

Д.К. Нгуен, В.Н. Малинин, С.М. Гордеева

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ВЫЛОВА ТУНЦА В ЮЖНО-КИТАЙСКОМ МОРЕ

D.K. Nguyen, V.N. Malinin, S.M. Gordeeva

STATISTICAL MODELING OF TUNA CATCH IN THE SOUTH CHINA SEA

Показано, что вылов тунца вьетнамскими судами в Южно-Китайском море в большей мере определяется экономическими факторами и в меньшей степени зависит от океанологических условий. Построена статистическая модель вылова тунца в зависимости от количества промысловых судов и аномалий температуры поверхности моря, описывающая 98 % дисперсии исходного ряда. Предложены модели экстраполяции для прогноза вылова тунца на 1–2 года.

Ключевые слова: Южно-Китайское море, вылов тунца, температура поверхности моря, статистическое моделирование, прогноз.

It is shown that economic factors are much more important than oceanographic conditions in tuna catch by Vietnamese fishing vessels in the South China Sea. The statistical model of the tuna catch was built depending on the number of fishing boats and the sea surface temperature anomalies, illustrating 98 % of the variance of the original series. The extrapolation models for 1–2 year prediction of the tuna catch for are proposed.

Key words: South China Sea, catch of tuna, sea surface temperature, statistical modeling, prediction.

Рыболовство в Южно-Китайском море имеет очень важное значение в экономическом развитии Вьетнама, особенно промысел океанического тунца, который занимает 1-е место в структуре экспорта морепродуктов из Вьетнама в более чем 60 стран мира [4]. Интенсивный промысел тунца во Вьетнаме начал развиваться с начала нынешнего столетия, поэтому статистические данные по вылову тунца имеются только с 2000 г., но даже они не являются вполне надежными, поскольку не существует хорошо отлаженной системы мониторинга промысла. Хотя оценки по вылову тунца в табл. 1 возможно несколько занижены из-за не учета в мониторинге данных многих мелких судов, однако уникальность их несомненна.

Как видно из табл. 1, суммарный годовой вылов тунца вьетнамскими судами за 15 лет увеличился более чем в 2 раза. Очевидными причинами этого является довольно быстрый прирост народонаселения, которое возросло примерно на 15 %, и рекордное увеличение валового внутреннего продукта, который за 15 лет вырос в 6 раз! Естественно, столь стремительный рост экономики Вьетнама позволил почти ежегодно наращивать рыбопромысловый флот и его мощности. Из табл. 1 видно, что число промысловых судов увеличилось более чем в 3 раза, а их суммарная мощность — более чем в 4 раза!

Таблица 1

**Статистические данные по вылову тунца вьетнамскими промысловыми судами
и некоторым экономическим факторам**

№ п/п	Год	Количество промысловых судов, шт.	Сумма мощности судов, 1000 л.с.	Вылов тунца, т	Валовой внутренний продукт, млрд USD	Количество населения, чел.
1	2000	9766	1385,1	33634	31	77630900
2	2001	14326	1613,3	35401	32	78620500
3	2002	15988	1947,5	37685	35	79537700
4	2003	17303	2192,9	37460	39	80467400
5	2004	20071	2641,8	50520	45	81436400
6	2005	20537	2801,1	52409	52	82392100
7	2006	21232	3046,9	52573	60	83311200
8	2007	21552	3051,7	54211	70	84218500
9	2008	22729	3342,1	49903	89	85118700
10	2009	24990	3721,7	44240	91	86025000
11	2010	25346	4498,7	41665	101	86932500
12	2011	23284	4132,7	49586	130	87840000
13	2012	30861	5477,5	65544	152	88775500
14	2013	31571	560,6	78633	168	90000000
15	2014	32833	5827,7	81778	184	90493300

Представляет интерес выявление взаимосвязи представленных в табл. 1 различных факторов и вылова тунца между собой. С этой целью были рассчитаны свободные от распределения непараметрические коэффициенты корреляции Спирмена (табл. 2). Из табл. 2 видно, что между всеми факторами наблюдается высокая корреляция, которая обусловлена главным образом наличием во временных рядах хорошо выраженных трендов. Вылов тунца имеет минимальную корреляцию с народонаселением ($r = 0,81$), а максимальную — с количеством промысловых судов ($r = 0,87$). Анализ табл. 2 позволяет сделать очевидный вывод, что вылов тунца в значительной степени определяется экономическими причинами, в то время как роль океанологических условий на вылов тунца является второстепенной. Итак, в общем случае вылов тунца (V) можно представить в виде аддитивной модели, состоящей из суммы двух компонент: $V_{\text{эконом}}$ и $V_{\text{океан}}$.

Таблица 2

**Оценки непараметрических коэффициентов корреляции Спирмена между различными факторами,
представленными в табл. 1**

	Количество промысловых судов	Суммарная мощность судов	Вылов тунца	ВВП
Суммарная мощность судов	0,977	1		
Вылов тунца	0,868	0,847	1	
ВВП	0,927	0,971	0,852	1
Количество населения	0,965	0,9798	0,812	0,913

Приведенные в табл. 2 оценки корреляции Спирмена свидетельствуют о возможности построения простой статистической модели зависимости вылова рыбы (V) от количества судов (N). С помощью метода наименьших квадратов было получено следующее уравнение:

$$V = V_{\text{эконом}} + V_{\text{океан}} = 7495 + 1,964N + \delta V, \quad (1)$$

где δV — остатки модели, которые определяются океанологическими условиями.

Данное уравнение и его коэффициенты значимы по критериям Фишера и Стьюдента, коэффициент детерминации $R^2 = 0,75$, а средняя квадратическая ошибка $\sigma = 7497$ т/год, что составляет 15 % от среднего значения вылова. Следующей задачей является идентификация компоненты δV .

Как было показано в работе [5], важнейшим океанологическим фактором, определяющим изменчивость различных характеристик биопродуктивности, является глубина изотермы $24\text{ }^\circ\text{C}$ (H_{24}). Но поскольку доступность и особенно оперативность ее получения создают серьезные затруднения в численных расчетах, особенно для многолетних периодов времени, то целесообразно использование температуры поверхности моря (ТПМ), которая определяется с высокой точностью, почти доступна в режиме он-лайн и имеет значимую корреляцию как с характеристиками биопродуктивности, так и с H_{24} [5].

На рис. 1 представлено распределение остатков регрессионной модели (1) вылова тунца за период 2000–2014 гг. Именно они послужили функцией отклика на межгодовые изменения ТПМ, которые брались из известного архива NOAA NCDC ERSSTv4 (<http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.ERSST/.version4/.SST>). Шаг сетки этого архива составляет 2° по широте и долготе, а продолжительность наблюдений — с 1854 г. по настоящее время. Для акватории ЮКМ, ограниченной $7\text{--}17^\circ$ с.ш. и $107\text{--}115^\circ$ в.д., где в основном и осуществляется вылов тунца вьетнамскими судами, сформированы временные ряды среднегодовых значений аномалий температуры ($\Delta\text{ТПМ}$) в 23 узлах двухградусной сетки за период 2000–2014 гг.

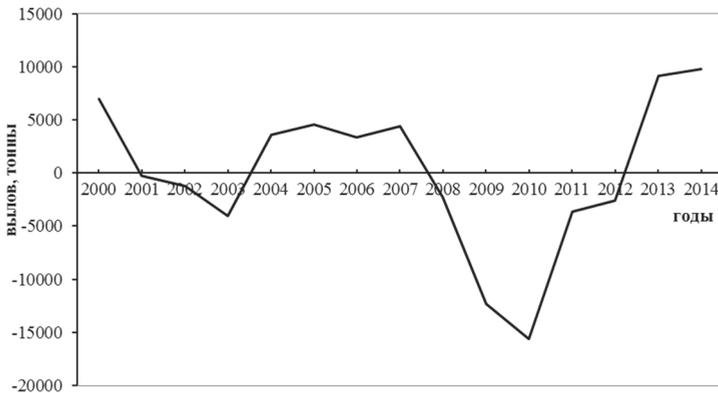


Рис. 1. Остатки регрессионной модели вылова тунца за период 2000–2014 гг.

В результате использования пошагового алгоритма модели множественной линейной регрессии [2] получено, что оптимальным может считаться уравнение с 8 переменными:

$$\delta V = b_0 + b_1 \Delta \text{ТПМ}_1 + \dots + b_8 \Delta \text{ТПМ}_8, \quad (2)$$

которое описывает 95 % дисперсии ряда δV , имеет малую среднеквадратическую ошибку вылова тунца ($\sigma = 2514$ т), значимый критерий Фишера для модели и значимые критерии Стьюдента для всех переменных.

Итак, в общем виде статистическая модель вылова тунца в ЮКМ имеет вид:

$$V = a_0 + a_1 N + \sum_{i=0}^8 b_i \Delta \text{ТПМ}_i. \quad (3)$$

Данная модель описывает 98 % дисперсии исходного временного ряда вылова тунца, а случайная ошибка вылова по модели составляет 1590 т/год, то есть 3 %. Из рис. 2 видно, что вычисленные по формуле (3) значения вылова рыбы почти в точности совпадают с их исходными значениями. Отметим также, что модель (3) носит диагностический характер и ее использование, например, в прогностических целях весьма проблематично.

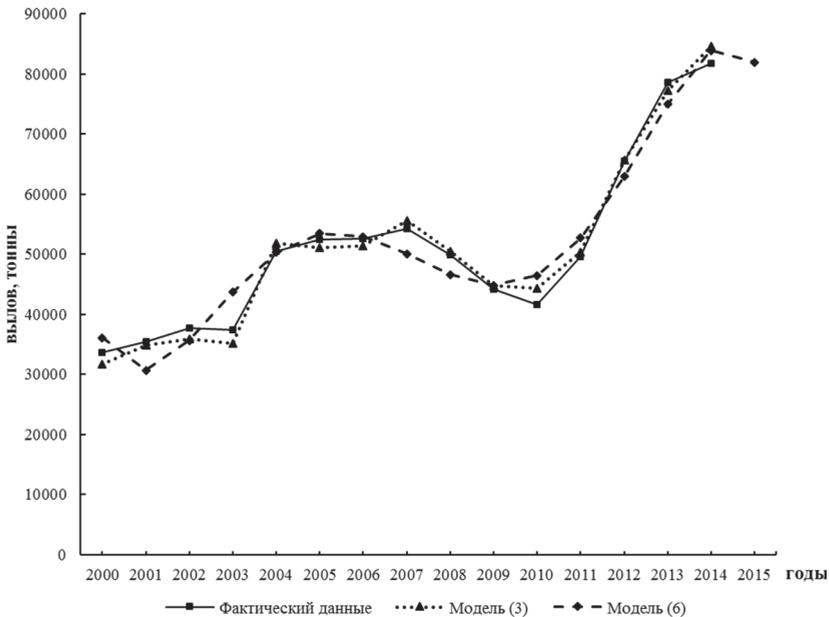


Рис. 2. Сопоставление фактических и вычисленных по моделям (3) и (6) значений вылова тунца в ЮКМ

С точки зрения планирования промысла, очень важной задачей является предварительный прогноз вылова рыбы. Однако сделать это чрезвычайно сложно, так как

для коротких временных рядов невозможно исследовать и выявить присущие им внутренние закономерности изменчивости и частотную структуру. Кроме того, для малых выборок использование формального статистического аппарата и, в частности, метода наименьших квадратов, имеет весьма ограниченное применение. Очевидно, в данном случае наиболее рациональным представляется использование экстраполяционного подхода, в соответствии с которым построение прогностических моделей возможно только для относительно «гладких» рядов, не содержащих экстремумы. По существу это эквивалентно задаче выделения и экстраполяции тренда временного ряда. Особенно полезным такой подход представляется, когда тренд описывает преобладающую долю дисперсии исходной выборки [3]. Для данной цели может быть использован программный комплекс «Аппроксимация», который позволяет оценивать параметры нескольких десятков нелинейных формул, а также авторегрессионные модели.

Рассчитанное по указанному программному комплексу уравнение линейного тренда имеет вид:

$$V = a_0 + a_1 t + \delta V^*, \quad (4)$$

где t — время; δV^* — остатки.

Уравнение (4) описывает 65 % дисперсии исходного ряда и имеет стандартную ошибку $\sigma = 8288$ т/год. Нетрудно видеть, что данное уравнение несколько уступает в точности уравнению (1). В принципе, даже формула (4) может быть использована для ориентировочного прогноза вылова тунца на 1–2 года. Для получения более точных прогностических оценок вылова необходимо построить статистическую модель для остатков δV^* в формуле (4). Использование комплекса «Аппроксимация» показало, что самым точным способом аппроксимации остатков δV^* является полином 5 степени, то есть

$$\delta V^* = b_0 + b_1 t + b_1 t^2 + \dots + b_5 t^5. \quad (5)$$

Коэффициент детерминации этой зависимости $R^2 = 0,84$, а среднеквадратическая ошибка $\sigma = 4236$ т/год. Суммируя теперь составляющие в формулах (4) и (5), получим вылов рыбы (рис. 2), функционально зависящий только от времени:

$$V = a_0 + b_0 + (a_1 + b_1)t + b_1 t^2 + \dots + b_5 t^5. \quad (6)$$

Данное уравнение описывает 95 % дисперсии исходного ряда и имеет стандартную ошибку $\sigma = 3250$ т/год. Если вычислить вылов тунца непосредственно по исходному ряду (табл. 1) полиномом 5 степени, то получим $R^2 = 0,95$ и $\sigma = 4222$ т/год. Нетрудно видеть, что коэффициенты детерминации полностью совпадают, в то время как величина σ в первом случае заметно ниже. Это связано с тем, что полиномы высоких степеней накапливают вычислительные ошибки. Но поскольку остатки по абсолютной величине значительно меньше исходных значений вылова рыбы, то естественно аппроксимация их полиномом 5 степени выполняется с меньшими ошибками.

Другой вариант построения экстраполяционной модели состоит в использовании применительно к остаткам в формуле (4) авторегрессионной модели вида:

$$X_o(t) = \alpha_1 X^o(t-1) + \alpha_2 X^o(t-2) + \dots + \alpha_p X^o(t-p) + Z(t), \quad (7)$$

где $X^o(t)$ — центрированный случайный процесс; $Z(t)$ — белый шум.

Главной задачей авторегрессии является оценка порядка p . Достаточно просто это можно сделать с помощью частной автокорреляционной функции (рис. 3).

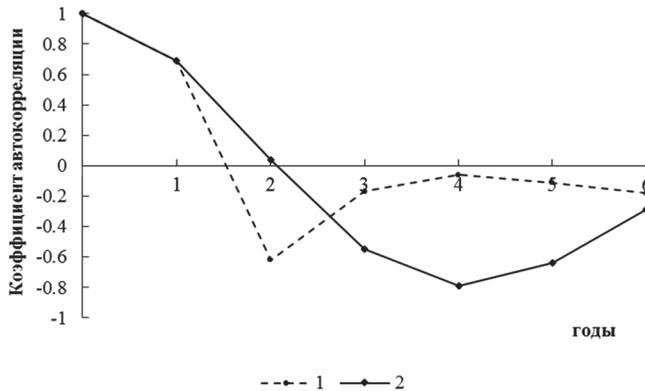


Рис. 3. График частной (1) и общей (2) автокорреляционной функции значений δV^*

Как видно из рис. 3, на первых двух сдвигах отмечается высокая корреляция, превышающая критическое значение коэффициента корреляции при уровне значимости $\alpha = 0,05$ ($r_{кр} = 0,50$). Однако на последующих сдвигах $\tau \geq 3$ она резко уменьшается и становится незначимой. Отсюда следует, что в принципе мы можем использовать модели авторегрессии первого и второго порядков, то есть $p = 1$ и $p = 2$. Естественно, модель второго порядка предпочтительнее, так как имеет более низкую среднеквадратическую ошибку. Отметим, что коэффициенты авторегрессии рассчитываются путем решения системы нормальных уравнений Юла-Уокера [1]. После этого по формуле (7) несложно рассчитать значения δV^* , складывая которые с трендовой компонентой получаем оценки вылова тунца.

Формула (6) легко поддается экстраполяции. Но выполнять прогноз более чем на 1 шаг рискованно. Прогноз на 2 шага возможен только на качественном уровне, то есть дается оценка увеличения (уменьшения) вылова по сравнению с предшествующим шагом. Прогноз по линейному тренду (4) на 1 шаг (на 2015 г.) дает оценку вылова рыбы, равную 72 100 т/год, а по уравнению (6) — 82 000 т/год, что почти совпадает с фактическим выловом рыбы в 2014 г. Прогноз с использованием модели авторегрессии первого и второго порядков дает оценку вылова тунца на 2015 г., соответственно равную 81 500 и 77 200 т/год. Отсюда видно, что модель авторегрессии второго порядка заметно занижает оценку вылова рыбы. Однако только сравнение с фактическим

выловом позволит определить, какие прогнозные оценки являются более точными. К сожалению, на момент сдачи статьи в редакцию журнала данные о вылове тунца в 2015 г. отсутствовали.

Что касается прогноза вылова тунца на 2016 г., то если ориентироваться на модели экстраполяции, он может снизиться. В заключение подчеркнем, что учитывая особенности исходных данных, их короткую продолжительность, полученные прогностические оценки вылова тунца вьетнамскими судами в ЮКМ носят приближенный характер. Однако получение более точных прогностических оценок вылова рыбы в настоящее время вряд ли представляется возможным.

Оценка вылова тунца за 2015 г. была получена во время редактирования статьи. Она составила 78 670 т, то есть практически совпадает с прогностической оценкой по модели авторегрессии второго порядка.

Литература

1. *Бокс Дж., Дженкинс Г.* Анализ временных рядов. Прогноз и управление. Вып. 1. — М.: Мир, 1974. — 406 с.
2. *Малинин В.Н.* Статистические методы анализа гидрометеорологической информации. — СПб.: РГГМУ, 2008. — 407 с.
3. *Малинина Ю.В.* Особенности морской деятельности в современных условиях и оценка негативных последствий от повышения морского уровня в XXI столетии: Автореф. дисс. ... канд. геогр. наук. — СПб.: РГГМУ, 2010. — 23 с.
4. *Нгуен Д.К., Бухарицин П.И.* Исследование влияния некоторых экологических факторов на распределение пелагических рыб в Южно-Китайском море (на примере тунца) // Вестник АГТУ. Сер.: Рыбное хозяйство, 2014, № 2, с. 13–20.
5. *Нгуен Д.К., Малинин В.Н., Гордеева С.М.* Влияние температуры воды на формирование биологической и промысловой продуктивности Южно-Китайского моря // Учёные записки РГГМУ, 2016, № 42, с. 64–73.