

# Revisiting a central prediction of the Gill Oxygen Limitation Theory: Gill area index and growth performance

Jennifer S. Bigman<sup>1</sup>  | Nicholas C. Wegner<sup>2</sup>  | Nicholas K. Dulvy<sup>1</sup> 

<sup>1</sup>Earth to Ocean Research Group, Simon Fraser University, Burnaby, British Columbia, Canada

<sup>2</sup>Fisheries Resources Division, Southwest Fisheries Science Center, NOAA, La Jolla, California, USA

## Correspondence

Jennifer S. Bigman, Earth to Ocean Research Group, Simon Fraser University, Burnaby, British Columbia, Canada.  
Email: [jennifersbigman@gmail.com](mailto:jennifersbigman@gmail.com)

## Present address

Jennifer S. Bigman, Alaska Fisheries Science Center, NOAA, Seattle, Washington, USA

## Funding information

National Science Foundation, Grant/Award Number: 2109411; Natural Sciences and Engineering Research Council of Canada

## Abstract

The Gill Oxygen Limitation Theory (GOLT) posits that a mismatch in oxygen supply and demand stemming from geometric constraints on gill surface area limits metabolic rate and energy available for biological processes. This theory has been suggested to explain numerous phenomena observed with warming yet is based upon a relationship among maximum size, growth, and gill surface area established over 40 years ago. However, the metric used in this relationship to characterize gill surface area, gill area index, fails to capture the known variability in the scaling of gill surface area and is biased by the sizes at which gills were measured. Here, we revisit a central prediction of the GOLT, asking four key questions that examine limitations in the original relationship. We find that gill area index does indeed explain variation in growth performance across 132 species of fish and this relationship is strikingly similar to the original relationship across 42 species. Yet, we argue that gill area index is not an adequate measure of gill surface area because (1) gill surface area has a non-linear relationship with size and, thus, changes ontogenetically as an individual grows over time and (2) because it is based on mean estimates of both gill surface area and body mass. Indeed, we show that the value of gill area index for a given species is variable depending on how it is calculated. We therefore suggest a pathway forward for assessing whether gill surface area is an important factor in explaining variation in growth performance.

## KEYWORDS

Bayesian hierarchical modelling, ecophysiology, Jensen's inequality, life history theory, metabolic ecology, von Bertalanffy growth

## 1 | INTRODUCTION

The Gill Oxygen Limitation Theory (GOLT) posits that in aquatic, water-breathing ectotherms, oxygen uptake at the gills limits the aerobic metabolic rate and ultimately, growth and other processes that rely on the energy produced by aerobic metabolism (Pauly, 1981, 2010, 2021). The basis of this theory is that the surface area of the gills (as a two-dimensional surface) cannot grow as fast as the body it must supply with oxygen (a three-dimensional volume;

Pauly, 1981, 2010, 2021). In other words, the ontogenetic scaling of gill surface area and body mass will always be less than one, resulting in a mismatch between oxygen supply and demand as an organism increases in size (Pauly, 1981, 2010, 2021). Thus, the ratio of gill surface area to body mass will decrease throughout an organism's lifetime and, eventually, will not be able to match the demand of a growing body, at which point the maximum size of the organism will be reached (Pauly, 1981, 2010, 2021). Because this theory posits that gill surface area constrains aerobic metabolic rate, and

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Fish and Fisheries* published by John Wiley & Sons Ltd.

thus processes related to or relying on metabolism, energy, and oxygen, it is multi-faceted and explicitly and implicitly generates a range of predictions. Such predictions include those surrounding growth and other aspects of life history and ecology (e.g. maximum size, timing of maturation and reproduction, spawning phenology, geographic distributions, activity level) and those based more on physiological processes (e.g. food consumption and conversion efficiency, the balance of oxidative versus glycolytic enzymes; Pauly, 1981, 2010, 2021; Pauly & Liang, 2022). Although originally proposed in the early 1980s, interest in the GOLT has experienced a recent resurgence in light of research that aims to predict how species will respond to continued environmental change, particularly warming temperatures and shifting oxygen availability (Cheung et al., 2013; Lefevre et al., 2017, 2018; Seibel & Deutsch, 2020). In particular, the maximum body size of fishes is expected to decline, or 'shrink,' as temperatures rise due to, in part, the proposed mismatch in oxygen supply and demand at the gills as predicted by the GOLT (Cheung et al., 2013; Pauly & Cheung, 2016).

A central prediction of the GOLT is the relationship between two indices: growth performance,  $\phi$ , and gill area index (Pauly, 1981, 2010, 2021). Growth performance is an index integrating the life history and mathematical trade-off between growth and maximum size and is calculated as  $\log_{10}(k * W_{\infty})$ , where  $k$  is the Brody growth coefficient and  $W_{\infty}$  is asymptotic size, both from a von Bertalanffy growth function (Juan-Jordá et al., 2013; Pauly, 1981, 1991). The gill area index is somewhat similar to the intercept of the ontogenetic relationship of gill surface area and body mass (the predicted gill surface area for a given body size resulting from a regression equation; Pauly, 1981, 2010). However, gill area index is not estimated from a regression relationship but calculated as  $G/W^d$ , where  $G$  = an estimate of a species-specific mean gill surface area,  $W$  = mean body mass estimate associated with the mean gill surface area estimate, and  $d$  is a scaling parameter that would ideally be the species-specific ontogenetic slope of the relationship between body mass and gill surface area (Pauly, 1981, 2010). The parameter  $d$  is included because gill surface area largely scales disproportionately with body mass (i.e. usually  $<1$ ), such that the ratio of gill surface area to body mass for a single individual changes throughout its lifetime (and thus is why a mean, relative [gill surface area at a given body mass or that predicted from a regression equation at a specific body mass], or mass-specific [gill surface area per gram of body mass] is not an ideal metric of gill surface area; Bigman et al., 2018; De Jager & Dekkers, 1975; Palzenberger & Pohla, 1992; Wegner, 2011). To calculate gill area index ( $G/W^d$ ), Pauly (1981) used a compilation from Hughes and Morgan (1973) that reported mean estimates of gill surface area and body mass data measured from a random sample of individuals for a given species (Table S1). These data were restricted to marine fishes with published von Bertalanffy growth parameters, yielding a total of 42 species. Gill area index was then calculated for each of these species using mean gill surface area, mean body mass data and (due to the paucity of individual gill surface area and body mass data for a given species) a predicted value of  $d$  (Pauly, 1981, 2010). For each species, Pauly (1981) predicted  $d$  from a previously estimated linear relationship between

1 INTRODUCTION	354
2 METHODS	356
2.1 How sensitive is the relationship between gill area index and growth performance to outliers?	356
2.2 What is the effect of flipping the axes and testing whether gill area index explains variation in growth performance?	357
2.3 Does the original gill area index and growth performance relationship still hold when examined across more species?	357
2.3.1 Additional data collection and sources	357
2.3.2 Gill surface area data	357
2.3.3 Life history data	358
2.3.4 Estimating gill area index and growth performance	358
2.3.5 Statistical analysis	358
2.4 How do the different ways of calculating gill area index affect the relationship between gill surface area and growth performance?	358
2.4.1 Data filtering	358
2.4.2 Statistical analysis	358
3 RESULTS	359
3.1 How sensitive is the relationship between gill area index and growth performance to outliers?	359
3.2 What is the effect of flipping the axes and testing whether gill area index explains variation in growth performance?	361
3.3 Does the original gill area index and growth performance relationship still hold when examined across more species?	361
3.4 How do the different ways of calculating gill area index affect the relationship between gill surface area and growth performance?	363
4 DISCUSSION	363
ACKNOWLEDGEMENTS	365
DATA AVAILABILITY STATEMENT	365
CONFLICT OF INTERESTS	365
REFERENCES	365

the ontogenetic slope of a relationship between body mass and gill surface area or metabolic rate versus maximum observed body mass ( $W_{\max}$ ) for 27 species or genera of fishes, as well as average values of  $d$  for (1) 'all freshwater fishes', (2) 'all marine fishes', (3) two different average values for 'fishes' and (4) 'Gray's intermediates [various marine teleosts]' (Pauly, 1981). Hereafter, we call this value 'predicted  $d$ ' as a value of  $d$  was calculated for each species using their maximum body mass and the established regression equation,  $d = 0.6742 + 0.03574 * \log(\text{maximum body mass})$  (see p. 264 in Pauly, 1981). Later, Pauly (2010) estimated gill area index using a constant value of  $d$  for all species,  $d = 0.8$ , instead of predicting it based on a species' maximum body mass. Hence, the gill area indices originally used by Pauly (1981)

and Pauly (2010) were likely biased by the sizes at which gills were measured (i.e. the sizes used to generate the species' means), the prediction of the parameter  $d$  and the assumption that the slope of gill surface area was constant across a broad array of species (an assumption we now know is false, Bigman et al., 2018; Jager & Dekkers, 1975; Palzenberger & Pohla, 1992; Wegner, 2011).

The relationship between growth performance (as the predictor variable) and gill area index (as the response variable) was fitted with a reduced major axis regression (called 'functional regression' in Pauly (1981)). A significant correlation between growth performance and gill area index, with  $r = .431$  (equivalent to an  $r^2$  of .19) and a  $p = .01$  [i.e. the slope does not equal zero] was reported across 40 species (with two species deemed outliers and removed from analysis; Pauly, 1981). Three additional species were determined to be outliers upon visual inspection and were removed from the analysis. After doing so, it was reported that the removal of these outliers, to arrive at 37 species, 'greatly improves the correlation, which increases to  $r = .661$  [equivalent to an  $r^2$  value of .44]' (Pauly, 1981). Notably, one key assumption of using reduced major axis regression is that the relationship between the predictor(s) and response variables is symmetric (McArdle, 2003; Smith, 2009). Thus, it does not matter which variable is the predictor and which is the response as the resulting coefficients and relationship will be identical (McArdle, 2003; Smith, 2009). However, the GOLT makes explicit predictions about the direction of causality between gill area index and growth performance (Pauly, 1981, 2010). Specifically, this theory argues that gill surface area is constraining growth and maximum size (Pauly, 1981, 2010). As such, reduced major axis regression is likely not an ideal method for assessing the relationship between gill area index and growth performance in the context of the GOLT and thus other regression methods are likely better suited for such biological questions. For example, ordinary least squares regression (OLS) makes an explicit hypothesis regarding the directionality of the relationship between the predictor and response variables and robust regression is a technique that can deal with the effect of leverage from data points, or outliers, on the model fit (Gelman & Hill, 2007; Hampel, 2001; Kruschke, 2015).

Here, we reanalyse a central prediction of the GOLT – the relationship between growth performance and gill area index. We ask four key questions that examine limitations and assumptions in the original analysis. Additionally, we incorporate gill surface area and body mass data across more species that have become available in the 40 years since this relationship was first established. Specifically, we ask (1) how sensitive is the relationship between gill area index and growth performance to outliers, (2) what is the effect of parameterizing the relationship between gill area index and growth performance based on the predictions made by the GOLT (i.e. flipping the axes and testing whether gill area index can explain variation in growth performance), (3) does the gill area index and growth performance relationship still hold when examined across more species and (4) how do the different ways of calculating gill area index affect the relationship between gill surface area and growth performance?

## 2 | METHODS

Because of the different data and methods associated with each of the four questions, we outline the methods for each question separately below.

### 2.1 | How sensitive is the relationship between gill area index and growth performance to outliers?

A statistical outlier can be defined as an outlying or extreme observation, one that appears to deviate markedly from other members of the sample, fall unusually far from the expected value based on the model or greatly influences model results (Fahrmeir et al., 2013; Gelman & Hill, 2007). Although checking for outliers is common practice, standardized methods across fields to identify and deal with outliers are rare (Burnham & Anderson, 2002; Fahrmeir et al., 2013; Hampel, 2001). Historically, outliers have often been removed from data sets to facilitate modelling by OLS or similar methods (e.g. reduced major axis regression), but because there is often no way to non-arbitrarily remove outliers, it is more commonly recommended to instead refine the model to accommodate outliers (Hampel, 2001; Kruschke, 2015). One way of identifying outliers is to use model diagnostics such as Cook's distance for frequentist models and Pareto  $k$  for Bayesian models (Fahrmeir et al., 2013; Gabry et al., 2019; Vehtari et al., 2017). If the value of a data point's Cook's distance or Pareto  $k$  is above the threshold (0.5 for Cook's distance, 0.7 for Pareto  $k$ ), it is recommended to employ robust regression, a class of regression models that relax the assumption of normality that is characteristic of the most common OLS regression models (Gelman & Hill, 2007; Kruschke, 2015; Vehtari et al., 2017). Robust regression downweights influential data points or employs a fat-tailed distribution, for example Student's  $t$ , to model the response distribution (Anderson et al., 2017; Gelman & Hill, 2007; Wang & Blei, 2018).

While many ways to estimate a robust regression exist, common frequentist methods are quantile regression and iteratively reweighted least squares regression (Fahrmeir et al., 2013; Fox & Weisberg, 2012; Rousseeuw, 1984). As used in our analysis, quantile regression simply models the median of the response variable as a linear function of the predictor variables (instead of the mean, as in OLS; Fahrmeir et al., 2013; Fox & Weisberg, 2012; Rousseeuw, 1984). Iteratively reweighted least squares regression downweights outliers according to the distance from the best-fit line and iteratively refits the model (Fox & Weisberg, 2012; Rousseeuw, 1984). In a Bayesian framework, robust regression simply involves allowing the response distribution to be less restrictive; in our case, this would entail using a Student's  $t$  distribution (Gelman et al., 2020; Lange et al., 1989; Wang & Blei, 2018). The Student's  $t$  distribution is a normal distribution but with the degrees of freedom parameter,  $nu$ , set to infinity;  $nu$  can either be estimated from the model directly or set to a specific value (Wang & Blei, 2018).

For this first question, we used model diagnostics to identify possible outliers in the original data set used by Pauly (1981; hereafter, 'Pauly data set'). To do so, we used Cook's distance and Pareto  $k$ , both of which are measures of the influence of a given observation on the model. Cook's distance values were estimated using OLS and Pareto  $k$  values were estimated using Bayesian simple linear regression (Fahrmeir et al., 2013; Vehtari et al., 2017). We then compared regression coefficients estimated by different methods of linear and robust linear regression to Pauly's (1981) original gill area index to growth performance relationship (which was estimated via a reduced major axis regression on 37 and 40 data points [species]). To do so, we fit several models using R (R Core Team, 2020): (1) we re-estimated the reduced major axis regression for the 37 and 40 data points as reported in Pauly (1981, 2010), (2) we fit a reduced major axis regression with all 42 species (*lmodel2* function in the *lmodel2* package; Legendre & Oksanen, 2018), and (3) we fit four types of robust regression models: (a) quantile regression (*rq* function in the *quantreg* package; Koenker, 2020), (b) iteratively reweighted least squares regression (*rlm* function in the *MASS* package; Venables & Ripley, 2002), and (c and d) Bayesian regression with a Student's  $t$  distribution (i.e. robust Bayesian regression; *brm* function in the *brms* package; Bürkner, 2017, 2018) with two priors that differed in the strength of the prior on the degrees of freedom parameter ( $\nu$ ), one model with a strong prior and the other with a weakly informative prior (Vehtari et al., 2017).

## 2.2 | What is the effect of flipping the axes and testing whether gill area index explains variation in growth performance?

Using the Pauly data set, we ask how the relationship between gill area index and growth performance would differ if parameterized according to the prediction of the GOLT. As this is a different model, we first estimated both Cook's distance and Pareto  $k$  (see Question 1) to identify any outliers using OLS and Bayesian linear regression, respectively, as outliers, as well as other diagnostics should be checked for any and all model runs. If any outliers were identified (Cook's distance  $>0.5$  or a Pareto  $k$  value  $>0.7$ ), we used robust regression in a Bayesian framework to estimate the slope and intercept as Bayesian regression offers more information than frequentist regression and thus we preferentially choose to use it when possible. As in Question 1, we used the *brm* function in the *brms* package to estimate two Bayesian linear models with a Student's  $t$  response distribution, one with a strong prior on  $\nu$  and one with a weak prior on  $\nu$  (Bürkner, 2017, 2018). Pareto smoothed importance sampling leave-one-out cross validation (PSIS-LOO) was used to compare the two Bayesian models with different priors on  $\nu$  to identify which prior provided the best fit to the data (Vehtari et al., 2017). Finally, we computed an  $r^2$  value using the *bayes\_R2* function in the *brms* package for the purpose of comparing it with Pauly's reported  $r^2$  values (Bürkner, 2017, 2018).

## 2.3 | Does the original gill area index and growth performance relationship still hold when examined across more species?

### 2.3.1 | Additional data collection and sources

We compiled a data set of additional fishes (teleost, elasmobranch and coelacanth) for which gill surface area, body mass associated with gill surface area (i.e. the body mass of the individual whose gill surface area was measured, hereafter 'measurement body mass'), and von Bertalanffy growth parameters were available. Data were first collated for species with estimates of gill surface area and available growth parameters in Fishbase (Froese & Pauly, 2020). This was further supplemented with published gill surface area data from other sources if a given species also had available growth parameters from Bigman et al. (2021), Jager and Dekkers (1975), Gray (1954), Hughes and Morgan (1973) and Palzenberger and Pohla (1992).

### 2.3.2 | Gill surface area data

Gill surface area estimates,  $\text{cm}^2$  or  $\text{mm}^2$ , and measurement body mass were extracted from the original study in which they were reported, if possible, otherwise were extracted from Fishbase, which was the case for three species. Prior to analyses, all gill surface area data were converted to  $\text{cm}^2$ . Both raw, that is estimates for multiple individuals of a species, and mean gill surface area data were included in our expanded data set. If more than one study reported raw data for a number of individuals for a given species, we included both data sets (this was only the case for three species: Common Thresher Shark (*Alopias vulpinus*, Alopiidae), Sandbar Shark (*Carcharhinus plumbeus*, Carcharhinidae) and Shortfin Mako (*Isurus oxyrinchus*, Lamnidae)). If a given species had both published raw and mean data, we preferentially chose the study that included raw data. All raw data were averaged per species to generate a species-specific mean of gill surface area and body mass for calculating gill area index. If more than one study reported mean data (this was the case for four species), we chose the study with the largest sample size. Any gill surface area estimate that was not directly measured (e.g. predicted from theoretical geometric relationships) was not included in this study (for further discussion, see Satora & Wegner, 2012). Additionally, the 42 species in the Pauly data set were included in our expanded data set, hereafter, 'full data set', with the exception of four species for which the gill surface area data could not be verified: European Anchovy [*Engraulis encrasicolus*, Engraulidae], Lined Seahorse [*Hippocampus hudsonicus* = *Hippocampus erectus*, Syngnathidae] and Black Scorpionfish [*Scorpaena porcus*, Scorpaenidae], or the gill surface area was predicted from a regression equation and not empirically measured (Spiny Dogfish [*Squalus acanthias*, Squalidae]). For the remaining 38 species, 25 of these did not have more recent or higher quality gill surface area data available (e.g. larger sample size, raw data) and thus the original data used by Pauly (1981) from Hughes and Morgan (1973) were included in our data set. For the remaining

13 species, either raw data were acquired or as the Pauly data set included averaged mean gill surface area data (i.e. means of means), only the mean from the study with the largest sample size was included in the full dataset.

### 2.3.3 | Life history data

Using the 'rfishbase' package for Fishbase, we extracted all observations of von Bertalanffy growth function parameters for each species in our dataset including  $k$  ( $\text{year}^{-1}$ ), the growth coefficient, and  $W_{\infty}$  (g) and  $L_{\infty}$  (cm), the asymptotic mass or length individuals in a population would reach if they were to grow indefinitely (Boettiger et al., 2012; Froese & Pauly, 2020). As most von Bertalanffy growth functions are estimated in terms of length and not weight, we used length-weight regressions (also from Fishbase) to convert  $L_{\infty}$  to  $W_{\infty}$  (Boettiger et al., 2012; Froese & Pauly, 2020).  $L_{\infty}$  was converted to  $W_{\infty}$  based on species-, length type- (total length [TL], fork length [FL]) and sex-specific length-weight regressions. If growth data were not available in Fishbase for a species, the primary literature was searched for published age and growth data. For eight species, growth parameters were not available in Fishbase but were found in the literature. For seven species, sex-specific length-weight coefficients were not available for sex-specific growth parameters, and so available length-weight coefficients were averaged and then used to estimate  $W_{\infty}$ . For 14 species, length-weight coefficients for the same length type as was used to estimate growth parameters were not available (i.e. the growth coefficient was estimated using fork length but no length-weight regression for fork length to weight or conversion from another length type was available), and thus matching type-specific length-weight regressions were collated from the literature.

### 2.3.4 | Estimating gill area index and growth performance

Following Pauly (1981, 2010) we estimated gill area index for each species in the full dataset ( $n = 132$ ). For simplicity in this question, we opted to use a constant value of  $d = 0.8$  for all species as used in Pauly (2010). We then re-estimated gill area index for the 42 species in the Pauly dataset using  $d = 0.8$  as these values of gill area index were not reported in Pauly (2010) and Pauly (1981) used the predicted  $d$ . Growth performance was calculated for each observation as  $\log_{10}(k * W_{\infty})$  following Pauly (1981). For analyses, a mean of growth performance was taken for each species.

### 2.3.5 | Statistical analysis

We used OLS regression estimated in a Bayesian framework using the *brm* function in the *brms* package to estimate regression coefficients for the relationship between gill area index and growth

performance for the Pauly dataset ( $n = 42$ ) and the full dataset (Bürkner, 2017, 2018; R Core Team, 2020). For these models, growth performance was the response variable and gill area index was the predictor variable. We examined the existence of outliers in all models using Cook's distance and Pareto  $k$ .

## 2.4 | How do the different ways of calculating gill area index affect the relationship between gill surface area and growth performance?

For this question, we assessed how the calculation of gill area index affected the relationship between growth performance and gill area index. We used the two ways gill area index has been calculated by Pauly (1981) and Pauly (2010) but restricted the data set to those species that have raw gill surface area and measurement body mass data, which allowed us to calculate gill area index as originally intended, where  $d$  is the empirically estimated slope of a species-specific ontogenetic allometry. Specifically, we compared the relationship between growth performance and gill area index with gill area index calculated in three different ways: (1) gill area index calculated following Pauly (1981) – that is, with the predicted  $d$  for each species, (2) gill area index calculated following Pauly (2010) – that is, with a constant  $d$  value across all species ( $d = 0.8$ ) and (3) gill area index calculated with an empirically estimated  $d$  value – that is, the  $d$  value is the scaling (slope) of gill surface and body mass.

### 2.4.1 | Data filtering

For this question, we filtered our full data set to include those species that had raw gill surface area data for at least eight individuals, hereafter, 'raw data set'. This threshold of eight individuals was based on simulations (see Bigman et al., [in review](#)) and other studies that have assessed the effect of sample size on regression parameters (Jenkins & Quintana-Ascencio, 2020). In order to facilitate comparison across models with and without a phylogeny, we further restricted our data set to include only those species that have a resolved phylogenetic position on the *Fish Tree of Life* or are included in a recently published Chondrichthyan phylogeny (Chang et al., 2019; Rabosky et al., 2018; Stein et al., 2018). Of the 132 species that have published gill surface area and life history traits in our full data set, 32 species met our criteria (raw gill surface area data with at least eight individuals, known growth parameters and resolved position on the phylogeny) for inclusion in our raw data set.

### 2.4.2 | Statistical analysis

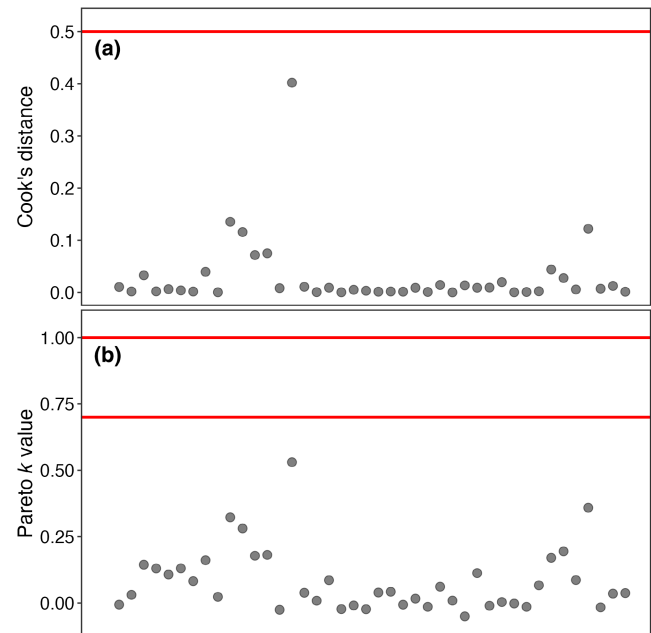
To assess whether growth performance varied with gill area index as calculated with empirically estimated slope values, we



employed a novel phylogenetic Bayesian multilevel modelling framework that included three levels. The first level of the model estimated the ontogenetic allometry of gill surface area and body mass for each species, resulting in a species-specific posterior distribution of the intercept and ontogenetic slope. The second level then used those species-specific slope values, as well as the species-mean gill surface area and body mass, to calculate the gill area index for that species. The third level of the model then examined whether this species-specific gill area index calculated with an empirically estimated  $d$  value explained variation in growth performance. Thus, this model estimated species-specific slopes in the first level, calculated gill area index for each species using the species-specific mean gill surface area, mean body mass associated with the mean gill surface area, and slope estimated in the first level of the model in the second level and then examined whether gill area index explained variation in growth performance in the third level. To ensure that intercepts were estimated accurately across the broad size range of species included in the data set, body mass data were centred on the mean value of body mass for all 32 species in the data set (300 g). The gill area index was  $\log_{10}$ -transformed and standardized using a z-score prior to the second level of the model. The strength of using such a multilevel modelling approach is that the uncertainty in the species-specific intercepts and ontogenetic slopes estimated in the first level of the model, and thus gill area index, is propagated across levels of the model as each iteration of all levels of the model happens in succession (Bigman et al., 2021). All models were fit in R using *rstan* (Stan Development Team, 2019; R Core Team, 2020; see the Appendix S1 for more detail on our modelling approach).

We then calculated gill area index for the same 32 species based on Pauly (1981) and Pauly (2010) and fit two additional models, for a total of three, to assess whether these gill area indices explained variation in growth performance. Finally, we compared the results from all three models to assess whether the relationship between growth performance and gill area index was sensitive to how gill area index is calculated. Note that these last two models differed from those in Question 3 by the number of data points: here, we only used species in the raw data set for purposes of comparison with the model that calculates gill area index using empirically estimated  $d$  values (which we note can only be done for species with raw gill surface area and body mass data).

There were four considerations that necessitated rerunning models in Question 4 on four subsets of data. First, we ran all models with and without a random effect of phylogeny to ensure our results were not biased due to species' sharing various parts of evolutionary trajectories (Felsenstein, 1985; Freckleton, 2009; Harmon, 2019). To do so, we constructed a new supertree with species from our data set using two published phylogenies – one for teleosts (Chang et al., 2019; Rabosky et al., 2018) and one for chondrichthyans (Stein et al., 2018). Second, we reran all models without those species traditionally used in aquaculture because the growth of 'aquacultured' species is known to differ from that of wild fishes due to food ad libitum, reduced predation and possibly



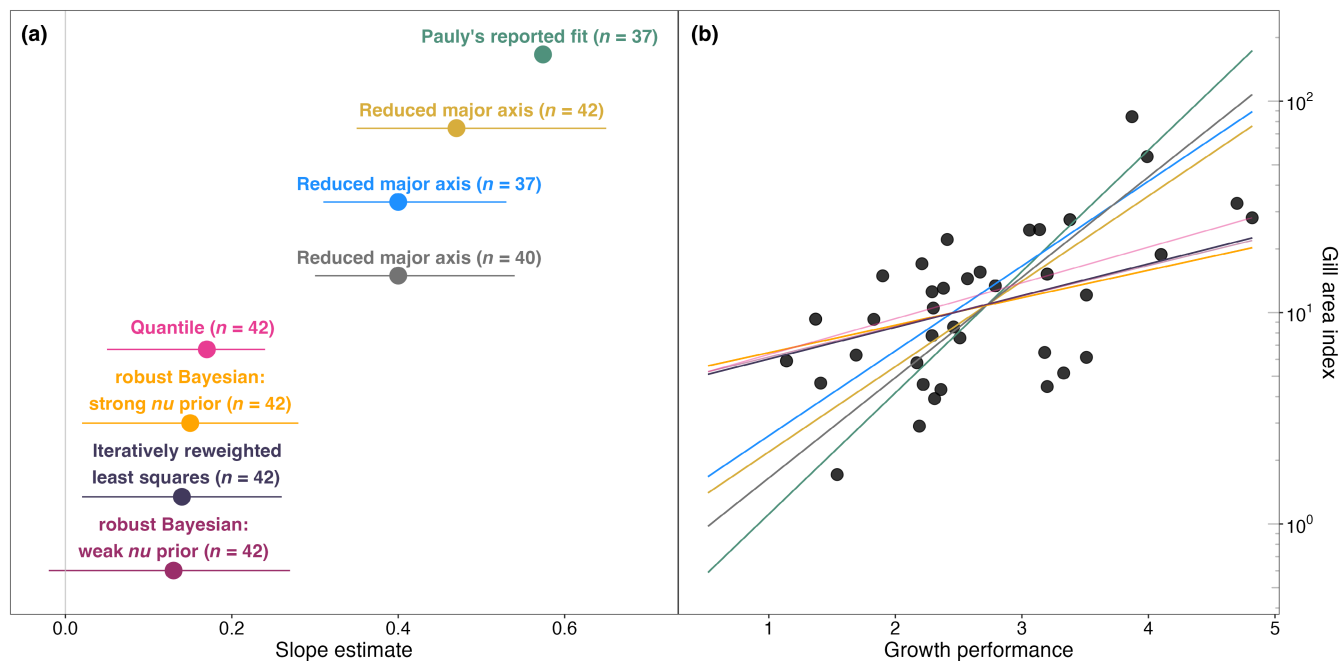
**FIGURE 1** No outliers were identified in the original gill area index and growth performance data set from Pauly (1981). The Cook's distance (a) and Pareto  $k$  values (b) for each of the 42 species in the original data set. The horizontal red line in each plot indicates the threshold for outliers based on Cook's distance (top) and Pareto  $k$  (bottom).

increased aeration of aquaculture ponds (Pauly, 2010). Third, we reran all models without air-breathing fishes because fishes that breathe air either by possessing an air-breathing organ or passive oxygen diffusion through the skin often have a lower gill surface area for a given body size compared to their non-air-breathing counterparts (Graham, 1997; Wegner, 2011). Finally, we also ensured that our regression coefficients were not sensitive to a threshold of eight species for estimating an allometry – for this, we limited our raw data set to those species with gill surface area measurements that ranged over an order of magnitude of body mass ( $n = 34$  species). Adding a random effect of phylogeny, removing aquacultured and air-breathing species or using a data set where allometries were estimated based on body size range did not change our results (Table S1).

### 3 | RESULTS

#### 3.1 | How sensitive is the relationship between gill area index and growth performance to outliers?

We did not identify any outliers in the Pauly data set based on both Cook's distance and the Pareto  $k$  values (Figure 1). This suggests that the five species suspected to be outliers that were subsequently removed from analysis did not affect the relationships between gill area index and growth performance reported in Pauly (1981) and Pauly (2010) and, thus, the results are very similar with and without



**FIGURE 2** The relationship between growth performance and gill area index depends on the type of regression used. (a) A coefficient plot of the mean slope (coloured circle) and 95% intervals (confidence intervals for all non-Bayesian models, credible intervals for all Bayesian models) for all models examined. The vertical grey line indicates zero. (b) The relationship of growth performance and gill area index depends on various regression methods (coloured according to those in panel (a)).

**TABLE 1** Comparison of coefficients estimated by various regression methods for the original model of gill area index and growth performance following Pauly (1981, 2010). In these models, gill area index is the response variable and growth performance is the predictor variable. The number of species used in each model is indicated by 'n' and thus indicates whether or not outliers were excluded when fitting the model.

Model name	Regression method	Intercept (95% CI)	Slope (95% CI)	n
Pauly (1981) reported fit with five potential outliers excluded	Reduced major axis (i.e. 'functional') regression	-0.53 (not given)	0.57 (not given)	37
Pauly (1981) reported fit with two potential outliers excluded	Reduced major axis regression	not given	not given	40
Reduced major axis: 37 species	Reduced major axis regression	-0.06 (-0.40-0.19)	0.40 (0.31-0.53)	37
Reduced major axis: 40 species	Reduced major axis regression	0.02 (-0.34-0.28)	0.40 (0.30-0.54)	40
Reduced major axis: 42 species	Reduced major axis regression	-0.26 (-0.71-0.08)	0.47 (0.35-0.65)	42
Quantile regression	Robust regression	0.63 (0.38-0.86)	0.17 (0.05-0.24)	42
Iteratively reweighted least squares	Robust regression	0.64 (0.31-0.98)	0.14 (0.02-0.26)	42
Robust Bayesian regression weak prior on nu	Bayesian robust regression	0.68 (0.29-1.08)	0.13 (-0.02-0.27)	42
Robust Bayesian regression strong prior on nu	Bayesian robust regression	0.63 (0.27-1.00)	0.15 (0.02-0.28)	42

the inclusion of these species (Figure 2, Table S2 shows which species were originally thought to be outliers).

When comparing model fits across regression methods (reduced major axis, quantile, iteratively reweighted least squares and Bayesian robust), the mean slope for the relationship of gill area index and growth performance for all regression methods was positive but depended on the type of regression (Figure 2, Table 1). The three reduced major axis regression models yielded the greatest mean slope values (~0.4, 95% confidence intervals [CI] did not overlap with zero).

The slope value from Pauly's original reported fit estimated via reduced major axis regression was also slightly greater than any of the other reduced major axis regression models (Figure 2, Table 1). Notably, the mean slopes of all models estimated with reduced major axis regression were significantly greater than the mean slopes estimated by robust regression, which had 95% CIs or Bayesian Credible Intervals [BCIs] that were only slightly non-zero (or for robust Bayesian regression with a weak prior on *nu*, just overlapping with zero). The four different robust regression models almost had identical mean slopes and 95% CIs or BCIs. Further, the choice of prior

on  $nu$  for the Bayesian models did not significantly affect the mean slope estimate (Figure 2, Table 1).

### 3.2 | What is the effect of flipping the axes and testing whether gill area index explains variation in growth performance?

When the relationship between gill area index and growth performance was parameterized according to predictions made by the GOLT, one outlier, West Indian Coelacanth (*Latimeria chalumnae*, Latimeriidae) was identified using Cook's distance; however, the Pareto  $k$  value of this species was below the outlier threshold of 0.7. To be conservative, we compared the results from a Bayesian robust regression to those estimated by a Bayesian simple linear regression (or, 'Bayesian regression with a Gaussian distribution') to ask whether gill area index can explain variation in growth performance for the 42 fish species in the Pauly data set. The model fit was equivalent regardless of the model used based on the leave-one-out information criterion (*looic*, similar interpretation as AIC; robust regression weak prior = 119.6, robust regression strong prior = 120.9, Bayesian simple linear regression = 119.2; Figure 3, Table 2).

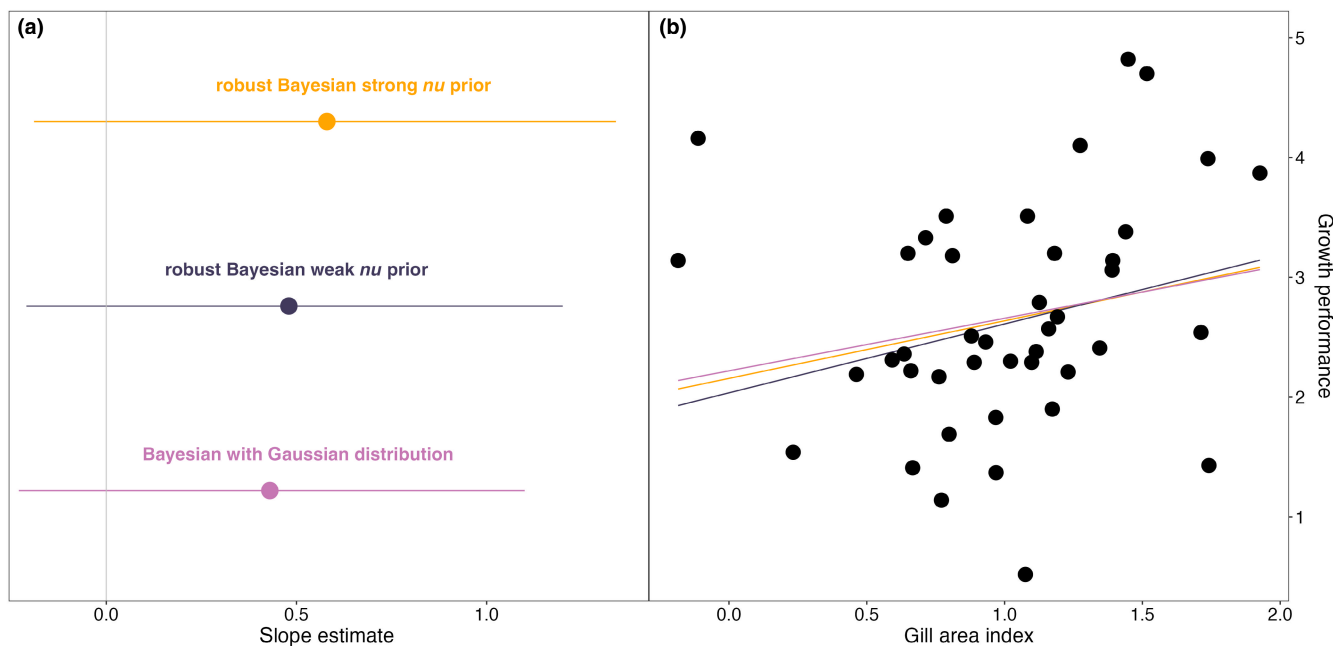
Flipping the axes such that gill area index is the predictor variable and growth performance is the response variable and, thus, testing whether gill area index explains variation in growth performance, results in a weakly positive relationship for all three models with the 95% BCIs overlapping with zero (mean effect sizes ranged from 0.43

to 0.58 depending on the model; Figure 3, Table 2). For the robust regression with a weak prior, 91.3% of the posterior distribution of the gill area index slope was greater than zero, for the robust regression with a strong prior, 92.9% of the posterior distribution of the gill area index slope was greater than zero, and for the Bayesian simple linear regression (that with a Gaussian distribution), 89.9% of the posterior distribution of the gill area index slope was  $>0$  (Figure 3, Table 2). The  $r^2$  values were low for all three models (.06–.09).

### 3.3 | Does the original gill area index and growth performance relationship still hold when examined across more species?

The full data set included 708 observations of gill surface area and associated body masses from a total of 132 fish species for which von Bertalanffy growth parameters were available (Table S3).

The relationship between gill area index calculated with  $d = 0.8$  for all species and growth performance parametrized according to the predictions of the GOLT was positive irrespective of the number of species included in the dataset (Figure 4). Indeed, we found a positive relationship between gill area index and growth performance across 132 fishes, a relationship that was strikingly similar to the original relationship across the 42 species in the Pauly data set. For the Pauly data set, the mean slope = 0.83 (95% BCI 0.19–1.47), with 99.4% of the posterior distribution greater than zero. For the full data set ( $n = 132$  species), the mean slope = 0.87 (95% BCI 0.52–1.23), with 100% of the

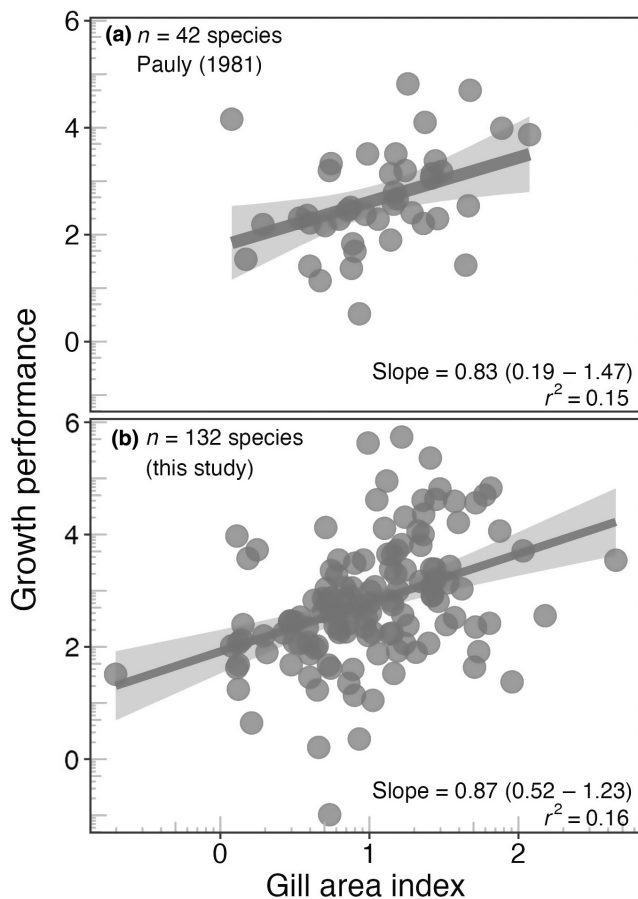


**FIGURE 3** The relationship between growth performance and gill area index for the 42 fish species in the Pauly (1981) data set as parameterized according to the GOLT's prediction that gill surface area constrains growth in fishes. (a) Coefficient plot of the mean slope (coloured circle) and 95% Bayesian Credible Intervals (BCIs; bar) for both robust Bayesian regression models (strong vs. weak prior on  $nu$ ) and the simple linear Bayesian regression (or, 'Bayesian regression with a Gaussian distribution'). The vertical grey line indicates zero. (b) The relationship between gill area index and growth performance for those same three models (coloured according to those in panel (a)). The model fit is equivalent regardless of the model; the 95% BCIs for all three models overlap with zero.



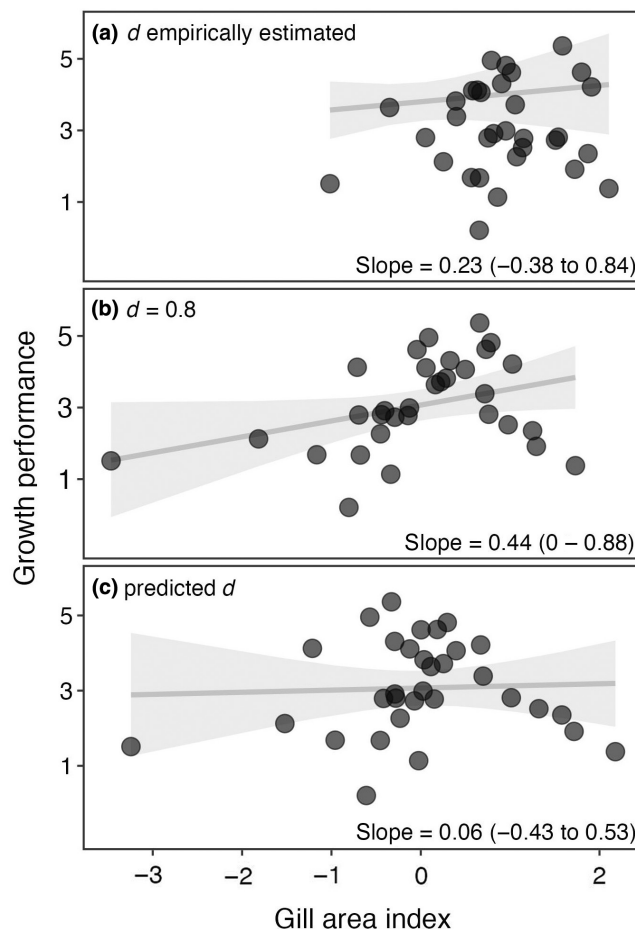
**TABLE 2** Comparison of Bayesian coefficients and their 95% Bayesian Credible Intervals (BCI) estimated for the relationship between gill area index and growth performance, as parameterized according to the predictions made by the GOLT (growth performance ~ gill area index). Model selection was conducted using Pareto-smoothing importance sampling leave-one-out cross validation (PSIS-LOO) using the *loo* package in R v.5.3.1 and v.4.0.1. *looic* = LOO information criterion value (similar to Akaike Information Criterion [AIC]).

Regression method	Intercept (95% BCI)	Slope (95% BCI)	Proportion of posterior distribution >0	<i>Looic</i>
Bayesian simple linear regression	2.22 (1.50–2.95)	0.43 (–0.23–1.10)	0.899	119.2
Bayesian robust regression with weak prior on <i>nu</i>	2.16 (1.38–2.91)	0.48 (–0.21–1.20)	0.913	119.6
Bayesian robust regression with strong prior on <i>nu</i>	2.03 (1.20–2.86)	0.58 (–0.19–1.34)	0.929	120.9



**FIGURE 4** The relationship between growth performance and gill area index was strikingly similar between the (a) 42 fish species in the Pauly (1981) data set and the (b) 132 fish species for which gill surface area data have become available in the 40 years since this relationship was first examined. The model was parameterized according to the GOLT's prediction that gill surface area constrains growth in fishes and with gill area index estimated using  $d = 0.8$ . The model fits were estimated using Bayesian linear regression using the *brm* function in the *brms* package in R.v.4.0.2.

posterior distribution greater than zero. The slopes were statistically indistinguishable from each other, as the slopes of both models fell within the 95% BCI of the other (Figure 4). We note that the  $r^2$  value is still low (.16 for the full data set, .15 for the Pauly data set) and that



**FIGURE 5** The relationship between growth performance and gill area index differed depending on how gill area index was calculated: (a) gill area index calculated from empirically estimated  $d$  values, (b) gill area index calculated using  $d = 0.8$  and (c) gill area index calculated from a predicted  $d$  value from the relationship of  $d$  and maximum size from Pauly (1981). All models were estimated using the 'raw data set' of 32 species with raw gill surface area and body mass data. The relationship in (a) was estimated using a Bayesian multilevel modelling framework (see text), and the relationships in (b) and (c) were estimated using Bayesian linear regression (see text). The fit lines represent the fitted growth performance for each value of gill area index, and the grey shaded region represents the 95% Bayesian Confidence Interval. The 95% BCIs for all models overlapped with zero.

there remains a great deal of variability in the relationship, as indicated by the range of growth performance values for a given gill area index.

### 3.4 | How do the different ways of calculating gill area index affect the relationship between gill surface area and growth performance?

Our raw data set included 457 observations of gill surface area and associated body masses from a total of 32 fish species (teleosts and elasmobranchs) that have a resolved phylogenetic position and for which von Bertalanffy growth parameters were available (Table S4).

The mean effect sizes of the slope for the relationship between growth performance and gill area index differed depending on how gill area index was parameterized. The mean effect size of the slope was greatest when gill area index was calculated with a constant value of  $d = 0.8$ , lower when gill area index was calculated with an empirically estimated  $d$ , and almost equal to zero when gill area index was calculated with the  $d$  value predicted from a relationship between maximum size and the scaling of gill surface area or metabolic rate for other species (Figure 5, Table 3).

When calculating gill area index with an empirically estimated  $d$  from the scaling of gill surface area data and body mass for the 32 species for which it was available, or a  $d$  value predicted from the relationship between maximum size and the slope of gill surface area or metabolic rate as used in Pauly (1981), the posterior distribution of the mean effect size overlapped with zero (empirical  $d$ : mean slope = 0.23, 95% BCI = -0.35 to 0.60, 69.4% of the posterior distribution greater than zero; predicted  $d$ : mean slope = 0.06, 95% BCI = -0.43 to 0.53, 59.6% of the posterior distribution greater than zero; Figure 5, Table 3). However, when gill area index was calculated with  $d = 0.8$  (i.e. assuming the scaling of gill surface area does not differ across species), the relationship was positive, although the lower end of the 95% BCI = 0 (mean slope = 0.44, 95% BCI = 0–0.88, 97.4% of the posterior distribution greater than zero, Figure 5, Table 3).

## 4 | DISCUSSION

Overall, we found that although a central prediction of the GOLT – the relationship between growth performance and gill area index – is positive, suggesting that gill area index does explain variation in growth performance, it is dependent on the method used to calculate gill area index and the type of regression used to fit the data. Indeed, the relationship (the mean effect size, the proportion of the posterior distribution that was  $>0$  and whether the posterior distribution [and thus the 95% BCI] crossed 0) between growth performance and gill area index differed depending on how gill area index was calculated (i.e. with an empirically estimated  $d$  value, a constant  $d$  value, or a predicted  $d$  value from a relationship between maximum size and the scaling of gill surface area and metabolic rate for other species). In addition, we found that this relationship was

less sensitive to the inclusion of outliers than originally suspected and that it was still positive when parameterized according to the prediction made by the GOLT (that gill area index explains variation in growth performance). Remarkably, we found that the relationship between growth performance and gill area index was similar whether fitted across Pauly's original data set (42 species) or an enhanced data set of 132 species. We focus our discussion on the gill area index metric, how different regression techniques affected the results, and finally, suggest future areas of research.

We found that the relationship between growth performance and gill area index was sensitive to how gill area index was calculated. The three methods of calculating gill area index resulted in relationships with growth performance that differed in terms of mean effect sizes, the proportion of the posterior distribution that was  $>0$  and, thus, whether the posterior distribution crossed zero, which is traditionally used to identify a nonsignificant relationship. Further, all three relationships between gill area index and growth performance estimated using our raw data set of 32 species (with high-quality gill surface area and body mass data) differed from Pauly's originally estimated relationship and our relationship estimated across more species (but using the same methods as Pauly (1981) to estimate gill area index). Pauly's original model (Pauly, 1981) and our model across 132 fishes both had steeper slopes between growth performance and gill area index than we found using our raw data set. This variability in gill area index, and subsequently, the resulting relationship between gill area index and growth performance is largely due to the nature of the gill area index metric. While intended to solve the issue of mean data (explained next), the scaling parameter in the gill area index calculation influences a given gill area index value for a species and, in turn, the relationship between gill area index and growth performance. We found that when gill area index was calculated as intended (with  $d =$  empirically estimated species-specific slope of gill surface area and body mass), it did not explain much variation in growth performance. However, the larger issue with the gill area index stems from its calculation with mean data.

As mentioned in the introduction, the gill area index is calculated using mean gill surface area and mean body mass data. For most fishes, gill surface area and the associated body mass data are from a non-random sample of individuals – whatever size range and number of individuals that can be logistically brought back from the field to the laboratory for dissection and measurement (Bigman et al., 2018; Carlson et al., 2004; VanderWright et al., 2020). This is an issue for many traits, especially when measuring them on ectothermic species, which largely grow indeterminately throughout their life (Berrigan & Charnov, 1994; Bigman et al., 2018; Kozłowski, 1996). Indeed, many traits, such as gill surface area or metabolic rate, scale non-linearly with body size and are described by a power law relationship (before the common log-transformation, see Bigman et al., 2018; White & Kearney, 2014). Thus, taking an average of such a trait results in underestimating the true average of this trait for a given species, even if data are available across ontogeny, which is rarely the case. This results in mean values of such traits being problematic and not representative of that trait for that species.

**TABLE 3** Comparison of coefficients and their 95% Bayesian Credible Intervals (BCI) for the relationship between growth performance and the gill area index as calculated using (1) empirically estimated  $d$  values, (2)  $d = 0.8$  and (3)  $d$  predicted from the relationship of  $d$  and maximum size in Pauly (1981; see text). The model in which gill area index was calculated using empirically estimated ontogenetic slope values was estimated using a Bayesian multilevel modelling framework in *rstan* in R v.4.0.2 and the models in which gill area index was calculated using either  $d = 0.8$  or  $d$  predicted from Pauly (1981) were estimated using a Bayesian linear regression using the *brms* package in R.v.4.0.2. All slopes were standardized in the model prior to calculating gill area index, which was  $\log_{10}$ -transformed prior to the second level of the model (see text and SI).

Estimation of gill area index in the model for growth performance ~ gill area index	Intercept (95% CI)	Slope (95% CI)	Proportion of posterior distribution > 0
Empirically estimated	3.80 (3.27–4.33)	0.23 (–0.38 to 0.84)	0.776
$d = 0.8$	3.07 (2.63 to 3.50)	0.44 (0.00 to 0.88)	0.974
Predicted $d$	3.07 (2.60 to 3.53)	0.06 (–0.43 to 0.53)	0.596

This includes the oft-calculated mass-specific or relative measure of a trait (e.g. relative gill surface area is the gill surface area at a given body mass or that predicted from a regression equation at a specific body mass and mass-specific gill surface area is the gill surface area per g of body mass). Indeed, an index based on mean data suffers from Jensen's inequality, which is also known as 'the fallacy of the average', or the problem of averaging over a non-linear relationship (Denny, 2017). We note that this problem with gill area index applies to the gill area indices we used here (i.e. regardless of how gill area index is calculated, it still does not capture how gill surface area changes with size).

Our results also suggest that the type of regression used to fit a relationship between two variables, and how a relationship is parameterized, affects the outcome more so than whether outliers were included or not. We found that different techniques to employ robust regression (here, quantile regression, iteratively reweighted least squares regression, and robust regression with both strong and weak priors) generally yielded the same results, although the mean effect sizes (and whether the 95% BCI crossed zero) did differ among regression techniques. Further, the results from the robust regression models differed from the original method used, reduced major axis regression. While modelling can be thought of as more art than science (*sensu* Box 1976, who famously stated, 'All models are wrong, but some are useful'), care should be taken to select a statistical technique—and parametrization of the model—that is appropriate for a given data set and questions (Gelman & Hill, 2007). In our case, recommendations for deciding between reduced major axis regression or robust regression in the literature are based on the biological question(s) at hand (Kilmer & Rodríguez, 2017; Smith, 2009). One key assumption of using reduced major axis regression is that the relationship between the predictor(s) and response variable is symmetric—it does not matter which variable is the predictor and which is the response, as the resulting coefficients and relationship will be identical (McArdle, 2003; Smith, 2009). As the GOLT makes an explicit prediction about directionality—that gill surface area is constraining maximum size—we suggest that the relationship between gill area index and growth performance should not be tested in a symmetrical manner, and thus favour using robust regression and the resulting conclusions that can be drawn from this framework. We note that to draw conclusions in an asymmetrical manner between

gill area index and growth performance following the GOLT, the response variable and predictor variable should be flipped. When we did so, we found that the relationship between growth performance and gill area index was stronger. However, gill area index was still used as a metric of gill surface area in these relationships.

To examine whether maximum size, growth, and even other life history traits are indeed related to oxygen limitation at the gills, more realistic measures of gill surface area are needed to help uncover whether and how oxygen is related to growth performance and other life history traits. Indeed, we argue that the gill area index, a metric based on mean data (an average across a non-random sample of individuals) is not an ideal metric for which to test this relationship. Our study showed that this metric was variable depending on how it was calculated, yielding different relationships with growth performance. To this end, future work in this area should employ more realistic measures of gill surface area, for example, the ontogenetic regression coefficients from an allometry of gill surface area and body mass, with respect to its relationship with growth performance and then expand from there to other life history traits (as the GOLT argues that oxygen limitation at the gills limits the aerobic metabolic rate and thus energy for traits related to survival, growth and reproduction). In a follow-up paper (Bigman et al., *in review*), we address the former part of this idea by bringing gill surface area into a scaling context and examining whether the gill surface area at a given size (the intercept of a gill surface area allometry—or the predicted gill surface area at a given size) and the rate at which gill surface area increases with body mass ontogenetically (the slope of a gill surface area allometry) are related to somatic growth, maximum size and growth performance. Specifically, we base our estimations of these intercepts and slopes on similar principles here—raw gill surface area for multiple individuals of a species. Additionally, we employ the novel phylogenetic Bayesian multilevel model used here to assess how individual variation in the gill surface area and body mass relationship confers patterns of gill surface area and life history traits across species. Indeed, understanding the role that oxygen plays in the physiology and ecology of fishes, including whether oxygen limitation is occurring at the gills, will be paramount to predicting how species will respond to increasing temperatures and reduced oxygen availability associated with climate change.

## AUTHOR CONTRIBUTION

J.S.B. and N.K.D. conceived and designed the project and analysis. J.S.B. collected the data and performed all analyses and visualizations. All authors contributed to the interpretation of results. J.S.B. drafted the manuscript and supplementary information with input from all authors. N.K.D. supervised the project.

## ACKNOWLEDGEMENTS

We would like to thank Daniel Pauly for the many enlightening and open-minded discussions on this topic as well as the members of the past and present Dulvy Lab for insightful comments over the years. This project was funded by the Natural Sciences and Engineering Research Council of Canada and the Canada Research Chairs Program. JSB also acknowledges funding from the National Science Foundation (NSF grant number 2109411).

## CONFLICT OF INTEREST

The authors declare that they have no competing interests.

## DATA AVAILABILITY STATEMENT

The data and code associated with this paper are freely available. The data are available for download on Figshare: <https://doi.org/10.6084/m9.figshare.22081052.v1> and the code is available for download from Github: <https://github.com/jennybigman/revisiting-the-GOLT>.

## ORCID

Jennifer S. Bigman  <https://orcid.org/0000-0001-8070-3061>

Nicholas C. Wegner  <https://orcid.org/0000-0002-8447-1488>

Nicholas K. Dulvy  <https://orcid.org/0000-0002-4295-9725>

## REFERENCES

- Anderson, S. C., Branch, T. A., Cooper, A. B., & Dulvy, N. K. (2017). Black-swan events in animal populations. *Proceedings of the National Academy of Sciences*, 114(12), 3252–3257. <https://doi.org/10.1073/pnas.1611525114>
- Berrigan, D., & Charnov, E. L. (1994). Reaction norms for age and size at maturity in response to temperature: A puzzle for life historians. *Oikos*, 70(3), 474. <https://doi.org/10.2307/3545787>
- Bigman, J. S., M'Gonigle, L. K., Wegner, N. C., & Dulvy, N. K. (2021). Respiratory capacity is twice as important as temperature in explaining patterns of metabolic rate across the vertebrate tree of life. *Science Advances*, 7(19), eabe5163. <https://doi.org/10.1126/sciadv.abe5163>
- Bigman, J. S., Pardo, S. A., Prinzing, T. S., Dando, M., Wegner, N. C., & Dulvy, N. K. (2018). Ecological lifestyles and the scaling of shark gill surface area. *Journal of Morphology*, 279(12), 1716–1724. <https://doi.org/10.1002/jmor.20879>
- Bigman, J. S., Wegner, N. C., & Dulvy, N. K. (in review). Gills, growth, and activity across fishes. *Fish and Fisheries*.
- Boettiger, C., Lang, D. T., & Wainwright, P. C. (2012). Rfishbase: Exploring, manipulating and visualizing FishBase data from R. *Journal of Fish Biology*, 81(6), 2030–2039. <https://doi.org/10.1111/j.1095-8649.2012.03464>
- Bürkner, P. C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28.
- Bürkner, P. C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1), 395–341.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: A practical information-theoretic approach* (2nd ed.). Springer.
- Carlson, J., Goldman, K., & Lowe, C. (2004). Metabolism, energetic demand, and Endothermy. In J. Musick, J. Carrier, & M. Heithaus (Eds.), *Biology of sharks and their relatives* (Vol. 20043354, pp. 203–224). CRC Press. <https://doi.org/10.1201/9780203491317>
- Chang, J., Rabosky, D. L., Smith, S. A., & Alfaro, M. E. (2019). An R package and online resource for macroevolutionary studies using the ray-finned fish tree of life. *Methods in Ecology and Evolution*, 10(7), 1118–1124. <https://doi.org/10.1111/2041-210X.13182>
- Cheung, W. W. L., Sarmiento, J. L., Dunne, J., Frölicher, T. L., Lam, V. W. Y., Deng Palomares, M. L., Watson, R., & Pauly, D. (2013). Shrinking of fishes exacerbates impacts of global ocean changes on marine ecosystems. *Nature Climate Change*, 3(3), 254–258. <https://doi.org/10.1038/nclimate1691>
- Denny, M. (2017). The fallacy of the average: On the ubiquity, utility and continuing novelty of Jensen's inequality. *Journal of Experimental Biology*, 220(2), 139–146. <https://doi.org/10.1242/jeb.140368>
- Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). *Regression*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-34333-9>
- Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist*, 125(1), 1–15.
- Fox, J., & Weisberg, S. (2012). Robust regression. *An R and S-Plus Companion to Applied Regression*, 91, 1–16.
- Freckleton, R. P. (2009). The seven deadly sins of comparative analysis. *Journal of Evolutionary Biology*, 22(7), 1367–1375. <https://doi.org/10.1111/j.1420-9101.2009.01757>
- Froese, R., & Pauly, D. (2020). *FishBase. World Wide Web electronic publication*. [www.fishbase.org](http://www.fishbase.org), version (12/2020).
- Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(2), 389–402. <https://doi.org/10.1111/rssa.12378>
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and other stories*. Cambridge University Press.
- Graham, J. (1997). *Air-breathing fishes: Evolution, diversity, and adaptation*. Academic Press.
- Gray, I. E. (1954). Comparative study of the gill area of marine fishes. *The Biological Bulletin*, 107(2), 219–225. <https://doi.org/10.2307/1538608>
- Hampel, F. R. (2001). *Robust Statistics: A Brief Introduction and Overview*. <https://doi.org/10.3929/ETHZ-A-004158597>
- Harmon, L. (2019). *Phylogenetic comparative methods*. Independent.
- Hughes, G. M., & Morgan, M. (1973). Supplement 90005 to the structure of fish gills in relation to their respiratory function. *Biological Reviews*, 48, 419–475.
- Jager, S. D., & Dekkers, W. (1975). Relation between gill structure and activity in fish. *Netherlands Journal of Zoology*, 25(3), 276–308.
- Jenkins, D. G., & Quintana-Ascencio, P. F. (2020). A solution to minimum sample size for regressions. *PLoS One*, 15(2), e0229345. <https://doi.org/10.1371/journal.pone.0229345>
- Juan-Jordá, M. J., Mosqueira, I., Freire, J., & Dulvy, N. K. (2013). Life in 3-D: Life history strategies in tunas, mackerels and bonitos. *Reviews in Fish Biology and Fisheries*, 23(2), 135–155. <https://doi.org/10.1007/s11160-012-9284-4>
- Kilmer, J. T., & Rodríguez, R. L. (2017). Ordinary least squares regression is indicated for studies of allometry. *Journal of Evolutionary Biology*, 30(1), 4–12. <https://doi.org/10.1111/jeb.12986>



- Koenker, R. (2020). *quantreg: Quantile Regression. R package version 5.73*. <https://CRAN.R-project.org/package=quantreg>
- Kozłowski, J. (1996). Optimal allocation of resources explains interspecific life-history patterns in animals with indeterminate growth. *Proceedings of the Royal Society B: Biological Sciences*, 263(1370), 559–566. <https://doi.org/10.1098/rspb.1996.0084>
- Kruschke, J. K. (2015). What's in this book (read this first!). In *Doing Bayesian Data Analysis* (pp. 1–11). Elsevier. <https://doi.org/10.1016/B978-0-12-405888-0.00001-5>
- Lange, K. L., Little, R. J. A., & Taylor, J. M. G. (1989). Robust statistical modeling using the *t* distribution. *Journal of the American Statistical Association*, 84(408), 881–896. <https://doi.org/10.1080/01621459.1989.10478852>
- Lefevre, S., McKenzie, D. J., & Nilsson, G. E. (2017). Models projecting the fate of fish populations under climate change need to be based on valid physiological mechanisms. *Global Change Biology*, 23(9), 3449–3459. <https://doi.org/10.1111/gcb.13652>
- Lefevre, S., McKenzie, D. J., & Nilsson, G. E. (2018). In modelling effects of global warming, invalid assumptions lead to unrealistic projections. *Global Change Biology*, 24(2), 553–556. <https://doi.org/10.1111/gcb.13978>
- Legendre, P., & Oksanen, M. J. (2018). *Package 'lmodel2'*. <https://CRAN.R-project.org/package=lmodel2>
- McArdle, B. H. (2003). Lines, models, and errors: Regression in the field. *Limnology and Oceanography*, 48(3), 1363–1366. <https://doi.org/10.4319/lo.2003.48.3.1363>
- Palzenberger, M., & Pöhla, H. (1992). Gill surface area of water-breathing freshwater fish. *Reviews in Fish Biology and Fisheries*, 2(3), 187–216. <https://doi.org/10.1007/BF00045037>
- Pauly, D. (1981). The relationships between gill surface area and growth performance in fish: A generalization of von Bertalanffy's theory of growth. *Berichte Der Deutschen Wissenschaftlichen Kommission Für Meeresforschung*, 28(4), 251–282.
- Pauly, D. (2010). *Gasping fish and panting squid oxygen, temperature and the growth of water breathing animals*. International Ecology Institute.
- Pauly, D. (2021). The gill-oxygen limitation theory (GOLT) and its critics. *Science Advances*, 7(2), eabc6050. <https://doi.org/10.1126/sciadv.abc6050>
- Pauly, D., & Cheung, W. W. L. (2016). Impacts and effects of ocean warming on marine fishes. In *Explaining ocean warming: Causes, scale, effects, and consequences* (pp. 239–253). IUCN.
- Pauly, D., & Liang, C. (2022). Temperature and the maturation of fish: A simple sine-wave model for predicting accelerated spring spawning. *Environmental Biology of Fishes*, 105(10), 1481–1487. <https://doi.org/10.1007/s10641-022-01212-0>
- Pauly, D. (1991). Growth performance in fishes: Rigorous description of patterns as a basis for understanding causal mechanisms. *Aqua*, 4(3), 3–6.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.Rproject.org/>
- Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T. J., Coll, M., & Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, 559(7714), 392–395. <https://doi.org/10.1038/s41586-018-0273-1>
- Rousseeuw, P. (1984). Least median of squares regression. *Journal of the American Statistical Association*, 79(388), 871–880.
- Satora, L., & Wegner, N. C. (2012). Reexamination of the Byczkowska-Smyk gill surface area data for European teleosts, with new measurements on the pikeperch, Sander lucioperca. *Reviews in Fish Biology and Fisheries*, 22(1), 1–9. <https://doi.org/10.1007/s11160-011-9213-y>
- Seibel, B. A., & Deutsch, C. (2020). Oxygen supply capacity in animals evolves to meet maximum demand at the current oxygen partial pressure regardless of size or temperature. *Journal of Experimental Biology*, 223(12), jeb.210492, 1–11. <https://doi.org/10.1242/jeb.210492>
- Smith, R. J. (2009). Use and misuse of the reduced major axis for line-fitting. *American Journal of Physical Anthropology*, 140(3), 476–486. <https://doi.org/10.1002/ajpa.21090>
- Stan Development Team. (2019). *Stan modeling language users guide and reference manual, Version 2.19.2*. <http://mc-stan.org>
- Stein, R. W., Mull, C. G., Kuhn, T. S., Aschliman, N. C., Davidson, L. N. K., Joy, J. B., Smith, G. J., Dulvy, N. K., & Mooers, A. O. (2018). Global priorities for conserving the evolutionary history of sharks, rays and chimaeras. *Nature Ecology & Evolution*, 2(2), 288–298. <https://doi.org/10.1038/s41559-017-0448-4>
- VanderWright, W. J., Bigman, J. S., Elcombe, C. F., & Dulvy, N. K. (2020). Gill slits provide a window into the respiratory physiology of sharks. *Conservation Physiology*, 8(1), coaa102. <https://doi.org/10.1093/conphys/coaa102>
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer.
- Wang, C., & Blei, D. M. (2018). A general method for robust Bayesian modeling. *Bayesian Analysis*, 13(4), 1163–1191. <https://doi.org/10.1214/17-BA1090>
- Wegner, N. (2011). Gill respiratory morphometrics. In *Encyclopedia of fish physiology: From genome to environment* (Vol. 2, pp. 803–811). Academic Press.
- White, C. R., & Kearney, M. R. (2014). Metabolic scaling in animals: Methods, empirical results, and theoretical explanations. *Comprehensive Physiology*, 4, 27.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Bigman, J. S., Wegner, N. C., & Dulvy, N. K. (2023). Revisiting a central prediction of the Gill Oxygen Limitation Theory: Gill area index and growth performance. *Fish and Fisheries*, 24, 354–366. <https://doi.org/10.1111/faf.12730>