2. The complex application of monitoring and express diagnosing for searching failures on common rail system units / I. Gritsuk et al. SAE Technical Paper. 2018-01-1773. 2018. DOI: https://doi.org/10.4271/2018-01-1773. [in English]

3. Information security risk management of vehicles / D. Klets et al. SAE Technical Paper. 2018-01-0015. 2018. DOI: https://doi.org/10.4271/2018-01-0015. [in English]

4. Матейчик В. П. Системний підхід до аналізу структурних схем енергоустановок транспортних засобів. Вісник НТУ «ХПІ». 2002. № 7. Т. 2. С. 162-167. [Ukrainian]

5. Metody systemnoho analizu vlastyvostej avtomobil'noi tekhniky: navch. posib. / Dmytrychenko M. F., Matejchyk V. P., Hryschuk O. K., Tsiuman M. P. Kyiv : NTU, 2014. 168 s. [Ukrainian]

6. Katrenko A. V. Systemnyj analiz ob'iektiv ta protsesiv komp'iuteryzatsii: Navchal'nyj posibnyk. L'viv : «Novyj svit – 2000», 2003. 424 s. [Ukrainian]

7. Matejchyk V.P., Navrots'kyj A.V. 2024. Systematyzatsiia skhem aparatnoho zabezpechennia system monitorynhu ekspluatatsii transportnykh zasobiv. Visnyk Pryazovs'koho Derzhavnoho Tekhnichnoho Universytetu. Seriia: Tekhnichni nauky. 48 (Cher 2024), 184–192. DOI:https://doi.org/10.31498/2225-6733.48.2024.310711. [Ukrainian]

FORMAL APPROACHES TO DEVELOPING AN EXPERT SYSTEM FOR EVALUATING A NAVIGATOR’S QUALIFICATION BASED ON SHIP TRAJECTORY DATA

Nosov, P.
Ph.D.,
Associate Professor
Kherson State Maritime Academy
Kherson, Ukraine
ORCID ID https://orcid.org/0000-0002-5067-9766
Ponomaryova, V.
Graduate student
Kherson State Maritime Academy

Kherson, Ukraine
Ben, A.
Ph.D., Professor
Kherson State Maritime Academy
Kherson, Ukraine
ORCID ID https://orcid.org/0000-0002-9029-3489
Prokopchuk, Y.
Dr.s, Associate Professor
National Academy of Sciences of Ukraine
Kyiv, Ukraine

Abstract. The article presents a mathematical model for a comprehensive assessment of the navigator's qualification based on the analysis of the vessel's trajectory. The model considers technical parameters such as speed and course, as well as additional factors like external conditions (weather, currents) and the navigator's psychological state (fatigue, stress). To enhance the accuracy and reliability of the assessment, the model integrates data processing methods, including the Kalman filter, and machine learning algorithms for the automatic optimization of weight coefficients. The model's adaptability to various operational conditions, its validation on different datasets, and its integration with other ship management systems for providing real-time recommendations are also explored. Furthermore, the model allows for future expansion by incorporating additional parameters, such as fuel consumption and maneuvering efficiency.

Keywords: navigator qualification, trajectory analysis, machine learning, Kalman filter, mathematical model, external conditions, parameter optimization.

Introduction. In modern maritime navigation, ensuring the high qualification of navigators is crucial for guaranteeing the safety of shipping and the efficiency of operations [1,2]. Traditional approaches to assessing a navigator's qualification often fail to account for the dynamic factors influencing a vessel's trajectory, such as course changes, speed variations, external conditions, and the navigator's psychological state [3,4]. To enhance the accuracy and reliability of such assessments, it is necessary to implement automated systems that consider these factors and utilize advanced data processing methods, including machine learning and filters for trajectory correction. This paper presents a mathematical model designed to evaluate the navigator's qualification level based on the analysis of vessel movement parameters, taking into account additional influencing factors.

Presentation of the core research material and its results

Under the proposed framework, a mathematical model for an automated expert system is suggested to evaluate a navigator's qualification by functionally analyzing the ship's trajectory in the Bosphorus Strait.

1. Data Collection and Preprocessing [5].

1.1. Input Data

The dataset consists of measurements of the ship's movement parameters:

Times tamps: $\{t_i\}_{i=1}^{N}$, where t_i - the observation time, i =1,2,...,N.

Position Coordinates: $\{(x_i, y_i)\}_{i=1}^{N}$, where x_i, y_i - are ship's coordinates at time t_i.

Ship Speed: $\{v_i\}_{i=1}^{N}$, where v_i - the actual speed at time t_i.

Ship Course: $\{\theta_i\}_{i=1}^{N}$, where θ_i - is the actual heading at time t_i.

1.2. Data Preprocessing

Noise and Anomaly Removal: A Kalman filter is used to smooth the data and eliminate noise.

Interpolation of Missing Data: If there are any data gaps, interpolation is applied.

2. Trajectory and Movement Dynamics Modeling [6].

2.1. Selection of Speed and Course Models

To model the speed $v(t)$ and course $\theta(t)$ several models are tested, and the optimal one is selected based on the best fit quality criteria.

2.1.1. Linear Model

Speed: $v(t) = a_v t + b_v$

a_v - is the speed change rate over time.

b_v - is the initial speed t = 0.

Course: $\theta(t) = a_\theta t + b_\theta$

a_θ - is the course change rate over time.

b_θ - is the initial course $t = 0$.

2.1.2. Second-Order Polynomial Model

Speed: $v(t) = a_v t^2 + b_v t + c_v$

a_v, b_v, c_v - are the model coefficients.

Course: $\theta(t) = a_\theta t^2 + b_\theta t + c_\theta$

a_θ, b_θ, c_θ - are the model coefficients.

2.1.3. Exponential Model

Speed: $v(t) = A_v e^{kvt}$

A_v - is the initial speed.

k_v - is the exponential growth/decay rate.

Course: $v(t) = A_v e^{k\theta t}$

A_θ - is the initial course.

k_θ - is the exponential rate of course change.

2.1.4. Logarithmic Model

Speed: $v(t) = a_v ln(t+\delta_t)+bv$

δ_t - is a small positive constant to avoid the logarithm of zero.

Course: $\theta\ (t) = a_\theta ln(t+\delta_t)+bv$

2.2. Model Parameter Fitting

Model parameters are selected by minimizing the loss function (least squares method):

For speed: $L_v = \sum_{i=1}^{N} \left[v_i - \hat{v}(t_i) \right]^2$

$\hat{v}(t_i)$ - is the modeled speed value.

For course: $L_\theta = \sum_{i=1}^{N} \left[\theta_i - \hat{\theta}(t_i) \right]^2$

$\hat{\theta}(t_i)$ - is the modeled heading value.

2.3. Optimal Model Selection

The coefficient of determination R^2 is calculated for each model. The model with the highest R^2 and the lowest root mean square error (RMSE) is selected. Cross-validation is performed to verify the model's generalization ability.

Cross-Validation for Model Generalization Testing [7].

We perform cross-validation to verify the model's generalization capability.

3. Noise and Anomaly Processing

3.1. Application of the Kalman Filter

The Kalman filter is used for data smoothing and system state prediction.

3.1.1. Recursive Kalman Filter Equations

State Prediction:

$$\hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1} + Bu_{k-1}.\ P_{k|k-1} = AP_{k-1|k-1}A^T + Q.$$

State Update:

$$K_k = P_{k|k-1}H^T\left(HP_{k|k-1}H^T + R\right)^{-1}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k\left(z_k - H\hat{x}_{k|k-1}\right)$$

$$P_{k|k} = \left(I - K_kH\right)P_{k|k-1}$$

Where:

$\hat{x}_{k|k-1}$ - is the predicted state at step k.

$\hat{x}_{k|k}$ - is the updated state.

A - is the state transition matrix.

B - is the control matrix.

u_{k-1} - is the control vector.

P - is the state error covariance matrix.

Q - is the process covariance matrix.

R - is the measurement covariance matrix.

K_k - is the Kalman gain.

z_k - is the measured vector.

H - is the observation matrix.

I - is the identity matrix.

3.2. Anomaly Detection

Statistical methods are used to detect anomalous deviations, such as points that fall beyond three standard deviations from the mean.

Clustering techniques can also be applied to detect atypical movement patterns.

4. Deviation and Stability Assessment [8].

4.1. Deviation Calculation

Speed Deviation: $\varepsilon_{v_i} = v_i - \hat{v}(t_i)$. Course Deviation: $\varepsilon_{\theta_i} = \theta_i - \hat{\theta}(t_i)$

4.2. Statistical Characteristics of Deviations

Mean Deviation:

Speed: $\overline{\varepsilon_v} = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{v_i}$. Course: $\overline{\varepsilon_\theta} = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{\theta_i}$

Root Mean Square Deviation (RMSD):

Speed: $\sigma_v = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon_{v_i} - \overline{\varepsilon_v}\right)^2}$, Course: $\sigma_\theta = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon_{v_i} - \overline{\varepsilon_\theta}\right)^2}$

Coefficient of Variation:

Speed: $CV_v = \frac{\sigma_v}{\overline{v}} \times 100\%,\ \overline{v} = \frac{1}{N}\sum_{i=1}^{N} v_i$.

Course: $CV_\theta = \frac{\sigma_\theta}{\overline{\theta}} \times 100\%,\ \overline{\theta} = \frac{1}{N}\sum_{i=1}^{N}\theta_i$

5. Model Fit Quality Assessment

5.1. Coefficient of Determination R^2

For speed: $R_v^2 = 1 - \frac{\sum_{i-1}^{N}\left(v_i - \hat{v}(t_i)\right)^2}{\sum_{i-1}^{N}\left(v_i - \overline{v}\right)^2}$. For course: $R_\theta^2 = 1 - \frac{\sum_{i-1}^{N}\left(\theta_i - \hat{\theta}(t_i)\right)^2}{\sum_{i-1}^{N}\left(\theta_i - \overline{\theta}\right)^2}$

5.2. Model Validation

Cross-validation: We split the data into training and test sets to evaluate the model’s generalization ability.

Residual Analysis: We check whether the residuals are randomly distributed, without systematic deviations.

6. Determining the Integral Qualification Indicator

6.1. Selection and Optimization of Weight Coefficients

Weight coefficients w_v, w_θ та α_v, β_v, α_θ, β_θ are selected based on expert evaluations or through optimization, such as the Analytical Hierarchy Process (AHP) or using machine learning techniques.

6.2. Calculation of Indicators for Speed and Course

Speed: $Q_v = \alpha_v R_v^2 + \beta_v \left(1 - CV_v\right)$, $\alpha_v + \beta_v = 1$

Course: $Q_\theta = \alpha_\theta R_\theta^2 + \beta_\theta \left(1 - CV_\theta\right)$, $\alpha_\theta + \beta_\theta = 1$

6.3. Integral Qualification Indicator

$Q = w_v Q_v + w_\theta Q_\theta$, $w_v + w_\theta = 1$

7. Considering Additional Factors

7.1. External Conditions

The correction coefficient K_{env} accounts for the influence of weather conditions, currents, etc., typically ranging from $0.8 \le K_{env} \le 1.00$.

7.2. Psychological and Physiological Factors

The psychological state coefficient K_{psych} considers the state of the navigator (fatigue, stress).

7.3. Maneuver Dynamics Coefficient

The coefficient $K_{dynamic}$ accounts for the complexity of the maneuvers performed.

7.4. Final Formula Considering Additional Factors

$Q_{final} = Q \times K_{env} \times K_{psych} \times K_{dynamic}$

8. Use of Machine Learning

8.1. Recurrent Neural Networks (RNN)

Used for modeling speed and course data sequences. They take into account nonlinear and complex temporal dependencies.

8.2. Clustering and Pattern Recognition

Clustering algorithms (e.g., k-means) help identify typical movement patterns and anomalies. They assist in creating navigator behavior profiles.

8.3. Optimization of Weight Coefficients

Machine learning methods are used to automatically optimize the weight coefficients based on historical data and outcomes.

9. Final Formula for the Comprehensive Indicator

$Q_{comprehensive} = Q_{final} = (w_v Q_v + w_\theta Q_\theta) \times K_{env} \times K_{psych} \times K_{dynamic}$

Conclusion. The presented mathematical model allows for a detailed assessment of a navigator's qualification based on the analysis of the vessel's trajectory, considering both technical parameters (speed, course) and additional factors (external conditions, psychological state). The inclusion of data processing methods, such as the Kalman filter, and machine learning algorithms increases the system's accuracy and reliability.

To improve the model's performance, it is essential to validate it across diverse datasets to ensure its generalizability, while also developing mechanisms for automatic parameter adjustments based on specific operational conditions. Integrating the model into the overall ship management system will enable real-time recommendations for navigators. Regular expert assessments should be conducted to refine the model based on real-world observations, and the model should be expanded to account for additional factors like fuel consumption and maneuvering efficiency for a more comprehensive evaluation.

List of references

1. Nosov P. S., Palamarchuk I.V., Safonov M.S., Novikov V.I. (2018) Modeling the manifestations of the human factor of the Maritime crew // Dnipropetrovsk National University of Railway Transport named after Academician V. Lazaryan (Dnipro) № 5 (77). – 2018. Pages 82-92. https://doi.org/10.15802/stp2018/ 147937. [in English]

2. Solovey, O., Ben, A., Dudchenko, S., Nosov, P. (2020). Development of control model for loading operations on Heavy Lift vessels based on inverse algorithm. Eastern European Journal of Enterprise Technologies, 5/2 (107), 48–56. https://doi.org/10.15587/1729-4061.2020.214856. [in English]

3. Zinchenko, S. M., Nosov, P. S., Mateychuk, V. M., Mamenko, P. P., & Grosheva, O. O. (2020). AUTOMAT-IC COLLISION AVOIDANCE WITH

MULTIPLE TARGETS, INCLUDING MANEUVERING ONES. Ra-dio Electronics, Computer Science, Control, (4), 211–222. https://doi.org/10.15588/1607-3274-2019-4-20. [in English]

4. Nosov, P. S., Ben, A. P., Matejchuk, V. N., & Safonov, M. S. (2019). IDENTIFICATION OF "HUMAN ER-ROR" NEGATIVE MANIFESTATION IN MARITIME TRANSPORT. Radio Electronics, Computer Science, Control, (4). https://doi.org/10.15588/1607-3274-2018-4-20. [in English]

5. Baesens, Bart & Van Vlasselaer, Véronique & Verbeke, Wouter. (2015). Data Collection, Sampling, and Preprocessing. 10.1002/9781119146841.ch2. [in English]

6. U., Osuagwu & I., Okenwe. (2024). Simple Regression Models: A Comparison using Criteria Measures. African Journal of Mathematics and Statistics Studies. 7. 131-143. https://doi.org/10.52589/AJMSS-VKLVNUU5.

7. Lyu, Boyang & Nguyen, Thuan & Scheutz, Matthias & Ishwar, Prakash & Aeron, Shuchin. (2023). A principled approach to model validation in domain generalization. https://doi.org/10.48550/arXiv.2304.00629. [in English]

8. Jin, Sheng & Xiao, Yuzhi & Han, Jiaxin & Huang, Tao. (2024). An Evaluation Model for Node Influence Based on Heuristic Spatiotemporal Features. Entropy. 26. 676. https://doi.org/10.3390/e26080676. [in English]