Predictability of South China Sea summer monsoon onset

Article

Accepted Version

1   **Predictability of South China Sea Summer Monsoon onset**

2   Gill M. MARTIN[1*], Amulya CHEVUTURI[2], Ruth E. COMER[1], Nick J. DUNSTONE[1],

3   Adam A. SCAIFE[1,3], and Daquan ZHANG[4]

4   *[1]Met Office Hadley Centre, Met Office, FitzRoy Road, Exeter, EX1 3PB, UK*

5   *[2]NCAS-Climate and Dept. of Meteorology, University of Reading, Reading, RG6 6BB,*

6   *UK*

7   *[3]College of Engineering, Mathematics and Physical Sciences, Exeter University, Exeter,*

8   *EX4 4QJ, UK*

9   *[4]Laboratory for Climate Studies, National Climate Center, China Meteorological*

10   *Administration, Beijing 100081, China*

11   **ABSTRACT**

12   Predicting monsoon onset is crucial for agriculture and socioeconomic planning in

13 countries where millions rely on the timely arrival of monsoon rains for their livelihoods.

14 In this study we demonstrate useful skill in predicting year to year variations in South China

15 Sea summer monsoon onset at up to 3 months lead time using the GloSea5 seasonal

16 forecasting system. The main source of predictability comes from skilful prediction of

17 Pacific sea surface temperatures associated with El Niño and La Niña. The South China

18 Sea summer monsoon onset is a known indicator of the broadscale seasonal transition that

19 represents the first stage of the onset of the Asian summer monsoon as a whole. Subsequent

20 development of rainfall across East Asia is influenced by sub-seasonal variability and

21 synoptic events that reduce predictability, but interannual variability in the broadscale

*Corresponding author : Gill MARTIN
Email: gill.martin@metoffice.gov.uk

22 monsoon onset for East Asian summer monsoon still provides potentially useful

23 information for users about possible delays or early occurrence of the onset of rainfall over

24 East Asia.

25 **Key words:** SCSSM, South China Sea Summer Monsoon, EASM, East Asian Summer

26 Monsoon

27

28 **1. Introduction**

29   The broadscale East Asian Summer Monsoon (EASM) onset occurs in two stages

30 (Wang et al, 2004; 2009): The first stage is a seasonal transition that occurs over the South

31 China Sea (SCS) and is characterised by an abrupt but sustained reversal of the lower

32 tropospheric zonal winds from easterlies to westerlies. Several studies have considered the

33 SCS Summer Monsoon (SCSSM) onset as the precursor for the EASM development (Tao

34 and Chen, 1987; Lau and Yang, 1997), with the formation and progression of the mei-yu

35 rainband forming the second salient phase (Wang et al., 2004). Predicting monsoon onset

36 is crucial for agriculture and socioeconomic planning in countries where millions rely on

37 the timely arrival of monsoon rains for their livelihoods.

38   Interannual variability in the seasonal transition that constitutes the broadscale

39 monsoon onset has been shown to be related to thermal conditions over the Tibetan Plateau

40 (Wu et al., 2012), El Niño-Southern Oscillation (ENSO) effects (Zhou and Chan, 2006; Hu

41 et al., 2014; Xie et al., 2015; Zhu and Li, 2017), regional air-sea interactions (He and Wu,

42 2013) and intraseasonal oscillations (ISO; Li et al., 2013; Wu 2010; Zhu and He, 2013;

43 Shao et al., 2014; Wang et al., 2017). He et al. (2017) carried out a comprehensive analysis

44 of the SCSSM onset in individual years between 1997 and 2014 and showed that the years

45 can be divided into "normal", "intermittent" and "delayed" onset years based on the

46 development of local circulations, thermodynamic conditions and rainfall patterns

47 following the seasonal transition. He et al. (2017) found that eight out of the 18 years they

48 analyzed exhibited intermittent rainfall onset (such that the seasonal dynamical transition

49 is not closely followed by the establishment of monsoon rains and maximum SCS surface

50 temperatures, with a delay caused by an active ISO or northern cold air entering the SCS),

51 and suggested that this reduces the potential predictability of local rainfall onset even if the

52 seasonal dynamical transition may be predictable. Wang et al. (2017) described the effects

53 of the tropical ISO on early, normal and late SCSSM onsets observed over 34 years. They

54 confirmed work from previous studies which showed that, before each onset, the SCS is

55 controlled by the dry phase of the ISO (Shao et al., 2014), and the SCS is warmed to

56 precondition the onset, while after each onset, the SCS is cooled by the wet phase of the

57 ISO (Wu, 2010). However, Wang et al. (2017) showed that the transition process is found

58 to be related to different ISO evolutions over the Indian Ocean for the three types of onsets.

59 Even in non-intermittent onset years, the progression of rainfall onset over East Asia

60 is rarely smooth. After an initial burst of rainfall over the SCS, the rain band rapidly

61 advances northward before stagnating over the Yangtze and Huai River valleys in the mei-

62 yu front (baiu in Japan). The mei-yu rainband exhibits large intra-seasonal and interannual

63 variability and has been the subject of extensive literature (see Ding and Chan (2005) for a

64 review). Its onset is associated with a northward shift of the Northwest Pacific Subtropical

65 High axis to about 25°N and the migration of the upper level westerly jet over Eurasia to

66 the north of the Tibetan Plateau (Sampe and Xie, 2010; Luo 2013). Li et al. (2018) showed

67 that the anticyclone in the upper troposphere over South Asia in April has a significant

68 relationship with the mei-yu onset dates, such that a stronger South Asian anticyclone in

69 April is followed by earlier onset dates of the mei-yu.

70 Despite the complexity associated with these multiple drivers, interannual variability

71 in the seasonal transition that constitutes the broadscale monsoon onset for the Asian

72 summer monsoon as a whole still provides useful information for forecasters about possible

73 delays or early occurrence of the onset of rainfall over East Asia. One of the most-used

74 indices for determining SCSSM onset is that proposed by Wang et al. (2004). This index

75 identifies the first pentad after 25th April in which the zonal wind at 850 hPa over the

76 southern part of the South China Sea (5°-15°N,110°-120°E) shifts from a mean easterly to

77 a mean westerly. Wang et al. (2004) demonstrate that this index is highly indicative of the

78 seasonal transition of the large-scale circulation. They showed that the onset variations

79 determined using this index matched the broadscale onset determined by the principal

80 component of the first empirical orthogonal function (EOF) of the low level winds over

81 East Asia and the Western North Pacific. They argued that this simple index avoids the

82 additional complications of the intraseasonal variability that is included in EOF analysis.

83 An alternative definition for SCSSM onset was proposed by Gao et al. (2001), and is

84 used in operational extended-range forecasting by the Chinese Meteorological

85 Administration (D. Zhang, personal communication). This includes an additional criterion

86 of a sustained increase of equivalent potential temperature at 850 hPa above 340K over the

87 SCS region 10°-20°N, 110°-120°E concurrent with the establishment of westerly winds

88 over the same region. The increase in equivalent potential temperature is considered to

89 indicate sea surface warming, monsoonal transport of moisture into the region and the

90   potential for increased convective activity (Gao et al., 2001; Luo et al., 2013; Li et al.,

91   2013). The region specified by Gao et al. (2001) is further north than that for the Wang et

92   al. (2004) index and includes the northern SCS.

93   In this paper, we investigate the prediction skill of the SCSSM onset on seasonal

94   timescales in the operational hindcast set of the GloSea5-GC2 seasonal forecasting system.

95   Section 2 outlines the data and methods used in our study; section 3 shows the analysis of

96   predictability of the two onset indicators, including tests of the robustness of the seasonal

97   forecast skill. Discussion and conclusions on the usefulness of the seasonal forecast skill

98   of the broadscale monsoon onset using these SCSSM onset indicators are included in

99   Section 4.

100 **2. Data and methods**

101   Daily and pentad timeseries of 850 hPa zonal winds (U_{850}), air temperature (T_{850}) and

102 specific humidity (q_{850}) from the 23-year set of hindcasts (1993-2015) made with the

103 GloSea5-GC2 operational long-range forecast system (MacLachlan et al, 2015; Williams

104 et al., 2015) are taken from four start dates (17th, 25th March, 1st, 9th April). This represents

105 a >1-month lead-time for the average SCSSM onset date of mid-May. The standard

106 operational hindcast set includes 7 members per start date. In order to investigate the

107 robustness of our results, and the dependence on ensemble size, we make use of an

108 additional hindcast ensemble, using the same model configuration and also with 7 members

109 per start date (except for 17th March, for which there are only 3 members). Further, to

110 investigate changes with lead-time, we repeat the analysis for a 56-member ensemble of

111 start dates 25th March, 1st, 9th, 17th April, and for 28-member ensembles generated using

112 the four start dates (1st, 9th, 17th, 25th) of January, February and March respectively.

113   Data representative of observations are taken from ERA-Interim reanalyses (Dee et

114   al., 2011) for the same years. Equivalent potential temperature is calculated from the

115   temperature and humidity fields at 850 hPa using the formula in Bolton (1980). Sea surface

116   temperatures for March (used in section 3.3) are taken from the HadISST1.1 dataset

117   (Rayner et al., 2005).

118   SCSSM onset is determined using the criteria established by Wang et al. (2004) and

119   Gao et al. (2001). According to Wang et al. (2004), the onset date is the first pentad after

120   25th April (i.e. pentad 24 onwards) when the area-averaged U_{850} over the southern SCS

121   (5°-15°N, 110°-120°E, denoted U_{SCS}) is (a) > 0 m s^{-1}; (b) in the subsequent four pentads

122   (including the onset pentad) U_{SCS} must be positive in at least three pentads, and (c) the

123   accumulated 4-pentad mean $U_{SCS} > 1$ m s^{-1}. Wang et al (2004), He and Zhu (2015) and Zhu

124   and Li (2017) have compared the SCSSM onset pentads between different reanalyses

125   (including the National Centers for Environment Prediction (NCEP) reanalyses versions I

126   (Kalnay et al., 1996) and II (Kanamitsu et al., 2002) as well as ERA-Interim) and show

127   reasonable correlations between them (generally >0.8).

128   Gao et al. (2001) suggested an onset criterion based on the area-averaged 850 hPa

129   pentad equivalent potential temperature (θ_e) and U_{850} over the region 10°-20°N, 110°-

130   120°E, with the onset date being the first pentad when $\theta_e > 340K^{\dagger}$ and the $U_{850} > 0.0$ m s^{-1}

131   stably (persists for at least three pentads followed by a break of no more than 2 pentads, or

132   for two pentads followed by a break of no more than one pentad). It should be noted that

133   the region of consideration for this index is slightly further north than that considered by

134   Wang et al. (2004).

---

† Originally specified as 335K by Gao et al. (2001) but revised to 340K by Ding and He (2006).

135 **3. Results**

136 *3.1 Prediction skill of SCSSM onset using the Wang et al. (2004) criterion*

137   Figure 1 shows the SCSSM onset pentads identified using the Wang et al. (2004)

138 criterion for each forecast member with start dates 17th, 25th March, 1st, 9th April in each

139 year, with the ensemble mean pentad and that identified in the reanalyses. The average

140 interannual standard deviation of onset dates from individual ensemble members is 2.2

141 pentads, which compares reasonably well with that of the reanalyses (2.6 pentads), and

142 there is a statistically significant (at the 0.75% level, for a one-tailed t-test) correlation of

143 0.5 between the interannual variations of the ensemble mean dates and those from the

144 reanalyses, indicating significant predictability. The hindcasts also predict the mean onset

145 pentad to match that of the reanalyses, i.e. pentad 28 (16th - 20th May).

146   Luo and Lin (2017) suggest that a more objective measure of the SCSSM onset can be

147 determined using a daily cumulative U_{SCS} and specifying the onset as where this time series

148 changes from decreasing to increasing (indicating that the flow is becoming predominantly

149 westerly). Wang et al. (2004) also checked their SCSSM onset dates against a cumulative

150 U_{SCS} criterion, DU, which compares the accumulated U_{SCS} in the 3 days prior to and after

151 the onset. They showed that although their onset criteria do not explicitly require an abrupt

152 change in westerly speed across the onset pentad, the resultant onset pentads were

153 coincident with such a change. We find that including the additional criterion of DU > 7 m

154 s^{-1} makes very little difference to our results (not shown).

155   We have carried out the same analysis for four start dates (1st, 9th, 17th, 25th) in January,

156 February and March taken from the standard operational hindcast ensemble of 7 members

157 per start date, and also for a 56 member combined ensemble using start dates of 25th March,

158   1st, 9th and 17th April (see Table 1). The correlation coefficient increases with decreasing

159   lead-time, becoming statistically significant at the 1.5% level (for a one-tailed t-test) from

160   February start dates onwards. Thus, there is significant skill in the SCSSM onset prediction

161   using the Wang et al. (2004) index at nearly 3 months lead-time over this hindcast period.

162   *3.2 Predictability of SCSSM onset using the Gao et al. (2001) criterion*

163   Figure 2 shows the SCSSM dates identified using the Gao et al. (2001) criterion in

164   each year by each of the 52 ensemble members with start dates 17th, 25th March, 1st, 9th

165   April, with the ensemble mean pentad and that identified in the reanalyses. In contrast with

166   the findings using the Wang et al. (2004) U_{SCS} index, we find low skill in onset prediction

167   using the Gao et al. (2001) index at >1 month lead time. Table 1 shows that the correlation

168   increases slightly if the lead-time is reduced to ~1 month, but remains barely statistically

169   insignificant at the 6% level (using a one-tailed t-test).

170   The difference in prediction skill between the two methods of determining SCSSM

171   onset may be in part related to the region used for the Gao et al. (2001) index; Wang et al.

172   (2004) commented that *"the northern SCS is open to the invasion of a cold front from the*

173   *north. The westerly flow occurring before the onset is located north of the subtropical ridge*

174   *and is not of tropical origin."* They state, therefore, that the northern part of the SCS should

175   be excluded when defining the tropical monsoon burst over the SCS. He et al. (2017) also

176   commented on the influence of northern cold air entering this region of the SCS

177   contributing to ambiguous or intermittent onset. They highlighted the case of 2009, where

178   the strong westerly flow established in mid-April was interrupted by easterlies propagating

179   from the northern SCS for several days in early May. Other examples of years where this

180   occurred were given in He et al. (2017, their Figures 1, 2) and include 2007, 2009, 2011.

181   Additionally, although He et al (2017) did not identify 2004 as an intermittent onset year,

182   the U_{850} averaged over the Gao et al. (2001) SCS box fluctuates between easterly and

183   westerly during May, making the onset ambiguous when the Gao et al. (2001) index is

184   used. He et al. (2017, their Figure 1) shows that this is related to variability of the winds in

185   the northern part of the SCS. In contrast, the U_{850} winds over the southern part of the SCS

186   (as covered by the Wang et al. (2004) box) do not fluctuate to the same extent. Chan et al.

187   (2000) showed that, in 1998, incursion of cold air into the northern SCS promoted release

188   of convectively available potential energy which helped to trigger the onset earlier than

189   may have been expected given the ENSO conditions. Liu et al. (2002) further linked the

190   cold air incursion to a Rossby wave train triggered over the Bay of Bengal.

191   The additional influence of variability from the subtropics in the northern SCS, which,

192   like the ISO, is unpredictable on seasonal timescales, is likely to be a contributing factor in

193   the reduced seasonal prediction skill for SCSSM onset using the Gao et al. (2001) criteria.

194   In recognition of this, forecasters at CMA release their SCSSM onset forecasts using the

195   Gao et al. (2001) criteria only on the extended range (11-30 day) timescale (D. Zhang,

196   personal communication, 30th March 2018), on which models have been show in previous

197   work to have skill for predicting intraseasonal variability (e.g. Lim et al., 2018; Lee et al.

198   2015).

199   *3.3 Drivers of SCSSM onset predictability using Wang et al. (2004) index*

200   Several studies have shown that ENSO is one of the main drivers of large-scale

201   interannual variability in the Asian monsoon region (e.g. Zhou and Chan, 2007; Luo et al.,

202   2016). Westerly (easterly) equatorial wind anomalies associated with El Niño (La Niña)

203   and a weaker (stronger) Walker circulation are typically associated with negative (positive)

204   sea surface temperature (SST) anomalies over the SCS and a delayed (advanced) seasonal

205   transition (He et al., 2017). This relationship is not symmetrical, however: He et al. (2017)

206   suggest that both intraseasonal oscillations (ISO) and changes in west–east thermal

207   contrasts across the Indian Ocean and western Pacific can influence the timing of onset in

208   La Nina years. Hardiman et al. (2018) found a similar asymmetry in the relationship

209   between seasonal mean Yangtze River rainfall and ENSO in observations and hindcasts.

210   We also show on Figure 1 the observed March Niño3.4 sea surface temperature (SST)

211   anomaly timeseries from HadISST1.1 (yellow line). The correlation coefficient between

212   the ensemble mean SCSSM onset pentad timeseries derived using the Wang et al. (2004)

213   index and the Niño3.4 SST timeseries is 0.9, indicating that the predictable component of

214   the hindcast SCSSM onset is driven mainly by ENSO, which itself is highly predictable on

215   this timescale in GloSea5 (MacLachlan et al., 2015; Scaife et al., 2014). The correlation

216   between observed estimates of SCSSM onset and the observed March Niño3.4 SST is

217   rather lower (0.41), indicating the influence of other drivers of SCSSM onset variability

218   that may not be predictable, particularly the ISO (e.g. Shao et al., 2014; Wang et al., 2017),

219   which is itself subject to inter-annual variations relating to large-scale modes such as the

220   Pacific-Japan teleconnection (Li et al., 2014). The skill of the ensemble (0.5) is therefore

221   marginally higher than using predicted ENSO conditions alone to predict monsoon onset,

222   though both are skilful.

223   Figure 3(a) provides additional insight by showing the correlation between the

224   ensemble mean SCSSM onset dates for the 23 years from the hindcast and observed global

225   monthly mean SSTs in March over the same period. This illustrates that the predictable

226   part of the SCSSM onset from the hindcast is strongly correlated with an ENSO-like pattern

227 of Pacific SSTs, consistent with the findings of Zhu and Li (2017). There is also a strong

228 positive correlation with SSTs in the equatorial Indian Ocean, again indicating that warmer

229 SSTs are associated with later SCSSM onset dates. For the observed onset dates derived

230 from ERA-interim (Fig. 3(b)), the correlations with SST are far smaller, due to the presence

231 of additional factors in the observations that are not predicted by the ensemble mean. The

232 average correlations between the SSTs and 1000 pseudo-timeseries of SCSSM onset

233 created by randomly choosing an individual ensemble member hindcast for each year (Fig.

234 3(c)) are naturally smaller than with the ensemble mean timeseries, but not as low as those

235 in observations (Fig. 3(b)). This suggests that some of the sub-seasonal variations (e.g.

236 intraseasonal oscillations) that affect SCSSM onset in reality may not be sufficiently well

237 represented by the model to capture such influences, even at the relatively high horizontal

238 resolution used by GloSea5 (N216; about 60 km at 50°N). This is consistent with findings

239 of Fang et al. (2016), who showed that while several aspects of the boreal summer ISO

240 were improved in the Met Office Unified Model at this resolution, difficulty remained in

241 realistic representation of the variance and propagation characteristics.

242 *3.4 Robustness of SCSSM wind onset predictability to ensemble size*

243 To assess the influence of ensemble size on the prediction skill using the Wang et al.

244 (2004) index, we randomly sample small ensembles of between 1 and 51 members from

245 the 52 members in our combined ensembles with start dates between 17th March and 9th

246 April, and re-calculate the correlation between the ensemble-mean timeseries and that from

247 the observations for different numbers of ensemble members. Figure 4 indicates that, for

248 this measure of monsoon onset, the prediction skill (black line) rises quickly with ensemble

249 size, reaching a mean value of 0.5 for a 28-member ensemble (which is the size of the

250 standard operational hindcast set), and is robust (correlation coefficients averaged over all

251 ensemble-mean timeseries are statistically significant at the 1% level for a one-tailed test)

252 for around 10 ensemble members or more. This is a reflection of the strong and predictable

253 influence of ENSO on wider tropical rainfall (Kumar et al., 2013; Scaife et al, 2017) and

254 here on the SCSSM onset dates in the hindcast: in most of the summers following strong

255 El Niño/La Niña years (e.g. 1998, 1999, 2000, 2001, 2005, 2008, 2010) the spread among

256 ensemble members is small and several members identify the same onset pentad (see

257 Figure 1), thereby constraining the values selected by random sampling of the ensemble

258 for those years.

259   Several authors (e.g. Scaife et al., 2014; Eade et al., 2014; Dunstone et al., 2016) have

260 demonstrated that the model's North Atlantic Oscillation is less predictable than that

261 observed, so that a large number of ensemble members is required for good prediction skill.

262 This was confirmed by repeatedly randomly selecting a single member to be the truth and

263 using the ensemble mean of the remaining members to predict that member. In contrast,

264 the dashed line on Figure 4 indicates that the model's SCSSM onset dates are more

265 predictable than those from reanalyses, *i.e.* that the model is *over-confident* in its

266 predictions, as is often found for tropical rainfall (Weisheimer and Palmer, 2014). This

267 again illustrates the dominant role of ENSO in providing the predictability in the model,

268 while the observed onset dates are also influenced by intraseasonal variations that are

269 unpredictable on the seasonal timescale.

270 **4. Conclusions**

271   SCSSM onset, as determined by the Wang et al. (2004) U_{850} wind index, is skilfully

272 predicted in GloSea5 at up to 3 months lead time, particularly during active ENSO years.

273 Since the SCSSM onset signifies the start of the broadscale EASM, its skilful prediction is

274 important for forecasters as an indicator of the possible characteristics of the season to

275 come. This complements the skill previously demonstrated for predicting seasonal mean

276 precipitation in the Yangtze River region (Li et al., 2016). The prediction skill for SCSSM

277 onset using this index is robust even with only around 10 ensemble members, consistent

278 with skill in prediction of rainfall in the deep tropics (e.g. Scaife et al., 2017). The skill is

279 largely related to ENSO SSTs which have been shown to be highly predictable in the

280 GloSea5 seasonal forecasting system.

281   In contrast, the Gao et al. (2001) SCSSM onset index, which includes an increase of

282 θ_e in the SCS region as a measure of thermodynamic onset alongside the change to westerly

283 winds, shows little predictability on seasonal timescales. We speculate that this is partly

284 due to the region used by Gao et al. (2001), as this includes the northern SCS which can be

285 influenced by incursions of cold air from the north. This additional influence is, like the

286 ISO, inherently unpredictable on the seasonal timescale, and thus its inclusion through the

287 northward extension of the box used for the Gao et al. (2001) index compared with that of

288 Wang et al. (2004) is, in our view, a contributing factor in the reduced seasonal prediction

289 skill. However, we propose that a seasonal forecast of the broadscale transition using the

290 Wang et al. (2004) index would provide some useful early information for forecasters, and

291 their guidance could later be refined, using other measures such as the Gao et al. (2001)

292 index, with medium-range forecasts that may capture the influence of intraseasonal

293 variations at shorter lead-times.

294   He and Zhu (2015) investigated the correlations between the SCSSM onset (as

295 determined by the Wang et al. (2004) criteria) and the subsequent EASM rainfall from May

296 to September in observations/reanalyses. They suggested that, in contrast with the

297 traditional view that a later onset date would be associated with a lower than normal total

298 seasonal rainfall amount, the region from the lower Yangtze River to Korea and southern

299 Japan shows a positive correlation between the SCSSM onset date and the seasonal mean

300 rainfall, i.e. early SCSSM onset tends to be followed by lower than normal seasonal mean

301 rainfall further north. He and Zhu (2015) associate this relationship with a persistent

302 Western North Pacific anticyclonic/cyclonic anomaly accompanied by decaying El

303 Niño/La Niña conditions in boreal spring to summer (Wu et al., 2010; Stuecker et al., 2013;

304 Hardiman et al, 2017). This suggests that skilful predictions of SCSSM onset could provide

305 an indication of the seasonal mean rainfall in parts of the EASM region.

306   To our knowledge, this is the first time that skill in predicting the broadscale transition

307 associated with the SCSSM onset on seasonal timescales in an operational dynamical

308 forecasting system has been demonstrated. We encourage other centres to investigate this

309 in their operational forecasting systems. While it is recognised that the onset and

310 progression of the SCSSM and EASM systems is complex and may be influenced by other

311 factors such as synoptic events, intraseasonal variability and regional air-sea interactions

312 with little or no predictability on the seasonal timescale, the ability to provide skilful

313 predictions of whether the broadscale seasonal transition is likely to be early, late or normal

314 provides useful, early information for local forecasters, particularly when combined with

315 other predictions, such as the Yangtze River basin rainfall, which have also been shown to

316 be skilful (Li et al., 2016) and are now provided in real time to CMA (Bett et al., 2018).

317

323

324   **REFERENCES**

325   Adler, R. F., G. J. Huffman, A. Chang, R. Ferraro, P.-P. Xie, J. Janowiak, B. Rudolf, U.

326   Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, and E. Nelkin, 2003:

327   The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation

328   analysis   (1979–present).   J.   Hydrometeor.,   4,   1147–1167,

329   https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

330   Bett, P., A. A. Scaife, C. Li, C. Hewitt, N. Golding, P. Zhang, N. Dunstone, D. M. Smith,

331   H. E. Thornton, R. Lu, and H. –L. Ren, 2018: Seasonal Forecasts of the Summer 2016

332   Yangtze River Basin Rainfall. *Adv. Atm. Sci.,* **35**, 918, doi.org/10.1007/s00376-018-

333   7210-y.

334   Bolton, D., 1980: The computation of equivalent potential temperature. Mon. Wea. Rev.,

335   108,   1046–1053,   https://doi.org/10.1175/1520-

336   0493(1980)108<1046:TCOEPT>2.0.CO;2.

337   Chan, J. C. L., Y. G. Wang, and J. J. Xu, 2000: Dynamic and thermodynamic characteristics

338   associated with the onset of the 1998 South China Sea summer monsoon. J. Meteor.

339   Soc. Japan, 78, 367–380

340   Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae,

341   U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van

342   de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J.,

343   Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P.,

344   Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park,

345   B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. 2011: The

346   ERA-Interim reanalysis: configuration and performance of the data assimilation

347   system. *Q.J.R. Meteorol. Soc.*, **137**, 553–597. doi:10.1002/qj.828

348 Ding, Y., and J. C. L. Chan, 2005: The East Asian summer monsoon: an overview. *Meteor.*

349   *Atmos. Phys*., **89**, 117–142, https://doi.org/10.1007/s00703-005-0125-z

350 Ding, Y., and C. He, 2006: The Summer Monsoon Onset over the Tropical Eastern Indian

351   Ocean: The Earliest Onset Process of the Asian Summer Monsoon. *Adv. Atmos. Sci*,

352   **23**, 940–950, doi: 10.1007/s00376-006-0940-2.

353 Dunstone, N., D. Smith, A. Scaife, L. Hermanson, R. Eade, N. Robinson, M. Andrews, and

354   J. Knight, 2016: Skilful predictions of the winter North Atlantic Oscillation one year

355   ahead. *Nature Geosci.* **9**, 809–814, doi:10.1038/ngeo2824.

356 Eade R., D. Smith, A.A. Scaife and E. Wallace, 2014: Do seasonal to decadal climate

357   predictions underestimate the predictability of the real world? *Geophys. Res. Lett.,* **41,**

358   5620-5628, DOI: 10.1002/2014GL061146.

359 Fang, Y., P. Wu, M. S. Mizielinski, M. J. Roberts, T. Wu, B. Li, P. L. Vidale, M. –E.

360   Demory, and R. Schiemann, 2016: High-resolution simulation of the boreal summer

361   intraseasonal oscillation in Met Office Unified model. *Q. J. R. Meteorol. Soc.,* doi:

362   10.1002/qj.2927.

363 Fong, S. K., and A. Y. Wang (Eds.), 2001: *Climatological Atlas for Asian Summer*

364   *Monsoon*. Meteorological and Geophysical Bureau and Macau Foundation, 318 pp.

365 Gao, H., J. He, Y. Tan, and J. Liu, 2001: Definition of 40-year onset date of South China

366   Sea Summer Monsoon. *J. Nanjing Inst. Meteor*., **24**, 379–383.

367 Hardiman, S. C., N. J. Dunstone, A. A. Scaife, P. E. Bett, C. Li, B. Lu, H.-L. Ren, D. M.

368   Smith, and C. C. Stephan, 2018: The asymmetric response of Yangtze river basin

369   summer rainfall to El Niño/La Niña. *Environ. Res. Lett*., **13(2)**, 024015,

370   doi:https://doi.org/10.1088/1748-9326/aaa172

371 He, B., Y. Zhang, T. Li and W.-T. Hu, 2017: Interannual variability in the onset of the

372   South China Sea summer monsoon from 1997 to 2014. *Atmos. and Oc. Sci, Lett.*, **10:1**,

373   73-81, DOI: 10.1080/16742834.2017.1237853

374 He, J. and Z. Zhu, 2015: The relation of South China Sea monsoon onset with the

375   subsequent rainfall over the subtropical East Asia. *Int. J. Climatol.*, **35**: 4547–4556.

376   doi: 10.1002/joc.4305

377 He, Z., and R. Wu, 2013: Seasonality of Interannual Atmosphere–Ocean Interaction in the

378   South China Sea. *J. Oceanography,* **69**, 699–712. doi:10.1007/s10872-013-0201-9

379 Hu, W. T., R. G. Wu, and Y. Liu, 2014: Relation of the South China Sea Precipitation

380   Variability to Tropical Indo-Pacific SST Anomalies during Spring-to-summer

381   Transition. *J. Climate,* **27**, 5451–5467. doi:10.1175/JCLI-D-14-00089.1

382 Kalnay, E., M. Kanamitsu, R. Kister W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha,

383   G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzakl, W. Higgins, J. Janowiak,

384   K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph,

385   1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.,* **77,** 437–

386   471.

387　Kanamitsu, M., W. Ebisuzaki, J. Woollen, S. –K. Yang, J. J. Hnilo, M. Fiorino, and G. L.

388　　　Potter, 2002: NCEP‐DOE AMIP‐II reanalysis (R-2). *Bull. Am. Meteorol. Soc.*, **83**,

389　　　1631-1643.

390　Kumar, A., M. Chen, and W. Wang, 2013: Understanding prediction skill of seasonal mean

391　　　precipitation over the Tropics. *J. Clim*. **26**, 5674–5681.

392　Lau, K. M., and S. Yang, 1997: Climatology and interannual variability of the Southeast

393　　　Asian summer monsoon. *Adv. Atmos. Sci*., **14**, 141–162

394　Lee, S. –S., B. Wang, D. E. Waliser, J. M. Neena, and J.-Y. Lee, 2015: Predictability and

395　　　prediction skill of the boreal summer intraseasonal oscillation in the Intraseasonal

396　　　Variability Hindcast Experiment. *Climate Dyn*., **45**, 2123–2135, doi:10.1007/s00382-

397　　　014-2461-5.

398　Li, C., A. A. Scaife, R. Lu, A. Arribas, A. Brookshaw, R. E. Comer, J. Li, C. MacLachlan,

399　　　and P. Wu, 2016: Skilful seasonal prediction of Yangtze river valley summer rainfall.

400　　　*Environ. Res. Lett.,* **11(9),** 094002, doi:10.1088/1748-9326/11/9/094002.

401　Li, H., He, S., Fan, K. and Wang, H., 2018: Relationship between the onset date of the

402　　　Meiyu and the South Asian anticyclone in April and the related mechanisms. *Clim.*

403　　　*Dyn.* https://doi.org/10.1007/s00382-018-4131-5

404　Li, K. P., W. Yu, T. Li, V. S. N. Murty, S. Khokiattiwong, T. R. Adi, and S. Budi, 2013:

405　　　Structures and mechanisms of the first-branch northward-propagating intraseasonal

406　　　oscillation over the tropical Indian Ocean. *Clim. Dyn.*, **40**, 1707–1720

407　Li, R. C. Y., W. Zhou, and T. Li, 2014: Influences of the Pacific–Japan Teleconnection

408　　　Pattern on Synoptic-Scale Variability in the Western North Pacific. *J. Clim.,* **27**, 140–

409　　　154, doi:10.1175/JCLI-D-13-00183.1

410   Lim, Y., S. Son, and D. Kim, 2018: MJO prediction skill of the Subseasonal-to-Seasonal

411   Prediction models. *J. Clim.*, **31**, 4075–4094, https://doi.org/10.1175/JCLI-D-17-

412   0545.1

413   Liu, Y., J. C. L. Chan, J. Mao, and G. Wu, 2002: The role of Bay of Bengal convection in

414   the onset of the 1998 South China Sea summer monsoon. Mon. Wea. Rev., 130, 2731–

415   2744, doi:https://doi.org/10.1175/1520-0493(2002)130<2731:TROBOB>2.0.CO;2

416   Luo, M. and L. Lin, 2017: Objective determination of the onset and withdrawal of the

417   South China Sea summer monsoon. *Atmos. Sci. Lett.*, **18**, 276–282.

418   doi:10.1002/asl.753

419   Luo, M., Y. Leung, H.F. Graf, M. Herzog, and W. Zhang, 2016: Interannual variability of

420   the onset of the South China Sea summer monsoon. *Int. Jou. Clim.,* **36**, 550–562

421   Luo, Y., H. Wang, R. Zhang, W. Qian, and Z. Luo, 2013: Comparison of rainfall

422   characteristics and convective properties of monsoon precipitation systems over south

423   China and the Yangtze and Huai River basin. *J. Climate*, **26**, 110–132,

424   doi:https://doi.org/10.1175/JCLI-D-12-00100.1.

425   MacLachlan, C., A. Arribas, A. K. Peterson, A. Maidens, D. Fereday, A. A. Scaife, M.

426   Gordon, M. Vellinga, A. Williams, R. E. Comer, J. P. Camp, Xavier, and G. Madec,

427   2015: Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal

428   forecast system. *Q.J.R. Meteorol. Soc.*, **141**, 1072–1084. doi:10.1002/qj.2396

429   Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell,

430   E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice,

431   and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*,

432   **108**, 4407

433 Saha K., 2010: Monsoon over Eastern Asia (Including China, Japan, and Korea) and

434   Adjoining Western Pacific Ocean. *Tropical Circulation Systems and Monsoons*. K.

435   Saha, Ed., Springer, Berlin, Heidelberg, 123-153.

436 Sampe, T., and Xie S. P. 2010: Large-scale dynamics of the Meiyu–Baiu rainband:

437   environmental forcing by the westerly jet. *J. Clim.*, **23**, 113–134

438 Scaife, A. A., A. Arribas, E. Blockley, A. Brookshaw, R. T. Clark, N. Dunstone, R. Eade,

439   D. Fereday, C. K. Folland, M. Gordon, L. Hermanson, J. R. Knight, D. J. Lea, C.

440   MacLachlan, A. Maidens, M. Martin, A. K. Peterson, D. Smith, M. Vellinga, E.

441   Wallace, J. Waters, A. Williams, 2014: Skilful long-range prediction of European and

442   North American winters, *Geophys. Res. Lett.*, **41**, 2514–2519,

443   doi:10.1002/2014GL059637.

444 Scaife, A.A., R. E. Comer, N. J. Dunstone, J. R. Knight, D. M. Smith, C. MacLachlan, N.

445   Martin, A. K. Peterson, D. Rowlands, E. B. Carroll, S. Belcher, and J. Slingo, 2017:

446   Tropical Rainfall, Rossby Waves and Regional Winter Climate Predictions. *Quart. J.*

447   *Roy. Met. Soc.*, **143**, 1-11. doi:10.1002/qj.2910.

448 Shao, X., P. Huang, and R. -H. Huang, 2014: Role of the phase transition of intraseasonal

449   oscillation on the South China Sea summer monsoon onset. *Clim. Dyn.,* **45**, 125–137.

450 Stuecker, M. F., A. Timmermann, F. –F. Jin, S. McGregor, and H. –L. Ren, 2013: A

451   Combination mode of the annual cycle and the El Niño/Southern Oscillation. *Nat*

452   *Geosci,* **6,** 540–544.

453 Tao, S-Y., and L-X. Chen, 1987: Review of recent research on the East Asian summer

454   monsoon in China. *Monsoon Meteorology*, C.-P. Chang and T. N. Krishnamurti, Eds.,

455   Oxford University Press, 60–92.

456 Wang, B., F. Huang, Z. Wu, j. Yang, X. Fu, and K. Kikuchi, 2009: Multi-scale climate

457   variability of the South China Sea monsoon: a review. *Dyn. Atmos. Oceans*., **47,** 15–

458   37

459 Wang, B., Y. Zhang, and M. M. Lu, 2004: Definition of South China Sea monsoon onset

460   and commencement of the East Asia summer monsoon. *J. Clim.,* **17**, 699–710.

461 Wang, H., F. Liu, B. Wang, and T. Li, 2017: Effects of intraseasonal oscillation on South

462   China Sea summer monsoon onset. *Clim. Dyn.,* doi:10.1007/s00382-017-4027-9.

463 Weisheimer, A., and T. N. Palmer, 2014: On the reliability of seasonal climate forecasts.

464   *J. R. Soc. Interface,* **11**, 20131162, doi: 10.1098/rsif.2013.1162.

465 Williams, K. D., C. M. Harris, A. Bodas-Salcedo, J. Camp, R. E. Comer, D. Copsey, D.

466   Fereday, T. Graham, R. Hill, T. Hinton, P. Hyder, S. Ineson, G. Masato, S. F. Milton,

467   M. J. Roberts, D. P. Rowell, C. Sanchez, A. Shelly, B. Sinha, D. N. Walters, A. West,

468   T. Woollings, and P. K. Xavier, 2015: The Met Office Global Coupled model 2.0 (GC2)

469   configuration, *Geosci. Model Dev.*, **8**, 1509-1524, https://doi.org/10.5194/gmd-8-

470   1509-2015.

471 Wu, R., 2010: Subseasonal variability during the South China Sea monsoon onset. *Clim.*

472   *Dyn.*, **34**, 629–642, doi:10.1007/s00382-009-0679-4.

473 Xie, S. P., Y. Kosaka, Y. Du, and G. Huang, 2015: Indo-western Pacific ocean capacitor

474   and coherent climate anomalies in post-ENSO summer: a review. *Adv. Atmos. Sci.*,

475   doi:10.1007/s00376-015-5192-6

476 Zhou, W., and J. C. L. Chan, 2007: ENSO and the South China Sea Summer Monsoon

477   Onset. *Int. Jou. Climatol.,* **27**, 157–167.10.1002/(ISSN)1097-0088

478   Zhu, Z., and J. He, 2013: The vortex over Bay of Bengal and its relationship with the

479   outbreak of South China Sea summer monsoon. *J. Trop. Meteor.*, **71 (3)**, 915-923

480   Zhu, Z., and T. Li, 2017: Empirical prediction of the onset dates of South China Sea

481   summer monsoon. *Clim. Dyn.*, **48 (5)**, 1633-1645, doi:10.1007%2Fs00382-016-3164-

482   x

483

484   **FIGURE CAPTIONS**

485   **Figure 1:** Predictability of the SCSSM wind onset: onset pentads derived using the method

486   proposed by Wang et al. (2004) from the GloSea5 ensemble predictions initialized on

487   17th, 25th March, 1st, 9th April (green dots represent individual members of the 52-

488   member ensemble, with the size of the dot scaled by the number of members predicting

489   the same onset pentad) and their ensemble mean (green line) compared with the

490   equivalent onset pentads derived from ERA-Interim (black line). The yellow line

491   shows the Niño3.4 SST anomaly in March for each year taken from the HadISST1.1

492   dataset. Pearson correlation coefficients are given in the legend: r(ens,obs) represents

493   the correlation between the GloSea5 ensemble mean and ERA-Interim; r(ens,sst)

494   represents the correlation between the GloSea5 ensemble mean SCS onset pentads and

495   the observed March Niño3.4 SST anomaly; r(obs,sst) represents the correlation

496   between the ERA-Interim SCS onset pentads and the observed March Niño3.4 SST

497   anomaly.

498   **Figure 2:** As Fig. 1 but for SCSSM thermodynamic onset as determined by a sustained

499   increased of $\theta_{e,SCS}$ above 340K accompanied by the establishment of westerly winds

500 over the region 10-20N 110-120E, as proposed by Gao et al. (2001) (with the threshold

501   modified by Ding and He, 2006).

502 **Figure 3:** Correlation coefficients between SCSSM onset pentad derived using the Wang

503 et al. (2004) index and observed March average sea surface temperatures from

504 HadISST1.1 for the period 1993-2015, using: (a) ensemble mean onset dates from the

505 hindcast; (b) onset dates from ERA-Interim, (c) 10,000 pseudo-timeseries of onset

506 dates created by randomly selecting an individual ensemble member from each year;

507 panel shows average over all correlations. Contours and darker shades indicate

508 correlations significant at the 1% (r=0.48) and 3% (r=0.40) levels respectively, for a

509 one-tailed t-test.

510 **Figure 4:** Effect of ensemble size on the skill of SCSSM onset predictions using the Wang

511 et al. (2004) index (solid line), denoted r(ens,obs), and the signal to noise ratio

512 (correlation of ensemble mean timeseries with a pseudo-timeseries created by

513 randomly selecting a single model ensemble member for each year, dashed line),

514 denoted r(ens,mod). In both cases, for each choice of ensemble size, up to 10,000

515 ensemble-mean timeseries are generated by randomly selecting the chosen number of

516 ensemble member onset dates (independently and without replacement) from the 52

517 onset dates diagnosed in each year in the combined ensemble and averaging over the

518 chosen number of ensemble members. Dot-dashed lines indicate the values of r that

519 are significant at the 1% and 0.1% levels for a one-tailed t-test.

520

521 **Table 1.** Pearson correlation coefficients between ensemble mean SCSSM onset dates

522 from GloSea5 and those from ERA-Interim, using the definitions of Wang et al. (2004)

523 and Gao et al. (2001), for different hindcast start dates. Note that the earliest observed

524 SCSSM onset date is pentad 25 (1st -5th May) and the mean onset date is pentad 28 (16th-

525 20th May). Where just the month is shown, start dates are 1st, 9th, 17th, and 25th of the month.

526 Correlation coefficients statistically significant (for a 23 year hindcast period) at <1.5%

527 level for a 1-tailed test are in *italics* and those significant at <1% level for a 1-tailed test

528 are in **bold.**

|				Ensemble start dates	
	January	February	March	17th, 25th March, 1st, 9th April	25th March, 1st, 9th, 17th April
Wang et al. (2004)	0.28	*0.46*	*0.45*	**0.50**	**0.53**
Gao et al. (2001)	-	-	-	0.27	0.30

529

530
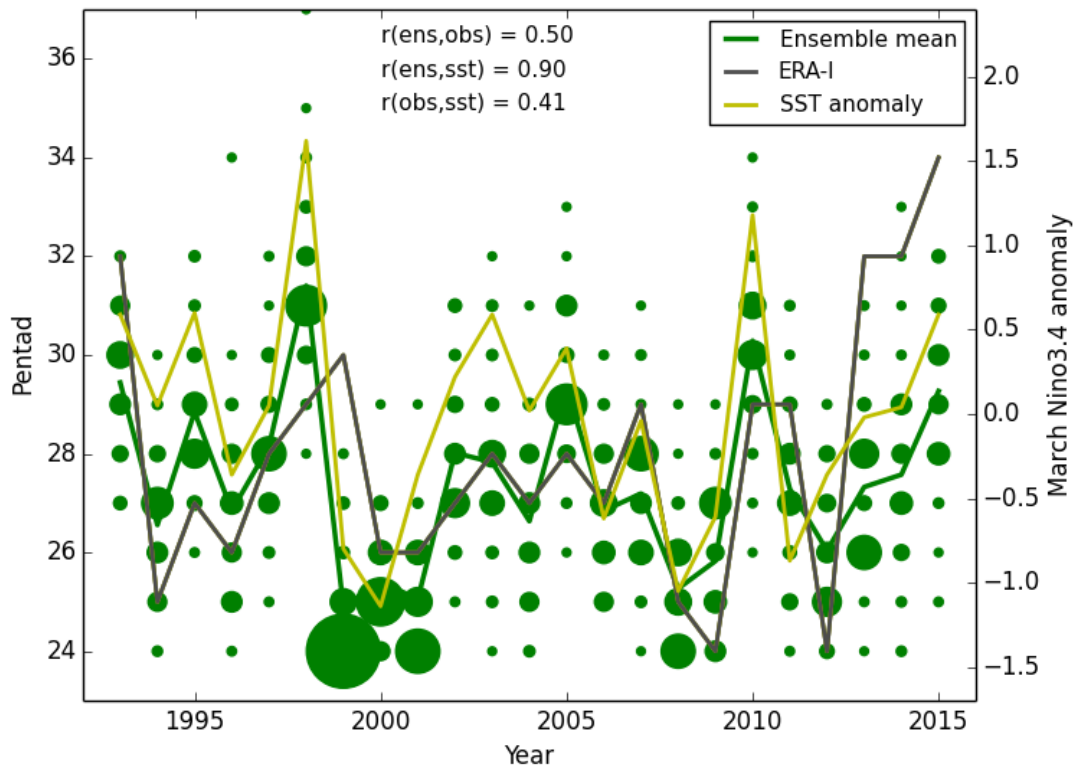
Fig. 1: Predictability of the SCSSM wind onset: onset pentads derived using the method proposed by Wang et al. (2004) from the GloSea5 ensemble predictions initialized on 17th, 25th March, 1st, 9th April (green dots represent individual members of the 52-member ensemble, with the size of the dot scaled by the number of members predicting the same onset pentad) and their ensemble mean (green line) compared with the equivalent onset pentads derived from ERA-Interim (black line). The yellow line shows the Niño3.4 SST anomaly in March for each year taken from the HadISST1.1 dataset. Pearson correlation coefficients are given in the legend: r(ens,obs) represents the correlation between the GloSea5 ensemble mean and ERA-Interim; r(ens,sst) represents the correlation between the GloSea5 ensemble mean SCS onset pentads and the observed March Niño3.4 SST

541   anomaly; r(obs,sst) represents the correlation between the ERA-Interim SCS onset pentads

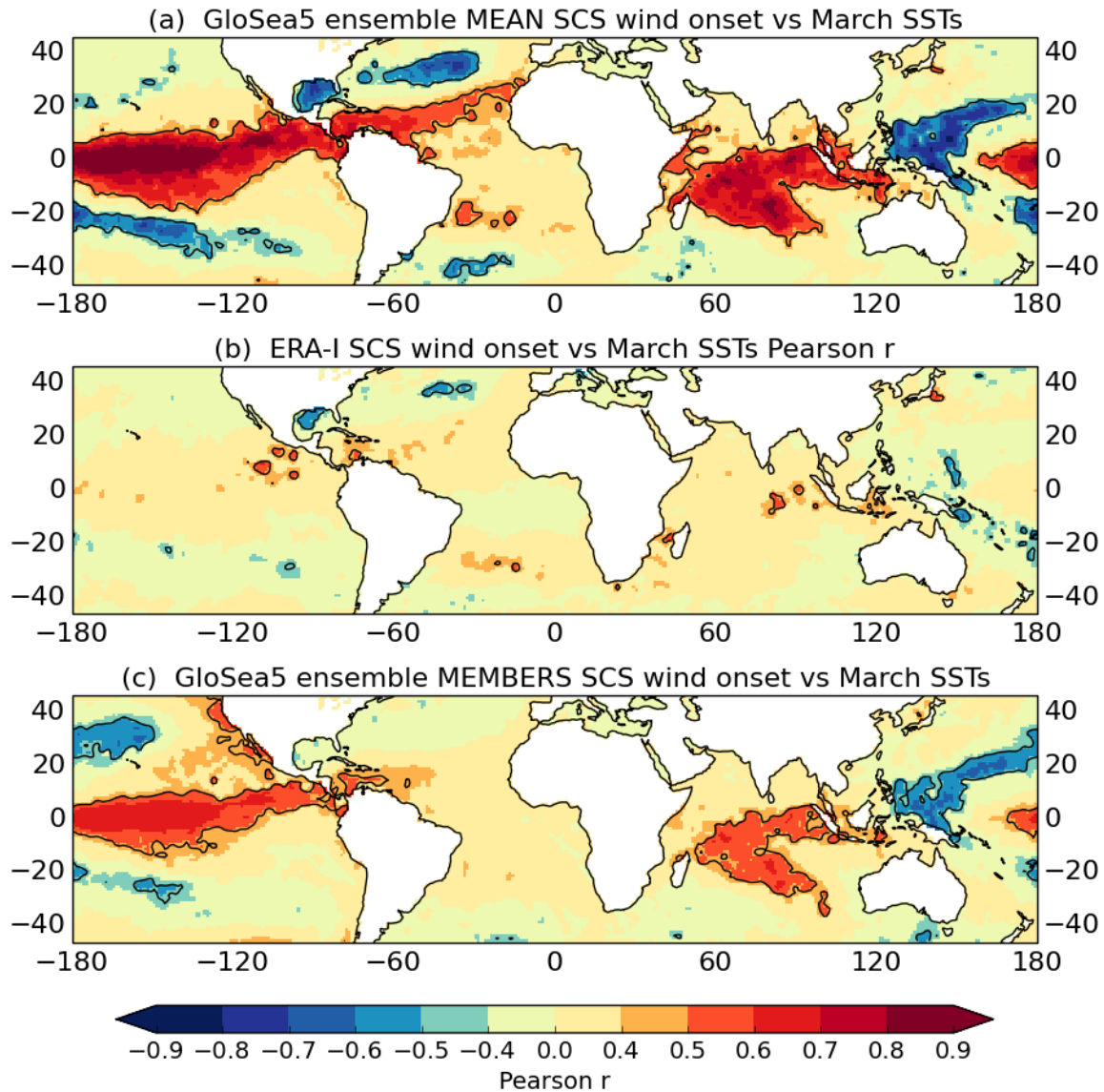542   and the observed March Niño3.4 SST anomaly.

543

544
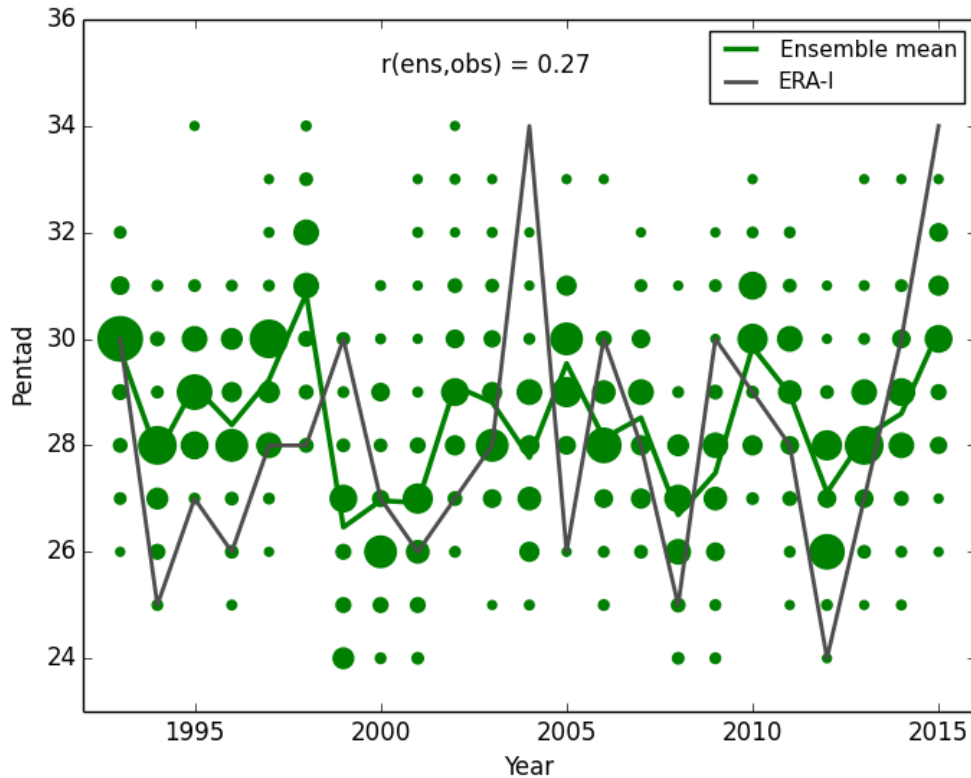
545

546   **Fig. 2:** As Fig. 1 but for SCSSM thermodynamic onset as determined by a sustained

547   increased of $\theta_{e,SCS}$ above 340K accompanied by the establishment of westerly winds over

548   the region 10-20N 110-120E, as proposed by Gao et al. (2001) (with the threshold modified

549   by Ding and He, 2006).

550

Fig. 3: Correlation coefficients between SCSSM onset pentad derived using the Wang et al. (2004) index and observed March average sea surface temperatures from HadISST1.1 for the period 1993-2015, using: (a) ensemble mean onset dates from the hindcast; (b) onset dates from ERA-Interim, (c) 10,000 pseudo-timeseries of onset dates created by randomly selecting an individual ensemble member from each year; panel shows average over all correlations. Contours and darker shades indicate correlations significant at the 1% (r=0.48) and 3% (r=0.40) levels respectively, for a one-tailed t-test.
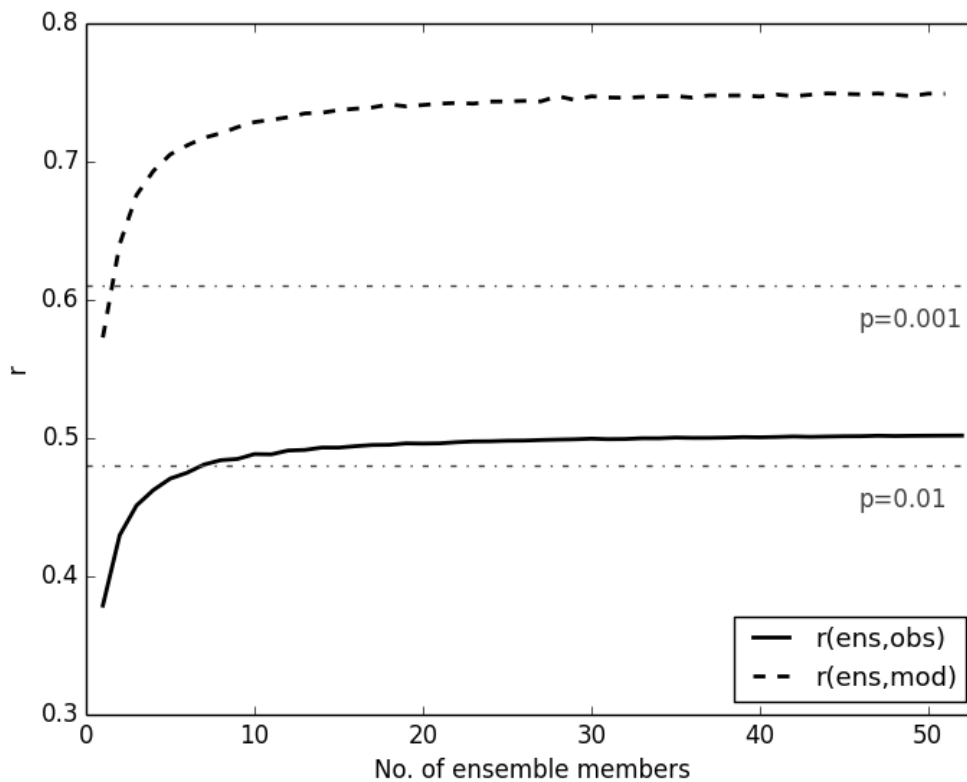
Fig. 4: Effect of ensemble size on the skill of SCSSM onset predictions using the Wang

et al. (2004) index (solid line), denoted r(ens,obs), and the signal to noise ratio

(correlation of ensemble mean timeseries with a pseudo-timeseries created by randomly

selecting a single model ensemble member for each year, dashed line), denoted

r(ens,mod). In both cases, for each choice of ensemble size, up to 10,000 ensemble-mean

timeseries are generated by randomly selecting the chosen number of ensemble member

onset dates (independently and without replacement) from the 52 onset dates diagnosed

in each year in the combined ensemble and averaging over the chosen number of

ensemble members. Dot-dashed lines indicate the values of r that are significant at the

1% and 0.1% levels for a one-tailed t-test.