УДК 629.4.014.66

В. В. СВИРИДЕНКО; старший научный сотрудник Белорусский государственный университет транспорта, г. Гомель

МЕТОДИКА ПРОГНОЗИРОВАНИЯ ПОКАЗАТЕЛЕЙ, ВЛИЯЮЩИХ НА ВЕЛИЧИНУ ПАРКА ПАССАЖИРСКИХ ВАГОНОВ, ПО ВРЕМЕННЫМ РЯДАМ

Одним из источников пополнения доходов Белорусской железной дороги являются пассажирские перевозки. Уровень доходности их во многом определяется рациональной величиной инвентарного парка пассажирских вагонов, техническим состоянием и комфортными условиями вагона. За последние шесть-семь лет материальная база, обеспечивающая пассажирские железнодорожные перевозки, претерпела большие изменения. Нет необходимого поступления нового подвижного состава, возросла зависимость от других государств по обеспечению запасными частями и материалами. Положение дел с пассажирским подвижным составом является достаточно тревожным. Постоянно повышается время нахождения вагона в эксплуатации, не всегда выдерживаются нормативные сроки поставки новых вагонов, техническое состояние имеющихся в эксплуатации вагонов ухудшается. В условиях необходимости обновления пассажирского подвижного состава Белорусской железной дороги, связанного в большей мере со старением вагонного парка, появилась необходимость принятия ряда мер по его обновлению. Выполненные исследования показали, что для получения прогнозных показателей ключевым вопросом в этой серьезной проблеме по обеспечению перевозочного процесса надежным и комфортабельным подвижным составом является установление рационального объема парка пассажирских вагонов как на существующую ситуацию, так и с разбивкой по периодам прогнозирования. Для этого, в свою очередь, необходима экспертная оценка факторов, которые потенциально оказывают влияние на величину пассажирского парка.

В теории и практике прогнозирования применяется широкий спектр математических методов: корреляционные функции и авторегрессионные модели, экспоненциального сглаживания, гармонического анализа, вероятностно-статистические модели. Все более широкое признание в последнее время получают вероятностно-статистические модели. Происходит это в силу того, что основа всех процессов, происходящих на транспорте, – вероятностная.

Динамическим (временным) рядом называется последовательность численных значений, характеризующих изменения исследуемого явления во времени. Наблюдаемые ряды могут быть дискретными или непрерывными. Дискретные ряды получаются при регистрации необходимых данных через определенные промежутки времени. Существуют следующие разновидности динамических рядов: моментальные, интервальные и ряды средних величин. Моментальные ряды – последовательность величин, относящихся к определенным датам, моментам времени. Интервальные ряды – это такие, когда используемые величины относятся к определенному промежутку времени. Ряды средних величин характеризуют изменение средних уровней исследуемого явления.

Непрерывные ряды получаются в том случае, если ведется непрерывная фиксация изучаемого явления.

Для анализа (изучения) характера, направления и интенсивности количественных изменений исследуемого явления дается оценка следующим понятиям: уровень ряда, средний уровень, абсолютные приросты (конечные разности), темпы роста, тренд.

Уровень ряда – это значение каждого члена ряда динамики. Средний уровень может быть определен в зависимости от характера ряда как средняя арифметическая или геометрическая. Абсолютный прирост характеризует изменение исследуемого явления и определяется разностью двух уровней. Необходимо различать базисные и цепные показатели прироста. Базисные приросты – это разности между уровнем любого члена и одним и тем же первоначальным (базисным) уровнем. Цепные абсолютные приросты получаются как разности последующего $(k + 1)$ и предыдущего k уровней. Темп роста определяет отношение двух уровней ряда динамики. Темпом прироста является отношение абсолютного прироста к базисному уровню. Трендом называется основная тенденция изменения явления.

Важнейшим требованием к временным рядам при построении моделей прогнозирования является сопоставимость уровня показателя по методам сбора информации, условиям развития тенденции и т. д. Использование динамических рядов для прогнозирования основано на инерционности экономических явлений: сохранении в будущем тех или иных тенденций прошлого периода. При резком изменении тенденций прогноз получится недостаточно обоснованным, другими словами, будет допущена большая ошибка прогноза [1].

Особенность моделирования по временным (динамическим) рядам состоит в том, что рассматриваемые уровни исследуемой величины выступают как случайные. В такой постановке динамический ряд реализует случайный процесс.

Для установления общей тенденции изменения исследуемого явления в течение изучаемого периода времени необходимо провести сглаживание (выравнивание) временного ряда. Эта необходимость обусловлена тем, что помимо влияния на уровни ряда главных факторов, под воздействием которых и формируется тренд динамического ряда, на них действуют еще и случайные факторы.

Для сглаживания ряда существуют две группы методов: механическое сглаживание и аналитическое выравнивание. Механическое сглаживание достигается применением скользящей средней (простой и взвешенной), аналитическое выравнивание – с помощью метода конечных разностей, наименьших квадратов, полиномов Лагранжа, авторегрессии и др. Метод скользящей средней преследует цель погасить случайные отклонения при определении средних значений. При сглаживании скользящей средней фактические значения динамического ряда заменяются средними значениями, характеризующими срединную точку периода сглаживания. При этом методе сглаживания искомый динамический ряд заменяется рядом, состоящим из средних арифметических для заданных промежутков времени.

Метод конечных разностей позволяет простейшим способом выявить приближенную номинальную зависимость исследуемого показателя от времени. Основная тенденция (тренд) показателя соответствует полиному P-й степени, если разности P-го порядка соседних членов ряда примерно равны. Если имеется ряд $Y_0, Y_1, Y_2, ..., Y_n$, то для него вычисляются первые, вторые и т. д. разности. Первые разности $\Delta_k^m = Y_{k+1} - Y_k$, вторые – $\Delta_k^{(2)} = \Delta_{k+1}^{(1)} - Y_k^{(1)}$ и т. д. при $k = 0, 1, 2, ..., n$–2. Если первые разности окажутся равными (примерно), то исследуемый показатель изменяется во времени линейно. Если равны между собой вторые разности, то предпочтительной зависимостью является парабола второго порядка и т. д.

Метод наименьших квадратов нашел наибольшее применение при сглаживании динамических рядов. Сущность метода заключается в нахождении таких параметров аппроксимирующей функции, при которых сумма квадратов отклонений расчетных значений уровней, вычисленных по полученной модели, от их фактических значений была бы минимальной.

С учетом сказанного выше простейшую модель прогнозирования можно получить, используя линейную зависимость вида

$$Y = a_0 + a_1 t, \tag{1}$$

где a_0, a_1 – коэффициенты, которые устанавливаются с использованием методов наименьших квадратов.

К зависимостям вида (1) можно привести и ряд нелинейных функций. Например, путем логарифмирования показательная $Y_1 = a_0 + a_1^t$ и дробностепенная $Y_2 = a_0 + t^{a_1}$ функции примут вид

$$\lg Y_1 = \lg a_0 + \lg t a_1; \; \lg Y_2 = \lg a_0 + a_1 \lg t.$$

Затем, обозначив $\lg Y_1 = Z_1$, $\lg Y_2 = Z_2$, $\lg a_0 = b_0$, $\lg a_1 = b_1$, $\lg t = X$, получим линейные уравнения

$$Z_1 = b_0 + b_1 t; \tag{2}$$

$$Z_2 = b_0 + b_1 X; \tag{3}$$

Неизвестные коэффициенты модели (3) определяются из условия

$$S = \sum_{i=1}^{n} (Y_i - a_0 - a_1 t_i)^2 \to \min, \tag{4}$$

где n – число наблюдений, т. е. длина временного ряда; Y_i – действительное количество пассажирских вагонов в инвентарном парке в i-й момент времени; t_i – год, предшествующий начальному году временного ряда (например, при временном ряде 1990–2002 гг., t_i = 1989).

Затем дифференцируя (4) по a_0 и приравнивая производную к нулю, получаем

$$\partial Y / \partial a_0 = 2\sum_{i=1}^{n} (Y_i - a_0 - a_1 t_i) = 0.$$

Учитывая, что $\sum_{i=1}^{n} a_0 = n a_0$, запишем

$$n a_0 + a_1 \sum_{i=1}^{n} t_i = \sum_{i=1}^{n} Y_i. \tag{5}$$

Аналогичным образом получаем

$$\partial Y / \partial a_1 = 2\sum_{i=1}^{n} (Y_i - a_0 - a_1 t_i) t_i = 0;$$

$$a_0 \sum_{i=1}^{n} t_i + a_1 \sum_{i=1}^{n} t_i^2 = \sum_{i=1}^{n} Y_i t_i. \tag{6}$$

Равенства (5) и (6) образуют систему нормальных уравнений, решение которой дает искомые значения a_0 и a_1.

$$a_0 = \frac{\sum_{i=1}^{n} Y_i \sum_{i=1}^{n} t_i^2 - \sum_{i=1}^{n} t_i \sum_{i=1}^{n} t_i Y_i}{A}; \tag{7}$$

$$a_1 = \frac{n\sum_{i=1}^{n} t_i Y_i - \sum_{i=1}^{n} t_i \sum_{i=1}^{n} Y_i}{A}; \tag{8}$$

$$A = n\sum_{i=1}^{n} t_i^2 - \left(\sum_{i=1}^{n} t_i \right)^2.$$

Расчет коэффициентов a_0 и a_1. по зависимостям (7) и (8) производится в предположении того, что ценность информации в начале периода наблюдения и в конце одинакова. Во многих исследованиях доказана необходимость учета экспоненциального убывания «веса» информации при внедрении внутрь ретроспективного ряда от прогнозного года до начального момента t_i. В таком случае в интервале времени $t_1 - t_п$ «вес» информации выступает как функция времени

$$\alpha_i = f(t) = \varphi(t) dt. \tag{9}$$

В интервале времени от t_{i-1} до t_{i+1}

$$\alpha_i = \int\limits_{t_{i-1}}^{t_{i+1}} \varphi(t)dt . \quad (10)$$

Для соблюдения условий нормирования должно выполняться равенство

$$\varphi(t_{\text{пр}}, t_{\text{п}}) = \int\limits_{t_{\text{пр}}}^{t_{\text{п}}} \varphi(t)dt = 1. \quad (11)$$

где $t_{\text{пр}}$ – номер года, в котором вес информации равен нулю.

В предложении об экспоненциальном убывании «веса» информации, получим

$$\varphi(t) = C - d\ln t . \quad (12)$$

Постоянные C и d могут быть определены, используя равенства

$$\varphi(t = t_{\text{пр}}) = 0;\ \ \varphi(t_1 = t_{\text{п}}) = 1. \quad (13)$$

Подставляя соответствующие значения в выражение (13), получим

$$\left.\begin{aligned} & C - d\ln t_{\text{пр}} = 0; \\ & \int\limits_{t_{\text{пр}}}^{t_{\text{п}}} (C - d\ln t)dt = 1. \end{aligned}\right\} \quad (14)$$

После соответствующих преобразований находим

$$d = \frac{1}{t_{\text{пр}} - \ln 2{,}718 t_{\text{пр}}};\ \ C = d\ln t_{\text{пр}}.$$

Тогда с учетом коэффициента α_i выражение (4) запишется в следующем виде:

$$S = \sum_{i=1}^{n} (Y_i - a_0 - a_1 t_i)^2 \alpha_i \to \min,$$

а система уравнений (5) и (6) примет вид

$$\left.\begin{aligned} & \sum_{i=1}^{n} Y_i\alpha_i - \sum_{i=1}^{n} (a_0 + a_1 t_i)\alpha_i = 0; \\ & \sum_{i=1}^{n} Y_i\alpha_i t_i - \sum_{i=1}^{n} (a_0 + a_1 t_i) t_i \alpha_i = 0. \end{aligned}\right\} \quad (15)$$

Решение системы (15) позволяет определить значения a_0 и a_1:

$$a_0 = \frac{\sum\limits_{i=1}^{n} Y_i\alpha_i \sum\limits_{i=1}^{n} (t_i\alpha_i)^2 - \sum\limits_{i=1}^{n} t_i\alpha_i \sum\limits_{i=1}^{n} t_i Y_i \alpha_i^2}{\sum\limits_{i=1}^{n} (t_i\alpha_i)^2 - \left(\sum\limits_{i=1}^{n} t_i\alpha_i\right)^2}; \quad (16)$$

$$a_1 = \frac{\sum\limits_{i=1}^{n} t_i Y_i \alpha_i^2 - \sum\limits_{i=1}^{n} t_i\alpha_i \sum\limits_{i=1}^{n} Y_i\alpha_i}{\sum\limits_{i=1}^{n} (t_i\alpha_i)^2 - \left(\sum\limits_{i=1}^{n} t_i\alpha_i\right)^2}. \quad (17)$$

Использование «веса» информации позволяет значительно повысить точность прогноза и ограничить влияние длительности ретроспективного ряда [2].

Расчет прогнозных моделей по оценке количества пассажирских вагонов в инвентарном парке Белорусской железной дороги проводился на основании данных службы статистики Управления Белорусской железной дороги и Министерства статистики и анализа Республики Беларусь [4, 5].

Динамические ряды подвергались анализу на подбор модели из трех возможных видов:

$$\left.\begin{aligned} & \text{линейная} - && Y = a_0 + a_1 t; \\ & \text{экспоненциальная} - && Y = a_0 e^{a_1 t}; \\ & \text{степенная} - && Y = a_0 t^{a_1}. \end{aligned}\right\} \quad (18)$$

Выбор модели из возможных трех определяется с помощью остаточной дисперсии

$$S^2_{\text{ост}} = \frac{1}{n}\sum_{i=1}^{n} (Y_{\text{в}i} - Y_{\text{ф}i})^2, \quad (19)$$

где $Y_{\text{в}i}$, $Y_{\text{ф}i}$ – соответственно вычисленное по полученной линии регрессии и фактическое значение количества пассажирских вагонов инвентарного парка в i-й точке.

Чем меньше величина остаточной дисперсии, тем наилучшим образом модель описывает динамический временной ряд. Для более полной характеристики соответствия модели временному ряду вычисляется коэффициент детерминации

$$R^2 = 1 - S^2_{\text{ост}} \big/ S^2_Y , \quad (20)$$

где S^2_Y – полная дисперсия показателя Y.

Полная дисперсия определяется, в свою очередь, по формуле

$$S^2_Y = \frac{1}{n}\sum_{i=1}^{n} (Y_{\text{ф}i} - \overline{Y})^2, \quad (21)$$

где $\overline{Y}$ – математическое ожидание (среднее) совокупности величины Y_i.

В результате выполненных исследований и расчетов была получена прогнозная модель трех видов (линейная, экспоненциальная и степенная) по временным рядам для каждого из факторов, влияющих на величину парка пассажирских вагонов ($X_1 - X_{11}$).

Модели по линейной функции:

$$X_1 = 5{,}53 + 0{,}003t, \quad S^2_{\text{ост}} = 0{,}0009;$$

$$X_2 = 105{,}13 - 6{,}54t, \quad S^2_{\text{ост}} = 246;$$

$$X_3 = 65{,}62 - 4{,}23t, \quad S^2_{\text{ост}} = 101{,}45;$$

$$X_4 = 167{,}13 - 0{,}97t, \quad S^2_{\text{ост}} = 195{,}45;$$

$$X_5 = 36{,}09 - 2{,}17t, \quad S^2_{\text{ост}} = 39{,}1;$$

$X_6 = 16{,}82 - 0{,}2t, \quad S^2_{\text{ост}} = 5{,}63;$

$X_7 = 102{,}15 - 1{,}63t, \quad S^2_{\text{ост}} = 17{,}47;$

$X_8 = 14{,}13 + 0{,}1t, \quad S^2_{\text{ост}} = 2{,}2;$

$X_9 = 10{,}31 - 0{,}017t, \quad S^2_{\text{ост}} = 0{,}006;$

$X_{10} = 19{,}18 - 1{,}39t, \quad S^2_{\text{ост}} = 67{,}04;$

$X_{11} = 18{,}97 - 1{,}39t, \quad S^2_{\text{ост}} = 67{,}49.$

Модели по экспоненциальной функции:

$X_1 = 5{,}53e^{0{,}00059t}, \quad S^2_{\text{ост}} = 0{,}0009;$

$X_2 = 1003{,}59e^{-0{,}086t}, \quad S^2_{\text{ост}} = 204{,}8;$

$X_3 = 64{,}62e^{-0{,}091t}, \quad S^2_{\text{ост}} = 84{,}05;$

$X_4 = 165{,}98e^{-0{,}006t}, \quad S^2_{\text{ост}} = 296{,}18;$

$X_5 = 37{,}69e^{-0{,}094t}, \quad S^2_{\text{ост}} = 43{,}34;$

$X_6 = 16{,}71e^{-0{,}014t}, \quad S^2_{\text{ост}} = 5{,}63;$

$X_7 = 102{,}22e^{-0{,}017t}, \quad S^2_{\text{ост}} = 17{,}22;$

$X_8 = 14{,}05e^{0{,}007t}, \quad S^2_{\text{ост}} = 2{,}2;$

$X_9 = 10{,}31e^{-0{,}002t}, \quad S^2_{\text{ост}} = 0{,}006;$

$X_{10} = 9{,}18e^{-0{,}037t}, \quad S^2_{\text{ост}} = 90{,}66;$

$X_{11} = 8{,}76e^{-0{,}034t}, \quad S^2_{\text{ост}} = 92{,}07.$

Модели по степенной функции:

$X_1 = 5{,}53t^{0{,}003}, \quad S^2_{\text{ост}} = 0{,}0009;$

$X_2 = 124t^{-0{,}44}, \quad S^2_{\text{ост}} = 78{,}59;$

$X_3 = 78{,}06t^{-0{,}46}, \quad S^2_{\text{ост}} = 38{,}11;$

$X_4 = 165{,}86t^{-0{,}021}, \quad S^2_{\text{ост}} = 300{,}16;$

$X_5 = 36{,}56t^{-0{,}034}, \quad S^2_{\text{ост}} = 67{,}41;$

$X_6 = 17{,}37t^{-0{,}077}, \quad S^2_{\text{ост}} = 5{,}46;$

$X_7 = 103{,}37t^{-0{,}072}, \quad S^2_{\text{ост}} = 21{,}43;$

$X_8 = 13{,}99t^{0{,}03}, \quad S^2_{\text{ост}} = 2{,}21;$

$X_9 = 10{,}28t^{-0{,}004}, \quad S^2_{\text{ост}} = 0{,}008;$

$X_{10} = 15{,}04t^{-0{,}45}, \quad S^2_{\text{ост}} = 63{,}65;$

$X_{11} = 14{,}54t^{-0{,}45}, \quad S^2_{\text{ост}} = 65{,}03.$

Проанализировав величину остаточной дисперсии при различных функциональных зависимостях $X(t)$, были приняты следующие модели для функции $X_i = f(t)$:

$X_1 = 5{,}53 + 0{,}003t;$

$X_2 = 124t^{-0{,}44};$

$X_3 = 78{,}06t^{-0{,}46};$

$X_4 = 167{,}13 - 0{,}97t;$

$X_5 = 36{,}09 - 2{,}17t;$

$X_6 = 17{,}37t^{-0{,}077};$

$X_7 = 102{,}22e^{-0{,}017t};$

$X_8 = 14{,}13 + 0{,}1t;$

$X_9 = 10{,}31 - 0{,}017t;$

$X_{10} = 15{,}04t^{-0{,}45};$

$X_{11} = 14{,}54t^{-0{,}45};$

Таким образом, наличие такой модели позволяет составить прогноз по временным рядам для факторов, характеризующих инвентарный парк пассажирских вагонов [3].

Список литературы

1 **Кобринский, Н. Е.** Точность экономико-математических моделей / Н. Е. Кобринский, В. И. Кузьмин. – М. : Финансы и статистика, 1981. – 255 с.

2 **Лукомский, Я. И.** Теория корреляции и ее применение к анализу производства / Я. И. Лукомский. – М. : Госстатиздат, 1961. – 375 с.

3 **Свириденко, В. В.** Экспертная оценка факторов, влияющих на потребный парк пассажирских вагонов Белорусской железной дороги / В. В. Свириденко // сб. науч. ст. / под ред. В. И. Сенько. – Гомель, 1998. – С. 54–56.

4 Статистический ежегодник Республики Беларусь. – Минск, 2006. – 586 с.

5 Транспорт и связь Республики Беларусь : стат. сб. – Минск, 1999. – 94 с.

Получено 19.09.2007

V. V. Svirydzenka. Technique of forecasting of the parameters influencing size of park of carriages, on time numbers.

One of sources of updating of incomes of the railway of Belarus are passenger transportations. The level of their profitableness in many respects is defined by rational size of inventory park of carriages, a technical condition and comfortable conditions of the car. For last six - seven years the material resources providing passenger rail transportation, has undergone the big changes. There is no necessary receipt of a new rolling stock, dependence on other states on maintenance with spare parts and materials has increased. The state of affairs with a passenger rolling stock is disturbing enough. Time of a presence of the car in operation constantly raises, normative terms of delivery of new cars are not always maintained, the technical condition of cars available in operation worsens. In conditions of necessity of updating of a passenger rolling stock of the Belarus railway connected in the greater measure with ageing of carload park necessity of acceptance of some measures on his updating has appeared. The executed researches have shown, that for reception The forecast parameters a key question in this serious problem on maintenance of transportation process with a reliable and comfortable rolling stock is the establishment of rational volume of park of carriages, both on an existing situation, and with breakdown on the periods of forecasting. The expert estimation of factors which potentially influence size of passenger park is in turn necessary for this purpose.