Impact of cruising speed on the ship-based sampling of marine fog frequency

Li Yi¹, King-Fai Li²

¹Frontiers Science Center for Deep Ocean Multispheres and Earth System, Key Laboratory of Physical Oceanography, Ocean University of China, Qingdao, Shandong, China
²Department of Environmental Sciences, University of California, Riverside, California 92521, USA

Correspondence to: King-Fai Li (king-fai.li@ucr.edu)

Abstract. Understanding secular changes in marine fog frequency is crucial for marine traffic planning under global change. Voluntary ship-based weather reports from community activities provide unique decadal records of marine weather conditions over world’s oceans, including visibility that implies the presence of marine fog. However, slowly changing external factors (such as the voyage technology, vessel types, etc.) may interfere with the secular changes in ship-based weather reports. Here we identify the cruising speed as an example of “target-induced” sampling biases in ship-based weather reports, where the fog itself causes the bias in its own sampling due to human’s decision. As a demonstration, we rectify the sampling bias in the marine fog frequency by multiplying the ratio of the cruising speeds under fog over the average cruising speeds under all weather conditions. The target-induced sampling biases may cause significant errors in the long-term trends of fog occurrences in the Okhotsk Sea, the Grand Banks, and the North Sea. Similar target-induced sampling biases may also be defined in the ship-based measurements of other weather phenomena.

1 Introduction

Marine fog causes low visibility and seriously affects human activities, such as marine traffic, fishery, port operation, etc., over remote ocean and coastal regions. As a special class of cloud, marine fog may also alter sea surface energy budget through changes in radiation and heat transfer (Koračin et al., 2005; Koračin et al., 2001; Lewis et al., 2003). Research-based operational marine fog observations are rare. Satellite measurements of marine fog is too short to for deducing the global change-related marine fog variations (Yi et al., 2019; Yi et al., 2016). To date, most of the operational fog detection are provided by weather stations located along coastlines or on islands. However, only marine fog that moves over land can be detected by these stations and the characteristics of marine and land fog may be very different (Dorman et al., 2020). Factors that determine the onshore movement of the marine fog include the directions of the sea breeze, the land height, and the land–sea temperature contrast (Cereceda et al., 2002; Findlater et al., 1989; Johnstone and Dawson, 2010; Lee and Chang, 2018).

Over remote oceans, ship-based fog observations during the voyage have been provided by commercial or public vessels since as early as the 1950s. These records have been archived in the International Comprehensive Ocean-
Atmosphere Data Set (ICOADS), which, albeit their sparse coverage in space and time, has become a unique dataset for the study of the variability of marine fog (Freeman et al., 2017; Koračin et al., 2014). A number of studies, however, are concerned about the applicability of ICOADS data on studying the long-term trends in fog, clouds, and other climate variables such as sea surface temperature (SST) due to various potential biases, including changes in the number of operating vessels (Norris, 1999), shipping routes, vessel types (Bajuk and Leovy, 1998), and observing techniques (James and Fox, 1972) that are usually not standardized for scientific purposes. Bajuk and Leovy (1998) showed that the anticorrelation of cumulonimbus frequency with the SST in the northern and southern hemispheres on the decadal time scale cannot be explained by any physical phenomenon but seemed to be consistent with the slow changes in aforementioned observational biases. Norris (1999) also proposed a similar explanation for the unexpectedly large trends in low cloud in the southern hemisphere that was inconsistent with the anthropogenic aerosol loading. Thomas et al. (2008) provided an example where the change in observational methods (e.g. visual and instrumental) and the increase in the anemometer height may lead to an artificial trend in the measured wind speed.

Figure 1. Photos illustrating the reduced visibility by fog during the voyage, which directly impact captain’s decision on the cruising speed. These photos were taken on-site by plot camera installed on the side of the Xuelong Research Vessel (Yi et al., 2022). (a) Visibility is 14,260 m located in 172.24°E, 55.73°N at 01:53 24th July, 2020; (b) Visibility is 4,220 m located in 177.83°E, 58.65°N at 20:12 24th July, 2020; (c) Visibility is 1,140 m located in 177.85°E, 58.47°N at 20:17 24th July, 2020; (d) Visibility is 960 m located in 177.87°E, 58.48°N at 20:22 24th July, 2020; (e) Visibility is 550 m located in 177.90°E, 58.49°N at 20:27 24th July, 2020; (f) Visibility is 380 m located in 177.97°E, 58.53°N at 20:43 24th July, 2020.
This paper provides another example of an observational bias related to marine traffic safety that may impact the long-term variability of marine fog frequency. This bias is self-induced by the target being measured (i.e., marine fog) because it is the presence of marine fog itself that leads to captain’s decision to vary the cruising speed for a safe voyage. Other severe weather phenomena (such as precipitation, wind speeds, ocean waves, etc.) could lead to similar biases in the observation of the weather phenomena per se.

To understand the origin of the “target-induced” bias in the marine fog observations, recall that the most intuitive definition of an ship-based marine fog frequency at a particular location is simply the number of weather reports signaling the presence of fog divided by the total number of weather reports at that location (e.g., Dorman et al., 2020). This definition, however, has a potential problem that may not have been considered widely in the literature: The fog occurrence is defined as a probability relative to the occurrence of non-fog events. Such a definition works well if the weather report is made by a stationary object like a fixed buoy, where the sampling of the fog events is uniform in time at a fixed location and the probability of fog occurrence at that location is not biased towards fog or non-fog events. However, for measurements over a moving vessel (or any moving objects), the time that the vessel spends in a particular region is dependent on the cruising speed. The cruising speed may, in turn for safety considerations, depend on the presence of fog, as well as other meteorological conditions. For example, many countries adopt marine traffic rules along their coastal regions to limit cruising speeds and reduce traffic accidents associated with low visibility (Figure 1). The U.S. Coast Guard’s Navigation Rules 83.19 (b) states that “Every vessel shall proceed at a safe speed adapted to the prevailing circumstances and conditions of restricted visibility”. The cruising speeds in remote oceans may also be subjectively adjusted by the captain due to other factors such as wave heights and wind speeds. The observed fog statistics may deviate from the real fog statistics due to the speed variations: If a vessel operates at a lower speed when there is fog relative to the speed when there is no fog, then the vessel would spend a longer time at the location under the fog and, consequently, the observed fog occurrences at this location would appear to be higher than it should have been because the vessel spends less time at the same location when there is no fog.

To see the target-induced bias mathematically, consider a 1-D track. Let \( n(x) \) be the number density of ship observations at location \( x \). If the grid size is \( \Delta x \), then the total number of ship-based observations in this grid is given by \( n(x) \Delta x \). \( n(x) \) is related to the cruising speed: Let \( v(x) \) be the cruising speed (typically measured relative to sea surface water) at \( x \) and \( \omega \) be the sampling frequency. \( \omega \) is assumed to be uniform (e.g., one record per minute, which is a common preset of commercial instruments) and independent of space and time. The number of observations to be made in the grid box is given by \( \frac{\omega \Delta x}{v(x)} \). Therefore, we can express \( n(x) \) in terms of \( v(x) \) and \( \omega \) as

\[
  n(x) \Delta x = \frac{\omega \Delta x}{v(x)}
\]  

or

\[
  n(x) = \frac{\omega}{v(x)}
\]
\[ n(x) = \frac{\omega}{v(x)}. \quad (2) \]

Let \( f_{\text{true}}(x) \) be the “true” fog frequency at location \( x \). Suppose the ship crosses the same grid box \( M \) times. Then \( f_{\text{true}} \times M \) times of the crossings observe fog and \( (1 - f_{\text{true}}) \times M \) times of the crossings do not observe fog. Let \( v_f(x) \) be the average cruising speed when there is fog and \( v_{\text{total}}(x) \) be the “grand” average cruising speed under all weather conditions (i.e. fog and fogless). Then the number of observations reporting fog is given by

\[ N_f(x) = f_{\text{true}}(x) \times M \times n(x) = \frac{f_{\text{true}}(x)M\omega}{v_f(x)}. \quad (3) \]

while the total number of observations is given by

\[ N_{\text{total}}(x) = \frac{M\omega}{v_{\text{all}}(x)}. \quad (4) \]

Finally, the observed fog frequency is given by

\[ f_{\text{obs}}(x) = \frac{N_f(x)}{N_{\text{all}}(x)} = \frac{f_{\text{true}}(x)v_{\text{all}}(x)}{v_f(x)} = \frac{f_{\text{true}}(x)}{r(x)}. \quad (5) \]

where

\[ r(x) = \frac{v_f(x)}{v_{\text{all}}(x)}. \quad (6) \]

Thus, \( f_{\text{obs}}(x) \) is not necessarily the same as \( f_{\text{true}}(x) \). Rather, it is scaled by a speed-dependent factor \( r(x) \). Figure 2 is a contour plot showing the dependence of \( f_{\text{obs}} \) on \( r \).

The above mathematical results can be understood intuitively: Assuming that for safety, the captain would reduce the speed when they are surrounded by fog. Because of the reduced cruising speed, the vessel would spend a longer time at the fog location than it would during another period when there is no fog at the same location. Thus, the weather report would be inherently biased with more fog records at this location [Eq. (3)] if the instrument sampling rate \( \omega \) is constant.

The special case, \( f_{\text{true}}(x) = f_{\text{obs}}(x) \), holds if and only if the cruising speed is the same \( [r(x) = 1] \) regardless of the fog condition. If the vessel sails at a lower speed in fog \( [r(x) < 1] \) than normal, then \( f_{\text{obs}}(x) > f_{\text{true}}(x) \), i.e. the actual fog
frequency would be overestimated. In the rest of the paper, we will apply the speed-correction to the ship-based fog observations and illustrate how the speed-correction might impact the long-term trends of the observed marine fog.

Figure 2. The difference between the “true” fog frequency $f_{true}$ and the observed fog frequency $f_{obs}$ due to the speed-correction factor $r$. The dashed line is the zero contour.

In the rest of the paper, we will calculate the average cruising speed on $1^\circ \times 1^\circ$ grids under non-fog and fog conditions using the ICOADS data as a preliminary investigation of the target-induced bias in the ship-based fog frequency. This provides an “Eulerian picture”, where the cruising speed in a grid box is simply the average of all speed data in the fog and non-fog groups, regardless of time and vessel types. In reality, different vessel types (e.g. oil tankers, container ships, or cruise ships) may adopt different average speeds in the open oceans. For example, the average speed of an oil tanker is 13–15 knots while the average speed of a cruise ship is 22–24 knots. As a result, in some $1^\circ \times 1^\circ$ grid boxes, the fog events may be observed by different vessel types cruising at different speeds, which our Eulerian picture does not consider. A possible drawback is that at some locations in the remote oceans, the Eulerian average of the cruising speed in fog may appear to be faster than the grand average under all weather conditions, which would appear to be against the safety consideration. A more accurate approach would be to identify individual ship tracks in the ICOADS record using the technique outlined by Carella et al. (2017) and correct the fog frequency using the speeds along each journey, which is beyond the scope of this paper.
2 Data and Method

2.1 Weather reports and cruising speed from ICOADS

This study uses ICOADS marine data in Summer (June, July and August) from 1968 to 2021 to extract fog and precipitation information from present weather records named as “WW” in ICOADS. A fog event is defined when the present weather code (a two-digit code from 01 to 99 based on SYNOP rules. WMO, 2009) lies between 40 and 49. Individual weather observations have been averaged over 1° × 1° latitude–longitude grid boxes.

The cruising speed is recorded by the variable “VS” in the ICOADS database. All the cruising speeds have been encoded with a 1-digit code from 0 to 9, which represent 10 speed intervals. However, the 10 speed intervals are different before and after January 1, 1968. In addition, we note that the cruising speed data were generally missing in the 1950s and in the early 1960s. Thus, in this work, we only focus on the records after 1st January 1968. The speed intervals after 1 January 1968 are 0–0, 1–3, 4–6, 7–9, 10–12, 13–15, 16–18, 19–21, 22–24, and over 24 knots. We use the central values of 0, 2, 5, 8, 11, 14, 17, 20, 23 knots of the first 9 intervals when averaging the individual records.

Figure 3 shows the average cruising speed in the northern mid-latitudes from June to August during 1968–2021 using all vessel records in the ICOADS database. The average speeds in most of coastal regions and latitudes greater than 60°N are less than 10 knots. In contrast, in the remote oceans such as the Pacific and Atlantic basins, the average speed is around 16–18 knots. The spatial pattern of the high-speed regions coincides with the North Pacific/Atlantic Drifts and the Golf Stream, indicating that vessels going eastward in these regions take the advantages of the currents.

2.2 “Eulerian” Correction Factor

At a location \( x \) (where \( x \) is the center of a 1°×1° grid box), we define the observed fog frequency \( f_{\text{obs}}(x) \) as the...
number of ICOADS records reporting visibility below 1 km (i.e. when the vessel is cruising in fog) divided by the total number of ICOADS records in the same grid box. To correct the bias in \( f_{\text{obs}}(x) \) due to the speed variation, we first calculate the time-averaged speeds under fog conditions as well as under all weather conditions to estimate \( v_f(x) \) and \( v_{\text{all}}(x) \), respectively. Then we define a corrected fog frequency \( f_{\text{corr}}(x) \) using the following equation:

\[
f_{\text{corr}}(x) = r(x) f_{\text{obs}}(x). 
\]  

Eq. 7 only explicitly shows the spatial dependence. However, the fog frequency may also change in time, e.g., due to global change. Let \( t \) be the summer year; that is, \( t \) may be 1968, 1969, ..., and 2021. If the evolution of the observed fog frequency is denoted by \( f_{\text{obs}}(x, t) \), then the evolution of the corrected fog frequency \( f_{\text{corr}}(x, t) \) is given by

\[
f_{\text{corr}}(x, t) = r(x, t) f_{\text{obs}}(x, t). 
\]

where \( r(x, t) \) is the average cruising speed in fog in a given summer year \( t \) divided by the average cruising speed under all weather conditions.

3 Results

3.1 Fog frequency and the speed-correction factor

The ICOADS data coverage in the southern hemisphere is sparser than in the northern hemisphere. Dorman et al. (2017) shows that the estimated fog frequency may be correlated with the number of observations available in a grid box when the number of observations is less than ~100, which generally holds in the southern hemisphere. For our purpose to illustrate the effect of the cruising speeds, it suffices for us to focus only on the ICOADS data in the northern hemisphere.

Figure 4a shows the raw average summer fog frequency between 1968 and 2021 derived from the ICOADS data in the northern hemisphere before the speed-correction is applied. Regions of high fog occurrences include the northwestern Pacific near the Kuril Islands and Okhotsk Sea, the northwestern Atlantic near the Labrador Sea, and the east coast of Greenland. Most of these frequent fog regions lie at intersections between subtropical and mid-latitude gyres. For example, the high fog frequency region in the northwestern Pacific is located at the intersection of the warm poleward Kuroshio current in West Pacific and the cold equatorward Oyashio Current in North Pacific. Similarly, the high fog frequency regions in the northwestern Atlantic are located at the intersection of the Gulf Stream and the Labrador Current and at the intersection of the East Greenland Current and Irminger Current. Steam fog is likely formed at these intersections when the warmth and the moisture from the south are mixed with the cold and dry air from the north. Furthermore, low air
temperature resulted from ocean-tidal cooling weakens the surface wind and lead to strong surface inversion up to 0.5 km in height, which is a typical precondition of fog formation (Tokinaga and Xie, 2009). The boundaries of the high fog regions may be correlated with the boundaries of the continental shelves (Dorman et al., 2020).

Figure 4. (a) The raw fog frequency \(f_{\text{obs}}\) over the period between 1968 and 2021, calculated directly using the weather report in ICOADS on \(1^\circ \times 1^\circ\) grids without the speed-correction; (b) The average speed-correction factor, \(r\) in Eq. (6), of fog frequency over the same period between 1968 and 2021 on the same grids. “ND” means “No Data”. 80% of the values of \(r\) lies between 0.85 and 1.15. A \(9^\circ \times 5^\circ\) spatial smoothing has been applied to highlight the large-scale variability.

Figure 4b shows the speed-correction factor \(r\). A \(9^\circ \times 5^\circ\) smoothing filter has been applied to both figures to highlight large-scale variabilities. Note that the spatial patterns of high and low \(r\) generally have a length scale greater than the size of the smoothing filter, so the spatial correlation is not an artifact of the smoothing. The blue and red shades represent areas where the cruising speed is lower and higher than the average, respectively. We emphasize that the changes in the cruising speed shown in Figure 4b do not necessarily imply the decisions made by the captains under fog conditions; part of the variations shown in Figure 4b could be results of the Eulerian picture that averages speeds over unevenly sampled vessel types when fog occurs, as we have discussed in Section 1. The speed-correction factor is thus defined only relative to the Eulerian mean and may be greater than 1 even if the captain has reduced the cruising speed. Figure 4b shows that the average cruising speeds in fog and non-fog conditions may vary by as much as ±14%. The corrections in the remote oceans,
e.g. along the North Pacific Drift where the average cruising speeds are high, are generally a few percent. Large corrections are mostly found in coastal regions (c.f. Figure 3).

### 3.2. Trend corrections

Dorman et al. (2020) commented that the trend in the ship-based marine fog observations may be erroneous. Below we illustrate how the cruising speed may affect some regional long-term trends of the fog frequency.

Given a time series in a $1^\circ \times 1^\circ$ grid box, we fit $f_{\text{obs}}(x, t)$ and $f_{\text{corr}}(x, t)$ individually with a linear trend:

$$f_{\text{obs}}(x, t) = a_{\text{obs}}(x) t + b_{\text{obs}}(x).$$

$$f_{\text{corr}}(x, t) = a_{\text{corr}}(x) t + b_{\text{corr}}(x).$$

![Map of fog frequency trends](image.png)

**Figure 5.** The linear trend of summer (June–August) fog frequency (expressed in the unit of %/decade) from 1968 to 2021. No spatial smoothing has been applied. Panel (a) shows the raw observed trend $a_{\text{obs}}$ based on the observed fog frequency $f_{\text{obs}}$; Panel (b) shows the revised trend $a_{\text{corr}}$ based on the speed-corrected fog frequency. The black dots indicate regions where the trends are statistically significant at the 95% confidence level. Four $10^\circ \times 10^\circ$ regions to be studied in Figure 6 are enclosed by the rectangular boxes.
Figure 6. Long term trends of observed fog frequency and corrected fog frequency in three 10°×10° box regions. (a) East of the New York State: 70°W–60°W, 35°N–45°N; (b) The Grand Banks: 50°W–40°W, 40°N–50°N; (c) South of Kuril Islands: 140°E–150°E, 35°N–45°N. The left y-axis presents the frequency value (%) and the right y-axis presents the number of reports in all weather conditions (i.e. fog + non-fog).
where \( a_{\text{obs}}(x) \) and \( a_{\text{corr}}(x) \), to be expressed in the unit of fog frequency per decade (%/decade), are the trends derived from the observed (Figure 5a) and speed-corrected (Figure 5b) fog frequencies, respectively. The dot-shaded areas in Figure 5 represent regions where the trends are significant at the 95% confidence level. Note that no spatial smoothing has been applied in Figure 5. Therefore, the spatial pattern of the trends is not due to artificial averaging.

The signs of the trends in most of the remote oceans are consistent before and after the speed-correction, except for those boxed regions which we will focus on in the next paragraph. There are statistically significant positive trends in most of the northwest Pacific, the Yellow Sea, the Grand Banks, and around Iceland. These regions coincide with the regions where the mean marine fog frequency is above 10% (c.f. Figure 4). The strongest positive trend of 4%/decade is found in the Okhotsk and Bering Seas surrounding the Kamchatka Peninsula. Weak (less than \(-2\%\)/decade) but statistically significant negative trends are found in the Gulfs of Alaska and St. Lawrence, the Japan Sea between Korea and Japan, and the North Sea between England and Norway. The mean marine fog frequency in most of these regions is less than 10%.

Figure 6 shows the change of the trends in 4 selected regions: the east of the New York State (Box 1), the Grand Banks (the east of Newfoundland; Box 2), the north of Scotland (Box 3), and the south of Kuril Islands (Box 4). These regions are enclosed by the \(10^\circ \times 10^\circ\) boxes shown in Figure 5. The raw and speed-corrected fog frequencies are shown as the red and blue solid lines, respectively. The ratio of the red solid line over the blue solid line is the annual value of the speed-correction factor \( r \). The annual numbers of ship reports being averaged in the \(10^\circ \times 10^\circ\) boxes are represented by the yellow bar histogram.

In Box 1 (Figure 6a) in the east of New York State, the raw fog frequency is relatively stable at \(~15\%\) before 1980, which then increased to \(~25\%\) in the mid-1980s and slowly decreased to a minimum of \(~5\%\) in 2015. If a line fit is applied to the raw fog frequency in this region over the period from 1968 to 2021, then the overall trend obtained from this line fit is \(-1.6\pm0.9\%/\text{decade}\), apparently implying that the fog occurrence might have reduced in the last 50 years. However, after correcting the speed effects, the decreasing trend becomes statistically insignificantly primarily due to the correction of the slower speeds during 1970–2005. In Box 2 (Figure 6b) in the Grand Banks, the raw fog frequency is biased high in the latter part of the time series (between 1980–1990 and after 2000), resulting in a strong increasing trend of \(2.8\pm1.0\%/\text{decade}\); the increasing trend remains statistically significant after correcting the slower speeds in those periods, but the trend reduced almost by a factor of 3 to \(1.0\pm0.6\%/\text{decade}\). In Box 3 (Figure 6c) in the north of Scotland, the raw decreasing trend of \(-1.8\pm0.9\%/\text{decade}\) becomes statistically insignificant to \(0.2\pm0.9\%/\text{decade}\) after the speed-correction, almost entirely due to the revised fog frequency in the last 10 years of the record during 2010–2021. Lastly, in Box 4 (Figure 6d) in the south of Kuril Islands, although the raw trend is statistically insignificant \((-0.5\pm1.4\%/\text{decade}\)), the revised trend after the speed correction becomes statistically significant at \(-1.6\pm0.8\%/\text{decade}\).

4 Summary and concluding remarks

We have discussed a target-induced bias in ship-based marine fog frequency, which can be corrected by the ratio of
the cruising speed in fog conditions over the cruising speed in all weather conditions. We showed that the observed trends in some high-fog regions may become statistically insignificant, or vice versa, after applying the speed-correction factor, so the varying cruising speed could impact our current understanding of long-term fog variability. We note that our adoption of the Eulerian picture, i.e. an all-time average of the cruising speed in fixed grid boxes, only provides an illustration of the effects of the target-induced bias in ship-based fog observations. A more quantitative study requires a Lagrangian picture where the speed-correction is done along individual journeys.

Any meteorological variables that would influence the captain’s decision would self-induce similar biases in the measurements during the voyage. For example, precipitation may also reduce the visibility like fog and therefore lead to variability in the cruising speed. We recalculate the speed-correction factor using the ICOADS precipitation data, which is given by the average cruising speed when there is precipitation divided the average cruising speed under all weather conditions. (Note that the average cruising speed under all weather conditions is the same as the one we used for marine fog.) A precipitation event is defined when the present weather code “WW” lies between 50 and 99. Figure 7 shows that the raw precipitation frequency (i.e. before the application of the correction factor) may have been overestimated in most oceans.
Therefore, the effects of the cruising speed in the precipitation should also be considered when interpreting the ship-based observation.

Acknowledgements. The authors thank Mao-Chang Liang, Ka-Kit Tung, and Xianyao Chen for helpful discussions.

Financial support. YL was supported by the National Key Research and Development Program of China (2019YFA0607000) and the National Natural Science Foundation of China (41975024).

Data availability. The ICOADS data are freely available at https://icoads.noaa.gov/products.html.

Author contributions. YL and KFL contributed equally to this research.

Competing interests. The authors declare that they have no conflict of interest.

References


