Ecological Indicators 111 (2020) 105989

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Original Articles

Ecoregional or site-specific lake nutrient criteria? Evidence from ecological fallacy

Zhongyao Liang[a], Feifei Dong[a], Song S. Qian[b], Yong Liu[a,*], Huili Chen[a], Wentao Lu[c]

[a] College of Environmental Sciences and Engineering, State Environmental Protection Key Laboratory of All Materials Flux in Rivers, Peking University, Beijing 100871, China

[b] Department of Environmental Sciences, The University of Toledo, Toledo, OH 43606-3390, USA

[c] Chinese Academy for Environmental Planning, Beijing 100012, China

ARTICLE INFO

Keywords:
Nutrient criteria
Site-specific
Ecoregion
Ecological fallacy
Bayesian hierarchical model

ABSTRACT

Ecoregional nutrient criteria are widely used but their validity has rarely been verified competed with site-specific criteria. In this study, we introduced ecological fallacy, which describes phenomenon that site-specific stressor-response relationships cannot be deducted from ecoregional relationship and vice versa, to explore the spatial scale of nutrient criteria. A long-term and nationwide water quality dataset of lakes and reservoirs was used to determine if ecological fallacy existed or not. Ecoregional and site-specific nutrient-Chlorophyll *a* relationships were built employing Bayesian linear model and hierarchical model, respectively. By comparing ecoregional and site-specific relationships, we found that ecological fallacy existed in each ecoregion. Ecoregional relationship may misidentify limiting nutrient or miscalculate the nutrient effect direction or magnitude. We found huge differences between estimated Chlorophyll *a* concentrations deduced from regional and site-specific relationships conditioning average TN or TP concentrations. Based on these results, we determined that lake nutrient criteria should be site-specific, primarily to avoid ecological fallacy rather than to improve their accuracy. These findings could guide the future nutrient criteria development. We further recommended partial pooling of data to develop stressor-response relationship facing with intensive environmental and ecological data.

1. Introduction

Nutrient criteria provide foundations for lake water quality management (USEPA, 1998). Proper spatial scale is essential to make sure the effectiveness of nutrient criteria (Soranno et al., 2008). In the past two decades, ecoregional scale is the most widely used (Huo et al., 2014, USEPA, 1998). An ecoregional nutrient criterion is applicable for all the lakes within the ecoregion and one implicit assumption is the homogeneity within the ecoregion. Under that assumption, space-for-time substitution, which has been comprehensively applied in data-poor ecosystems (Celentano and Defeo, 2016, Lester et al., 2014), is applicable for ecoregional nutrient criteria development. That is, data from lakes within the same ecoregion are aggregated together to develop the unique ecoregional nutrient criterion, solving the data-poor problem. Stressor-response model is one of the most widely used methods to develop nutrient criteria (USEPA, 2010). Nutrient is treated as the stressor and nutrient criteria are deduced based on the stressor-response relationship by keeping the response (management endpoint, e.g. Chlorophyll *a* (Chla) concentration) at a certain level.

Recently, some ecological researches revealed that there might be a big bias between regional and site-specific stressor-response relationship. For example, Cha et al. (2016) found that in Nakdong River, while Chla and flow were positively correlated at each site, the aggregated data showed negative correlations between them. Results in Martay et al. (2016) indicated that aggregating data spatially overestimated the regression slope of temperature on butterfly Community Temperature Index.

Above cases could be explained by ecological fallacy (Robinson, 1950), an item describing the phenomenon that individual-level relationships cannot be deducted from the group-level relationship and vice versa (Genser et al., 2015, Maas-Hebner et al., 2015). Ecological fallacy happens when some site-specific characters significantly impact the stressor-response relationship. Typically, ecological fallacy includes the misidentification of limiting factor and the miscalculation of relationship direction or magnitude (Hamil et al., 2016). If ecological fallacy exists in developing nutrient criteria, then ecoregional nutrient

* Corresponding author.
E-mail address: yongliu@pku.edu.cn (Y. Liu).

https://doi.org/10.1016/j.ecolind.2019.105989
Received 13 June 2019; Received in revised form 29 November 2019; Accepted 4 December 2019

criteria are not applicable anymore and nutrient criteria should be site-specific. Therefore, determining whether ecological fallacy exists or not is critically important to set a proper spatial scale of nutrient criteria. However, whether ecological fallacy exists or not in nutrient criteria development has been neglected and remains unexploited and unclear.

Previously, water quality data were too rare to build site-specific stressor-response relationships, which had further hindered related explorations. Recently, with the accumulation of monitoring data, environmental and ecological data are becoming intensive (Hampton et al., 2017). Long-term and spatial water quality data are available (Soranno et al., 2017, Zhou et al., 2017). Under the data-intensive era, determining if ecological fallacy exists in nutrient criteria development becomes possible.

In this study, a long-term and nationwide water quality dataset is used to explore the existence of ecological fallacy in lake nutrient criteria development. Because of increasing attentions of lake eutrophication (Vinçon-Leite and Casenave, 2019), Chla was selected as the response variable and both total nitrogen (TN) and total phosphorus (TP) were selected as stressors (Malve and Qian, 2006). Bayesian linear model (BLM) (Mostafa, 2010) and Bayesian hierarchical model (BHM) (Cha et al., 2016) were employed to build ecoregional and site-specific nutrient-Chla relationships, respectively. Then, we compared differences between ecoregional and site-specific stressor-response relationships to determine if ecological fallacy exists or not.

2. Materials and methods

2.1. Types of ecological fallacy

The definition of ecological fallacy is narrative and has rarely been discussed in nutrient criteria development. To understand the meaning of ecological fallacy visually, we showed some examples of different types of ecological fallacy in Fig. 1. Ideally, when the ecological fallacy does not exist, ecoregional and site-specific relationships would be identical (Fig. 1a). In practice, some factors might influence the relationship minimally and ecoregional relationship may be slightly different from site-specific relationships (Fig. 1b), which will not impact the applicability of ecoregional nutrient criteria.

In contrast, ecological fallacy includes the misidentification of limiting factor, the miscalculation of relationship direction, and the miscalculation of relationship magnitude. The misidentification of limiting factor can be false determination of non-significance (Fig. 1c) or significance (Fig. 1d). The miscalculation of relationship direction is typically shown in Fig. 1d, in which the positive relationships are falsely determined as an overall negative relationship. The miscalculation of relationship magnitude can be the entire overestimation (the slope, Fig. 1f), entire underestimation (Fig. 1g), and partial misestimation of nutrient effect (Fig. 1h). Relatively large differences among site-specific relationships in Fig. 1c–h represent effects of some site-specific factors on the stressor-response relationship.

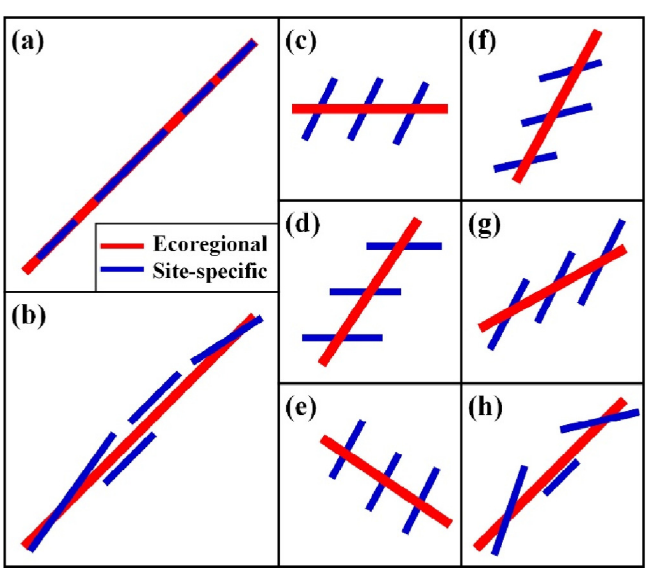

Fig. 1. Examples of patterns of ecoregional and site-specific relationships. The x-axis and y-axis are the stressor and response variable, respectively. Modified from Genser et al. (2015).

2.2. Data description

Monthly water quality data from 93 key monitoring lakes and reservoirs in China was collected by National Environmental Centre of China. Note that lakes and reservoirs are recognized as the same waterbody type to develop nutrient criteria (Huo et al., 2018, USEPA, 2000). The monitoring period ranges among 2010 and 2016. Besides, Environmental Monitoring Centre of Yunnan Province collected data of nine key monitoring lakes in Yunnan Province between 2006 and 2016. We removed observations with concentration values under method detection limits (TN $\leq$ 0.1 mg/L, TP $\leq$ 5 μg/L, and Chla $\leq$ 1 μg/L). Next, outliers were identified and removed, following a combination of discordancy tests and visual examination of probability plots (Qian and Lyons, 2006). Finally, the nationwide dataset includes 102 lakes or reservoirs with a total sample size of 4,448 (see Fig. S1 & Table S1 for details). Locations of these lakes and reservoirs cover all the eight ecoregions in China (Huo et al., 2014). The sites are not evenly distributed among ecoregions and about three-fourths of them are located at ecoregion V, VI, and VII (Table 1).

2.3. Model development

The model development process was shown is Fig. 2. To determine if ecological fallacy exists or not, we compared the ecoregional relationship and site-specific relationships by three aspects: (1) the sign or magnitude of regression slope, (2) the linear nutrient-Chla relationship on the log-log scale, and (3) the median of predicted Chla concentration conditioning average TN or TP concentration in some sites. Regression slope represents the effect of nutrient on Chla. More precisely, it represents the percent change in Chla concentration with per 1% change in nutrient concentration (Qian, 2016). Regression slopes and relationships give qualitative evidence, and medians of predicted Chla concentration will quantitatively show differences between the ecoregional and site-specific relationships.

Observed data of nutrients and Chla were firstly $\log_e$ transformed to accommodate the normality and homoscedasticity assumption (Malve and Qian, 2006, Oliver et al., 2017). Although lake-year mean (averaging data from the same lake in the same year) or lake mean (averaging data from the same lake) are always used to develop ecoregional relationship (Phillips et al., 2008), it has been recognized that averaging data may narrow the range of variables and mislead the cross-

Table 1
Summary of nutrient limitation conditions in eight ecoregions.

Ecoregion	Number of Sites	Limiting Nutrient	
		TN	TP
I	8	0	3
II	4	3	0
III	7	0	0
IV	2	0	0
V	20	2	3
VI	13	5	5
VII	41	6	5
VIII	7	0	2
Total	**102**	**16**	**18**

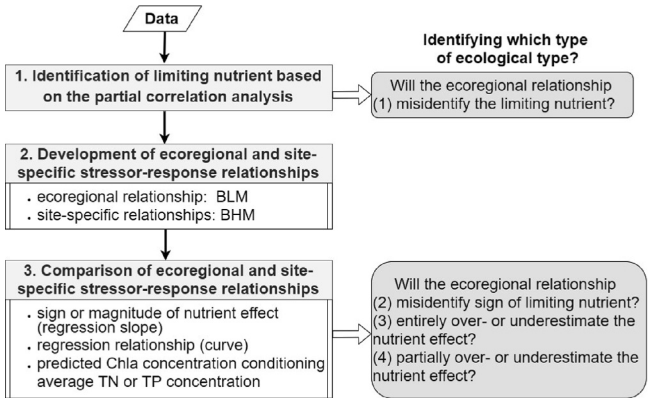

Fig. 2. Flowchart of the model development process.

sectional pattern (Dimberg, 2014, Jones et al., 1998). Therefore, data were not averaged in this study.

Figuring out the limiting nutrient is essential to nutrient criteria development (Dodds and Welch, 2000). We then determined the limiting nutrient in each site based on the partial correlation analysis (Liang et al., 2018b). Partial correlation analysis can eliminate the effect of other independent variables and reflect the relationship veritably (Wang et al., 2016). A *t*-statistic was established to test the significance of partial correlation coefficients (Kim, 2015). If a nutrient is positively significant correlated with Chla, it is selected as a limiting nutrient.

Next, we built stressor-response models to simulate TN-Chla and TP-Chla relationships, separately. BLM and BHM are employed to build the ecoregional and site-specific stressor-response relationships, respectively. Both of two methods have been widely used to build nutrient-Chla relationships (Lamon and Qian, 2008, Liang et al., 2018a, Liang et al., 2019). The ecoregional relationship was expressed as

$$y_k \sim N(\alpha + \beta x_k, \tau^2) \tag{1}$$

where y and x are the natural log of observed Chla and nutrient (TN or TP) concentration, k is the index of observed data in the same ecoregion, α and β are the regression intercept and slope, and τ^2 is the residual variance, respectively.

While it is hard to list all the site-specific factors influencing the relationship, BHM accounts for site-specific heterogeneity by latent variables (Hamil et al., 2016). In addition, BHM also improves the overall estimation accuracy by partially pooling of data (Qian et al., 2015). Therefore, BHM is employed to obtain site-specific relationships. Within each ecoregion, data from sites with the same limiting nutrient were partially pooled to obtain site-specific relationships (Cha et al., 2016):

$$y_{ij} \sim N(a_i + b_i x_{ij}, \sigma^2) \tag{2}$$

$$a_i \sim N(a_0, \sigma_a^2) \tag{3}$$

$$b_i \sim N(b_0, \sigma_b^2) \tag{4}$$

where i is the index of site, j is the index of observed data in each site, a_i and b_i represent the regression intercept and slope for each site-specific relationship, and σ^2 is the residual variance, respectively. Each group of site-specific parameters shares a common normal prior distribution waiting for estimation, which leads to the partial pooling of data (Eqs. (3) and (4)).

All the computations were conducted in R software (Version 3.4.2). Non-informative prior distributions were used for all the parameters. Posterior distributions of parameters for the ecoregional relationship and site-specific relationships were obtained via Hamiltonian Monte Carlo (HMC) algorithm implemented in the *rstan* package (Stan Development Team, 2016). Four HMC chains were set with random initials. Each chain ran 100,000 iterations, with the first half for burn-in period and the last half to obtain posterior distributions. Convergences were assured by R-hat statistic (Gelman and Hill, 2007).

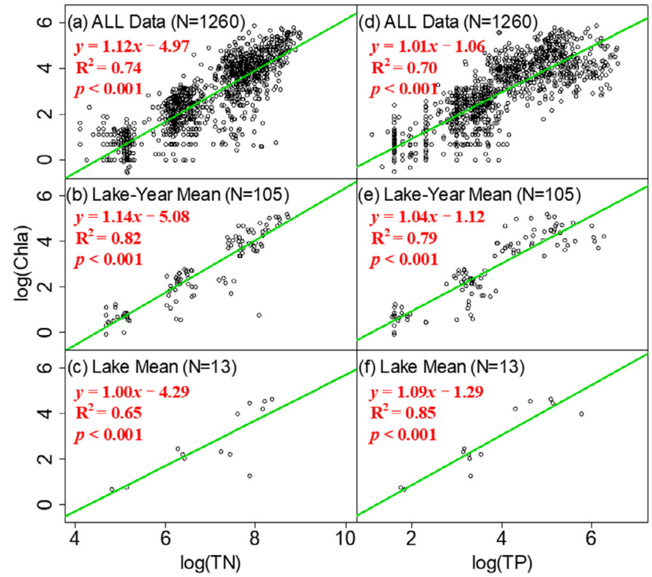

Fig. 3. Ecoregional nutrient-Chla relationships in ecoregion VI. Plots in the left panel show relationships between log (TN) and log (Chla) based on (a) all the data, (b) lake-year mean, and (c) lake mean. Plots in the right panel show relationships between log (TP) and log (Chla) based on (d) all the data, (e) lake-year mean, and (f) lake mean. Black points are observed data. Blue lines are linear regression curves. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Nutrient limitation conditions

At the national scale, correlation coefficients between TN or TP and Chla are all significantly positive using data without averaging, lake-year mean data, and lake mean data (Fig. S2), which is also the case at the ecoregional scale, e.g. in ecoregion VI (Fig. 3). However, at the site-specific scale, only in a few sites TN or TP is determined as the limiting nutrient (Table 1). More precisely, only 16% and 18% of sites are limited by TN and TP, respectively. Sites with TN or TP limitation are not evenly distributed among ecoregions, potentially caused by the uneven distribution of monitoring sites. Obviously, data aggregation either from national or ecoregional scale misleads the determination of limiting nutrient in most sites. That is a typical type of ecological fallacy, under which condition the limiting factor is misidentified (refer to Fig. 1d).

3.2. Relationship comparisons

If an ecoregion only has one lake limited by a certain nutrient, the ecoregional and site-specific stressor-response relationship would be identical. However, that could not be treated as the evidence supporting the ecoregional nutrient criteria. For ecoregions with more than one sites limited by a certain nutrient, we can build both the ecoregional and site-specific stressor-response relationships. We took the ecoregional and site-specific nutrient (TN or TP)-Chla relationships in the same ecoregion as a pair of relationship. Therefore, in total nine pairs of relationship are built (Table 1), namely, the TN-Chla relationship in ecoregion II, V, VI, and VII, and the TP-Chla relationship in ecoregion I, V, VI, VII, and VIII. Posterior distributions of parameters are summarized in supporting materials.

We found that each pair of relationship showed a certain type of ecological fallacy. Nine pairs of relationship could be clustered into

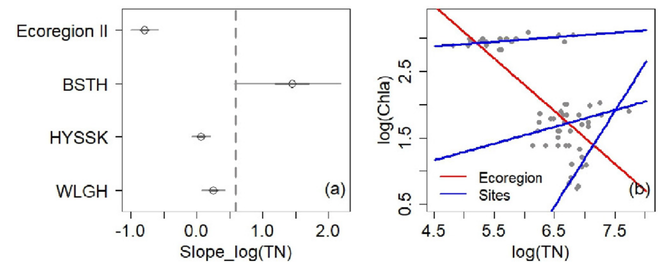

Fig. 4. Ecoregional and site-specific TN-Chla relationships in ecoregion II. Circle dots, thick lines, fine lines, and dashed vertical line in (a) represent average values, 50% credible intervals, 90% credible intervals of regression slopes and the overall mean of all the site-specific slopes. Grey points in (b) are observed data from the ecoregion or a certain site, the same below. The left plot shows regression slopes of the ecoregional relationship and site-specific relationships. The right plot shows regression curves of relationships. This pair of relationship shows a pattern of ecological fallacy (Fig. 1e) by miscalculating the relationship direction at the ecoregional scale.

three types. The first type is the miscalculation of regression slope direction, as showed by the TN-Chla relationship in ecoregion II (Fig. 4). Based on the aggregated data, a negative regression slope was obtained, while all the site-specific relationships had positive regression slopes.

The second type is the miscalculation of all the site-specific nutrient effects, including five pairs of relationship. For example, for the TN-Chla relationships in ecoregion VII, the regression slope of ecoregional relationship is much smaller than all the slopes of site-specific relationships (Fig. 5), indicating entire underestimations of site-specific nutrient effects. There are huge differences between the ecoregional and site-specific relationships and the median of predicted Chla concentrations in some sites, e.g. in TuoLinHu and HuangDaHu. This type of ecological fallacy also exists for the TN-Chla relationship in

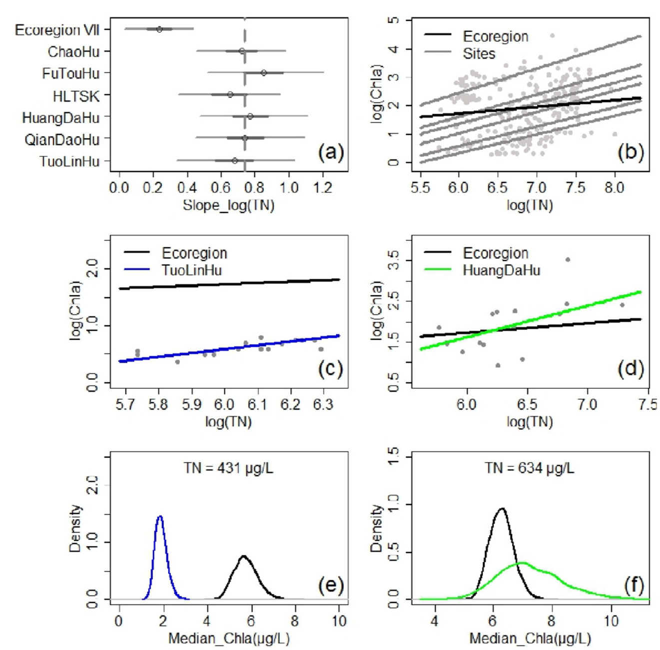

Fig. 5. Ecoregional and site-specific TN-Chla relationships in ecoregion VII. Grey points in (c) & (d) are observed data from the ecoregion or a certain site, the same below. Plots in the top panel show a pattern of ecological fallacy by underestimating the nutrient effect (regression slope) at the ecoregional scale (refer to Fig. 1g). Plots in the middle panel compare the ecoregional relationship and site-specific relationships in TuoLinHu and HuangDaHu. Plots in the bottom panel show posterior distributions of predicted median Chla concentration in TuoLinHu and HuangDaHu, conditioning the average TN concentrations (431 μg/L and 634 μg/L, respectively).

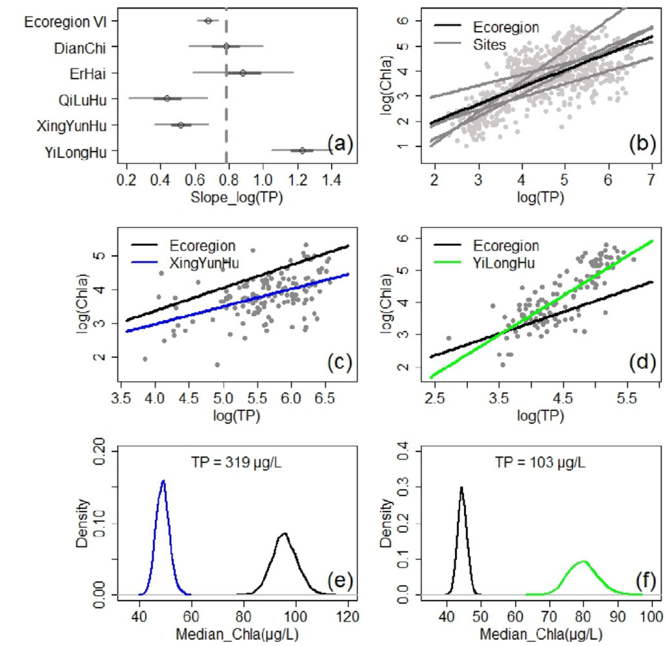

Fig. 6. Ecoregional and site-specific TP-Chla relationships in ecoregion VI. Plots in the top panel show a pattern of the ecological fallacy by misestimating the nutrient effect (regression slope) at the ecoregional scale (refer to Fig. 1h). Plots in the middle panel compare the ecoregional relationship and site-specific relationships in XingYunHu and YiLongHu. Plots in the bottom panel show posterior distributions of predicted median Chla concentration in XingYunHu and YiLongHu, conditioning the average TP concentrations (319 μg/L and 103 μg/L, respectively).

ecoregion V (Fig. S3), and the TP-Chla relationship in ecoregion V, VII, and VIII (Fig. S4–S6).

The third type is the miscalculation of nutrient effect in some sites, including three pairs of relationship. A typical example is the TP-Chla relationship in ecoregion VI (Fig. 6). Regression slope of ecoregional relationship is close to the overall slope of all the site-specific relationships. However, the ecoregional relationship dramatically underestimates the nutrient effect in YiLongHu and overestimates the nutrient effect in XingYunHu. Differences of median Chla concentration predicted by the ecoregional relationship and the site-specific relationships are also apparent in these two sites. While the first two types were impressive and had been well-recognized in other ecological fields (Hamil et al., 2016), the third type was easily overlooked, in which ecoregional relationship might also lead to huge biases for some certain sites. This type of ecological fallacy also exists for the rest two pairs of relationship (Figs. S7 and S8).

4. Discussion

4.1. Spatial scale of nutrient criteria

Results of nutrient limitation conditions showed that the limiting nutrient in some lakes were misidentified by aggregating data from the ecoregion. In many lakes, we did not identify any limiting nutrient. Actually, except for nutrient, many other variables, such as water temperature, Secchi depth, and wind, could be influencing factors instead of nutrients for the long-term variation of Chla concentrations (Liu et al., 2016, Mcquatters-Gollop et al., 2007). The significant positive correlation coefficient (Fig. 3, Fig. S2) might be caused by the data aggregation or data averaging. A large sample size always easily brings a statistically significant regression slope (Bryhn and Dimberg, 2011). Besides, it has been proved that averaging data could eliminate information on data variability and therefore strengthen the relevance

(Jones et al., 1998). Therefore, a national or ecoregional relationship would always have a significant regression slope but might fail to correctly identify the limiting nutrient. Note that for sites without any limiting nutrient, it is impossible to develop nutrient criteria based on the nutrient-Chla relationship. For the purpose of nutrient criteria development, other management endpoints, e.g. the macroinvertebrate or fish community metrics (Evans-White et al., 2013) could be selected.

Results of relationship comparisons showed that ecological fallacy existed in each pair of relationship. There are large variations among site-specific relationships within the same ecoregion, which might be caused by the heterogeneity of site-specific factors. Many site-specific factors, such as the catchment area (Lohman and Jones, 1999), lake mean depth, and the percentage of wooded wetlands (Wagner et al., 2011), have been determined to impact the nutrient-Chla relationship. Exploring the exact drivers is expected to be of great importance (Soranno et al., 2016). However, it is beyond the scope of this study for the limitation of relatively small numbers of valid sites (sites with the same limiting nutrient) within ecoregions. Apparently, the heterogeneity of site-specific factors was ignored in the ecoregional nutrient criteria when developing ecoregional stressor-response relationship using the ecoregion-aggregated data.

While many studies focus on approaches to determine ecoregional nutrient criteria values (Evans-White et al., 2013), we here explored the spatial scale of nutrient criteria, which is a more fundamental and easily neglected issue. Based on the above analysis, different types of ecological fallacy, including the misidentification of limiting nutrient, and miscalculation of nutrient effect direction, could happen when developing ecoregional stressor-response relationship, which limits the applicability of ecoregional nutrient criteria. Therefore, we determine that the ecoregional nutrient criteria is not applicable for all the lakes within the ecoregion. The spatial scale of lake nutrient criteria should be the site-specific scale.

We emphasize that the primary reason for developing site-specific instead of ecoregional nutrient criteria is to avoid ecological fallacy rather than to improve accuracy. Site-specific nutrient criteria are expected to be a naturally better way with data accumulation for its higher accuracy (Mclaughlin, 2014). However, when considering the existence of ecological fallacy, we recognize that the ecoregional stressor-response relationship might be wrong for many sites. To avoid the ecological fallacy, site-specific stressor-response relationship is required, and site-specific nutrient criteria should be used. Ecoregional nutrient criteria even cannot be used as a rough estimation of site-specific criteria for any site (including any data-poor site). The proper thing to do for data-poor site is to collect enough data rather than taking the ecoregional nutrient criteria as a replacement, because such a replacement takes a risk of wrong deduction when the ecological fallacy exists.

Our results also emphasize the important impact of site-specific factors in nutrient criteria development. As did in a few ecological researches recently (Huang et al., 2017, Mimet et al., 2016), for example, Roitberg and Shoshany (2017) found that space-for-time substitution overestimates consequences of rainfall decline on properties of vegetation patterns between semi-arid and arid rainfall regions, the space-for-time substitution was challenged here. The underlying mechanism is that the spatial heterogeneity of site-specific factors makes the response variable vary unequally with the stressor in space (Mimet et al., 2016). The space-for-time substitution should not be used in nutrient criteria development.

By recognizing the importance of site-specific factors, a few studies provided evidence supporting site-specific nutrient criteria based on the reference conditions method. Read et al. (2015) found that lake-specific characters were more important for explaining water quality (56% and 60% variance explained for TP and TN respectively) than regional scale drivers. Olson and Hawkins (2013) revealed that the site-specific nutrient criteria accounted for natural variation among sites better than criteria based on regional average conditions. Therefore, site-specific nutrient criteria were supported both by the stressor-response model in this study and by the reference conditions method from other studies.

Previously, environmental and ecological data were rare. Testifying the existence of ecological fallacy in nutrient criteria is hardly possible. In the current study, using the spatio-temporal dataset we explored the existence of ecological fallacy and determine the proper spatial scale of nutrient criteria. Obviously, this provides essential evidence to guide the future nutrient criteria development in the new era with intensive environmental and ecological data.

Moreover, our study is also instructive for other researches in environmental and ecological fields, where stressor-response relationship should be built. Our results remind researchers to pay attentions to the possibility of ecological fallacy when aggregating data to develop ecoregional stressor-response relationship and further to extend the ecoregional relationship to sites. Researchers should be aware of the risk of deduction based on the ecoregional stressor-response relationship. We encourage that researchers should always build site-specific stressor-response relationship.

4.2. Partial pooling of data in data-intensive era

Definitely, in the data-intensive era, data is becoming adequate to developing reliable site-specific stressor-response relationship. However, does that mean we should develop site-specific stressor-response relationship based solely on the data of a certain site and abandon the constrain of ecoregion? Our answer might be NO!

We recommend the partial pooling of data strategy in data-intensive era to develop site-specific stressor-response relationships. Reasons are generally from two aspects. Firstly, the delineation of ecoregions is always based on the similarity and contiguity of many indicators of climate and landscape (Cheruvelil et al., 2017, Omernik, 1987). Although that cannot guarantee the consistency of stressor-response relationship within the ecoregion as showed in our study, the homogeneity of drivers at the regional scale would always make the site-specific relationships within the ecoregion closely related (Wagner et al., 2011) and the rationality of ecoregions as a good tool for ecosystem management has been verified (Smith et al., 2018). Secondly, developing site-specific stressor-response relationship based solely on the data of a certain site is a no pooling of data strategy (aggregating all the ecoregional data to build an ecoregional stressor-response relationship is a complete pooling of data strategy). Researches have showed that the partial pooling of data could improve the overall prediction accuracy and reduce the uncertainty (Cha et al., 2016, Qian et al., 2015). Therefore, the partial pooling of data is a better strategy than no pooling strategy. As an typical partial pooling method, BHM is an effective tool to build site-specific stressor-response relationship and in the meanwhile to use ecoregional information (Genser et al., 2015). Another similar method is the multilevel model (Qian et al., 2010). Note that such a recommendation is certainly applicable for the nutrient criteria development in data-intensive era.

5. Conclusions

Our study focused on determining a proper spatial scale of lake nutrient criteria, which is essential to guarantee the applicability of nutrient criteria. Benefitting from long-term and nationwide dataset, we found that ecological fallacy commonly existed when developing nutrient criteria and thus determined that lake nutrient criteria should be site-specific. The findings provide important information to inform the future nutrient criteria development. Inspired by our results, we recommended the partial pooling strategy to build stressor-response relationship in data-intensive era in the environmental and ecological researches.

CRediT authorship contribution statement

Zhongyao Liang: Conceptualization, Methodology, Software, Formal analysis, Visualization, Supervision, Writing - review & editing. **Feifei Dong:** Visualization, Writing - review & editing. **Song S. Qian:** Visualization, Writing - review & editing. **Yong Liu:** Resources, Funding acquisition, Writing - review & editing. **Huili Chen:** Writing - review & editing. **Wentao Lu:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge Sifeng WU for all the enlightening discussion on this manuscript. The manuscript is financially supported by high-performance Computing Platform of Peking University and National Natural Science Foundation of China (51779002, 51721006).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2019.105989.

References

Bryhn, A.C., Dimberg, P.H., 2011. An operational definition of a statistically meaningful trend. PLoS ONE 6 (4), e19241.
Celentano, E., Defeo, O., 2016. Climate effects on the mole crab Emerita brasiliensis on a dissipative beach of Uruguay. Marine Ecology Progress 552, 211–222.
Cha, Y.K., Soon Park, S., Won Lee, H., Stow, C.A., 2016. A Bayesian hierarchical approach to model seasonal algal variability along an upstream to downstream river gradient. Water Resour. Res. 52 (1), 348–357.
Cheruvelil, K.S., Yuan, S., Webster, K.E., Tan, P.N., Lapierre, J.F., Collins, S.M., Fergus, C.E., Scott, C.E., Henry, E.N., Soranno, P.A., 2017. Creating multithemed ecological regions for macroscale ecology: Testing a flexible, repeatable, and accessible clustering method. Ecol. Evol. 7 (9), 3046–3058.
Dimberg, P.H., 2014. Defining a new parameter for regression models with aggregated data in aquatic science. Environmetrics 25 (2), 97–106.
Dodds, W.K.K., Welch, E.B., 2000. Establishing nutrient criteria in streams. J. North Am. Benthol. Soc. 19 (1), 186–196.
Evans-White, M.A., Haggard, B.E., Scott, J.T., 2013. A Review of Stream Nutrient Criteria Development in the United States. J. Environ. Qual. 42 (4), 1002–1014.
Gelman, A., Hill, J., 2007. Data Analysis Using Regression andMultilevel/Hierarchical Models. Cambridge University Press, New York, NY.
Genser, B., Teles, C.A., Barreto, M.L., Fischer, J.E., 2015. Within- and between-group regression for improving the robustness of causal claims in cross-sectional analysis. Environmental Health 14 (1), 1–10.
Hamil, K.A.D., Iii, B.V.I., Huang, W.K., Fei, S., Zhang, H., 2016. Cross-scale contradictions in ecological relationships. Landscape Ecol. 31 (1), 7–18.
Hampton, S.E., Jones, M.B., Wasser, L.A., Schildhauer, M.P., Supp, S.R., Brun, J., Hernandez, R.R., Boettiger, C., Collins, S.L., Gross, L.J., 2017. Skills and knowledge for data-intensive environmental research. Bioscience 67 (6), 546–557.
Huang, M., Piao, S., Janssens, I.A., Zhu, Z., Wang, T., Wu, D., Ciais, P., Myneni, R.B., Peaucelle, M., Peng, S., Yang, H., Penuelas, J., 2017. Velocity of change in vegetation productivity over northern high latitudes. Nat. Ecol. Evol. 1, 1649–1654.
Huo, S., Ma, C., Xi, B., Gao, R., Deng, X., Jiang, T., He, Z., Su, J., Wu, F., Liu, H., 2014. Lake ecoregions and nutrient criteria development in China. Ecol. Ind. 46, 1–10.
Huo, S., Ma, C., Xi, B., Zhang, Y., Wu, F., Liu, H., 2018. Development of methods for establishing nutrient criteria in lakes and reservoirs: a review. J. Environ. Sci. 67, 54–66.
Jones, J.R., Knowlton, M.F., Kaiser, M.S., 1998. Effects of aggregation on chlorophyll-phosphorus relations in Missouri reservoirs. Lake Reservoir Manage. 14 (1), 1–9.
Kim, S., 2015. ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. Commun. Stat. Appl. Methods 22 (6), 665–674.
Lamon III, E.C., Qian, S.S., 2008. Regional scale stressor-response models in aquatic ecosystems. J. Am. Water Resour. Assoc. 44 (3), 771–781.
Lester, R.E., Close, P.G., Barton, J.L., Pope, A.J., Brown, S.C., 2014. Predicting the likely response of data-poor ecosystems to climate change using space-for-time substitution across domains. Glob. Change Biol. 20 (11), 3471–3481.
Liang, Z., Chen, H., Wu, S., Zhang, X., Yu, Y., Liu, Y., 2018a. Exploring Dynamics of the Chlorophyll a-Total Phosphorus Relationship at the Lake-Specific Scale: a Bayesian Hierarchical Model. Water Air Soil Pollut. 229 (1), 21.
Liang, Z., Liu, Y., Chen, H., Ji, Y., 2019. Is ecoregional scale precise enough for lake nutrient criteria? Insights from a novel relationship-based clustering approach. Ecol. Indic. 97, 341–349.
Liang, Z., Wu, S., Chen, H., Yu, Y., Liu, Y., 2018b. A probabilistic method to enhance understanding of nutrient limitation dynamics of phytoplankton. Ecol. Model. 368, 404–410.
Liu, X., Qian, K., Chen, Y., Gao, J., 2016. A comparison of factors influencing the summer phytoplankton biomass in China's three largest freshwater lakes: Poyang, Dongting, and Taihu. Hydrobiologia 1–20.
Lohman, K., Jones, J.R., 1999. Nutrient – sestonic chlorophyll relationships in northern Ozark stream. Can. J. Fish. Aquat. Sci. 56 (1), 124–130.
Maas-Hebner, K.G., Harte, M.J., Molina, N., Hughes, R.M., Schreck, C., Yeakley, J.A., 2015. Combining and aggregating environmental data for status and trend assessments: challenges and approaches. Environ. Monit. Assess. 187 (5), 278.
Malve, O., Qian, S.S., 2006. Estimating nutrients and chlorophyll a relationships in Finnish lakes. Environ. Sci. Technol. 40 (24), 7848–7853.
Martay, B., Monteith, D.T., Brewer, M.J., Brereton, T., Shortall, C.R., Pearce-Higgins, J.W., 2016. An indicator highlights seasonal variation in the response of Lepidoptera communities to warming. Ecol. Ind. 68, 126–133.
Mclaughlin, D.B., 2014. Maximizing the accuracy of field-derived numeric nutrient criteria in water quality regulations. Integr. Environ. Assess. Manage. 10 (1), 133.
Mcquatters-Gollop, A., Raitsos, D.E., Edwards, M., Pradhan, Y., Mee, L.D., Lavender, S.J., Attrill, M.J., 2007. A Long-Term Chlorophyll Data set Reveals Regime Shift in North Sea Phytoplankton Biomass Unconnected to Nutrient Trends. Limnol. Oceanogr. 52 (2), 635–648.
Mimet, A., Pellissier, V., Houet, T., Julliard, R., Simon, L., 2016. A holistic landscape description reveals that landscape configuration changes more over time than composition: implications for landscape ecology studies. PLoS One 11 (3), e0150111.
Mostafa, M.M., 2010. A Bayesian approach to analyzing the ecological footprint of 140 nations. Ecol. Ind. 10 (4), 808–817.
Oliver, S.K., Collins, S.M., Soranno, P.A., Wagner, T., Stanley, E.H., Jones, J.R., Stow, C.A., Lottig, N.R., 2017. Unexpected stasis in a changing world: Lake nutrient and chlorophyll trends since 1990. Glob. Change Biol. 23 (12), 5455–5467.
Olson, J.R., Hawkins, C.P., 2013. Developing site-specific nutrient criteria from empirical models. Freshwater Science 32 (3), 719–740.
Omernik, J.M., 1987. ECOREGIONS OF THE CONTERMINOUS UNITED-STATES. Ann. Assoc. Am. Geogr. 77 (1), 118–125.
Phillips, G., Pietiläinen, O.P., Carvalho, L., Solimini, A., Solheim, A.L., Cardoso, A.C., 2008. Chlorophyll–nutrient relationships of different lake types using a large European dataset. Aquat. Ecol. 42 (2), 213–226.
Qian, S.S., 2016. Environmental and Ecological Statistics with R, Second Edition. Chapman & Hall/CRC.
Qian, S.S., Cuffney, T.F., Alameddine, I., McMahon, G., Reckhow, K.H., 2010. On the application of multilevel modeling in environmental and ecological studies. Ecology 91 (2), 355–361.
Qian, S.S., Lyons, R.E., 2006. Characterization of background concentrations of contaminants using a mixture of normal distributions. Environ. Sci. Technol. 40 (19), 6021–6025.
Qian, S.S., Stow, C.A., Cha, Y., 2015. Implications of Stein's Paradox for Environmental Standard Compliance Assessment. Environ. Sci. Technol. 49 (10), 5913–5920.
Read, E.K., Patil, V.P., Oliver, S.K., Hetherington, A.L., Brentrup, J.A., Zwart, J.A., Winters, K.M., Corman, J.R., Nodine, E.R., Woolway, R.I., 2015. The importance of lake-specific characteristics for water quality across the continental United States. Ecological Applications A Publication of the Ecological Society of America 25 (4), 943.
Robinson, W.S., 1950. Ecological correlations and the behavior of individuals. Am. Sociol. Rev. 15 (3), 351–357.
Roitberg, E., Shoshany, M., 2017. Can spatial patterns along climatic gradients predict ecosystem responses to climate change? Experimenting with reaction-diffusion simulations. Plos One 12 (4), e0174942.
Smith, J.R., Letten, A.D., Ke, P.-J., Anderson, C.B., Hendershot, J.N., Dhami, M.K., Dlott, G.A., Grainger, T.N., Howard, M.E., Morrison, B.M., 2018. A global test of ecoregions. Nat. Ecol. Evol. 2 (12), 1889.
Soranno, P.A., Bacon, L.C., Beauchene, M., Bednar, K.E., Bissell, E.G., Boudreau, C.K., Boyer, M.G., Bremigan, M.T., Carpenter, S.R., Carr, J.W., Cheruvelil, K.S., Christel, S. T., Claucherty, M., Collins, S.M., Conroy, J.D., Downing, J.A., Dukett, J., Fergus, C.E., Filstrup, C.T., Funk, C., Gonzalez, M.J., Green, L.T., Gries, C., Halfman, J.D., Hamilton, S.K., Hanson, P.C., Henry, E.N., Herron, E.M., Hockings, C., Jackson, J.R., Jacobson-Hedin, K., Janus, L.L., Jones, W.W., Jones, J.R., Keson, C.M., King, K.B.S., Kishbaugh, S.A., Lapierre, J.-F., Lathrop, B., Latimore, J.A., Lee, Y., Lottig, N.R., Lynch, J.A., Matthews, L.J., McDowell, W.H., Moore, K.E.B., Neff, B.P., Nelson, S.J., Oliver, S.K., Pace, M.L., Pierson, D.C., Poisson, A.C., Pollard, A.I., Post, D.M., Reyes, P.O., Rosenberry, D.O., Roy, K.M., Rudstam, L.G., Sarnelle, O., Schuldt, N.J., Scott, C. E., Skaff, N.K., Smith, N.J., Spinelli, N.R., Stachelek, J.J., Stanley, E.H., Stoddard, J. L., Stopyak, S.B., Stow, C.A., Tallant, J.M., Tan, P.-N., Thorpe, A.P., Vanni, M.J., Wagner, T., Watkins, G., Weathers, K.C., Webster, K.E., White, J.D., Wilmes, M.K. and Yuan, S. (2017) LAGOS-NE: A multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of U.S. lakes. GigaScience.
Soranno, P.A., Cheruvelil, K.S., Bissell, E.G., Bremigan, M.T., Downing, J.A., Fergus, C.E., Filstrup, C.T., Henry, E.N., Lottig, N.R., Stanley, E.H., 2016. Cross-scale interactions: quantifying multi-scaled cause–effect relationships in macrosystems. Front. Ecol. Environ. 12 (1), 65–73.
Soranno, P.A., Cheruvelil, K.S., Stevenson, R.J., Rollins, S.L., Holden, S.W., Heaton, S., Torng, E., 2008. A framework for developing ecosystem-specific nutrient criteria: Integrating biological thresholds with predictive modeling. Limnol. Oceanogr. 53 (2), 773–787.
Stan Development Team Stan Modeling Language Users Guide and Reference Manual

2016.
USEPA, 1998. National Strategy for the Development of Regional Nutrient Criteria, EPA 822-R-98-002. U.S Environmental Protection Agency, Washington, D.C.
USEPA Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs. EPA-822-B-00-001. US EPA, Office of Water 2000 Washington, DC.
USEPA Using Stressor-Response Relationships to Derive Numeric Nutrient Criteria. EPA-820-S-10-001, U.S. Environmental Protection Agency, Office of Water 2010 Washington, D.C.
Vinçon-Leite, B., Casenave, C., 2019. Modelling eutrophication in lake ecosystems: a review. Sci. Total Environ. 651, 2985–3001.
Wagner, T., Soranno, P.A., Webster, K.E., Cheruvelil, K.S., 2011. Landscape drivers of regional variation in the relationship between total phosphorus and chlorophyll in lakes. Freshw. Biol. 56 (9), 1811–1824.
Wang, C., Wang, Z., Wang, P., Zhang, S., 2016. Multiple effects of environmental factors on algal growth and nutrient thresholds for harmful algal blooms: application of response surface methodology. Environ. Model. Assess. 21 (2), 247–259.
Zhou, Y., Ma, J., Zhang, Y., Qin, B., Jeppesen, E., Shi, K., Brookes, J.D., Rgm, S., Zhu, G., Gao, G., 2017. Improving water quality in China: Environmental investment pays dividends. Water Res. 118, 152–159.