On the 2008 World Fly Fishing Championships

Thomas Yee

University of Auckland

30 November 2009

t.yee@auckland.ac.nz

http://www.stat.auckland.ac.nz/~yee

Outline of This Talk

1. Introduction I

The 2008 World Fly Fishing Championships (WFFC)

- *When* 26–28 March 2008.
- *Where* Central North Island of NZ.
- *Background* Event organized annually by the FIPS-MOUCHE[1] and held successively in different countries. It attracts top anglers from all over the world including some professional teams.

1981 . . .	2007 Finland
2005 Sweden	2008 NZ
2006 Portugal	2009 Scotland

[1] The competition is organized by the Fédération Internationale de la Pêche Sportive Mouche or International Federation of Sport Fly Fishing, which is a federation of the umbrella organization Confédération Internationale de Pêche Sportive, or CIPS.

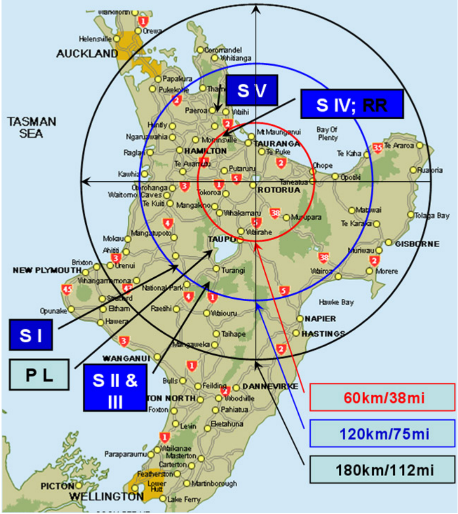
1. Introduction II

Venues

Sector	Water	Description
I	Whanganui River†	Larger free stone bottom
II	Lake Otamangakau†	2 km^2; maximum depth 12 m
III	Lake Rotoaira	13 km^2
IV	Waihou River	Small spring-fed stream
V	Waimakariri River	Small spring-fed stream

Table: The five trout-infested sectors. All have rainbow trout, and those daggered hold brown trout.

1. Introduction III



1. Introduction IV

Trout species

All sectors have naturally *wild Rainbow trout* (*Oncorhynchus mykiss*) while Lake Otamangakau and the Whanganui River also hold *Brown trout* (*Salmo trutta*). Only these species were targetted.

By 'wild', both species were introduced into the Taupo region from overseas in the late 1800s and since then the populations have become self-sustaining.

1. Introduction V

WFFC regulations and some competition details

1. *Catch-and-release* $\Longrightarrow$ "*catch reduction*", not "*fish depletion*". Barbless hooks used.
2. 19 countries/teams, one was a composite. Each team had five individual competitors labelled A, B, C, D, E. About 100 competitors in total.

   Some teams had reserves which could replace a member who was sick etc.

Session	Day	Time	Session	Day	Time	Session	Day	Time
1	1	morning	3	2	morning	5	3	morning
2	1	afternoon	4	2	afternoon	6	3	afternoon

Table: The sessions. Morning and afternoon sessions were 9.00am–12.00pm and 2.30pm–5.30pm. Each sector had a unique *resting session*.

1. Introduction VI

Table: Abbreviations for the countries represented.

Australia	AUS	Japan	JPN
Canada	CAN	Malta	MAL
Croatia	CRO	New Zealand	NZL
Czech Republic	CZE	Poland	POL
England	ENG	Portugal	POR
Finland	FIN	South Africa	RSA
France	FRA	Slovakia	SVK
Holland	NED	USA	USA
Ireland	IRE	Wales	WAL
Italy	ITA		

1. Introduction VII

3. A large part of each river was divided into 19 "*beats*"—contiguous downstream stretches of approximate length of 400 m.

   All beats had one competitor fishing during every (non-resting) session.

   The concept of a beat does not exist on a lake under WFFC rules. Instead, two competitors shared the same boat, which was piloted by a controller to whatever part of the lake desired by the competitors.

   Thus "beat/boat" specifies the location within each sector the competitors fished.

4. The (competitor, session, sector) combinations were randomized the night before Day 1.

1. Introduction VIII

5. WFFC scoring system: 100 points to each eligible fish (minimum length was 18 cm) and 20 points for each cm of its length (rounded up to the nearest centimeter).

   At each (sector, session) combination, the number of points was summed, then ranked into 1, 2, . . . , 19 (*placings*). These placings were summed over the sessions to give *total placings*.

6. Gold, silver and bronze medals were awarded at the team and individuals levels.

7. For want of time, there are lots of (mainly small) details not mentioned in this talk, e.g.,
   - Impartial (local and overseas) judges were on hand to handle disputes.
   - Missing data: $< 0.6\%$ of cases needed adjustment.
   - After each day of competition, data entry from the score sheets to a spreadsheet was performed efficiently, along with queries decided by sector and international judges. The competitors knew their comparative ranking late that evening.

1. Introduction IX

Two types of analyses

1. *Fish length analysis*

   What distribution do the fish lengths have in each of the sectors?

   How does the distribution vary as a function of number of fish caught per competitor?

2. *Catch reduction analysis*

   Did the fish populations suffer from measurable catch reduction over the duration of the competition?

   In particular, is there any evidence that the smaller rivers suffered more pronounced catch reduction over successive days of fishing?

   If so, by how much and how can the effects be ameliorated?

All regression analyses in this talk were performed using the VGAM package for R.

2. VGLMs and VGAMs I

Data $(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_n, \mathbf{y}_n)$ on n independent "individuals".

Definition Conditional distribution of $\mathbf{y}$ given $\mathbf{x}$ is

$$f(\mathbf{y}|\mathbf{x}; \boldsymbol{\beta}) = h(\mathbf{y}, \eta_1, \ldots, \eta_M),$$

where for $j = 1, \ldots, M$,

$$\begin{aligned} \eta_j &= \eta_j(\mathbf{x}) = \boldsymbol{\beta}_j^T \mathbf{x}, \qquad \text{(VGLM)} \\ \boldsymbol{\beta}_j &= (\beta_{(j)1}, \ldots, \beta_{(j)p})^T, \\ \boldsymbol{\beta} &= (\boldsymbol{\beta}_1^T, \ldots, \boldsymbol{\beta}_M^T)^T. \end{aligned} \tag{1}$$

Often $g_j(\theta_j) = \eta_j$ for parameters θ_j and link functions g_j.

Nb. $-\infty < \eta_j < \infty$.

Given the covariates, the conditional distribution of the response is purposely as general as possible.

2. VGLMs and VGAMs II

VGLM Examples

1. The standard *two parameter gamma distribution*

   Density function

$$f(y;r,s) \;=\; \frac{e^{-ry}\,y^{s-1}\,r^{s}}{\Gamma(s)}, \qquad y > 0, \tag{2}$$

   where $r > 0$ and $s > 0$ (the rate and shape parameters). Then $E(Y) = \mu = s/r$ with $\mathrm{Var}(Y) = \mu^2/s = s/r^2$.

   Default:

$$\boldsymbol{\eta} \;=\; \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \;=\; \begin{pmatrix} \log\, r \\ \log\, s \end{pmatrix}.$$

   We fit (2) to $Y =$ fish length $-$ 18 cm since only eligible-sized fish were recorded.

2. VGLMs and VGAMs III

2. The *bivariate (logistic) odds-ratio model*

   Data: (Y_1, Y_2) where $Y_j = 0$ or 1.

   $Y_1 = 1/0$ if fish caught in Waihou/Waimakariri River,
   $Y_2 = 1/0$ if fish caught in afternoon/morning.

$$\begin{aligned} p_j &= P(Y_j = 1), \quad \text{marginal probabilities,} \\ p_{rs} &= P(Y_1 = r, Y_2 = s), \quad r, s = 0, 1, \quad \text{joint probabilities,} \\ \psi &= p_{00}\, p_{11} / (p_{01}\, p_{10}) \qquad \text{(Odds ratio).} \end{aligned}$$

   Model:

$$\begin{aligned} \operatorname{logit} p_j(\mathbf{x}) &= \eta_j(\mathbf{x}), \qquad j = 1, 2\,, \\ \log\ \psi(\mathbf{x}) &= \eta_3(\mathbf{x}). \end{aligned}$$

   Recover p_{rs}'s from p_1, p_2 and ψ. Here, $\mathbf{x} = (1, \text{length})^T$.

2. VGLMs and VGAMs IV

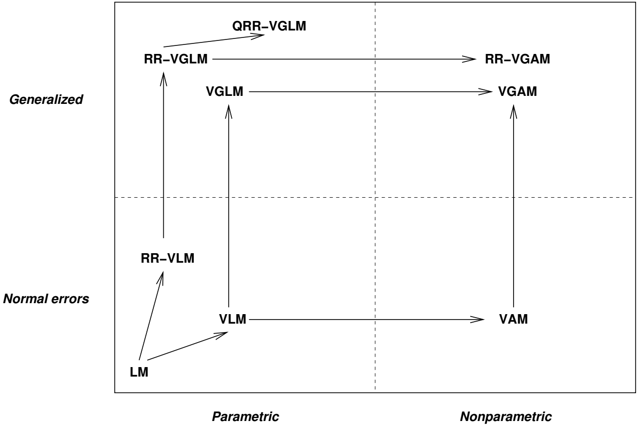


Figure: Flowchart for different classes of models. Legend: LM = linear model $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$, V = vector, G = generalized, A = additive, RR = reduced-rank.

2. VGLMs and VGAMs V

t

$\boldsymbol{\eta} = (\eta_1, \ldots, \eta_M)^T$	**Model**	**S function**	**Reference**
$\mathbf{B}_1^T \mathbf{x}_1 + \mathbf{B}_2^T \mathbf{x}_2 \ (= \mathbf{B}^T \mathbf{x})$	VGLM	`vglm()`	Yee & Hastie (2003)
$\mathbf{B}_1^T \mathbf{x}_1 + \sum_{k=p_1+1}^{p_1+p_2} \mathbf{H}_k \, \mathbf{f}_k^*(x_k)$	VGAM	`vgam()`	Yee & Wild (1996)
$\mathbf{B}_1^T \mathbf{x}_1 + \mathbf{A} \, \boldsymbol{\nu}$	RR-VGLM	`rrvglm()`	Yee & Hastie (2003)
$\mathbf{B}_1^T \mathbf{x}_1 + \mathbf{A} \, \boldsymbol{\nu} + \begin{pmatrix} \boldsymbol{\nu}^T \mathbf{D}_1 \boldsymbol{\nu} \\ \vdots \\ \boldsymbol{\nu}^T \mathbf{D}_M \boldsymbol{\nu} \end{pmatrix}$	QRR-VGLM	`cqo()`	Yee (2004)
$\mathbf{B}_1^T \mathbf{x}_1 + \sum_{r=1}^{R} \mathbf{f}_r(\nu_r)$	RR-VGAM	`cao()`	Yee (2006)

Table: VGAM & its framework. The latent variables $\boldsymbol{\nu} = \mathbf{C}^T \mathbf{x}_2$, $\mathbf{x}^T = (\mathbf{x}_1^T, \mathbf{x}_2^T)$. More abbreviations: C = constrained, O = ordination, Q = quadratic.

2. VGLMs and VGAMs VI

VGLM:

$$\eta_j(\mathbf{x}) \;=\; \boldsymbol{\beta}_j^T \mathbf{x} \;=\; \beta_{(j)1}\, x_1 + \cdots + \beta_{(j)p}\, x_p \tag{3}$$

VGAM:

$$\eta_j(\mathbf{x}) \;=\; \beta_{(j)1} + f_{(j)2}(x_2) + \cdots + f_{(j)p}(x_p), \tag{4}$$

a sum of arbitary smooth functions.

The VGAM package for R

`VGAMrefcard.pdf` is a summary.

VGAM is on CRAN, and has the data frames `wffc`, `wffc.nc`, `wffc.indiv` and `wffc.teams`.

3. Fish length analyses I

Table: Some basic summary statistics for the WFFC. Lengths are in cm. Mean *CPUE (catch per unit effort)* is the number caught per fishing hour, averaged over the entire competition. The bottom portion of the table are CPUEs for some other New Zealand waters fished by the general population in 2005/06; Source: Venman (2006).

	Waihou	Waimakariri	Whanganui	Otam.	Rotoaira	Total
Number of fish caught	1208	1276	1310	273‡	201†	4268
Mean length	22.8	23.0	26.1	33.2	43.8	25.53
Median length	21.5	22.1	21.5	23.6	45.0	22.00
SD length	4.61	4.50	9.54	14.57	8.68	8.97
Mean CPUE	4.24	4.48	4.60	0.96	0.71	3.00
No. of Rainbow females				164‡	125†	
No. of Rainbow males				46‡	54†	
No. of Brown females				1‡	1†	
No. of Brown males				10‡	2†	
Missing values				52‡	19†	
Lake Taupo CPUE						0.17–0.35
Tongariro R. CPUE						0.24–0.38
Tauranga-Taupo R. CPUE						0.28–0.29

3. Fish length analyses II

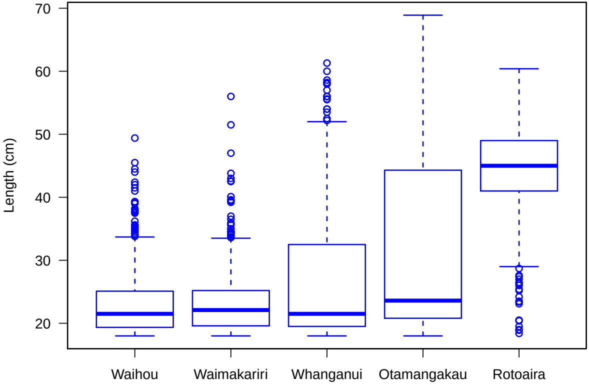


Figure: Length of the fish (cm), for each sector.

3. Fish length analyses III

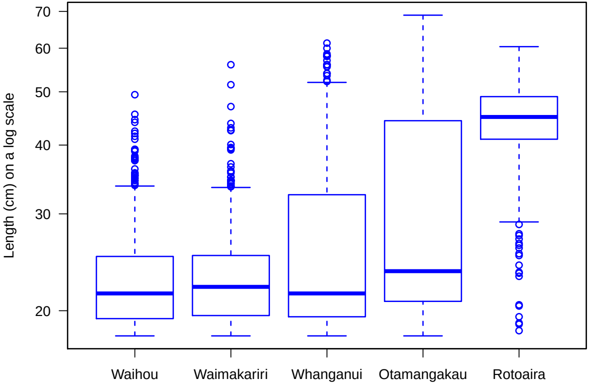


Figure: Length of the fish (cm), on a log scale, for each sector.

3. Fish length analyses IV

Density estimation

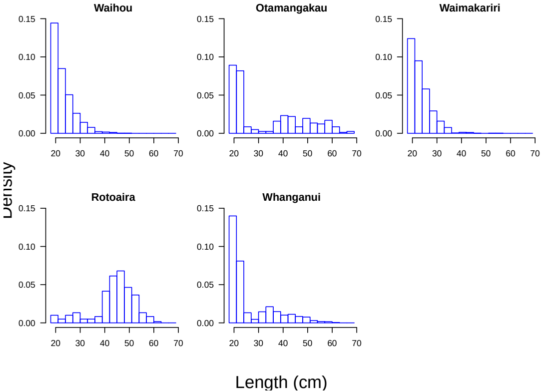


Figure: Histogram of fish lengths, by sector.

3. Fish length analyses V

Density estimation

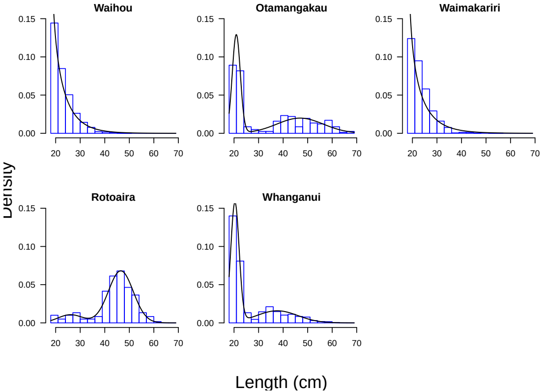


Figure: Histogram of fish lengths, by sector. The estimated pdfs are two 2-parameter gamma, and a mixture of two normal distributions.

3. Fish length analyses VI

Extreme value analysis

We investigate the relationship between $Y =$ longest fish caught and $X =$ number of captures, and compare the results with quantile regression.

Consider a fixed sector (Waihou River) and let N be the number of captures obtained by a particular competitor. Let Y_j be the length of the jth fish for that competitor. Then $\max\{Y_1, \ldots, Y_N\}$ is the competitor's longest fish and has an approximate *generalized extreme value (GEV)* distribution if N is sufficiently large.

The GEV cumulative distribution function is

$$G(y; \mu, \sigma, \xi) \;=\; \exp\left\{ - \left[1 + \xi \left(\frac{y - \mu}{\sigma} \right) \right]_{+}^{-1/\xi} \right\}, \quad \sigma > 0, \quad -\infty < \mu < \infty,$$

$1 + \xi(y - \mu)/\sigma > 0$, where $x_+ = \max(x, 0)$. The μ, σ and ξ are known as the *location*, *scale* and *shape* parameters respectively.

3. Fish length analyses VII

Quantile regression

VGAM fits *three* classes of quantile regression methods. One is the LMS-normal method: it assumes a Box-Cox power transformation of $Y|X = x$ is $N(0, 1)$. That is,

$$Z = \left[(Y/\mu(x))^{\lambda(x)} - 1\right] / \{\sigma(x)\,\lambda(x)\}\,, \qquad \lambda(x) \neq 0. \tag{6}$$

Default: $\boldsymbol{\eta}(x) = (\lambda(x),\ \mu(x),\ \log(\sigma(x)))^T$.

We have $\eta_1 = \beta_{(1)1}$, $\eta_3 = \beta_{(3)1}$,

$$\eta_2 = \beta_{(2)1} + f_{(2)2}(x_2), \tag{7}$$

where $x_2 = x =$ number of captures. Applications include detecting cheating in fishing competitions, e.g., Tolonen and Lappalainen (1999).

3. Fish length analyses VIII

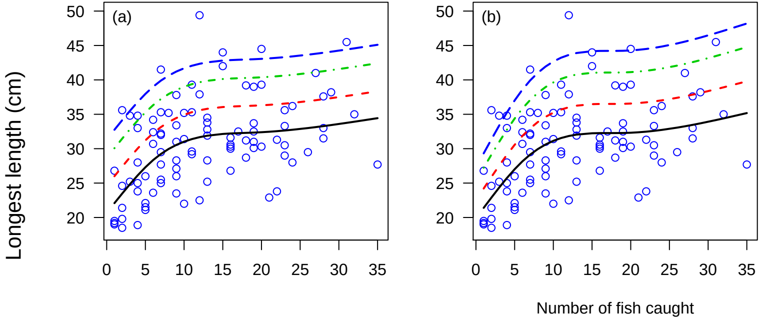


Figure: For each competitor on the Waihou River. The curves are the 50, 75, 90 and 95 percentiles of a fitted (a) GEV model with $\mu(x)$ modelled with a regression spline with 4 df; (b) LMS-Box-Cox-normal model using a regression spline with 4 df.

4. Catch reduction analysis I

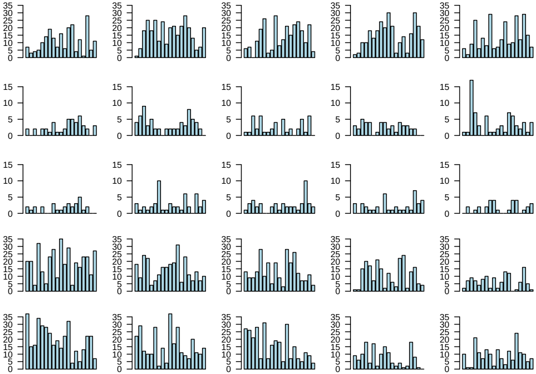


Figure: Numbers of captures at all the sectors. Sectors I–V are the rows. The 5 fishing sessions are the columns, and there are 19 beats.

4. Catch reduction analysis II

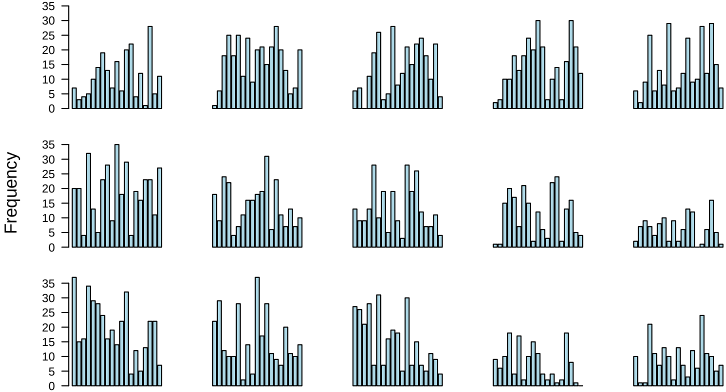


Figure: Subset from Slide 26. Rivers only (rows; Sectors I, IV, V).

4. Catch reduction analysis III

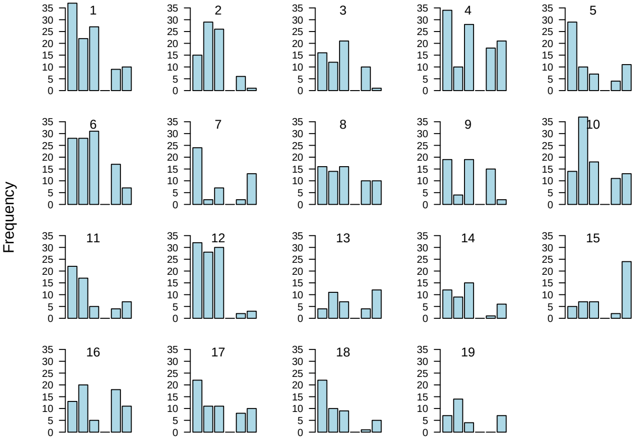


Figure: Numbers of captures, for Sector V. There are 19 beats and 6 sessions.

4. Catch reduction analysis IV

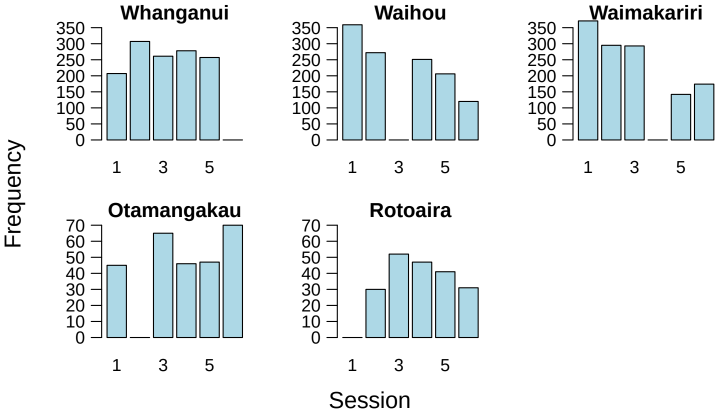


Figure: Numbers of captures by session, for all sectors. Each bar is summed over the 19 beats. One of the sessions is resting.

4. Catch reduction analysis V

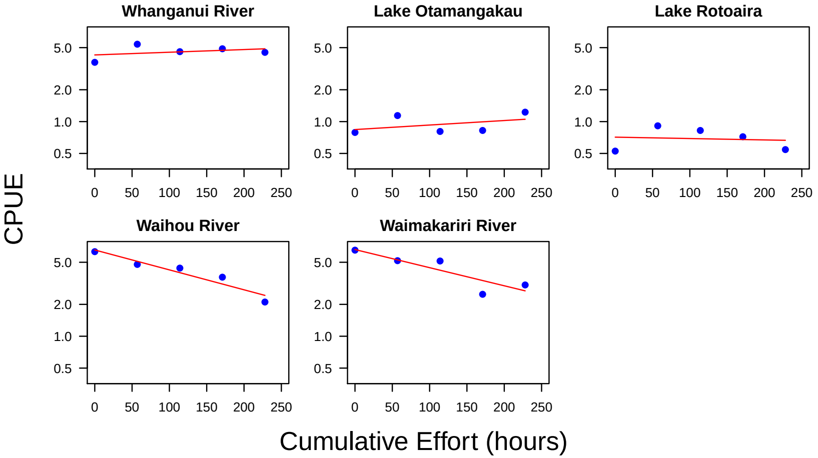


Figure: The DeLury model. Note: the assumptions do not hold with WFFC data!

4. Catch reduction analysis VI

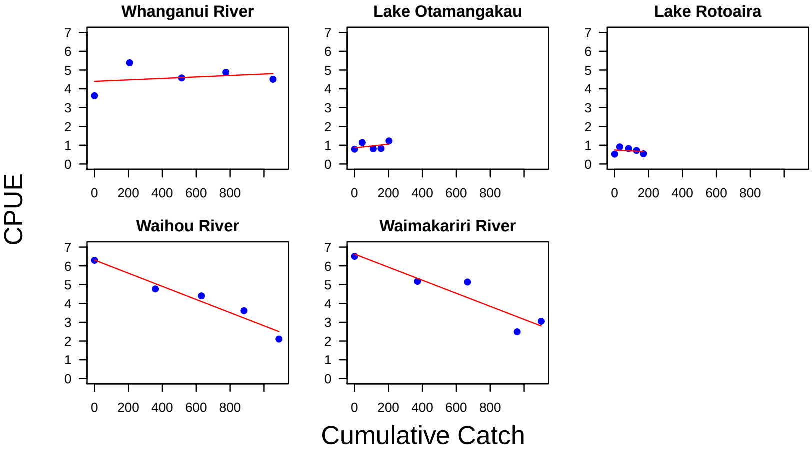


Figure: The DeLury model. Note: the assumptions do not hold with WFFC data!

4. Catch reduction analysis VII

The DeLury model assumptions

1. Catch and effort records are available for a series of consecutive time intervals. The catch for a given time interval, specified by t, is $c(t)$, and the corresponding effort by $e(t)$. The *catch per unit effort (CPUE)* for the time interval t is $C(t) = c(t)/e(t)$. Let $d(t)$ represent the proportion of the population captured during the time interval t. Then $d(t) = k(t)e(t)$ so that $k(t)$ is the proportion of the population captured during interval t by one unit of effort. Then $k(t)$ is called the *catchability*, and the *intensity* of effort is $e(t)$. Let $E(t)$ and $K(t)$ be the total effort and total catch up to interval t, and $N(t)$ be the number of individuals in the population at time t. It is good idea to plot $\log(C(t))$ against $E(t)$ and/or $C(t)$ versus $K(t)$.
2. *The catch is removed from the fishery (or at the very least tagged and not recorded again if captured twice). WFFC rulings that ensure captive fish are returned to the water with minimal trauma implies that the assumption is unmet.*
3. The population is closed—the population must be closed to sources of animals such as recruitment and immigration and losses of animals due to natural mortality and emigration.

4. Catch reduction analysis VIII

4. Catchability is constant over the period of removals.
5. The units of effort are independent, i.e., the individual units of the method of capture (i.e., nets, traps, etc) do not compete with each other.
6. All fish are equally vulnerable to the method of capture—source of error may include gear saturation and trap-happy or trap-shy individuals.
7. Enough fish must be removed to substantially reduce the CPUE.
8. The catches may remove less than 2% of the population.

Also, the usual assumptions of simple regression such as

9. random sampling,
10. the independent variable(s) are measured without error—both catches and effort should be known, not estimated,
11. a line describes the data,
12. the errors are independent and normally distributed.

4. Catch reduction analysis IX

Loglinear analyses

Fit Poisson and negative binomial regressions at each competitor-session combination. Both models had the log-linear relationship

$$\log \mu_{adsc} \;=\; \eta \;=\; \beta_{(1)1} + \alpha_s + \beta_a + \gamma_d + \delta_c \;\; \text{where} \tag{8}$$

$\mu \equiv E(Y)$	is the mean number caught,
$\beta_{(1)1}$	is the intercept,
α_s	are the *sector effects* for $s = 1, \ldots, 5$ sectors,
δ_c	are the "*competitor effects*" for $c = 1, \ldots, 91$ competitors,
β_a	are the morning ($a = 1$) and afternoon ($a = 2$) effects,
γ_d	are the *day effects* for day $d = 1, 2, 3$.

Note: $\alpha_1 = \beta_1 = 0$ etc. Unused: $b = 1, \ldots, 19$.

4. Catch reduction analysis X

Table: Selected loglinear regression coefficients for the quasi-Poisson and negative binomial models. Standard errors are in parentheses and “Wald” denotes the Wald statistic. Eight competitors who did not fish all 5 sessions were excluded from the models. The Whanganui River, Mornings and Day 1 are the baseline levels of the factors.

	Poisson		Negative binomial	
Coefficient	Estimate (SE)	Wald	Estimate (SE)	Wald
Intercept	3.058 (0.213)	14.34	2.996 (0.208)	14.41
L. Otamangakau	−1.546 (0.104)	−14.89	−1.524 (0.083)	−18.35
L. Rotoaira	−1.814 (0.119)	−15.21	−1.800 (0.092)	−19.61
Waihou R.	−0.055 (0.063)	−0.87	−0.055 (0.063)	−0.87
Waimakariri R.	−0.046 (0.062)	−0.75	−0.030 (0.062)	−0.48
Afternoon	−0.172 (0.051)	−3.35	−0.154 (0.049)	−3.15
Day 2	−0.075 (0.062)	−1.21	−0.045 (0.059)	−0.76
Day 3	−0.293 (0.065)	−4.53	−0.295 (0.059)	−5.00

4. Catch reduction analysis XI

A small summary:

1. The two smaller rivers were not noticeably different from the Whanganui River, but the lakes were.
2. Afternoon fishing was less productive than the morning, and each successive day had poorer fishing than the previous day (although only the third day was statistically significantly different from the opening day).
3. The model is a poor one! Add an indicator variable to “lake versus river” and interact it with “session”. This is because `session` via (day, time-of-day) has a different meaning for lakes compared to rivers.

4. Catch reduction analysis XII

Bivariate logistic odds-ratio model (BLOM) analysis

$$\text{logit } P[Y_{ij} = 1 \big| \mathbf{x}_i] = \eta_j(\mathbf{x}_i), \qquad j = 1, 2, \tag{9}$$
$$\log \psi(\mathbf{x}_i) = \eta_3(\mathbf{x}_i), \tag{10}$$

where $\eta_j(\mathbf{x}_i) = \boldsymbol{\beta}_j^T \mathbf{x}_i$

The odds ratio (OR),

$$\psi = \frac{P(Y_1 = 0, Y_2 = 0)\, P(Y_1 = 1, Y_2 = 1)}{P(Y_1 = 0, Y_2 = 1)\, P(Y_1 = 1, Y_2 = 0)},$$

is a natural measure of the association between Y_1 and Y_2; a value of unity denotes statistical independence, and a value greater/less than unity means a positive/negative association.

4. Catch reduction analysis XIII

We fit a model to the two small rivers with

$$
\begin{aligned}
Y_1 &= 1 \text{ for the Waihou River,} \\
Y_1 &= 0 \text{ for the Waimakariri River, and} \\
Y_2 &= 0 \text{ for the morning,} \\
Y_2 &= 1 \text{ for the afternoon,} \\
x_2 &= \text{the fish length (cm).}
\end{aligned}
$$

The model can be used to investigate whether there are catch reduction differences between the two rivers with respect to morning versus afternoon fishing, as a function of fish length. Specifically, the model (9)–(10) is

$$
\begin{aligned}
\eta_j(\mathbf{x}_i) &= \beta_{(j)1} + f_{(j)2}(x_2), \qquad j = 1, 2, && (11) \\
\eta_3(\mathbf{x}_i) &= \beta_{(3)1}. && (12)
\end{aligned}
$$

4. Catch reduction analysis XIV

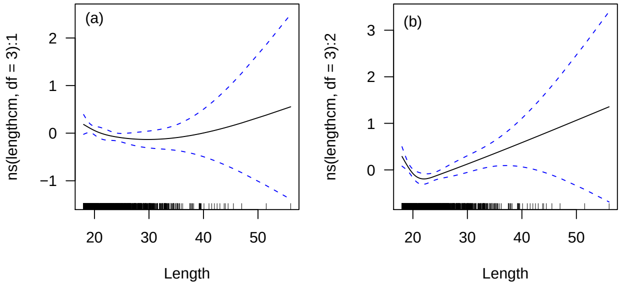


Figure: VGAM plots for the BLOM. Each centered component function is modelled by a regression spline with 3 degrees of freedom. The dashed lines are ± 2 standard error bands about the estimated curves. The plots are, up to a constant, (a) $\text{logit}\, P(\text{Fish caught in the Waihou River}|\text{length})$, (b) $\text{logit}\, P(\text{Fish caught in the afternoon}|\text{length})$, respectively (the alternatives are the Waimakariri River and the morning).

4. Catch reduction analysis XV

Some interpretation

- Figure (a) indicates some downward trend from 18 to 25 cm for $\widehat{f}_{(1)2}(x_2)$: it is easier to catch small (18 cm) fish in the Waihou River than mid-sized (25 cm) fish, relative to the Waimakariri River.
- Figure (b) suggests fish around 21 cm in length are easiest caught in the morning because $\widehat{f}_{(2)2}(x_2)$ attains its minimum there; shorter and longer ones are more easily caught in the afternoon.

4. Catch reduction analysis XVI

- $\log \widehat{\psi} = -0.287(0.092)$; this is strongly statistically significant, i.e., there is strong evidence against Y_1 and Y_2 being independent. With $\widehat{\psi} = 0.75$, the estimated odds of the event $(Y_1 = 1 | Y_2 = 1, x_2)$ is 0.75 times the estimated odds of $(Y_1 = 1 | Y_2 = 0, x_2)$, or more simply, $\widehat{P}[Y_1 = 1 | Y_2 = 1, x_2] < \widehat{P}[Y_1 = 1 | Y_2 = 0, x_2]$.

  This means, for a given length of fish, the probability of catching an afternoon fish in the Waihou is significantly less than catching a morning fish in the Waihou. Similarly, $\widehat{P}[Y_1 = 0 | Y_2 = 1, x_2] > \widehat{P}[Y_1 = 0 | Y_2 = 0, x_2]$, i.e., for a given length of fish, the probability of catching an afternoon fish in the Waimakariri is significantly greater than catching a morning fish in the Waimakariri.

5. Suggestions on the WFFC regulations I

Selective catching It is interesting to analyze, at both individual and team levels, the association between fish size and the total number of competition points awarded.

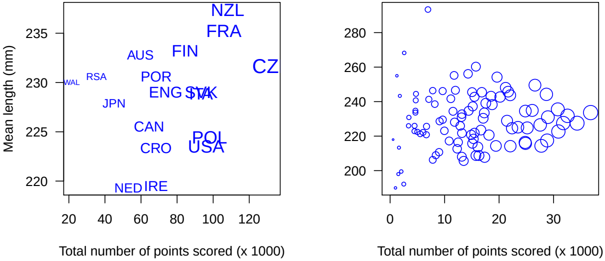


Figure: Mean length of fish versus total points scored. The size of each text/circle is approximately proportional to the number of fish caught. Data from the two small rivers only. (a) For each country. (b) For each individual.

5. Suggestions on the WFFC regulations II

Q: Given the competition scoring system, is there any evidence that certain teams selectively avoided catching larger fish for strategic purposes?

It was speculated that the professional teams targetted smaller sized fish because of their light tackle (which lowers line visibility and therefore increases the strike rate). Also, the number of competition points awarded for large fish can be heavily offset by the time it would take to bring it in and the decreased probability of a successful landing. It was therefore thought that some competitors purposely avoided catching the bigger fish; if so this strategy might be accentuated on the two smaller rivers where sight-fishing (catching specific fish seen by the angler, or stalking) is more likely.

A: We present the results here on the two small rivers only because the number of fish caught was large, and sight-fishing is not nearly as practical on the other three sectors.

5. Suggestions on the WFFC regulations III

To test this, the figure on Slide 42(a) plots, for the two small rivers only, the mean length of the fish and the points awarded. There appears to be little to no association between mean points and length per fish. A weighted linear least squares regression shows a two-sided *p*-value of 0.11, indicating weak evidence that, in fact, bigger fish are caught by the better teams.

But a repeat analysis (not given here) showed no evidence of any association between mean points and length per fish in the Whanganui River but there was statistical significance in one lake.

In summary, there appears to be no evidence here to suggest the bigger fish are being avoided. Or perhaps all the teams are avoiding the bigger fish equally.

5. Suggestions on the WFFC regulations IV

The present WFFC scoring system and a new proposal

Currently

$$P_1(y) \;=\; I(y \geq 0.18) \times \{100 + 20 \times \lceil 100\, y \rceil\} \qquad (13)$$

where

y	is fish length in metres,
$I(\cdot)$	is the *identity function*, and,
$\lceil x \rceil$	is the *ceiling* of x, e.g., $\lceil 3.1 \rceil = 4$.
	Thus each fish gave at least 460 points.

Approximate (13) by the continuous version (plotted in Slide 48).

$$P_1^{*}(y) \;=\; I(y \geq 0.18) \times \{100 + 2000\, y\} \qquad (14)$$

5. Suggestions on the WFFC regulations V

Competitors were ranked according to their *placings* at each sector-session combination. Then these placings were summed (*total placings*) over the sessions. Those with the minimum total placings were the winners. Thus it was not necessarily those who had the maximum points who won.

For example, in Session 1 at the Waihou River, each of the 19 competitors was ranked 1 (best) to 19 (worst) according to the point system. This is the "placing" for that session. These placings were added up over the 5 fishing sessions to give the "total placings".

Consider (13) and (14) more closely. Under this scheme, two fish of minimum legal length is equivalent to one fish of length 41 cm. However, a 41 cm trout is *much* harder to land than two barely legal ones, even taking into account of the time to hook-up both.

Therefore it behoves instigating regulation changes.

5. Suggestions on the WFFC regulations VI

Complexities of thinking and strategies:

A good strategy requires some thinking, e.g.,

- It was speculated that the professional teams targetted smaller sized fish because of their light tackle (which lowers line visibility and therefore increases the strike rate).
- The number of competition points awarded for large fish can be heavily offset by the time it would take to bring it in and the decreased probability of a successful landing.
- It was therefore thought that some competitors purposely avoided catching the bigger fish; if so this strategy might be accentuated on the two smaller rivers where *sight-fishing* is more likely.

We present the results here on the two small rivers only because the number of fish caught was large, and sight-fishing is not nearly as practical on the other three sectors.

5. Suggestions on the WFFC regulations VII

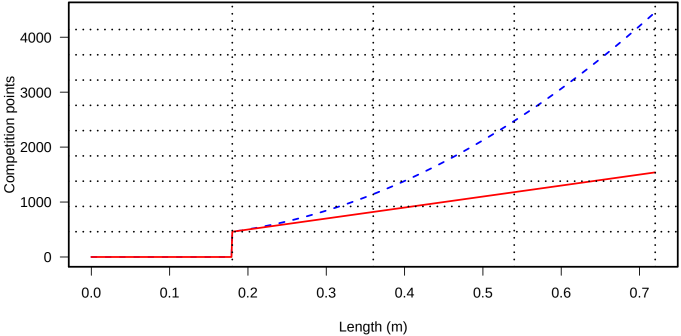


Figure: Proposed quadratic point system (top blue dashed curve is (16)). The red solid line is (14). The dotted grid represents integer multiples of the points given to a fish of length equal to integer multiples of the minimal length (0.18 m), under the existing rules, e.g., a minimal length fish is worth 460 points whereas one 3 times longer is worth 1180 points.

5. Suggestions on the WFFC regulations VIII

Let's look at the ratios:

```
> cbind(`Tiddler multiples` = 1:5, `Points ratio` = round(P1star(0.18 *
+     (1:5))/P1star(0.18), 1), `Proposed points ratio` = round(P2star(0.18
+     (1:5))/P2star(0.18), 1))

     Tiddler multiples Points ratio Proposed points ratio
[1,]                 1          1.0                   1.0
[2,]                 2          1.8                   2.5
[3,]                 3          2.6                   5.4
[4,]                 4          3.3                   9.7
[5,]                 5          4.1                  15.4
```

For example, landing a fish 4 times longer than the minimal size results in approximately 9.7 times the number of points given to one of minimum length. Under the current regulations this ratio is 3.3. The proposal attempts to compensate for the much lower probability of successfully landing a big fish relative to a tiddler.

5. Suggestions on the WFFC regulations IX

Rather than a linear relationship, a quadratic is suggested:

$$
\begin{aligned}
P_2(y) &= I(y \geq 0.18)\left\{100 + 20\lceil 100\,y\rceil + 1.0\left(\lceil 100\,y - 18\rceil\right)^2\right\} \quad (15)\\
\approx\ P_2^*(y) &= I(y \geq 0.18)\left\{100 + 2000\,y + 10,000\,(y - 0.18)^2\right\}. \quad (16)
\end{aligned}
$$

where the "1.0" and "10,000" can be replaced by some other comparable positive constant depending on the species (however, the number of points should ideally be integer-valued). Function (16) is plotted in Slide 48.

Adoption of (15)–(16) would add a slight level of complexity to the rules but competitors only need to know that landing a big fish would be awarded handsomely.

5. Suggestions on the WFFC regulations X

Q: does the new proposed scoring system make any change to the present rankings?

When applied to the 2008 WFFC data, the team rankings change as follows.

Ranking	Existing	Proposed	
1	CZE	CZE	team
2	NZL	FRA	team
3	FRA	NZL	team
1	CZE	FRA	individual
2	FRA	CZE (2=)	individual
3	CZE	ENG (2=)	individual

Overall the rankings changed a little but not markedly. A Spearman correlation coefficient of approximately 0.985 revealed high correlation but little change.

6. Yet to do . . .

There are lots and lots of things yet to do . . .

- Present the results from the random effects models.
- Fit the loglinear models (8) correctly.
- Add random effect to linear predictors, e.g.,

$$g_j(\theta_j) \;=\; \eta_j \;=\; \boldsymbol{\beta}_j^T \mathbf{x} + \boldsymbol{\gamma}_j^T \mathbf{z} \tag{17}$$

  where $\boldsymbol{\gamma}_j \sim N_g(\mathbf{0}, \boldsymbol{\Sigma})$, say.

  Obtain the class of *vector generalized linear mixed models (VGLMMs)* to add random effects to the VGLM class.

  Here, could treat the competitor and beat/boat effects $\alpha_{b(s)}$ and δ_c as random effects.

- Obtain the 2009 WFFC data to examine the competitors effects.

7. References

- DeLury, D. B., 1947. On the estimation of biological populations. Biometrics 3 (4), 145–167.
- Tolonen, A., Lappalainen, J., 1999. Origin of the large burbot (*Lota lota* (L.)) caught in an Arctic ice-fishing competition: a case study. Journal of Applied Ichthyology 15 (3), 122–126.
- Venman, M., 2006. Fish harvest results. Target Taupo: A Newsletter for Taupo Anglers (53), 4–11. URL `http://www.doc.govt.nz`
- Yee, T. W., 2009. Some issues raised by an analysis of the 2008 World Fly Fishing Championships data. In preparation .
- Yee, T. W., 2010. VGLMs and VGAMs: an overview for applications in fisheries research. Fish. Res. 101 (1–2), 116–126.

8. Closing Comments

- VGLMs and VGAMs are a very large class of models; VGLMs are model-driven while VGAMs are data-driven.
- The 2008 WFFC data are interesting,

8. Closing Comments

- VGLMs and VGAMs are a very large class of models; VGLMs are model-driven while VGAMs are data-driven.
- The 2008 WFFC data are interesting, *especially to flyfishermen.*

8. Closing Comments

- VGLMs and VGAMs are a very large class of models; VGLMs are model-driven while VGAMs are data-driven.
- The 2008 WFFC data are interesting, *especially to flyfishermen.*
- *I wish I had more time . . .*

8. Closing Comments

- VGLMs and VGAMs are a very large class of models; VGLMs are model-driven while VGAMs are data-driven.
- The 2008 WFFC data are interesting, *especially to flyfishermen.*
- *I wish I had more time . . .*

Thanks for your attention and tight lines!