Applied Coastal and Ocean Engineering

Volume 1 No. 1
June 2026

E-ISSN XXXX-XXXX (Online)
https://acoe.parpi.id/

Automated Hydro-Oceanographic Data Pipeline for Marine Spatial Decision Support (Prototype)

Muhammad Fadly

Makassar Marine Spatial Planning Centre, Makassar South of Sulawesi, Indonesia.
muhammad.fadly@kkp.go.id

Abstract

The implementation of risk-based marine spatial licensing demands accurate hydro-oceanographic data to mitigate structural and ecological hazards. However, reliance on costly in-situ measurements or computationally heavy global models—which are often plagued by shallow-water anomalies and complex binary formats—poses significant challenges for local regulatory bodies. This study presents the development of a Hybrid Intelligent Decision Support System (HIDSS), a Python-based automated spatial data pipeline designed to bridge this technical gap. The backend architecture employs automated spatial clipping to eliminate land mask boundary noise and utilizes robust vectorized operations for high-speed, 20-year tidal hindcasting using eight major harmonic constituents. Wave climates, current circulation, and dynamic bathymetric profiles are instantaneously extracted and systematically compressed into a lightweight, Parquet-formatted Data Lake. Crucially, this highly structured storage seamlessly integrates with Large Language Models (LLMs) via API and Fuzzy Logic protocols, automatically translating massive numerical arrays into comprehensive descriptive analyses and quantitative risk indices. The developed the Prototype HIDSS platform successfully empowers evaluators to conduct instantaneous, evidence-based technical audits, fundamentally optimizing the efficiency, transparency, and scientific rigor of the marine spatial planning ecosystem.

Keywords— *Decision Support System, Hydrodynamics, Large Language Models, Marine Spatial Planning, Python*

I. INTRODUCTION

The governance of coastal zones and small islands faces multidimensional challenges due to the intensifying use of marine spaces for human activities (Halpern et al., 2012; Ehler and Douvere, 2009). In archipelagic nations, spatial planning instruments—such as the Marine Spatial Suitability Permit (KKPRL) in Indonesia—are absolute requirement for ensuring ecological sustainability and mitigating spatial conflicts (Kementerian Kelautan dan Perikanan RI, 2021). The paradigm shift towards risk-based licensing requires policymakers and technical evaluators to conduct precise assessments using Decision Support Tools (DSTs) grounded in physical oceanographic dynamics, including bathymetry, tides, waves, and currents (Pınarbaşı et al., 2017). In other hand, the implementation of risk-based licensing is frequently hindered by the lack of in-situ hydro-oceanographic data. Field measurements, such as bathymetric sounding or the deployment of Acoustic Doppler Current Profilers (ADCP) and Wave Recorder, provide massive operational costs and lengthy observation periods, making them impractical for evaluating small to medium-scale activities (Rayner et al., 2019). Consequently, leveraging secondary data derived from global models, such as the Copernicus Marine Environment Monitoring Service (CMEMS) for physical circulation and Global Tide Mo dels, has become the industry standard (Egbert and Erofeeva, 2002; Le Traon et al., 2019).

Models increasingly offer high spatial and temporal resolutions, largely driven by satellite altimetry data assimilation. Despite offering massive data availability, a significant technical gap remains when utilizing global models for local-level decision support systems. First, there is a high degree of hydrodynamic uncertainty in shallow waters and narrow straits. Global models typically struggle to accurately represent tidal energy attenuation in complex coastal zones, a problem exacerbated by land mask boundary noise (Stammer et al., 2014). Second, the simplified calculation of tidal ranges often ignores the significance of diurnal components (K_1, O_1) or shallow-water constituents, leading to deviations in extreme datum elevations, such as the Highest and Lowest Astronomical Tides (Pugh and Woodworth, 2014). Third, more data does not guarantee the availability of the specific information required to address key scientific inquiries or to formulate evidence-based policies for sustainable ocean management. To maximize the utility of marine big data, these datasets must be openly accessible, interoperable, and synthesized through advanced multidisciplinary analytics, potentially leveraging artificial intelligence (Guidi et al., 2020).

This study aims to bridge these gaps by developing a Hybrid Intelligent Decision Support System (HIDSS) architecture based on Python. The proposed framework performs automated downscaling of global model data through three key innovations: (1) automated spatial clipping and coordinate interpolation using geospatial libraries to eliminate land boundary anomalies with high precision; (2) brute-force computation for long-term tidal hindcasting using eight major harmonic constituents to reduce coastal hydrodynamic uncertainty; and (3) the transformation of massive datasets into a lightweight Parquet-formatted Data Lake architecture. By automating this spatial data pipeline, the complexity of hydro-oceanographic and bathymetric data can be streamlined into an interactive visual interface, thereby supporting an evidence-based policy approach to marine spatial audits.

II. RESEARCH METHOD

This study adopts a geospatial software engineering approach to construct an automated data pipeline. The framework addresses hydrodynamic uncertainty and data complexity in marine spatial licensing through three primary phases: data acquisition, spatial calibration, and analytical mathematical modeling.

A. Global Data Acquisition

The system integrates three primary secondary datasets representing crucial physical oceanographic parameters: High-resolution national bathymetry data (BATNAS) at a 6 arc-second spatial resolution, provided by the Geospatial Information Agency of Indonesia (Badan Informasi Geospasial, 2018).Tide data extracted from the Global Tide Model, comprising eight major harmonic constituents: M_2, S_2, N_2, K_2, K_1, O_1, P_1,and Q_1 (Egbert and Erofeeva, 2002). Time-series data of wave energy spectra and physical current circulation from the Copernicus Marine Environment Monitoring Service (CMEMS) (Le Traon et al., 2019).

B. Spatial Calibration and Filtering

Raw numerical data from global models typically lack absolute spatial references. Therefore, the matrix indices are first translated into the WGS 1984 absolute coordinate system (EPSG:4326). Following this, an automated spatial clipping algorithm is executed using the *GeoPandas* library. The system performs a *Spatial Join* between the model's grid points and high-resolution coastal boundary shapefiles. Grid points that intersect or overlap

History of article:
Received: 5 June 2026, Revised: 5 June 2026, Published: 5 June 2026

with the land polygon are deterministically dropped. This filtering mechanism is critical for eliminating land mask anomalies that frequently distort hydrodynamic representations in shallow waters (Stammer *et al.*, 2014).

C. Hydro-Oceanographic Mathematical Modeling

C.1 Wave Analysis.

The significant wave height (H_s) is represented by the H_{m0} estimate, which is derived from the zero-order spectral moment (m_0) of the wave energy variance spectrum (Holthuijsen, 2007). The zero-order moment is formulated as the total integration of spectral energy:

$$m_0 = \int_0^{\infty} \int_0^{2\pi} E\,(f,\theta) df d\theta \tag{1}$$

Assuming the sea surface elevation follows a Rayleigh distribution under narrow-banded spectrum conditions, H_s is calculated using the following equation:

$$H_s \approx 4\sqrt{m_0} \tag{2}$$

Through vectorized *NumPy* array operations, temporal aggregate values are extracted to obtain the Maximum Significant Wave Height (H_{smax}) for extreme design parameters and the Mean Operational Wave Height (H_{smean}).

C.2 Current Circulation Analysis:

Hydrodynamic circulation velocities are extracted from their orthogonal zonal (u) and meridional (v) components. The magnitude of the current velocity (V) is computed using the Euclidean formulation (Pond and Pickard, 2013):

$$V = \sqrt{u^2 + v^2} \tag{3}$$

The current direction (θ_c) is converted into a meteorological azimuth using the four-quadrant arctangent function:

$$\theta_c = \left(270 - \frac{180}{\pi} \arctan 2\,(v, u)\right) \bmod 360 \tag{4}$$

C.3 Tidal Hindcasting Analysis

A brute-force computational approach is utilized to synthesize tidal predictions over 175,320 hours (approximately 20 years) to fully encompass the lunar nodal cycle. The instantaneous water elevation Z(t) is formulated as (Pugh and Woodworth, 2014).

$$Z(t) = Z_0 + \sum_{i=1}^{8} A_i \cos(\omega_i t - \phi_i) \tag{5}$$

From this long-term time series, extreme Admiralty datums—namely the Highest Astronomical Tide (HAT), Mean Sea Level (MSL), and Lowest Astronomical Tide (LAT)—along with the Formzahl number, are accurately extracted.

C.4 Bathymetric Profiling:

To evaluate seabed topography, dynamic bathymetric profiling optimizes the *k-d tree* data structure from the *SciPy* library for high-speed Nearest Neighbor searches (Virtanen *et al.*, 2020). This algorithm precisely projects seabed elevations along user-digitized transect lines on the map.

D. Data Lake Architecture and Prototype Dashboard Integration

To handle massive spatial computations without overwhelming the client side, all model outputs are partitioned by provincial boundaries and compressed into the *Apache Parquet* format. This columnar storage architecture drastically minimizes file sizes and accelerates I/O operations (Vohra, 2016). The frontend interface is integrated using the *Streamlit* framework and the *Folium* spatial library. This setup allows for the instantaneous rendering of tidal cycles, Wave Roses, and bathymetric profiles without causing high latency in the browser.

E. System Flowchart

The integration protocol, from secondary data acquisition to the final user interface visualization, is systematically represented in Figure 1.

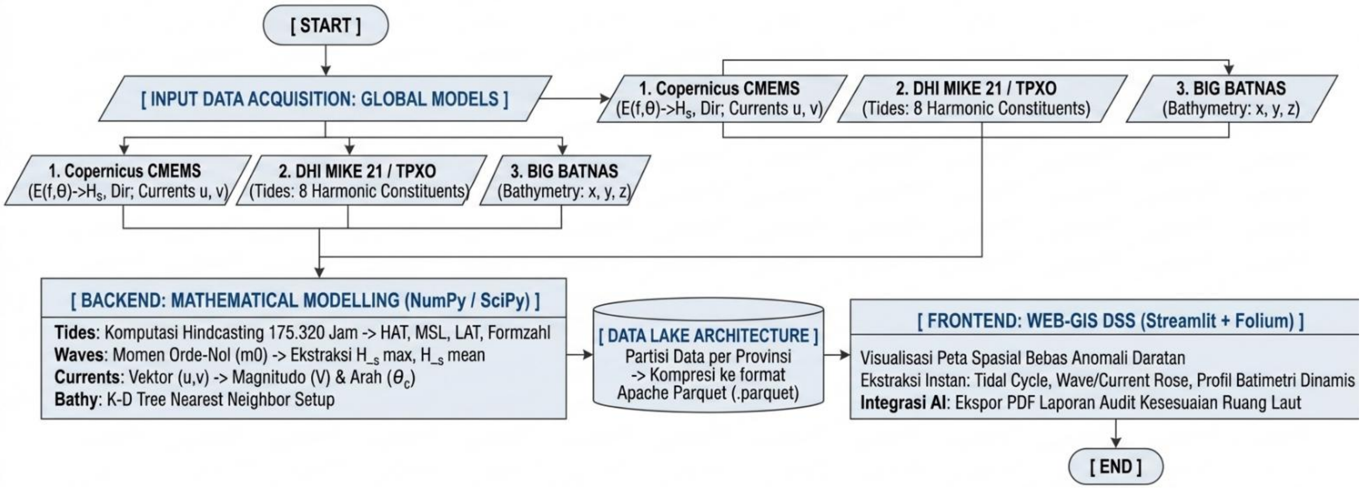


Figure 1. Flowchart

Source : Analysis results (2026)

III. RESULTS AND DISCUSSION

The implementation of the Python-based framework described in the methodology has yielded an operational platform based Decision Support System (DSS) for the Central Indonesia region. This section presents the performance of the automated data pipeline, the validation of the spatial data cleaning process, an in-depth analysis of each hydro-oceanographic and bathymetric parameter, and a discussion regarding the implications of this system for the marine spatial licensing process.

History of article:
Received: 5 June 2026, Revised: 5 June 2026, Published: 5 June 2026

HIDSS WORKSTATION

Hybrid Intelligent Decision Support System

Developed by : Ir. Muhammad Fadly, S.T., M.T

Marine Spatial Planning and Licensing • BPRL Makassar

Figure 1 HIDSS Workstation Splash Screen

Source : Analysis results (2026)

A. Automated Data Pipeline Performance (Backend)

The system successfully integrates and processes massive datasets (big data) from various global sources with high computational efficiency. Vectorized matrix operations using the NumPy library enable the synthesis of 175,320 hours (~20 years) of tidal time-series data for thousands of grid points in just a few minutes on standard hardware.

Transforming the data into a columnar storage architecture using the Apache Parquet format has proven to reduce storage size by up to 80% compared to conventional CSV formats, while simultaneously accelerating query times (I/O operations) on the frontend. This optimization is crucial to ensure the platform interface remains lightweight and highly responsive (low latency) whenever a user clicks on any coordinate point on the map.

B. Spatial Clipping Validation

A primary novelty of this research is the GeoPandas-based automated spatial clipping algorithm designed to resolve hydrodynamic uncertainty at land mask boundaries. Figure 2 illustrates a sharp visual comparison between the raw data distribution from the global model and the system's spatial filtering results.

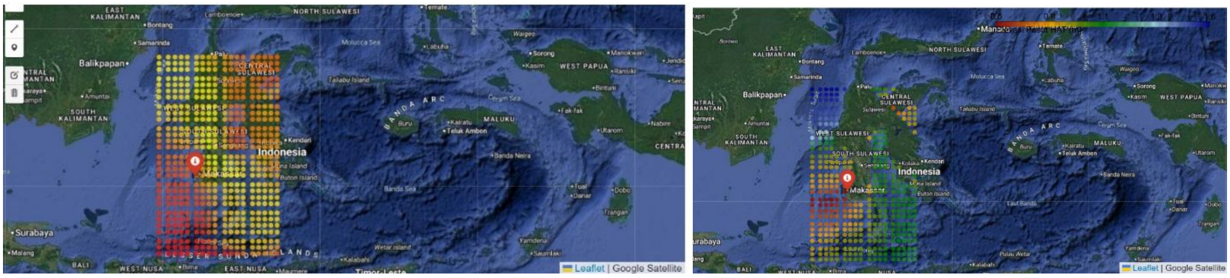


Figure 2 comparison between the raw data distribution from the global model tide and the system's spatial filtering results

Source : Analysis results (2026)

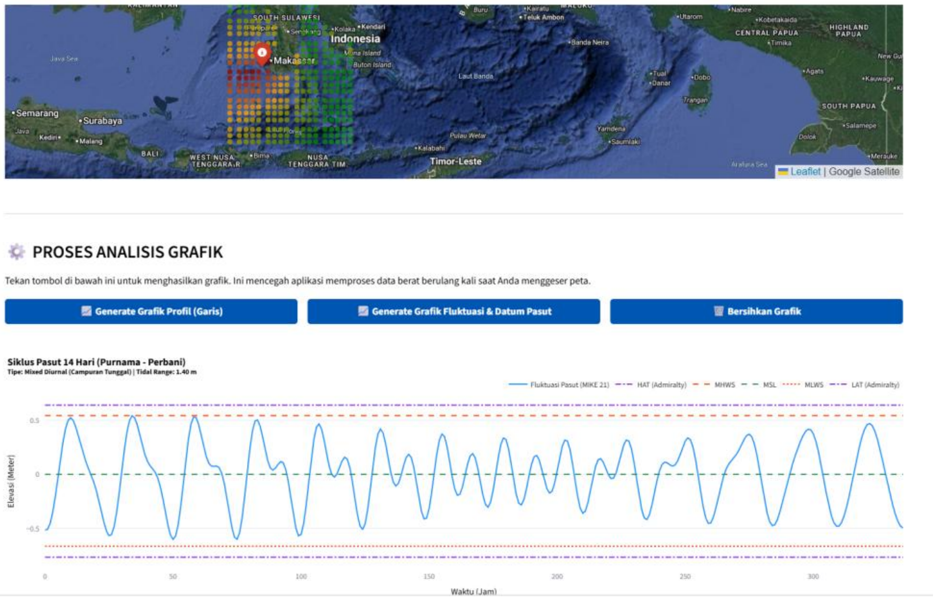
As seen in Figure 2, the algorithm successfully identifies and drops all global model grid points that intersect with the high-resolution coastline polygon, particularly in morphologically complex areas such as the narrow straits and small islands of Central Indonesia. This visual validation confirms that the parameter values displayed on the HIDSS dashboard are pure representations of marine conditions, effectively reducing the risk of analytical bias caused by land mask noise during the technical risk assessment of coastal infrastructure.

C. Interactive Hydro-Oceanographic and Bathymetric Analysis

This section discusses the analytical outputs for each parameter, which are generated in real-time through user interactions on the platform interface.

1) Long-Term Tidal Hindcasting Analysis:

Brute-force computation over a 20-year span successfully extracted extreme Admiralty datums with high precision. Figure 3 displays the 14-day (spring-neap) tidal fluctuation graph complete with extreme datum reference lines (HAT and LAT).



Tekan tombol di bawah ini untuk menghasilkan grafik. Ini mencegah aplikasi memproses data berat berulang kali saat Anda menggeser peta.

Figure 3 displays the 14-day (spring-neap) tidal fluctuation graph

Source : Analysis results (2026)

Unlike the simplified approaches often utilized in bureaucratic practices, this system leverages 8 major harmonic constituents. The results in Figure 3 demonstrate its capability to capture the actual tidal range, including energy variations between spring and neap tides. The instantaneous extraction of HAT and LAT datums provides highly accurate reference values for licensing evaluators to assess the elevation feasibility of proposed marine structures (e.g., docks or land reclamation) against extreme inundation risks, thereby minimizing hydrodynamic uncertainty in coastal zones.

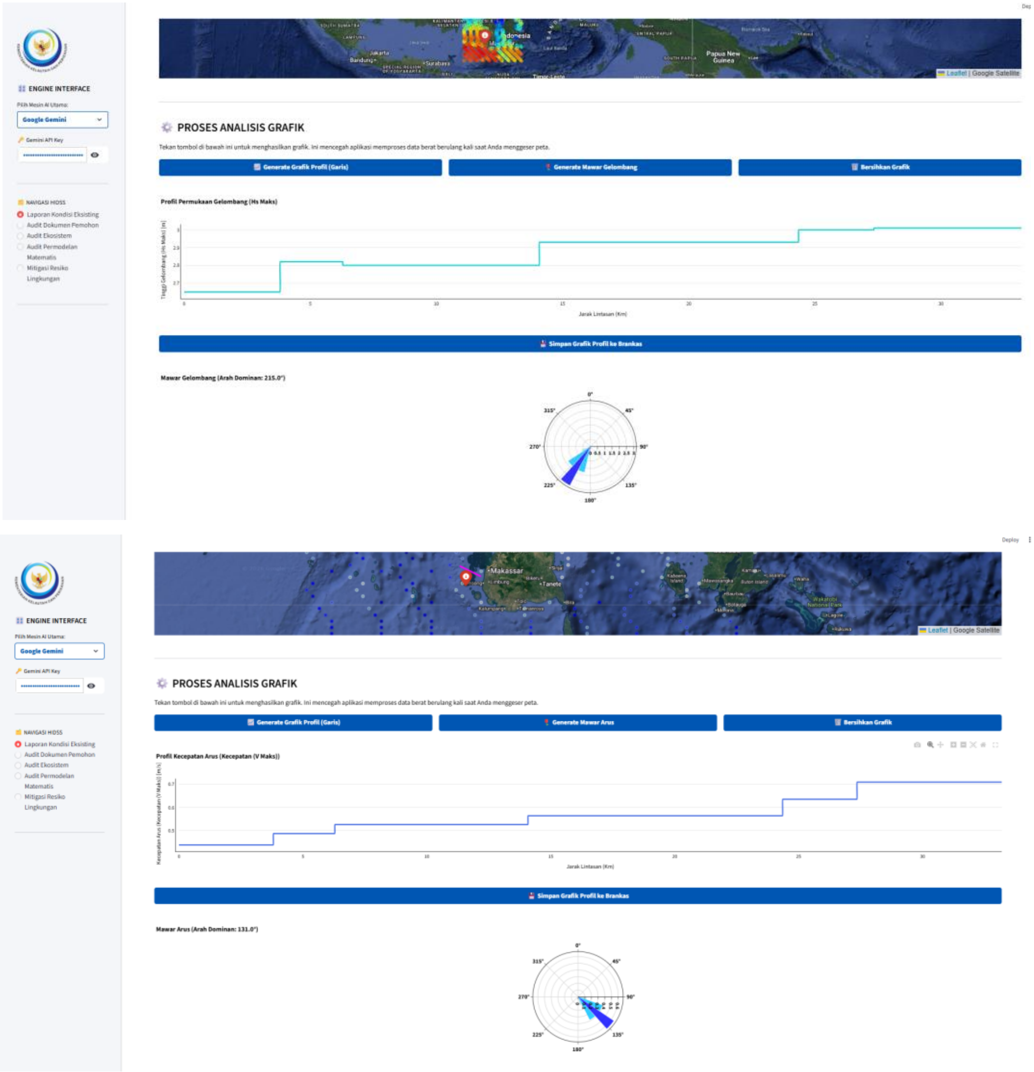
2) Wave Climate and Current Circulation Analysis:

The extraction of wave and current parameters focuses on presenting extreme and operational statistics through dynamic visualizations. Figure 4 displays the interactive Wave Rose visualization and related parameter statistics on the system interface.

Through the Wave Rose visualization in Figure 4, users can immediately identify the dominant incoming wave direction and the most frequently occurring wave height ranges. The instant availability of historical Maximum Significant Wave Height (H_{smax}) and Maximum Current Velocity (V_{max}) values is highly valuable for initial risk

History of article:
Received: 5 June 2026, Revised: 5 June 2026, Published: 5 June 2026

assessments regarding potential hydrodynamic forces acting on marine structures, as well as identifying potential erosion or sedimentation risks around the project site.



PROSES ANALISIS GRAFIK

Tekan tombol di bawah ini untuk menghasilkan grafik. Ini mencegah aplikasi memproses data berat berulang kali saat Anda menggeser peta.

Profil Permukaan Gelombang (Hs Maks)

Mawar Gelombang (Arah Dominan: 215.0°)

PROSES ANALISIS GRAFIK

Tekan tombol di bawah ini untuk menghasilkan grafik. Ini mencegah aplikasi memproses data berat berulang kali saat Anda menggeser peta.

Profil Kecepatan Arus (Kecepatan (V Maks))

Mawar Arus (Arah Dominan: 131.0°)

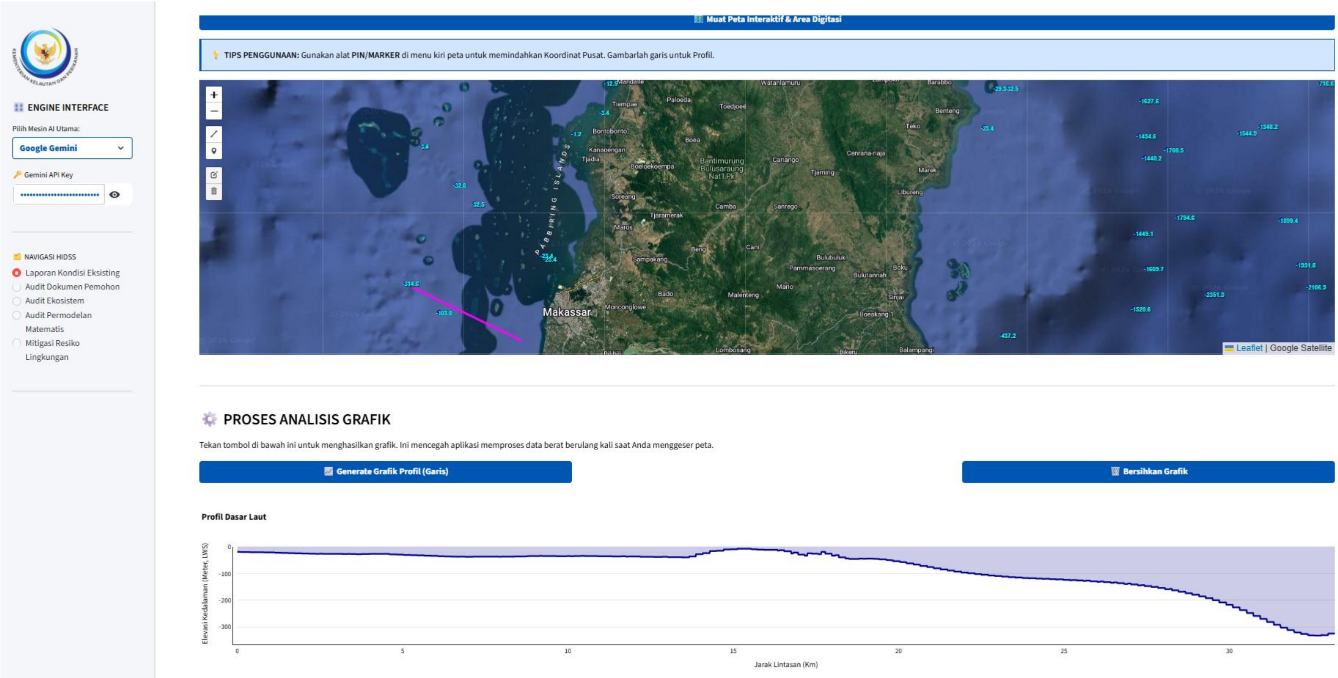
Figure 4 Wave Climate and Current Circulation Analysis

Source : Analysis results (2026)

3) Dynamic Bathymetric Profiling:

The depth analysis optimizes the *k-d tree* algorithm from SciPy for high-speed nearest neighbor searches. Figure 5 demonstrates the dynamic bathymetric profiling feature.

This feature revolutionizes initial technical feasibility assessments. As shown in Figure 5, users can draw a proposed infrastructure trajectory (e.g., submarine cables or outfall pipelines) directly on the map, and the system instantly generates the seabed topography profile along that line. The availability of this rapid bathymetric profile is crucial for identifying potential topographical hazards, extreme slopes, or areas requiring additional protection, eliminating the need to wait for expensive field bathymetric surveys during the pre-feasibility stage.



TIPS PENGGUNAAN: Gunakan alat PIN/MARKER di menu kiri peta untuk memindahkan Koordinat Pusat. Gambarlah garis untuk Profil.

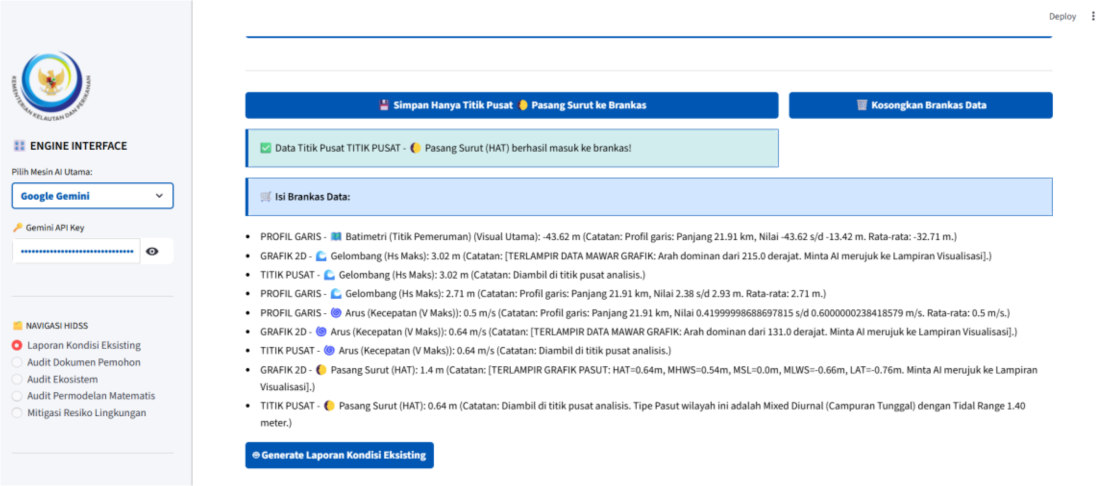
Tekan tombol di bawah ini untuk menghasilkan grafik. Ini mencegah aplikasi memproses data berat berulang kali saat Anda menggeser peta.

Figure 5 Dynamic Bathymetric Profiling

Source : Analysis results (2026)

D. Large Language Model Integration for Descriptive Analysis

The robust "storage" mechanism provided by the Parquet-formatted Data Lake architecture developed in this study (Module 1) is fully primed for future integration with Artificial Intelligence (AI). By utilizing this highly compressed and structured columnar storage, the system can seamlessly feed extracted hydro-oceanographic parameters into advanced Large Language Models (LLMs) via API integration, specifically utilizing Gemini and DeepSeek. The structured storage acts as the critical bridge, translating massive arrays of spatial data into context-rich prompts. This capability allows the system to automatically generate descriptive analyses, converting raw numerical data—such as extreme Admiralty datums, wave heights, and current velocities—into readable, comprehensive narratives tailored for bureaucratic evaluators who are often not oceanographic experts.



✅ Data Titik Pusat TITIK PUSAT - Pasang Surut (HAT) berhasil masuk ke brankas!

Isi Brankas Data:

- PROFIL GARIS - Batimetri (Titik Pemeruman) (Visual Utama): -43.62 m (Catatan: Profil garis: Panjang 21.91 km, Nilai -43.62 s/d -13.42 m. Rata-rata: -32.71 m.)
- GRAFIK 2D - Gelombang (Hs Maks): 3.02 m (Catatan: [TERLAMPIR DATA MAWAR GRAFIK: Arah dominan dari 215.0 derajat. Minta AI merujuk ke Lampiran Visualisasi].)
- TITIK PUSAT - Gelombang (Hs Maks): 3.02 m (Catatan: Diambil di titik pusat analisis.)
- PROFIL GARIS - Gelombang (Hs Maks): 2.71 m (Catatan: Profil garis: Panjang 21.91 km, Nilai 2.38 s/d 2.93 m. Rata-rata: 2.71 m.)
- PROFIL GARIS - Arus (Kecepatan (V Maks)): 0.5 m/s (Catatan: Profil garis: Panjang 21.91 km, Nilai 0.41999998688697815 s/d 0.6000000238418579 m/s. Rata-rata: 0.5 m/s.)
- GRAFIK 2D - Arus (Kecepatan (V Maks)): 0.64 m/s (Catatan: [TERLAMPIR DATA MAWAR GRAFIK: Arah dominan dari 131.0 derajat. Minta AI merujuk ke Lampiran Visualisasi].)
- TITIK PUSAT - Arus (Kecepatan (V Maks)): 0.64 m/s (Catatan: Diambil di titik pusat analisis.)
- GRAFIK 2D - Pasang Surut (HAT): 1.4 m (Catatan: [TERLAMPIR GRAFIK PASUT: HAT=0.64m, MHWS=0.54m, MSL=0.0m, MLWS=-0.66m, LAT=-0.76m. Minta AI merujuk ke Lampiran Visualisasi].)
- TITIK PUSAT - Pasang Surut (HAT): 0.64 m (Catatan: Diambil di titik pusat analisis. Tipe Pasut wilayah ini adalah Mixed Diurnal (Campuran Tunggal) dengan Tidal Range 1.40 meter.)

Figure 6 Storage Data

History of article:
Received: 5 June 2026, Revised: 5 June 2026, Published: 5 June 2026

Source : Analysis results (2026)

D.1 Core Prompt for LLM (Gemini / DeepSeek API):

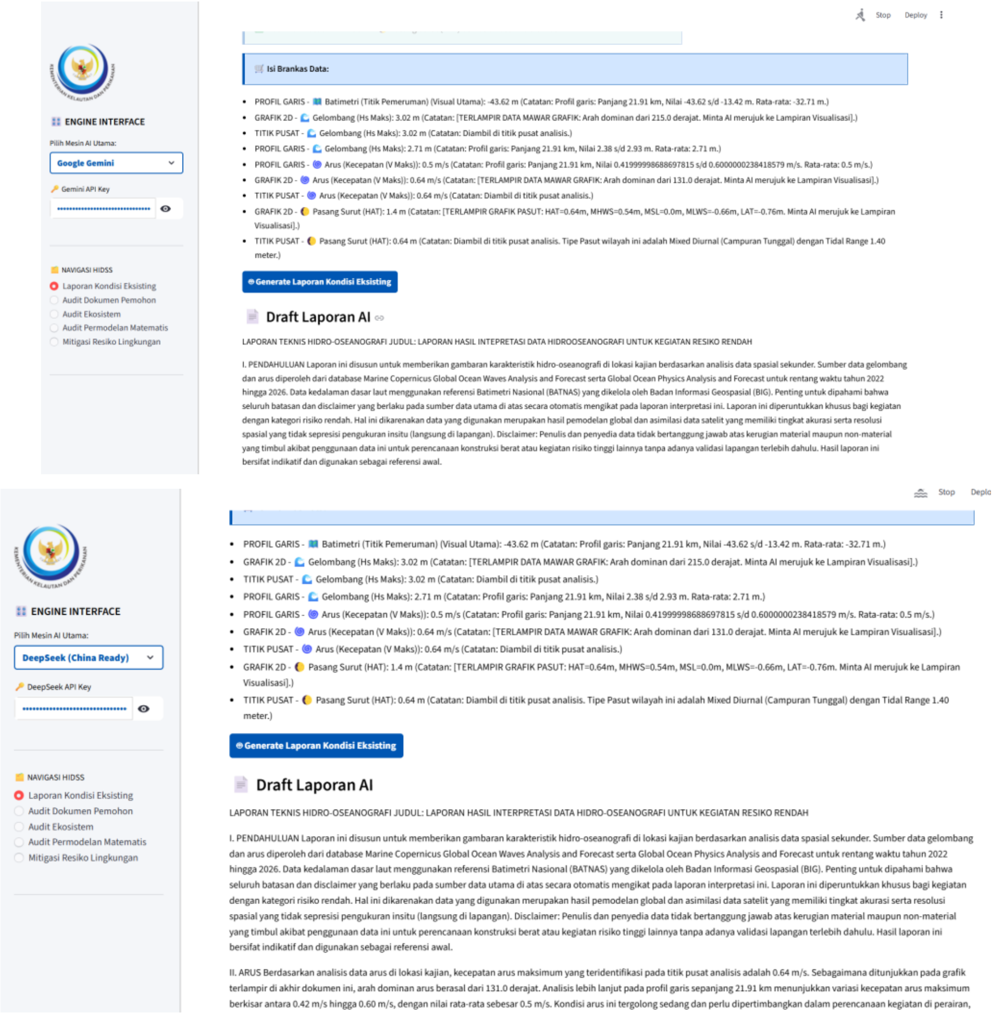
To effectively harness the analytical capabilities of these Large Language Models, the HIDSS platform employs a structured prompt engineering strategy. Instead of merely feeding raw numerical datasets to the API, the system programmatically constructs a contextual instruction set. This automated prompt assigns a specific expert persona to the AI and dynamically injects the aggregated statistical values—retrieved directly from the localized Parquet storage—into predefined variables. By standardizing this query formulation, the system ensures that the generated descriptive analyses remain objective, scientifically rigorous, and strictly aligned with the regulatory evaluation criteria of marine spatial planning. The foundational architecture of this automated query is established as follows:

"You are an expert oceanographer and technical evaluator integrated into the HIDSS Marine Spatial Planning system. Based on the following hydro-oceanographic data extracted from the system's Parquet storage for a proposed coastal infrastructure site:

- Highest Astronomical Tide (HAT): [Insert Value]
- Lowest Astronomical Tide (LAT): [Insert Value]
- Maximum Significant Wave Height (H_{smax}) : [Insert Value]
- Maximum Current Velocity (V_{max}): [Insert Value]
- Bathymetric Slope: [Insert Value]

Generate a comprehensive descriptive analysis of the marine environmental conditions. Assess the potential technical risks for structural development (e.g., docks, reclamation) and provide a clear, evidence-based recommendation to support risk-based licensing decisions."

Based on the spatial data retrieved from the system's centralized storage, the targeted coastal site exhibits a significant tidal range, with a pronounced discrepancy between the Highest Astronomical Tide (HAT) and the Lowest Astronomical Tide (LAT). This indicates a high exposure to extreme inundation risks during spring tide cycles. The historical Maximum Significant Wave Height (H_{smax}) and Maximum Current Velocity (V_{max}) extracted from the Wave Rose and Current Circulation analysis indicate moderate-to-high hydrodynamic stress, which could accelerate coastal erosion or structural fatigue over time. Furthermore, the dynamic bathymetric profile reveals a steep seabed gradient near the coastal boundary, increasing the complexity of foundation installations.



Pilih Mesin AI Utama: Google Gemini

Gemini API Key: ••••••••••••••••••••••••••••••

Isi Brankas Data:

- PROFIL GARIS - Batimetri (Titik Pemeruman) (Visual Utama): -43.62 m (Catatan: Profil garis: Panjang 21.91 km, Nilai -43.62 s/d -13.42 m. Rata-rata: -32.71 m.)
- GRAFIK 2D - Gelombang (Hs Maks): 3.02 m (Catatan: [TERLAMPIR DATA MAWAR GRAFIK: Arah dominan dari 215.0 derajat. Minta AI merujuk ke Lampiran Visualisasi].)
- TITIK PUSAT - Gelombang (Hs Maks): 3.02 m (Catatan: Diambil di titik pusat analisis.)
- PROFIL GARIS - Gelombang (Hs Maks): 2.71 m (Catatan: Profil garis: Panjang 21.91 km, Nilai 2.38 s/d 2.93 m. Rata-rata: 2.71 m.)
- PROFIL GARIS - Arus (Kecepatan (V Maks)): 0.5 m/s (Catatan: Profil garis: Panjang 21.91 km, Nilai 0.41999998688697815 s/d 0.6000000238418579 m/s. Rata-rata: 0.5 m/s.)
- GRAFIK 2D - Arus (Kecepatan (V Maks)): 0.64 m/s (Catatan: [TERLAMPIR DATA MAWAR GRAFIK: Arah dominan dari 131.0 derajat. Minta AI merujuk ke Lampiran Visualisasi].)
- TITIK PUSAT - Arus (Kecepatan (V Maks)): 0.64 m/s (Catatan: Diambil di titik pusat analisis.)
- GRAFIK 2D - Pasang Surut (HAT): 1.4 m (Catatan: [TERLAMPIR GRAFIK PASUT: HAT=0.64m, MHWS=0.54m, MSL=0.0m, MLWS=-0.66m, LAT=-0.76m. Minta AI merujuk ke Lampiran Visualisasi].)
- TITIK PUSAT - Pasang Surut (HAT): 0.64 m (Catatan: Diambil di titik pusat analisis. Tipe Pasut wilayah ini adalah Mixed Diurnal (Campuran Tunggal) dengan Tidal Range 1.40 meter.)

Draft Laporan AI

LAPORAN TEKNIS HIDRO-OSEANOGRAFI JUDUL: LAPORAN HASIL INTEPRETASI DATA HIDROOSEANOGRAFI UNTUK KEGIATAN RESIKO RENDAH

I. PENDAHULUAN Laporan ini disusun untuk memberikan gambaran karakteristik hidro-oseanografi di lokasi kajian berdasarkan analisis data spasial sekunder. Sumber data gelombang dan arus diperoleh dari database Marine Copernicus Global Ocean Waves Analysis and Forecast serta Global Ocean Physics Analysis and Forecast untuk rentang waktu tahun 2022 hingga 2026. Data kedalaman dasar laut menggunakan referensi Batimetri Nasional (BATNAS) yang dikelola oleh Badan Informasi Geospasial (BIG). Penting untuk dipahami bahwa seluruh batasan dan disclaimer yang berlaku pada sumber data utama di atas secara otomatis mengikat pada laporan interpretasi ini. Laporan ini diperuntukkan khusus bagi kegiatan dengan kategori risiko rendah. Hal ini dikarenakan data yang digunakan merupakan hasil pemodelan global dan asimilasi data satelit yang memiliki tingkat akurasi serta resolusi spasial yang tidak sepresisi pengukuran insitu (langsung di lapangan). Disclaimer: Penulis dan penyedia data tidak bertanggung jawab atas kerugian material maupun non-material yang timbul akibat penggunaan data ini untuk perencanaan konstruksi berat atau kegiatan risiko tinggi lainnya tanpa adanya validasi lapangan terlebih dahulu. Hasil laporan ini bersifat indikatif dan digunakan sebagai referensi awal.

Pilih Mesin AI Utama: DeepSeek (China Ready)

DeepSeek API Key: ••••••••••••••••••••••••••••••

- PROFIL GARIS - Batimetri (Titik Pemeruman) (Visual Utama): -43.62 m (Catatan: Profil garis: Panjang 21.91 km, Nilai -43.62 s/d -13.42 m. Rata-rata: -32.71 m.)
- GRAFIK 2D - Gelombang (Hs Maks): 3.02 m (Catatan: [TERLAMPIR DATA MAWAR GRAFIK: Arah dominan dari 215.0 derajat. Minta AI merujuk ke Lampiran Visualisasi].)
- TITIK PUSAT - Gelombang (Hs Maks): 3.02 m (Catatan: Diambil di titik pusat analisis.)
- PROFIL GARIS - Gelombang (Hs Maks): 2.71 m (Catatan: Profil garis: Panjang 21.91 km, Nilai 2.38 s/d 2.93 m. Rata-rata: 2.71 m.)
- PROFIL GARIS - Arus (Kecepatan (V Maks)): 0.5 m/s (Catatan: Profil garis: Panjang 21.91 km, Nilai 0.41999998688697815 s/d 0.6000000238418579 m/s. Rata-rata: 0.5 m/s.)
- GRAFIK 2D - Arus (Kecepatan (V Maks)): 0.64 m/s (Catatan: [TERLAMPIR DATA MAWAR GRAFIK: Arah dominan dari 131.0 derajat. Minta AI merujuk ke Lampiran Visualisasi].)
- TITIK PUSAT - Arus (Kecepatan (V Maks)): 0.64 m/s (Catatan: Diambil di titik pusat analisis.)
- GRAFIK 2D - Pasang Surut (HAT): 1.4 m (Catatan: [TERLAMPIR GRAFIK PASUT: HAT=0.64m, MHWS=0.54m, MSL=0.0m, MLWS=-0.66m, LAT=-0.76m. Minta AI merujuk ke Lampiran Visualisasi].)
- TITIK PUSAT - Pasang Surut (HAT): 0.64 m (Catatan: Diambil di titik pusat analisis. Tipe Pasut wilayah ini adalah Mixed Diurnal (Campuran Tunggal) dengan Tidal Range 1.40 meter.)

Draft Laporan AI

LAPORAN TEKNIS HIDRO-OSEANOGRAFI JUDUL: LAPORAN HASIL INTERPRETASI DATA HIDRO-OSEANOGRAFI UNTUK KEGIATAN RESIKO RENDAH

I. PENDAHULUAN Laporan ini disusun untuk memberikan gambaran karakteristik hidro-oseanografi di lokasi kajian berdasarkan analisis data spasial sekunder. Sumber data gelombang dan arus diperoleh dari database Marine Copernicus Global Ocean Waves Analysis and Forecast serta Global Ocean Physics Analysis and Forecast untuk rentang waktu tahun 2022 hingga 2026. Data kedalaman dasar laut menggunakan referensi Batimetri Nasional (BATNAS) yang dikelola oleh Badan Informasi Geospasial (BIG). Penting untuk dipahami bahwa seluruh batasan dan disclaimer yang berlaku pada sumber data utama di atas secara otomatis mengikat pada laporan interpretasi ini. Laporan ini diperuntukkan khusus bagi kegiatan dengan kategori risiko rendah. Hal ini dikarenakan data yang digunakan merupakan hasil pemodelan global dan asimilasi data satelit yang memiliki tingkat akurasi serta resolusi spasial yang tidak sepresisi pengukuran insitu (langsung di lapangan). Disclaimer: Penulis dan penyedia data tidak bertanggung jawab atas kerugian material maupun non-material yang timbul akibat penggunaan data ini untuk perencanaan konstruksi berat atau kegiatan risiko tinggi lainnya tanpa adanya validasi lapangan terlebih dahulu. Hasil laporan ini bersifat indikatif dan digunakan sebagai referensi awal.

II. ARUS Berdasarkan analisis data arus di lokasi kajian, kecepatan arus maksimum yang teridentifikasi pada titik pusat analisis adalah 0.64 m/s. Sebagaimana ditunjukkan pada grafik terlampir di akhir dokumen ini, arah dominan arus berasal dari 131.0 derajat. Analisis lebih lanjut pada profil garis sepanjang 21.91 km menunjukkan variasi kecepatan arus maksimum berkisar antara 0.42 m/s hingga 0.60 m/s, dengan nilai rata-rata sebesar 0.5 m/s. Kondisi arus ini tergolong sedang dan perlu dipertimbangkan dalam perencanaan kegiatan di perairan,

Figure 7 Report generated by LLM (Gemini and Deepseek)

Source : Analysis results (2026)

E. System Integration and Implications for Marine Spatial Planning (MSP)

The successful integration of the four primary physical parameters (Tides, Waves, Currents, Bathymetry) into a single, lightweight, and interactive platform framework carries significant policy implications for the Makassar Marine Spatial Management Office (BPRL) in administering the KKPRL.

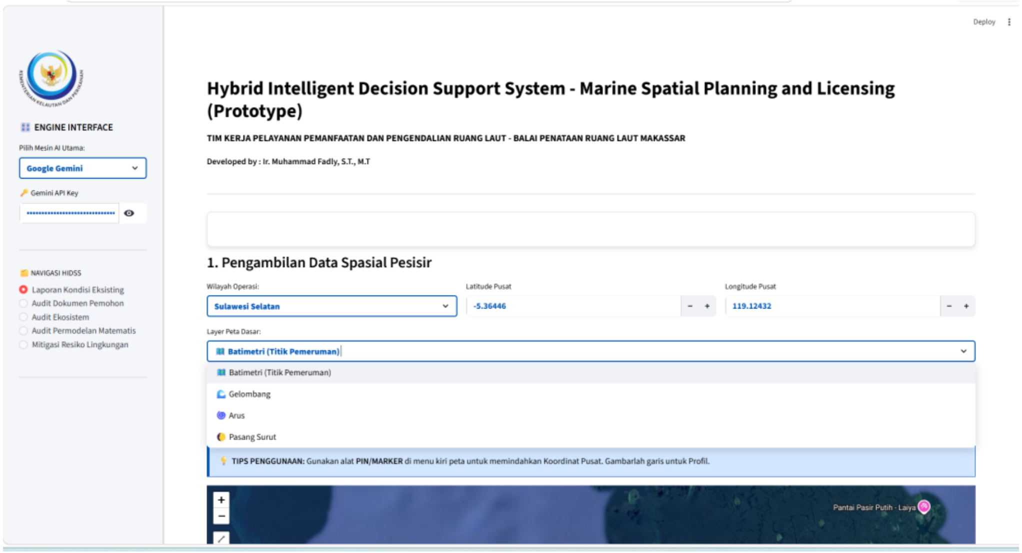
First, this system redefines transparency and accuracy in the technical audits of applicant documents. Bureaucratic evaluators, who are often not oceanographic experts, are now supported by a robust scientific instrument to validate applicants' technical claims in a matter of seconds, fostering a true evidence-based policy environment.

Second, the system is highly reliable for supporting risk-based licensing. For low-risk marine activities (e.g., the installation of Fish Aggregating Devices or small-scale mariculture), data derived from this system serves as a sufficient primary technical reference for issuing approvals. This significantly accelerates bureaucratic services without compromising fundamental coastal engineering principles.

Third, the Parquet-formatted Data Lake architecture developed in this study (Module 1) is fully primed for future integration with Artificial Intelligence (AI) for document auditing (Module 2) and automated risk assessment systems utilizing Fuzzy Logic (Modules 4 and 5). Consequently, this framework establishes a fundamental

History of article:
Received: 5 June 2026, Revised: 5 June 2026, Published: 5 June 2026

stepping stone toward the full automation of intelligent marine spatial planning decision support systems in the future.



Hybrid Intelligent Decision Support System - Marine Spatial Planning and Licensing (Prototype)

TIM KERJA PELAYANAN PEMANFAATAN DAN PENGENDALIAN RUANG LAUT - BALAI PENATAAN RUANG LAUT MAKASSAR

Developed by : Ir. Muhammad Fadly, S.T., M.T

1. Pengambilan Data Spasial Pesisir

Pilih Mesin AI Utama: Google Gemini

Gemini API Key: ••••••••••••••••••••••••••••••

Wilayah Operasi: Sulawesi Selatan

Latitude Pusat: -5.36446

Longitude Pusat: 119.12432

Layer Peta Dasar: Batimetri (Titik Pemeruman)

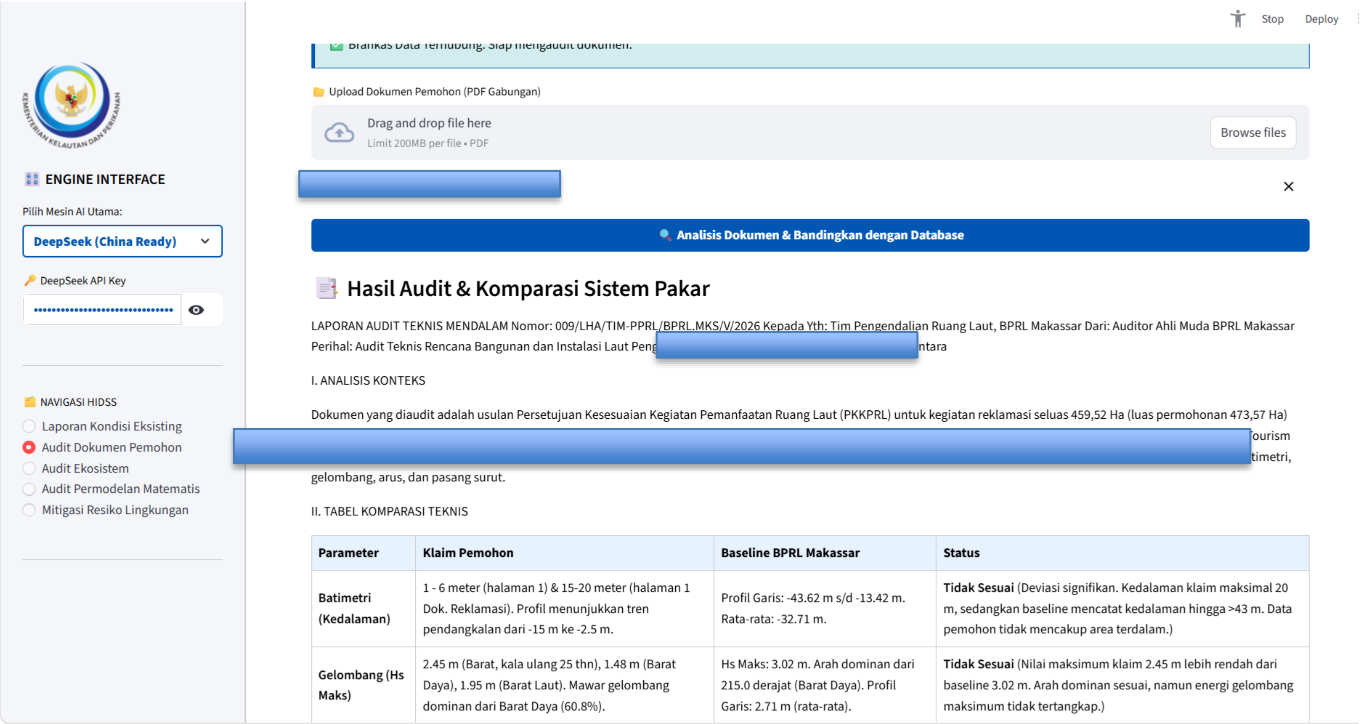
💡 **TIPS PENGGUNAAN:** Gunakan alat **PIN/MARKER** di menu kiri peta untuk memindahkan Koordinat Pusat. Gambarlah garis untuk Profil.

Figure 8 System Integration and Implications for Marine Spatial Planning (MSP)

Source : Analysis results (2026)

E.1 Module 2: Automated Applicant Document Auditing

The HIDSS framework extends its capabilities beyond spatial visualization by incorporating an automated document auditing system, designated as Module 2. As illustrated in Figure 9, this module serves as the primary interface for evaluating the technical feasibility documents submitted by applicants for marine spatial utilization approvals. A defining feature of this module is its dynamic "Engine Interface," which grants evaluators the flexibility to select their preferred Large Language Model (LLM) backend—such as DeepSeek—and securely input the required API key for processing.



[redacted]

Hasil Audit & Komparasi Sistem Pakar

LAPORAN AUDIT TEKNIS MENDALAM Nomor: 009/LHA/TIM-PPRL/BPRL.MKS/V/2026 Kepada Yth: Tim Pengendalian Ruang Laut, BPRL Makassar Dari: Auditor Ahli Muda BPRL Makassar Perihal: Audit Teknis Rencana Bangunan dan Instalasi Laut Peng[redacted]ntara

I. ANALISIS KONTEKS

Dokumen yang diaudit adalah usulan Persetujuan Kesesuaian Kegiatan Pemanfaatan Ruang Laut (PKKPRL) untuk kegiatan reklamasi seluas 459,52 Ha (luas permohonan 473,57 Ha) [redacted] ourism timetri, gelombang, arus, dan pasang surut.

II. TABEL KOMPARASI TEKNIS

Parameter	Klaim Pemohon	Baseline BPRL Makassar	Status
Batimetri (Kedalaman)	1 - 6 meter (halaman 1) & 15-20 meter (halaman 1 Dok. Reklamasi). Profil menunjukkan tren pendangkalan dari -15 m ke -2.5 m.	Profil Garis: -43.62 m s/d -13.42 m. Rata-rata: -32.71 m.	**Tidak Sesuai** (Deviasi signifikan. Kedalaman klaim maksimal 20 m, sedangkan baseline mencatat kedalaman hingga >43 m. Data pemohon tidak mencakup area terdalam.)
Gelombang (Hs Maks)	2.45 m (Barat, kala ulang 25 thn), 1.48 m (Barat Daya), 1.95 m (Barat Laut). Mawar gelombang dominan dari Barat Daya (60.8%).	Hs Maks: 3.02 m. Arah dominan dari 215.0 derajat (Barat Daya). Profil Garis: 2.71 m (rata-rata).	**Tidak Sesuai** (Nilai maksimum klaim 2.45 m lebih rendah dari baseline 3.02 m. Arah dominan sesuai, namun energi gelombang maksimum tidak tertangkap.)

Figure 9 Module 2 interface for evaluating the technical feasibility documents submitted by applicants

Source : Analysis results (2026)

Upon navigating to the "Applicant Document Audit" (*Audit Dokumen Pemohon*) menu, the system immediately initializes and verifies its connection to the centralized Parquet-formatted Data Lake. This successful integration is indicated by a system confirmation stating that the data vault is connected and ready for auditing (*"Brankas Data Terhubung. Siap mengaudit dokumen."*).

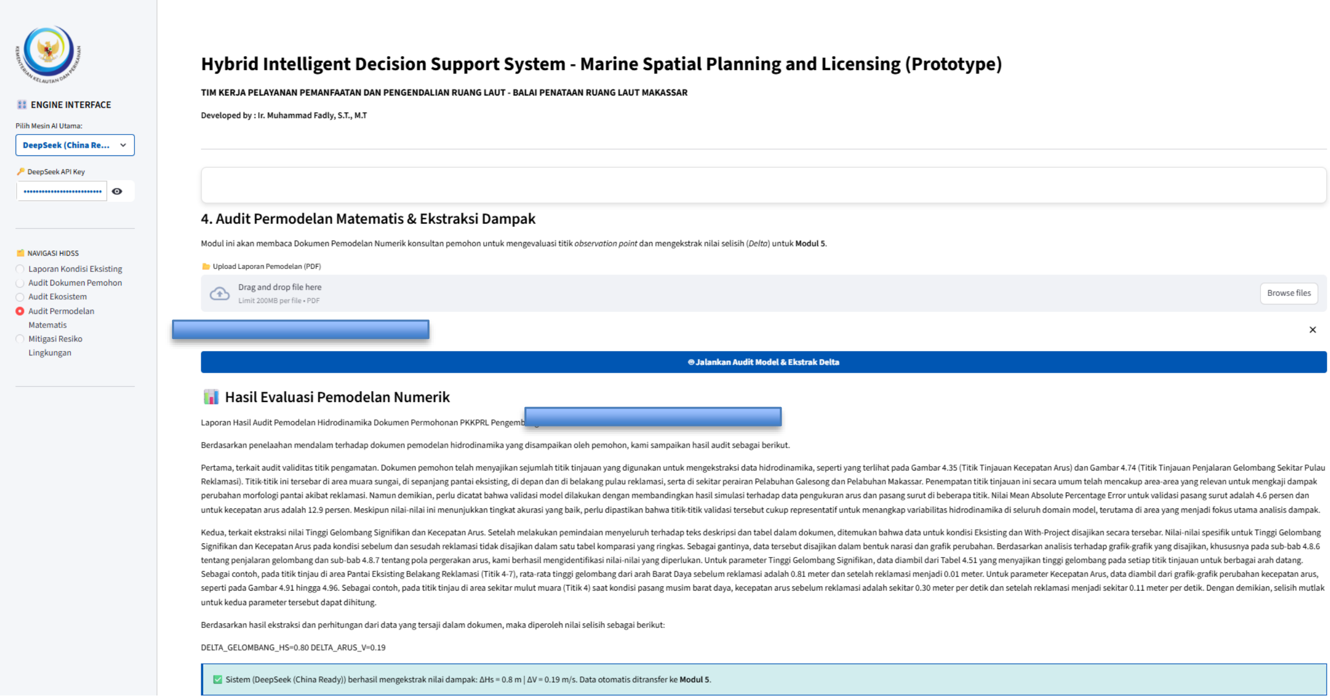
The interface facilitates a streamlined workflow where evaluators can easily upload consolidated feasibility study documents in PDF format—accommodating comprehensive reports up to 200MB—via a drag-and-drop mechanism. This module acts as the critical gateway where the uploaded textual and tabular claims made by applicants are systematically extracted by the selected AI engine. The AI then cross-references these claims against the highly precise hydro-oceanographic statistical data generated by the backend spatial pipeline (Module 1). This automation drastically expedites the bureaucratic verification process, ensuring that the technical parameters proposed by the applicant align strictly with the actual physical constraints of the marine environment.

E.2 Module 4: Mathematical Modeling Audit and Impact Extraction (Audit Permodelan Matematis & Ekstraksi Dampak)

Advancing the system's analytical depth, the HIDSS framework features Module 4, which is specifically dedicated to the rigorous evaluation of hydrodynamic numerical modeling reports submitted by the applicant's consultants. As depicted in Figure 10, this module maintains the flexible AI Engine interface (e.g., DeepSeek) while focusing on highly technical engineering documents.

The primary objective of Module 4 is to automate the extraction of complex technical metrics from large PDF reports (up to 200MB). According to the system interface, the integrated AI is programmed to read the numerical modeling documents and specifically evaluate the defined "observation points."

History of article:
Received: 5 June 2026, Revised: 5 June 2026, Published: 5 June 2026



Hybrid Intelligent Decision Support System - Marine Spatial Planning and Licensing (Prototype)

TIM KERJA PELAYANAN PEMANFAATAN DAN PENGENDALIAN RUANG LAUT - BALAI PENATAAN RUANG LAUT MAKASSAR

Developed by : Ir. Muhammad Fadly, S.T., M.T

4. Audit Permodelan Matematis & Ekstraksi Dampak

Modul ini akan membaca Dokumen Pemodelan Numerik konsultan pemohon untuk mengevaluasi titik *observation point* dan mengekstrak nilai selisih (*Delta*) untuk **Modul 5.**

[redacted]

Hasil Evaluasi Pemodelan Numerik

Laporan Hasil Audit Pemodelan Hidrodinamika Dokumen Permohonan PKKPRL Pengemb[redacted]

Berdasarkan penelaahan mendalam terhadap dokumen pemodelan hidrodinamika yang disampaikan oleh pemohon, kami sampaikan hasil audit sebagai berikut.

Pertama, terkait audit validitas titik pengamatan. Dokumen pemohon telah menyajikan sejumlah titik tinjauan yang digunakan untuk mengekstraksi data hidrodinamika, seperti yang terlihat pada Gambar 4.35 (Titik Tinjauan Kecepatan Arus) dan Gambar 4.74 (Titik Tinjauan Penjalaran Gelombang Sekitar Pulau Reklamasi). Titik-titik ini tersebar di area muara sungai, di sepanjang pantai eksisting, di depan dan di belakang pulau reklamasi, serta di sekitar perairan Pelabuhan Galesong dan Pelabuhan Makassar. Penempatan titik tinjauan ini secara umum telah mencakup area-area yang relevan untuk mengkaji dampak perubahan morfologi pantai akibat reklamasi. Namun demikian, perlu dicatat bahwa validasi model dilakukan dengan membandingkan hasil simulasi terhadap data pengukuran arus dan pasang surut di beberapa titik. Nilai Mean Absolute Percentage Error untuk validasi pasang surut adalah 4.6 persen dan untuk kecepatan arus adalah 12.9 persen. Meskipun nilai-nilai ini menunjukkan tingkat akurasi yang baik, perlu dipastikan bahwa titik-titik validasi tersebut cukup representatif untuk menangkap variabilitas hidrodinamika di seluruh domain model, terutama di area yang menjadi fokus utama analisis dampak.

Kedua, terkait ekstraksi nilai Tinggi Gelombang Signifikan dan Kecepatan Arus. Setelah melakukan pemindaian menyeluruh terhadap teks deskripsi dan tabel dalam dokumen, ditemukan bahwa data untuk kondisi Eksisting dan With-Project disajikan secara tersebar. Nilai-nilai spesifik untuk Tinggi Gelombang Signifikan dan Kecepatan Arus pada kondisi sebelum dan sesudah reklamasi tidak disajikan dalam satu tabel komparasi yang ringkas. Sebagai gantinya, data tersebut disajikan dalam bentuk narasi dan grafik perubahan. Berdasarkan analisis terhadap grafik-grafik yang disajikan, khususnya pada sub-bab 4.8.6 tentang penjalaran gelombang dan sub-bab 4.8.7 tentang pola pergerakan arus, kami berhasil mengidentifikasi nilai-nilai yang diperlukan. Untuk parameter Tinggi Gelombang Signifikan, data diambil dari Tabel 4.51 yang menyajikan tinggi gelombang pada setiap titik tinjauan untuk berbagai arah datang. Sebagai contoh, pada titik tinjau di area Pantai Eksisting Belakang Reklamasi (Titik 4-7), rata-rata tinggi gelombang dari arah Barat Daya sebelum reklamasi adalah 0.81 meter dan setelah reklamasi menjadi 0.01 meter. Untuk parameter Kecepatan Arus, data diambil dari grafik-grafik perubahan kecepatan arus, seperti pada Gambar 4.91 hingga 4.96. Sebagai contoh, pada titik tinjau di area sekitar mulut muara (Titik 4) saat kondisi pasang musim barat daya, kecepatan arus sebelum reklamasi adalah sekitar 0.30 meter per detik dan setelah reklamasi menjadi sekitar 0.11 meter per detik. Dengan demikian, selisih mutlak untuk kedua parameter tersebut dapat dihitung.

Berdasarkan hasil ekstraksi dan perhitungan dari data yang tersaji dalam dokumen, maka diperoleh nilai selisih sebagai berikut:

```
DELTA_GELOMBANG_HS=0.80 DELTA_ARUS_V=0.19
```

✅ Sistem (DeepSeek (China Ready)) berhasil mengekstrak nilai dampak: ΔHs = 0.8 m | ΔV = 0.19 m/s. Data otomatis ditransfer ke **Modul 5**.

Figure 10 Module 4 interface for Mathematical Modeling Audit and Impact Extraction

Source : Analysis results (2026)

Crucially, the module is engineered to extract the "Delta" values—the quantitative differences representing the environmental impact of the proposed project (such as changes in wave energy, current velocity, or sedimentation patterns before and after construction). Rather than acting as a standalone extraction tool, Module 4 serves as an intelligent data feeder; the extracted Delta values are systematically forwarded as direct inputs into Module 5 (Environmental Risk Mitigation). This interconnected architecture ensures that the evaluation of physical environmental impacts is seamlessly transitioned into automated, quantitative risk mitigation protocols, thereby eliminating human error in transferring highly technical engineering data during the licensing audit.

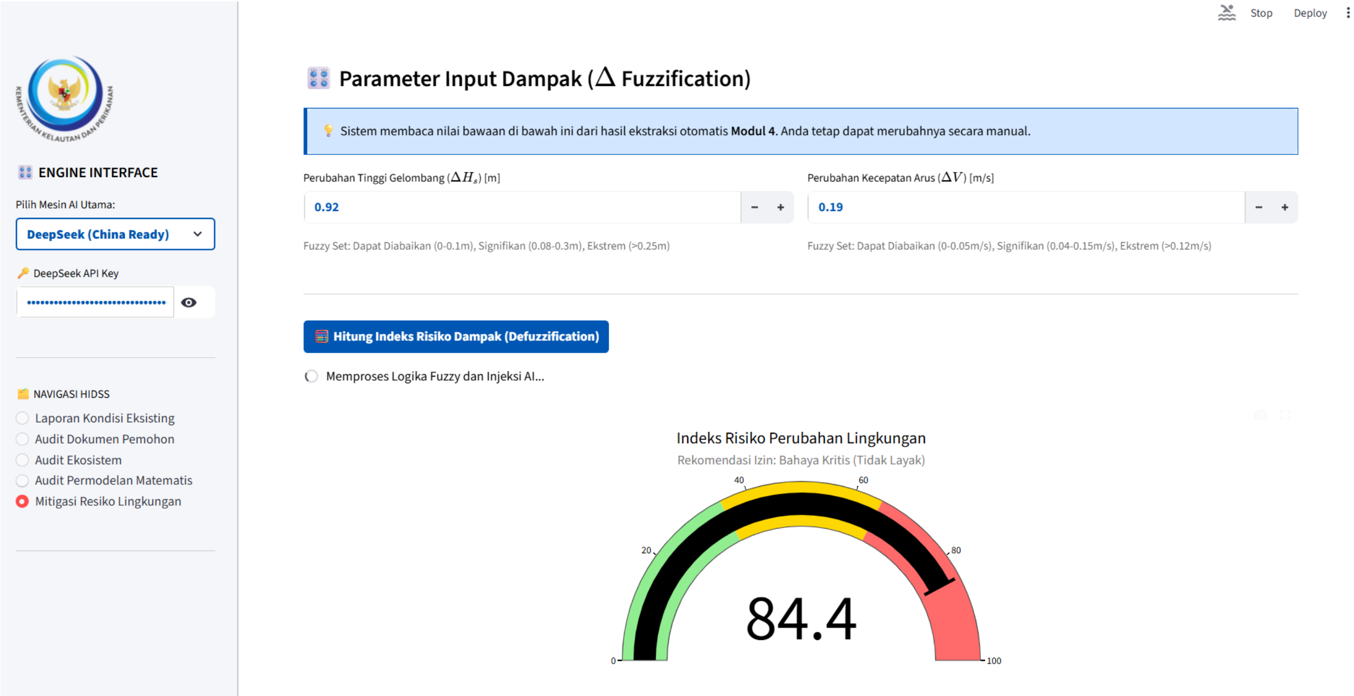
E.3 Module 5: Environmental Risk Mitigation and Fuzzy Logic Assessment (Mitigasi Risiko Lingkungan)

Serving as the culmination of the analytical pipeline, Module 5 functions as the quantitative decision-making core of the HIDSS framework. As illustrated in Figure 11, this module operates on a Fuzzy Logic architecture designed to evaluate the environmental impacts of proposed marine infrastructure. The system seamlessly imports the "Delta" values (Δ) automatically extracted by Module 4, specifically the projected changes in significant wave height (ΔH) and current velocity (ΔV). While these parameters are auto-populated to maintain workflow efficiency, the interface provides evaluators with the flexibility to manually adjust the values for calibration purposes.

The risk assessment methodology begins with the Fuzzification process, where the absolute numerical changes are categorized into linguistic risk sets based on predefined oceanographic thresholds. According to the system's logic, the impact parameters are classified into subsets such as Negligible (*Dapat Diabaikan*), Significant (*Signifikan*), and Extreme (*Ekstrem*). In the specific scenario depicted in Figure 11, the input values reflect a severe environmental alteration $\Delta H_s = 0.80$ m and $\Delta V = 0.19$m/s), both of which far exceed the threshold for the "Extreme" fuzzy subset.

Upon executing the Defuzzification protocol, the system aggregates these fuzzy inputs into a single, actionable metric: the Environmental Change Risk Index (*Indeks Risiko Perubahan Lingkungan*). As shown in the output gauge chart, the compounding extreme variables result in a severe risk score of 84.4. Crucially, the system instantly translates this numerical index into a definitive policy recommendation—in this instance, "Critical Hazard (Not Feasible)" (*Bahaya Kritis - Tidak Layak*).

By integrating mathematically rigorous Fuzzy Logic with the automated data extraction from previous modules, Module 5 effectively eliminates subjectivity from the environmental impact assessment process. It provides policymakers and licensing authorities with an immediate, evidence-based verdict, empowering them to confidently reject high-risk spatial utilization applications or mandate substantial engineering mitigations prior to approval.



Pilih Mesin AI Utama: DeepSeek (China Ready)

DeepSeek API Key: ••••••••••••••••••••••••••••••

○ Laporan Kondisi Eksisting
○ Audit Dokumen Pemohon
○ Audit Ekosistem
○ Audit Permodelan Matematis
● Mitigasi Resiko Lingkungan

Parameter Input Dampak (Δ Fuzzification)

Sistem membaca nilai bawaan di bawah ini dari hasil ekstraksi otomatis **Modul 4**. Anda tetap dapat merubahnya secara manual.

Perubahan Tinggi Gelombang (ΔH_s) [m]: 0.92

Fuzzy Set: Dapat Diabaikan (0-0.1m), Signifikan (0.08-0.3m), Ekstrem (>0.25m)

Perubahan Kecepatan Arus (ΔV) [m/s]: 0.19

Fuzzy Set: Dapat Diabaikan (0-0.05m/s), Signifikan (0.04-0.15m/s), Ekstrem (>0.12m/s)

Memproses Logika Fuzzy dan Injeksi AI...

Indeks Risiko Perubahan Lingkungan

Rekomendasi Izin: Bahaya Kritis (Tidak Layak)

84.4

Figure 11 Environmental Risk Mitigation and Fuzzy Logic Assessment

Source : Analysis result (2026)

IV CONCLUSION

This study successfully achieved its primary objective of bridging the technical gap between complex global oceanographic big data and the practical needs of local-level marine spatial planning. By developing the Python-based Hybrid Intelligent Decision Support System (HIDSS) framework, the research effectively transformed massive, computationally heavy hydro-oceanographic datasets into a streamlined, highly interactive platform interface. The automated spatial data pipeline demonstrated that leveraging geospatial libraries for spatial clipping successfully eliminates land mask boundary noise, thereby significantly reducing hydrodynamic uncertainty in complex, shallow coastal waters. The key findings highlight the system's capability to deliver precise, real-time analytical outputs critical for risk-based licensing. The brute-force computation of 20-year tidal

History of article:
Received: 5 June 2026, Revised: 5 June 2026, Published: 5 June 2026

hindcasting utilizing eight major harmonic constituents, combined with dynamic bathymetric profiling and instantaneous wave and current climate visualizations, provides evaluators with highly accurate extreme datums and operational statistics. Furthermore, the implementation of a Parquet-formatted Data Lake architecture proved to be a pivotal optimization, drastically minimizing storage requirements and ensuring low-latency performance during real-time spatial queries. These technological advancements directly empower regulatory bodies to conduct rapid, transparent, and strictly evidence-based technical audits for marine spatial suitability approvals. Looking forward, the established automated pipeline lays a robust foundation for the next generation of intelligent maritime bureaucracy. The highly structured columnar storage enables seamless API integration with Large Language Models, such as Gemini and DeepSeek, to automatically translate raw numerical data into comprehensive descriptive risk assessments. Future developments will focus on fully operationalizing this Artificial Intelligence integration for automated document auditing and implementing comprehensive risk assessment modules utilizing Fuzzy Logic, ultimately paving the way for a fully autonomous, intelligent marine spatial licensing ecosystem.

ACKNOWLEDGMENT

The authors express their deepest gratitude to the Makassar Marine Spatial Management Office (BPRL), Ministry of Marine Affairs and Fisheries of the Republic of Indonesia, for providing the institutional support and the practical regulatory framework necessary for developing this spatial decision support system. Sincere appreciation to the Perkumpulan Ahli Rekayasa Pantai Indonesia (PARPI) for their invaluable insights, constructive feedback, and collaborative discussions regarding coastal engineering standards and mathematical modeling frameworks. Furthermore, the authors acknowledge the Copernicus Marine Environment Monitoring Service (CMEMS) and the Geospatial Information Agency of Indonesia (BIG) for their commitment to open science by providing access to the essential hydro-oceanographic and bathymetric datasets utilized throughout this study.

REFERENCE

Badan Informasi Geospasial, 2018. *Buku Panduan Spesifikasi Produk Data Batimetri Nasional (BATNAS)*. Cibinong: Pusat Pemetaan Kelautan dan Lingkungan Pantai.

Egbert, G. D. and Erofeeva, S. Y., 2002. Efficient Inverse Modeling of Barotropic Ocean Tides. *Journal of Atmospheric and Oceanic Technology*, 19(2), pp.183-204.

Ehler, C. and Douvere, F., 2009. *Marine Spatial Planning: a step-by-step approach toward ecosystem-based management*. Intergovernmental Oceanographic Commission and Man and the Biosphere Programme. IOC Manual and Guides No. 53. Paris: UNESCO.

Guidi, L., Fernandez Guerra, A., Canchaya, C., Curry, E., Foglini, F., Irisson, J. O., Malde, K., Marshall, C. T., Obst, M., Ribeiro, R. P. and Tjiputra, J., 2020. *Big Data in Marine Science*. Future Science Brief 6 of the European Marine Board, Ostend, Belgium.

Halpern, B. S., Longo, C., Hardy, D., McLeod, K. L., Samhouri, J. F., Katona, S. K., Kleisner, K., Lester, S. E., O'Leary, J., Ranelletti, M., Rosenberg, A. A., Scarborough, C., Selig, E. R., Best, B. D., Brumbaugh, D. R., Chapin, F. S., Crowder, L. B., Daly, K. L., Doney, S. C., Elfes, C., Fogarty, M. J., Gaines, S. D., Jacobsen, K.

I., Karrer, L. B., Leslie, H. M., Neeley, E., Pauly, D., Polasky, S., Risien, B., St Martin, K., Stone, M., Sumaila, U. R. and Zeller, D., 2012. An index to assess the health and benefits of the global ocean. *Nature*, 488(7413), pp.615-620.

Holthuijsen, L. H., 2007. *Waves in Oceanic and Coastal Waters*. Cambridge: Cambridge University Press.

Kementerian Kelautan dan Perikanan RI, 2021. *Peraturan Menteri Kelautan dan Perikanan Nomor 28 Tahun 2021 tentang Penyelenggaraan Penataan Ruang Laut*. Jakarta.

Le Traon, P. Y., Reppucci, A., Alvarez Fanjul, E., Aouf, L., Behrens, A., Belmonte, M. and Bentamy, A., 2019. From Observation to Information and Users: The Copernicus Marine Service Perspective. *Frontiers in Marine Science*, 6, p.234.

Pınarbaşı, K., Galparsoro, I., Borja, Á., Stelzenmüller, V., Ehler, C. N. and Garmendia, J. M., 2017. Decision support tools in marine spatial planning: Present applications, gaps and future perspectives. *Marine Policy*, 83, pp.83-91.

Pond, S. and Pickard, G. L., 2013. *Introductory Dynamical Oceanography*. Edisi ke-2. Oxford: Butterworth-Heinemann.

Pugh, D. and Woodworth, P., 2014. *Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes*. Cambridge: Cambridge University Press.

Rayner, R., Jolly, C. and Gouldman, C., 2019. Ocean Observing and the Blue Economy. *Frontiers in Marine Science*, 6, p.330.

Stammer, D., Ray, R. D., Andersen, O. B., Arbic, B. K., Bosch, W., Carrère, L., Cheng, Y., Chinn, D. S., Dushaw, H. O., Egbert, G. D. and Erofeeva, S. Y., 2014. Accuracy assessment of global barotropic ocean tide models. *Reviews of Geophysics*, 52(3), pp.243-282.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J. and van der Walt, S. J., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), pp.261-272.

Vohra, D., 2016. Apache Parquet. Dalam: *Practical Hadoop Ecosystem*. Berkeley, CA: Apress.

History of article:
Received: 5 June 2026, Revised: 5 June 2026, Published: 5 June 2026