

Global distribution and predictive modeling of *Vibrio vulnificus* abundance

Corresponding Author: Professor Matthias Labrenz

This file contains all editorial decision letters in order by version, followed by all author rebuttals in order by version.

Version 0:

Decision Letter:

**** Please ensure you delete the link to your author home page in this e-mail if you wish to forward it to your coauthors ****

Dear Professor Labrenz,

Your manuscript titled "Poleward spread, environmental modulation, and predictive modeling of global *Vibrio vulnificus* abundance" has now been seen by 3 reviewers, whose comments are appended below. You will see that they find your work of some potential interest. However, they have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication in its current form, but would be interested in considering a revised version that fully addresses these serious concerns. More specifically, the revised manuscript should 1) provide better data-based evidence for the poleward spread of *Vibrio vulnificus*, 2) better justify the global significance of your results (considering the limited spatial and temporal resolutions), provide additional detail regarding your predictive model, and overall, better substantiate your approaches.

We hope you will find the reviewers' comments useful as you decide how to proceed. Should additional work allow you to address these criticisms, we would be happy to look at a substantially revised manuscript. If you choose to take up this option, please either highlight all changes in the manuscript text file, or provide a list of the changes to the manuscript with your responses to the reviewers.

When resubmitting, please provide a point-by-point response to the reviewers' comments. Please submit your responses as a separate file, distinct from your cover letter where you can add responses to the Editors' comments that you do not want to be made available to the reviewers. Word files are preferred. We recommend that any figures, tables or graphs that are included in the response to reviewers are also included in the main article or Supplementary Information.

Please bear in mind that we will be reluctant to approach the reviewers again in the absence of substantial revisions.

If the revision process takes significantly longer than three months, we will be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the reviewers' comments with a list of your changes to the manuscript text (which should be in a separate document to any cover letter), a tracked-changes version of the manuscript (as a PDF file) and any completed checklist:

Link Redacted

**** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first ****

Please do not hesitate to contact us if you have any questions or would like to discuss the required revisions further. Thank you for the opportunity to review your work.

Best regards,

Annie Bourbonnais, PhD
Editorial Board Member
Communications Earth & Environment
orcid.org/0000-0001-7247-5230

Alice Drinkwater, PhD
Associate Editor
Communications Earth & Environment

EDITORIAL POLICIES AND FORMAT

If you decide to resubmit your paper, please ensure that your manuscript complies with our editorial policies and complete and upload the checklist below as a Related Manuscript file type with the revised article:

Editorial Policy Policy requirements
(Download the link to your computer as a PDF.)

For Manuscripts that fall into the following fields:

- Behavioural and social science
- Ecological, evolutionary & environmental sciences
- Life sciences

An updated and completed version of our Reporting Summary must be uploaded with the revised manuscript

You can download the form here:

<https://www.nature.com/documents/nr-reporting-summary.zip>

For your information, you can find some guidance regarding format requirements summarized on the following checklist: (<https://www.nature.com/documents/commsj-phys-style-formatting-checklist-article.pdf>) and formatting guide (<https://www.nature.com/documents/commsj-phys-style-formatting-guide-accept.pdf>).

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

In this study, the authors aim to assess the current global distribution of the potentially pathogenic bacterium *Vibrio vulnificus* and identify environmental predictors of its relative abundance. They intend to accomplish this by conducting a global reanalysis of archived 16S rRNA gene sequencing data. To predict the bacterium's abundance, they trained random forest models using either earth observation data. The study also seeks to understand how climate variability may be influencing a poleward shift in *V. vulnificus* presence, and which environmental factors, such as temperature, salinity, chlorophyll-a levels, and surface currents, are likely to contribute to this trend. The ultimate goal is to integrate these findings into predictive models to better estimate future risks associated with *V. vulnificus* in the context of climate change. However, the article has some significant limitations, which are mentioned below:

1. It is not clearly explained in the manuscript whether global prevalence data or sample data was used in this study. The manuscript emphasizes on the data around the Baltic Sea, but this information is later generalized to a global scale, which is not clearly justified or explained.
2. The information on the prediction model is not provided in sufficient detail. For example, the scale of predictive suitability for *Vibrio vulnificus* ranges from very low to medium, but the authors do not provide any quantification for this scale. What does "medium" mean in terms of suitability? Can this be quantified in terms of pathogen concentration or the number of infection cases?
3. The manuscript suggests that two years of satellite data were used to develop the near-real-time predictive model (if that is indeed the case; if not, this needs to be clarified) using machine learning techniques. However, this represents a very limited dataset that may not capture the full range of extreme events in a climatic cycle. While machine learning techniques are effective at interpolation, they often fail at extrapolation, which could mean that this model might not perform well under extreme events or outside the range of the training data.
4. Only positive dataset was used for modeling, and absence data was not incorporated. Despite being mentioned by the authors, this limitation introduces biasness into the model. I suggest that the authors change the narrative from predictive modeling to a tool that can be used under certain conditions.

Minor comment: Many abbreviations are used in the text without proper references or explanations.

Reviewer #2 (Remarks to the Author):

The authors applied an original approach to study the global distribution of *Vibrio vulnificus* (Vv) on a global scale using available data from publicly accessible repositories. The experimental design and analytical strategy are innovative and the general approach applied to the study is apparently sound. However, there are several major caveats in the work that need deep analysis and revision.

The authors put the focus of the work on the poleward spread of Vv, an aspect that has been documented in other works for some particular regions, but providing a global scale in the present work. This question is central in the study and extremely relevant, and a reference to the poleward spread has been included in the title and most of the discussion is around this point. However, this aspect has not been analysed in the study and there is not a single result showing any evidence of this trend. It has not even been subjected to analysis in the Results sections, and there is not a single piece of evidence supporting this pattern. Moreover, the poleward expansion is based on the location of individual studies, without any temporal series or the use of longitudinal analysis. Detection of Vv at a single site in a particular study at high latitudes does not substantiate the poleward expansion of Vv (or change in the ecological suitability for Vv), since it is not possible to rule out the presence of this organism in the area in previous periods. Secondly, Figure 1b shows a lower number of studies in recent years, both in the north and south hemispheres, which represents an additional constraint to reach conclusions about changes in the geographical distribution of Vv.

A second critical element in the study is the selection of the 16S rRNA gene to identify Vv at species level. The use of the 16S gene to discriminate *Vibrio* at species level has been a permanent question of debate, and many studies have shown that the sequence of 16S gene (the whole gene) does not provide enough level of discrimination to identify species within *Vibrio*. However, the authors support here the high discriminatory power of the 16S based on just a few variable sites in a short fraction of the gene few hundred base pairs long. Given the well-documented limitations of the use of this gene to assign taxonomy in *Vibrio* species, the authors should include some additional data to substantiate the approach applied in this study and probe the efficiency of the 16S gene to identify Vv and avoid wrong or imprecise taxonomic assignments (e.g., a preliminary test with different *Vibrio* species to validate the methodology).

Absence of a global taxonomic context in relation to the whole *Vibrio* group. Data of Vv presence is provided at a species level, with relative abundances of 0.0056%, which means that there are 5.6 reads of Vv every 1000 reads. According to the map provided in figure 1, positive samples for Vv were identified in the middle of the ocean, initially assumed as ecological adverse for this species (the presence of Vv in the open ocean has not been documented so far), and with a relatively high abundance according to the results displayed, which may suggest some level of misidentification. A more complete description of the sampling areas and environmental conditions for the different experiments and datasets would be necessary, where results are displayed in relation to the different ecological or environmental contexts (coastal, offshore, oceanic, ...) with data of the relative abundance of Vv in these different areas, and how the model works for these different areas. It is difficult to assume that Vv presence in the ocean or offshore waters and in estuaries are subjected to the influence of similar environmental variables, or they respond similarly to fluctuations of local environmental conditions.

Finally, the spatial resolution of the environmental data is also a constraint. According to the limited information provided in the manuscript. Line 350: "A 50 km was used, where the values were averaged over a 10x10km square surrounding the original sampling point", considering the Vv ecological preferences for waters with moderate salinity, the grid used in the study could not be enough to capture changes in salinity, which is a critical factor for Vv distribution, resulting in the significance of the temperature as the main factor (performance in the model and best predictor). No information is provided about how environmental data was processed, whether temporal trends were considered, or how it was linked to the sampling site or dataset used for the analysis.

Other comments:

Line 68. Amplicon sequencing cannot be considered a "high-throughput sequencing method" which allows for a "detailed" composition analysis of prokaryotic communities. Amplicon sequencing methods have several limitations, introduce bias through the PCR amplification, and amplification is rarely universal, with primers showing higher affinity for one species than others. This is even more relevant when the target gene is the 16S, with problems to identify species in many genera due to high similarity and limited discriminatory power.

Line 146. It is not clear the suitability of the Baltic to test the model, and no data supporting this decision is provided. Even though there has been documented a rise in the number of infections in the Baltic region, these have been related to different species, not only Vv. If in this study the authors are using a dataset with Vv cases for this region, this information needs to be incorporated into the text and full details provided.

Line 189. Temperature is a secondary element in the dynamics of Vv, and only affects those areas with salinity levels suitable for the presence of this species (low to moderate). Even though temperature is high in tropical areas, the high salinity in many areas can be a constraint to the presence of Vv. Moreover, Vv has not been reported in these areas, and infections caused by Vv have been documented in a few sites beyond the USA.

Line 199. The effect of currents on the presence of Vv is not evident as described in the text. As a coastal or estuarine organism, it would be expected that currents would have a very limited impact on the life cycle of Vv.

Line 294. "Datasets further than 50km from the coast were removed using a high-resolution coastline (10m) dataset from natural earth (V. 4.1.0)". The selection of sites 50 km away from shorelines is not a realistic approach to studying Vv, with ecological preferences for moderate salinity levels. Salinity is the major factor governing the distribution of Vv and areas outside estuaries have been considered hostile for this organism because of the high salinity prevailing in offshore waters. It is not clear in the manuscript how this critical factor has been considered in the study (and the influence of the lack of spatial resolution for salinity in cells of 10 x10 km).

The Discussion is too long, and several aspects discussed in the manuscript are rather speculative, without any data or result supporting the conclusions.

Reviewer #3 (Remarks to the Author):

This is an interesting and well written paper with appropriate diagrams. There are two key messages from the paper the first being that the 16S rRNA data is a new way to observe trends in *V. Vulnificus*. This is backed up with an analysis in which a reasonable agreement is observed between the data and satellite derived observations opening up the possibility of real-time modelling. Although just how much better is uncertain. There is also a more descriptive section of graphs which is less convincing, mostly based upon visual assessment. The second is that a poleward shift observed in 16S rRNA data. Given that they only have 10 years of data, based upon the data presented in this paper I was not convinced by this claim. Furthermore over such short time periods other environmental oscillations may be leading to any trends observed.

Introduction:

Line 63 this is all focussed on wound infections but food-borne are globally important as well

Line 65 your reference is a case report so if you are going to say "fatalities often occurring" you need another reference

Overall this section would benefit from a clear discussion of the difference between hazard, what this paper is able, and risk to human health with may or may not concur.

Results

Line 89: For this map you should highlight that absences are not analysed

Figure 1a How do you account for date of sampling, if inconsistent between samples then may confound the relationship e.g.

line 106 you have a sample from March in the UK (when sea temperatures are relatively cold)

Figure 1a. What are the white circles on this map and why are they not on other figures. If we are only looking at relative abundance then is 3 categories really enough?

Figure 1b With only 10 years of data claiming an upward trajectory needs more explanation / analysis. It is also hard to see this claim from the graph presented.

Figure legend. What range do the 1,2,3 represent.

Line 114: As I interpret these graphs this is hard to spot? i.e. larger circles at higher temperatures?

Line 142: There is no information on where the health information has come from or how it was processed to produce "normalised to official monitoring .." The Gyraite reference is not numbered.

Line 146: I can see the value of the prediction but your evaluation or predictive ability is descriptive and not quantitative? Actually the results look very similar to what the ECDC vibrio viewer says for this month.

Discussion

Line 166: Dates should be 2012-2022

Line 190: Given you only have 10 years of data and given the patterns in your data this is not a well backed up statement

Line 221: To be useful we would need more information on spatial and temporal resolution of data as well as timescales between when the data are collected and when they are available

Line 237: This is a somewhat misleading comparison as this paper is validating on a different outcome to your paper i.e. human infection data

Line 240: Accepting that your model is more complex, does it produce better predictions than say the ECDC vibrio viewer. Does more complex mean better?

Line 254: For comparison how does your work compare to the 2018 paper in Lancet Planetary health on "Future scenarios of risk of Vibrio infections in a warming planet: a global mapping study"

Line 267: I don't think you have shown comprehensively that your results align to health data. There is virtually no health data in the paper

Methods

Line 297: What is the spatial accuracy of the reed data?

Line 350: Why did you have to average the data to 10x10km squares. Why not use a finer resolution. 10km will have a large variety of environmental conditions within.

** Visit Nature Research's author and referees' website at www.nature.com/authors for information about policies, services and author benefits**

Communications Earth & Environment is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System prior to acceptance. ORCID helps

the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the Manuscript Tracking System by clicking on 'Modify my Springer Nature account' and following the instructions in the link below. Please also inform all co-authors that they can add their ORCIDs to their accounts and that they must do so prior to acceptance.

<https://www.springernature.com/gp/researchers/orcid/orcid-for-nature-research>

For more information please visit <http://www.springernature.com/orcid>

If you experience problems in linking your ORCID, please contact the [Platform Support Helpdesk](http://platformsupport.nature.com/).

Version 1:

Decision Letter:

**** Please ensure you delete the link to your author home page in this e-mail if you wish to forward it to your coauthors ****

Dear Professor Labrenz,

Please accept our apologies for the delay in sending a decision on your manuscript. Your manuscript titled "Global distribution and predictive modeling of *Vibrio vulnificus* abundance" has now been seen by our reviewers, whose comments appear below. In light of their advice we are delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment, provided you fully acknowledge the limitations of your dataset with regards to capturing climatic cycles and extreme events.

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers, which pertain to acknowledging data limitations. At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at <https://www.nature.com/documents/commsj-file-checklist.pdf>.

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication. For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit <https://www.nature.com/commsenv/open-access>

At acceptance, you will be provided with instructions for completing the open access licence agreement on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:

Link Redacted

**** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first ****

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Alice Drinkwater, PhD
Associate Editor
Communications Earth & Environment

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

In the revised manuscript, the authors have not adequately addressed the concerns raised in the initial review. Specifically:

1) To effectively capture the impact of climate change, the analysis should incorporate at least 30 years of data. Climatic cycles typically span periods longer than 30 years, and using a dataset shorter than this may result in false trends or misleading conclusions.

2) The manuscript does not account for extreme events, which are critical for interpolation and model robustness. Without incorporating extreme events, particularly when using a dataset spanning only 10–11 years, the model risks poor extrapolation performance. The likelihood that extreme events are underrepresented in the current dataset is a significant limitation.

These two issues represent major concerns that need to be addressed.

Reviewer #2 (Remarks to the Author):

All the comments and concerns have been addressed and authors have provided additional information and amend the text in response to comments from reviewers.

Reviewer #3 (Remarks to the Author):

This paper has progressed nicely and is now suitable for publication in my opinion.

** Visit Nature Research's author and referees' website at <http://www.nature.com/authors> for information about policies, services and author benefits**

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

REVIEWER COMMENTS:

We would like to thank the reviewers for the opportunity of revision, and for their comments, which improved the manuscript a lot. Please find our comments and changes in relation to the individual points of criticism marked in bold.

Reviewer #1 (Remarks to the Author):

In this study, the authors aim to assess the current global distribution of the potentially pathogenic bacterium *Vibrio vulnificus* and identify environmental predictors of its relative abundance. They intend to accomplish this by conducting a global reanalysis of archived 16S rRNA gene sequencing data. To predict the bacterium's abundance, they trained random forest models using either earth observation data. The study also seeks to understand how climate variability may be influencing a poleward shift in *V. vulnificus* presence, and which environmental factors, such as temperature, salinity, chlorophyll-a levels, and surface currents, are likely to contribute to this trend. The ultimate goal is to integrate these findings into predictive models to better estimate future risks associated with *V. vulnificus* in the context of climate change. However, the article has some significant limitations, which are mentioned below:

1. It is not clearly explained in the manuscript whether global prevalence data or sample data was used in this study. The manuscript emphasizes on the data around the Baltic Sea, but this information is later generalized to a global scale, which is not clearly justified or explained.

Thank you for this comment. All models are trained on global datasets, with the Baltic Sea only serving as an example use case due to its relatively large amount of infection data and wide range of salinity levels. The performance on global data is indicated in Fig. 2.

We emphasize this global prevalence in:

- **Line 129 – 130: “Performance and feature importance of the global *V. vulnificus* model based on remote sensing data (a + b) and on the prokaryotic community (c + d).”**

and now emphasize this exemplary role by specifying it in:

- **Line 151 – 152: “Using satellite data from August 2021, the global model shown in Figure 2a+b was employed, as an example, for the Baltic Sea (Fig. 3)”**

2. The information on the prediction model is not provided in sufficient detail. For example, the scale of predictive suitability for *Vibrio vulnificus* ranges from very low to medium, but the authors do not provide any quantification for this scale. What does "medium" mean in terms of suitability? Can this be quantified in terms of pathogen concentration or the number of infection cases?

That's a good point. The goal was to make it understandable for readers who may not be familiar with the typical relative abundances of *Vibrio vulnificus* in natural environments. While 0.3% of the population might be considered high for some bacteria, it is fairly common for others. While we had already included the actual figures in the methods section, we aimed to maintain this clarity while providing the more detailed data as requested in the figure legend:

- Line 148 – 150: “Very low suitability is defined as 0.0014 %, low suitability as 0.002%, and medium suitability as 0.0057% of the bacterial community predicted to be *V. vulnificus*”.

Defining it in infection cases is unfortunately not possible. Our model predicts relative abundances, and cannot be 1:1 equated to infection cases.

3. The manuscript suggests that two years of satellite data were used to develop the near-real-time predictive model (if that is indeed the case; if not, this needs to be clarified) using machine learning techniques. However, this represents a very limited dataset that may not capture the full range of extreme events in a climatic cycle. While machine learning techniques are effective at interpolation, they often fail at extrapolation, which could mean that this model might not perform well under extreme events or outside the range of the training data.

The model based on remote sensing data is trained on all available data from 2012-2022 (provided in Figure 1). Thus, we are not sure where the reference to 'two years' comes from.

By incorporating global data, including from the tropics and the northern Baltic Sea, the model captures a wide range of climatic conditions. This has been clarified now in

- Line 74 –77: “This approach also enables the determination of the relative abundance of *V. vulnificus* within the overall bacterial population and assesses the extent of its spread and co-occurrence with other species at an increased spatiotemporal scale across a wide range of environmental conditions—from tropical regions to latitudes as high as the northern Baltic Sea”.

4. Only positive dataset was used for modeling, and absence data was not incorporated. Despite being mentioned by the authors, this limitation introduces biasness into the model. I suggest that the authors change the narrative from predictive modeling to a tool that can be used under certain conditions.

We have described this limitation more explicit by including it more clearly in the results section now and in the discussion explicitly refer to high risk regions, as the base conditions for the presence of *V. vulnificus* need to be met to achieve the performance stated in Fig. 2:

- Line 135 - 137: “The model based on the remote sensing data explained almost 50% of the variance in samples where *V. vulnificus* was detected, both on the test dataset and using 10-fold cross-validation (Fig. 2a, SI Table 1)”.
- Line 231 -233: “The comparison of predicted environmental suitability for *V. vulnificus* in the Baltic Sea, averaged over August—the month with the highest infection risk—with reported infection cases demonstrates a possible use case for high risk regions”

Minor comment: Many abbreviations are used in the text without proper references or explanations.

We have amended the abbreviations: SST (line 110), chl-a (line 114) and have removed the abbreviation SRA by replacing it with sequence read archive

Reviewer #2 (Remarks to the Author):

The authors applied an original approach to study the global distribution of *Vibrio vulnificus* (Vv) on a global scale using available data from publicly accessible repositories. The experimental design and

analytical strategy are innovative and the general approach applied to the study is apparently sound. However, there are several major caveats in the work that need deep analysis and revision.

The authors put the focus of the work on the poleward spread of Vv, an aspect that has been documented in other works for some particular regions, but providing a global scale in the present work. This question is central in the study and extremely relevant, and a reference to the poleward spread has been included in the title and most of the discussion is around this point. However, this aspect has not been analysed in the study and there is not a single result showing any evidence of this trend. It has not even been subjected to analysis in the Results sections, and there is not a single piece of evidence supporting this pattern. Moreover, the poleward expansion is based on the location of individual studies, without any temporal series or the use of longitudinal analysis. Detection of Vv at a single site in a particular study at high latitudes does not substantiate the poleward expansion of Vv (or change in the ecological suitability for Vv), since it is not possible to rule out the presence of this organism in the area in previous periods. Secondly, Figure 1b shows a lower number of studies in recent years, both in the north and south hemispheres, which represents an additional constraint to reach conclusions about changes in the geographical distribution of Vv.

As already mentioned above, we re-evaluated our data and agree with the reviewer that the current data situation worldwide is insufficient to really prove that poleward spread of *V. vulnificus*. Thus, this statement is principally deleted and we have amended the manuscript accordingly. We have made the following changes in the text:

- **Title: Has been changed to “Global distribution and predictive modeling of *Vibrio vulnificus* abundance”**
- **Line 96 –97: Datasets where *V. vulnificus* was detected were found increasingly northward every year between 2013 and 2021 except in 2015 and 2020.**
- **Line 185 – 187 “Indications for the poleward expansion of *V. vulnificus* are visible both on the northern and southern hemisphere on a limited temporal scale. This is provided by environmental data reflecting the predicted increase in *V. vulnificus* occurrence in high-latitude regions, for example in the northern Baltic Sea”**
- **Line 188 – 192: “Sampling bias might contribute to the possible observed poleward shift in *V. vulnificus* detection, as the bacterium may have been present already in areas where it was detected or in more polar regions that went unsampled in earlier years. However, the substantial number of analyzed datasets where *V. vulnificus* was absent in polar regions each year compared to the northernmost positive detection (Fig. 1a & Suppl. Fig. 2) suggests a possible shift”**

We aim to more openly highlight the limitations of this approach. However, as shown in Suppl. Fig. 2, there are many samples collected each year from locations farther north than the northernmost positive sample, which were negative, suggesting that sampling bias is unlikely responsible for the observed patterns. Therefore, we maintain that, while we can conclusively show a northward spread on this temporal scale, there is at least an indication of a trend that aligns with the observed infection rates.

A second critical element in the study is the selection of the 16S rRNA gene to identify Vv at species level. The use of the 16S gene to discriminate *Vibrio* at species level has been a permanent question of debate, and many studies have shown that the sequence of 16S gene (the whole gene) does not provide enough level of discrimination to identify species within *Vibrio*. However, the authors support here the high discriminatory power of the 16S based on just a few variable sites in a short fraction of the gen few hundred base pairs long. Given the well-documented limitations of the use of this gene to assign taxonomy in *Vibrio* species, the authors should include some additional data to substantiate the approach applied in this study and probe the efficiency of the 16S gene to identify Vv and avoid

wrong or imprecise taxonomic assignments (e.g., a preliminary test with different *Vibrio* species to validate the methodology).

We have performed a preliminary test with the closest relatives of *V. vulnificus*, which we now describe in detail in

- **Line 279 – 286 : “Using ARB (V. 7.0)⁴⁶, positions were identified where *V. vulnificus* could be differentiated from its closest relatives, such as *V. pommerensis* and *V. navarrensis*. This was achieved by comparing *V. vulnificus* type strains to known *V. pommerensis* and *V. navarrensis* sequences and type strains of sister branches (*Vibrio thalassae*, *Vibrio madracius*, *Vibrio mediterranei*, and *Vibrio shilonii*). Position 10266 within the SILVA alignment was identified as the most divergent between *V. vulnificus* and the other sequences. Position 15968 in the SILVA alignment further allows for the distinction between *V. vulnificus* and *V. pommerensis* and *V. navarrensis* only while 16421 to 16427 allow for the differentiation between *V. vulnificus* and type strains of the sister branches mentioned above.”**

Absence of a global taxonomic context in relation to the whole *Vibrio* group. Data of Vv presence is provided at a species level, with relative abundances of 0.0056%, which means that there are 5.6 reads of Vv every 1000 reads. According to the map provided in figure 1, positive samples for Vv were identified in the middle of the ocean, initially assumed as ecological adverse for this species (the presence of Vv in the open ocean has not been documented so far), and with a relatively high abundance according to the results displayed, which may suggest some level of misidentification. A more complete description of the sampling areas and environmental conditions for the different experiments and datasets would be necessary, where results are displayed in relation to the different ecological or environmental contexts (coastal, offshore, oceanic, ...) with data of the relative abundance of Vv in these different areas, and how the model works for these different areas. It is difficult to assume that Vv presence in the ocean or offshore waters and in estuaries are subjected to the influence of similar environmental variables, or they respond similarly to fluctuations of local environmental conditions.

We are sorry, obviously the resolution of Figure 1 led to confusion. In fact, the stations that appear to be in the middle of the ocean are actually located near the shorelines of French Polynesian islands, the Galapagos, Hawaii, and other small islands. There are no oceanic stations in the processed datasets, as outlined in the methods and shown in the provided code. Consequently, the environmental context for all datasets is either coastal, estuarine or brackish marginal sea. To clarify this misconception we have added

- **Line 88 – 89: “The analyzed datasets, all within 50 km of a coast, ranged in latitude from 78° S to 83.1° N and spanned to globe laterally.”**

Misidentification is minimized by applying stringent cutoffs and is detailed in the methods section

- **Line 299 – 301: “Only zOTUs with either 100% query coverage and similarity to the custom database or those with > 99.5% similarity and no other species hit in SILVA were classified as *V. vulnificus*”.**

While we cannot entirely rule out the possibility that some reads originate from a close relative of *V. vulnificus* and contain a sequencing error at one of the key sites identified in the ARB analysis, leading to misclassification as *V. vulnificus*, this is unlikely and would have minimal impact on the overall results.

Finally, the spatial resolution of the environmental data is also a constraint. According to the limited information provided in the manuscript. Line 350: "A 50 km was used, where the values were averaged over a 10x10km square surrounding the original sampling point", considering the Vv ecological preferences for waters with moderate salinity, the grid used in the study could not be enough to capture changes in salinity, which is a critical factor for Vv distribution, resulting in the significance of the temperature as the main factor (performance in the model and best predictor). No information is provided about how environmental data was processed, whether temporal trends were considered, or how it was linked to the sampling site or dataset used for the analysis.

We recognize the limitation of the grid based approach, this was however the only feasible way to acquire data for all 1208 datasets for which a sampling date was known. This increase in data volume was necessary for the machine learning approaches.

We are mentioning this limitation explicitly in the discussion:

- Line 223 –230: “Despite the model performance, issues with using remote sensing data to predict *V. vulnificus* remain. Due to a preference for brackish waters, *V. vulnificus* are often found in (small) estuaries where conditions can rapidly change with the tides, making daily remote sensing measurement imprecise. Additionally, the applied bounding box method, while necessary to ensure consistent input data across all datasets, can result in the inclusion of satellite data from outside the estuary. However, the extensive amount of data available from archived and satellite sources helps to compensate for the increased noise in the dataset.”

No further processing of the satellite data was performed, besides what is now indicated more clearly in the methods

- Line 316 – 317 : “A bounding box method was applied by averaging satellite-derived values found within a 10x10 km square surrounding the original sampling point”

Other comments:

Line 68. Amplicon sequencing cannot be considered a "high-throughput sequencing method" which allows for a "detailed" composition analysis of prokaryotic communities. Amplicon sequencing methods have several limitations, introduce bias through the PCR amplification, and amplification is rarely universal, with primers showing higher affinity for one species than others. This is even more relevant when the target gene is the 16S, with problems to identify species in many genera due to high similarity and limited discriminatory power.

These terms have been removed.

Line 146. It is not clear the suitability of the Baltic to test the model, and no data supporting this decision is provided. Even though there has been documented a rise in the number of infections in the Baltic region, these have been related to different species, not only Vv. If in this study the authors are using a dataset with Vv cases for this region, this information needs to be incorporated into the text and full details provided.

We have explained our choice for the Baltic Sea in more detail and specified that we are only using *V. vulnificus* infections in the text:

- Lines 151 – 156: “Using satellite data from August 2021, the global model shown in Figure 2a+b was employed, as an example, for the Baltic Sea (Fig. 3). The Baltic Sea provides excellent environmental conditions for *V. vulnificus* in summer, over the full**

known salinity range of *V. vulnificus* and has relatively comprehensive *V. vulnificus* infection data sets available for Sweden and Germany and to a lesser extent other Baltic Sea countries (Gyraite et al. 2024). The 85 recorded infection cases can thus be used to indicate possible utility of our model for infection risk prediction for a wide range of environmental conditions”.

We have provided the infections per year and now also the total number of infections used to test this model, where and how these infection cases have occurred and have been recorded has this has been thoroughly documented by Gyraite et al.

Line 189. Temperature is a secondary element in the dynamics of Vv, and only affects those areas with salinity levels suitable for the presence of this species (low to moderate). Even though temperature is high in tropical areas, the high salinity in many areas can be a constraint to the presence of Vv. Moreover, Vv has not been reported in these areas, and infections caused by Vv have been documented in a few sites beyond the USA.

We improved the probable explanation why we find very high *V. vulnificus* concentrations at many tropical areas and added a possible explanation why very few infections are known from these regions:

- Lines 181 – 184: “This explains why the majority of datasets with high relative abundance originated from the tropics, where runoff, estuaries or sewage can provide suitable salinities locally combined with very high temperatures. Known infections in this region are possibly low due to lower reporting standards compared to areas like the US and the Baltic Sea region.”

Line 199. The effect of currents on the presence of Vv is not evident as described in the text. As a coastal or estuarine organism, it would be expected that currents would have a very limited impact on the life cycle of Vv.

You are right, out of all predictors, current speed has the lowest predictive power. This and a possible reason it might have an impact are addressed in:

- Line 200 – 205. “This would also explain why lower current speeds, while having limited importance in the model, might favor a higher relative abundance of *V. vulnificus*. The more stagnant waters are usually more stratified and, consequently, they foster larger surface blooms which after depletion of the local inorganic nutrients, decompose. The organic nutrients, presumably important for the rapid growth of *V. vulnificus*, subsequently lead to an increase in the relative abundance.”

Line 294. “Datasets further than 50km from the coast were removed using a high-resolution coastline (10m) dataset from natural earth (V. 4.1.0)”. The selection of sites 50 km away from shorelines is not a realistic approach to studying Vv, with ecological preferences for moderate salinity levels. Salinity is the major factor governing the distribution of Vv and areas outside estuaries have been considered hostile for this organism because of the high salinity prevailing in offshore waters. It is not clear in the manuscript how this critical factor has been considered in the study (and the influence of the lack of spatial resolution for salinity in cells of 10 x10 km).

The choice of a 50 km range allows for the detection of *V. vulnificus* in marginal seas further from the coast, where salinity conditions remain suitable, such as the Baltic Sea. As such, the purpose of the filter is one of excluding data set that are unlikely to contain *V. vulnificus*,

while retaining all that have the possibility of retaining *V. vulnificus*. The filter was supposed to have a high sensitivity at the cost of a low specificity. It is purely for data volume reduction. Actual *V. vulnificus* detection is then based on the analysis of each individual data set. If environmental conditions are unfavorable for *V. vulnificus*, it will not be detected in the sequencing data. We acknowledge the limitations of the 10 x 10 km grid and the impact of tidal influence on the salinity of estuaries. While the resolution is constrained by the nature of remote sensing, it remains the most effective option currently available for rapid, large-scale global predictions.

The Discussion is too long, and several aspects discussed in the manuscript are rather speculative, without any data or result supporting the conclusions.

The discussion is shortened and majorly rewritten based on your and other reviewer's comments.

Reviewer #3 (Remarks to the Author):

This is an interesting and well written paper with appropriate diagrams. There are two key messages from the paper the first being that the 16S rRNA data is a new way to observe trends in *V. Vulnificus*. This is backed up with an analysis in which a reasonable agreement is observed between the data and satellite derived observations opening up the possibility of real-time modelling. Although just how much better is uncertain. There is also a more descriptive section of graphs which is less convincing, mostly based upon visual assessment. The second is that a poleward shift observed in 16S rRNA data. Given that they only have 10 years of data, based upon the data presented in this paper I was not convinced by this claim. Furthermore over such short time periods other environmental oscillations may be leading to any trends observed.

Thank you for your comments. In response we have removed the focus on the perceived poleward spread by removing it from the title and changing the discussion accordingly (please see also comments above). In addition to above mentioned changes we are just discussing an indication for a poleward trend on a limited temporal scale:

- **Line 185 – 187: “Indications for the poleward expansion of *V. vulnificus* are visible both on the northern and southern hemisphere on a limited temporal scale. This is provided by environmental data reflecting the predicted increase in *V. vulnificus* occurrence in high-latitude regions, for example in the northern Baltic Sea.”**

Introduction:

Line 63 this is all focussed on wound infections but food-borne are globally important as well

We rewrote it accordingly:

- **Line 61 – 67: “Climate change will likely increase both the geographic range and prolong the growth period of *V. vulnificus*, leading to an increase in both skin and food-borne infections in temperate regions at higher latitudes. Foodborne infections can result in symptoms such as fever, chills, hypotensive septic shock, and may ultimately lead to death, while skin infections can swiftly progress to necrosis, necessitating urgent surgical tissue removal or limb amputation in approximately 10 % of cases and has mortality rates as high as 18 %. Fatalities may occur within 48 hours of exposure”**

Line 65 your reference is a case report so if you are going to say "fatalities often occurring" you need another reference

Rewrote to:

- **Line 66 – 67: “Fatalities may occur within 48 hours of exposure¹².”**

Overall this section would benefit from a clear discussion of the difference between hazard, what this paper is able, and risk to human health with may or may not concur.

We clarified this:

- **Line 70 – 71: “While not all *V. vulnificus* are pathogenic, there is a clear relationship between overall abundance and pathogenic strains” now indicating that while we are not predicting pathogens, the prediction is still relevant.**

Which combined with

- **Line 81 –83 “This allowed the modelling of relative *V. vulnificus* abundances based on environmental factors at the global scale, the localization of global hotspots, and the observation of shifts in spatial distribution over time”**

indicates what we are modelling and what the associated potential health hazard is.

Results

Line 89: For this map you should highlight that absences are not analysed

Clarified:

line 123 – 124 : “Worldwide spatial distribution and relative abundance over time, with white dots representing all 70,640 initially selected coastal data points”

in combination with the methods

line 262 – 263 : “In total 70,640 runs, for which Fastq files were retrieved, were retained for initial analysis. These are referred to as datasets.”

Figure 1a How do you account for date of sampling, if inconsistent between samples then may confound the relationship e.g. line 106 you have a sample from March in the UK (when sea temperatures are relatively cold)

The sampling dates are inconsistent, as we utilized all archived 16S rRNA amplicon sequencing data that met the criteria outlined in the methods and code in the repository as a starting point for our analysis. However, this inconsistency is not relevant for assessing the global range of *V. vulnificus* detections in archived data. In fact, it allows for a broader representation of environmental conditions under which *V. vulnificus* was present, enhancing our subsequent modeling efforts. Within these models, sampling day was not used as one of the features and can thus only be derived by the model from environmental conditions.

Figure 1a. What are the while circles on this map and why are they not on other figures. If we are only looking at relative abundance then is 3 categories really enough?

Line 123 – 124: “Worldwide spatial distribution and relative abundance over time, with white dots representing all 70,640 initially selected coastal data points”, they do not appear in the other plots as the either amplified the wrong region or did not contain *V. vulnificus* according to our methodology.

It's a continuous scale, this is now indicated:

Line 121 - 122: “The size of the dots, as a continuous scale, represents % *V. vulnificus* of the sequences bacterial community”

Figure 1b With only 10 years of data claiming an upward trajectory needs more explanation / analysis. It is also hard to see this claim from the graph presented.

Thank you for the comment. Upon revision we agree that the data is insufficient for such a claim and have restructured the results and discussion accordingly.

Figure legend. What range do the 1,2,3 represent.

We have clarified it in the figure legend:

- **Line 121 – 122:** “The size of the dots, as a continuous scale, represents % *V. vulnificus* of the total bacterial community as assessed by 16S rRNA gene sequencing”

Line 114: As I interpret these graphs this is hard to spot? i.e. larger circles at higher temperatures?

Yes, we have removed the claim about the elevated chl_a concentrations being related to higher *V. vulnificus* abundances at temperatures between 20-25 C as it is a purely visual observation with no statistical analysis behind it. We maintain that there is a general relationship based on the ML models.

Line 142: There is no information on where the health information has come from or how it was processed to produce "normalised to official monitoring .. " The Gyraite reference is not numbered.

The paper has now been published, numbered, and is referenced accordingly. We have also clarified our data processing methodology:

- **Line 146 – 148:** “Normalized by dividing the total recorded cases by the official monitoring years, resulting in a per-year figure to allow for country-to-country comparisons”

Line 146: I can see the value of the prediction but your evaluation or predictive ability is descriptive and not quantitative? Actually the results look very similar to what the ECDC vibrio viewer says for this month.

It looks very similar because temperature and salinity remain two of the dominant predictors, so we wouldn't expect significant deviations from the ECDC Vibrio Viewer. The model's quantitative performance is shown in Figure 2a, as it is the same model being used. Since the model predicts abundance, not infection cases, we cannot quantitatively validate with the recorded infections. A separate quantification for the Baltic Sea, is also not an option as many of the datapoints from the Baltic sea were used in training, leaving too small of a test dataset to be meaningful.

Discussion

Line 166: Dates should be 2012-2022

This section has been completely reworked

Line 190: Given you only have 10 years of data and given the patterns in your data this is not a well backed up statement

We have removed that part and interpret our results later in the discussion as both models indicate that blooms may stimulate *V. vulnificus*:

- **Line 200 – 203: “This would also explain why lower current speeds, while having limited importance in the model, might favor a higher relative abundance of *V. vulnificus*. The more stagnant waters are usually more stratified and consequently, they foster larger surface blooms. “**

Line 221: To be useful we would need more information on spatial and temporal resolution of data as well as timescales between when the data are collected and when they are available

Unfortunately, we are not quite sure what the reviewer is trying to say with this statement. We have made a table with date, ENA accessions and latitude and longitude available on the repository under: `accession_date_location_pos_samples.csv`. Spatio-temporal resolutions are visualized in Fig. 1.

Line 237: This is a somewhat misleading comparison as this paper is validating on a different outcome to your paper i.e. human infection data

We have removed the comparison to the Baker-Austin paper.

Line 240: Accepting that your model is more complex, does it produce better predictions than say the ECDC vibrio viewer. Does more complex mean better?

More complex is generally a negative term in modelling, but given the growing evidence that primary production influences *V. vulnificus* dynamics (Brumfield et al., 2023a, Brumfield et al., 2023b), this factor should be considered when modelling *V. vulnificus* occurrence to accurately capture the dynamics in the Baltic Sea, for example. However, a direct comparison between the models is challenging because they predict different outcomes. While chlorophyll-a (chl-a) may positively affect *V. vulnificus* concentrations, it could also reduce the number of swimmers at risk of infection for example. This would limit the usefulness of chl-a in infection modeling, although it proves valuable in predicting relative abundance, as we have done. To avoid confusion, we have removed the comparison with the Baker-Austin / ECDC Vibrio Viewer model.

Brumfield, Kyle D., et al. (a) "Genomic diversity of *Vibrio* spp. and metagenomic analysis of pathogens in Florida Gulf coastal waters following Hurricane Ian." *MBio* 14.6 (2023): e01476-23.

b Brumfield, Kyle D., et al. (b) "Environmental factors influencing occurrence of *Vibrio parahaemolyticus* and *Vibrio vulnificus*." *Applied and environmental microbiology* 89.6 (2023): e00307-23.

Line 254: For comparison how does your work compare to the 2018 paper in Lancet Planetary health on "Future scenarios of risk of Vibrio infections in a warming planet: a global mapping study"

Thanks a lot, we have referenced the paper now:

- Line 175 – 178: **"An increase in infections at these hotspots has been noted on the US coast, the Baltic Sea, and the East Asian coast. An aging population in the Baltic Sea region, East Asia, and the USA will exacerbate these infection rates, aligning with previous at-risk predictions."**
- And line 185 – 186: **"This provides environmental data that reflects the predicted increase in *V. vulnificus* occurrence in high-latitude regions, for example in the northern Baltic Sea"**

Line 267: I don't think you have shown comprehensively that your results align to health data. There is virtually no health data in the paper

The patterns in our prediction for the Baltic Sea align with the recorded infections in the Baltic Sea. We have rewritten the final paragraph of the discussion to illustrate this better:

Line 230 – 240: The comparison of predicted environmental suitability for *V. vulnificus* in the Baltic Sea, averaged over August—the month with the highest infection risk—with reported infection cases demonstrates a possible use case for high risk regions. Assuming SSS³⁸, SST³⁹, phytoplankton dynamics, and current speed⁴⁰ patterns are relatively regular, a similar risk profile for the Baltic Sea can be assumed each summer. The risk profile predicted by the model aligns with the recorded infections per year. Monitoring efforts remain incomplete in numerous countries and are entirely absent in significant portions of the eastern Baltic Sea region. Gyraite et al., 2024¹⁷ predicted unreported cases in Lithuania and Poland over the last three decades, identifying regions that align with suitable conditions according to our model. Additional validation is advisable using the increasing number of infection cases with a known origin as they become available. The similarity in regional patterns does, however, indicate that the proposed model may have a role in infection prediction as well as *V. vulnificus* environmental suitability prediction.

-

Methods

Line 297: What is the spatial accuracy of the reed data?

Unfortunately, we are unsure about the point the reviewer is making with this statement. The spatial accuracy of our data is as good as the lat / lon coordinates that have been supplied by other researchers when archiving their reed data. This can be found in the repository under: : accession_date_location_pos_samples.csv

Line 350: Why did you have to average the data to 10x10km squares. Why not use a finer resolution. 10km will have a large variety of environmental conditions within.

This was the minimum resolution at which we had valid input data for all datasets containing *V. vulnificus*. While this approach increases noise, it maximizes the amount of training data available. If someone else wishes to run this model in a coastal region, they may not obtain a

prediction without the 10x10 square method, as certain input parameters, such as current, might not be available at a higher resolution in their region. By using this approach, the model can be applied globally with the indicated performance.

Line 226 – 227: “Additionally, the applied bounding box method, while necessary to ensure consistent input data across all datasets, can result in the inclusion of satellite data from outside the estuary.”

We would like to thank the editor and the three reviewers for their reevaluation of our manuscript. Below are the specific responses to the remaining critiques.

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

In the revised manuscript, the authors have not adequately addressed the concerns raised in the initial review. Specifically:

1) To effectively capture the impact of climate change, the analysis should incorporate at least 30 years of data. Climatic cycles typically span periods longer than 30 years, and using a dataset shorter than this may result in false trends or misleading conclusions.

2) The manuscript does not account for extreme events, which are critical for interpolation and model robustness. Without incorporating extreme events, particularly when using a dataset spanning only 10–11 years, the model risks poor extrapolation performance. The likelihood that extreme events are underrepresented in the current dataset is a significant limitation.

These two issues represent major concerns that need to be addressed.

Response: We thank the reviewer for their critique of the modeling approach. As is often the case, the quality and availability of input data can pose a significant bottleneck to effective modeling. We acknowledge the limitations of the chosen approach; however, due to the relatively recent development and adoption of 16S rRNA amplicon sequencing, the dataset available for global *V. vulnificus* occurrence is inherently limited in its temporal scope.

- 1) To address this, we have further emphasized the problems of the timespan covered by data versus the length of climatic cycles in

Lines 167-168: “Indications for the poleward expansion of *V. vulnificus*²⁵ are visible both on the northern and southern hemisphere on a limited, compared to climatic cycles, temporal scale.”

And

Lines 174 – 175: “. This possible shift cannot be directly tied to climate change due to the limited timespan of the available data.”

We have removed “due to climate change” to further reduce any implication that our data directly indicates a climate change signal in:

Lines 198 – 199:” Following the findings of our models, it is reasonable to expect that the poleward trend of *V. vulnificus* ~~due to climate change~~ will be exacerbated by predicted increases in (harmful) algal blooms at higher latitudes^{36,37”}

Extreme events play a significant role in shaping the distribution of microbial communities and their representation in the dataset is limited. However, the dataset’s broad geographical

coverage provides insights into a variety of conditions that constitute extreme weather events for different regions. For instance, tropical temperatures represent extreme heatwaves in marginal seas such as the Baltic Sea, whereas Baltic Sea temperatures in summer are equivalent to extreme cold for tropical regions. Consequently, the dataset, encompassing a temperature range of 5.4°C to 32°C with *V. vulnificus* detected, and sea surface salinity ranging from 4.33 to 39.80, already captures extreme conditions for many localities.

We have made this more explicit in

Lines 212-215: “The broad geographical distribution of the dataset allows the model to also approximate the effects of extreme weather events for many locations. For instance, temperatures in the tropics represent extreme heatwaves for marginal seas like the Baltic Sea, while Baltic Sea temperatures are equivalent to extreme cold for tropical regions.”

Reviewer #2 (Remarks to the Author):

All the comments and concerns have been addressed and authors have provided additional information and amend the text in response to comments from reviewers.

Response: Thank you for the previous helpful comments

Reviewer #3 (Remarks to the Author):

This paper has progressed nicely and is now suitable for publication in my opinion.

Response: Thank you for the response

IN ADDITION:

In response to the editorial request table to abstract has been modified to:

Lines 25-35: “The proliferation of the potentially pathogenic bacterium *Vibrio vulnificus* due to climate change poses a growing public health threat worldwide. Here we present a global reanalysis of archived 16S ribosomal RNA gene sequencing data, suitable for the identification of bacteria and archaea, to evaluate its current distribution and determine predictors of relative abundance. Two random forest models were trained to predict relative abundance using either satellite data or the prokaryotic community. The models indicated that, besides temperature and salinity, both chlorophyll-a and surface currents likely played a role in relative *V. vulnificus* abundance and that non-vulnificus *Vibrio* spp. and *Pseudoaltermonas* spp. were bacterial predictors of *V. vulnificus*. Collectively, this indicates decaying phytoplankton blooms stimulate *V. vulnificus* relative abundance. A poleward trend could accelerate due to climate change-induced modifications of phytoplankton dynamics and currents; factors that should be integrated into predictive modelling to estimate future risks.”

