

Authors

Ruo-Yu Pan

Institute of Marine Affairs and Resource Management, National Taiwan Ocean University, Taiwan 202301

elthina02017@gmail.com

Ting-Chun Kuo

Institute of Marine Affairs and Resource Management, National Taiwan Ocean University, Taiwan 202301

cielocot@gmail.com

Chih-hao Hsieh

Institute of Oceanography, National Taiwan University, Taiwan 106319

chsieh@ntu.edu.tw

SizeAggregTend_data

Here store the data used to estimate size-based Taylor's exponents and the relationship between Taylor's exponents and body size in the North Sea. We sorted data in three groups, the raw data (folder "raw_data"), the organized data (folder "compiled"), and the output data (folder "output"). In the output data, the data are further sorted into five groups based on the analysis process:

- (1) Size-based Taylor's power law and the relationship between aggregation tendency and body size (the folder "1_size_TL"),
- (2) Spawning effect on relationship between Taylor's exponents and body size (the folder "2_spwan_effect"),
- (3) Hump-shaped characteristics for 9 species (the folder "3_humpshape_characteristics"),
- (4) Potential factors effect on species-specific hump-shaped relationship (the folder "4_potential_factor_effect_on_humpshape"), and
- (5) The relationship between sex ratio and length.

In the tables below (Table 1-3), we listed and gave the description of each file.

Table 1. The name and the description of the raw data.

Raw data download from the website include:			
File name or file name format	Data information	The information used in this study	Source
CPUE_per_length_per_subarea.csv	The quarterly CPUE data of each subarea in the North Sea and each length bin for 9 species	The quarterly CPUE data by subarea and by length bin in the North Sea for 9 species	ICES, DATRAS (https://datras.ices.dk/Data_products/Download/Download_Data_public.aspx)

[Year].csv e.g. 1995.csv	Sea temperature by depth from 1995 to 2015	Sea bottom temperature (SBT) and sea surface temperature (SST)	ICES (https://ocean.ices.dk/HydChem/HydChem.aspx?plot=yes)
SMALK_data	The sex, maturity, and age-length key for 9 species in 10 areas in the North Sea	Maturation stage by length bin for 9 species in the North Sea	ICES, DATRAS (https://datras.ices.dk/Data_products/Download/Download_Data_public.aspx)
[SP.NAME]_stkases.xml e.g. herring_stkases.xml	Stock assessment of each of 9 species conducted by ICES	Fishing mortality	ICES (http://ices.dk/marine-data/assessment-tools/Pages/stock-assessment-graphs.aspx)
Exchange_Data_Q1_dataneed.csv Exchange_Data_Q3_dataneed.csv	The raw IBTS data that includes the data as reported by the national institute	The gender by length	ICES, DATRAS (https://datras.ices.dk/Data_products/Download/Download_Data_public.aspx)

Table 2. The name and the description of the organized data.

Organized or Recalculated data:			
File name or file name format	Data information	Variable (unit)	Data or research the file based on
subarea_symb.csv	The latitude and longitude of central point for each subarea	3 columns: subarea -- subarea defined by ICES, lat (degree), long (degree) 193 rows: 193 subarea in the North Sea	CPUE_per_length_per_subarea.csv
CPUE_per_length_per_subarea_clean.csv	The quarterly CPUE data of each subarea in the North Sea and each length bin for 9 species used	7 columns: Year, Quarter, Area, Subarea, Species, Length class (mm), CPUE (number/hr) 771397 records	CPUE_per_length_per_subarea.csv
allsp_length_mid_16class.txt	the midpoint of the redefined length class in the study	the midpoint of the redefined length class (mm)	CPUE_per_length_per_subarea_clean.csv
allsp_length_range_16class.txt	the break points of the redefined length classes in this study	the break points of the redefined length classes (mm)	CPUE_per_length_per_subarea_clean.csv
fishingM.csv	the annual fishing mortality for 9 species	the annual fishing mortality	[SP.NAME]_stkases.xml

mature_L.txt	the length at maturation	the length at maturation (mm)	Thorpe et al. (2015)
size_at_maturityStage_byQ.csv	the length at 5% mature, length at 50% mature, and length at 95% mature for 9 species and two quarters	5 columns: L05 (mm), L50 (mm), L95 (mm), Species, Quarter 18 rows: 9 species * 2 quarters	SMALK_data
Temp_1991_2015.csv	Compile the temperature data from 1995 to 2015	10 columns: Cruise, Station, Type, date and time, Latitude (degree), Longitude (degree), bottom depth (m), pressure (db), temperature (deg.C), salinity (psu) 1048575 records	[Year].csv
Temperature_info.csv	The mean and CV of the SST and SBT for each length class in the species-specific habitat for two quarters	4 columns: mean of habitat temperature (deg.C), cv of habitat temperature (deg.C), species, quarter 288 rows: 16 size classes * 2 quarters * 9 species	Temp_1991_2015.csv
Thorpe2015_Life_history_traits.csv	The life history traits of 9 fish species, including life style, asymptotic length, von Bertalanffy growth rate, and length at maturation	5 columns: species name, life style, asymptotic length (cm), von Bertalanffy growth rate, length at maturation (cm) 9 rows: 9 species	Thorpe et al. (2015), FishBase and ICES-FishMap

Table 3. The name and the description of the output data.

Output data (Result):		
File name or file name format	Data information	Detail
(1) Size-based Taylor's power law and the relationship between aggregation tendency and body size		
[SP.NAME]__[Quarter]_[SIZECLASS].csv e.g. Cod__Q1_1	999 bootstrapped result of the Taylor's power law for each of 16 size classes, each quarters, and each of 9 species	1 column: bootstrapped Taylor's exponents 999 rows: 999 bootstraps

all_TL_result_16class.csv	The summary of Taylor's power law fitting to 16 classes and 2 quarters	<p>9 columns: slope of Taylor's power law (<i>b</i>), intercept of Taylor's power law (<i>a</i>), R-square of fitted Taylor's power law, the average CPUE across 25 years, the lower bound of the Bca of the intercept (<i>a</i>), the upper bound of the Bca of the intercept (<i>a</i>), the lower bound of the Bca of the slope (<i>b</i>), the upper bound of the Bca of the slope (<i>b</i>), the size class</p> <p>288 rows: 16 size classes * 2 quarters * 9 species</p>
all_TL_result_16_compile.csv	The data include the information of the summary of Taylor's power law fitting to 16 classes, as well as the length-specific information for the further analysis, including standardized length, species, quarters, spawning season or not, mean habitat temperature, cv of the habitat temperature, and the maturation group	<p>19 columns: slope of Taylor's power law (<i>b</i>), intercept of Taylor's power law (<i>a</i>), R-square of fitted Taylor's power law, the average CPUE across 25 years, the lower bound of the Bca of the intercept (<i>a</i>), the upper bound of the Bca of the intercept (<i>a</i>), the lower bound of the Bca of the slope (<i>b</i>), the upper bound of the Bca of the slope (<i>b</i>), the size class, standardized length, species, quarters, spawning season or not, mean of the habitat temperature, cv of the habitat temperature, the maturation group, dummy variable (used for exclude random effect)</p> <p>288 rows: 16 size classes * 2 quarters * 9 species</p>
resample_TL_bs_data.csv	Resample the Taylor's exponenets for 16 length classes 999 times. This data is for following bootstrap analysis.	<p>999 columns: 999 times bootstraps</p> <p>288 rows: 16 size classes * 2 quarters * 9 species</p>
bs_fit_para_summary.csv	Fitting relationship between Taylor's exponents and body size using the best model 999 times. This data contains 999 bootstrapped <u>parametric</u> result of GAMM.	<p>4 columns: Estimated value, standard error, t value, p value</p> <p>999 rows: 999 bootstraps</p>
bs_fit_smooth_summary.csv	Fitting relationship between Taylor's exponents and body size using the best model 999 times. This data contains 999 bootstrapped p-value of <u>nonparametric</u> result of GAMM.	<p>8 columns: p-value for 8 terms in the model, the 8 terms are cv of the habitat temperature in Q1, cv of the habitat temperature in Q3, size in Q1, size in Q3, the interaction of temperature index and size in Q1, the interaction of temperature index and size in Q3, the random intercept for species, and the random slope for species.</p> <p>999 rows: 999 bootstraps</p>

bs_fit_quarter_difference.csv	Fitting relationship between Taylor's exponents and body size using the best model 999 times. This data contains 999 bootstrapped difference of the smooth trend between two quarters.	999 columns: 999 bootstraps 200 rows: the predicted difference of the smooth trend between two quarters at 200 standardized length points.
bs_fit_predicted_data_Q1_way1.csv	The 999 predicted relationships between Taylor's exponents and body size using the best model.	999 columns: 999 bootstraps
bs_fit_predicted_data_Q3_way1.csv	These data are used for calculating the 95% bootstrapped confidence intervals (with acceleration).	200 rows: the predicted Taylor's exponents at 200 standardized length points.
bs_lin_matstage_byQ_QART_1.csv	Fitting relationship between Taylor's exponents and body size for three maturation group using GLM 999 times. This data contains 999 bootstrapped	4 columns: the intercept of fitted GLM, the slope of fitted GLM, maturation groups, quarter
bs_lin_matstage_byQ_QART_3.csv	slope for three maturation groups for each of 9 species.	2997 rows: 3 maturation groups * 999 bootstraps

(2) Spawning effect on relationship between Taylor's exponents and body size

SPWN_bs_fit_para_summary.csv	Fitting relationship between Taylor's exponents and body size with consideration of spawning effect 999 times. This data contains 999 bootstrapped <u>parametric</u> result of GAMM.	4 columns: Estimated value, standard error, t value, p value 999 rows: 999 bootstraps
SPWN_bs_fit_smooth_summary.csv	Fitting relationship between Taylor's exponents and body size with consideration of spawning effect 999 times. This data contains 999 bootstrapped <u>nonparametric</u> result of GAMM.	8 columns: p-value for 8 terms in the model, the 8 terms are cv of the habitat temperature in spawn season (Y), cv of the habitat temperature in resting season (N), size in spawn season (Y), size in resting season (N), the interaction of temperature index and size in spawn season (Y), the interaction of temperature index and size in resting season (N), the random intercept for species, and the random slope for species. 999 rows: 999 bootstraps
bs_fit_spawn_diff.csv	Fitting relationship between Taylor's exponents and body size with consideration of spawning effect 999 times. This data contains 999 bootstrapped difference of the smooth trend between the spawning and resting season.	999 columns: 999 bootstraps 200 rows: the predicted difference of the smooth trend between spawn season and rest season at 200 standardized length points.

bs_fit_predicted_data_Y.csv	The 999 predicted relationships between Taylor's exponents and body size with consideration of spawning effect. These data are used for calculating the 95% bootstrapped confidence intervals (with acceleration).	999 columns: 999 bootstraps 200 rows: the predicted Taylor's exponents at 200 standardized length points.
bs_fit_predicted_data_N.csv		
(3) Hump-shaped characteristics for 9 species		
sp_resp_humppeak.csv	The 999 bootstrapped maximum peak and peak location of the hump-shaped relationship for two quarters fitted for each species respectively	999 columns: 999 times bootstraps 36 rows: [maximum b value in Q1, peak location in Q1, maximum b value in Q3, peak location in Q3]* 9 species
sp_resp_slope_Q1.csv	The summary for 999 bootstrapped GLM relationship and original GLM relationship between Taylor's exponents and body size for each of three maturation group.	21 columns: 3 maturation groups * [Estimated slope, Standard error, t-value, p-value, R-square, Deviance explained, # of data point]
sp_resp_slope_Q3.csv		9000 rows: 9 species*(1 original+999 bootstraps)
sp_bs_pred_Q1.RData	The Rdata contains a list with 999 predicted relationships between Taylor's exponents and body size fore each of 9 species. These data are used for calculating the 95% bootstrapped confidence intervals (with acceleration).	9 item in the list: 9 species
sp_bs_pred_Q3.RData		Each item contains a data frame with 1000 columns and 200 rows. 1000 columns: 1 original+999 bootstraps 200 rows: the predicted Taylor's exponents at 200 standardized length points.
sp_resp_peak_slope_95CI.csv	The bootstrapping 95% CI (with acceleration) of 5 humped shape characteristics in two quarters	5 columns: lower bound of the BCa, upper bound of the Bca, species, quarter, humped shape characteristics 90 rows: 5 humped shape characteristics * 2 quarters * 9 species
(4) Potential factors effect on species-specific hump-shaped relationship		
Potential_factor_effect_result.csv	The result of 999 bootstrapped plus 1 original correlation test or M-W U test for 5 protential factors' effect on 3 humped shape characteristics, which are the slopes for 3 maturation groups.	5 columns: statistics value of the test (correlation or U value), p-value, potential factor, maturation group, quarter 30000 rows: 2 quarters * 5 potential factors * (999 bootstraps + 1 origin) * 3 maturation groups

Potential_factor_effect_res ult.csv	The bootstrapping 95% CI (with acceleration) of the statistics values and p-value of correlation test or M-W U test between 5 potential factors and 3 humped shape characteristics, which are the slopes for 3 maturation groups.	7 columns: lower bound of the BCa of statistics value, upper bound of the BCa of statistics value, lower bound of the BCa of p value, upper bound of the BCa of p value, potential factor, maturation group, quarter 30 rows: 5 potential factors * 3 maturation groups * 2 quarters
(5) The relationship between sex ratio and length		
sexratio_vs_lengthclass.csv	The result for linear relationship between sex ratio and length bins for 8 species	4 columns: Estimated value, standard error, t value, p value 32 rows: [Intercept summary, Slope summary] * 8 species * 2 quarters