

1 **Impacts of anemometer changes, site relocations and processing**

2 **methods on wind speed trends in China**

3 Yi Liu¹, Lihong Zhou¹, Yingzuo Qin¹, Cesar Azorin-Molina², Cheng Shen³, Rongrong
4 Xu^{1*}, Zhenzhong Zeng^{1*}

5 ¹ School of Environmental Science and Engineering, Southern University of Science and Technology,
6 Shenzhen, China

7 ² Centro de Investigaciones sobre Desertificación, Consejo Superior de Investigaciones Científicas
8 (CIDE, CSIC-UV-Generalitat Valenciana), Climate, Atmosphere and Ocean Laboratory (Climatoc-Lab),
9 Moncada, Valencia, Spain

10 ³ Regional Climate Group, Department of Earth Sciences, University of Gothenburg, Gothenburg,
11 Sweden

12
13 * Correspondence: xurr@sustech.edu.cn (R. X); zengzz@sustech.edu.cn (Z. Zeng)

14 Manuscript for *Atmospheric Measurement Techniques*

15 September 3, 2023

16

17 **Abstract**

18 *In-situ* surface wind observation is a critical meteorological data source for various
19 research fields. However, data quality is affected by factors such as surface friction
20 changes, station relocations, and anemometer updates. Previous methods to address
21 discontinuities have been insufficient, and processing methods have not always adhered
22 to World Meteorological Organization (WMO) World Climate Programme guidelines.
23 We analyzed data discontinuity caused by anemometer changes and station relocations
24 in China's daily *in-situ* near-surface (~ 10m) wind speed observations and the impact
25 of the processing methods on wind speed trends. By comparing the wind speed
26 discontinuities with the recorded location changes, we identified 90 stations that
27 showed abnormally increasing wind speeds due to relocation. After removing those
28 stations, we followed a standard quality control method recommended by the World
29 Meteorological Organization to improve the data reliability and applied Thiessen
30 Polygons to calculate the area-weighted average wind speed. The result shows that
31 China's recent reversal of wind speed was reduced by 41% after removing the
32 problematic stations, with an increasing trend of $0.017 \text{ m s}^{-1} \text{ year}^{-1}$ ($R^2 = 0.64, P < 0.05$),
33 emphasizing the importance of robust quality control and homogenization protocols in
34 wind trend assessments.

35 **Keywords.** wind speed trends; anemometer changes; station relocations; processing
36 methods; quality control; data homogenization

37

38 **1. Introduction**

39 *In-situ* surface wind observation is a key meteorological data that has been used in
40 various avenues of research, e.g., wind power evaluation (Tian et al., 2019; Zeng et al.,
41 2019; Liu et al. 2022a), extreme wind hazard monitoring and prevention (Zhou et al.,
42 2002; Tamura, 2009; Liu et al., 2022b), and evapotranspiration analysis (Rayner, 2007;
43 McVicar et al., 2012), to name but a few. The application of robust quality control and
44 homogenization protocols are crucial for generating reliable wind speed time series for
45 further trend and variability analyses (Azorin-Molina et al., 2014; Azorin-Molina et al.,
46 2019).

47 Wind data quality is affected by surrounding surface friction change, station
48 location issues, and anemometer changes in type and height (Masters et al., 2010; Wan
49 et al., 2010; Cao & Yan, 2012; Hong et al., 2014; He et al., 2014; Azorin-Molina et al.,
50 2018; Camuffo et al., 2020). Surrounding surface friction changes are mainly associated
51 with urbanization (Zhang et al., 2022) and vegetation growth (Vautard et al., 2010),
52 which modify wind speed fields around the stations. Because of these issues, stations
53 are relocated to satisfy observing criteria (Trewin, 2010). Station relocation is quite
54 common in rapidly developing countries. For instance, about 60% of stations in China
55 experienced relocation (Sohu, 2004). Some relocation-caused breakpoints have been
56 corrected by parallel observations (i.e. operating observations for an overlapping period
57 at both the old and new observing stations; CMA, 2011; CMA, 2012; WMO, 2020), but
58 not all (Feng et al., 2004; Fu et al., 2011; Patzert et al., 2016; Tian et al., 2019; Yang et
59 al., 2021). Besides relocation caused by rapid urbanization (or vegetation growth),
60 updates to automatic anemographs at the beginning of the 21st century in China also
61 caused discontinuities in wind series (Fu et al., 2011).

62 Scientists have tried different methods to handle discontinuities. Tian et al. (2019)
63 and Yang et al. (2021) deleted stations with recorded changes in latitudes, longitudes or
64 altitudes, but they omitted to check whether those recorded relocations caused an abrupt
65 discontinuity in the time series or if parallel observations have corrected them. This
66 results in some stations being mistakenly deleted and significantly reduced the number

67 of available stations. Other research used statistical methods to detect or correct the
68 time series' abnormal breakpoint (Feng et al., 2004; Wang, 2008). However, without
69 examining the causes behind the discontinuity, this may also mistakenly delete stations
70 with natural abrupt climatic changes (Bathiany et al., 2003). Combining those two
71 methods by matching discontinuity with recorded station relocation is needed. Li et al.
72 (2018) have manually checked the station histories for nine stations in North West
73 China, but an algorithm is required to apply this approach to large datasets.

74 Besides data discontinuities, the processing method also affects the wind series.
75 There are two critical steps in the processing: 1) selecting qualified stations and 2)
76 calculating the average value. As for the first step, World Meteorological Organization
77 (WMO) World Climate Programme suggests deleting stations with either too much
78 missing data or continuous missing data (WMO, 2003; WMO, 2017). Previous studies
79 only constrained the number of missing values monthly (Zeng et al., 2019), yearly (Tian
80 et al., 2019) or even in the whole period (Yang et al., 2021) but didn't check whether
81 the missing values were continuous. As for the second step, most studies used the
82 station average as the mean wind speed (Li et al., 2017; Zeng et al., 2019; Tian et al.,
83 2019; Yang et al., 2021; Shen et al., 2021; Zha et al., 2021). However, given station
84 distribution and wind speed spatial variation are often inhomogeneous with larger wind
85 but fewer station in Northwest while smaller wind but more stations in Southeast (Feng
86 et al., 2004; Fu et al., 2011; Liu et al., 2019), the wind variation in Northwest is
87 underrepresented because of few stations. An improved average method (area weighted
88 average) to rearrange the weight for each station based on the area it represents is
89 needed and Thiessen Polygon (Thiessen, 1991) is widely used in which the area is only
90 determined by the station locations while other method like grids is sensitive to the
91 grids chosen.

92 Herein, taking stations in China as an example, we analyzed the existing data
93 discontinuities and their potential causes. Furthermore, we propose an improved
94 solution by using an algorithm to compare the statistic breakpoint with the recorded
95 relocation to double-check the discontinuity caused by relocation. Then using WMO's

96 quality control criteria and Thiessen Polygon (Thiessen, 1911), we generated wind
97 speed time series without temporal bias caused by heterogeneous missing values and
98 spatial biases caused by uneven station distribution.

99

100 **2. Dataset and methodology**

101 **2.1 WMO quality control method**

102 We used the China Surface Climatic Data Daily Data Set (CSD) (Version 3.0) from
103 the China Meteorological Data Service Center (<http://data.cma.cn/en/?r=data/>; last
104 accessed March 2020). The quality control method is recommended by WMO (2017),
105 which required the following criteria before using the daily mean values in a month as
106 monthly mean values: (1) <11 missing daily values in a month; (2) <5 consecutive
107 missing daily values in a month; (3) Complete monthly values for every month during
108 the study period. The station excluded by each criterion can be found in Table S1.

109

110 **2.2 Station location changes in record**

111 CSD provides daily wind speed and location information for 840 stations for 1961-
112 2019. But there are some mistakes in the daily location records. For example, if the
113 station location changed from A to B and back to A within a month, B is potentially a
114 mistaken record. Therefore, we first use mode (the statistic term meaning the value that
115 appears most often, here referring to the location with the highest frequency in a month)
116 to resample the daily location to the monthly location. Second, considering that
117 recorded longitude and latitude has the same spatial resolution of minutes, we defined
118 the threshold of location change as the minimum accuracy of the longitude and latitude
119 record, i.e., one minute. That is 1.85 km for longitude and $1.85 \text{ km} \times \cos\varphi$ for latitude,
120 where φ is the latitude. Third, as for altitude, we allow a 20m measuring error following
121 Tian et al. (2019). A station with more than 20m change in altitude will be considered
122 as relocation. It is noteworthy that CSD labels uncertain altitude records by adding 10
123 km to the raw data (CMA, 2017), which are considered as no observations in our
124 analysis. This way, we identified 432 stations as relocations from the 601 qualified

125 stations after applying the WMO quality control (details in Table S2).

126

127 **2.3 Breakpoint detection and the comparison with recorded relocation**

128 We used Pruned Exact Linear Time (PELT) method (Killick, Fearnhead & Eckley,
129 2012) to detect the jumps in the mean level in the monthly wind speed time series (Fig
130 4a, Fig 4c). This method is a wrapped function named *findchangepts* in Matlab. PELT
131 is essentially a traversing method. For a time series with N values ($x_1, x_2 \dots x_N$), the
132 function uses equations 1 & 2 to calculate the total residual errors (J) for each point (k)
133 assumed as a breakpoint. The point with the most significant change in the mean (lowest
134 total residual errors, J) is reported as the breakpoint. The breakpoints here can be caused
135 by artificial relocations or natural climate changes.

$$136 \quad J(k) = \sum_{i=1}^{k-1} (x_i - \text{mean}([x_1 \dots x_{k-1}]))^2 + \sum_{i=k}^N (x_i - \text{mean}([x_k \dots x_N]))^2 \quad (1)$$

$$137 \quad \text{mean}([x_m \dots x_n]) = \frac{1}{n-m+1} \sum_{r=m}^n x_r \quad (2)$$

138 Then we use relocation records to separate changes brought by artificial relocation
139 from changes in natural climate. If the breakpoint and one of the relocation dates (some
140 stations have more than one relocation record) happened in the same two months, we
141 will consider that the time series is significantly affected by the relocation, and the
142 station will be deleted. Stations with natural-climate-caused location changes will be
143 reserved.

144 The change point in the trend of the annual national average wind speed (Fig 2b,
145 Fig 5b) is detected following the method used by Wang et al. (2011). All the trends
146 reported are based on the least square fits.

147

148 **3. Results and discussion**

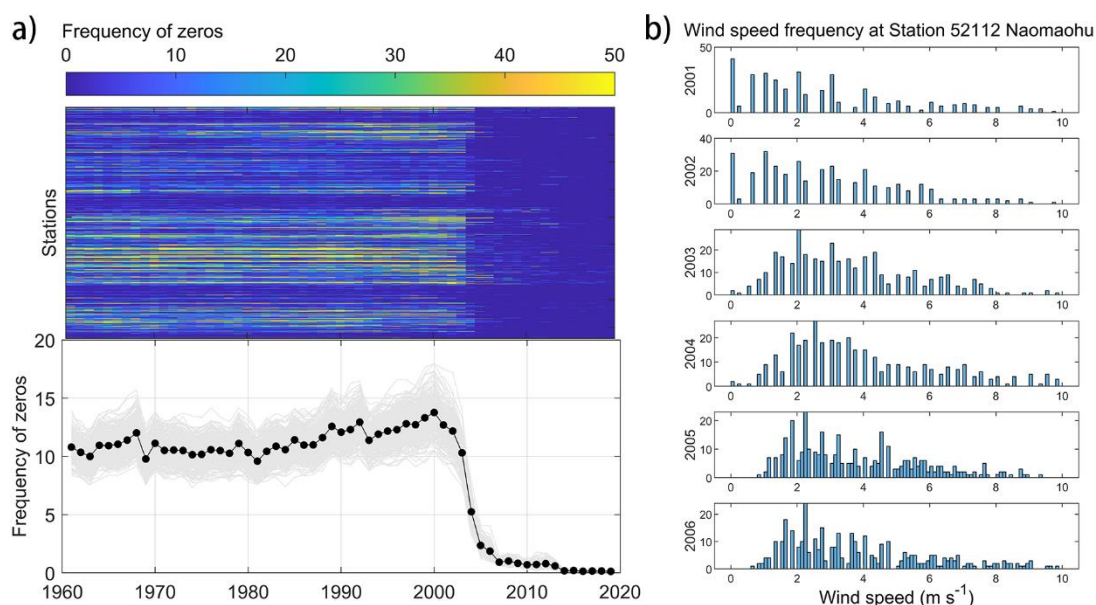
149 **3.1 Data issue due to anemometer changes**

150 We found a clear decline in the frequency of zeros (zero wind speed) in most CSD
151 stations between 2002 and 2007 (Figure 1a), from 10-14 days per year to less than two
152 days per year. This clear drop is not a result of wrongly taking zero values as no

153 observations (NaN) as happened in the Integrated Surface Dataset (ISD, Dunn et al.,
154 2022), as no abrupt increase in NaN frequency was observed (Supplementary Figure
155 S1). Instead, the decline is accompanied by a decrease in measure uncertainty: i.e., the
156 measurement intervals became narrow (from 0, 0.3, 0.5, 0.7, 0.8, 1.0 m s⁻¹, etc. to 0,
157 0.1, 0.2, 0.3 m s⁻¹, etc.; Figure 1b and Supplementary Figure S2). Taking Station
158 Naomaohu in Xinjiang (station ID: 57432) as an example, from 2002 to 2003, zero
159 values decreased from more than 30 days per year to less than five days per year and
160 wind speed records changed from 0, 0.3, 0.7, 1.0 m s⁻¹, etc. to 0, 0.3, 0.5, 0.8, 1.0 m s⁻¹,
161 etc. Since 2004, the measurement was further improved to 0.1, 0.2, 0.3... m s⁻¹ and
162 zeros values almost disappeared (Figure 1b).

163 This change is potentially caused by the transformation in measure frequency,
164 anemometer type and data logging, based on the station history recorded by Xin et al.
165 (2012). As for measurement frequency, in 2003, Station Naomaohu changed from 3
166 observations per day (i.e., 8:00, 14:00 and 20:00, China Standard Time) to four times
167 per day (2:00, 8:00, 14:00 and 20:00, China Standard Time). The increase in the
168 frequency of measurements decreases zeros in daily wind data, as only if all
169 observations report zero wind speeds, will the daily data (i.e., the average of all
170 observations in a day; CMA, 2017) be recorded as zero. Then in 2005, the EL contact
171 anemograph (Yang, 1986; Jin, 2011; Xin et al., 2012, Zhang et al., 2020) requiring
172 manual recording was changed to EC photoelectric encoder self-recording type (Kuang,
173 2016; Jin, 2011; Xin et al., 2012). Both EL and EC type anemographs use cup
174 anemometers to measure wind speed. This anemograph change further decreases the
175 likelihood of recording zero daily wind speed because the updated new anemometers
176 are more sensitive, and even very low wind speeds will be measured with a value
177 instead of recorded as zero (Azorin-Molina et al., 2018). The smooth increasing
178 frequency of zero values from 1960 until 2000 also supports this statement (Figure 1a):
179 the longer the anemometer is used, the less sensitive it will become, and hence a greater
180 wind speed will be required to record a non-zero value (Azorin-Molina et al., 2018),
181 overall increasing the zero values. As for the change in data accuracy, there are two

182 reasons: 1) EL type anemograph only measures the times of electronic contact (200
 183 meters rotation distance per contact) in 10 mins, therefore it has discrete records. For
 184 example, one contact means 0.3 m s^{-1} (200m/600s) and two contacts means 0.7 m s^{-1}
 185 ($400\text{m}/600\text{s}$) (Hu et al., 2009) while EC type has more accurate records using the Grey
 186 Code; 2) the data logging changed from manual reading, calculating and rounding to
 187 instrument automatically calculating and retaining one decimal place. This example
 188 shows us the importance of recording siting criteria, required functional specifications
 189 of wind sensors and maintenance policy. However, those records are missing for most
 190 of the stations which hindered the quality classification and data processing.

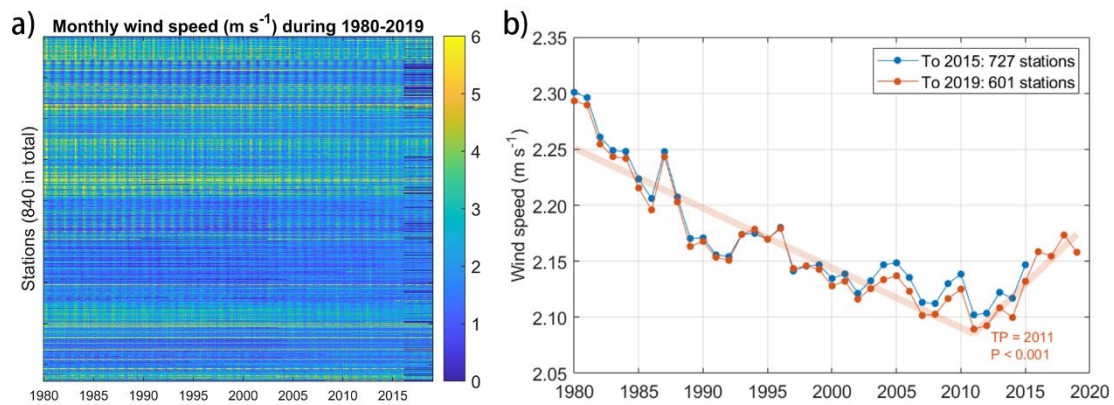


191
 192 **Figure 1. Changes in wind speed data caused by anemometer updates. a)** Decrease of
 193 frequency of zeros. Each horizontal bar in the upper figure represents one station and
 194 there are 840 stations in total. The color indicates the frequency of zeros (days per year).
 195 The black dotted line in the lower figure is the average annual frequency of zeros of all
 196 the stations. The 300 grey lines are sample averages, each containing 40% amount of
 197 the total stations. **b)** Frequency (days per year) of daily wind speed measurements
 198 between 2001 and 2006 for Station #52112 Naomaohu ($43^{\circ}45'N$, $94^{\circ}59'E$, 479.0 m
 199 a.s.l.)

200

201 3.2 Quality-controlled series

202 Following WMO's criteria, we generated the monthly average wind speed for each
 203 station (Figure 2a). We found that since January 2016, there have been 126 stations that
 204 no longer have records (distribution see Figure S3). We compared the time series with
 205 and without these stations and found the difference is not significant (t-test $P < 0.001$,
 206 Figure 2b). To obtain a longer time series including recent years' data, we deleted the
 207 126 stations and only used the 601 stations with complete monthly average wind speeds
 208 for 1980-2019. The breakpoint was detected in 2011 ($P < 0.001$) with a decreasing trend
 209 of $-0.011 \text{ m s}^{-1} \text{ year}^{-1}$ ($R^2 = 0.84$, $P < 0.001$) before the breakpoint and an increasing
 210 trend of $+0.022 \text{ m s}^{-1} \text{ year}^{-1}$ ($R^2 = 0.87$, $P < 0.001$) after.
 211

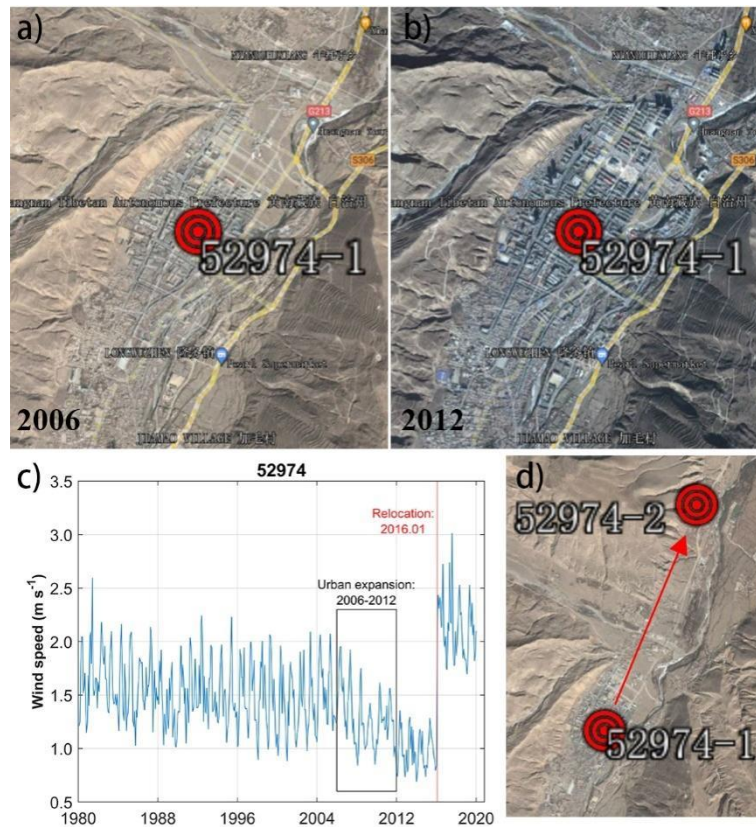


212 **Figure 2. Monthly average wind speed after being filtered by WMO's criteria. a)**
 213 Each horizontal bar represents one station. Months with no data (NaNs) are represented
 214 by the deepest blue. **b)** Comparison of the monthly average wind speed for the short-
 215 (1980-2015; 727 stations) and long-period (1980-2019; 601 stations)
 216

217
 218 **3.3 Station relocations caused by urbanization**

219 Another key factor influencing wind speed measurements is the relocation of
 220 stations. We found that there is a clear data jump caused by relocations in some of the
 221 stations. Taking the station located in Qinghai (station ID 52974) as an example, we
 222 detected an abrupt jump in wind speed in January 2016. This date coincides with the
 223 relocation of the station from $35^{\circ}31'N$, $102^{\circ}01'E$ (ID 52974-1) in December 2015 to
 224 $35^{\circ}33'N$, $102^{\circ}02'E$ (ID 52974-2) in January 2016 (Figures 3c & 3d). The relocation is
 225 potentially attributed to the urban growth around the station. As viewed by satellite

226 images from Google Earth Pro, there is a rapid urban expansion from 2006 (Figure 3a)
227 to 2012 (Figure 3b), especially towards the Northeast of the station, during wind speed
228 records also experienced a decrease (Figure 3c). A similar decrease in both daily mean
229 wind speed and maximum wind speed caused by urbanization was also reported in the
230 Yangtze River region (Zhang et al., 2022). To eliminate the effect of buildings on the
231 wind speed measurements, Station 52974 was moved to 4 km away from its previous
232 location (Figure 3d) so that wind speed is properly measured without artificial obstacles
233 in the surroundings. However, this estimation of roughness change based on satellite
234 data is rough. A more proper way as required by the World Climate Data and Monitoring
235 Programme is to record the change in the station logbook (WMO, 2021), which will
236 provide more reliable information about the quality of the data. But most stations don't
237 have such a record. Despite the absence of mete data, we used an established global
238 roughness model through satellite albedo observations to monitor alterations in surface
239 roughness. For the selected station, we employed the roughness estimation technique
240 devised by Chappell & Webb (2016) to analyze changes in roughness across a 5 km x
241 5 km area encompassing the station's location. Our quantitative examination of
242 roughness alterations aligns with the findings derived from satellite imagery analysis,
243 affirming a pronounced increase in roughness between 2000 and 2010 (Supplementary
244 Figure S4). This increase in roughness likely contributed to the observed decline in
245 wind speed and ultimately compelled the relocation of the station.
246



247

248 **Figure 3. Example of station relocation caused by rapid urbanization growth. a-b)**

249 Landsat images crop from Google Earth near Station 52974 in 2006 and 2012,

250 respectively. c) the wind speed change with urbanization and relocation. d) Landsat

251 images of the station relocation crop from Google Earth.

252

253 Though some stations were influenced by station relocation as shown in Figure 3,

254 a larger fraction (79%) of stations show no change in wind speed after the relocation.

255 Further checking the raw record of locations for those stations, we find that one reason

256 is that some “relocations” result from wrong location records. For example, Station

257 52974 is mistakenly detected with three relocations (Figure 4a). However, only the first

258 relocation is real and the latter two are results of location encoding change from 10202

259 (interpreted as 102°02′) to 1022 (interpreted as 10°22′) and back. Another possible

260 reason is that the relocation did happen but the data has been corrected. According to

261 the *Provisional Regulations on Relocation, Construction and Removal of National*

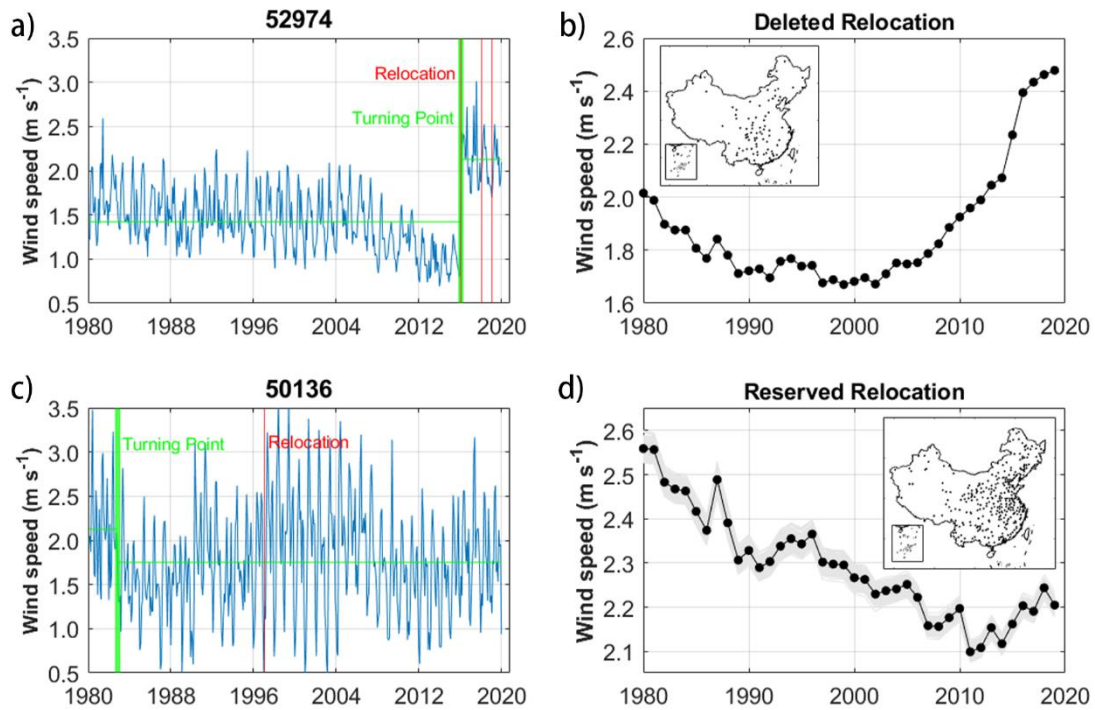
262 *Ground Meteorological Observation* announced by China’s government in 2012,

263 station relocations should have 1-2 years of parallel observations for data correction

264 (CMA, 2012). This process may fix some of those discontinuities but not all (Feng et
265 al., 2004; Fu et al., 2011; Patzert et al., 2016; Tian et al., 2019; Yang et al., 2021). For
266 example, Station 59287, the only national basic weather station in Guangzhou,
267 experienced two relocations in both 1996 and 2011, which is confirmed by the metadata
268 (CMA, 2011). After correction, the 1996 relocation doesn't show a sharp breaking point
269 but the 2011 one does (Supplementary Figure S5).

270 To examine whether the relocation caused a substantial change in the wind speed
271 record, we identified the most abrupt change in the wind speed time series and checked
272 whether a relocation happened near the change point (see details in *Methods 2.3*). Out
273 of the 432 relocated stations, 90 were deleted because the most significant shift in mean
274 is at the time of the relocation, and hence this is the most likely cause. We then took the
275 average of the "deleted relocation" stations and "reserved relocation" stations
276 separately. The "deleted relocation" group shows an abnormally rapid increase in the
277 recent two decades (Figure 4b). While the "reserved relocation" group is similar to
278 stations without relocation (Supplementary Figure S6). To exclude the impact of
279 different station counts in each category (fewer stations mean higher sensitivity to the
280 individual abnormal station), we performed 300 samples using a random draw of 90
281 stations from the "reserved relocation" group and showed them in grey lines in Figure
282 4d. None of the grey lines shows an abnormal trend as the "deleted relocation" group.
283 This proves that our method is efficient in identifying problematic stations.

284



285

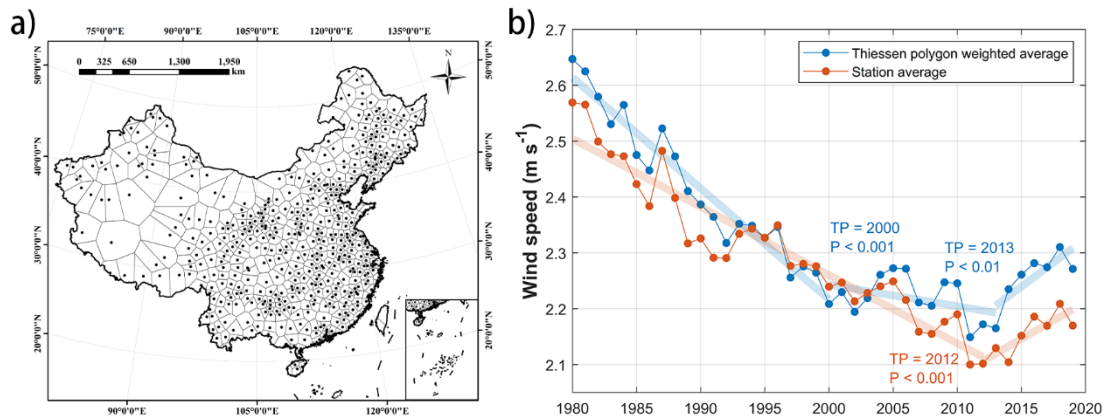
286 **Figure 4. Comparison of deleted relocated stations and reserved ones. a)** The wind
 287 speed data breakpoint and relocations of one example of deleted relocation, Station
 288 52974. **b)** The station average wind speed of 90 deleted relocated stations. The inset
 289 shows the station distribution across China. **c)** One example of reserved relocation,
 290 Station 50136. **d)** The station average wind speed of 342 reserved relocated stations.
 291 The grey lines are the averages of 300 samples, each with 90 randomly drawn reserved
 292 relocated stations. Maps information are from Department of Natural Resources
 293 standard map service system of China.

294

295 3.4 Average method used to calculate the national average

296 In the station average time series, the breakpoint was detected in 2012 ($P < 0.001$)
 297 with a trend of $-0.012 \text{ m s}^{-1} \text{ year}^{-1}$ ($R^2 = 0.90$, $P < 0.001$) before and $+0.013 \text{ m s}^{-1} \text{ year}^{-1}$
 298 ($R^2 = 0.70$, $P < 0.01$) after (Figure 5b). The increasing trend decreased by 41% after
 299 deleting those relocation-affected stations, compared with the $+0.022 \text{ m s}^{-1} \text{ year}^{-1}$ in
 300 Figure 2b (also reported by Liu et al., 2022a). But the trend is larger than the $+0.011 \text{ m}$
 301 $\text{s}^{-1} \text{ yr}^{-1}$, reported by Yang et al. (2021), with all the recorded location changed stations
 302 deleted without checking whether the station is affected by the relocation.

303 We further used Thiessen Polygon (Thiessen, 1911) to give different weights to
304 each station according to their representing area, i.e., large weight for stations located
305 in sparse stations area (Figure 5a) and compare the result with the station average
306 (Figure 5b). The Thiessen Polygon method, also known as the Voronoi Diagram, is a
307 spatial analysis technique often employed in hydrology and climatology. It involves
308 tessellating a region into polygons based on point data, such that each polygon
309 encompasses only one data point, and every location within a polygon is closer to its
310 associated point than any other. This method is particularly useful for interpolating
311 values across a region when the exact nature of change between points is unknown or
312 when changes are abrupt. By drawing perpendicular bisectors between adjacent data
313 points, the entire area is divided, with each polygon assuming the value of its associated
314 data point. While straightforward and clear in its delineation, the Thiessen Polygon
315 method assumes uniform variation within each polygon. The Thiessen polygon
316 weighted average is overall higher than the station average. This can be explained by
317 the increasing weight of stations in North West China with higher wind speeds (Liu et
318 al., 2019). While in the Thiessen polygon weighted average time series, there are two
319 breakpoints in 2000 ($P < 0.001$) and 2013 ($P < 0.01$). The trend changes from quick
320 decrease ($-0.020 \text{ m s}^{-1} \text{ year}^{-1}$, $R^2 = 0.94$, $P < 0.001$) to unstable moderate decrease ($-$
321 $0.004 \text{ m s}^{-1} \text{ year}^{-1}$, $R^2 = 0.17$, $P = 0.14$) and quickly increase ($+0.017 \text{ m s}^{-1} \text{ year}^{-1}$, $R^2 =$
322 0.64 , $P < 0.05$). The increasing trend in the recent decade increased by 31% (from
323 $+0.013 \text{ m s}^{-1} \text{ year}^{-1}$ to $+0.017 \text{ m s}^{-1} \text{ year}^{-1}$) after using the Thiessen polygon approach.
324 This is because the weights of stations in North West and South West are increased
325 when calculating the average and those area has strong increasing wind speed trend
326 (Figure S7). Despite the Thiessen polygon approach already utilizing the nearest station
327 observation to represent wind speed in locations lacking direct observations, it remains
328 unsatisfactory due to the intricate spatial variability of wind speed attributed to complex
329 terrains. To enhance the accuracy of wind speed interpolation, a more comprehensive
330 model necessitates additional observations within areas characterized by complex
331 terrain.



333

334 **Figure 5. Thiessen polygons and the comparison between Thiessen polygon**
 335 **weighted average and station average. a)** The Thiessen polygon map of the 511
 336 qualified stations. **b)** The comparison of station average wind speed (orange line) and
 337 Thiessen polygon weighted average (blue line) across China for 1980-2019. The linear
 338 fitting models are shown in translucent thick lines accordingly. Maps information are
 339 from Department of Natural Resources standard map service system of China.

340

341 4. Conclusions

342 Continuity is crucial for meteorological observation data. However, either the
 343 updates in the anemograph, the relocation caused by urbanization or the methods of
 344 data logging will affect wind speed data continuity. In this study, we comprehensively
 345 examined the discontinuity in wind speed data using a Chinese dataset. We found that
 346 updates to the automatic anemometer improved the observation frequency and
 347 instrument sensitivity, decreasing the zero-value daily wind speed data and increasing
 348 data accuracy. We also propose comparing the discontinuity in time series with recorded
 349 station relocation to check whether a relocation caused a breakpoint. We found that 90
 350 stations were affected by the relocation and show a quickly increasing wind speed in
 351 the recent two decades. After excluding those problematic stations, the wind speed
 352 reversal trend is reduced by 41% but still strong ($P < 0.001$, with an increasing trend of
 353 $+0.013 \text{ m s}^{-1} \text{ year}^{-1}$). The increasing trend reaches $+0.017 \text{ m s}^{-1} \text{ year}^{-1}$ ($R^2 = 0.64$, $P <$
 354 0.05) after using Thiessen Polygon, which gives the stations in North West China a

355 larger weight because their small number but located in a large area.

356 Though lots of methods (Masters et al., 2010; Wan et al., 2010; Cao & Yan, 2012;
357 Hong et al., 2014; He et al., 2014; Azorin-Molina et al., 2018; Camuffo et al., 2020)
358 were proposed to handle those problems, a comprehensive summary of them is missing.
359 Also, it is hard for external researchers to provide a better solution without a
360 collaboration with National Weather Services and the access to station data records
361 and/or metadata. Therefore, we hope National Weather Services could improve the data
362 quality based on these feedbacks and World Climate Data and Monitoring Programme's
363 guides, and complete the process by introducing an R package with open-source code
364 on GitHub and publishing the metadata. This way, not only the data is easier to get and
365 process, but also researchers can contribute to improve the dataset. One such example
366 is the "rnpn" package to access and process USA National Phenology Network data
367 (<https://github.com/usa-npn/rnpn>). Anyway, all raw data processing has limitations and
368 adds additional uncertainty. As we keep reporting problems in datasets and improving
369 our processing method, we should also pay more attention to increasing the quality and
370 homogeneity of the wind data. This requires raising awareness of the importance of
371 protecting the environment around the observation station and avoiding relocations.

372

373 **Supplemental Information**

374 Document S1. Supplemental Information, Table S1, Figures S1 – S7.

375

376 **Acknowledgements**

377 Thank Robert Dunn (UK Met Office) for discussions and comments on the manuscript.

378 Thank Adrian Chappell for providing the surface roughness data. The authors wish to
379 acknowledge the reviewers for their detailed and helpful comments to the original
380 manuscript.

381

382 This study was supported by the National Natural Science Foundation of China (grant
383 no. 42071022), the Swedish Formas (2019–00509 and 2017–01408) and VR (2021–
384 02163 and 2019–03954), and the start-up fund provided by Southern University of
385 Science and Technology (no. 29/Y01296122). C. A-M. was supported by VENTS
386 (GVA-AICO/2021/023), the CSIC Interdisciplinary Thematic Platform (PTI) Clima
387 (PTI-CLIMA), the 2021 Leonardo Grant for Researchers and Cultural Creators, BBVA
388 Foundation, and the “Unidad Asociada CSIC-Universidad de Vigo: Grupo de Física de
389 la Atmosfera y del Océano”.

390

391 **Data availability statement**

392 The data that support the findings of this study are available upon request from the
393 authors.

394

395 **Author contributions**

396 **Zhenzhong Zeng:** Conceptualization, Methodology **Yi Liu:** Methodology, Software,
397 Writing – Draft **Lihong Zhou:** Methodology, Data Curation **All other authors:** Writing
398 – Review & Editing

399

400 **Declaration of interest**

401 The authors declare no competing financial interests.

403 **References.**

- 404 1. Azorin-Molina, C., Asin, J., McVicar, T. R., Minola, L., Lopez-Moreno, J. I.,
405 Vicente-Serrano, S. M., & Chen, D. (2018). Evaluating anemometer drift: A
406 statistical approach to correct biases in wind speed measurement. *Atmospheric*
407 *research*, **203**, 175-188.
- 408 2. Azorin-Molina, C., Guijarro, J. A., McVicar, T. R., Trewin, B. C., Frost, A. J., &
409 Chen, D. (2019). An approach to homogenize daily peak wind gusts: An application
410 to the Australian series. *International Journal of Climatology*, **39**(4), 2260-2277.
- 411 3. Azorin-Molina, C., Vicente-Serrano, S. M., McVicar, T. R., Jerez, S., Sanchez-
412 Lorenzo, A., López-Moreno, J., ... & Espírito-Santo, F. (2014). Homogenization
413 and Assessment of Observed Near-Surface Wind Speed Trends over Spain and
414 Portugal, 1961–2011, *Journal of Climate*, **27**(10), 3692-3712.
- 415 4. Bathiany, S., Scheffer, M., Van Nes, E. H., Williamson, M. S., & Lenton, T. M.
416 (2018). Abrupt climate change in an oscillating world. *Scientific reports*, **8**(1), 1-
417 12.
- 418 5. Camuffo, D., della Valle, A., Becherini, F., & Zanini, V. (2020). Three centuries of
419 daily precipitation in Padua, Italy, 1713–2018: history, relocations, gaps,
420 homogeneity and raw data. *Climatic Change*, **162**(2), 923-942.
- 421 6. Cao, L. & Yan, Z. (2012). Progress in research on homogenization of climate data.
422 *Advances in Climate Change Research*, **3**(2), 59-67.
- 423 7. Chappell, A., & Webb, N. P. (2016). Using albedo to reform wind erosion
424 modelling, mapping and monitoring. *Aeolian Research*, **23**, 63-78.
- 425 8. China Meteorological Administration (CMA, 2017). *Meteorological data set*
426 *description document*. [In Chinese]
- 427 9. China Meteorology Administration (CMA, 2011). National Basic Meteorological
428 Station in Guangzhou was relocated four times in 62 years, *Media attention*,
429 Retrieved from:
430 http://www.cma.gov.cn/2011xwzx/2011xmtjj/201110/t20111026_121807.html.
431 [In Chinese]

- 432 10. China Meteorology Administration (CMA, 2012). Notice of the China
433 Meteorological Administration on the Issuance of Provisional Regulations on
434 Relocation and Removal of National ground meteorological observation Stations,
435 *State Council Bulletin*, Retrieved from:
436 http://www.gov.cn/gongbao/content/2013/content_2344560.htm. [In Chinese]
- 437 11. Dunn, R. J. H., Azorin-Molina, C., Menne, M. J., Zeng, Z., Casey, N. W., & Shen,
438 C. (2022). Reduction in reversal of global stilling arising from correction to
439 encoding of calm periods. *Environmental Research Communications*, **4**(6), 061003.
- 440 12. Feng, S., Hu, Q., & Qian, W. (2004). Quality control of daily meteorological data
441 in China, 1951–2000: a new dataset. *International Journal of Climatology: A*
442 *Journal of the Royal Meteorological Society*, **24**(7), 853-870.
- 443 13. Fu, G., Yu, J., Zhang, Y., Hu, S., Ouyang, R., & Liu, W. (2011). Temporal variation
444 of wind speed in China for 1961–2007. *Theoretical and Applied Climatology*. **104**,
445 313–324.
- 446 14. He, Y. C., Chan, P. W., & Li, Q. S. (2014). Standardization of raw wind speed
447 data under complex terrain conditions: a data-driven scheme. *Journal of Wind*
448 *Engineering and Industrial Aerodynamics*, **131**, 12-30.
- 449 15. Hong, H. P., Mara, T. G., Morris, R., Li, S. H., & Ye, W. (2014). Basis for
450 recommending an update of wind velocity pressures in Canadian design codes.
451 *Canadian Journal of Civil Engineering*, **41**(3), 206-221.
- 452 16. Hu, W., Kong, L., Zhu, X., & Xue, W. (2009). Accuracy analysis on contact
453 anemometer self–recording records digitization processing system. *Journal of Arid*
454 *Meteorology*, **27**, 168-171. [In Chinese]
- 455 17. Jin, Rui. (2011). Difference of wind measurement data between automatic station
456 and manual station. *Meteorological, Hydrological, and Oceanographic*
457 *Instruments*, **3**, 16-18. [In Chinese]
- 458 18. Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of
459 changepoints with a linear computational cost. *Journal of the American Statistical*
460 *Association*, **107**(500), 1590-1598.

- 461 19. Kuang, Yan. (2016). Application of magnetic rotary encoder in wind direction
462 anemometer. *Civil Aviation Management*, **311**(9), 52-53. [In Chinese]
- 463 20. Li, Y., Chen, Y., Li, Z., & Fang, G. (2018). Recent recovery of surface wind speed
464 in northwest China. *International Journal of Climatology*, **38**(12), 4445-4458.
- 465 21. Liu, F., Sun, F., Liu, W., Wang, T., Wang, H., Wang, X., & Lim, W. H. (2019). On
466 wind speed pattern and energy potential in China. *Applied Energy*, **236**, 867–876.
- 467 22. Liu, X. (2000). The homogeneity test on mean annual wind speed over China.
468 *Journal of Applied Meteorological Science*, **11**(1), 27-34.
- 469 23. Liu, Y. et al. (2022a). Increases in China’s wind energy production from the
470 recovery of wind speed since 2012, *Environmental Research Letters*, **17**(11).
- 471 24. Liu, Y., Xu, R., Ziegler, A. D. & Zeng, Z. (2022b). Stronger winds increase the
472 sand-dust storm risk in northern China, *Environmental Science: Atmospheres*,
473 online.
- 474 25. Masters, F. J., Vickery, P. J., Bacon, P., & Rappaport, E. N. (2010). Toward
475 objective, standardized intensity estimates from surface wind speed observations.
476 *Bulletin of the American Meteorological Society*, **91**(12), 1665-1682.
- 477 26. McVicar, T. R., Roderick, M. L., Donohue, R. J., Li, L. T., Van Niel, T. G., Thomas,
478 A., ... & Dinpashoh, Y. (2012). Global review and synthesis of trends in observed
479 terrestrial near-surface wind speeds: Implications for evaporation. *Journal of*
480 *Hydrology*, **416**, 182-205.
- 481 27. Patzert, W., LaDochy, S., Ramirez, P., & Willis, J. (2016). Los Angeles Weather
482 Station's relocation impacts climate and weather records.
- 483 28. Rayner, D. P. (2007). Wind run changes: the dominant factor affecting pan
484 evaporation trends in Australia. *Journal of Climate*, **20**(14), 3379-3394.
- 485 29. Shen, C., Zha, J., Wu, J., & Zhao, D. (2021). Centennial-Scale Variability of
486 Terrestrial Near-Surface Wind Speed over China from Reanalysis. *Journal of*
487 *Climate*, **34**(14), 5829–5846. <https://doi.org/10.1175/JCLI-D-20-0436.1>
- 488 30. Sohu (2004 9.21). Due to lack of sufficient attention, 60% of the national ground
489 meteorological observation stations were forced to relocate, *Domestic news*,

- 490 Retrieved from: <https://news.sohu.com/20040921/n222160625.shtml>. [In Chinese]
- 491 31. Tamura, Y. (2009). Wind-induced damage to buildings and disaster risk reduction.
492 *Proceedings of the APCWE-VII, Taipei, Taiwan.*
- 493 32. Thiessen, A. H. (1911). Precipitation averages for large areas. *Monthly weather*
494 *review*, **39**(7), 1082-1089.
- 495 33. Thiessen, A. H. (1911). Precipitation averages for large areas. *Monthly weather*
496 *review*, **39**(7), 1082-1089.
- 497 34. Tian, Q., Huang, G., Hu, K., & Niyogi, D. (2019). Observed and global climate
498 model based changes in wind power potential over the Northern Hemisphere
499 during 1979–2016. *Energy*, **167**, 1224-1235.
- 500 35. Trewin, B. (2010). Exposure, instrumentation, and observing practice effects on
501 land temperature measurements. *Wiley Interdisciplinary Reviews: Climate Change*,
502 **1**(4), 490-506.
- 503 36. Vautard, R., Cattiaux, J., Yiou, P., Thépaut, J. N., & Ciais, P. (2010). Northern
504 Hemisphere atmospheric stilling partly attributed to an increase in surface
505 roughness. *Nature Geoscience*, **3**(11), 756-761.
- 506 37. Wan, H., Wang, X. L., & Swail, V. R. (2010). Homogenization and trend analysis
507 of Canadian near-surface wind speeds. *Journal of Climate*, **23**(5), 1209-1225.
- 508 38. Wang, X. L. (2008). Accounting for autocorrelation in detecting mean shifts in
509 climate data series using the penalized maximal t or F test. *Journal of Applied*
510 *Meteorology and Climatology*, **47**(9), 2423-2444.
- 511 39. Wang, X., Piao, S., Ciais, P., Li, J., Friedlingstein, P., Koven, C., & Chen, A. (2011).
512 Spring temperature change and its implication in the change of vegetation growth
513 in North America from 1982 to 2006. *Proceedings of the National Academy of*
514 *Sciences*, **108**(4), 1240-1245.
- 515 40. World Meteorological Organization (WMO, 2003). Guidelines on Climate
516 Metadata and Homogenization. *WMO/TD - No. 1186; WCDMP - No. 53.*
517 https://library.wmo.int/doc_num.php?explnum_id=10751
- 518 41. World Meteorological Organization (WMO, 2017). WMO guidelines on the

- 519 calculation of climate normals. *WMO-No. 1203*.
520 https://library.wmo.int/doc_num.php?explnum_id=4166
- 521 42. World Meteorological Organization (WMO, 2020). Guidelines on homogenization.
522 *WMO-No. 1245*. https://library.wmo.int/doc_num.php?explnum_id=10352
- 523 43. World Meteorological Organization (WMO, 2021). Guide to Instruments and
524 Methods of Observation. *WMO-No. 8*.
525 https://library.wmo.int/doc_num.php?explnum_id=11612
- 526 44. Xin, Y., Chen, H., & Li, Y. (2012). Homogeneity adjustment of annual mean
527 wind speed and elementary calculation of fundamental wind pressure over
528 Xinjiang meteorological stations. *Climatic and Environmental Research*, **17**(2),
529 184-196. [In Chinese]
- 530 45. Yang, Jihua. (1986). The repair and maintenance of EL anemometer.
531 *Meteorology of Xinjiang*, **8**, 46-47. [In Chinese]
- 532 46. Yang, M., Zhang, L., Cui, Y., Yang, Q., & Huang, B. (2019). The impact of wind
533 field spatial heterogeneity and variability on short-term wind power forecast errors.
534 *Journal of Renewable and Sustainable Energy*, **11**(3), 033304.
- 535 47. Yang, Q., Li, M., Zu, Z., & Ma, Z. (2021). Has the stilling of the surface wind speed
536 ended in China? *Science China Earth Sciences*, **64**(7), 1036-1049.
- 537 48. Zeng, Z., Ziegler, A. D., Searchinger, T., Yang, L., Chen, A., Ju, K., ... & Wood, E.
538 F. (2019). A reversal in global terrestrial stilling and its implications for wind
539 energy production. *Nature Climate Change*, **9**(12), 979-985.
- 540 49. Zha, J., Shen, C., Zhao, D., Wu, J., & Fan, W. (2021). Slowdown and reversal of
541 terrestrial near-surface wind speed and its future changes over eastern China.
542 *Environmental Research Letters*, **16**(3), 034028. [https://doi.org/10.1088/1748-](https://doi.org/10.1088/1748-9326/abe2cd)
543 [9326/abe2cd](https://doi.org/10.1088/1748-9326/abe2cd)
- 544 50. Zhang, G. et al. (2022). Rapid urbanization induced daily maximum wind speed
545 decline in metropolitan areas: a case study in the Yangtze River Delta (China).
546 *Urban Climate*, **43**, 101147.
- 547 51. Zhang, G., Azorin-Molina, C., Chen, D., Guijarro, J. A., Kong, F., Minola, L.,

- 548 McVicar, T. R., Son, S., & Shi, P. (2020). Variability of Daily Maximum Wind
549 Speed across China, 1975–2016: An Examination of Likely Causes, *Journal of*
550 *Climate*, **33**(7), 2793-2816.
- 551 52. Zhou, Z., W, X., & Niu, R. (2002). Climate characteristics of sandstorm in China
552 in recent 47 years. *Journal of Applied Meteorological Science*, **13**(2): 193-200.
553