Multidecadal Oscillation. *Journal of Climate*, *31*(8), 3061–3076. doi: 10.1175/JCLI-D-17-0448.1

Geen, R., Bordoni, S., Battisti, D. S., & Hui, K. (2020). Monsoons, ITCZs and the Concept of the Global Monsoon. *Reviews of Geophysics*, *Accepted*.

Geen, R., Lambert, F. H., & Vallis, G. K. (2018). Regime Change Behavior During Asian Monsoon Onset. *Journal of Climate*, *31*, 3327–3348. Retrieved from `http://journals.ametsoc.org/doi/10.1175/JCLI-D-17-0118.1` doi: 10.1175/JCLI-D-17-0118.1

Geen, R., Lambert, F. H., & Vallis, G. K. (2019). Processes and Timescales in Onset and Withdrawal of Aquaplanet Monsoons '. *J. Atmos. Sci.*.

He, J., & Zhu, Z. (2015). The relation of South China Sea monsoon onset with the subsequent rainfall over the subtropical East Asia. *International Journal of Climatology*, *35*(15), 4547–4556. doi: 10.1002/joc.4305

Hill, S. A. (2019). Theories for Past and Future Monsoon Rainfall Changes. *Current Climate Change Reports*, *5*, 160–171.

Hu, P., Chen, W., Huang, R., & Nath, D. (2018). On the weakening relationship between the South China Sea summer monsoon onset and cross-equatorial flow after the late 1990s. *Int. J. Climatol.*, *38*, 3202–3208.

Huang, R., Gu, L., Zhou, L., & Wu, S. (2006). Impact of the Thermal State of the Tropical Western Pacific on Onset Date and Process of the South China Sea Summer Monsoon. *Adv. Atmos. Sci.*, *23*, 909–924.

Hurley, J. V., & Boos, W. R. (2013). Interannual variability of monsoon precipitation and local subcloud equivalent potential temperature. *Journal of Climate*, *26*(23), 9507–9527. doi: 10.1175/JCLI-D-12-00229.1

Japan Meteorological Agency. (2006). Characteristics of Global Sea Surface Temperature Analysis Data (COBE-SST) for Climate Use. *Monthly Report on Climate System Separated Volume*, *12*, 116.

Japan Meteorological Agency. (Ongoing). *COBE-SST.* Author. Retrieved from `http://ds.data.jma.go.jp/tcc/tcc/products/elnino/cobesst/cobe-sst.html`

Japan Meteorological Agency/Japan. (2013). *JRA-55: Japanese 55-year Reanalysis, Monthly Means and Variances.* Boulder CO: Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. Retrieved from `https://doi.org/10.5065/D60G3H5B`

Johnstone, R. M., Johnson Jr, C. R., Bitmead, R. R., & Anderson, B. D. O. (1982). Exponential convergence of recursive least squares with exponential forgetting factor. *Systems & Control Letters*, *2*(2), 77–82.

Kajikawa, Y., & Wang, B. (2012). Interdecadal Change of the South China Sea Monsoon Onset. *Journal of Climate*, *25*, 3207–3218.

Kajikawa, Y., Yasunari, T., Yoshida, S., & Fujinami, H. (2012). Advanced Asian summer monsoon onset in recent decades. *Geophysical Research Letters*, *39*(3), 1–5. doi: 10.1029/2011GL050540

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., ... Joseph, D. (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, *77*(3), 437-472. doi: 10.1175/1520-0477(1996)077⟨0437:TNYRP⟩2.0.CO;2

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., & Potter, G. L. (2002). NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, *83*(11), 1631-1644. doi: 10.1175/BAMS-83-11-1631

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., ... Takahashi, K. (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *J. Meteorol. Soc. Jpn.*, *93*(1), 5-48. doi: 10.2151/jmsj.2015-001

Lau, K. M., & Yang, S. (1997). Climatology and Interannual Variability of the Southeast Asian Summer Monsoon. *Adv. Atmos. Sci.*, *14*, 141–162.

Lin, A., Zhang, R., & He, C. (2017). The relation of cross-equatorial flow during winter and spring with South China Sea summer monsoon onset. *Int. J. Climatol.*, *37*, 4576–4585.

Liu, B., & Zhu, C. (2020). Boosting Effect of Tropical Cyclone "Fani" on the Onset of the South China Sea Summer Monsoon in 2019. *JGR Atmospheres*, *125*, e2019JD031891.

Liu, P., Qian, Y., & Huang, A. (2009). Impacts of land surface and sea surface

temperatures on the onset date of the South China Sea summer monsoon. *Adv. Atmos. Sci.*, *26*, 493–502.

Ma, D., Sobel, A. H., Kuang, Z., Singh, M. S., & Nie, J. (2019). A Moist Entropy Budget View of the South Asian Summer Monsoon Onset. *Geophysical Research Letters*, *46*(8), 4476-4484. doi: 10.1029/2019GL082089

MacLachlan, C., & Coauthors. (2015). Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Q. J. R. Meteorol. Soc.*, *141*, 1072-1084. doi: 10.1002/qj.2396

Mao, J., & Wu, G. (2008). Influences of Typhoon Chanchu on the 2006 South China Sea summer monsoon onset. *Geophys. Res. Lett.*, *35*, L12809.

Martin, G. M., Chevuturi, A., Comer, R. E., Dunstone, N. J., Scaife, A. A., & Zhang, D. (2019). Predictability of South China Sea Summer Monsoon Onset. *Advances in Atmospheric Sciences*, *36*(3), 253–260. doi: 10.1007/s00376-018-8100-z

Nie, J., Boos, W. R., & Kuang, Z. (2010). Observational evaluation of a convective quasi-equilibrium view of monsoons. *Journal of Climate*, *23*(16), 4416–4428. doi: 10.1175/2010JCLI3505.1

Privé, N. C., & Plumb, R. A. (2007). Monsoon Dynamics with Interactive Forcing. Part I: Axisymmetric Studies. *Journal of the Atmospheric Sciences*, *64*(5), 1417–1430. doi: 10.1175/JAS3916.1

Schneider, T., & Bordoni, S. (2008). Eddy-Mediated Regime Transitions in the Seasonal Cycle of a Hadley Circulation and Implications for Monsoon Dynamics. *Journal of the Atmospheric Sciences*, *65*(1), 915–934. doi: 10.1175/2007JAS2415.1

Schwendike, J., Govekar, P., Reeder, M. J., Wardle, R., Berry, G. J., & Jakob, C. (2014). Local partitioning of the overturning circulation in the tropics and the connection to the Hadley and Walker circulations. *Journal of Geophysical Research*, *119*(3), 1322–1339. doi: 10.1002/2013JD020742

Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V. D. C., Fiorino, M., ... others (2005). The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012.

Wang, B., Huang, F., Wu, Z., Yang, J., Fu, X., & Kikuchi, K. (2009). Multi-scale climate variability of the South China Sea monsoon: A review. *Dynamics of Atmospheres and Oceans*, *47*(1), 15 - 37. doi: https://doi.org/10.1016/j.dynatmoce.2008.09.004

Wang, B., & LinHo. (2002). Rainy Season of the Asian-Pacific Summer Monsoon. *Journal of Climate*, *15*, 386–398. doi: 10.1175/1520-0442(2002)015%3C0386:RSOTAP%3E2.0.CO;2

Wang, B., LinHo, Zhang, Y., & Lu, M. M. (2004). Definition of South China Sea monsoon onset and commencement of the East Asian summer monsoon. *Journal of Climate*, *17*(4), 699–710. doi: 10.1175/2932.1

Williams, K. D., & Coauthors. (2015). The Met Office Global Coupled model 2.0 (GC2) configuration. *Geoscientific Model Development*, *8*(5), 1509–1524. doi: 10.5194/gmd-8-1509-2015

Wu, G., & Zhang, Y. (1998). Tibetan Plateau Forcing and the Timing of the Monsoon Onset over South Asia and the South China Sea. *Monthly Weather Review*, *126*(4), 913–927. doi: 10.1175/1520-0493(1998)126⟨0913:TPFATT⟩2.0.CO;2

Wu, R. (2010). Subseasonal variability during the South China Sea summer monsoon onset. *Climate Dynamics*, *34*, 629–642.

Young, P., & Ord, J. K. (1985). The use of discounted least squares in technological forecasting. *Technological forecasting and social change*, *28*(3), 263–274.

Zhou, W., & Chan, J. C. L. (2005). Intraseasonal oscillations and the South China Sea summer monsoon onset. *Int. J. Climatol.*, *25*, 1585–1609.

Zhou, W., & Chan, J. C. L. (2007). ENSO and the South China Sea summer mon-

443 soon onset. *Int. J. Climatol.*, *27*, 157–167.
444 Zhu, Z., & Li, T. (2017). Empirical prediction of the onset dates of South China Sea
445 summer monsoon. *Climate Dynamics*, *48*(5-6), 1633–1645. doi: 10.1007/s00382
446 -016-3164-x

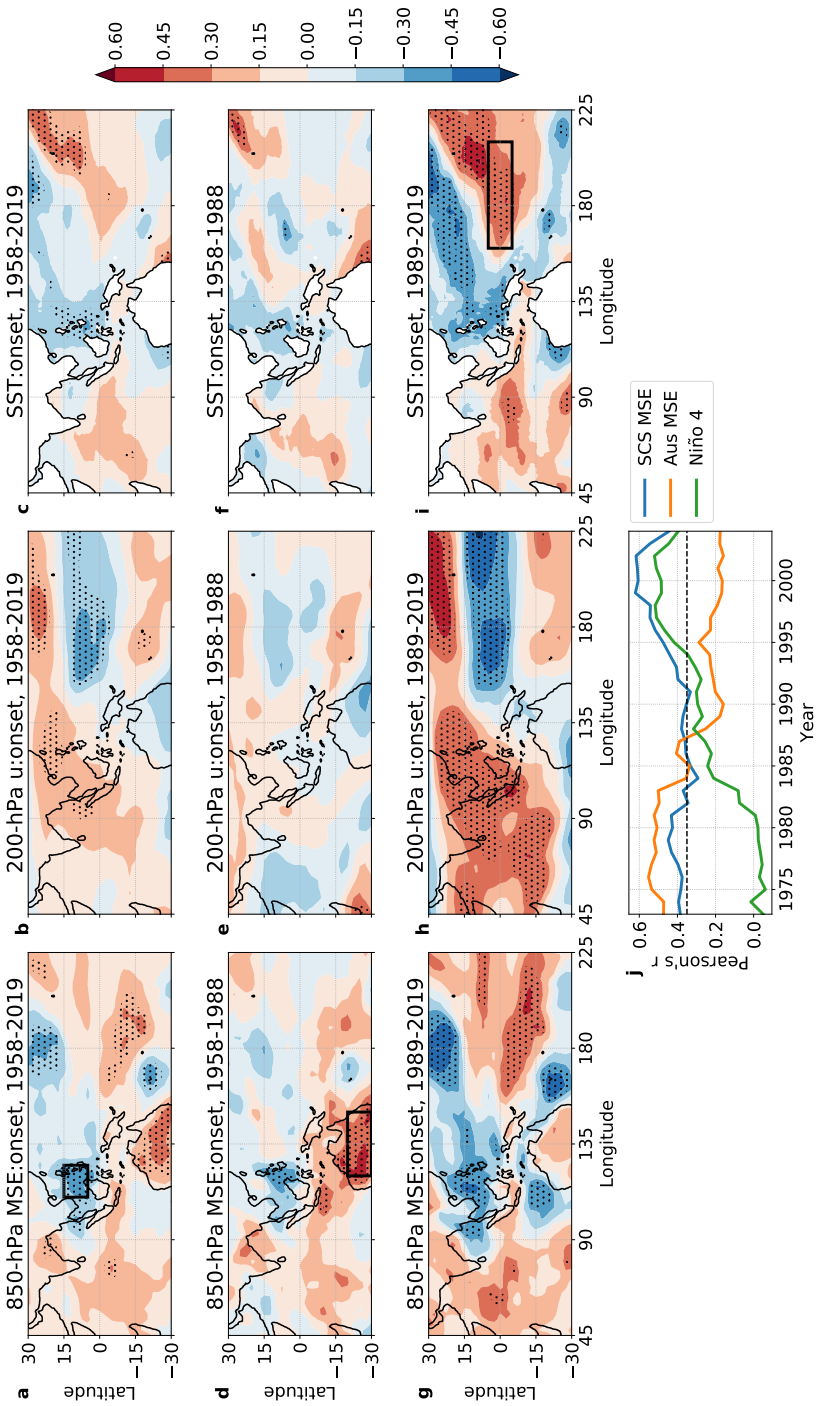

Figure 2. Correlations between potential predictors and the SCSM onset pentad. Panels (a-h) show maps of Pearson correlation coefficient between SCSM onset pentad and Jan-March averaged 850-hPa MSE (left column), 200-hPa zonal wind (centre column) and SST (rightmost column). The top row shows the correlations for the full dataset, 2nd and 3rd rows show correlations for early and late periods respectively (see titles). Bottom centre (j) shows running correlations of SCSM onset pentad with: MSE averaged over 5 – 15 °N, 110 – 125 °E (blue); MSE averaged over -30 – -20 °N, 120 – 150 °E (orange); SST averaged over the Niño 4 region (green). Averaging areas for MSE are indicated by boxes in (a) and (d), and for SST by the box in (i). Stippling on maps and the dashed line in (j) indicate the threshold where the correlation is significantly different from 0 ($p < 0.05$). For all panels data has been detrended relative to an 11-year rolling mean.

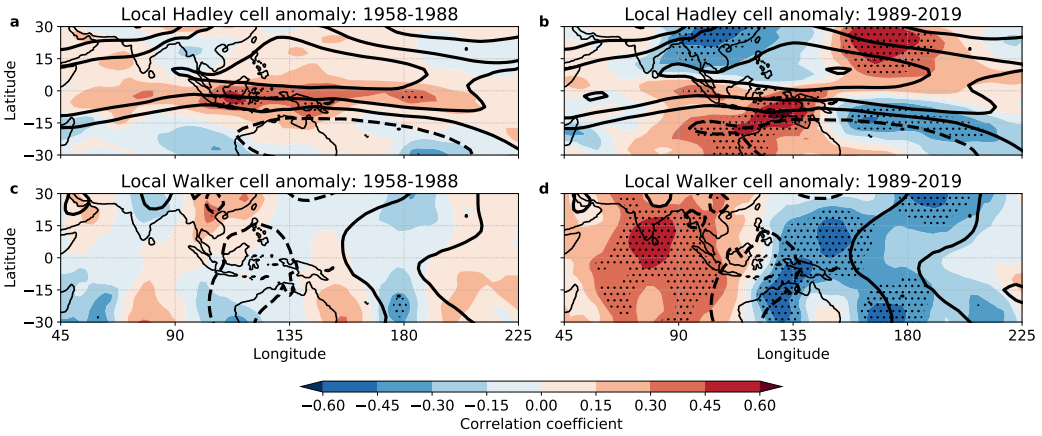

Figure 3. Correlations of Jan-Feb averaged local Hadley (top) and Walker (bottom) cell overturning (Schwendike et al., 2014) with the SCSM onset pentad. Left and right columns show correlations for 1958–1988 and 1989–2019 respectively. Thick dark lines show 500-hPa cross sections of the time-mean Hadley and Walker overturning circulations over the two periods, with contour interval 3000 $\mathrm{kgs^{-1}m^{-1}}$. Stippling indicates where the correlation is significantly different from 0 ($p < 0.05$). For all panels data has been detrended relative to an 11-year rolling mean.