

события, способствуя или препятствуя их осуществлению, так или иначе модифицируя их». Поэтому наилучшим является сочетание экстраполяционного прогноза производительности и программно-целевого моделирования критерия соответствия целевых, плановых и директивных показателей с фактически достигнутыми результатами.

Обсуждение проблемы критериев завершения следующей рекомендацией: «Когда проектируется и создается самолет, то его можно пустить в эксплуатацию тогда и только тогда, когда он будет полностью закончен. Иное дело - система моделей. Её можно и нужно начинать использовать уже тогда, когда закончены лишь некоторые составные части» [10, с.22]. Методология идентификации биотехнического закона это вполне позволяет.

Статья опубликована при поддержке гранта 3.2.3/4603 МОН РФ

СПИСОК ЛИТЕРАТУРЫ:

1. Мазуркин, П.М. Статистическое моделирование. Эвристико - математический подход / П.М. Мазуркин. - Научное издание. - Йошкар-Ола: МарГТУ, 2001. - 100с.
2. Скурихин В.И., Шифрин В.Б., Дубровский В.В. Математическое моделирование / В.И. Скурихин, В.Б. Шифрин, В.В. Дубровский. - Киев: Техника. 1983. -270 с.
3. Поляк, Б.Т. Введение в оптимизацию / Б.Т. Поляк. - М.: Наука, 1983. - 384 с.
4. Пасхавер И.С., Яблочник А.Л. Общая теория статистики. Для программированного обучения / И.С. Пасхавер, А.Л. Яблочник. - М.: Финансы и статистика, 1983. - 432 с.
5. Современный синтез критериев в задачах принятия решений. - М.: Радио и связь, 1992. - 120 с.
6. Гойзман, Э.И. Моделирование производственных процессов на шахтах / Э.И. Гойзман. - М.: Недра, 1977. - 192 с.

7. Олдак, П.Г. Равновесное природопользование и экономический рост / П.Г. Олдак. - ЭКО. - 1984. - № 8. - С.44-61

8. Федулов А.А. Проблемы создания экономико-математических моделей в АСУНТ / А.А. Федулов, В.Н. Растегаев. - М.: Информэлектро, 1980. - 78 с.

9. Елисеева И.И. Логика прикладного статистического анализа / И.И. Елисеева, В.О. Рукавишников. - М.: Финансы и статистика, 1982. - 192 с.

10. Проектирование самолетов: Учебник. - 3-е изд., перераб. и дополн. - М.: Машиностроение, 1983. - 616 с.

БИОТЕХНИЧЕСКИЙ ЗАКОН И АДЕКВАТНОСТЬ ГОТОВОЙ МОДЕЛИ

Мазуркин П.М.

*Марийский государственный технический университет
Йошкар-Ола, Россия*

Пусть несколькими сеансами идентификации биотехнического закона создана некая готовая модель. В статье приведем несколько подходов к выявлению уровня её адекватности к исходным данным по статистической выборке. Пусть вычислена при этом средняя арифметическая остатков $\bar{\varepsilon}$. Тогда чем ближе значение $\bar{\varepsilon}$ к нулю, тем точнее приближена модель к фактическим данным. Далее расчетные уравнения приведены применительно к процедуре «Расчет статистических показателей остатков» алгоритма идентификации [1].

По остаткам размах вариации (изменчивости) R по совокупности остатков равен

$$R = \varepsilon_{\max} - \varepsilon_{\min} \quad (1)$$

Среднее линейное отклонение \bar{d} составляет
$$\bar{d} = \frac{1}{n} \sum_{i=1}^n |\varepsilon_i - \bar{\varepsilon}| \quad (2)$$

Дисперсия σ^2 остатков с учетом малой выборки, когда $[m] \times [m'] \leq 20$, будет равной

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2 \quad (3)$$

При $[m] \times [m'] > 20$ впереди правой части уравнения (3) принимается отношение $1/n$. Среднее квадратичное отклонение остатков равно

$$\sigma = \sqrt{\sigma^2} \quad (4)$$

Коэффициент изменчивости (вариации) V_{ε} , дающий относительную оценку вариабельности остатков, определяется из выражения

$$V_{\varepsilon} = 100\sigma / \bar{\varepsilon} \quad (5)$$

Нормированное отклонение t_{ε} наблюдения определяется по формуле

$$t_{\varepsilon} = (\varepsilon - \bar{\varepsilon}) / \sigma \quad (6)$$

Принимается значение доверительного интервала - t - сигмальная единица. Например, пусть класс надежности расчетов принят B , то есть расчеты необходимы с малым риском. Поэтому $[t] = 1,96$. Для решения вопроса об исключении наблюдения и повторном поиске значений параметров модели по рекомендациям [2, с.260] необходимо выяснить сущность отклонения (так называемые «особые» точки, которые нами рассмотрены отдельно). Формально, если выполнено соотношение

$$t_i \leq [t], \quad (7)$$

то такое наблюдение оставляется в статистической совокупности. При невыполнении условия (7) наблюдение исключается и поиск проводится по новому сеансу на ПЭВМ.

Асимметрия (скошенность) A_s распределения остатков вычисляется по формуле

$$A_s = \frac{1}{\sigma^3} \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^3 \quad (8)$$

Если $A_s > 0$, то наблюдается правосторонняя асимметрия (вершина распределения расположена слева среднего значения), при условии $A_s < 0$ - левосторонняя асимметрия. Чаще всего асимметрия выражает ненормальность распределения или наличие нескольких, сдвинутых вместе, распределений.

Экссесс (островершинность) E будет равен

$$E = \frac{1}{\sigma^4} \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^4 - 3 \quad (9)$$

При островершинности $E > 0$, то эксцесс положительный. Если $E < 0$, то распределение остатков плосковершинное. По значениям t и σ определяются параметры функции кривой распределения. Уравнение нормированной (стандартной) плотности нормального распределения будет иметь вид [3, с.41]

$$\varphi(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \quad (10)$$

Другие виды распределения, как χ^2 (хи - квадрат), Стьюдента (t - распределение), биномиальное, логарифмическое, логнормальное и другие в технических исследованиях встречаются редко. В исследованиях могут встречаться известные устойчивые законы распределения, а также их комбинации.

Предельная ошибка выборки остатков вычисляется по формуле

$$\Delta_{\varepsilon} = t_{\max} \sqrt{\sigma^2 / n}, \quad (11)$$

из которой вычисляется допустимое значение n по формуле

$$n \geq [t]^2 \sigma^2 / [\Delta_{\varepsilon}]^2 = [n], \quad (12)$$

где допустимое значение t принимается по известным данным. Полученное значение n сравнивается со значением расчетного числа наблюдений.

Такие сложные расчеты необходимы при аппроксимации. В них нет необходимости при процедурах идентификации. Потому при поиске параметров модели можно воспользоваться еще более простыми формулами:

$$\text{- ошибка среднеквадратичного отклонения } m_{\sigma} = \sigma / \sqrt{\bar{\varepsilon}}; \quad (13)$$

$$\text{- показатель точности наблюдений } P_{\varepsilon} = V_{\varepsilon} / \sqrt{n}; \quad (14)$$

$$\text{- доверительные интервалы } \varepsilon_{\cdot} = \bar{\varepsilon} - [t]\sigma; \quad \varepsilon_{\cdot p} = \bar{\varepsilon} + [t]\sigma. \quad (15)$$

Адекватность модели. Для вычислений критериев адекватности необходимо определить:

$$\sigma_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \quad (16)$$

- дисперсию теоретических значений показателя

где

$$\bar{\hat{y}} = \frac{1}{n} \sum_{i=1}^n \hat{y}_i \quad (17)$$

- дисперсию фактических значений показателя

$$\sigma_{\hat{y}}^2 = \frac{1}{n-1} \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 \quad (18)$$

- дисперсию остатков

$$\sigma_{\varepsilon}^2 = \frac{1}{n-1} \sum_{i=1}^n \varepsilon_i^2 \quad (19)$$

Значение дисперсии остатков можно также вычислять по правилу сложения дисперсией, поэтому

$$\sigma_y^2 = \sigma_{\hat{y}}^2 + \sigma_{\varepsilon}^2 \quad (20)$$

Для измерения тесноты связи принимается коэффициент детерминации B_{yx} , равный [3, с.102]

$$B_{yx} = \sigma_{\hat{y}}^2 / \sigma_y^2 \quad (21)$$

Чем ближе $B_{yx} \rightarrow 1$, то тем больше совпадают теоретические и фактические выходные резуль-

таты. Мера неопределенности регрессии, то есть влияние «необъясненной» дисперсии σ_{ε}^2 , определится из выражения

$$U_{yx} = 1 - B_{yx} = \sigma_{\varepsilon}^2 / \sigma_y^2 \quad (22)$$

Исправленный коэффициент детерминации B_{yx}^* определится из формулы [3, с.108]

$$B_{yx}^* = 1 - (1 - B_{yx}) \frac{n-1}{n-N-1} \quad (23)$$

С вероятностью $1-\alpha$, где α - уровень значимости по табл. 2, коэффициент детерминации проверяется по критерию Фишера [3, с.202]

$$F_{p \in} = \frac{B_{yx} (n - N - 1)}{N(1 - B_{yx})} > F_{\hat{p}}(\alpha, n, n - N - 1) \quad (24)$$

Существенность дисперсионной связи вычисляется по критерию Фишера [25, с.33] по формуле

$$F'_{p \in} = \sigma_{\hat{y}}^2 / \sigma_{\varepsilon}^2 > F_{\hat{p}}(\alpha, n, n - N - 1) \quad (25)$$

где $F_{\hat{p}}$ - табличное значение критерия Фишера, n - число наблюдений, N - число параметров модели.

Регрессионная модель считается работоспособной, если $F'_{p \in} > 3$ [4, с.34]. По критерию Фишера

$F'_{p \in}$ сравниваются несколько альтернативных регрессионных моделей, если таковые возникнут при обсуждении проблемы построения конструкций моделей (способами аппроксимации). Например, лучше обычная или нормализованная формы модели, предпочтительнее арифметические или алгебраические оперативные константы.

Автоматический поиск параметров модели. Для оценки сходимости и адекватности модели автоматического поиска параметров модели целесообразно вычислять статистические показатели, соотношенные к выходным результатам объекта исследования:

- сумма квадратов отклонений

$$\sum \varepsilon^2 = \sum_{i=1}^n \varepsilon_i^2 \quad (26)$$

- среднеквадратичное отклонение

$$\sigma_{\varepsilon} = \sqrt{\sigma_{\varepsilon}^2}; \quad (27)$$

- ошибка среднеквадратичного отклонения

$$m_{\sigma} = \sigma_{\varepsilon} / \sqrt{\hat{y}}; \quad (28)$$

- коэффициент изменчивости

$$\hat{V} = 100\sigma_{\varepsilon} / \hat{y}; \quad (29)$$

- показатель точности

$$\hat{P} = \hat{V} / \sqrt{n}; \quad (30)$$

- относительное отклонение выходных результатов

$$100(\hat{y}_i - y_i) / \hat{y}_i; \quad (31)$$

- среднее относительное отклонение выходных результатов

$$\hat{\Delta} = \frac{1}{n} \sum_{i=1}^n |\Delta_i| \quad (32)$$

Максимальное отклонение Δ_{\max} выбирается из значений (31) по абсолютной величине из множества $|\Delta_i|$.

Сравнение моделей по критериям адекватности. Все существующие методы анализа адекватности моделей статистическим выборкам предполагают однородность экспериментов, включенных в одно множество. При этом классическая статистика четко предполагает только наличие закона нормального распределения.

Однако полезно сравнить предложенный **метод оценки адекватности по максимальной относительной погрешности** с известными методами, в частности, с методом Фишера. Он наиболее распространен в сельскохозяйственной и инженерно-экологической науке.

Средняя относительная погрешность по остаткам от модели. Поэтому наиболее простому

$$\bar{\Delta} = 100 \frac{1}{n} \sum_{i=1}^n |\varepsilon_i| / y_{\phi i} \quad (33)$$

Однако любая средняя арифметическая величина относится к нормальному закону распределения, поэтому не применима к неоднородным статистическим выборкам, каждый член которых имеет собственную индивидуальность. Поэтому по формуле (33) можно получить одно и то же значение относительной погрешности для нескольких выборок, имеющих различные друг от друга законы распределения. При этом размах $|\varepsilon_{\max}| - |\varepsilon_{\min}|$ может быть сильно различающимся. Поэтому лучше всего определять относительную погрешность по формуле (33), а затем выделить $|\Delta_{\max}|$. Тогда получается, что требование

известному способу суммируются все отношения остатков ε_i по модулю (без учета знаков, то есть $|\varepsilon_i|$) к фактическим значениям изучаемого показателя $y_{\phi i}$. В результате исключается проблема сравнимости моделей и одновременно «проклятие» размерности показателя. Затем вычисляется результат как сумма \sum и она делится на число n членов (наблюдений) статистической выборки.

Тогда получается сопоставимый для любых статистических моделей, полученных аппроксимацией произвольных формул или идентификацией устойчивых законов и закономерностей критерий адекватности - **средняя относительная погрешность** $\bar{\Delta}$ (%)

адекватности по максимальной относительной погрешности в несколько раз (равное отношению $|\Delta_{\max}| / \bar{\Delta}$) строже по сравнению с известным методом.

Средняя квадратичная погрешность. Чтобы исключить влияние знаков, было придумано возвести в квадрат остатки, то есть получить выборку в виде ряда ε^2 .

Но вначале придется определить среднее арифметическое от всех значений изучаемого показателя, то есть применить формулу для нормального распределения случайных событий

$$\bar{y}_\phi = \frac{1}{n} \sum_{i=1}^n y_{\phi i} \quad (34)$$

Среднеарифметического объекта в природе физически не существует. Поэтому эта физическая величина весьма условная, придуманная математиками для сравнения статистических выборок друг с другом. Однако много существует однородных выборок с членами, равными по сущности неотличительных признаков, у которых с различным размахом $y_{\max} - y_{\min}$ может оказаться одинаковое значение среднего арифметического.

$$\bar{\Delta}_{кв} = \frac{100}{\bar{y}_\phi} \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{\phi i} - \bar{y}_\phi)^2} \quad (35)$$

F-критерий Фишера. Этот показатель адекватности относится к дисперсионному анализу. Он оказывает статистическую значимость уравнения регрессии (в нашем случае статистически устойчивой закономерности, построенной на основе устойчивых законов) в целом. Анализ адекватности модели выполняется сравнением фак-

Таким образом, уже изначально ясно, что метод среднеквадратичного отклонения (погрешности) не применим для выявления адекватности идентификации закономерностей. Однако до сих пор применяют для выявления адекватности аппроксимированных уравнений, не имеющих содержательного смысла по параметрам и отдельным составляющим, формулу (при этом к методу наименьших квадратов не относится)

тического F_ϕ и табличного (критического) F значений критерия Фишера.

При этом фактическое значение критерия вычисляется, с использованием дисперсий на одну степень свободы, по формуле

$$F_\phi = \frac{n-m-1}{m} \times \frac{S_\phi}{S} \quad (36)$$

где n - число членов (единиц) статистической совокупности, шт., m - количество объясняющих факторов в уравнении (модели), шт. S_ϕ - факторная дисперсия (по размерности изучаемого показателя), S - остаточная от модели дисперсия (по размерности показателя).

В формуле (36) обращает внимание тот факт, что по выражению в числителе $n-m-1$ количество объясняющих переменных должно соответствовать условию $m \leq n-1$.

Многие примеры идентификации показали, что это условие по Фишеру при использовании устойчивых законов необязательное, поэтому количество одновременно учитываемых факторов может быть и больше числа наблюдений (напри-

мер. 14 факторов при 10 сложных по количеству влияющих факторов наблюдениях).

При этом было замечено, что конструктивная связь между факторами требует наименьшего количества параметров модели. С переходом на мультиколлениарные связи потребуется большее количество параметров модели, а при аддитивной связи будет наблюдаться наибольшее количество параметров у искомой статистической закономерности.

Факторная дисперсия, то есть сумма квадратов отклонений фактических значений показателя от их среднеарифметического по формуле (7.25), вычисляется по уравнению

$$S_\phi = \sum_{i=1}^n (y - \bar{y}_\phi)^2 \quad (37)$$

где y - расчетное по модели значение изучаемого показателя, \bar{y}_ϕ - среднее арифметическое у фактических значений показателя. Остаточная дисперсия, то есть сумма квадратов отклонений по остаткам \mathcal{E}_i , определяется по выражению

$$S = \sum_{i=1}^n (y_\phi - y) \quad (38)$$

Общая вариация (дисперсия) результата определяется по формуле

$$S_{общ} = S_\phi + S \quad (39)$$

Примеры сравнения. Для сопоставления различных критериев адекватности используем табличные данные по трем гидрометрическим створам и пяти показателям свойств у 18 проб травы [5]. Табличное критическое значение критерия Фишера равно 3,68 для всех пяти двухфакторных моделей при степенях свободы для расчета фактического критерия Фишера $k_1 = n - m - 1 = 18 - 2 - 1 = 15$ и при этом для двух учтенных факторов $k_2 = m = 2$.

Удельная масса сырой травы. В табл. 1 приведены результаты расчетов различных критериев адекватности одной и той же модели. Предлагаемый нами критерий адекватности приведен в первом столбце. Он не требует никаких дополнительных расчетов и прост в применении тем, что нужно просмотреть относительную погрешность по всем строкам и выбрать максимальную из значений погрешности, обозначив как Δ_{\max} (%).

Поверхность отклика на рис. 1 сложна и она в среднем характеризуется статистической моделью вида

$$m = 29,917H^{-3,96422} \exp(+2,94285H^{1,01429} - 6,55787 \cdot 10^{-7} L^{3,00957}) \quad (39)$$

На площадке № 7 показатель по формуле (39) получил наибольшую погрешность $\Delta_{\max} = 54,24\%$. Относительную погрешность $\Delta = 38,21\%$ получила травяная проба на пробной площадке № 15. Средняя относительная погреш-

ность $\bar{\Delta}$ (%) требует небольших вычислений. Вначале определяется сумма погрешностей по модулю (второй столбец табл. 1), а затем вычисляется средняя погрешность.

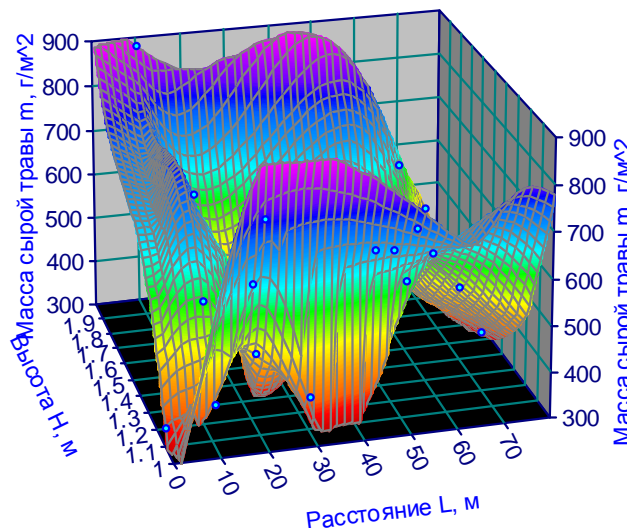


Рис. 1. Поверхность отклика массы срезанной травы в среднем на трех створах реки Ировка

Таким образом, формула (39) получила по предложенному нами критерию адекватности $\Delta_{\max} = 54,24\%$ для седьмой пробной площадки и при этом второй статистический показатель $\bar{\Delta} = 21,14\%$. Средняя квадратичная погрешность равна $\bar{\Delta}_{кв} = 23,80\%$. Тогда получается, что максимальная относительная погрешность больше

средних величин соответственно в $54,24 / 21,14 \approx 2,6$ и $54,24 / 23,80 \approx 2,3$ раза.

Распределение погрешностей только частично подчиняется нормальному закону Гаусса-Лапласа. Однако биотехнический закон адекватнее. Например, нормальное и ранговое распределения значений из второго столбца табл. 1 определяются формулами:

- закон нормального распределения погрешностей по модулю

$$|\Delta| = 39,1996 \exp(-0,0092209(r - 0)^2) \quad (40)$$

- распределение по биотехническому закону проф. П.М. Мазуркина

$$|\Delta| = 54,2247 \exp(-0,29032r^{0,26634}) - 1,53538r \quad (41)$$

Сравнение показывает, что нормальное распределение по закону Гаусса-Лапласа при условии $r=0$ (здесь этот закон приводится напрямую по параметрам) относительно оси ординат

биотехнический закон (закон гибели как его частный случай) имеет более высокое значение коэффициента корреляции $0,9962 > 0,9425$ (рис. 2).

Таблица 1.

Критерии адекватности закономерности (39) массы проб сырой травы

№ п/п	Δ_{\max} , %	$\bar{\Delta}$, %	Среднее квадратичное		F -критерий Фишера	
			\bar{y}_{ϕ}	$\bar{\Delta}_{\text{кв}}$, %	$F_{\phi} = 3,20 < 3,68$	S
1	16,91	16,91	880	22141,4	48796,8	22141,4
2	8,76	8,76	580	2580,6	14520,3	2580,6
3	30,26	30,26	420	16154,4	1354,2	16154,4
4	1,45	1,45	475	47,6	806,6	47,6
5	13,92	13,92	600	6972,3	38,4	6972,3
6	21,70	21,70	440	9120,3	635,0	9120,3
7	54,24	54,24	340	34003,4	198,8	34003,4
8	37,08	37,08	390	20909,2	590,5	20909,2
9	29,60	29,60	780	53314,8	1505,4	53314,8
10	29,05	29,05	630	33489,0	4006,9	33489,0
11	19,04	19,04	490	8704,9	12905,0	8704,9
12	6,52	6,52	315	506,3	20391,8	2756,3
13	11,42	11,42	565	4160,3	96,0	4160,3
14	10,91	10,91	460	2520,0	0,0	2520,0
15	38,21	38,21	380	21083,0	222,0	21083,0
16	25,28	25,28	595	22620,2	4316,5	22620,2
17	8,00	8,00	415	1102,2	3856,4	1102,2
18	18,26	18,26	430	6162,3	3,2	6162,3
Сумма		380,61	9185	265592,0	114244,0	267842,1
Критерий		21,15	510,3	23,80		

Вроде бы асимметрия небольшая, но она имеет колоссальное креативное значение. Поэтому формула (39) может быть применена, то есть, не отброшена как статистически недостоверное, только из-за того, что её конструкция идентична

конструкции моделей других показателей у свойств луговой травы. Однако уравнение с конкретными параметрами модели (39) не может быть использованы в различных технологических или иных расчетах.

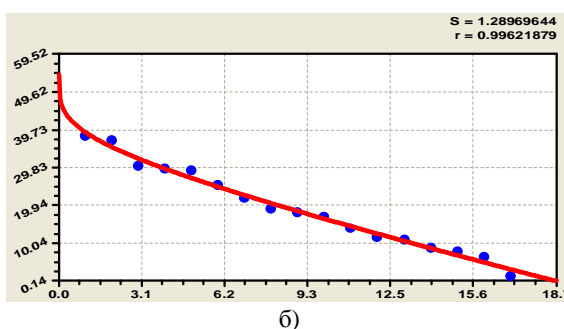
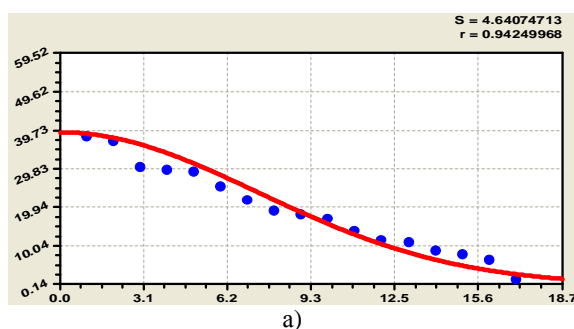


Рис. 2. Графики нормального (а) и биотехнического (б) распределения ошибок

Результаты сравнительных расчетов приведены в табл. 2.

Полужирным шрифтом выделены наилучшие значения того или иного критериев адекватности. Заметим здесь, что эмпирическим путем в эконометрике достигнуто требование к средней погрешности $\bar{\Delta}$ эконометрической модели в 8,0%.

Этот уровень вполне совпадает с погрешностью 8,81% табл. 2 для относительной влажности.

Сравнение показало, что наипростейшим критерием адекватности и более точным по сути идентифицируемым природных и природно-антропогенных явлений и процессов является статистический показатель Δ_{\max} - **максимальная относительная погрешность** (в процентах).

В заключение приведем слова Н. Винера [6, с.76]: «... современный аппарат малых выборок, как только он выходит за рамки простого подсчета своих собственных, специально определенных параметров и превращается в метод положительных статистических выводов для новых

случаев, уже не внушает мне доверия, Исключение составляет случай, когда этот аппарат применяется статистиком, который явно знает или хотя бы неявно чувствует основные элементы динамики исследуемой ситуации».

Таблица 2.

Критерии адекватности по моделям изменения показателей луговой травы

Параметр травы	Δ_{\max} , %	$\bar{\Delta}$, %	\bar{y}_ϕ	$\bar{\Delta}_{\text{кв}}$, %	F_ϕ	R^2	R
m , г/м ²	54.24	21.15	510.3	23.80	3.20	0.299	0.547
m_c , г/м ²	51.31	19.22	138.6	20.62	3.82	0.461	0.679
m_{60} , г/м ²	59.87	22.95	364.2	26.26	3.12	0.294	0.542
W , %	26.21	8.81	250.1	10.32	4.59	0.379	0.616
\bar{v} , г/(м ² ч)	46.66	18.77	1.312	21.44	1.87	0.200	0.447

Статья опубликована при поддержке гранта 3.2.3/4603 МОН РФ

СПИСОК ЛИТЕРАТУРЫ:

1. Мазуркин, П.М. Статистическое моделирование. Эвристико-математический подход / П.М. Мазуркин. - Научное издание. - Йошкар-Ола: МарГТУ, 2001. - 100с.
2. Пасхавер, И.С. Общая теория статистики. Для программированного обучения / И.С. Пасхавер, А.Л. Яблочник. М.: Финансы и статистика, 1983. - 432 с.
3. Фёрстер, Э. Методы корреляционного и регрессионного анализа: Руководство для экономистов / Э. Фёрстер, Б. Рёнц. - М.: Экономика и статистика, 1983. - 302 с.
4. Елисеева, И.И. Логика прикладного статистического анализа / И.И. Елисеева, В.О. Рукавишников. - М.: Финансы и статистика, 1982. - 192 с.
5. Мазуркин, П.М. Измерение продуктивности травяного покрова пойменного луга / П.М. Мазуркин, С.И. Михайлова // Современные наукоемкие технологии: материалы заочной электронной конференции. - 2008. - № 7. - С.91-92.
6. Винер, Н. Кибернетика или управление и связь в животном и машине / Н. Винер. - 2-е изд. - М.: Наука, 1983. - 344 с.

БИОТЕХНИЧЕСКИЙ ЗАКОН И ПОДГОТОВКА ИСХОДНЫХ ДАННЫХ

Мазуркин П.М.

Марийский государственный технический университет
Йошкар-Ола, Россия

Чаще всего статистическое моделирование выполняется по прошлой количественной информации (качественные значения преобразуются в коды, ранги, отношения), которая зафиксирована в виде текстового описания (эвристическая модель) и таблицы исходных для моделирования

данных (табличной модели), а также графиков (чаще всего при однофакторных статистических выборках).

Модель может идентифицироваться и по текущей информации в реальном режиме времени, но при этом процесс идентификации не должен превышать одной трети промежутка времени между получением каждой порции сведений. Однако и здесь на первом сеансе значения параметров искомой модели необходимо вычислить по прошлым данным, то есть необходим анализ некоторой предыстории явления или процесса. Последующие сеансы параметрической идентификации выполняются гораздо быстрее из-за использования готовых после первого сеанса моделей.

На рис. 1 приведены условные примеры аппроксимации (рис. 1а) и параметрической идентификации (рис. 1б).

В первом случае логарифмированием получаем вместо показательного закона $y = a_1 x^{a_2}$ линейную модель $\ln y = \ln a_1 + a_2 \ln x$. Во втором случае точная линеаризация невозможна. Исследователи пытаются практически решить эту задачу статистического моделирования с помощью уравнения $y - a_0 = a_1 x^{a_2}$ путем принятия ориентировочных значений a_0 . Однако результат такого решения может оказаться некорректным.

Тогда снова подбирают значение a_0 до тех пор, пока график на рис. 1б не будет глазомерно максимально приближен к множеству экспериментальных данных.

Этот простейший пример неавтоматизированного выбора конструкции модели показывает, что необходим некоторый перебор вариантов значений a_0 . Если принятое значение a_0 удовлетворяет критериям сходимости и адекватности модели к фактическим данным, то он запоминается. Так шаг за шагом выполняется случайный