Manuscript: Statistical exploration of gaseous elemental mercury (GEM) measured at Cape Point from 2007 to 2011 (Ref. No.: acp-2014-971), Atmos. Chem. Phys. Discuss., 15, 4025–4053, 2015, doi:10.5194/acpd-15-4025-2015.

Anonymous Referee #1

Received and published: 05 March 2015

We thank Anonymous Referee #1 for the positive review of this paper and recommending that the paper can be published in its current form.

Anonymous Referee #2

Received and published: 04 May 2015

We thank Anonymous Referee #2 for a detailed review of our manuscript. The comments and remarks have been processed and explained in the manuscript, which we believe has gained in clarity and scientific soundness. Below is a point-by-point reply (in blue) to the comments of Referee #2 (in black font).

Venter and co-authors present a statistical exploration of a 4-year record (2007-2011) of gaseous elemental mercury (GEM) concentrations measured at Cape Point, South Africa. Firstly, this paper aims at identifying the origin of high and low mercury concentrations events using a dataset already presented and discussed elsewhere (e.g., Slemr et al., 2013; Slemr et al., 2015) and based on back-trajectories and cluster analysis. Secondly, multi-linear regression analysis was used to predict GEM concentrations from other atmospheric parameters measured at the station. The paper is clearly organized, easy to follow and well written but overall lacks robust statistics. I recommend major revisions.

Major comments

1.Cluster analysis

I generally agree with the comments of B. Denzler and will avoid duplication. Briefly, I would focus the analysis on the extreme data points. Indeed using only two clusters, most of the points lie very close to the 0.904 ng/m3 threshold and results from the two clusters are very close (Fig.4).

Since Referee #2 and B Denzler had similar comments relating to the cluster analysis we addressed these comments together.

We agree with Referee #2 and B Denzler that a two cluster analysis results in points lying very close to the 0.904 ng/m^3 threshold. However, in this study the aim was for statistical analysis of the data to lead to certain deductions/conclusions. Therefore, it was not decided beforehand that a two grouping solution was going to be used. Interpretation of the cluster analysis resulted in identifying two clusters as the best solution with the optimum separation and representation of the data. Although a 5, 6 and 7 cluster solution was obtained with slightly better separation, some of these clusters represented only a small fraction of the total GEM distribution at Cape Point. Clusters representing smaller fractions of the GEM data are considered to be extreme cases that can be evaluated as separate case studies. Studies of this nature can be an important future perspective, which was indicated in the Conclusions as follows:

"Data indicated as extreme events, as indicated by 5, 6 and 7 cluster solutions should be investigated as special case studies."

Furthermore, the 0.904 ng/m^3 threshold is close to, but not the same as the mean and median of the data, which are 0.917 and 0.925 ng/m^3, respectively. This threshold value was statistically determined with cluster analysis and is specific for this dataset. If the dataset was only separated according to mean/median value, extreme data points would have been included in calculations. After cluster analysis data not included in the clusters (that can include extreme data points) was not considered in further processing.

We agree with B Denzler that the use of quartiles would have been a feasible alternative. However, this approach would have required a pre-classification of the data into four groups and using two of these (highest and lowest) in further processing. As mentioned, in this study the aim was for statistical analysis of the data to lead to certain deductions/conclusions and not pre-classification. In our approach cluster analysis was performed that resulted in a 2 cluster optimum solution. Cluttering was dealt with by subtracting the back trajectory analysis of each of the two clusters identified, which resulted in more detailed plots from which two distinct observations could be made, i.e. GEM values associated with ship routes and air masses of continental origin.

2.Trend sign at Cape Point

The authors report a decline in GEM concentrations at Cape Point over the 2007-2011 period. In contrast, Slemr et al. (2015) reported a change in the trend sign at the same station from decreasing mercury concentrations in 1996-2004 to increasing concentrations over the 2007-2013 period. How do the authors explain these contrary conclusions?

The trends identified by Slemr et al. (2015) and in this paper cannot be directly compared. Slemr et al. (2015) utilised pre-processed, i.e. de-trended and de-seasonilised data, while in this study all the data was considered for statistical evaluation. In order to clarify this to the readers the following text was added:

"In contrast, Slemr et al. (2015) reported an increase in GEM concentration at CPT GAW. However, this increase was calculated by utilising pre-processed, i.e. de-trended and de-seasonilised data, which was not the case in this study. Therefore these different approaches cannot be directly compared."

3.Back-trajectories as an alternative tool to distinguish continental/marine GEM contributions

Several studies (e.g., Slemr et al. (2013) and Brunke et al. (2004)) used 222Rn measurements to determine the continental/maritime origin of air masses reaching the Cape Point station. In this paper, the authors used back-trajectories as an alternative tool in order to distinguish continental and marine GEM contributions. The hourly arriving back trajectories were divided into groups according to the time that these air masses had spent over the continent. This work needs to include a more critical discussion of results obtained by both methods and associated uncertainties. According to the authors, the errors accompanying a single trajectory are 15-30% of the trajectory distance travelled. How does it compare with 222Rn measurements uncertainties? Could 222Rn concentrations also be a tool to determine the time spent by a trajectory over the continent? Lacking any of the above, it is not clear to me what new results this paper brings to the topic.

Although ^{222}Rn is considered to be a good tracer for studies of emissions from terrestrial surfaces, according to Jacob et al. (1997), the assumption of a uniform ^{222}Rn emission rate of 1 atom cm^{-2} s^{-1} is accurate to roughly 25% globally, or by a factor of 2 regionally. Therefore the 15 – 30% error associated with back trajectory analysis is in the same range as the uncertainties associated with ^{222}Rn as tracer. Fig. 6 was augmented by adding Fig. 6(b) that indicates the ^{222}Rn concentration range associated with air masses classified by back trajectory analysis. It is evident from comparison between Fig. 6(a) (originally Fig. 6) and Fig. 6(b) that back trajectory analysis provides a more sensitive method of characterising GEM according to time that air masses spent over the continent up to 11 hours, while ^{222}Rn classification only allows separation within three hours that air masses spent over the continent. The text was modified accordingly to:

"However, the ^{222}Rn classification method only allows for the separation of the CPT GAW GEM data into relatively few classes, i.e. marine background, mixed and continentally influenced, while the back trajectory analysis methods provide a more quantified classification based on the length of time that air masses spent over the continent resulting in increased GEM concentrations. It is evident from comparison between Fig. 6(a) and Fig. 6(b) that back trajectory analysis provides a more sensitive method of characterising GEM according to time that air masses spent over the continent up to 11 hours (where GEM concentrations reached a plateau), while ^{222}Rn classification only allows separation within three hours that air masses spent over the continent.The difference in average GEM concentrations between air masses that had spent one hour or less over the continent, i.e. 0.92 ng m^{-3} and air masses that had spent more than 11 hours on the continent, i.e. 1.09 ± 0.150 ng m^{-3}, therefore provides some quantified indication of the possible continental contribution of GEM at CPT GAW. When GEM concentrations were classified according to ^{222}Rn levels, i.e. ^{222}Rn levels > 1000 mBq m^{-3} indicating continentally influenced air masses (Slemr et al., 2013), 50% of the data was greater than 0.92 ng m^{-3}. This value is somewhat lower than the average concentration value determined for air masses spending more than 11 hours over the continent, i.e. 1.09 ng m^{-3}.

According to Jacob et al. (1997), the assumption of a uniform ^{222}Rn emission rate of 1 atom cm^{-2} s^{-1} is accurate to roughly 25% globally, or by a factor of 2 regionally. Therefore the 15 – 30% error associated with back trajectory analysis is in the same range as the uncertainties associated with ^{222}Rn as tracer."

Minor comments

p. 4026

l. 8: please define SA

We have replaced "SA" with "South Africa" in the text as follows:

"...semi-arid interior of South Africa and..."

l. 17-19: "Both measured and MLR calculated data confirm a decline in GEM concentrations at CPT GAW over the period evaluated". See major comment #2.

This was addressed in major comment #2.

p.4028

l. 3-5: Angot et al. (2014) and Slemr et al. (2015) should be included as references in addition to Ebinghaus et al. (2002).

These references were included in the text:

"...for the Southern Hemisphere (Ebinghaus et al. (2002); Slemr et al., 2011; Angot et al. (2014); Slemr et al. (2015))."

Angot et al. (2014) and Slemr et al. (2015) were added to the Bibliography. Ebinghaus et al. (2002) was included in the Bibliography.

p. 4031

l. 24-28: “Eight-day back trajectories with hourly arrival times at an arrival height of 100m (....). An arrival height of 100m was chosen since the orography in HYSPLIT is not very well defined, and therefore lower arrival heights could result in increased error margins”. I wonder why the authors used an arrival height of 100m given that measurements are were carried out higher, on the top of a cliff at 230m a.s.l.

The 100m arrival height mentioned here is above ground level of the location from where the back trajectory is calculated and not above sea level. This was clarified in the text:

“...at an arrival height of 100 m (above ground level) were calculated...”

p. 4033

l.21-23: “However, significant differences between these two overlay trajectory maps (: : :) are not that evident”. I agree, see major comment #1.

As mentioned at comment #1 subtraction of the back trajectory analysis of each of the two clusters resulted in more detailed maps from which distinct observations could be made. This is also explained in the text:

“Therefore, a third overlay trajectory map (Fig. 4c) was drawn, which represents the difference between the two individual maps, i.e. subtracting the percentage of trajectories passing over each correlating 0.2° x 0.2° grid cell in Fig. 4b from the percentage of trajectories passing over each 0.2° x 0.2° grid cell in Fig. 4a. In Fig. 4c, positive values (red) correspond with areas over which cluster one’s (> 0.904 ng m^{-3}) air masses dominated, whereas negative values (dark blue) indicate areas over which air mass movement of cluster two (< 0.904 ng m^{-3}) were dominant. From this map (Fig. 4c), two observations can be made. Firstly, oceanic regions along both the east- and west coast around CPT GAW correspond with air masses mostly related to cluster one (higher GEM values), which could potentially indicate the influence of shipping routes on GEM measured at CPT GAW. Secondly, air masses that had passed over the very sparsely populated semi-arid Karoo region, almost directly to the north of CPT GAW, were mostly associated with cluster two (lower GEM values).”

p. 4035

l. 12-13: "An evident trend is observed in Fig. 6, i.e. an increase of GEM concentrations for air masses that spent more time over the continent". Air masses spending less than 10 hours over the continent are associated with highly variable GEM concentrations. Is the mean statistically different from one group to another? This should be tested statistically.

The graph presented in Fig. 6 is statistical representation of the Hg data with mean, median and quartile values, as well as 99.3% of the data coverage indicated. This is explained by the caption of the figure:

"Figure 6: The statistical distribution of GEM concentrations as a function of time spent over the continent. The mean is indicated by the black stars, the median by the red line, the 25- and 50 percentile by the blue box and the whiskers indicating 99.3 % data coverage (if a normal distribution is assumed), while the black line connects the mean values to provide an indication of the trend observed"

From the whiskers of the plot it is evident that there are more lower Hg levels associated with air masses that spent shorter periods over the continent. In addition, considering the typical atmospheric Hg concentrations measured at Cape Point, there is also a significant difference between the mean/median values of air mass spending less than 10 hours over the continent.

p. 4036

l. 1-3: "The average marine background GEM concentration for the entire sampling period according to the 222Rn level classification (<350 mBq/m3 –as proposed by Slemr et al. (2013) and Brunke et al. (2004)) was 0.92 ±0.275 ng/m3." I believe they rather used a 100-250 mBq/m3 threshold. Does it affect the calculated mean marine background GEM concentration?

We thank Reviewer #2 for pointing out this misquote. The threshold value must be 100 mBq/m3 as indicated by Brunke et al. (2004). The reference to Slemr et al . (2013) was removed from the text, since these authors did not consider marine classification. The average marine background GEM was recalculated for the 100 mBq/m3 threshold as 0.89 ± 0.106 ng/m3. The text was changed as follows:

"The average marine background GEM concentration for the entire sampling period according to the ^{222}Rn level classification (< 100 mBq m^{-3} – as proposed by Brunke et al., 2004) was 0.89 ± 0.106 ng m^{-3}, while the average GEM level for air masses that spent one hour or less over the continent (Fig. 6) was 0.92 ± 0.300 ng m^{-3}."

l. 17-20: "When GEM concentrations were classified according to 222Rn levels, i.e. 222Rn levels > 1200 mBq/m3 indicating continentally influenced air masses ((Slemr et al., 2013) and (Brunke et al., 2004)), 50% of the data was greater than 0.99 ng/m3". Same as above, Slemr et al. (2013) used a threshold of > 1000 mBq/m3 rather than > 1200 mBq/m3. Does it affect the calculated mean GEM concentration of continentally influenced air masses?

Similar to the previous comment the threshold was changed to > 1000 mBq/m3 and Brunke et al., (2004) was removed from the text. The median continental GEM was recalculated for the 1000 mBq/m3 threshold as 0.92 ± ng/m3. The text was changed as follows:

"When GEM concentrations were classified according to ^{222}Rn levels, i.e. ^{222}Rn levels > 1000 mBq m^{-3} indicating continentally influenced air masses (Slemr et al., 2013), 50% of the data was greater than 0.92 ng m^{-3}."

p. 4037

l. 7-9: "Minimization of the RSME was attained when the number of independent variables included in the optimum solution of the equation was increased to eight, and had a RMSE of 0.1205". Values of RMSE are very close to each other. How do you know if the small difference is statistically significant?

We agree with Reviewer #2 that this was not explained adequately. The text was changed to clarify as follows:

"...if the optimum MLR solution contained more independent variables. The optimised RMSE was attained when the number of independent variables included in the optimum solution of the equation was increased to eight, and had an RMSE of 0.1205. The measure of optimisation was taken as at least 1% contribution to the overall reduction of RMSE. Table 2 indicates the identity of the independent parameters determined for each of the optimum MLR solutions."

p.4039

l.14: "a slight decrease of GEM concentrations at CPT GAW over the evaluated period". Please see major comment #2.

This was addressed in major comment #2.

p.4040

l.2-3: "such analyses could be used as an alternative tool to distinguish between continental and marine GEM contributions". Please see major comment #3.

This was addressed in major comment #3.

Figure 1: It is hard to see anthropogenic point sources. Please consider using different colors.

We agree that it is difficult to see the different types of sources. On this scale it is also difficult to improve differentiation, since point sources spatially overlap in certain areas. This implies that the use of different markers and/or colours would not improve the legibility. However, our intention was to indicate the concentrations of major point sources in this region of southern Africa and not necessarily the different types. Therefore we have grouped the three different source types to be represented with the same marker.

Table 1: What about the eight-, nine- and ten-cluster solutions?

As explained in our response to major comment #1, a two cluster solution was considered to be the optimum separation and representation of the Cape Point Hg data. Fig. 2 was augmented to only indicate 7 clusters to avoid confusion as pointed out by Reviewer #1.

Table 2: Please define WGS.

WGS was defined in the text on p 4037 l. 27.

B. Denzler

Received and published: 24 March 2015

We thank B. Denzler for a detailed review of our manuscript. The comments and remarks have been processed and explained in the manuscript, which we believe has gained in clarity and scientific soundness. Below is a point-by-point reply (in blue) to the comments of B. Denzler (in black font).

Summary

In the article the authors analyze a gaseous elemental mercury (GEM) time series from the Cape Point Global Atmosphere Watch (CPT GAW) station ranging from 2007 until 2012. Different statistical methods and back-trajectory analysis were applied to identify the origin of high and low mercury concentrations. Furthermore, multiple linear regression (MLR) was used to predict mercury concentrations at CPT GAW from trace gas concentration and other atmospheric parameters. The regression was also used to gain insight into the relation of the parameters with mercury concentrations.

General impression

I regard the measurement series especially at this location in the southern hemisphere as highly important. Therefore, the analysis of this series is of great interest. Generally I would argue that the methods used to either identify source regions or estimating GEM concentrations are not suitable and not well enough applied to draw concise conclusions. I encourage the authors to reconsider their methods before resubmitting the manuscript. The data-set is highly interesting and worth being published.

Major comments

1. Cluster analysis

The cluster analysis was used to distinguish between high an low concentrations. Strangely only two clusters were formed. The authors justify this choice with a high silhouette number for two clusters (fig. 2) and significant amount of GEM data. But since only two groups are formed I would assume the separation at 0.904 ng/m3 amounts about to the mean or median concentration, which could serve as a separation equally good.

The problem with using only two clusters is visible when looking at the source region analysis. Here they compare all the values above 0.904 ng/m3 with the ones below. Yet most of the measurement points lie very close to this line and are certainly not containing much valuable information and still dominate the plots (fig. 4). I would argue that using quantiles on their data and comparing for example data below the first quartile (low concentrations) with data points above the third quartile (high concentrations) would result in much more detailed plots. It would focuse the analysis on the extreme data points, not on the majority of data points lying in the middle. I therefore question the cluster analysis as the adequate method in this case.

Since Referee #2 and B Denzler had similar comments relating to the cluster analysis we addressed these comments together.

We agree with Referee #2 and B Denzler that a two cluster analysis results in points lying very close to the 0.904 ng/m^3 threshold. However, in this study the aim was for statistical analysis of the data to lead to certain deductions/conclusions. Therefore, it was not decided beforehand that a two grouping solution was going to be used. Interpretation of the cluster analysis resulted in identifying two clusters as the best solution with the optimum separation and representation of the data. Although a 5, 6 and 7 cluster solution was obtained with slightly better separation, some of these clusters represented only a small fraction of the total GEM distribution at Cape Point. Clusters representing smaller fractions of the GEM data are considered to be extreme cases that can be evaluated as separate case studies. Studies of this nature can be an important future perspective, which was indicated in the Conclusions as follows:

"Data indicated as extreme events, as indicated by 5, 6 and 7 cluster solutions should be investigated as special case studies."

Furthermore, the 0.904 ng/m^3 threshold is close to, but not the same as the mean and median of the data, which are 0.917 and 0.925 ng/m^3, respectively. This threshold value was statistically determined with cluster analysis and is specific for this dataset. If the dataset was only separated according to mean/median value, extreme data points would have been included in calculations. After cluster analysis data not included in the clusters (that can include extreme data points) was not considered in further processing.

We agree with B Denzler that the use of quartiles would have been a feasible alternative. However, this approach would have required a pre-classification of the data into four groups and using two of these (highest and lowest) in further processing. As mentioned, in this study the aim was for statistical analysis of the data to lead to certain deductions/conclusions and not pre-classification. In our approach cluster analysis was performed that resulted in a 2 cluster optimum solution. Cluttering was dealt with by subtracting the back trajectory analysis of each of the two clusters identified, which resulted in more detailed plots from which two distinct observations could be made, i.e. GEM values associated with ship routes and air masses of continental origin.

2. Multiple linear regression

When looking at part 3.3. I question that multiple linear regression (MLR) has been adequately applied here. Since the root mean square error (RMSE) is always decreasing with increasing variables, the choice of eight variables for the MLR comes at random. The choice of number of variables must be made according to a criterion which penalizes an MLR with many variables (expl. Akaike Information Criterion (AIC)). However, the relationship they obtain from the MLR can also be obtained by simply doing individual linear regression of the chosen parameters with the GEM measurements.

We agree with B Denzler that this was not explained adequately. The text was changed to clarify as follows:

"...if the optimum MLR solution contained more independent variables. The optimised RMSE was attained when the number of independent variables included in the optimum solution of the equation was increased to eight, and had an RMSE of 0.1205. The measure of optimisation was taken as at least 1% contribution to the overall reduction of RMSE. Table 2 indicates the identity of the independent parameters determined for each of the optimum MLR solutions."

Furthermore, MLR is a mathematical calculation where one dependent variable is related to a number of independent variables simultaneously, which is based on nearest Euclidian distances calculated for each data point entry. Therefore we consider MLR to be more advanced calculation compared to simple linear regression of individual parameters.

3. Conclusion

In Section 4 the authors present a summary of their work and an outlook, but the conclusions are missing. It is not clear what processes; ships or cities are responsible for GEM emissions.

We thank B Denzler for pointing this out. We have rewritten the Conclusions and also added future perspectives in which it is clarified why specific sources could not be identified in this statistical evaluation of GEM data. The Conclusion section was changed as follows:

"As far as the authors could assess, this is the first study that has evaluated continuous high resolution GEM data of CPT GAW with different statistical analysis techniques. Cluster analysis on the dataset indicated that the GEM data could be divided into two clusters, separated at atmospheric concentrations of 0.904 ng m^{3}. Trajectory analyses of the individual clusters, as well as the differences between these clusters, indicated that shipping around Cape Point could be a significant source of GEM. In contrast, low GEM concentrations originated from the southern oceanic background and terrestrial areas with very low anthropogenic activities/population density. Correlation of the time that back trajectories spent over the African continent and GEM concentration, proved that such analyses could be used as an alternative tool to distinguish between continental and marine GEM contributions.

It was also demonstrated that MLR analysis could be used to determine an equation that can be used to predict GEM at CPT GAW. Moreover, this equation provided some insight into the complex nature of GEM chemistry. Lastly, the evaluation of both continuously measured and calculated (with the determined MLR Eq. 1) GEM concentrations seem to indicate a decline in GEM concentrations over the period evaluated in this paper. It remains to be seen whether this decline continues, which would reflect a positive response to global Hg emission reductions, or if it is only part of a longer-term cycle with a temporary decline.

From this statistical study of continuous GEM measurement at Cape Point additional research questions and/or perspectives were identified. Data indicated as extreme events, as indicated by 5, 6 and 7 cluster solutions should be investigated as special case studies. Further research quantifying the contribution of shipping should be undertaken, not only for the southern African region, but also for other busy shipping routes. In addition, source apportionment should be conducted in order quantify the contribution of specific sources."

Minor comments

Some acronyms are not defined in the manuscript, or too late

We have rectified this matter.

US and British English is not used consistently.

We have read through the manuscript and could not find any inconsistencies relating to the UK English used in the text. However, if considered necessary, we would appreciate it if the reviewer could point out specific instances of inconsistent use of UK English.

Fig. 6, showing the GEM concentration against duration above the ocean in the same plot would be interesting

We thank B Denzler for this suggestion. However, marine air masses that contribute to most of the air masses at Cape Point are considered to be representative of the Southern Hemispheric background levels of Hg, which remains relatively constant over a 24 hour period.

on fig. 5 it would be interesting to enlarge the point of interest; South Africa

We agree that an enlargement of South Africa in this figure would be interesting. However, this image was obtained from the Automated Mutual-Assistance Vessel Rescue System (AMVER) website, sponsored by the United States Coast Guard. Enlargement of the South African region resulted in an image with poor quality. Since we do not have access to detailed data of shipping routes, it is unfortunately not possible for us to construct a figure with higher resolution.

a figure of the whole GEM series with lower resolution and plotted mean concentration would be interesting

Although a figure of this nature would be very interesting we believe that the two current figures (Fig. 3 and 8) representing GEM trends for the entire sampling period in the manuscript adequately present the complete GEM data series. Another figure of this type could be considered repetitive representation of the dataset.

on p.4037, l.27, WD (wind direction) is not mentioned as a parameter included on the MLR (eq. 1). However, eq. 1 does not include WD.(?)

We excluded the text:

"...WD the wind direction in degrees..."

from the manuscript, since it was not included as a parameter in the equation obtained with MLR. Furthermore, based on this comment we also noticed that UVb was also mentioned in the text, but not in the equation. This text:

"and UVb the ultraviolet radiation in minimum erythema dose (MED) units"

was also excluded.