



HAL
open science

AquaDesign: A tool to assist aquaculture production design based on abiotic requirements of animal species

Grégoire Butruille, Marielle Thomas, Alain Pasquet, Nellya Amoussou, Lola Toomey, Axel Rosenstein, Sandrine Chauchard, Thomas Lecocq

► To cite this version:

Grégoire Butruille, Marielle Thomas, Alain Pasquet, Nellya Amoussou, Lola Toomey, et al.. AquaDesign: A tool to assist aquaculture production design based on abiotic requirements of animal species. PLoS ONE, 2022, 17 (8), pp.e0272508. 10.1371/journal.pone.0272508 . hal-04126514

HAL Id: hal-04126514

<https://hal.science/hal-04126514v1>

Submitted on 13 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

RESEARCH ARTICLE

AquaDesign: A tool to assist aquaculture production design based on abiotic requirements of animal species

Grégoire Butruille¹, Marielle Thomas^{1,2}, Alain Pasquet¹, Nellya Amoussou^{1,2},
Lola Toomey¹, Axel Rosenstein¹, Sandrine Chauchard³, Thomas Lecocq^{1,2*}

1 University of Lorraine, URAFPA, INRAE, Nancy, France, **2** LTSER France, Zone Atelier du Bassin de la Moselle, Vandœuvre-lès-Nancy, France, **3** University of Lorraine, AgroParisTech, INRAE, UMR Silva, Nancy, France

* thomas.lecocq@univ-lorraine.fr



OPEN ACCESS

Citation: Butruille G, Thomas M, Pasquet A, Amoussou N, Toomey L, Rosenstein A, et al. (2022) AquaDesign: A tool to assist aquaculture production design based on abiotic requirements of animal species. PLoS ONE 17(8): e0272508. <https://doi.org/10.1371/journal.pone.0272508>

Editor: Andreas C. Bryhn, Swedish University of Agricultural Science, SWEDEN

Received: February 2, 2022

Accepted: July 20, 2022

Published: August 1, 2022

Copyright: © 2022 Butruille et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All required datasets can be downloaded from GBIF (www.gbif.org), EarthEnv (www.earthenv.org/streams), FLO1K (DOI: [10.1038/sdata.2018.52](https://doi.org/10.1038/sdata.2018.52)), WorldClim (www.worldclim.org), and R-package geosphere (cran.r-project.org/web/packages/geosphere/index.html) using our R-script available on Figshare (DOI: [10.6084/m9.figshare.17170958.v3](https://doi.org/10.6084/m9.figshare.17170958.v3)).

Funding: NA is supported by the French Ministère de l'Enseignement Supérieur et de la Recherche (France). This study was supported by the research

Abstract

Farming new species and promoting polyculture can enhance aquaculture sustainability. This implies to define the rearing conditions that meet the ecological requirements of a target species and/or to assess if different species can live in the same farming environment. However, there is a large number of rearing conditions and/or taxon combinations that can be considered. In order to minimise cumbersome and expensive empirical trials to explore all possibilities, we introduce a tool, AquaDesign. It is based on a R-script and package which help to determine farming conditions that are most likely suitable for species through *in silico* assessment. We estimate farming conditions potentially suitable for an aquatic organism by considering the species niche. We define the species n-dimensional niche hypervolume using a correlative approach in which the species niche is estimated by relating distribution data to environmental conditions. Required input datasets are mined from several public databases. The assistant tool allows users to highlight (i) abiotic conditions that are most likely suitable for species and (ii) combinations of species potentially able to live in the same abiotic environment. Moreover, it offers the possibility to assess if a particular set of abiotic conditions or a given farming location is potentially suitable for the monoculture or the polyculture of species of interest. Our tool provides useful pieces of information to develop freshwater aquacultures. Using the large amount of biogeographic and abiotic information available in public databases allows us to propose a pragmatic and operational tool even for species for which abiotic requirements are poorly or not available in literature such as currently non-produced species. Overall, we argue that the assistant tool can act as a stepping stone to promote new aquatic productions which are required to enhance aquaculture sustainability.

Introduction

The world's human population has increased dramatically since the 1950s, which has triggered a surge in the demand for food products [1]. This demand has been mainly fulfilled by upgrading already existing agricultural productions and developing new ones. Aquaculture is one of

programmes SEPURE (funded by the European Fund for Maritime Affairs and Fisheries, FEAMP, UAEE-FEAMP-47-SEPURE) and Interreg GR VA Perciponie-073-4-08-191 (2020-2022). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

the agricultural sectors whose productions have skyrocketed these last decades [2]. The aquaculture contribution to total human-consumed aquatic products has increased from less than 5% in 1970 to more than 50% in 2018 [2]. Aquaculture is expected to take an even more prominent role in human food security and nutrition in the near future, especially as wild fisheries increasingly fail to meet the ever-growing demand for aquatic products [2]. Nevertheless, aquaculture is often criticised due to its negative environmental impacts and its potential unsustainability [3, 4]. Indeed, the development of aquaculture has triggered habitat destructions, biodiversity loss, biological invasions, and pathogen spill-overs, as well as water-quality degradation and eutrophication of wild aquatic ecosystems, notably due to a poorly efficient use of inputs and intrinsic resources of farmed systems [4]. Moreover, aquaculture production relies mainly on few species. For instance, only 27 species comprised over 90% of total finfish production in 2018 [2]. This low diversity of produced species jeopardises (i) human food security because the heavy dependence on few taxa puts at risk aquaculture production if, for instance, an epizootic outbreak happens in farmed species, and (ii) the future of productions since poorly diversified production limits the adaptive potential of aquaculture to face environment or consumer demand changes [5–7].

In order to overcome these issues, international organisations and scientists are strongly advocating for (i) production diversification in regard to species [7, 8] and (ii) application of agroecology concepts to minimise aquaculture environmental impacts, maximise the use of farming system resources, and promote the production of local species [4]. In this context, starting new domestication programs for currently unfarmed species and promoting development of polyculture (i.e. rearing/breeding two or more species in a particular production) have been pointed out as a way to design the tomorrow's aquaculture [4, 8, 9].

Developing new species productions is challenging because it implies to use an appropriate farming environment that suits the species ecological requirements [9]. Identifying a suitable farming environment has often been based on trial-and-error approaches through empirical experimental attempts. However, such approaches are cumbersome and expensive. Therefore, they can be ineffective and inefficient since, for a particular species or species combination, there is a large number of farming environment set-ups that can be considered. Moreover, this can raise concerns about animal welfare (e.g. for fish) if unsuitable farming conditions are tested. This calls for the development of *in silico* approach that allows limiting the number of possibilities that should be assessed in subsequent experimental attempts.

The ecological niche is the match of a species to a specific set of environmental conditions [10]. The niche can be interpreted as a n-dimensional hypervolume in a space defined by several independent axes corresponding to species requirements (e.g. food size, temperature) [11]. The hypervolume boundaries indicate the conditions that permit the species growth and reproduction [11]. This interpretation has provided an operational concept to use species niches in applied sciences. Biological conservation is a stunning example of wide application of niche analyses to project consequences of habitat loss (e.g. [12]), climate change (e.g. [13]), or biological invasion (e.g. [14]), and to provide guidelines and recommendations to manage wild species and ecosystems (e.g. [15]). Similarly, ecological niche analyses can be regarded as a useful tool for development of new productions in aquaculture. Indeed, the ability of species to live in a particular farming environment and/or with other species could be evaluated by comparing the taxon niche with the farming environment and/or with niches of other species. Moreover, species niche analyses could provide useful pieces of information to design suitable farming environment from scratch. Nevertheless, such an approach has not been implemented in aquatic farming development. In order to move it to an operational approach for aquaculture development, an easy-to-use packaged workflow to estimate and compare species niches should be provided to aquatic farming stakeholders.

In this paper, we introduce an open source R-script and package [16] designed as an assistant tool, AquaDesign, to facilitate development of new animal species productions in monoculture and polyculture based on niche analyses [17, 18]. Our goal is to use the large amount of biogeographic and environmental information available in public databases to provide a pragmatic and operational way to highlight species or species combination that could be further considered in new aquaculture developments.

Methodological and ecological assumptions to assess species niche

Several methods are available to define species niche, but most studies use a correlative approach where the species niche is estimated by relating distribution data to environmental conditions [19]. This approach is especially useful and pragmatic to estimate niches of large sets of species because it does not rely on time- and money-consuming experimental assessment in controlled conditions nor on detailed knowledge about ecological requirements of a particular species. Moreover, it relies on dataset mined from increasingly available ecological databases.

Using species niche for aquaculture purpose implies first to determine which environmental variables must be considered as relevant for aquatic species. These relevant variables can be different between marine and freshwater species. Since current aquaculture is dominated by inland productions [2], we designed our assistant tool for freshwater production development and focused on variables relevant for this specific production. These latter encompass biotic and abiotic variables. Although biotic variables are important shaping factors of species niche, the primary condition for a taxon to be reared in a particular farming environment is its ability to withstand the abiotic conditions. Regardless of biotic components, any aquaculture environment cannot be regarded as potentially suitable for a target species if it does not match with the taxon abiotic requirements. Therefore, we focused on abiotic variables to develop our assistant tool.

In order to select abiotic variables for our assistant tool, we first established an initial list of relevant variables to be considered for various freshwater aquaculture systems (i.e. pond, race-way, recirculating aquaculture system, pen, and cage) according to scientific literature about ecological and physiological requirements of aquatic animal species (e.g. [20–25]) (S1 Table). For all of these variables, a temporal variation occurs in the wild (i.e. seasonality). Therefore, we selected variables reflecting these variabilities (e.g. maximum temperature of the warmest month and minimum temperature of the coldest month). Then, we restricted the initial list according to the data availability of (i) each abiotic variable or (ii) proxy allowing inference of an abiotic variable (Table 1). Since correlative approach implies overlaying occurrence records with locally observed environmental variables, we chose to use only variables for which datasets can be mined from ecological databases with a worldwide coverage (Table 1).

Baseline datasets and data preparation

Abiotic datasets are obtained from three databases, EarthEnv [26], FLO1K [27], and WorldClim [28], as well as from one R-package, *geosphere* [29] (Table 1). Occurrence records are mined from the online Global Biodiversity Information Facility (GBIF, <http://www.gbif.org>) using the R-package *rgbif* [30]. The assistant tool allows users to select which species is/are downloaded by indicating species taxonomic names and returns a proposed correction in case of misspelling based on the GBIF taxonomical baseline. Occurrence data quality issues are minimised [31, 32] by removing potential spatial errors in GBIF using the R-package *CoordinateCleaner* [33]. It eliminates suspicious data (e.g. isolated points, mismatch between coordinates and country code, see discarding criteria in [33]). Since our tool is designed for freshwater aquaculture, points located in the sea are removed.

Table 1. Abiotic variables considered in the assistant tool.

Variable	Unit	Variable description and/or justification	Source
Annual upstream air temperature	°C*10	Proxy for water temperature [26].	Earthenv (WorldClim)
Maximum air temperature of the warmest month	°C*10	Maximum temperature observed in the species' distribution area.	Earthenv (WorldClim)
Minimum air temperature of the coldest month	°C*10	Minimum temperature observed in the species' distribution area.	Earthenv (WorldClim)
Upstream air temperature annual range	°C*10	Temperature long-term variability in the species' distribution area.	Earthenv (WorldClim)
Mean air temperature of the driest quarter	°C*10	Temperature when the thermal inertia is supposed to be minimal.	Earthenv (WorldClim)
Maximum pH of the soil	pH*10	Proxy for the pH of water.	Earthenv (ISRIC)
Annual upstream precipitation	mm	Proxy for parameters with varying concentrations (Ammonia, nitrate. . .).	WorldClim
		Can affect water pH due to runoff or acid rainfalls. Controls water availability in ponds (floods and droughts) and water conductivity.	
		Effects on turbidity [25].	
Upstream precipitation of wettest month	mm	Maximum precipitation observed in the species' distribution area during the wettest month.	WorldClim
Upstream precipitation of driest month	mm	Maximum precipitation observed in the species' distribution area during the driest month.	WorldClim
Slope average	[°] *100	Proxy for flow, water current, and dissolved oxygen concentration.	Earthenv (HydroSHEDS)
Annual average solar radiation	kJ.m ⁻² .day ⁻¹		WorldClim
Annual average water vapor pressure	kPa		WorldClim
Minimum daylength	Hours	Minimum sunshine duration observed in the species' distribution area.	geosphere
Maximum daylength	Hours	Maximum sunshine duration observed in the species' distribution area.	geosphere
Daylength range	Hours	Photoperiod range in the species' distribution area.	geosphere
Average elevation	m	Proxy for temperature, pressure (and dissolved oxygen concentration), streamflow. . .	Earthenv (HydroSHEDS)
Average annual flow	m ³ .s ⁻¹	Proxy for water renewal.	FLOIK
Maximum annual flow	m ³ .s ⁻¹	Maximum water flow observed in the species' distribution area.	FLOIK
Minimum annual flow	m ³ .s ⁻¹	Minimum water flow observed in the species' distribution area.	FLOIK

<https://doi.org/10.1371/journal.pone.0272508.t001>

Characterising niche by correlative approach requires that all occurrence and abiotic datasets share a same coordinate system (WGS84), spatial extent, and geographic scale. Therefore, our R-script limits spatial extent to 56°S to 60°N and 145°W to 180°E, which is the current coverage of EarthEnv. Then, it aggregates datasets at the same resolution grid by (i) recording only species absence/presence per grid cell to account for differences in local sampling effort and to obtain reliable absence data [34] and (ii) merging abiotic dataset by calculating the mean of each abiotic variable within each grid cell, except for daylength values for which the value at the cell centre is attributed to the whole cell. The user can choose to rescale the data either at 10 or 30 arcminutes (~18 or ~55 km at the equator). It is worth noting that the highest resolution might be more precise, but also increases the risk of generating false absence data, which could bias assistant tool results [35]. We recommend using 10 arcminutes only for species with large occurrence point datasets.

Data processing according to user's objectives

Assessing potential suitability of an available farming environment

Assessing whether a particular farming environment is potentially suitable for a species of interest can be done by checking that abiotic conditions of this environment are included in

the species hypervolume. The assistant tool allows performing this assessment for aquaculture facilities relying on local abiotic conditions (i.e. outdoor aquaculture) through three steps. First, species niche hypervolume is computed using the support vector machine method, which is appropriate when dimensionality is high as in our abiotic datasets (R-package *hypervolume*, [36]). Nevertheless, the method accuracy can be negatively impacted if the dimension number exceeds the log of occurrence sum [36]. Therefore, the assistant tool reduces the dimensionality of the initial dataset while minimising collinearity between variables through a principal component analysis (PCA) in which only dimensions with an eigenvalue superior to 1 are retained (i.e. Guttman-Kaiser criterion). Nevertheless, despite this dimensionality reduction, if the dimension number still exceeds the log of occurrence sum, the assistant tool returns a warning to users. During the PCA processing, abiotic variables are standardised. Second, the abiotic conditions corresponding to a given farming location are obtained by extracting the abiotic conditions occurring at the farming location (point or area) with the same procedure as the one to match species occurrences with local abiotic variable values. Alternatively, the user can directly provide values of abiotic variables, a useful option if farming conditions are regulated by farmers and modified from those occurring locally (e.g. use of thermal water from power plant for farming). Third, an inclusion test checks whether the projection of the given farming location/system in the n-dimensional space is included in the species niche hypervolume (R-package *hypervolume*, [36]).

Assisting definition of new farming environment from scratch

When designing a new farming environment for a species from scratch is aimed, knowing abiotic conditions occurring across the distribution range of this species can facilitate the design process. However, a species can undergo quite different abiotic conditions across its distribution range, especially for widespread taxa. This can shape geographic differentiation in abiotic requirements between conspecific populations due to local adaptations (e.g. [37]). This raises the question of which abiotic conditions should be applied for production of a target species. The choice of these conditions will depend on strategy used to create the initial organism stock that will be used in production. Two alternative paradigms are commonly used to create this stock [38, 39]: (i) using only one population or (ii) mixing several geographically distant populations. For the former, the assistant tool can provide the abiotic conditions occurring at the population location. For the latter, we consider that the most commonly observed values of each abiotic variable across the species range are those which are the most likely suitable conditions for most of the populations that will be mixed. The assistant tool highlights these commonly observed conditions by displaying density diagrams for each variable. These diagrams are obtained using the `geom_density` function from the R-package *ggplot2* [40].

Assessing polyculture potential feasibility

The ability of several species to live in the same abiotic environment can be assessed through the intersection between species niche hypervolumes: species with non-null intersection can be potentially co-farmed. Nevertheless, larger intersect means that species can be co-farmed in a larger set of abiotic conditions. This makes their polyculture (i) feasible in a larger set of farming conditions and (ii) still possible in the event of abiotic variations in the farming environments (e.g. extreme climatic events in ponds). Therefore, the assistant tool allows (i) computing the intersection hypervolume between species and (ii) ranking all tested combinations according to the extent of the intersection hypervolume. Since the R-package *hypervolume* [36] did not allow comparing more than two hypervolumes at once, we designed the function `hypervolume_set_n_intersection` to compare n species. The assistant tool also provides the

possibility for users (i) to show the overlap of density diagrams of all species of a particular combination to highlight common suitable abiotic parameters or (ii) to assess that a particular farming facility/location is included in the intersection hypervolume according to procedures detailed previously.

Test cases

We assessed if the assistant tool achieves its intended purpose through *in silico* evaluation of already known instances of polyculture.

Firstly, we selected as test cases three scientific references in which authors studied polyculture in an outdoor aquaculture production (i.e. ponds and cages) at a particular location. Test case 1: a pond system with a polyculture of common carp (*Cyprinus carpio*), roach (*Rutilus rutilus*), and tench (*Tinca tinca*) located at 48.1203°N, 1.7925°W in France [41]. Test case 2: ponds with a polyculture of catla (*Catla catla*), common carp (*C. carpio*), grass carp (*Ctenopharyngodon idella*), Java barb (*Barbonymus gonionotus*), mrigal (*Cirrhinus mrigala*), rohu (*Labeo rohita*), and silver carp (*Hypophthalmichthys molitrix*) in an area located at 23.7000°N–23.9833°N, 88.9167°E–89.0667°E in Bangladesh [42]. Test case 3: earthen ponds with polyculture of African sharptooth catfish (*Clarias gariepinus*) and Nile tilapia (*Oreochromis niloticus*) in an area located at 1.1667°S–1.6000°S, 34.1333°E–35.0167°E in Tanzania. Each test case is hereafter analysed independently. To our knowledge, farming locations of the test cases have not been recorded as species occurrences in GBIF. Therefore, we considered these polyculture farming locations as an effective external dataset to validate the relevance of the assistant tool.

Secondly, we assessed using the assistant tool if species of each test case can be potentially co-farmed. This assessment is based on intersections between the niche hypervolumes of all species from the same test case. We obtained occurrence records from GBIF for each species and abiotic datasets from EarthEnv, FLO1K, WorldClim, and the R-package geosphere (databases accessed on January 22, 2022). Potential occurrence data quality issues were fixed with the R-package CoordinateCleaner. For each species, the number of occurrences were (before/after fixing data quality issues): *B. gonionotus* = 508/69; *C. catla* = 23/0; *C. mrigala* = 426/41; *C. gariepinus* = 4,461/4,140; *C. idella* = 7,459/7,133; *C. carpio* = 179,600/170,826; *H. molitrix* = 5,230/5,179; *L. rohita* = 640/56; *O. niloticus* = 5,188/4,721; *R. rutilus* = 706,001/536,114; *T. tinca* = 145,701/133,242. We thus excluded *C. catla* from the test case 2. Then, we aggregated datasets at the 30 arcminutes resolution. We computed species niche hypervolume after reducing with a PCA and keeping only dimensions with an eigenvalue superior to 1. Since three species (*B. gonionotus*, *C. mrigala*, and *L. rohita*) from the test case 2 had still a dimension number exceeding the log of occurrence sum, we excluded them from the assessment. Then, in each test case, we computed the intersection between niche hypervolumes of all species. For all test cases, the intersection between all species niche hypervolumes was not null, meaning that polyculture of analysed species is feasible according to our approach.

Thirdly, we assessed using the assistant tool if the different polycultures can be performed at locations of each test case. We obtained the abiotic conditions at location of each farming place based on dataset mined from EarthEnv, FLO1K, WorldClim, and the R-package geosphere (databases accessed on January 22, 2022). Since we chose a 30 arcminutes resolution, the used coordinates of farming location were: test case 1 = 48.25°N; -1.75°E; test case 2 = 23.75°N; 88.75°E; test case 3 = -1.25°N; 34.75°E. Then, we performed inclusion tests to assess if each given farming location is included in the intersection hypervolume between all analysed species of each test case. For each test case, farming location was included in the intersection hypervolume. This means that already known instances of polyculture are considered as feasible at their specific location according to the assistant tool. Finally, we displayed

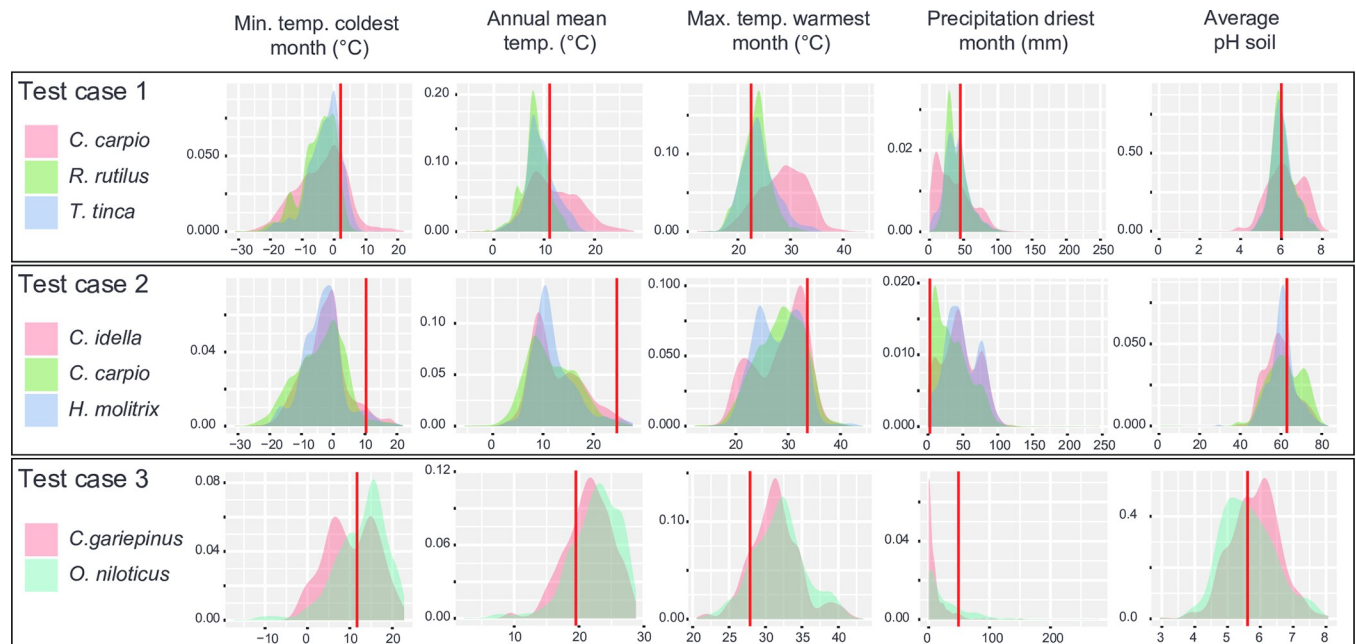


Fig 1. Density plots showing the distribution of a set of abiotic parameters based on occurrences used to determine species niche hypervolumes for the three test cases. Abiotic parameters are in columns: Min. temp. coldest month = Minimal temperature of the coldest month; Annual mean temp. = Annual mean temperature; Max. temp. warmest month = Maximal temperature of the warmest month; Precipitation driest month = Precipitation of the driest month; Average pH soil = Average pH observed in the soil. Other abiotic parameters were considered by the assistant tool (see Table 1), but only a subset of abiotic parameters is shown in the figure. Test cases are in lines. For each test case, a colour chart identifies density diagrams of each analysed species. The analysed species are *Clarias gariepinus*, *Ctenopharyngodon idella*, *Cyprinus carpio*, *Hypophthalmichthys molitrix*, *Oreochromis niloticus*, *Rutilus rutilus*, and *Tinca tinca*. The red lines correspond to position of the farming site of each test case in abiotic density plots.

<https://doi.org/10.1371/journal.pone.0272508.g001>

the density diagrams of all analysed species for key abiotic parameters along with the position of the farming site in these diagrams for each analysed species of each test case (Fig 1).

Limits and prospects

Overall, three main limitations of our assistant tool should be stated.

Firstly, although GBIF hosts occurrence datasets for several tens of thousands of fish species (*sensu* all taxa included in Myxini [hagfishes], Cephalaspidomorphi [lampreys], Holocephali [chimaeras], Elasmobranchii [sharks and rays], Sarcopterygii [lobe-finned fishes], and Actinopterygii [ray-finned fishes]), the quantity of information is uneven across species (see test cases). Special attention should be paid to quantity, distribution, and density of occurrences available for each species prior to using the assistant tool. Indeed, low occurrence dataset quality can result in biased assessment for the three here above described method applications. Nevertheless, assuming that the assistant tool will be mainly used for species of interest, which have thus been the focus of an abundant research, quantitative and qualitative datasets is likely to be available for users.

Secondly, using correlative approach for niche analyses allows estimating realised niche (i.e. all the environmental conditions where the species currently lives) rather than fundamental niche (i.e. the full range of environmental conditions in which a species is able to live, without any other limiting factors that could constrain the species distribution). Therefore, we cannot rule out that assistant tool results would be too restrictive due to an underestimation of species fundamental niche. However, determining fundamental niche is still hard-to-achieve without

detailed knowledge on species (e.g. [43]). We argue thus that the realised niche remains the most operational approach to propose a pragmatic tool for aquaculture development.

Finally, we designed our assistant tool according to worldwide biogeographic and abiotic databases currently available. It implied to (i) discard potentially key abiotic variables (S1 Table) due to their poor availability at large geographic scale or to (ii) use proxies which are only partially correlated to key variables. This can potentially decrease the relevance of our approach. However, we speculate that future development of databases will make available more abiotic variables which will be integrated in our approach and will reinforce the relevance of assessment performed by our assistant tool.

The current version of our assistant tool mainly allows assessing or designing farming environments relying on local abiotic conditions for freshwater productions. However, it can be applied to facilitate the development of indoor highly human-controlled systems (e.g. indoor recirculated aquaculture systems) or marine species productions provided that the list of abiotic variables is adjusted to the farming system specificities (e.g. altitude is not a relevant proxy for marine production development) and to the taxon group. Beside aquaculture purpose, our assistant tool could be applied for further purposes such as *ex-situ* conservations of aquatic species in order to design hosting facilities for endangered taxa.

Availability

The current stable version of the package requires R 4.1.1 and can be downloaded from github (<https://github.com/GregoireButruille/AquaDesign>). The R-script is available in figshare [17]. A tutorial of the assistant tool is available in figshare [18].

Supporting information

S1 Table. Abiotic component selected as relevant considering environmental requirements and physiology of aquatic organisms.

(XLSX)

Acknowledgments

We are grateful to Benjamin Blonder (University of California, Berkeley, USA) for his help and advice for development of data processing and to Denis Toomey for the English proof-reading.

Author Contributions

Conceptualization: Grégoire Butruille, Marielle Thomas, Alain Pasquet, Nellya Amoussou, Lola Toomey, Sandrine Chauchard, Thomas Lecocq.

Formal analysis: Marielle Thomas, Alain Pasquet, Sandrine Chauchard.

Funding acquisition: Marielle Thomas, Alain Pasquet, Thomas Lecocq.

Methodology: Grégoire Butruille, Alain Pasquet, Nellya Amoussou, Lola Toomey, Sandrine Chauchard, Thomas Lecocq.

Project administration: Marielle Thomas, Thomas Lecocq.

Software: Grégoire Butruille, Thomas Lecocq.

Supervision: Marielle Thomas, Sandrine Chauchard, Thomas Lecocq.

Validation: Grégoire Butruille, Marielle Thomas, Alain Pasquet, Nellya Amoussou, Lola Toomey, Axel Rosenstein, Sandrine Chauchard, Thomas Lecocq.

Writing – original draft: Grégoire Butruille, Thomas Lecocq.

Writing – review & editing: Grégoire Butruille, Marielle Thomas, Alain Pasquet, Nellya Amoussou, Lola Toomey, Axel Rosenstein, Sandrine Chauchard, Thomas Lecocq.

References

1. Foley JA, Ramankutty N, Brauman KA, Cassidy ES, Gerber JS, Johnston M, et al. Solutions for a cultivated planet. *Nature*. 2011; 478: 337–342. <https://doi.org/10.1038/nature10452> PMID: 21993620
2. FAO. The State of World Fisheries and Aquaculture 2020. Rome, Italy: Food and Agriculture Organization of the United Nations—Licence: CC BY-NC-SA 3.0 IGO; 2020. <https://doi.org/10.4060/ca9229en>
3. Diana JS, Egna HS, Chopin T, Peterson MS, Cao L, Pomeroy R, et al. Responsible Aquaculture in 2050: Valuing Local Conditions and Human Innovations Will Be Key to Success. *Bioscience*. 2013; 63: 255–262. <https://doi.org/10.1525/bio.2013.63.4.5>
4. Thomas M, Pasquet A, Aubin J, Nahon S, Lecocq T. When more is more: taking advantage of species diversity to move towards sustainable aquaculture. *Biol Rev*. 2021; 96: 767–784. <https://doi.org/10.1111/brv.12677> PMID: 33320418
5. Godoy MG, Aedo A, Kibenge MJ., Groman DB, Yason C V, Grothusen H, et al. First detection, isolation and molecular characterization of infectious salmon anaemia virus associated with clinical disease in farmed Atlantic salmon (*Salmo salar*) in Chile. *BMC Vet Res*. 2008; 4: 28. <https://doi.org/10.1186/1746-6148-4-28> PMID: 18680586
6. Medeiros M V., Aubin J, Camargo AFM. Life cycle assessment of fish and prawn production: Comparison of monoculture and polyculture freshwater systems in Brazil. *J Clean Prod*. 2017; 156: 528–537. <https://doi.org/10.1016/j.jclepro.2017.04.059>
7. Metian M, Troell M, Christensen V, Steenbeek J, Pouil S. Mapping diversity of species in global aquaculture. *Rev Aquac*. 2020; 12: 1090–1100. <https://doi.org/10.1111/raq.12374>
8. Harvey B, Soto D, Carolsfeld J, Beveridge M, Bartley DM. Planning for aquaculture diversification: the importance of climate change and other drivers. FAO Fish Aquac Proc No 47. Rome, Italy: Fisheries and Aquaculture Proceedings, FAO; 2017.
9. Teletchea F, Fontaine P. Levels of domestication in fish: implications for the sustainable future of aquaculture. *Fish Fish*. 2014; 15: 181–195. <https://doi.org/10.1111/faf.12006>
10. Pocheville A. The Ecological Niche: History and Recent Controversies. In: Heams T, Huneman P, Lecointre G, Silberstein M, editors. *Handbook of Evolutionary Thinking in the Sciences*. Dordrecht, Netherlands: Springer Netherlands; 2015. pp. 547–586. <https://doi.org/10.13140/RG.2.1.3205.8405>
11. Hutchinson GE. Concluding remarks. *Cold Spring Harb Symp Quant Biol*. 1957; 22: 415–427. <https://doi.org/10.1101/SQB.1957.022.01.039>
12. O’Neil ST, Coates PS, Brussee BE, Ricca MA, Espinosa SP, Gardner SC, et al. Wildfire and the ecological niche: Diminishing habitat suitability for an indicator species within semi-arid ecosystems. *Glob Chang Biol*. 2020; 26: 6296–6312. <https://doi.org/10.1111/gcb.15300> PMID: 32741106
13. Kolanowska M, Rewicz A, Baranow P. Ecological niche modeling of the pantropical orchid *Polystachya concreta* (Orchidaceae) and its response to climate change. *Sci Rep*. 2020; 10: 14801. <https://doi.org/10.1038/s41598-020-71732-1> PMID: 32908206
14. Eckert S, Hamad A, Kilawe CJ, Linders TEW, Ng W, Mbaabu PR, et al. Niche change analysis as a tool to inform management of two invasive species in Eastern Africa. *Ecosphere*. 2020; 11. <https://doi.org/10.1002/ecs2.2987>
15. Sohn N, Fernandez MH, Papes M, Anciães M. Ecological niche modeling in practice: Flagship species and regional conservation planning. *Oecologia Aust*. 2013; 17: 429–440. <https://doi.org/10.4257/oeco.2013.1703.11>
16. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
17. Lecocq T, Butruille G, Thomas M, Pasquet A. R-script to design of new aquaculture production based on animal species abiotic requirements. 2021. <https://doi.org/10.6084/m9.figshare.17170958.v3>
18. Lecocq T, Butruille G. A tutorial to design farming system for aquaculture (monoculture and polyculture) with AquaDesign. 2021. <https://doi.org/10.6084/m9.figshare.17170979.v2>

19. Guisan A, Zimmermann NE. Predictive habitat distribution models in ecology. *Ecol Modell.* 2000; 135: 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
20. Bhatnagar A, Devi P. Water quality guidelines for the management of pond fish culture. *Int J Environ Sci.* 2013; 3: 1980–2009.
21. Boeuf G, Le Bail P-Y. Does light have an influence on fish growth? *Aquaculture.* 1999; 177: 129–152. [https://doi.org/10.1016/S0044-8486\(99\)00074-5](https://doi.org/10.1016/S0044-8486(99)00074-5)
22. Boyd CE. Water quality management for pond fish culture. Amsterdam, The Netherlands: Elsevier Scientific Publishing Co.; 1982.
23. Colt J. Water quality requirements for reuse systems. *Aquac Eng.* 2006; 34: 143–156. <https://doi.org/10.1016/j.aquaeng.2005.08.011>
24. Jackson DA, Peres-Neto PR, Olden JD. What controls who is where in freshwater fish communities—the roles of biotic, abiotic, and spatial factors. *Can J Fish Aquat Sci.* 2001; 58: 157–170. <https://doi.org/10.1139/cjfas-58-1-157>
25. Ahmed N, Thompson S, Glaser M. Global Aquaculture Productivity, Environmental Sustainability, and Climate Change Adaptability. *Environ Manage.* 2019; 63: 159–172. <https://doi.org/10.1007/s00267-018-1117-3> PMID: 30460481
26. Domisch S, Amatulli G, Jetz W. Near-global freshwater-specific environmental variables for biodiversity analyses in 1 km resolution. *Sci Data.* 2015; 2: 150073. <https://doi.org/10.1038/sdata.2015.73> PMID: 26647296
27. Barbarossa V, Huijbregts MAJ, Beusen AHW, Beck HE, King H, Schipper AM. FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015. *Sci Data.* 2018; 5: 180052. <https://doi.org/10.1038/sdata.2018.52> PMID: 29583139
28. Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol.* 2017; 37: 4302–4315. <https://doi.org/10.1002/joc.5086>
29. Hijmans RJ. geosphere: Spherical Trigonometry. R package version 1.5–10. 2019.
30. Chamberlain S, Barve V, Mcglinn D, Oldoni D, Desmet P, Geffert L, et al. rgbif: Interface to the Global Biodiversity Information Facility API—R package version 3.6.0. 2021. Available: <https://cran.r-project.org/package=rgbif%3E>
31. Graham CH, Elith J, Hijmans RJ, Guisan A, Townsend Peterson A, Loiselle BA. The influence of spatial errors in species occurrence data used in distribution models. *J Appl Ecol.* 2007; 45: 239–247. <https://doi.org/10.1111/j.1365-2664.2007.01408.x>
32. Liu C, White M, Newell G. Detecting outliers in species distribution data. *J Biogeogr.* 2018; 45: 164–176. <https://doi.org/10.1111/jbi.13122>
33. Zizka A, Silvestro D, Andermann T, Azevedo J, Duarte Ritter C, Edler D, et al. CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods Ecol Evol.* 2019; 10: 744–751. <https://doi.org/10.1111/2041-210X.13152>
34. Zhang Z, Xu S, Capinha C, Weterings R, Gao T. Using species distribution model to predict the impact of climate change on the potential distribution of Japanese whiting *Sillago japonica*. *Ecol Indic.* 2019; 104: 333–340. <https://doi.org/10.1016/j.ecolind.2019.05.023>
35. Tyre AJ, Tenhumberg B, Field SA, Niejalke D, Parris K, Possingham HP. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. *Ecol Appl.* 2003; 13: 1790–1801. <https://doi.org/10.1890/02-5078>
36. Blonder B, Morrow CB, Maitner B, Harris DJ, Lamanna C, Violle C, et al. New approaches for delineating n-dimensional hypervolumes. McMahon S, editor. *Methods Ecol Evol.* 2018; 9: 305–319. <https://doi.org/10.1111/2041-210X.12865>
37. Yu D, Zhang Z, Shen Z, Zhang C, Liu H. Regional differences in thermal adaptation of a cold-water fish *Rhynchocypris oxycephalus* revealed by thermal tolerance and transcriptomic responses. *Sci Rep.* 2018; 8: 11703. <https://doi.org/10.1038/s41598-018-30074-9> PMID: 30076386
38. Gjedrem T. Breeding Plans. In: Gjedrem T, editor. *Selection and Breeding Programs in Aquaculture.* Berlin/Heidelberg: Springer, Dordrecht; 2005. pp. 251–277. https://doi.org/10.1007/1-4020-3342-7_16
39. Toomey L, Fontaine P, Lecocq T. Unlocking the intraspecific aquaculture potential from the wild biodiversity to facilitate aquaculture development. *Rev Aquac.* 2020; 12: 2212–2227. <https://doi.org/10.1111/raq.12430>
40. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Cham: Springer International Publishing; 2016. <https://doi.org/10.1007/978-3-319-24277-4>
41. Jaeger C, Aubin J. Ecological intensification in multi-trophic aquaculture ponds: an experimental approach. *Aquat Living Resour.* 2018; 31: 36. <https://doi.org/10.1051/alr/2018021>

42. Asadujjaman M, Hossain MA. Fish growth, yield and economics of conventional feed and weed based polyculture in ponds. *J Fish.* 2016; 4: 353. <https://doi.org/10.17017/jfish.v4i1.2016.170>
43. Jiménez L, Soberón J. Estimating the fundamental niche: accounting for the uneven availability of existing climates. *bioRxiv.* 2021; 2021.01.25.428165. <https://doi.org/10.1101/2021.01.25.428165>