Application of Step-Wise Weighted Least Squares Procedure and Piece-Wise Mathematical Models in Study of Time Series of Growth Characteristics of Atlantic Salmon Smolts

M. Khots, K. Haya, L. E. Burridge, S. B. Brown, W.L. Fairchild

Fisheries and Oceans Canada,
Science Branch
Maritimes Region
St. Andrews Biological Station,
531 Brandy Cove Rd.
St. Andrews, New Brunswick, Canada
E5B 2L9

2010

Canadian Manuscript Report of Fisheries and Aquatic Sciences 2931

Rapport Manuscript Canadien des Sciences Halieutiques et Aquatiques 2931

Fisheries and Oceans Canada
Pêchés et Océans Canada
Canada

Canadian Manuscript Report of Fisheries and Aquatic Sciences

Manuscript reports contain scientific and technical information that contributes to existing knowledge but which deals with national or regional problems. Distribution is restricted to institutions or individuals located in particular regions of Canada. However, no restriction is placed on subject matter, and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Manuscript reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Manuscript reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-900 in this series were issued as Manuscript Reports (Biological Series) of the Biological Board of Canada, and subsequent to 1937 when the name of the Board was changed by Act of Parliament, as Manuscript Reports (Biological Series) of the Fisheries Research Board of Canada. Numbers 1426 - 1550 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Manuscript Reports. The current series name was changed with report number 1551.

Rapport manuscrit canadien des sciences halieutiques et aquatiques

Les rapports manuscrits contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui traitent de problèmes nationaux ou régionaux. La distribution en est limitée aux organismes et aux personnes de régions particulières du Canada. II n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports manuscrits peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports manuscrits sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques.*

Les rapports manuscrits sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 900 de cette série ont été publiés à titre de Manuscrits (série biologique) de l'Office de biologie du Canada, et après le changement de la désignation de cet organisme par décret du Parlement, en 1937, ont été classés comme Manuscrits (série biologique) de l'Office des recherches sur les pêcheries du Canada. Les numéros 901 à 1425 ont été publiés à titre de Rapports manuscrits de l'Office des recherches sur les pêcheries du Canada. Les numéros 1426 à 1550 sont parus à titre de Rapports manuscrits du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 1551.

Canadian Manuscript Report of
Fisheries and Aquatic Sciences 2931

2010

Application of Step-Wise Weighted Least Squares Procedure and Piece-Wise Mathematical Models in Study of Time Series of Growth Characteristics of Atlantic Salmon Smolts

by

M. Khots, K. Haya, L. E. Burridge, S. B. Brown[1], W.L. Fairchild[2]

[1] Environment Canada, National Water Research Institute, PO Box 5050, Burlington, Ontario, L7R 4A6, Canada

[2] Fisheries and Oceans Canada, PO Box 5030, 343 Université Ave., Moncton, New Brunswick, E1C 9B6, Canada

Cat. No. FS97-4/2931E   ISSN 0706 – 6473 (print version)
Cat. No. FS97-4/2931E-PDF ISSN 1488-5387 (online version)

Correct citation for this publication:

Khots, M., Haya K., Burridge L.E., Brown S.B., Fairchild W.L. 2010.

Application of step-wise weighted least squares procedure and piece-wise mathematical models in study of time series of growth characteristics of Atlantic salmon smolts. Can. Manuscr. Rep. Fish. Aquat. Sci. 2931: v + 16 p.

TABLE OF CONTENTS

ABSTRACT

Khots, M., Haya K., Burridge L.E., Brown S.B., Fairchild W.L. 2010. Application of step-wise weighted least squares procedure and piece-wise mathematical models in study of time series of growth characteristics of Atlantic salmon smolts. Can. Manuscr. Rep. Fish. Aquat. Sci. 2931: v + 16 p.

Growth of Atlantic salmon (*Salmo salar*) during the latter stages of parr to smolt transformation (PST) and the early months in seawater is described statistically. Dependences of the length and the weight on time were studied. Statistical processing of data included several steps: grouping of experimental data; building of the distribution curves for length and weight for different periods of time; detection and removal of outliers; and construction of functional dependences of length and weight on time. This paper describes statistically the growth characteristics of salmon smolts in fresh and in seawater under our laboratory conditions.

RÉSUMÉ

Khots, M., Haya K., Burridge L.E., Brown S.B., Fairchild W.L. 2010. Application of step-wise weighted least squares procedure and piece-wise mathematical models in study of time series of growth characteristics of Atlantic salmon smolts. Can. Manuscr. Rep. Fish. Aquat. Sci. 2931: v + 16 p.

Le présent document décrit d’un point de vue statistique la croissance du saumon atlantique (*Salmo salar*) durant les derniers stades de la transformation des tacons en saumoneaux et durant leurs premiers mois en eau de mer. Les chercheurs ont étudié les dépendances des facteurs de la longueur et du poids sur le temps. Le traitement statistique des données comprenait plusieurs étapes : regroupement des données expérimentales, création des courbes de distribution pour ce qui est de la longueur et du poids à différentes périodes, détection et suppression des valeurs aberrantes, et enfin, création de dépendances fonctionnelles de la longueur et du poids sur la période de temps. Le document décrit d’un point de vue statistique les caractéristiques de la croissance des saumoneaux en eau douce et en eau salée, en laboratoire.

INTRODUCTION

Indigenous populations of Atlantic salmon (Salmo salar L.) have been decreasing throughout the northwest Atlantic area for a number of years. The cause of this decline remains unclear. Fairchild et al. (1999) showed a strong relationship between pesticide spray events in New Brunswick, Canada and reduced adult salmon returns to the same rivers. They hypothesized that a putative endocrine disruptor, p-nonylphenol (NP), used as an inert ingredient in the pesticide formulation may have contributed to the low returns (Fairchild et al., 1999).

Atlantic salmon is an anadromous fish species which spends its early life in freshwater, migrates to seawater and returns to freshwater to reproduce. The morphological and physiological changes that take place prior to seaward migration are collectively referred to as Parr Smolt Transformation (PST), smolt being the word used to describe salmon that have physiologically adapted for life at sea. A number of authors have described these changes (McCormick and Saunders, 1987). Smoltification is generally characterized by increased salinity tolerance, metabolic rate and scope for growth (McCormick et al., 1998). It was shown (Carey and McCormick, 1998) that Atlantic salmon smolts are more sensitive to stressors than younger, non-adapted, fish. This suggests that smolts, during their seaward migration, could be adversely affected by chemical stressors as hypothesized in (Fairchild et al., 1999).

To test this hypothesis, Atlantic salmon smolts were exposed to NP. Initial results indicate that smolts survive exposure to NP in freshwater but exposed fish may not grow as well when they are transferred to seawater.

In order to discuss changes in growth as a consequence of experimental manipulation, it is necessary to understand normal growth of salmon in our lab. It has been established that growth o Atlantic salmon smolts is affected by water temperature and by salinity (Hansen and Quinn, 1998, Handeland et al. 2008)We present here the statistical analysis of the growth (length and weight) of salmon during freshwater holding and after transfer to seawater. In subsequent papers we will describe experimental results from studies designed to assess the effects of chemical exposure on growth. The statistical approach employed and the results of our analyses will be used to help assess the effects of chemical treatment on the growth of Atlantic salmon smolts.

Previous studies have described statistical methods for assessing fish growth (for example see Miller 2004, Newman 2000, Beckman et al., 2004). These studies have focused on wild stocks or stocks of salmon being held and raised for aquaculture purposes. In this article we propose a new statistical approach to describe the trends in growth of Atlantic salmon smolts held under lab conditions at the St. Andrews Biological Station (SABS). The salmon were monitored from their fresh water phase, through PST and seawater transfer and for several months thereafter.

This approach herein described contains preliminary data processing (grouping of experimental data, detection and removal of outliers, and estimation of the character of error in experimental data), and step-wise procedure of building of mathematical models. Our main goal was to design methodology for reliable quantitative estimation of the tendency of the growth of smolts.

METHODS

Fourteen-month post-hatch Atlantic salmon parr were obtained from the Huntsman Marine Science Centre, Chamcook Hatchery, St. Andrews, NB, Canada, in the beginning of January, 1999 and were transferred to SABS. On January 5-7, 1999, 800 salmon parr were anesthetized in 1% tert-amyl alcohol and individually tagged with passive integrated transponder (PIT) tags (Biomark, Boise, Idaho). Fish were randomly distributed into 16 fiberglass tanks (400 L, 50 fish per tank).

All fish were held until May, 1999 in dechlorinated St. Andrews, NB, municipal water at ambient temperature. The tanks were covered and enclosed in individual compartments, each with their own water supply and light. The flow rate was maintained at approximately 5 $L \cdot min^{-1}$. Photoperiod was regulated to simulate natural photoperiod. The fish were fed by hand twice daily to satiation with a premium quality formula diet (Moore-Clark (a division of Nutreco Canada Inc.) St. Andrews, NB).

Fish in 11 tanks were exposed to chemicals over the period from mid May to early June at three different occasions: early window (12-16 May), middle window (26-30 May), and late window (9-13 June). Beginning from 12-14 days after the onset of the treatment at each time window, fish were gradually acclimated to filtered sea water over a five-day period. The flow was maintained at approximately 5 $L \cdot min^{-1}$. Also, salmon in fresh water fresh water (in five tanks without chemical treatment of fish) were transferred to seawater in the same manner: for two tanks – at the beginning of June, for two tanks – at the middle of June, and for one tank – at the end of June. At the end of August 1999 the fish were moved to two large tanks and held until the final sampling in October.
The description of chemical treatment procedure is presented in (Arsenault et al., 2004).

For the present article, we used only the data obtained from fish that were not exposed to chemicals. This includes data from all 16 tanks prior to the onset of chemical treatment (January to May-June) and the data from the five tanks without chemical treatment in the second period of experiment (June to October).

Length and weight data was collected from all fish during PIT tagging. Some fish were euthanized and tissues were sampled for biochemical analyses during the period from May until November. (288 days). Assessment of growth (length and weight) was made by comparing their size at sampling to their size at PIT tagging. All surviving fish were measured in late July, mid September and at the end of the experiment.

The program packages STATA and MS EXCEL were used to perform the necessary calculations and to draw figures.

Statistical Methods Development

Statistical processing of data was performed for length and weight measurements independently. It included two parts:
Part 1. Preliminary processing: the grouping of experimental data in time series, detection and removal of outliers, and evaluation and modeling of measurement error.
Part 2. Development and validation of mathematical models for the study of dependences of the growth characteristics on time.

Number of Coordinate	Day	Number of Coordinate	Day	Number of Coordinate	Day
1	19 (3)	13	113 (6)	25	161 (32)
2	20 (10)	14	114 (18)	26	195 (32)
3	21 (3)	15	115 (5)	27	196 (27)
4	47 (3)	16	125 (2)	28	197 (36)
5	48 (10)	17	126 (4)	29	202 (18)
6	49 (3)	18	127 (2)	30	203 (19)
7	82 (3)	19	132 (4)	31	230 (19)
8	83 (10)	20	140 (4)	32	231 (71)
9	84 (3)	21	141 (4)	33	279 (9)
10	109 (3)	22	154 (2)	34	280 (34)
11	110 (8)	23	155(2)	35	287 (10)
12	111 (3)	24	160 (10)	36	288 (39)

Table 1. Dates of measurements [number of days after beginning of experiment]. Number of measurements for each coordinate is written in brackets

The measurements of length and weight were taken successively for many days (see Table 1). We treated experimental data as time series where each measurement corresponds to a certain point on the time scale. The accuracy of measurements was the following: 0.1 centimeter (cm) for length; 0.1 gram (g) for weight and 1 day for time.
The distribution of length and weight of the salmon at first measurement (during PIT tagging, T=0) is shown in Figures 1A and 1B.

Part 1. Preliminary processing of data

The procedure of grouping experimental data in time series contained two steps:
- creation of the initial time point of experiment with the evaluation of the differences between all 16 tanks at the start of experiment (we had 50 measurements for each of 16 tanks);
- definition of the other ranges or periods of time on the basis of selection of reasonable adjacent time intervals with insignificant differences of means of measurements.

The grouping of data is an essential step of the preliminary data treatment. Growth of salmon smolts is a complex biological process that has certain rate. While the length and weight increase daily, researchers need a couple of weeks to detect an actual increase.

Without grouping data together, we may find that, on successive days, growth characteristics vary somewhat, even showing a decrease over short time frames. The other goals of grouping the data were to detect and remove outliers, and evaluate properties of the experimental error.

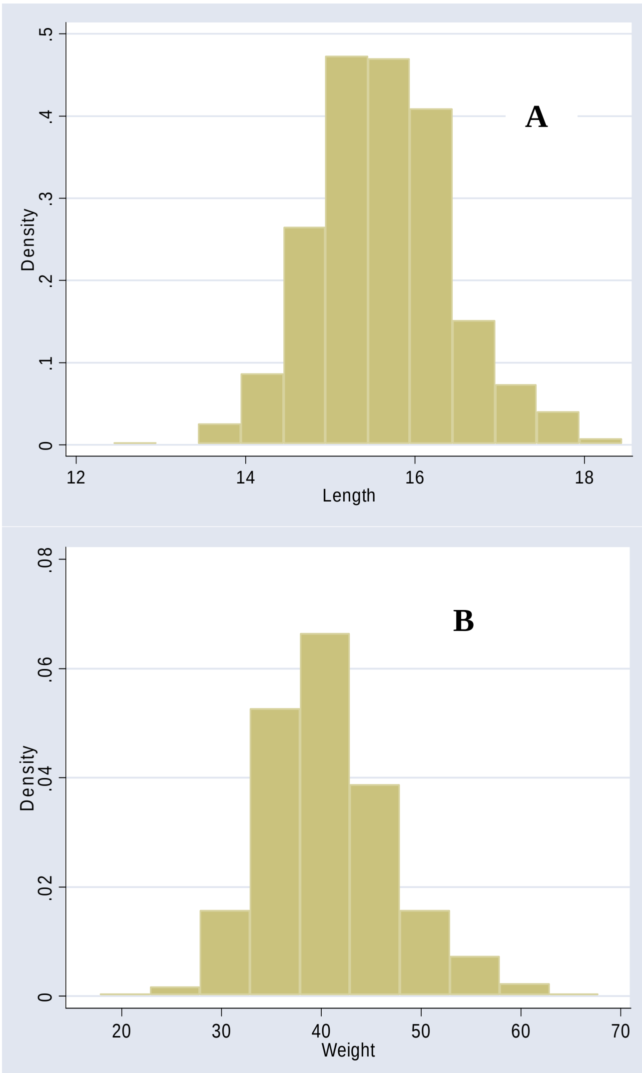
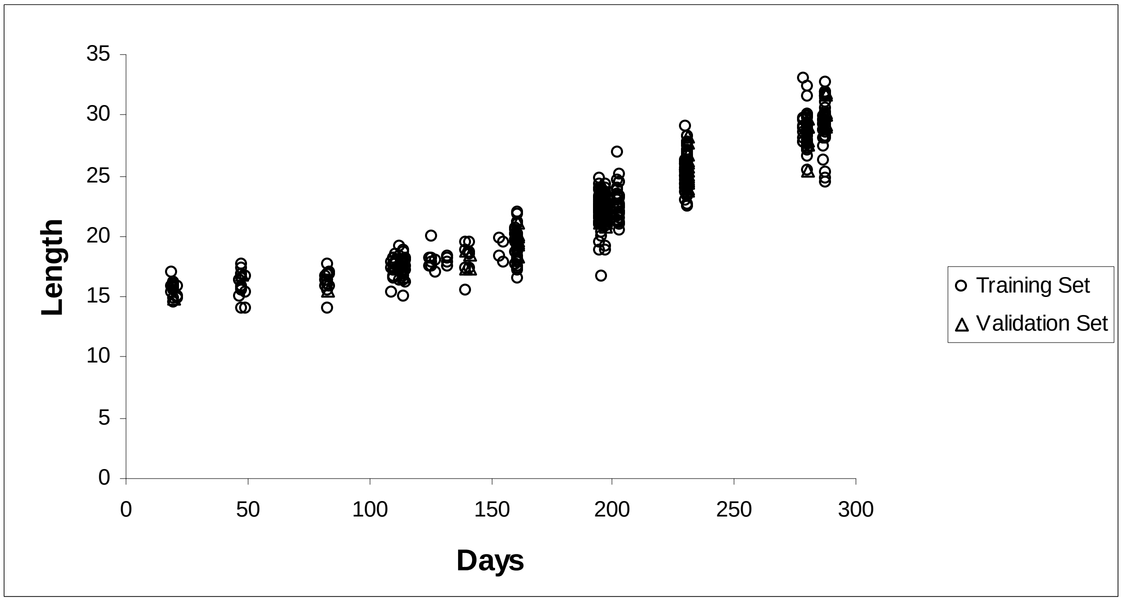


Figure 1: The histogram for T = 0 of: **A**) Length (cm); **B**) Weight (g)

Figures 2A and 2B show the change in length and weight, respectively, of salmon from January through October. The data used for validation of developed mathematical models is also shown.

To realize the procedure of grouping of experimental data we used statistical tests for:
- evaluation of homogeneity of variances (Cochran test or F-test) (Draper and Smith, 1998)
- estimation of the differences between means (t-test) (Draper and Smith, 1998).

A



B

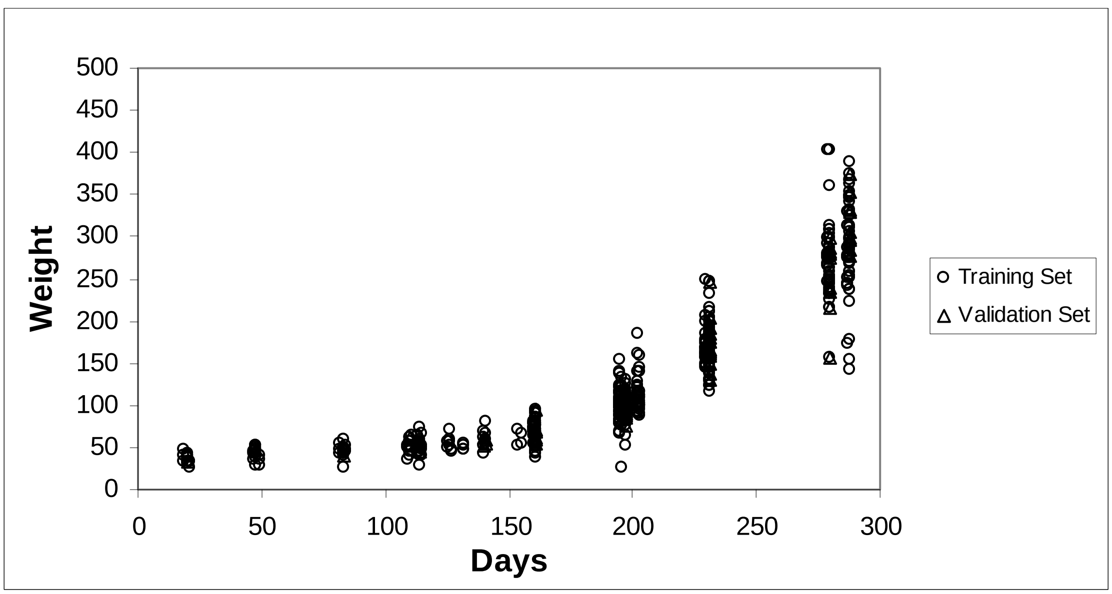


Figure 2. Growth (length in cm (**A**) and weight in grams (**B**)) of Atlantic salmon over the full experimental period. Training set = predicted growth from mathematical models; Validation set = actual measurements.

As result of grouping of data we identified the following times:
T = (0, 20, 48, 83, 112, 133, 160, 198, 231, 284) **(1)**
The grouping procedure reduced by about four times the initial dimension of experimental data, and facilitated further study of dependences time-length and time-weight.

While we recognized that size of fish was naturally variable in populations, our goal was to assess trends in growth with time. Outliers (atypical measurements) were therefore removed. The detection and removal of outliers was done in accordance with traditional methods on the basis of histogram technique (Barnett and Lewis, 1994). To approximate histograms of length and weight of smolts, we used the β-distribution because they exhibited mono-modal and asymmetric shape with the mean skewed right or left.

On the whole, the number of outliers was small and did not exceed 2 percent of all measurements.

Considering the relationship "mean – standard deviation" we evaluated the character of experimental error. The relationship was different between measurements of weight and measurements of length:
For weight, the standard deviation is an increasing linear function of mean; thus, experimental error is relative for measurements of weight.
SD (weight (g)) = .50 + .14*(Mean) , $38.0 \le \text{mean} \le 290.0$ **(2)**

For length, the relationship is more complicated. On Figure 3 are presented values of means and standard deviations for length measurements, resembling a series of increasing steps. To approximate this function, we performed statistical processing of data including the successive test of the homogeneity of variances, F-test, to determine the beginning and the end of each "step" and to estimate their variances:

SD (length (cm)) = 0.8 when $15.6 \le \text{mean} \le 15.9$
= 1.0 when $16.0 \le \text{mean} \le 17.9$
= 1.4 when $18.0 \le \text{mean} \le 29.0$ **(3)**

These results indicate that we had a mixture of absolute and relative error in the length measurements.

Part 2. Development and validation of mathematical models for the study of dependences of the growth characteristics of Atlantic salmon smolt on time.

We needed to construct equations to describe the dependences $Y_1(t)$ as "Time – Length" and $Y_2(t)$ as "Time – Weight" and to develop a technique for the statistical evaluation of these equations. For these goals, we applied regression analysis techniques.

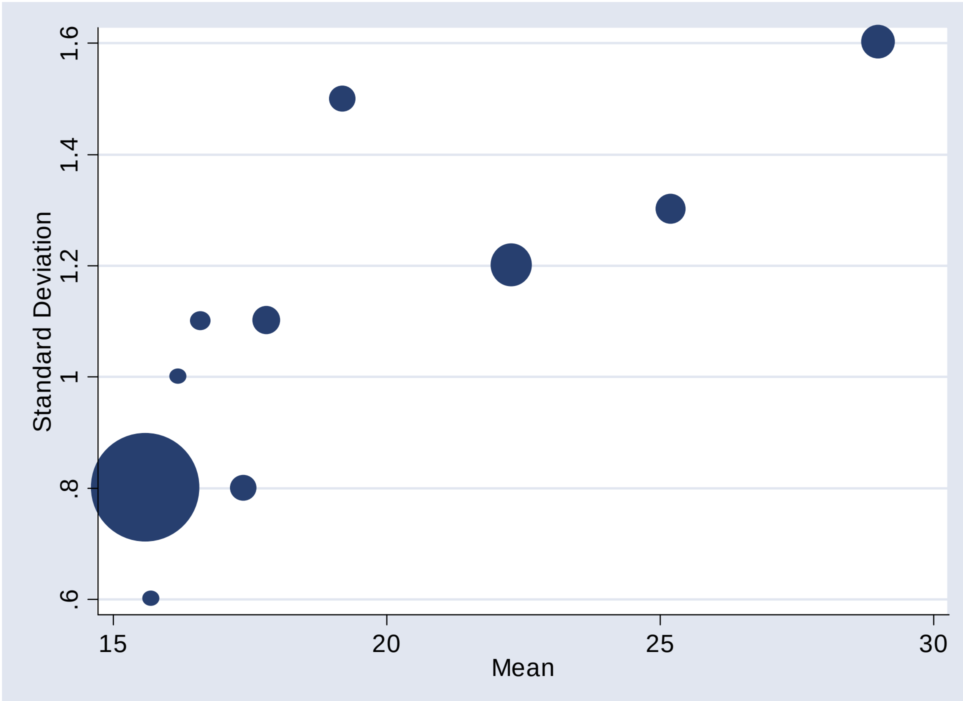


Fig. 3: Values of means and standard deviations for length (cm) measurements (with frequency weight)

For our equations, we used the mean values of length and weight calculated for each of the identified time periods (1). These mean values had different standard deviations.
As a result, we calculated the statistical weights (RW) of each point.
RW is calculated in three steps:

1. Estimation of the standard deviations of measurements (SD_i) with the use of equations (2) for weight and expressions (3) for length;
2. Division of SD_i by the square root of the number of measurements to calculate SD_i(mean);
3. Calculation of RW as inverse values of SD_i(mean).

RW were markedly different for identified times (1). For example, RW was seven times greater for point T = 0 than for point T = 83 (in calculation of coefficients of $Y_1(t)$); RW for point T = 160 was four times greater than for point T = 284 (in calculation of coefficients of $Y_2(t)$).

After obtaining RW for each point, we tried to approximate the dependences $Y_1(t)$ and $Y_2(t)$, $0 \le t \le 284$, by means of polynomial functions of the first and the second order with the method of weighted least squares (WLS). See (Barnhart and Williamson, 2002, Khots et al., 1995). Unfortunately, the results of approximations of both dependences were not satisfactory.

In the case of polynomials of the first order we got $F_{obs}(Y_1(t))= 158.5$, $F_{obs}(Y_2(t))= 238.4$, and $F_{crit}(0.95, 8, \infty)= 1.94$. In the case of polynomial of the second order we got

$F_{obs}(Y_1(t))= 11.5$, $F_{obs}(Y_2(t))= 28.7$, and $F_{crit}(0.95, 7, \infty)= 2.01$.

In F_{crit} (0.95, m, n), first number degrees of freedom (m) is equal to the difference between the number of experimental points and the number of polynomial coefficients. Second number degrees of freedom (n) for the variance was taken as infinite since the actual number n was greater than several hundred and would not have changed any results. P = 0.95 is the confidence level that we apply throughout this research.

The fact that we get an unsatisfactory approximation of the dependences $Y_1(t)$ and $Y_2(t)$ for all t ε [0,284] indicates that there exists at least two non-similar stages (periods of time) of growth of Atlantic salmon smolts during the present study.

A more complicated statistical approach was used to determine the boundaries of regions (critical regions) separating non-similar periods of growth. The main idea behind this approach was to obtain a procedure with step-by-step satisfactory data approximation:
a) Selection of initial set of points as the beginning of time series to build the first segment
b) Calculation of the first regression equation by means of WLS;
c) Test of the equation using F-test;
d) Inclusion of first adjacent point in the initial set followed by steps b) and c);
e) Determination of the end of the first segment after adding all adjacent points which yield satisfactory data approximation;
f) Selection of initial set of points for the second segment in time series.
Then we repeat parts a) through e) for the second segment and so on.

In our experiment, the critical region separating non-similar periods of growth is located between the end of the first segment and the beginning of the second segment.

We considered polynomials of the first order to approximate $Y_1(t)$: initial set contained the first three points of time series: t = 0, 20, 48. Then, we included the point t = 83, and afterwards, point t = 112. The next adjacent point t = 133 was the last one included in the first segment since the calculated regression equation after including point t = 160 (adjacent right point of t = 133) did not give us satisfactory data approximation (Table 2). The next step was to determine the second time segment. After calculation we found the segment to contain the points t = 160, 198, 231, 284. As a result, the critical region was located between t = 133 and t = 160.

A similar procedure was followed for $Y_2(t)$: Preliminary statistical analysis of data showed that a polynomial of the first order would not give satisfactory results. We then used the polynomial of the second order to approximate the data and found the initial set to include the first four points: t = 0, 20, 48, 83. After following the steps outlined above, we determined the two segments to be:

$$[0, 133] \text{ and } [160, 284] \quad \textbf{(4)}$$

Table 2 shows the results of statistical testing of successive equations for $Y_2(t)$.

We would like to emphasize that for both length and weight, we obtained piece-wise polynomial functions for approximation of dependences under study (Barakat et al., 2004).

Also, the critical period separating the two segments was the same for both length and weight.

Set of points	Model for Length Polynomial of 1st Order		Model for Weight Polynomial of 2nd order	
	m	Fobs/Fcrit (0.95, m, ∞)	m	Fobs/Fcrit (0.95, m, ∞)
(0, 20, 48)	1	0.07		
(0, 20, 48, 83)	2	0.09	1	0.16
(0, 20, 48, 83, 112)	3	0.42	2	0.32
(0, 20, 48, 83, 112, 133)	4	0.47	3	0.49
(0, 20, 48, 83, 112, 133, 160)	5	2.41	4	1.25

Table 2. Use of F-test for evaluation of regression equations for length (cm) and weight (g)

We denoted functions under consideration as $Y_{ij(t)}$, i = 1, 2, j = 1, 2, first index (i) represents the type of the growth characteristics: i = 1 corresponds to length, i = 2 corresponds to weight. The second index (j) represents the segment number: j = 1 corresponds to the first segment, j = 2 corresponds to the second segment.

To validate the developed models, we applied the Holdout method (Krizek et al., 2007). We performed the following successive steps:
- selection of validation and training sets (data from one of the tanks with no chemical treatment of smolts was considered as validation set; data from the other tanks was considered as training set), with the training set containing 15 tanks for the first segment and four tanks for the second segment;
- construction of mathematical models on the basis of training sets $Y_{ij,\ trng\ (t)}$, i = 1, 2, j = 1, 2;
- testing of developed models by F-test (Draper and Smith, 1998);
- comparison of experimental data in validation sets with results obtained by mathematical models; we determined $F_{ij\ obs\ val}$ i = 1, 2, j = 1, 2, as the sums of squares of differences between experimental and calculated values divided by the number of degrees of freedom;
- comparison of $F_{ij\ obs\ val}$ with $F_{crit}(0.95, m, \infty)$, where 0.95 is confidence level, m is the number degrees of freedom.
- combination of training and validation sets to produce the mathematical models for all data.

In Figures 4A and 4B are presented piece-wise functional dependences time-length and time-weight.

We estimated the dispersion of functions $Y_{ij(t)}$, i = 1, 2, j = 1, 2, by applying the confidence band technique (Draper and Smith, 1998). For this goal we calculated lower boundary (LB) and upper boundary (UB) of these functions:

$$LB(Y_{ij}(t)) = Y_{ij}(t) - 1.96*(Var(Y_{ij}(t))\verb|^|0.5 \qquad \textbf{(5)}$$

$$UB(Y_{ij}(t)) = Y_{ij}(t) + 1.96*(Var(Y_{ij}(t))\verb|^|0.5 \qquad \textbf{(6)}$$

$$Var(Y_{1,j}(t)) = Var(a_{0,j}) + 2*Cov(a_{0,j}, a_{1,j})*t + Var(a_{1,j})*t\verb|^|2,$$

where $a_{0,j}$ and $a_{1,j}$ are the intercepts and slopes of Y_{1j}, j = 1, 2.

$$Var(Y_{2,j}(t)) = Var(b_{0,j}) + 2*Cov(b_{0,j}, b_{1,j})*t + [Var(b_{1,j}) + 2*Cov(b_{0,j}, b_{2,j})]*t\verb|^|2 + 2*Cov(b_{1,j}, b_{2,j})*t\verb|^|3 + Var(b_{2,j})*t\verb|^|4,$$

where b_{0j} are constant terms in polynomials, b_{1j} are coefficients of linear terms, and b_{2j} are coefficients of quadratic terms, j = 1,2.

The number 1.96 in (5) and (6) corresponds to confidence level p = 0.95.

To evaluate the rate of the weight change, we use the derivatives of functions $Y_{2j(t)}$, j=1,2, and their statistical characteristics, i.e.
$Y'_{2j}(t) = b_{1j} + 2*b_{2j}*t$
$Var(Y'_{2j}(t)) = Var(b_{1j}) + 4*Cov(b_{1j}, b_{2j})*t + 4*Var(b_{2j})*t^2$, j = 1, 2

RESULTS AND DISCUSSION

The statistical approach described here allowed us to study complex biological process of growth of Atlantic salmon smolts in detail.
First of all, we could confirm that there are two essentially different periods of growth of fish because the two time segments (4) were the same for both growth characteristics. It is important that the critical region separating these two periods was located between time t = 133 and t = 160 and corresponded to the period of time between the end of spring and the beginning of summer and to transference of most of the smolts from fresh to sea water. Therefore, statistical inference about piece-wise character of the growth of Atlantic salmon smolts carries important biological sense: the dependences of growth characteristics as functions of time are different for fresh and seawater. Nevertheless, a split of the entire time period into two segments based simply on when the smolt is in fresh and seawater would be too rough and ambiguous: the experimental design allowed for transfer of fish from fresh to salt water over a 4 week period in June. It is impossible to identify a transfer time *per se* and statistical techniques and mathematical models were employed to define the split.

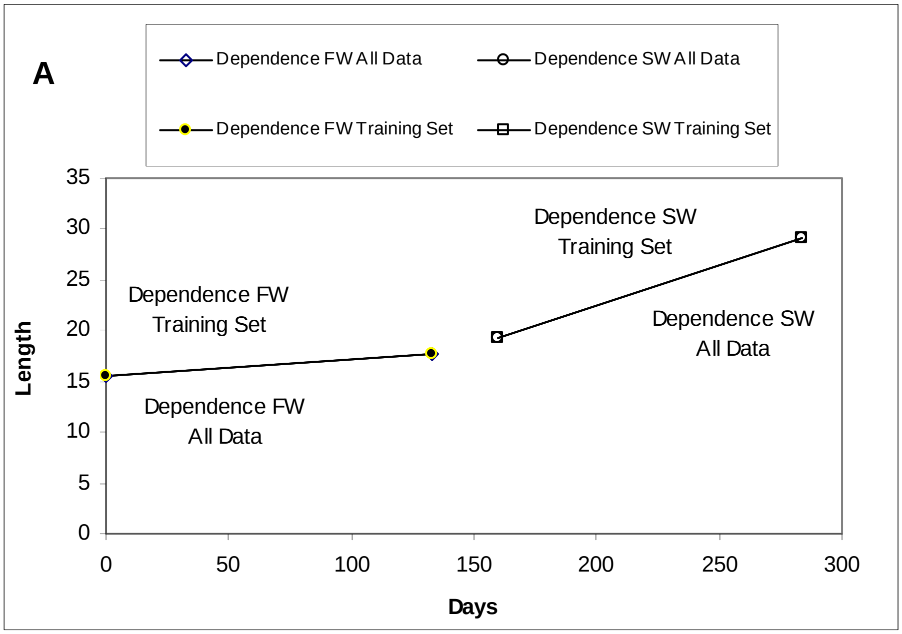
For quantitative analysis of the growth of smolts, we used the corresponding regression equations and their statistical characteristics (Tables 3 and 4).
The independent variables t_1 (for the first segment) and t_2 (for the second segment) were defined in the following manner:
- t_1 was calculated after scaling, i.e. as number of days (t) divided by the length of the first time segment:

$$t_1 = t/133 \quad \textbf{(7)}$$

- t_2 was calculated after shifting and scaling:

$$t_2 = (t-160)/124 \quad \textbf{(8)}$$

(The length of the second time segment is equal to 124 days; number 160 corresponds to the beginning of the second time segment).
For both time segments, t_1 and t_2 ε [0,1].



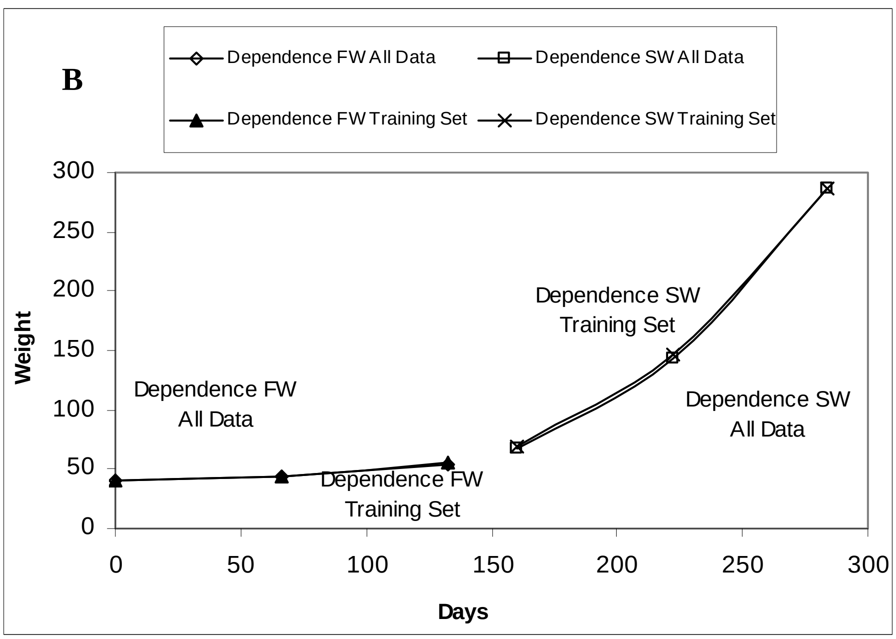


Figure 4. Relationship between growth (length (cm) (**A**); weight (g) (**B**)) and time. Predicted (training) equations are separated for freshwater and seawater conditions.

Observed values $F_{obs}Y_{ij}$, i=1, 2, j=1, 2, were less than $F_{crit}(0.95, m, \infty)$ for the regression equations under study (Tables 3 and 4); also, $F_{obs\ val}\ Y_{ij\ trng}$, i=1, 2, j=1, 2 were less than $F_{crit}(0.95, m, \infty) = 1.34$ (Table 3). In the validation procedure, the number of degrees of freedom (m) ranged from 55 to 60.

To compare the growth of smolts in fresh and seawater, we considered the rate of change of length and weight of fish using all available data (Table 4 & 5).

Models	Constant Term	Coefficient of the Term of the First Order	Coefficient of the Term of the Second Order	Fobs trng	F_{crit} (0.95, m, ∞)	Fobs val	F_{crit} (0.95, m, ∞)
Y11, trng(t1)	15.57	2.2		2.3	2.37 (m=4)	0.94	1.34
Y21, trng(t1)	40.4	-3.9	19.4	1.89	2.60 (m=3)	1.18	1.34
Y12, trng(t2)	19.31	9.93		0.88	3.00 (m=2)	0.59	1.34
Y22, trng(t2)	69.1	90	128.1	2.54	3.84 (m=1)	0.81	1.34

Table 3. Validation of models Time - Length (cm) and Time - Weight (g)

Models	Constant Term	Coefficient of the Term of the First Order	Coefficient of the Term of the Second Order	Fobs	F_{crit} (0.95, m, ∞)
Y11(t1)	15.59	2.12		1.1	2.37 (m=4)
Y21(t1)	40.6	1.2	12.8	1.28	2.60 (m=3)
Y12(t2)	19.32	9.79		1.9	3.00 (m=2)
Y22(t2)	67.6	83.9	134.7	2.02	3.84 (m=1)

Table 4. Models Time – Length (cm) and Time – Weight (g) and their statistical characteristics **A**.

Models	Var(Constant Term)	2*Cov(Constant Term, Coefficient of the Term of the First Order)	Var(Coefficient of the Term of the First Order)	2*Cov(Constant Term, Coefficient of the Term of the Second Order)	2*Cov(Coefficient of the Term of the First Order, Coefficient of the Term of the Second Order)	Var(Coefficient of the Term of the Second Order)
Y11(t1)	0.0007	-0.0001	0.0083			
Y21(t1)	0.04	-0.03	6.12	0.04	-11.45	6.82
Y12(t2)	0.0171	-0.0187	0.0268			
Y22(t2)	3.19	-1.17	23.09	1.86	-79.17	87.55

Table 5. Models Time – Length (cm) and Time – Weight (g) and their statistical characteristics **B**.

Study of length

As length is modeled by linear equations, we use the slopes of these lines to evaluate the difference between the rates of the length change in both time segments. In accordance with data in Table 4, slope a_{11} of the straight line on the first time segment is equal to 2.12 and covered by confidence interval (1.89, 2.35). Slope a_{12} of the straight line on the second time segment is equal to 9.79 and covered by confidence interval (9.38, 10.20). To compare the slopes of these straight lines between the first and the second time segments, we first scale variable t_2 using (7) and (8) by multiplying t_2 by the ratio 133/124. Then, we obtain the slope a_{12} equal to 10.50 and confidence interval of (10.07, 10.94). The rate of the length increase is 0.016 cm/day for the first segment and is 0.079 cm/day for the second segment. This means that for laboratory conditions under study the length of salmon in seawater increased several times faster than when the fish were held in freshwater.

Study of weight

As weight is modeled by the polynomials of the second order, the rate of the weight change are the linear functions of independent variables t_1 and t_2 in both time segments. Confidence bands of functions $Y'_{21}(t_1)$ and $Y'_{22}(t_2)$ are presented in Figure 5. After scaling variable t_2 (applying (7) and (8)), we conclude that the rate of the weight change in seawater was statistically more than in freshwater and varied from:

- Small in the very beginning of the experiment (less than 0.01 g/day);
- Increased to about 0.2 g/day to the end of the first time segment;
- Increased greatly (more than 3 times) in the beginning of the second time segment to approximately 0.6 g/day;
- Increased steadily to more than 2.6 g/day to the end of the second time segment.

In accordance with the data in Table 4, the acceleration of the change of weight on the first time segment is equal to 25.6 and is covered by confidence interval (12.5, 38.7). The acceleration of the change of weight on the second time segment is equal to 269.5 and is covered by confidence interval (222.7, 316.2). To compare the accelerations of weight change on both time segments, we again scale variable t_2 using (7) and (8) and obtain the value of acceleration as 289.0 and its confidence interval (238.9, 339.2). Obviously, the difference between mean values of acceleration for the first and the second segment is statistically significant. The mean acceleration of the weight change in freshwater (for laboratory conditions) was eleven times less than in seawater.

These quantitative results can be supported by qualitative biological observations. It is known in principle that Atlantic salmon grow relatively quickly in seawater (Hansen and Quinn, 1998). The reason for this is unclear but is likely related to the consistency of the water temperature and the increased availability of food when the fish are at sea (McCormick et al., 1998).

Handeland et al. (2008) described the effect of temperature on the growth of Atlantic salmon post-smolts for 12 weeks following transfer of fish to seawater. Optimal temperature of growth for different fish sizes was between 12°C and 14°C. Fish weighing 75-150g grew optimally at 12.8°C; fish weighing 150-300g grew optimally at 14°C.

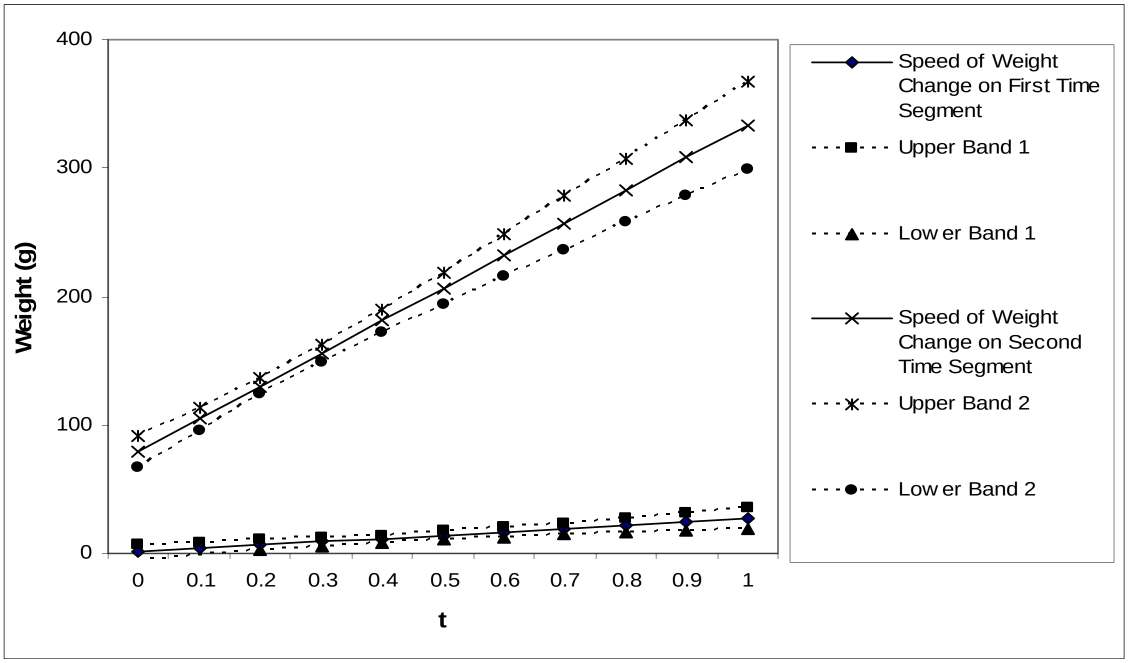


Figure 5: Confidence bands for equations $Y'_{21}(t)$ and $Y'_{22}(t)$

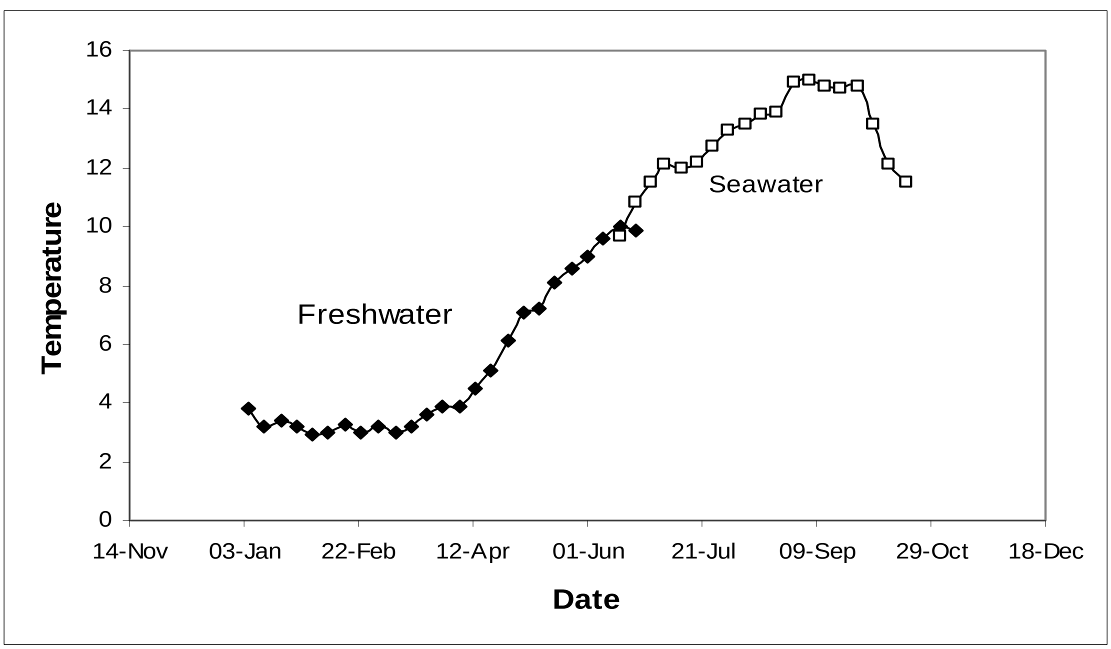


Figure 6: Dependences “Time - Freshwater Temperatures” and “Time - Seawater Temperatures”

Under our lab conditions, ample food was provided to the fish in both cases; freshwater temperatures in the first part of the experiment were almost always less than 6°C. Handeland et al. (2008) did not address growth in freshwater. The temperature of seawater in which our fish were raised (Figure 6) is quite close to the optimum described in (Handeland et al., 2008).

CONCLUSION

A procedure for developing a mathematical model of the growth of Atlantic salmon smolts as a function of time has been proposed. The procedure included preliminary processing of data (grouping of data in time series, detection and removal of outliers, evaluation and modeling of measurements error), calculation of dependences time-length and time-weight as piece-wise functions of time. As expected, growth of smolts is increased in fresh and seawater at periods of time likely related to water temperature in our laboratory, as feed is available at all times. The approach employed in this paper will be useful in analyzing data from experiments designed to evaluate the influence of environmental factors and chemicals on growth of Atlantic salmon smolts.

ACKNOWLEDGEMENTS

Dr. Scott Brown, a co-author on this manuscript, passed away unexpectedly in 2006. He was a driving force behind this research and a key member of our research team. His knowledge, insight and friendship are missed.
The authors wish to acknowledge the assistance of Ms. Monica Lyons, Ms Dawn Sephton, Ms. Jacqueline Arsenault and Mr. Ken MacKeigan in conducting these experiments. This work was funded by the Toxic Substances Research Initiative and by the Environmental Sciences Strategic Research Fund.

REFERENCES

Arsenault J.T.M., Fairchild W.L., MacLatchy D.L., Burridge L., Haya K., Brown S.B. 2004. Effects of water-borne 4-nonylphenol and 17β-estradiol exposures during parr-smolt transformation on growth and plasma IGF-I of Atlantic salmon (Salmo salar L.). Aquatic Toxicol. 66 (3): 255-265.

Barnett V. and Lewis T. 1994. Outliers in Statistical Data (3rd ed). John Wiley & Sons. New York. 604 p.

Barnhart H. X., Williamson J. M. 2002. Weighted least squares approach for comparing correlated Kappa. Biometrics. 58 (4):1012-1019.

Barakat N.A.M., Jiang J.-H., Liang Y.-Z., Yu R.-Q. 2004. Piece-wise quasi-linear modeling in QSAR and analytical calibration based on linear substructures detected by genetic algorithm. Chemom. Intell. Lab. Syst. 72 (1): 73-82.

Beckman B.R., Fairgrieve W., Cooper K.A., Mahnken M.V.W., Beamish R.J. 2004. Evaluation of endocrine indices of growth in individual postmolt Coho Salmon. Trans. Am. Fish. Soc. 133 (5): 1057-1067.

Carey J.B. and McCormick S.D. 1998. Atlantic salmon smolts are more responsive to an acute handling and confinement stress than parr. Aquaculture. 168:237-253.

Draper N.R.and Smith H. 1998. Applied Regression Analysis (3rd ed). John Wiley & Sons. New York. 736 p.

Fairchild W.L., Swansburg E.O., Arsenault, J.T., Brown, S.B. 1999. Does an association between pesticide use and subsequent declines in catch of Atlantic salmon (Salmo salar) represent a case for endocrine disruption? Environ. Health Perspect. 107 (5): 349-357.

Handeland S., Imsland A., Stefansson S. 2008. The effect of temperature and fish size on growth, feed intake, food conversion efficiency and stomach evacuation rate of Atlantic salmon post-smolts. Aquaculture. 283 (1-4): 36-42.

Hansen L.P. and Quinn T.P. 1998. The marine phase of the Atlantic salmon (Salmo salar) life cycle, with comparisons to Pacific salmon. Can. J. Fish. Aquat. Sci. 55 (Suppl. 1): 104-118.

Khots M., Nazarov V., Emelyanov V. 1995. Application of regression analysis for motor gasoline classification. J. of Chemom. 9 (5): 411-422.

Krizek P., Kittler J., Hlavac V. 2007. Improving Stability of Feature Selection Methods. In: Lecture Series Notes in Computer Science, Proceedings of the 12th International Conference on Computer Analysis of images and patterns. Springer-Verlag, Berlin- Heidelberg. pp. 929-936.

McCormick S.D. and Saunders R.L. 1987. Preparatory Physiological Adaptations for Marine life of Salmonids: Osmoregulation, Growth, and Metabolism. American Fisheries Society Symposium 1: 211-229.

McCormick S.D, Hansen L.P., Quinn, T.P., Saunders, R.L. 1998. Movement, Migration and Smolting of Atlantic Salmon (Salmo salar) Can. J. Fish. Aquat. Sci. 55 (Suppl. 1): 77-92.

Millar R. B. 2004. Sensitivity of Bayes estimators to hyper-parameters with an application to maximum yield from fisheries. Biometrics. 60 (2): 536-542.

Newman K. B. 2000. Hierarchic Modeling of Salmon Harvest and Migration. Journal of Agric. Biol. Environ. Stat. 5 (4): 430-455.