# **ON FORECASTING NAVIGATION SEASONS WITH MARKOV CHAINS**



Filippova, Nadezhda A., Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia.

Bogumil, Veniamin N., Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia.

Belyaev, Vladimir M., Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia.

## ABSTRACT

The transport network in the regions of the North of the Russian Federation basically remains seasonal (waterways, winter roads). The duration of river shipping season, depending on the climatic conditions, is 110–160 days, and the time of operation of winter roads varies within 120–210 days. Under these conditions, the accuracy of predicting the beginning and end of shipping season for the northern rivers plays a very important role. The article proposes a method for forecasting the duration of ice phenomena in the areas of shipping routes based on the use of the mathematical apparatus of Markov chains. An estimate of probability of an accurate forecast is given, taking into account the conformity with Bayes theorem and related dependencies. Verification of the method on the basis of real data proved that the forecast accuracy and probability of its implementation were sufficient for timely and effective organisation of preparatory operations for next shipping season on northern navigable rivers.

<u>Keywords:</u> northern river shipping, mathematical model, Markov chains, Bayes theorem, transport, cargo transportation, forecast, relative probability of events.

Background. The increase in air temperature occurring in recent decades on a global scale and in the territory of the Russian Federation has an impact on many natural processes, including the hydrological regime of rivers [1]. Since ice regime is the element, which is the most sensitive to warming, it becomes extremely important to identify long-term changes in the duration of freeze-up and in thickness of the ice cover. For obvious reasons, these studies are of particular importance in the conditions of the regions of the North of the Russian Federation. The changes in the timing of the beginning and end of freezing of revers. thickness of ice influence the duration of operation of winter roads, ice crossings, as well as of river shipping season [2, p. 226]. Under these conditions, it is important to clearly organize and control transportation of goods and passengers using modern satellite navigation and mobile communications [6-9, 14-17].

**Objective.** The objective of the authors is to forecast shipping seasons for northern rivers using Markov chains.

**Methods.** The authors use mathematical methods, theory of Markov random processes, methods of homogeneous Markov chains, data and statistics analysis, probability theory, Bayes theorem.

#### Results. 1. Statement of the task

The object of the research comprises a part of the natural system of the North of the Russian Federation, the indicators of which state change and to describe whose conditions are characterized by changing indicators, and it is suggested to describe the changes there-of with the help of Markov process theory.

Definition: in our case, the system will be understood as the natural system of the region, in a narrower sense as the nature of the water basin with conditions that characterize ice phenomena on the navigable river.

It is necessary to introduce the concept of the state of the system, particularly of the state associated with ice phenomena on the river. Then the system considered by us can get two states:

1) navigation – there is a process of transportation of goods and passengers;

2) lack of navigation – the process of transportation of goods and passengers is not carried out.

These two states involve two random events:

- · beginning of navigation;
- · end of navigation.

So-called ice phenomena on the river are associated with these random events. In spring, the following natural events of interest occur:

- beginning of the spring ice drift;
- end of the spring ice drift.
- In the autumn, events of opposite nature occur:
- · beginning of the autumn freeze-up;
- end of the autumn freeze-up.

Since each calendar date of the year corresponds to a day number, which varies from 1 to 365, we can say that random variables are associated with the above random events, respectively:

• the number of the day in the year when navigation began;

• the number of the day in the year when navigation ended.

The transition to the new state of the system for the spring period leads to formation of a new actual number of the day of beginning of navigation in the new season.

Transition to the new state of the system for the autumn period leads to formation of a new actual number of the end date for navigation in the new season.

Meteorological services have been collecting statistics for decades on beginning and end of the described natural phenomena. To use the available statistical data to forecast the timing of beginning and end of navigation, we make the following assumptions:

1. The beginning of navigation on the navigable river of the region coincides with the date of the end of ice phenomena in spring, which corresponds to the actual number of the day of the beginning of navigation.

2. The end of navigation on the navigable river of the region coincides with the date of the beginning of ice phenomena in the autumn, which corresponds to the actual number of the day of the end of navigation.

The accepted assumptions are confirmed by the analysis and comparison of the data of actual navigation periods and the specified periods of ice phenomena.

As a result, for the navigable river we have two statistical sequences of random numbers arranged chronologically. These random numbers are numbers

• WORLD OF TRANSPORT AND TRANSPORTATION, Vol. 17, Iss. 2, pp. 16–25 (2019)

Filippova, Nadezhda A., Bogumil, Veniamin N., Belyaev, Vladimir M. On Forecasting Navigation Seasons with Markov Chains



of calendar days, respectively, of beginning and end of navigation on the navigable river, arranged by years.

The set of possible values of the random sequence for the spring period is the whole set of possible numbers of the days of the beginning of navigation (coinciding with the set of dates for the end of spring ice phenomena) on the navigable river of the system.

The set of possible values of a random sequence for the autumn period is the whole set of possible numbers of the days of the end of navigation (coinciding with the set dates of the beginning of the autumn ice phenomena) on the navigable river of the system.

Conclusion: these two sequences of random numbers can legitimately be regarded as implementation of two random processes, the results of the analysis of which can be used to predict the beginning and end dates of navigation on the navigable river.

#### 2. Using the theory of Markov processes

Let us evaluate the possibility of applying the theory of Markov processes to the problem of forecasting the beginning of navigation. First of all, we note that the day of beginning of navigation is formed once a year and, in accordance with the accepted assumption, coincides with the number of the day of the end of ice phenomena. It follows that the process under consideration can be attributed to random processes with a fixed time (season) for transition from state to state. The set of possible values of a random variable is also limited, since the numbers of dates of transition to the expected state are within a limited set of numbers.

To use the theory of Markov processes, it is necessary that the probability of transition from one state to another should depend only on the current state in which the system is located, and should not depend on the trajectory along which the system came to this state. An analysis of the available statistical data on the dates of the beginning and end of navigation on the northern rivers shows that the indicated hypothesis is being fulfilled, and, therefore, the Markov process theory is applicable.

#### 3. Description of the forecasting task

Let's call the sequence of numbers of dates of opening of navigation arranged in chronological order as the random process «Beginning of navigation» and show that it can be considered as a Markov chain. We denote the Markov process «Beginning of navigation» considered by us as U(t), where  $t_i$  is the number of the day when navigation began in the i-th year, i.e. the subscript of the variable t will take the value of the number of the year to which the value of the random number «day number» belongs. For example, for the year 1940, the state of the Markov process will be denoted as  $U(t_{1940})$ . The value of the Markov process will be the

The value of the Markov process will be the number of the day on which opening of navigation occurred (the end of ice events) in the corresponding year. Let's suppose, according to statistics, that the end of the ice events fell on May 20. The number of the day of May 20 is «140». Then the value of the Markov process in 1940 will be determined as  $U(t_{1940}) = 140$ , and in that case the day of May 21 the value of the Markov process will receive the value of the day number «2», etc.

The process we are considering is development of an ordered set of random variables, the value of which depends on the actual date of opening of navigation. The described process refers to random processes with discrete time and a discrete finite set of states [3, 4].

Let in the general case the number of possible states of the process U(t) be n.

Transitions from one state to another can occur only at a fixed point in time. In our case, this is a new navigation year, which has its own number, counted from the beginning of the new era: 1, 2,..., k,... Thus, we get a step-by-step process in which the conditional year number is the system step number. The number of the day of the beginning of navigation is denoted in the k-th navigation year as  $S_k(i)$ .

In accordance with the accepted definition [3, p. 106], a random sequence is called a Markov chain if the conditions are met:

1. At any time t, a random sequence takes one of possible states  $S_{i}, S_{2^{i}}, ..., S_{n^{i}}$ 

2. For each step  $\tilde{k} = 1, 2,...$  the events  $S_1(k), S_2(k), ..., S_n(k)$  are inconsistent and form a complete group of events.

3. For each step, the probability of transition from any state  $S_i$  to any state  $S_j$  does not depend on when and how the system S was in the state of  $S_i$ .

The fulfillment of the first condition can be ensured if the moment of the change in the state of the random process  $U(t_i)$  introduced by us is superposed with the moment of beginning of navigation in each year. Then the condition that at any moment of time t the process





can remain in only one state is satisfied. The second condition is also satisfied, since the events of the random process U(t) i = 1, 2, 3,..., are by definition inconsistent and form a complete group of events.

As a hypothesis (which we formulate from the analysis of available data) we assume that the third condition is also satisfied.

Thus, the random process  $U(t_i)$  «Beginning of navigation», consisting in a random step-by-step change of the date of beginning of navigation, is considered to be a Markov chain. The set of possible values of the Markov chain is the set of possible numbers of dates of beginning of navigation.

Analysis of available data on beginning of navigation on the northern rivers shows that the process can be viewed as a homogeneous Markov one, in which the probability of transition to another state depends only on the current state of the process, but does not depend on the step number (in our case, on the year number) [3].

### 4. Building a matrix of transition probabilities

The baseline data for determining the transition probabilities matrix of a homogeneous Markov chain is the set of states  $S_1, S_2, ..., S_n$ , which in our case are statistical data on the starting dates of navigation on the navigable river.

Using the conditional number of the day instead of the beginning date of navigation provides the possibility to solve two problems:

1. Building a matrix of transition probabilities.

2. Obtaining an objective estimate of the probability for the state of the system in the next step based on its current state.

In fact, the second task is the task of generating a forecast for beginning of next year shipping season using information on the date of navigation in the current year and the matrix of transition probabilities.

The transition probability matrix (transition probabilities) characterizes the transition probabilities of the process with the current state  $S_i$  to the state  $S_j$  in the next step. This is a square matrix (1) with the number of rows and columns equal to the number of possible states. Each row of the matrix corresponds to one possible state. Each column is one possible transition state. The element of the matrix  $p_{ij}$  corresponds to the probability of transition of the system from state i to state j. The physical meaning of the probability  $p_{ij}$  means the probability of the beginning of the navigation period in the current year the navigation began on the *i*-th day.

In general, the Markov process has n possible states, the number of which in our case depends on the number of dates when navigation began in the navigable river in the region. Since the analysis shows that the transition probabilities do not depend on the step number, the Markov process, that we consider, is homogeneous. Thus, the dimension of the matrix will be  $n \cdot n$ .

General view of the matrix of transition probabilities:

 $P = [p_{ij}] = \begin{vmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{vmatrix}.$  (1)

On the main diagonal of the matrix (1) there are the probabilities that the system remains in the corresponding state. Since at each step the system can only be in one of the mutually exclusive states, for any non-zero row of the matrix, the sum of the probabilities  $p_{ij}$  will be equal to one:

$$\sum_{j=1}^{n} p_{ij} = 1 , \qquad (2)$$

where  $p_{ij}$  – probability of transition of the system from state *i* to state *j* at any step.

In accordance with the designated theoretical approach [3, 4], the forecast of the state of the system is estimated by the probabilities of possible states of the system at the next (k + 1) step, with a known state of the system at the k-th step represented by the homogeneous Markov chain model. The forecast can be calculated as follows.

Let the current state of the system be  $S_r$ . Then the *i*-th row of the matrix of transition probabilities  $p_{i1}, p_{i2},..., p_{in}$  shows the conditional probabilities of the onset of the state  $S_1, S_1,..., S_n$  in the next step, if the current state is  $S_r$ .

In our case, the interpretation of these events is as follows. The state of  $S_i$  is the number of the day in the current year when navigation began. The transition probabilities  $p_{i1}$ ,  $p_{i2}$ ,...,  $p_{in}$  are the probabilities that next shipping season will start on the day with the number  $S_n$ ,  $S_n$ ,...,  $S_n$  respectively.

5. Estimation of the probability of an accurate forecast

This kind of estimation has its meaning in terms of effectiveness of the method itself. To obtain it, it is necessary to enter the concept of an accurate forecast. Based on the already tested dependencies and assumptions:

1. The forecast using the apparatus of homogeneous Markov chains is constructed in the current year for the next year after beginning of navigation in the current year, i.e. after receiving information about the actual date of beginning of navigation.

2. We will consider the forecast accurate if the actual date of beginning of navigation in the next year will not differ from the forecast date by more than one day.

Let the system in the current year be in state  $S_{i}$ , which corresponds to the actual date of beginning of navigation in the current year. Let's choose some state  $S_i$  (it is the date of beginning of navigation) as a forecast of beginning of navigation next year, assuming that the total relative probability of transition from the current state  $S_i$  to state  $S_j$  or to neighboring states  $S_{j,1}$  and  $S_{j+1}$  is maximum. Let's denote the transition states for each current i-th state as  $S_{jmax}$ ,  $S_{jmax-1}$ ,  $S_{jmax+1}$ , and the probability of an accurate forecast as  $p(F_{exec})$ . Then, in accordance with the Bayes theorem [5, p. 42], the absolute probability of an accurate forecast is determined as follows:

$$p(F_{exec}) = \sum_{i=1}^{n} \frac{p(S_i)[p(S_{ij\max\to 1}) + p(S_{ij\max\to 1}) + p(S_{ij\max\to 1})]}{p(S_{ij\max\to 1})},$$
(3)

where  $p(S_i)$  – absolute probability of the event  $S_i$ :  $p(S_{ijmax})$ ,  $p(S_{ijmax-1})$ ,  $p(S_{ijmax-1})$  – relative probabilities of occurrence of events (consecutive dates of beginning of navigation in the next season) with the maximum total relative probability.

3. We will consider the forecast accuracy to be satisfactory if the actual date of beginning of navigation in the next year will not differ from the forecast date by more than three days.

Then, in accordance with the Bayes theorem [5], the absolute probability of an acceptable forecast is determined as follows:

• WORLD OF TRANSPORT AND TRANSPORTATION, Vol. 17, Iss. 2, pp. 16-25 (2019)

Filippova, Nadezhda A., Bogumil, Veniamin N., Belyaev, Vladimir M. On Forecasting Navigation Seasons with Markov Chains

$$p(F_{exec}) = \sum_{i=1}^{n} \frac{p(S_i)[p(S_{ijmn-s}) + p(S_{ijmn-s}) + p(S_{ijmn$$

where  $p(S_i)$  – absolute probability of the event  $S_i$ ;  $p(S_{ijmax})$ ,  $p(S_{ijmax-1})$ ,  $p(S_{ijmax-2})$ ,  $p(S_{ijmax-3})$ ,  $p(S_{ijmax+1})$ ,  $p(S_{ijmax+2})$ ,  $p(S_{ijmax+3})$  – relative probabilities of occurrence of events (groups of consecutive dates of beginning of navigation in the next season) with the maximum total relative probability.

4. We will consider the accuracy of the forecast unsatisfactory if the actual date of beginning of navigation in the next year will differ from the forecast date by more than three days.

**Conclusion.** The use of the Markov random processes theory makes it possible to develop a scientifically based forecast of the start of next river shipping season in the north of the Russian Federation long before resumption of cargo and passenger shipping. A similar approach can be applied to forecast the timing of completion of navigation. This helps to eliminate material losses associated with uncertainty of the period of operation of waterways in the zone of extreme climatic conditions [11–13]. Testing the method on real data showed that the forecast accuracy and the probability of its implementation are sufficient for effective organization and conduct of preparatory work before the next navigation on the northern navigable river.

#### REFERENCES

1. Vlasov, V. M., Efimenko, D. B., Zhankaziev, S. V. Use of GIS in dispatching technology of routed transport [*Ispolzovanie GIS v tekhnologii dispetcherskogo upravleniya marshrutizirovannym transporom*]. Moscow, MADI, 2007, 72 p.

2. Obyazov, V. A., Smakhtin, V. K. The ice regime of the Transbaikal rivers under the conditions of a changing climate [*Ledoviy rezhim rek Zabaikalya v usloviyakh izmenyayushchegosya klimata*]. *Vodnie resursy*, 2014, Iss. 3, pp. 227–234.

3. Ventzel, E. S., Ovcharov, L. A. The theory of random processes and its engineering applications [*Teoriya sluchainykh protsessov i ee inzhenernie prilozheniya*]. 2<sup>nd</sup> ed. Moscow, Vysshaya shkola publ., 2000, 383 p.

4. Ventzel, E. S., Theory of Probability: Textbook [*Teoriya veroyatnostei: Uchebnik*]. 4<sup>th</sup> ed., stereotypical. Moscow, Nauka publ., 1969, 576 p.

5. Gnedenko, B. V., Khinchin, A. Ya. An elementary introduction to probability theory [*Elementarnoe vvedenie v teoriyu veroyatnostei*]. 2<sup>nd</sup> ed. Moscow, Nauka publ., 1970, 169 p.

6. Vlasov, V. M., Efimenko, D. B. Coordinatetemporal and navigational support (KVNO) as a unified information basis for automation of basic technologies in transport [Koordinatno-vremennoe I navigatsionnoe obespechenie (KVNO) kak edinaya informatsiionnaya osnova avtomatizatsii bazovykh tekhnologii na transporte]. Fundamental and applied coordinate-temporal and navigational support: Proceedings of the 2<sup>nd</sup> All-Russian Conference. St. Petersburg, IPA RAS, 2007. [Electronic resource]: https://search.rsl.ru/ru/record/01003374619. Last accessed 14.02.2019.

7. Vlasov, V. M., Efimenko, D. B., Bogumil, V. N. Information technologies for automobile transport: A textbook for students of higher education institutions [Informatsionnie tekhnologii na avtomobilnom transporte: Uchebnik dlya studentov uchrezhdenii vysshego obrazovaniya]. Moscow, Academia publ., 2014, 256 p.

8. Vlasov, V. M., Bogumil, V. N., Efimenko, D. B. Use of digital infrastructure and telematic systems in urban passenger transport: Textbook [*Primenenie tsifrovoi infrastruktury i telematicheskih system na gorodskom passazhirskom transporte: Uchebnik*]. Moscow, Infra-M publ., 2018, 352 p.

9. Vlasov, V. M., Maktas, B. Ya., Bogumil, V. N., Konin, I. V. Wireless technologies in road transport. Global navigation and vehicle positioning: Study guide [Besprovodnie tekhnologii na avtomobilnom transporte. Globalnaya navigatsiya i opredelenie mestopolozheniya transportnykh sredstv: Ucheb. posobie]. Moscow, Infra-M publ., 2017, 184 p.

10. Filippova, N. A., Belyaev, V. M. Analysis of the process of managing the mass delivery to Northern regions in modern market conditions [*Analiz protsessa upravleniya severnym* zavozom v sovremennykh rynochnykh usloviyakh]. Gruzovoe i passazhirskoe avtokhozyaistvo, 2010, Iss. 9, pp. 17–20.

11. Filippova, N. A., Belyaev, V. M. Adaptive mathematical model for optimizing the delivery of goods in the conditions of the North [*Adaptivnaya matematicheskaya model dlya optimizatsii zavoza gruzov v usloviyakh Severa*]. Gruzovoe i passazhirskoe avtokhozyaistvo, 2013, Iss. 11, pp. 17–20.

12. Filippova, N. A. Methodology of organization and functioning of the systems of cargo delivery to the northern regions: Monograph [*Metodologiya organizatsii ifunktsionirovaniya system dostavki gruzov v severnie regiony: Monografiya*]. Moscow, Tekhpoligraftsentr, 2015, 208 p.

13. Belyaev, V. M., Filippova, N. A. Fundamentals of Transport System Organization in Northern Regions. *World of Transport and Transportation*, Vol. 15, 2017, Iss. 1, pp. 162–167.

14. Efimenko, D. B., Maksimychev, O. I., Ostroukh, A. V., Zbavitel, P. Yu., Ivakhnenko, A. M., Karelina, M. Y. Technology of Monitoring and Control Algorithm Design for Earth-Moving Machine. *International Journal of Applied Engineering Research*, 2016, Vol. 11, Iss. 9, pp. 6430–6434.

15. Efimenko, D., Ostroukh, A., Nuruev, Ya., Zhankaziev, S., Moroz, D. Automated dispatching control system of the mobile concrete batching plants. *ARPN Journal of Engineering and Applied Sciences*, Vol. 11, Iss. 11, June 2016, pp. 6733–6737.

16. Bogumil, V., Efimenko, D. Urban Transport Dispatch Control System Helps to Increase Intelligent Transport Systems Effectiveness. In: *Proceedings of the 11<sup>th</sup> European transport congress*, Prague, September 19–20, 2013, pp. 20–25.

17. Bhouri, N., Balbo, F., Pinson, S. An agent-based computational approach for urban traffic regulation. *Progress in Artificial Intelligence*, 2012, pp. 139–147, DOI: 10.1007/s13748–012–0011–0.

Information about the authors:

Filippova, Nadezhda A. – Ph.D. (Eng), associate professor of Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia, madizp@mail.ru.

**Bogumil, Veniamin N.** – Ph.D. (Eng), associate professor of Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia, v bogumil@mail.ru.

**Belyaev, Vladimir M.** – D.Sc. (Eng), professor of the department of management of Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia, belyaev-v@mail.ru.

Article received 06.11.2018, accepted 05.02.2019.

• WORLD OF TRANSPORT AND TRANSPORTATION, Vol. 17, Iss. 2, pp. 16–25 (2019)

Filippova, Nadezhda A., Bogumil, Veniamin N., Belyaev, Vladimir M. On Forecasting Navigation Seasons with Markov Chains

