UDC 504.064 + 004.9
ICSTI 20.53.21
DOI 10. 56525/PATZ1830

INTEGRATED ENVIRONMENTAL MONITORING INFORMATION SYSTEM FOR THE CASPIAN SEA

Kenzhebaeva Zh., Kazieva G.
Yessenov University, Aktau, Kazakhstan
e-mail: zhanat.kenzhebayeva@yu.edu.kz, gulnara.kaziyeva@yu.edu.kz
***Corresponding author**: zhanat.kenzhebayeva@yu.edu.kz

Abstract. This article proposes a holistic Environmental Monitoring Information System (EMIS) designed to assess and manage ecological conditions along the northern coast of the Caspian Sea. By integrating multiple data sources—sensor networks, remote sensing, meteorological data, and machine learning algorithms—the EMIS delivers real-time insights into water quality parameters, hydrocarbon pollution, and habitat stability. A scalable, modular architecture ensures that diverse datasets are synthesized into unified, actionable the information, enabling stakeholders to detect anomalies, forecast pollution trends, and identify potential ecological risks. Rigorous data validation processes mitigate discrepancies arising from sensor drift or transmission failures, thereby enhancing reliability. The system also features a cloud-based infrastructure and role-based user interfaces tailored to local requirements, the fostering collaborative decision-making across government agencies, industry operators, and local communities. This study delineates the technical blueprint of the EMIS, documents pilot test outcomes, and highlights implications for sustainable resource utilization and regional environmental policy development.

Keywords. Environmental Monitoring Information System (EMIS), Northern Caspian Sea, Sensor Networks, Remote Sensing, Data Fusion, Predictive Modeling, Sustainable Resource Management, Water Quality and Hydrocarbon Pollution, Real-Time Analytics, Coastal Habitat Protection.

Introduction.

Context and Importance

The Caspian Sea, the largest landlocked body of water in the world, has significant geopolitical and ecological importance. Its northern coast, largely under the jurisdiction of Kazakhstan and neighboring countries, combines unique biodiversity with expansive industrial operations, particularly in the oil and gas sectors. This region's ecological balance is fragile, confronted by sea-level fluctuations, variations in salinity, and anthropogenic stresses from ongoing hydrocarbon extraction. The need for accurate and timely environmental monitoring is therefore paramount to ensuring ecological resilience and safeguarding local communities that depend on the Caspian's resources.

An Environmental Monitoring Information System (EMIS) serves as a unifying platform for real-time data acquisition, processing, and interpretation. Traditional monitoring approaches typically rely on sporadic in situ sampling or limited remote sensing snapshots, resulting in data fragmentation. By contrast, a well-designed EMIS collates multiple data streams—sensor nodes, satellite imagery, meteorological information—into a cohesive framework. This integration supports faster and more informed decision-making, potentially averting severe ecological consequences.

Motivation and Objectives

Ongoing industrial activities such as oil extraction, maritime shipping, and coastal development impose significant stress on the Caspian Sea's northern coast. Operators, regulators, and local communities often face the challenge of recognizing ecological threats before they escalate into large-scale crises. The motivation behind this research is to close gaps in environmental monitoring by providing a robust, scalable, and user-centric information system. Key objectives include:

1. **Comprehensive Data Integration**

Merge heterogeneous data sources—sensor networks, satellite images, meteorological inputs—into a single, harmonized dataset.

2. **Real-Time Analytics and Predictive Modeling**

Employ machine learning algorithms to detect anomalies, forecast pollutant dispersal, and anticipate critical shifts in water quality or habitat conditions.

3. **Stakeholder Engagement**

Facilitate access to processed environmental data for government agencies, industrial stakeholders, and local communities, thereby promoting transparent and collaborative decision-making.

4. **Sustainability and Scalability**

Ensure that the system architecture can be adapted or expanded to other regions of the Caspian Sea, thus maximizing its long-term utility and cost-effectiveness [1].

5. Scope and Limitations

Although designed for the Caspian's northern coast, many system components—data ingestion pipelines, machine learning frameworks—can be reconfigured for other geographies. The study does not address socioeconomic or geopolitical aspects of resource exploitation, focusing primarily on the technical underpinnings of monitoring. In addition, the EMIS's performance is contingent on uninterrupted power supply, stable communication channels, and long-term maintenance funding. Ensuring local expertise and political will to sustain the system remains a practical challenge outside the scope of this immediate research.

Literature review.

Historical Overview of Caspian Sea Monitoring

Studies of the Caspian Sea's environment date back several decades, emphasizing biodiversity assessments, salinity gradients, and pollution tracking (Gao & Ma, 2018). Many of these efforts, however, were constrained to fragmented data collection, such as discrete sampling or annual surveys. These methods, while valuable for establishing baseline conditions, do not capture rapid changes or emergent threats in real time [2].

Remote Sensing in Coastal Monitoring

Recent advances in satellite technologies, including higher-resolution sensors on platforms like Sentinel-2 and Landsat 8, have broadened the scope of coastal monitoring. Researchers have used multispectral and thermal data to detect oil slicks, eutrophication events, or variations in turbidity (Mansour, Singh, & Chawla, 2019). Nevertheless, limitations exist in temporal resolution and cloud interference, underscoring the necessity of complementary in situ measurements.

Sensor Networks and IoT

In situ sensor networks offer granular, high-frequency measurements of parameters such as temperature, salinity, pH, and dissolved oxygen. Low-power wide-area network (LPWAN) protocols, including LoRaWAN, have enabled cost-effective data transmission from remote locations (Park, Chae, & Jung, 2019). However, sensor nodes in marine environments face hardware degradation due to salt exposure and biofouling, necessitating frequent maintenance.

Data Fusion and Big Data Analytics

A growing body of research underscores the value of data fusion in marine monitoring. Techniques like Kalman filtering help reconcile discrepancies between satellite and sensor data, while big data platforms facilitate parallel processing of massive datasets (Zhan, Li, & Chen, 2020). Machine learning models, including Random Forests and deep neural networks, have been employed for anomaly detection in water quality, species distribution prediction, and climate impact analysis (Liu, Wang, & Zhao, 2021).

Gaps in Existing Systems

Although various components—sensors, remote sensing, machine learning—are individually well-documented, few end-to-end solutions integrate all these elements into a single, region-focused

EMIS specifically tailored for the Caspian Sea. Moreover, user interface design and stakeholder engagement strategies remain underexplored areas, leading to a gap between technological capabilities and real-world utilization (Brown, Chen, & Zhang, 2022).

Methodology

Research Framework

The research followed a structured, iterative methodology combining system design, prototyping, and validation. Collaboration was established with local universities, government environmental agencies, and industry partners to define use cases and performance benchmarks [3].

Step 1: Requirement Analysis

Workshops and interviews identified critical parameters (temperature, salinity, hydrocarbon pollution, etc.) and user requirements (alerts, real-time dashboards, predictive analytics).

Step 2: System Design

A modular approach delineated distinct layers: data acquisition, data processing, and information delivery. Design documents included specifications for sensor hardware, communication protocols, cloud infrastructure, and user interfaces.

Step 3: Prototyping and Implementation

Software prototypes were developed for real-time data ingestion, preprocessing, and machine learning tasks. Sensors were deployed for pilot testing in select coastal areas, each with distinct ecological and industrial footprints.

Step 4: Validation and Refinement

Field trials evaluated system reliability, sensor accuracy, data latency, and user experience. Feedback from stakeholders led to iterative improvements, such as adding bilingual support and enhancing alert thresholds [4].

3.2 Pilot Area Selection

The northern Caspian region exhibits shallow waters, variable salinity, and significant ice coverage in winter. Sites were chosen to represent a spectrum of environmental pressures: near industrial outfalls, major population centers, and relatively pristine coastal stretches. Each site's local conditions—wave exposure, shipping routes, or discharge pipelines—provided insights into the EMIS's adaptability and resilience.

3.3 Sensor Deployment

To capture critical environmental parameters, sensor arrays measured temperature, conductivity (or salinity), pH, dissolved oxygen, and hydrocarbons (oil contamination). Each array was installed on a fixed mooring structure within about 3 km of the coastline. Data were logged at intervals ranging from 15 to 60 minutes, depending on power constraints and bandwidth availability.

3.4 Data Handling

Communication

Sensor arrays transmitted data to onshore gateways via LPWAN modules. The gateways used cellular (4G/5G) or satellite connections to send the data to a cloud-based server. Buffering mechanisms safeguarded data integrity during network outages [5].

Preprocessing and Storage

Raw sensor readings underwent quality checks to remove obvious outliers or sensor glitches. A metadata tagging system associated each reading with sensor ID, timestamp, and geospatial coordinates. Processed data were stored in a cloud-hosted database optimized for time-series queries [6].

Data Fusion

Remote sensing images (Sentinel-2, Landsat 8) were periodically retrieved and processed for surface temperature, turbidity, and chlorophyll-a indices. A Kalman filter algorithm combined satellite-based and in situ measurements, reducing uncertainties arising from sensor noise or cloud-covered satellite images [7].

Machine Learning and Predictive Models

Supervised learning models, including Gradient Boosted Trees and Long Short-Term Memory (LSTM) networks, were trained to forecast short-term fluctuations in water quality, specifically for oil pollution or algae bloom events. Historical data informed model calibration, and cross-validation ensured robust hyperparameter tuning. Alerts were triggered when predicted values exceeded threshold levels established by environmental standards or empirical distributions [8].

User Interface

A web dashboard, developed using open-source frameworks, presented time-series graphs, tabular summaries, and event-based notifications. Role-based access differentiated between government agencies (policy formulation), industry users (compliance monitoring), and academic researchers (advanced analytics). Mobile compatibility allowed data viewing and alert management from smartphones or tablets.

Ethical and Regulatory Compliance

Since the EMIS collected environmental and limited personal data from participating field teams, the project followed applicable regulations for data protection. Consent was obtained from local authorities to install sensors, and information regarding data usage was disclosed to community representatives.

Results and Discussion

Pilot Implementation Outcomes

Over a six-month pilot phase, the system collected and processed more than 2 million data points, spanning water temperature, salinity, pH, dissolved oxygen, and hydrocarbon levels. Despite occasional communication dropouts due to harsh weather or network congestion, the built-in buffering protocols ensured a data completeness rate of 96%. Sensor drift remained negligible, with monthly calibrations ensuring stable accuracy.

Anomaly Detection

Machine learning models identified multiple spikes in hydrocarbon concentrations near industrial sites, correlating with operational discharge schedules. The system also flagged a sudden drop in dissolved oxygen during a suspected algal bloom, prompting immediate field investigation. Early detection enabled timely mitigation strategies, such as adjusting wastewater treatment protocols or issuing local advisories.

Stakeholder Engagement

Interviews and surveys revealed that 78% of stakeholders considered the dashboard’s real-time charts and alert system crucial for their operations. Government regulators indicated that the EMIS provided a credible data trail for enforcement actions, while industry participants appreciated transparent benchmarks for environmental compliance. Community representatives valued the open-access summary of coastal health indicators, reporting increased trust in monitoring efforts.

Performance Metrics

1. **Data Capture Rate**

Achieved a 96% success rate in retrieving sensor data.

2. **Alert Accuracy**

Over 90% of triggered alerts were validated as genuine environmental events by subsequent lab analyses or field inspections.

3. **Response Time**

Average delay from sensor reading to dashboard update was under 5 minutes, enabling near-real-time situational awareness.

Technical Challenges

- **Sensor Biofouling**: Algae and sediments required regular maintenance to keep sensor readings accurate.
- **Power Reliability**: Solar charging capacity diminished during winter months, necessitating backup batteries or external power.
- **Network Coverage**: Some remote coastal stretches had limited cellular connectivity, leading to reliance on satellite modems with higher latency and costs.

Technical Feasibility

The pilot results affirm that a multi-layered EMIS can operate effectively in the demanding climate of the northern Caspian Sea. Modular design facilitated swift revisions to hardware or software components. Furthermore, adopting widely supported protocols (MQTT, LoRaWAN) allowed smooth integration with third-party sensors and data systems [8].

Environmental Impact

Rapid anomaly detection proved instrumental in mitigating pollution events, underscoring the practical value of real-time monitoring. By forecasting potential ecological disturbances—such as algal blooms or oil spills—the system also supports preventive measures that preserve biodiversity and maintain water quality for local fisheries.

Policy and Governance

The EMIS provides objective, data-driven insights that can guide policymaking in areas like industrial emissions regulation or marine conservation. Transparent, publicly accessible data channels can enhance accountability, ensuring that industry operators meet environmental standards. Collaboration among academic, governmental, and private sectors is crucial to institutionalize this approach.

Scalability and Replicability

While tailored to the northern Caspian context, the architecture can be extended to other coastal or inland water bodies facing similar ecological pressures. The cloud-based data storage and machine learning pipelines can be reconfigured to incorporate diverse sensor arrays and region-specific parameters. Nevertheless, any large-scale rollout requires robust funding, stakeholder cooperation, and a clear legal framework.

Limitations

- **Technical Infrastructure**: The EMIS depends on stable power sources and communication networks, which may be absent in more isolated regions.
- **Cost Constraints**: High-quality sensors and cloud resources can be prohibitively expensive for smaller municipalities or developing areas.
- **Human Capital**: Sustaining the system demands ongoing training for local technicians, data scientists, and administrators.

Conclusion. This article demonstrates the design, deployment, and validation of a comprehensive Environmental Monitoring Information System (EMIS) for the northern coast of the Caspian Sea. By unifying sensor networks, remote sensing imagery, and predictive analytics under a modular architecture, the system delivers real-time insights into water quality, hydrocarbon pollution, and ecological integrity. Pilot results confirm that the EMIS not only detects anomalies with high accuracy but also facilitates rapid stakeholder response, thereby mitigating adverse environmental impacts.

The system's success depends on several critical factors—reliable infrastructure, robust data fusion algorithms, and strong stakeholder engagement. While the northern Caspian served as an initial testbed, the core methodology can be adapted to other marine or freshwater environments. Long-term sustainability will require supportive policies, dedicated funding, and continuous technological updates. Future enhancements include expanding sensor arrays (e.g., heavy metals, sediment load), employing advanced deep learning models for event forecasting, and incorporating more user-friendly dashboards that address the linguistic and cultural needs of local communities. Ultimately, a well-designed EMIS can become a cornerstone of ecological preservation, resource management, and sustainable development in vulnerable coastal regions.

REFERENCES

1. Gao, J., & Ma, R. (2018). Marine Monitoring Systems: A Review of Sensor Networks and Remote Sensing Applications. *Ocean Engineering*, 157, 25–38.

2. Mansour, R., Singh, A., & Chawla, G. (2019). Utilizing Satellite Imagery for Coastal Water Quality Assessment. *Remote Sensing of Environment*, 233, 111–118.

3. Park, S., Chae, K., & Jung, J. (2019). Low-Power Wide-Area Networks for Environmental Monitoring: A Case Study in Coastal Regions. *Sensors*, 19(5), 1234.

4. Zhan, Y., Li, X., & Chen, X. (2020). Detection of Oil Spills Using SAR Data: A Comparative Study. *Marine Pollution Bulletin*, 153, 110–123.
5. Liu, Q., Wang, S., & Zhao, Z. (2021). Machine Learning for Marine Ecosystems: Trends and Future Prospects. *Ecological Informatics*, 61, 101–109.
6. Brown, A., Chen, T., & Zhang, M. (2022). Data-Driven Approaches to Water Quality Prediction: Integrating Remote Sensing and IoT. *Environmental Modelling & Software*, 144, 105–123.
7. Ivanov, S., Petrova, O., & Kravtsov, V. (2021). Hybrid Kalman Filters in Multi-Sensor Marine Environments. *IEEE Access*, 9, 67856–67866.
8. Smith, L., & Rivera, D. (2023). Cloud Computing for Marine Data Integration: Approaches and Implementations. *Journal of Ocean Technology*, 48(2), 29–45.

ЛИТЕРАТУРА

1. Gao, J., & Ma, R. (2018). Marine Monitoring Systems: A Review of Sensor Networks and Remote Sensing Applications. *Ocean Engineering*, 157, 25–38.
2. Mansour, R., Singh, A., & Chawla, G. (2019). Utilizing Satellite Imagery for Coastal Water Quality Assessment. *Remote Sensing of Environment*, 233, 111–118.
3. Park, S., Chae, K., & Jung, J. (2019). Low-Power Wide-Area Networks for Environmental Monitoring: A Case Study in Coastal Regions. *Sensors*, 19(5), 1234.
4. Zhan, Y., Li, X., & Chen, X. (2020). Detection of Oil Spills Using SAR Data: A Comparative Study. *Marine Pollution Bulletin*, 153, 110–123.
5. Liu, Q., Wang, S., & Zhao, Z. (2021). Machine Learning for Marine Ecosystems: Trends and Future Prospects. *Ecological Informatics*, 61, 101–109.
6. Brown, A., Chen, T., & Zhang, M. (2022). Data-Driven Approaches to Water Quality Prediction: Integrating Remote Sensing and IoT. *Environmental Modelling & Software*, 144, 105–123.
7. Ivanov, S., Petrova, O., & Kravtsov, V. (2021). Hybrid Kalman Filters in Multi-Sensor Marine Environments. *IEEE Access*, 9, 67856–67866.
8. Smith, L., & Rivera, D. (2023). Cloud Computing for Marine Data Integration: Approaches and Implementations. *Journal of Ocean Technology*, 48(2), 29–45.

ИНТЕГРИРОВАННАЯ ИНФОРМАЦИОННАЯ СИСТЕМА МОНИТОРИНГА ОКРУЖАЮЩЕЙ СРЕДЫ КАСПИЙСКОГО МОРЯ: КОМПЛЕКСНЫЙ ПОДХОД

***Кенжебаева Ж. Е., Казиева Г.Д.**
Университет Есенова, Актау Казахстан
e-mail:zhanat.kenzhebayeva@yu.edu.kz, gulnara.kaziyeva@yu.edu.kz

Аннотация. В этой статье предлагается целостная информационная система мониторинга окружающей среды (EMIS), предназначенная для оценки и управления экологическими условиями вдоль северного побережья Каспийского моря. Благодаря интеграции нескольких источников данных — сенсорных сетей, дистанционного зондирования, метеорологических данных и алгоритмов машинного обучения — EMIS обеспечивает получение информации в режиме реального времени о параметрах качества воды, загрязнении углеводородами и стабильности среды обитания. Масштабируемая модульная архитектура гарантирует, что разнообразные наборы данных синтезируются в единую, полезную информацию, что позволяет заинтересованным сторонам обнаруживать аномалии, прогнозировать тенденции загрязнения и определять потенциальные экологические риски. Строгие процессы проверки данных смягчают несоответствия, возникающие из-за дрейфа датчиков или сбоев передачи, тем самым повышая надежность. Система также имеет облачную инфраструктуру и ролевые пользовательские интерфейсы, адаптированные к местным требованиям, что способствует совместному принятию решений между государственными учреждениями, промышленными операторами и местными сообществами. В этом исследовании описывается технический план EMIS, документируются результаты пилотных испытаний и подчеркиваются последствия для устойчивого использования ресурсов и разработки региональной экологической политики.

Ключевые слова: Информационная система мониторинга окружающей среды (EMIS), Северный Каспий, Сенсорные сети, Дистанционное зондирование, Объединение данных, Прогностическое моделирование, Устойчивое управление ресурсами, Качество воды и загрязнение углеводородами, Аналитика в реальном времени, Защита прибрежных местообитаний.

КАСПИЙ ТЕҢІЗІНІҢ ЭКОЛОГИЯЛЫҚ МОНИТОРИНГІНІҢ ИНТЕГРАЦИЯЛАНҒАН АҚПАРАТТЫҚ ЖҮЙЕСІ: КЕШЕНДІ ТӘСІЛ

***Кенжебаева Ж. Е., Казиева Г. Д.**
Есенов университеті, Ақтау, Қазахстан
e-mail:zhanat.kenzhebayeva@yu.edu.kz, gulnara.kaziyeva@yu.edu.kz

Аңдатпа. Бұл мақалада Каспий Теңізінің солтүстік жағалауындағы экологиялық жағдайларды бағалауға және басқаруға арналған экологиялық Мониторингтің біртұтас Ақпараттық Жүйесі (ЭМАЖ) ұсынылған. Көптеген деректер көздерін—сенсорлық желілерді, қашықтықтан зондтауды, метеорологиялық деректерді және машиналық оқыту алгоритмдерін біріктіру арқылы EMI нақты уақыт режимінде су сапасының параметрлері, көмірсутектердің ластануы және тіршілік ету ортасының тұрақтылығы туралы түсінік береді. Масштабталатын модульдік архитектура мүдделі тараптарға ауытқуларды анықтауға, ластану тенденцияларын болжауға және ықтимал экологиялық қауіптерді анықтауға мүмкіндік беретін әртүрлі деректер жиынының бірыңғай, іске асырылатын ақпаратқа синтезделуін қамтамасыз етеді. Деректерді тексерудің қатаң процестері сенсорлардың жылжуынан немесе беріліс ақауларынан туындайтын сәйкессіздіктерді азайтады, осылайша сенімділікті арттырады. Жүйе сонымен қатар бұлтқа негізделген инфрақұрылыммен және жергілікті талаптарға бейімделген рөлдік пайдаланушы интерфейстерімен жабдықталған, бұл мемлекеттік органдарда, салалық операторларда және жергілікті қауымдастықтарда бірлескен шешім қабылдауға ықпал етеді. Бұл зерттеу EMI техникалық жоспарын анықтайды, пилоттық сынақтардың нәтижелерін құжаттайды және ресурстарды тұрақты пайдалану мен аймақтық экологиялық саясатты дамытудың салдарын көрсетеді.

Түйін сөздер: қоршаған ортаны бақылаудың ақпараттық жүйесі (EMI), Солтүстік Каспий теңізі, сенсорлық желілер, қашықтықтан зондтау, деректерді біріктіру, болжамды модельдеу, ресурстарды тұрақты басқару, судың сапасы мен көмірсутектердің ластануы, нақты уақыттағы аналитика, жағалаудағы тіршілік ету ортасын қорғау.