

Article

Study on the Oil Spill Transport Behavior and Multifactorial Effects of the Lancang River Crossing Pipeline

Jingyang Lu *, Liqiong Chen and Duo Xu

School of Petroleum and Natural Gas Engineering, Southwest Petroleum University, Chengdu 610500, China

* Correspondence: 202111000128@stu.swpu.edu.cn

Abstract: As the number of long-distance oil and gas pipelines crossing rivers increases, so does the risk of river oil spills. Previous research on oil spills in water mainly focuses on the oceans, and there are relatively few studies on oil spills in rivers. This study established two-dimensional hydrodynamic and oil spill models for the Lancang River crossing pipeline basin and verified the model's accuracy. The oil spill transport process under different scenarios was simulated, and the oil spill transport state data set was established. The effects of river flow, wind, and leakage mode on the transport behavior of oil spills were studied. The results show that an increase in flow rate accelerates the migration, diffusion, and longitudinal extension behavior of oil spills; Changes in wind speed have less effect on the transport behavior of oil spills under downwind and headwind conditions. The mode of leakage mainly affects the diffusion and longitudinal extension of the oil spill. The oil spill transport state prediction model was established using machine learning combination algorithms. The three combined machine learning algorithms, PSO-SVR, GA-BPNN, and PSO-BPNN, have the best performance in predicting the oil spill migration distance, oil spill area, and the length of the oil spill contamination zone, respectively, with the coefficient of determination (R^2) and the 1-Mean Absolute Percentage of Error (1-MAPE) above 0.971, and the prediction model has excellent accuracy. This study can provide support for the rapid development of emergency response plans for river crossing pipeline oil spill accidents.

Keywords: oil spill; transport behavior; oil spill transport state prediction model; river crossing pipeline



Citation: Lu, J.; Chen, L.; Xu, D. Study on the Oil Spill Transport Behavior and Multifactorial Effects of the Lancang River Crossing Pipeline. *Appl. Sci.* **2024**, *14*, 3455. <https://doi.org/10.3390/app14083455>

Academic Editor: Micòl Mastrocicco

Received: 6 March 2024

Revised: 16 April 2024

Accepted: 17 April 2024

Published: 19 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The rising demand for oil and gas in society implies a greater need for transportation via pipelines, vessels, and other means, thereby escalating the frequency and scale of energy transportation activities. With the amplification of energy transport operations, the probability of oil spill incidents during transportation also increases. Factors contributing to this include natural environmental conditions during navigation, potential equipment and facility malfunctions, as well as human error [1–3]. On 26 July 2010, a crude oil pipeline rupture of the Canadian pipeline company Enbridge located in the territory of Michigan resulted in the flow of at least 4000 cubic meters of crude oil into the Kalamazoo River, which has become one of the largest crude oil spills in the history of the United States [4]. On 22 November 2013, an oil pipeline rupture occurred at the Qingdao Oil Transmission and Storage Company (QOTSC), with some of the crude oil entering the Jiaozhou Bay along a stormwater pipeline, resulting in the contamination of approximately 3000 square meters of sea surface in the Jiaozhou Bay. Subsequently, a huge explosion occurred, resulting in 63 deaths and 156 injuries [5]. On 5 July 2020, the Japanese cargo ship Wakashio spilled off the southeast coast of Mauritius, with more than 1000 tonnes of fuel oil leaking into the Indian Ocean, threatening many rare species [6]. Oil spills not only cause serious pollution to the environment but also cause serious damage to the flora and fauna in the watershed and even to human life [7–9].

The study on oil spills in waters began in the 1960s. With the long-term development of hydrodynamic modeling, oil spill modeling, and numerical methods, the observation

of the physical behavior of oil spills in watersheds using numerical simulation has been widely applied [2,10–19]. Zhen and Li et al. [20] established a two-dimensional hydrodynamic model of Daya Bay in the South China Sea and studied the behavior and fate of oil spills in both clockwise and anticlockwise currents. Liu et al. [5] established a model of the oil spill in Jiaozhou Bay and verified the accuracy of the model with the “11.22” oil spill in Huangdao Island, and studied the effects of wind speed, tidal currents, and the location of leakage on the drifting and diffusion movements of oil spills. Perri  nez [21] established the Red Sea Lagrangian oil spill transport model to simulate the migration of oil spills in winter and summer seasons, different oil spill locations, and different current environments. Bozkurtoglu [22] developed a trajectory modeling of the oil spill in the Bosphorus Strait, investigated the short-term spreading behavior of the oil spill under different wind directions, and made recommendations for the oil spill contingency plan in the basin. Kuang [18] developed a two-dimensional hydrodynamic and oil spill transport model for the Yangtze River estuary, which mainly investigated the effects of wind and wave current interactions on the drift and diffusion behavior of oil spills under strong wind conditions. Da Cunha and De Abreu et al. [23] developed a hydrodynamic model and Lagrangian transport model for the Santana channel watershed in the lower Amazonian River flow, which simulated the spatial and temporal displacements of oil spills and the contaminated area of the oil sheen in the dry and rainy seasons and pointed out the critical contaminated zones and sensitive areas under the spill. Jiang and Tong et al. [24] established an oil spill model for an inland waterway and calibrated and validated the model by using flume experiments. The model was applied to the Luoqi section of the Yangtze River, and the migration trajectories of the oil spill were simulated under two scenarios of high flow and low wind speed and low flow and high wind speed. Although the volume of oil spills in inland rivers is relatively small, they serve as habitats for numerous wildlife species, and spills in these rivers can still have significant impacts on local ecosystems and biodiversity. Additionally, inland riverwater bodies are often used for drinking water, irrigation, fishing, and tourism. Thus, oil spill contamination may directly threaten the health and livelihoods of local residents. Despite the smaller scale of inland river pollution, the cleanup and restoration efforts may be constrained by geographical location and environmental conditions, resulting in more challenging and complex cleanup processes. Therefore, in-depth research on inland river oil spills is imperative [2,24,25].

With the massive construction of oil and gas pipelines, the increase in the number of pipelines laid along rivers and crossing rivers has increased the risk of river oil spill accidents [26,27]. Unlike the marine environment, the migration trajectory of oil spills in rivers is relatively unidirectional, the river channel is relatively narrow, the flow and water level of the river are highly variable, and the transport behavior of oil spills is more significantly affected by water flow and wind [28]. In most of the existing river oil spill studies, the oil spill transport behavior of transient spills under some specific scenarios is addressed. The river flow and the wind conditions in the basin usually vary seasonally [29], and the spills from long-distance pipelines are continuous. Therefore, the study of oil spill transport behavior in rivers under the effects of multiple factors can provide more effective support for emergency response to oil spill accidents. In addition, after the occurrence of a sudden oil spill, emergency response measures in the short term can largely reduce the consequences of the accident [23,30]. The use of numerical simulation to predict the transport state of oil spills takes up a lot of time, which prolongs the emergency response time. With the rapid development in the field of data science and AI, machine learning modeling approaches have been heavily researched in various fields, and new methods have been provided for accurate and efficient prediction of river oil spill transport state [31–33].

In this study, a two-dimensional hydrodynamic and oil spill model of the Lancang River crossing pipeline section basin was established based on Mike 21 software to simulate the oil spill transport process in the river. The oil spill head location, oil film diffusion area, and pollution zone length were obtained over time. The effects of river flow, wind, and leakage mode on oil spill transport behavior were discussed. An oil spill transport state

dataset was established, and an oil spill transport state prediction model was developed using a machine learning algorithm. This study can support the rapid development of emergency response plans for river crossing pipeline oil spill accidents.

2. Materials and Methods

2.1. Hydrodynamic Model

MIKE 21 is a simulation tool for physical, chemical, or biological processes in two-dimensional waters developed by the Danish Hydrodynamic Institute [34,35]. It is widely used in ocean, bay, lake, and river environments. The system is based on the numerical solution of the two-dimensional incompressible Reynolds-averaged Navier-Stokes equations, invoking the Boussinesq assumption and the hydrostatic pressure assumption [34]. Integrating the horizontal momentum and continuity equations in the depth direction yields the following shallow water equations, and the integral form of the shallow water system of equations can be written as:

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{U}) = \mathbf{S}(\mathbf{U}) \tag{1}$$

where \mathbf{U} is the vector of conserved variables; \mathbf{F} is the flux vector function, and \mathbf{S} is the vector of source terms.

In Cartesian coordinates, the system of 2D shallow water equations can be written as:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial (\mathbf{F}_x^I - \mathbf{F}_x^V)}{\partial x} + \frac{\partial (\mathbf{F}_y^I - \mathbf{F}_y^V)}{\partial y} = \mathbf{S} \tag{2}$$

where the superscripts I and V denote the inviscid (convective) and viscous fluxes.

$$\begin{aligned} \mathbf{U} &= \begin{bmatrix} h \\ h\bar{u} \\ h\bar{v} \end{bmatrix}, \mathbf{F}_x^I = \begin{bmatrix} h\bar{u}^2 + \frac{1}{2}g(h^2 - d^2) \\ h\bar{u}\bar{v} \end{bmatrix}, \mathbf{F}_x^V = \begin{bmatrix} 0 \\ hA \left(2\frac{\partial \bar{u}}{\partial x} \right) \\ hA \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x} \right) \end{bmatrix}, \\ \mathbf{F}_y^I &= \begin{bmatrix} h\bar{v} \\ h\bar{u}\bar{v} \\ h\bar{v}^2 + \frac{1}{2}g(h^2 - d^2) \end{bmatrix}, \mathbf{F}_y^V = \begin{bmatrix} 0 \\ hA \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x} \right) \\ hA \left(2\frac{\partial \bar{v}}{\partial x} \right) \end{bmatrix}, \\ \mathbf{S} &= \begin{bmatrix} 0 \\ g\eta \frac{\partial d}{\partial x} + f\bar{v}h - \frac{h}{\rho_0} \frac{\partial p_a}{\partial x} - \frac{gh^2}{2\rho_0} \frac{\partial p}{\partial x} - \frac{1}{\rho_0} \left(\frac{\partial s_{xx}}{\partial x} + \frac{\partial s_{xy}}{\partial y} \right) + \frac{\tau_{sx}}{\rho_0} - \frac{\tau_{bx}}{\rho_0} + hu_s \\ g\eta \frac{\partial d}{\partial y} + f\bar{u}h - \frac{h}{\rho_0} \frac{\partial p_a}{\partial y} - \frac{gh^2}{2\rho_0} \frac{\partial p}{\partial y} - \frac{1}{\rho_0} \left(\frac{\partial s_{yx}}{\partial x} + \frac{\partial s_{yy}}{\partial y} \right) + \frac{\tau_{sy}}{\rho_0} - \frac{\tau_{by}}{\rho_0} + hv_s \end{bmatrix} \end{aligned} \tag{3}$$

where x and y are the Cartesian co-ordinates; h is the total water depth, $h = \eta + d$, η and d are the surface elevation and the still water depth; u and v are the velocity components in the x and y direction; A is the horizontal eddy viscosity; f is the Coriolis parameter, $f = 2\omega \sin \varphi$, ω and φ are the angular rate of revolution and the geographic latitude; ρ and ρ_0 are the density of water and the reference density of water; p_a and g are the atmospheric pressure and the gravitational acceleration; (τ_{sx}, τ_{bx}) and (τ_{sy}, τ_{by}) are the x and y components of the surface wind and bottom stresses; (s_{xx}, s_{xy}) and (s_{yx}, s_{yy}) are components of the radiation stress tensor; (u_s, v_s) is the velocity by which the water is discharged into the ambient water; the depth-averaged velocities \bar{u} and \bar{v} are defined by:

$$h\bar{u} = \int_{-d}^{\eta} u dz, h\bar{v} = \int_{-d}^{\eta} v dz \tag{4}$$

Integrating Equation (1) over the i th cell and rewriting the flux integral using Gauss’s theorem yields:

$$\int_{A_i} \frac{\partial \mathbf{U}}{\partial t} d\Omega + \int_{\Gamma_i} (\mathbf{F} \cdot \mathbf{n}) ds = \int_{A_i} S(\mathbf{U}) d\Omega \tag{5}$$

where A_i is the area of the cell and Ω is the integration variable defined on A_i ; Γ_i is the boundary of the i th cell and ds is the integration variable along the boundary; \mathbf{n} is the unit outward normal vector along the boundary. Evaluating the area integrals by a one-point quadrature rule, the quadrature point being the centroid of the cell, and evaluating the boundary integral using a mid-point quadrature rule, Equation (5) can be written as:

$$\frac{\partial U_i}{\partial t} + \frac{1}{A_i} \sum_j^{NS} \mathbf{F} \cdot \mathbf{n} \Delta \Gamma_j = S_i \tag{6}$$

where U_i and S_i are the average values of U and S over the i th cell and stored at the cell center; NS is the number of sides of the cell; $\Delta \Gamma_j$ is the length of the j th interface.

The 2D model uses an approximate Riemann solver (Roe’s scheme) to compute the convective flux at the cell interface. The problem is reduced to a scalar fluctuation problem by linearising the conservation equation. The approximate solution to the problem is achieved by representing the problem in terms of eigenvalues, right eigenvectors, and left eigenvectors using linear algebra [36]. The mean gradient is estimated using the method proposed by Jawahar and Kamath, 2000 [37]. To avoid numerical oscillations, a second-order total variation diminishing (TVD) slope limiter is used [38].

2.2. Oil Spill Model

The MIKE 21 Oil Spill Model is used to simulate the fate of oil spilled or discharged into waters, including its transport processes and changes in physical and chemical properties. The model uses the Lagrangian method to track the motion of each oil spill particle and the follow-along method to describe the turbulent diffusive motion caused by shear and turbulence.

2.2.1. Diffusion and Drift of Oil Film

Crude oil enters the river and rapidly expands into an oil film under the action of gravity, inertial forces, viscous forces, and interfacial tension. A modification of Fay’s theory of oil film expansion proposed by Mackay and Paterson et al. [39] was used to calculate the diffusion behavior of the oil film:

$$\frac{dA}{dt} = K_{\text{Spread}} \cdot A^{\frac{1}{3}} \cdot \left[\frac{V}{A} \right]^{\frac{4}{3}} \tag{7}$$

where K_{Spread} is a rate constant, taken as 150 s^{-1} ; A and V are t area and volume of oil particles in river.

The drift behavior of the oil spill in a river is driven by a combination of current and wind forces, and the drift velocity of its oil particles is calculated by the following equation:

$$\begin{aligned} U_p &= U_c + C_w \cdot U_w \sin(\alpha_w - \pi + \theta_w) \\ V_p &= V_c + C_w \cdot U_w \cos(\alpha_w - \pi + \theta_w) \end{aligned} \tag{8}$$

where (U_p, U_c) is the particle velocity along x and y directions; (V_p, V_c) is the surface current speed in x and y directions; C_w , U_w , and α_w are the wind drift coefficient, the wind speed, and direction; θ_w is the wind drift angle, which can be calculated by the following formulation:

$$\theta_w = \beta \exp\left(\frac{\alpha |U_w|^3}{g \gamma_w}\right) \tag{9}$$

where γ_w and g are the kinematic viscosity and the gravity acceleration; α and β are the empirical coefficients, -0.3×10^{-8} and $28^\circ 38'$, respectively.

In addition, there is a turbulent diffusion behavior of oil particles in rivers, which can be viewed as stochastic. Assuming that the turbulent diffusion is isotropic, the diffusion distance of oil particles per unit time can be expressed as:

$$S_o = [R]_{-1}^1 \sqrt{6D_o \Delta t} \tag{10}$$

where $[R]_{-1}^1$ is a random number from 1 to -1 ; D_o is the diffusion coefficient.

2.2.2. Weathering of Oil Spill

Oil spills entering the river are affected by factors such as water flow, wind, and light radiation, and undergo processes such as evaporation, emulsification, dissolution, sedimentation, photo-oxidation, and biodegradation, resulting in changes in the physicochemical properties of the oil [20]. This study focuses on the evaporation and emulsification behavior of oil spills.

Where oil evaporation is a short-term weathering process, evaporation will be more obvious in the early stage after the oil spill enters the water body. The Reed model [40] was used to calculate the evaporation of an oil spill with the following equation:

$$EVAP = \frac{K_2 \cdot P_{vp} \cdot A}{R \cdot T} \cdot f \cdot MW \tag{11}$$

where P_{vp} and A are the vapor pressure and the slick area of each particle in contact with the water surface; R is the gas constant, $R = 8.206 \times 10^{-5}$; T , f , and MW are the temperature, the fraction of the evaporative oil component, and the molecular weight; K_2 is mass transfer coefficient, which can be calculated by:

$$K_2 = 0.0292 \cdot w^{0.78} \cdot D^{-0.11} \cdot Sc^{-0.67} \cdot \sqrt{\frac{MW + 29}{MW}} \tag{12}$$

where w and D are the wind speed and the diameter of each particle area in contact with the water surface; Sc is Schmidt number, $Sc = 2.7$.

Emulsification is the process by which oil and water mix to form a mixture, and it is encouraged by wind and river turbulence, which usually occurs a few hours after the oil spill enters the water column. The emulsification process uses a first-order water release equation [41]:

$$wateruptake = K_{em} * (U + 1)^2 \cdot \frac{(Y_{max} - Y_w)}{Y_{max}} \tag{13}$$

$$waterrelease = -\alpha \cdot Y_w \tag{14}$$

where Y_w and Y_{max} are the water fraction and the maximum water fraction; K_{em} is the emulsification rate constant, $K_{em} = 2 \times 10^{-6} \text{ s/m}^2$; U is the wind speed; α is water release rate which is related to the emulsion stability index S .

$$\alpha = \begin{cases} \alpha_0 - (\alpha_0 - \alpha_{0.67})S/0.67 & \text{for } S < 0.67 \\ \alpha_{0.67}[(1.22 - S)/(1.22 - 0.67)] & \text{for } 0.67 \leq S < 1.22 \\ 0 & \text{for } S \geq 1.22 \end{cases} \tag{15}$$

where α_0 and $\alpha_{0.67}$ are the water release for unstable emulsion with $S = 0$ and the mesostable emulsion with $S = 0.67$. The stability index S is given by:

$$S = X_a \cdot \exp [K_{ao} \cdot (1 - X_a - X_w)^2 + K_{aw} \cdot X_w^2] \cdot \exp [-0.04 \cdot (T - 293)] \tag{16}$$

where subscript a , w , and o represent asphaltene, wax, and other chemical components; K_{ao} and K_{aw} are 3.3 and 200 at 293 K; X_a and X_w are the fractions of asphaltenes and wax; T is the temperature.

2.3. Model Construction

The China-Myanmar Crude Oil Pipeline Lancang River crossing pipeline is located in Baoshan City, Yunnan Province, China. The China-Myanmar pipeline is an important strategic energy corridor for China. The pipeline starts from Kyaukphyu City on the west coast of Myanmar, enters China from Ruili City in Yunnan Province, and ends in Chongqing City. The Lancang River crossing pipeline basin is located at the western edge of the Yunnan-Guizhou Plateau and in the longitudinal valley of western Yunnan in the southern section of the Hengduan Mountains. The mountains on both sides of the river are high and steep, the riverbed is narrow, and the water flow is turbulent, which belongs to the rapid-flow type of rivers in high mountain canyons. The valley is “V” shaped, and the slope angles of the left and right banks of the mountain are 53° and 60° , respectively; the real picture is shown in Figure 1. This zone has a subtropical monsoon climate with an average annual rainfall of more than 800 mm. The environmental conditions in this basin are complex, and in the event of disasters such as mudslides or landslides during the rainy season, the pipeline will likely be damaged, and crude oil will leak into the river.



Figure 1. Actual view of Lancang River crossing pipeline.

2.3.1. Hydrodynamic Model Construction

The length of the modeled reach for this study is 33 km, and the crossing point is located in the middle and upper reaches of the modeled reach. Riparian boundary data and river elevation scatter data were obtained from actual measurements. Hydrological data, including flow, level, and velocity data, were provided by hydroelectric and hydrological stations in the basin, and wind data were provided by the local airport. Firstly, the terrain data was converted to map projection coordinates, imported into the mesh generator in Mike Zero to create an unstructured triangular mesh, and the bathymetry data was interpolated into the mesh. The maximum triangular grid area is 2600 square meters, the minimum allowable angle is 30° , there are 3580 nodes and 5884 units in the computational area, and the time step is set to 30 s. The overall topography and local grid map of the river channel in the simulation area are shown in Figure 2, in which the local grid map includes crossing point and validation points (M and N are the validation sites of the hydrodynamic model in the following section) and three oil spill interception and recovery points.

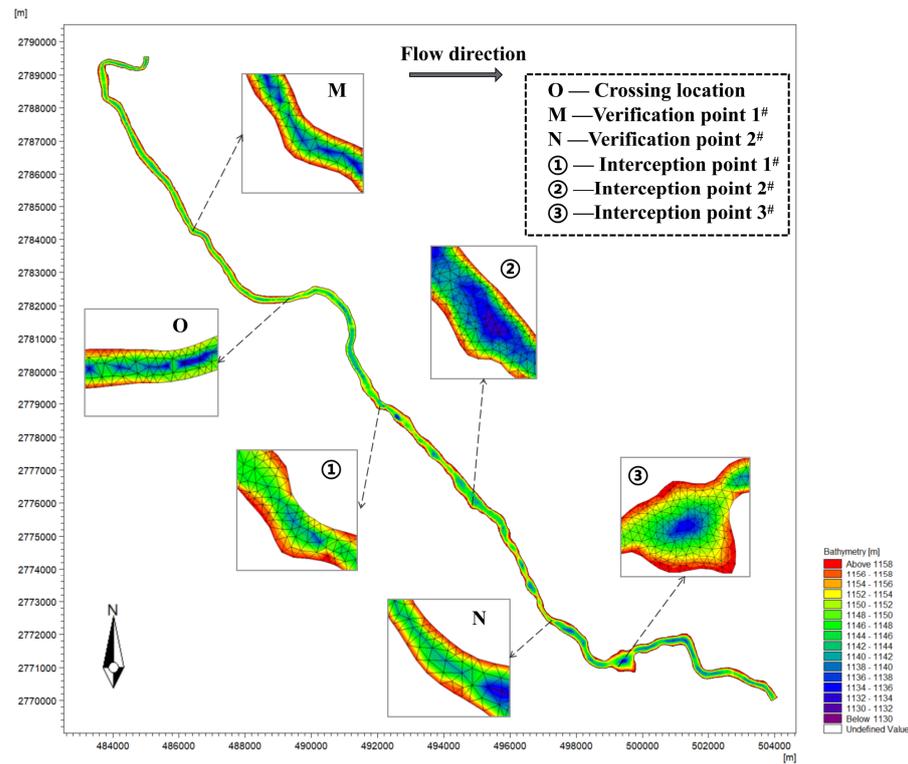


Figure 2. Overall topography and local grid map of the river channel.

2.3.2. Model Validation

To verify the reliability of the established hydrodynamic model, the measured hydrological data on 26 April and 5 August 2022 are selected for comparison with the model calculation results. The average flow rate on 26 April is $1210 \text{ m}^3/\text{s}$, and the average flow rate on 5 August is $5130 \text{ m}^3/\text{s}$. The validation results are shown in Figure 3, which shows that the simulated values of flow rate and water level at M, N are in good agreement with the measured values, and the mean absolute error (MAE) of the flow rate is 0.054 m/s , and that of the water level is 0.062 m . The results show that the established hydrodynamic model can be used for the oil spill in this river basin. The hydrodynamic model can provide reliable hydrodynamic conditions for the study of oil spill transport behavior in this basin.

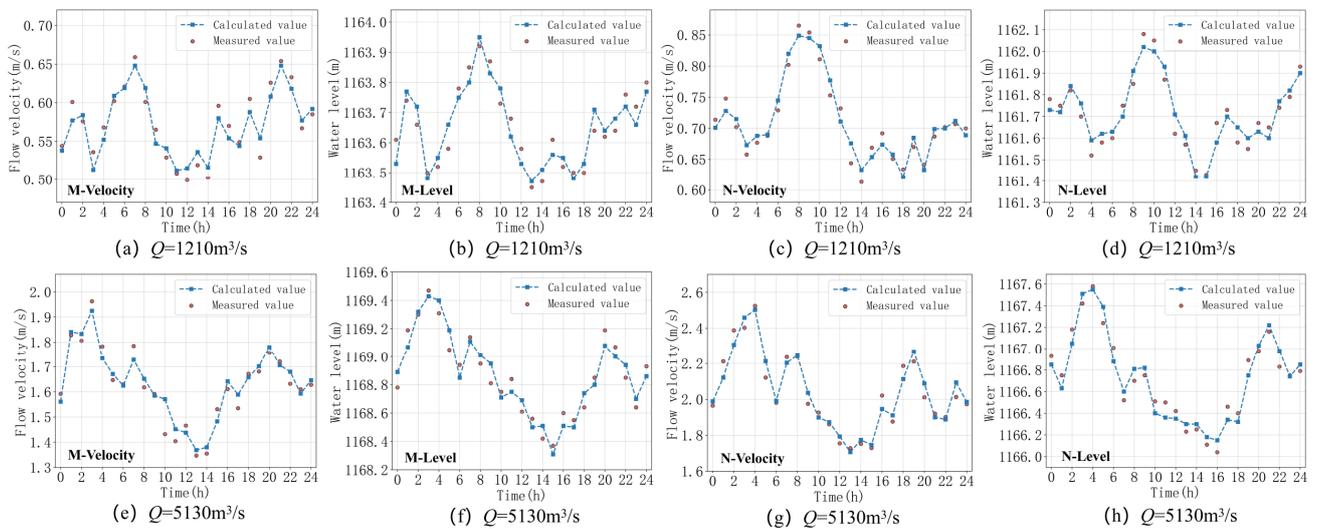


Figure 3. Flow velocity and water level validation.

3. Results and Discussions

3.1. Simulation

3.1.1. Simulation Scenario Setting

To study the transport behavior of oil spills in the basin of the Lancang River crossing pipeline section, three influencing factors, including river flow, wind (wind speed and direction), and leakage mode, were considered according to the hydrological and environmental conditions and the pipeline itself in the basin. The range of values for each of these parameters is given below:

1. River flow. As a typical monsoon river, the Lancang River has a large variation of flow under different water periods. According to the hydrological data of the basin, the value of the river flow in this study is selected between the minimum flow in the dry season and the maximum flow in the abundant season, ranging from $500 \text{ m}^3/\text{s}$ to $7000 \text{ m}^3/\text{s}$;
2. Wind, including wind speed and direction. Based on the wind data provided for 2021 and 2022, a rose wind map was produced, as shown in Figure 4. The average annual wind speed in the region is 2.1 m/s , and the main wind direction is southwest. The variables of wind speed were set from 1 m/s to 7 m/s , and the variables of wind direction were set in 8 directions;
3. Leakage modes. Leakage modes include fracture leakage and perforation leakage. According to the researched literature [42], when the pipeline section is fractured, from the beginning of the accident to the upstream and downstream valve closure, all the crude oil in the pipeline will be leaked into the Lancang River within 0.5 h ; when the pipeline is perforated and leaked, the maintenance personnel can complete the emergency blocking operation within 2 h . The amount of leakage under the fracture leakage mode was calculated to be 2600 t ; under the perforation leakage mode, the leakage hole size was set to be $5\%D$, $10\%D$, and $15\%D$ (D is the outer diameter of the pipe of 813 mm), and the amount of leakage was 180 t , 720 t , and 1630 t , respectively.

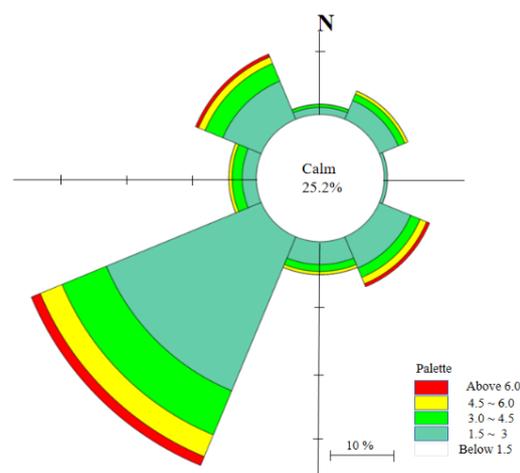


Figure 4. Rose wind map of the basin.

3.1.2. Simulation Results

Taking the spill scenario with a river flow of $5000 \text{ m}^3/\text{s}$, a south-west wind of 2.1 m/s , and a rupture spill (spill volume of 2600 t , spill time of 0.5 h) as an example, Figure 5 shows the state of oil spill transport at each time point. After the crude oil was spilled into the river, the oil spill continued to migrate downstream along the north bank of the river under the action of the water current and the southwest wind. In the early stage of the spill at $t = 1 \text{ h}$, the oil film is more aggregated, the area and length of the oil spill contamination zone are $A = 8.19 \times 10^4 \text{ km}^2$ and $L = 3.37 \text{ km}$, respectively, and the average thickness of the oil film is $T = 1.07 \text{ mm}$. In the migration process, the oil spill will diffuse and extend

continuously along the river channel under the hydrodynamic force of the river, the pulling force of the wind, and the obstructing force of the river bank. When $t = 5$ h, the oil spill head migrates to the tail of the simulated river section; at this time, the area and length of the oil spill pollution zone are $A = 77.41 \times 10^4 \text{ km}^2$ and $L = 10.11 \text{ km}$, respectively, and the average thickness of the oil film $T = 0.40 \text{ mm}$.

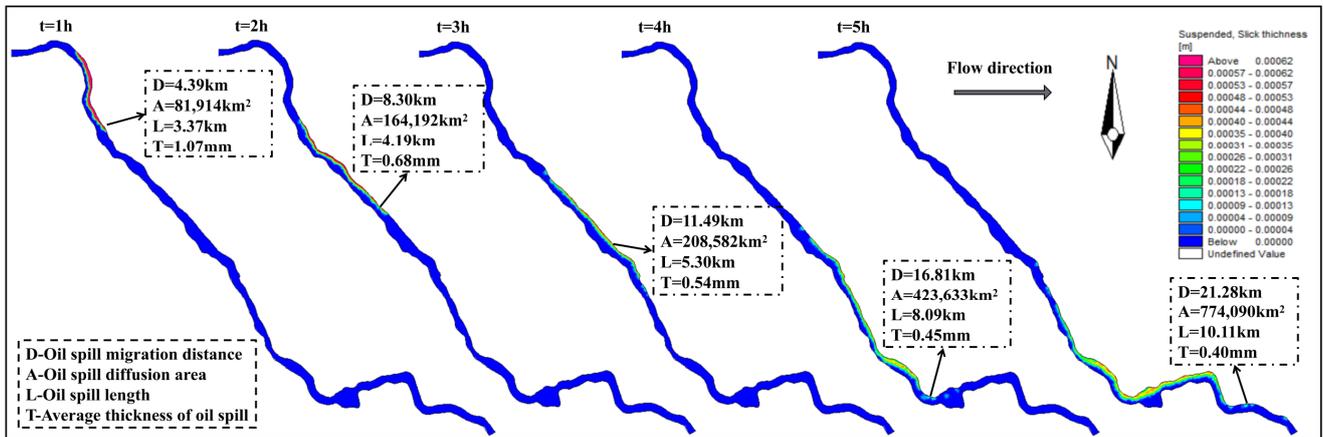


Figure 5. Oil spill transport state at each moment in time.

Figure 6 shows the variation of oil spill head migration distance (D), oil film area (A), oil spill contamination length (L), and average oil film thickness (T) with time (from the moment when the crude oil spill stops to the moment when the oil spill head flows out of the modeled river section). It can be seen that D, A, and L vary monotonically and incrementally with time, and the existence of fluctuations is caused by the irregularity of the river topography. T is monotonically decreasing with time. In the pre-spill period, the oil spill aggregation degree is large, and the easy diffusibility of the crude oil itself makes the oil film thickness decrease rapidly, and the decrease rate slows down obviously after 1 h.

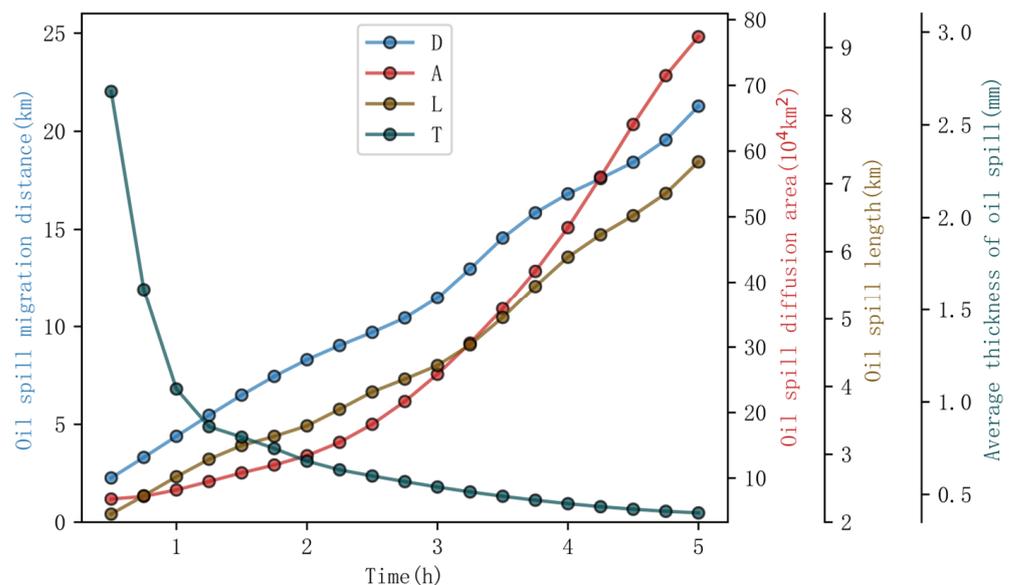


Figure 6. Variation of D, A, L, and T with time during the oil spill migration process.

3.2. Oil Spill Transport Behavior under Various Influencing Factors

The changes in D, A, L, and T are all time-dependent, with different spill scenarios having different times for the oil spill head to flow out of the modeled river section. In this paper, the average migration rate of oil spill head— v (m/s), the growth rate of oil

spill contaminated zone area per unit time— a ($10^4 \text{ km}^2/\text{h}$), and the growth rate of oil spill contaminated zone length per unit time— l (km/h) are used as the labels for analyzing the migration, diffusion and longitudinal extension behaviors of the oil spills under the various influencing factors. Because of the small degree of change in T in the middle and late periods, it is not considered in the analysis of the impact pattern.

3.2.1. River Flow

To study the effects of flow rate on oil spill behavior, the model only changes the flow rate data, and the wind and leakage parameters are kept fixed at 0 m/s and fracture leakage-2600 t (0.5 h), respectively. Figure 7 shows the variation of v , a , and l with river flow, with all three increasing with flow. As a carrier of oil spill migration, the higher the flow rate of the river, the faster the water velocity, the stronger the turbulence of the water flow, the faster the migration of the oil spill, and the more obvious the behavior of the oil spill in terms of diffusion and longitudinal extension. The increase of a is accelerated because the water level of the river will rise with the increase of flow; the increase of the basin area makes the oil spill diffusion space bigger, so the oil spill pollution area increases faster; l is slowed down, the increase of the flow, the water will overcome more resistance of the river bank to the migration of the oil spill, so the increase of the length of the oil spill pollution zone will be slowed down.

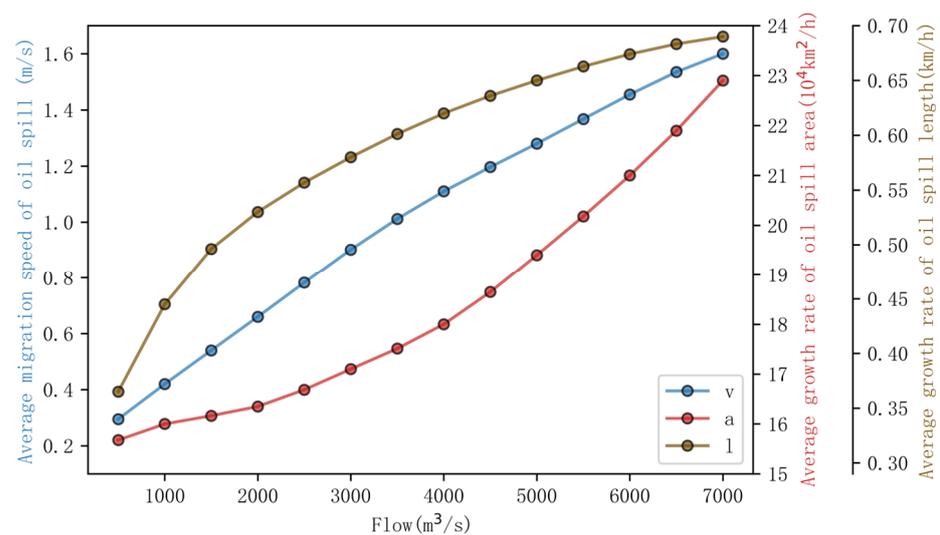


Figure 7. Variation of v , a , and l with river flow.

To take emergency response measures as soon as possible after the occurrence of a sudden oil spill accident, the pipeline's subsidiary units have set up three oil spill interception and recovery points along the river and stored the required emergency materials, the specific location of which is shown in Figure 2. After the accident, it would take three hours for emergency teams to reach the various interception points and commence interception operations. For sudden oil spill accidents under the condition of static wind, take the flow rate of $1000 \text{ m}^3/\text{s}$ and $5000 \text{ m}^3/\text{s}$ as an example; when the flow rate $Q = 1000 \text{ m}^3/\text{s}$, the time of oil spill migrating to the three oil spill interception points is 3.5 h, 6.5 h, and 15 h respectively; and when the $Q = 5000 \text{ m}^3/\text{s}$, it is 1.2 h, 2.2 h, and 3.6 h. Under the two flow rates, the rescue team should be rushed to the 1# and 3# interception points to carry out interception and recovery operations, respectively. In addition, the higher the river flow, the shorter it takes for the oil spill pollution zone to pass through the interception point. When $Q = 1000 \text{ m}^3/\text{s}$, the time for the oil spill pollution zone to pass through the 1# interception point is 1.3 h; when $Q = 5000 \text{ m}^3/\text{s}$, the time for the oil spill pollution zone to pass through the 3# interception point is 0.8 h. When the flow of the river is higher, the impact of the water on the boom is greater, and the interception efficiency of the boom will

be reduced, requiring more human and material resources to be invested in emergency operations. Therefore, the size of river flow directly affects the difficulty of emergency response to oil spill accidents. Managers should select oil spill interception and recovery points according to the river flow and reasonably equip personnel and emergency supplies.

3.2.2. Wind

The effects of wind on the oil spill behavior of rivers include two factors: wind speed and wind direction. Firstly, to explore the effects of wind speed, the wind direction was fixed as the main wind direction southwest in the basin, three river flow rates were set, 1000 m³/s, 3000 m³/s, 5000 m³/s, and the leakage parameter was fracture leakage-2600 t (0.5 h). Figure 8 shows the curves of v , a , and l as a function of wind speed for three river flows. In (a) and (b), the v and a both decrease continuously with the increase of wind speed. Under the southwest wind, the oil spill will continuously migrate and diffuse downstream along the upper boundary of the river channel. The higher the wind speed, the faster the oil spill accumulates to the upper boundary of the river bank, and the wind's pulling force makes the river bank more obstructive to the oil spill, thus limiting the migration and diffusion of the oil spill. In (b), the downward trend of a is faster at $Q = 1000$ m³/s because the smaller the river flow, the faster the oil spill accumulates upward to the upper boundary, and it also indicates that the smaller the river flow, the greater the effects of wind on the oil spill diffusion behavior. In (c), l increases and then decreases with increasing wind speed. The combined effect of water flow, wind, and riverbank obstruction causes the oil spill to extend longitudinally faster in the river channel compared to the no-wind condition, so the length of the oil spill contamination zone is longer. Although under the same flow rate, the larger the wind speed is, the longer the oil spill contamination zone is, the rate of change is related to the migration rate of the oil spill, and v decreases to a certain extent, l will also be smaller.

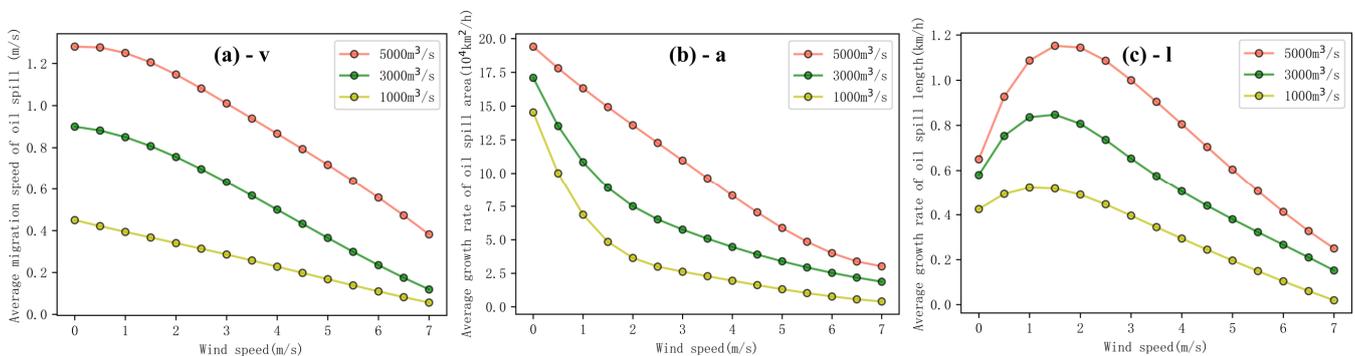


Figure 8. Variation of v , a , and l with wind speed.

Exploring the effects of wind direction, the wind speeds were fixed at 2.1 m/s and 5 m/s, the flow rate was 5000 m³/s, and the leakage parameter was 2600 t (0.5 h). Figure 9 shows v , a , and l for different wind directions. The modeled river section is oriented northwest to southeast. For the direction of oil spill transport, northwest winds can be defined as downwind; north, west, and southwest winds as partial downwind, and vice versa for headwind and partial headwind. In (a), (b), and (c), it can be seen that the values of v , a , and l are higher than those of the opposite wind direction for the same wind speed in downwind and partial downwind conditions, indicating that the migration, diffusion and longitudinal extension behaviors of oil spills are faster in downwind and partial downwind conditions; The reason that v , a under downwind and headwind is larger and l is smaller compared to partial downwind and partial headwind and does not change much with wind speed is that under downwind and headwind conditions, the oil spill accumulates to a lesser extent on the river bank side and is less obstructed by the river bank, so the oil spill migrates and diffuses faster and extends along the river channel more slowly. Combining the effects of wind speed and wind direction on the oil spill transport behavior under

downwind and headwind conditions, the effects of wind speed on the oil spill transport behavior are relatively small; under the rest of the wind directions, the migration and diffusion behaviors of the oil spill will slow down with the increase of wind speed, and the longitudinal extension behaviors of the oil spill will be accelerated firstly and then slowed down with the increase of wind speed.

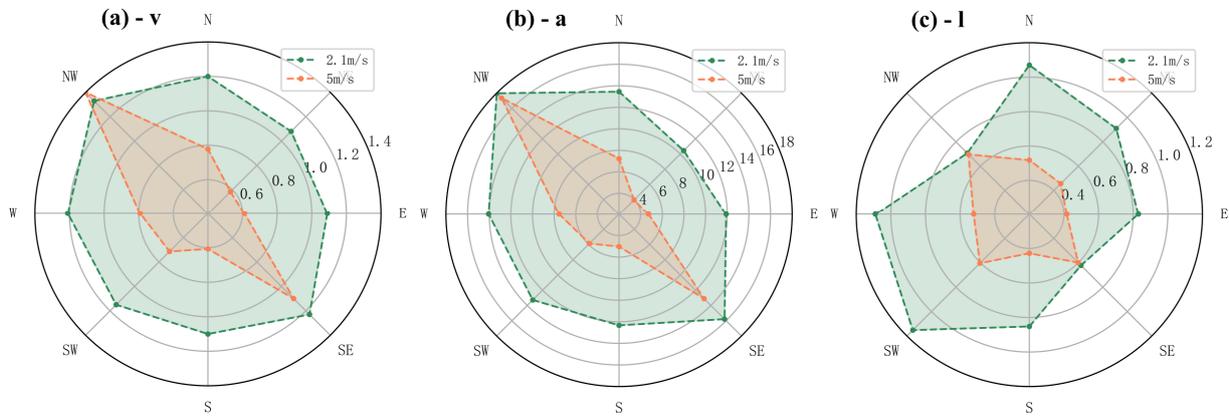


Figure 9. *v*, *a*, and *l* for different wind directions.

Through the above analysis, it can be seen that the wind will affect the migration speed and transport state of the oil spill. Taking the scenarios of $Q = 5000 \text{ m}^3/\text{s}$ and wind speed of 5 m/s with northeast, southwest, and northwest winds as an example, as shown in Figure 10, under northeast and southwest winds, the oil spill will migrate along the south and north banks of the river channel respectively, and the migration time to the 2# interception point will be 4.2 h and 3.8 h; under the northwest winds, the oil spill will migrate to the 3# interception point will be 3.2 h. Emergency personnel can select the interception and recovery point according to the migration speed of the oil spill under the current wind conditions and reasonably deploy personnel and equipment on both sides of the river according to the gathering orientation of the oil spill to improve the efficiency of interception and recovery of the oil spill.

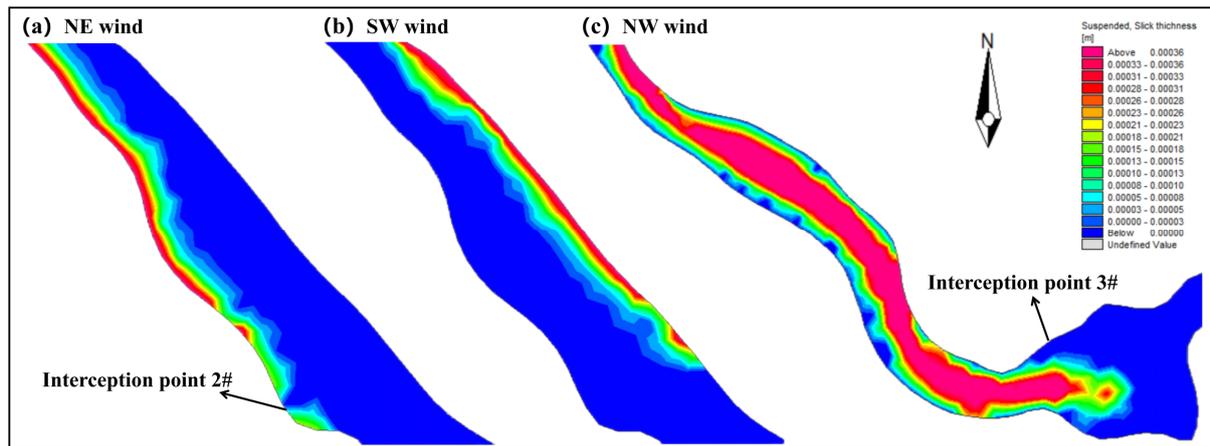


Figure 10. Transportation state of oil spill under NE and SW winds reaching the 2# interception point.

3.2.3. Leakage Mode

To explore the oil spill transport behavior under different leakage modes, three leakage volumes of 180 t (2 h), 720 t (2 h), and 1630 t (2 h) were assumed under perforated leakage mode, and 2600 t (0.5 h) under fracture leakage. The river flows are $1000 \text{ m}^3/\text{s}$, $3000 \text{ m}^3/\text{s}$, and $5000 \text{ m}^3/\text{s}$, and the wind speed is 0 m/s. Figure 11 shows *v*, *a*, and *l* for different leakage parameters, and it can be seen that the variation of *v* is not significant in both

leakage modes. Under perforated leakage, a and l increase with the increase of leakage volume. This indicates that the spill mode has less effect on the oil spill migration rate; under perforated leakage, the increase in oil spill volume causes the oil spill to diffuse and extend longitudinally in the river channel more quickly. In (b), the value of an under the scenario of perforated leakage-1630 t (2 h) is larger than that of fractured leakage-2600 (0.5 h) for $Q = 5000 \text{ m}^3/\text{s}$ and $3000 \text{ m}^3/\text{s}$ because, under perforated leakage, the initial dispersion of the oil spills is large. There is more space for the diffusion of the oil film, so the rate of diffusion of the oil spills is faster within a short period.

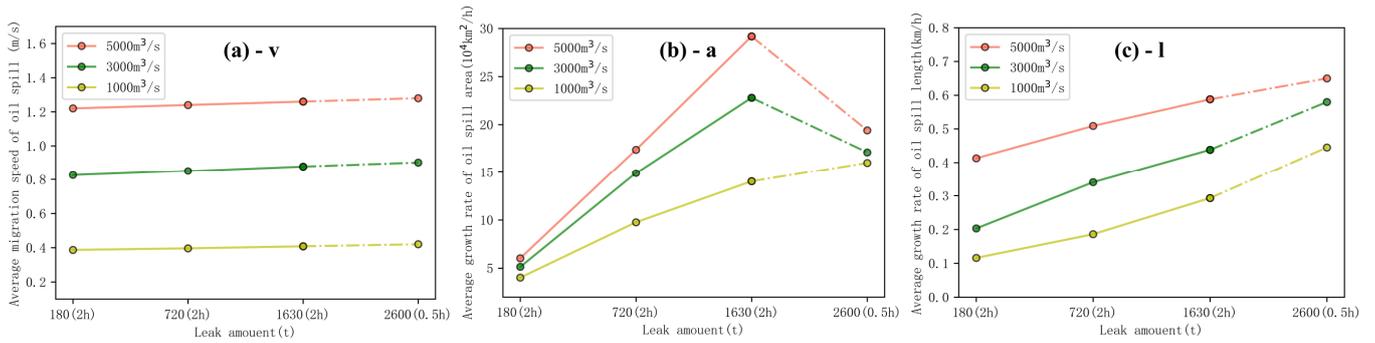


Figure 11. v , a , and l for different leakage parameters.

The difference in the timing of the oil spill in different spill modes results in different lengths of the oil spill contamination zone. For the oil spill accident under two kinds of leakage modes, under static wind conditions, with $Q = 3000 \text{ m}^3/\text{s}$, fracture leakage-2600 t (0.5 h) and perforation leakage-1630 t (2 h), for example, the head of the oil spillage migrates to 2# oil spill interception point after 3.3 h, and the distribution of oil spills in the river is shown in Figure 12. At this time, the length of the oil spill pollution zone is 1.97 km and 6.92 km, respectively. In the case of perforated leaks, emergency personnel can go to both 1# and 2# interception points to conduct interception operations. Therefore, emergency managers can develop emergency response plans for simultaneous interception and recovery operations at multiple interception points according to different spill modes.

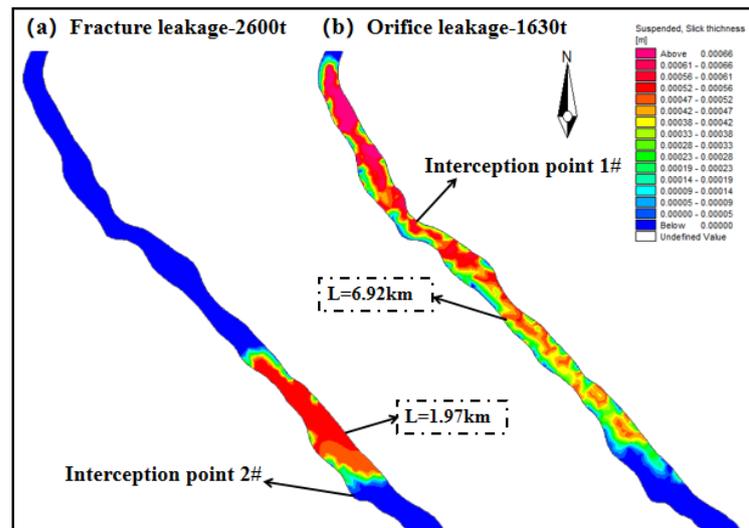


Figure 12. Transportation state of an oil spill reaching the 2# interception point under two leakage modes.

3.3. Prediction of Oil Spill Transport Behavior Based on Machine Learning

In the case of a sudden river oil spill, in order to intercept and recover the oil spill quickly and efficiently, it is necessary to accurately predict the location of the oil spill and

the extent of pollution. Traditional numerical simulation calculations are complex and time-consuming, which undoubtedly prolongs the emergency response time. Given the advantages of machine learning, such as fast computation speed, good model stability, and low dependence on physical mechanisms, this paper adopts the machine learning method to establish the oil spill transport state prediction model to guide the decision-making support of oil spill accidents.

In this paper, based on the value range of each variable in Section 3.1.1, a total of 320 sets of numerical tests were conducted, and eight times node data were recorded for each set of tests, which resulted in a total of 2560 sets of oil spill transport state data, and an oil spill transport state data set was established. Among them, the characteristic variables include a total of six dimensions: flow rate, wind speed, wind direction, leakage mode, leakage amount, and time; and the labels include the oil spill migration distance D , oil spill area A , and oil spill contamination zone length L , as shown in Table 1.

Table 1. Data set description table.

	Flow Rate (m ³ /s)	Wind Speed (m/s)	Wind Direction (°)	Leakage Mode	Leakage Amount (t)	Timing (h)	D (km)	A (10 ⁴ km ²)	L (km)
Mean	4380	2.1	-	-	1721	13.5	9.97	170.08	6.21
Min	500	0	0	0	180	0.5	0.81	4.29	0.81
25%	1500	1	135	0	720	4.5	4.99	47.65	2.31
50%	4000	2	225	1	1630	12	9.83	150.39	5.49
75%	5000	4.5	225	1	2600	33	15.94	226.38	9.45
Max	7000	7	315	1	2600	112	21.6	433.25	13.45

Because of the small amount of sample data, training with deep learning models is prone to lead to the problem of overfitting, which will instead impair the final model results [43]. Therefore, in this paper, traditional machine algorithms are used for predictive modeling. In this paper, using a posteriori thinking to select the optimal combination of algorithms, the dataset is divided into training and testing sets in the ratio of 8:2, and four typical learners, namely Random Forest (RF), Back Propagation Neural Network (BPNN), Extreme Gradient Boosting (XG Boost), and Support Vector Regression (SVR), are used for modeling and analysis [44–47], and the coefficient of determination (R^2) and 1-Mean Absolute Percentage Error ($1-MAPE$) are used as the model evaluation metrics to evaluate the prediction performance of the model. Table 2 shows the computation process of the four learners.

The modeling scheme is as follows:

Step 1: Four learners, RF, BPNN, XG Boost, and SVR, are used for modeling respectively, and the Random Search algorithm is applied to select the optimal hyper-parameters, analyze the prediction performance of different learners, and achieve the preliminary selection of learners.

Step 2: To further improve the prediction performance of the learner, an intelligent optimization algorithm is used instead of the traditional Random Search algorithm for the optimization of hyperparameters. The prediction performance of different combinations of machine learning algorithms is analyzed, and the optimal algorithm is selected to achieve the prediction model.

The predictions of the four training models are shown in Figure 13. For the three labels D , A , and L , the average R^2 of the four training models are 0.677, 0.873, 0.717, 0.880; the average $1-MAPE$ is 0.713, 0.860, 0.707, 0.863. The prediction performance of the BPNN and SVR is significantly better than that of the RF and the XG Boost, and so the BPNN and SVR are initially chosen as the benchmark. Therefore, BPNN and SVR are initially selected as the benchmark models.

Table 2. The computation process of the RF, XG Boost, BPNN, and SVR.

Arithmetic	Basic Principle and Computation Process	Hyperparameter
RF	<p>Classification or regression by constructing multiple decision trees and combining their results.</p> <p>Step 1: Random sampling.</p> <ul style="list-style-type: none"> The random sampling process is expressed as: $D_i = \{(x_{i1}, y_{i1}), (x_{i2}, y_{i2}), \dots, (x_{in}, y_{in})\}, i = 1, 2, \dots, T.$ <p>Step 2: Construct a decision tree.</p> <ul style="list-style-type: none"> Each decision tree is constructed through a decision tree algorithm, which includes steps such as feature selection, tree growth and pruning. <p>Step 3: Calculate the predictions (regression task averaged).</p> <ul style="list-style-type: none"> The final prediction is: $\hat{y} = \frac{1}{T} \sum_{i=1}^T y_i.$ 	<p>Number of trees, maximum depth of tree, minimum number of split nodes, minimum number of leaf nodes, random seeds, etc.</p>
XG Boost	<p>The model performance is gradually improved by serially training multiple decision trees and using gradient boosting.</p> <p>Step 1: Model initialization.</p> <ul style="list-style-type: none"> Initialize the model parameters, including the number of trees, tree depth, learning rate, etc. <p>Step 2: Construct the initialized prediction model.</p> <ul style="list-style-type: none"> A simple model (e.g., a tree containing only one leaf node) is used as the initial predictive model. <p>Step 3: Repeat until the maximum number of iterations is reached or the loss function converges.</p> <ul style="list-style-type: none"> Calculate the negative gradient (residual) of the loss function: $r_{ii} = -\frac{\partial}{\partial \hat{y}_{ii}} L(y_{ii}, \hat{y}_{ii}).$ A tree was fitted to represent the residuals of the model using the negative gradient as the target variable. Update the predictions of the model: $\hat{y}_{t+1,i} = \hat{y}_{ti} + \eta \cdot f_t(x_i).$ 	<p>Learning rate, number of trees, maximum depth of the tree, leaf nodes per tree, regularization parameters, sampling ratio, etc.</p>
BPNN	<p>The input signal is passed to the output layer through forward propagation, and then the network weights and biases are adjusted to minimize the prediction error using a back-propagation algorithm.</p> <p>Step 1: Model initialization.</p> <ul style="list-style-type: none"> Randomly initialize network parameters including weights and biases for multiple hidden layers. <p>Step 2: Repeat until convergence or the maximum number of iterations is reached.</p> <p>(1) Forward propagation</p> <ul style="list-style-type: none"> Compute the inputs and outputs of the 1st hidden layer: $z^{(1)} = W^{(1)}x + b^{(1)}; h^{(1)} = \sigma(z^{(1)}).$ Compute the inputs and outputs of each subsequent hidden layer: $z^{(l)} = W^{(l)}h^{(l-1)} + b^{(l)}; h^{(l)} = \sigma(z^{(l)}).$ Calculate the inputs and outputs of the output layer: $z^{(L)} = W^{(L)}h^{(L-1)} + b^{(L)}; h^{(L)} = \sigma(z^{(L)}).$ <p>(2) Back propagation</p> <ul style="list-style-type: none"> Calculate the output layer and each hidden layer error: $\delta^{(L)} = (y - o) \odot \sigma'(z^{(L)}); \delta^{(l)} = \sigma'(z^{(l)}) \odot (W^{(l+1)T} \delta^{(l+1)}).$ Update the weights and biases of the output layer and each hidden layer: $W_{new}^{(L)} = W^{(L)} + \eta \delta^{(L)} h^{(L-1)T}, b_{new}^{(L)} = b^{(L)} + \eta \delta^{(L)}; W_{new}^{(l)} = W^{(l)} + \eta \delta^{(l)} h^{(l-1)T}, b_{new}^{(l)} = b^{(l)} + \eta \delta^{(l)}$ 	<p>Number of hidden layers, number of neurons per layer, activation function type, learning rate, etc.</p>
SVR	<p>The data is fitted by finding the maximum interval in the feature space by means of a support vector machine.</p> <p>Step 1: Kernel function selection.</p> <ul style="list-style-type: none"> Select the kernel function to map the input features to the high dimensional space and acquire the mapped feature vector $\phi(x).$ <p>Step 2: Construct and solve the optimization problem.</p> <ul style="list-style-type: none"> The objective function of the optimization problem is: $Loss = \frac{1}{2} \ w\ ^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*);$ The constraints are: $\begin{cases} y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i \\ w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, n \end{cases} .$ <p>Step 3: Model predictions.</p> <ul style="list-style-type: none"> After the model training is completed, the new samples can be predicted by Equation After mapping the new samples to the high dimensional space using the optimized model parameters and, and kernel function: $f(x) = w^T \phi(x) + b.$ 	<p>Kernel functions, regularization parameters, $\varepsilon,$ etc.</p>

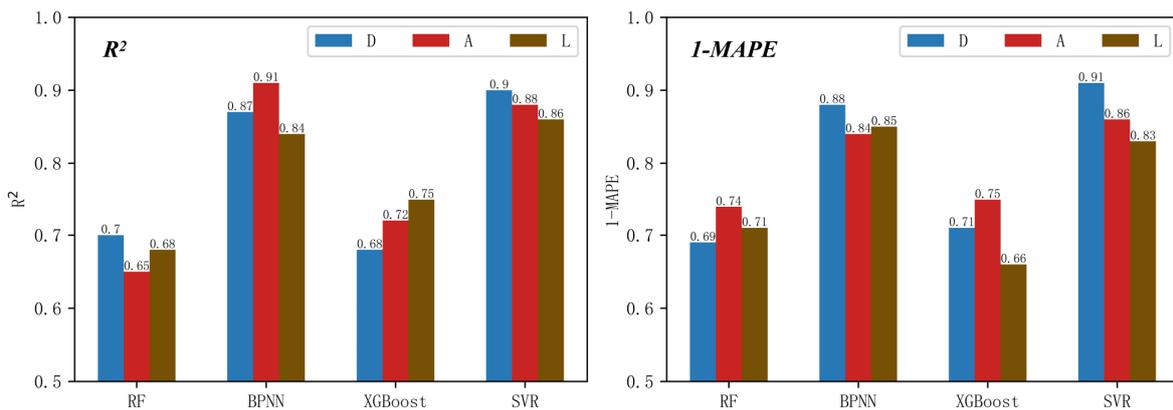


Figure 13. R² and 1-MAPE for the four prediction models.

The selection of hyperparameters in the above models all use the traditional Random Search algorithm, and the model accuracy is relatively low. To further improve the prediction performance of the learner, this paper adopts an intelligent optimization algorithm for hyperparameter preference. Considering the problems of local optimum and insufficient search capability, three intelligent optimization algorithms, Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Differential Evolution (DE), are introduced [48–50], which form a combination of six classes with BPNN and SVR respectively algorithms. Table 3 shows the computation of the three optimization algorithms. The prediction effect of the combined model is shown in Figure 14.

As can be seen from the figure, the prediction performance of the model is significantly improved after the hyperparameters are preferred by the intelligent optimization algorithm. For label D, PSO-SVR has the highest R² and 1-MAPE, 0.980 and 0.977, respectively. For label A, GA-BPNN has the highest R², and DE-SVR has the highest 1-MAPE. The R² and 1-MAPE of GA-BPNN are 0.982 and 0.971, respectively, and those of DE-SVR are 0.929 and 0.979, respectively. Combining the two model evaluation metrics of R² and 1-MAPE, the GA-BPNN has a predictive performance is better. For label L, PSO-SVR has the highest R², and PSO-BPNN has the highest 1-MAPE. The R² and 1-MAPE of PSO-SVR are 0.978 and 0.940, respectively, and those of PSO-BPNN are 0.972 and 0.975, respectively, so that PSO-BPNN has a better prediction performance. In summary, the PSO-SVR, GA-BPNN, and PSO-BPNN combination models were finally selected as the prediction models for the oil spill transport state in the Lancang River Crossing Pipeline Section Basin in this study, respectively. The differences between the predicted and true values of D, A, and L by the three combined models are shown in Figure 15.

Table 3. The computation process of the PSO, GA, DE.

Arithmetic	Basic Principle and Computation Process
PSO	<p>Simulating the behaviour of a group of organisms such as a flock of birds or a school of fish, each individual (particle) adapts to its own experience and information about its neighbours in the search space in order to find the optimal solution.</p> <p>Step 1: Initialise the position and velocity of the particle swarm.</p> <p>Step 2: Calculate the fitness of each particle (objective function value).</p> <p>Step 3: Update the velocity and position of each particle.</p> <ul style="list-style-type: none"> The velocity of the particles updates Eq: $v_{i,j}(t+1) = w \cdot v_{i,j}(t) + c_1 \cdot r_1 \cdot (pbest_{i,j}(t) - x_{i,j}(t)) + c_2 \cdot r_2 \cdot (gbest_j(t) - x_{i,j}(t))$ The position of the particle updates Eq: $x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1)$ <p>Step 4: Repeat the above steps until the stop condition is met.</p>

Table 3. Cont.

Arithmetic	Basic Principle and Computation Process
GA	<p>Simulating the process of biological evolution, populations evolve through operations such as selection, crossover and mutation to produce better individuals.</p> <p>Step 1: Initialise the population.</p> <p>Step 2: Evaluate the fitness of each individual.</p> <p>Step 3: Select the operation.</p> <ul style="list-style-type: none"> Parent individuals were selected based on fitness: $P(i) = \frac{f(i)}{\sum_{j=1}^N f(j)}$. <p>Step 4: Crossover operation.</p> <ul style="list-style-type: none"> Crossover was performed on the parent individuals to produce offspring individuals. $Offspring(i) = \begin{cases} Parent1(i) & \text{if } rand() < CrossoverRate \\ Parent2(i) & \text{otherwise} \end{cases}$ <p>Step 5: Mutation operation.</p> <ul style="list-style-type: none"> Variation on offspring individuals. $Offspring(i) = \begin{cases} 1 - Offspring(i) & \text{if } rand() < MutationRate \\ Offspring(i) & \text{otherwise} \end{cases}$ <p>Step 6: Repeat the above steps until the stop condition is met.</p>
DE	<p>The optimal solution is found by generating candidate solutions and updating them using a difference operation based on the results of the fitness evaluation.</p> <p>Step 1: Initialise the population.</p> <p>Step 2: Generate new candidate solutions based on the difference operation.</p> <ul style="list-style-type: none"> Three different individuals from the population are randomly selected as base vectors, and then a new candidate solution is generated using a difference operation: $x_{i,G+1} = x_{r1,G} + F \cdot (x_{r2,G} - x_{r3,G})$ <p>Step 3: Evaluate the fitness of the candidate solution.</p> <p>Step 4: Update the candidate solution based on the result of the fitness evaluation.</p> <p>Step 5: Repeat the above steps until the stopping condition is satisfied.</p>

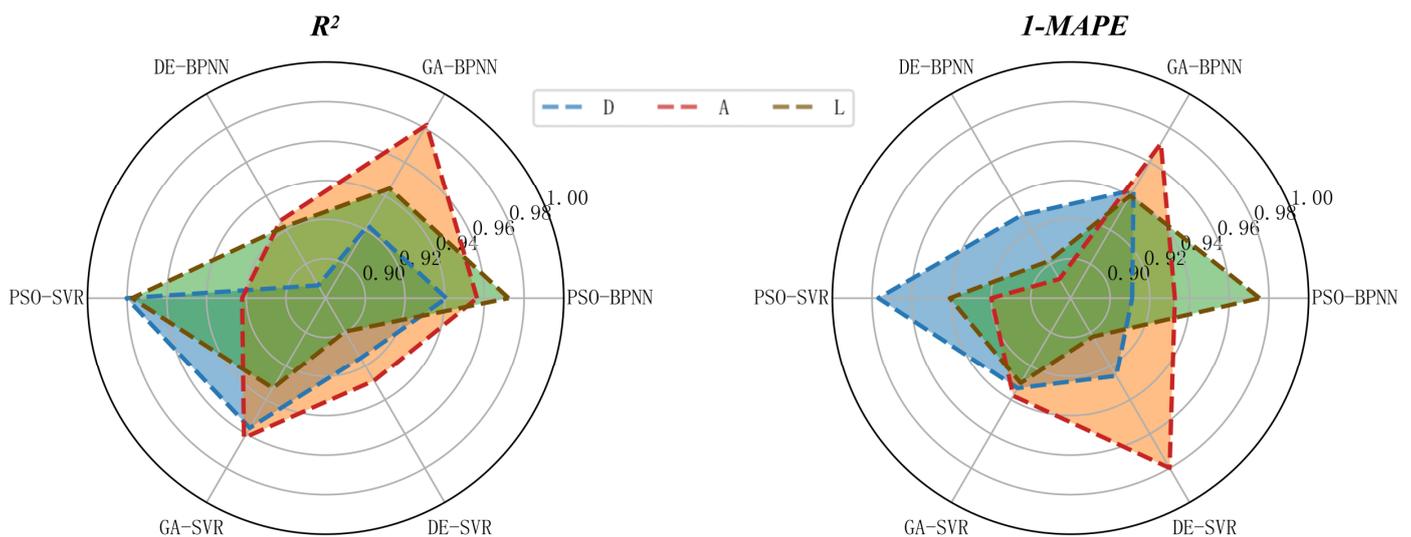


Figure 14. 6 R² and 1-MAPE for six combined models.

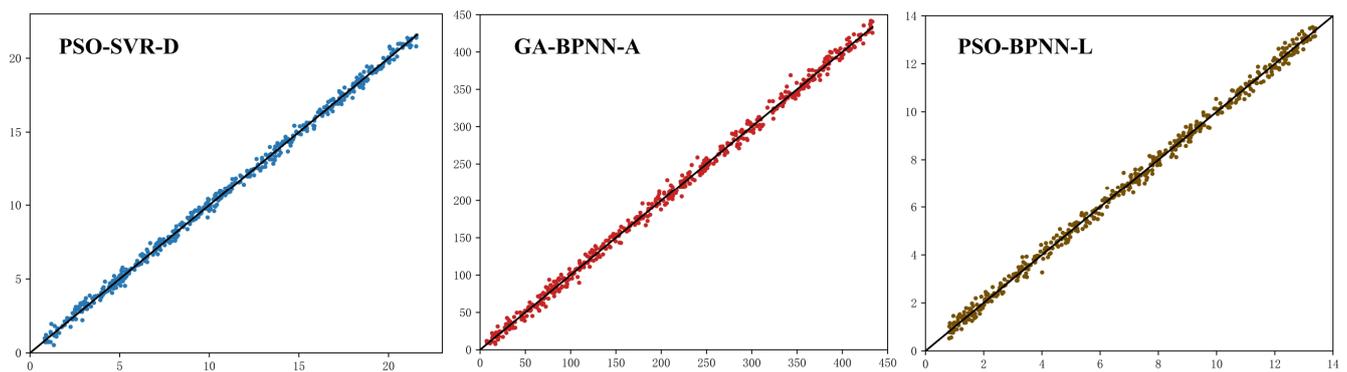


Figure 15. Difference between predicted and actual values of the three combined models.

4. Conclusions

In this study, a two-dimensional hydrodynamic model and an oil spill model of the Lancang River crude oil crossing pipeline basin were developed based on Mike 21. The oil spill transport behavior was simulated under different river flow rates, wind speeds and directions, and spill modes. The oil spill transport state dataset was established, and the oil spill transport state prediction model was built using a combined machine learning algorithm. The main conclusions are as follows:

- (1) For the two-dimensional hydrodynamic model of the Lancang River crude oil crossing pipeline basin, the flow velocity and water level of the model were verified according to the measured hydrological data. The *MAE* of the flow velocity is 0.054 m/s, and the *MAE* of the water level is 0.062 m. The simulated values are in good agreement with the actual values. It shows that the model is suitable for the study of oil spill transport behavior in this basin.
- (2) The transport behavior of oil spills under different spill scenarios was simulated. It was found that the higher the river flow, the faster the migration, diffusion, and longitudinal extension behavior of the oil spill; under downwind and headwind conditions, the effects of wind speed on the oil spill transport behavior are small. In the remaining wind directions, the migration and diffusion behaviors of the oil spill will slow down with the increase of wind speed, and the longitudinal extension behaviors of the oil spill will speed up firstly and then slow down with the increase of wind speed; the effects of leakage mode on the migration rate of the oil spill are small. Under the perforation leakage, the increase in oil spill volume will make the oil spill diffuse and extend longitudinally in the river channel faster. The simulation results can provide guidance for the development of emergency response plans for river crossing pipeline oil spill accidents.
- (3) An oil spill transport state dataset was established through 320 sets of numerical experiments. The oil spill transport state prediction model was established by machine learning combination algorithms. The three machine learning algorithms, PSO-SVR, GA-BPNN, and PSO-BPNN, have the best performance in predicting the oil spill migration distance, oil spill area, and oil spill contaminated zone length, respectively, with the R^2 and $1-MAPE$ above 0.971, and the accuracy of the prediction model is high. The model can provide support for rapid emergency response to sudden oil spill accidents of river crossing pipelines.

Author Contributions: Conceptualization, J.L.; Methodology, L.C.; Software, J.L.; Validation, J.L.; Formal analysis, D.X.; Data curation, D.X.; Writing—original draft, J.L.; Supervision, L.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets presented in this article are not readily available because the data is part of the subject being researched. Requests to access the datasets should be directed to 107901310@qq.com.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, J.; Zhang, W.; Wan, Z.; Li, S.; Huang, T.; Fei, Y. Oil spills from global tankers: Status review and future governance. *J. Clean. Prod.* **2019**, *227*, 20–32. [[CrossRef](#)]
2. Kvočka, D.; Žagar, D.; Banovec, P. A Review of River Oil Spill Modeling. *Water* **2021**, *13*, 1620. [[CrossRef](#)]
3. Kabył, A.; Yang, M.; Shah, D.; Ahmad, A. Bibliometric Analysis of Accidental Oil Spills in Ice-Infested Waters. *Int. J. Environ. Res. Public Health* **2022**, *19*, 15190. [[CrossRef](#)] [[PubMed](#)]
4. Yang, Y.-F.; Wang, S.; Zhu, Z.-D.; Jin, L.-Z. Prediction model and consequence analysis for riverine oil spills. *Front. Environ. Sci.* **2022**, *10*, 1054839. [[CrossRef](#)]
5. Liu, Z.; Chen, Q.; Zhang, Y.; Zheng, C.; Cai, B.; Liu, Y. Research on transport and weathering of oil spills in Jiaozhou Bight, China. *Reg. Stud. Mar. Sci.* **2022**, *51*, 102197. [[CrossRef](#)]
6. Lewis, D.; Sauzier, J. Cleaning up after Mauritius oil spill. *Nature* **2020**, *585*, 172. [[CrossRef](#)] [[PubMed](#)]
7. Chen, J.; Di, Z.; Shi, J.; Shu, Y.; Wan, Z.; Song, L.; Zhang, W. Marine oil spill pollution causes and governance: A case study of Sanchi tanker collision and explosion. *J. Clean. Prod.* **2020**, *273*, 122978. [[CrossRef](#)]
8. Aljaroudi, A.; Khan, F.; Akinturk, A.; Haddara, M.; Thodi, P. Risk assessment of offshore crude oil pipeline failure. *J. Loss Prev. Process Ind.* **2015**, *37*, 101–109. [[CrossRef](#)]
9. Aa, I.; Op, A.; Ujj, I.; Mt, B. A critical review of oil spills in the Niger Delta aquatic environment: Causes, impacts, and bioremediation assessment. *Environ. Monit. Assess.* **2022**, *194*, 816. [[CrossRef](#)] [[PubMed](#)]
10. Ganju, N.K.; Brush, M.J.; Rashleigh, B.; Aretxabaleta, A.L.; del Barrio, P.; Grear, J.S.; Harris, L.A.; Lake, S.J.; McCardell, G.; O'donnell, J.; et al. Progress and Challenges in Coupled Hydrodynamic-Ecological Estuarine Modeling. *Estuaries Coasts* **2016**, *39*, 311–332. [[CrossRef](#)]
11. Shen, Y.X.; Jiang, C.B. A comprehensive review of watershed flood simulation model. *Nat. Hazards* **2023**, *118*, 875–902. [[CrossRef](#)]
12. Munar, A.M.; Cavalcanti, J.R.; Bravo, J.M.; Fan, F.M.; da Motta-Marques, D.; Fragoso, C.R., Jr. Coupling large-scale hydrological and hydrodynamic modeling: Toward a better comprehension of watershed-shallow lake processes. *J. Hydrol.* **2018**, *564*, 424–441. [[CrossRef](#)]
13. De Goede, E.D. Historical overview of 2D and 3D hydrodynamic modelling of shallow water flows in the Netherlands. *Ocean Dyn.* **2020**, *70*, 521–539. [[CrossRef](#)]
14. Keramea, P.; Kokkos, N.; Zodiatis, G.; Sylaios, G. Modes of Operation and Forcing in Oil Spill Modeling: State-of-Art, Deficiencies and Challenges. *J. Mar. Sci. Eng.* **2023**, *11*, 1165. [[CrossRef](#)]
15. Spaulding, M.L. State of the art review and future directions in oil spill modeling. *Mar. Pollut. Bull.* **2017**, *115*, 7–19. [[CrossRef](#)] [[PubMed](#)]
16. Hou, X.; Hodges, B.R.; Feng, D.; Liu, Q. Uncertainty quantification and reliability assessment in operational oil spill forecast modeling system. *Mar. Pollut. Bull.* **2017**, *116*, 420–433. [[CrossRef](#)] [[PubMed](#)]
17. Azevedo, A.; Oliveira, A.; Fortunato, A.B.; Zhang, J.; Baptista, A.M. A cross-scale numerical modeling system for management support of oil spill accidents. *Mar. Pollut. Bull.* **2014**, *80*, 132–147. [[CrossRef](#)] [[PubMed](#)]
18. Kuang, C.; Chen, J.; Wang, J.; Qin, R.; Fan, J.; Zou, Q. Effect of Wind-Wave-Current Interaction on Oil Spill in the Yangtze River Estuary. *J. Mar. Sci. Eng.* **2023**, *11*, 494. [[CrossRef](#)]
19. Liu, D.; Mu, L.; Ha, S.; Wang, S.; Zhao, E. An innovative coupling technique for integrating oil spill prediction model with finite volume method-based ocean model. *Mar. Pollut. Bull.* **2022**, *185*, 114242. [[CrossRef](#)]
20. Zhen, Z.; Li, D.; Li, Y.; Chen, S.; Bu, S. Trajectory and weathering of oil spill in Daya bay, the South China sea. *Environ. Pollut.* **2020**, *267*, 115562. [[CrossRef](#)]
21. Periañez, R. A Lagrangian oil spill transport model for the Red Sea. *Ocean Eng.* **2020**, *217*, 107953. [[CrossRef](#)]
22. Bozkurtoglu, S.N.E. Modeling oil spill trajectory in Bosphorus for contingency planning. *Mar. Pollut. Bull.* **2017**, *123*, 57–72. [[CrossRef](#)] [[PubMed](#)]
23. Da Cunha, A.C.; De Abreu, C.H.M.; Crizanto, J.L.P.; Cunha, H.F.A.; Brito, A.U.; Pereira, N.N. Modeling pollutant dispersion scenarios in high vessel-traffic areas of the Lower Amazon River. *Mar. Pollut. Bull.* **2021**, *168*, 112404. [[CrossRef](#)]
24. Jiang, P.; Tong, S.; Wang, Y.; Xu, G. Modelling the oil spill transport in inland waterways based on experimental study. *Environ. Pollut.* **2021**, *284*, 117473. [[CrossRef](#)]
25. Wang, J.; Wang, S.; Zhu, Z.; Yang, Y.; Zhang, Q.; Xu, S.; Yan, J. Integrated framework for assessing the impact of inland oil spills on a river basin: Model and case study in China. *Ecol. Indic.* **2024**, *158*, 111576. [[CrossRef](#)]
26. Mendoza-Cantú, A.; Heydrich, S.C.; Cervantes, I.S.; Orozco, O.O. Identification of environmentally vulnerable areas with priority for prevention and management of pipeline crude oil spills. *J. Environ. Manag.* **2011**, *92*, 1706–1713. [[CrossRef](#)] [[PubMed](#)]

27. García, H.; Nieves, C.; Colonia, J.D. Integrity management program for geo-hazards in the ocesa pipeline system. In Proceedings of the ASME International Pipeline Conference, Calgary, AB, Canada, 27 September–1 October 2010; Volume 4, p. 415.
28. Zhao, X.B.; Xu, B.; Huang, X.L. *The Situation and Countermeasures of Oil Spill Control of Inland River Ships, 018(001)*; China Water Transp.: Wuhan, China, 2018; pp. 146–148.
29. Qi, Y.; Li, T.; Zhang, R.; Chen, Y. Interannual relationship between intensity of rainfall intraseasonal oscillation and summer-mean rainfall over Yangtze River Basin in eastern China. *Clim. Dyn.* **2019**, *53*, 3089–3108. [[CrossRef](#)]
30. Amir-Heidari, P.; Raie, M. Response planning for accidental oil spills in Persian Gulf: A decision support system (DSS) based on consequence modeling. *Mar. Pollut. Bull.* **2019**, *140*, 116–128. [[CrossRef](#)]
31. Najafizadegan, S.; Danesh-Yazdi, M. Variable-complexity machine learning models for large-scale oil spill detection: The case of Persian Gulf. *Mar. Pollut. Bull.* **2023**, *195*, 115459. [[CrossRef](#)] [[PubMed](#)]
32. Liu, D.R.; Li, Y.; Mu, L. Parameterization modeling for wind drift factor in oil spill drift trajectory simulation based on machine learning. *Front. Mar. Sci.* **2023**, *10*, 1222347. [[CrossRef](#)]
33. Burmakova, A.; Kalibatiene, D. Applying Fuzzy Inference and Machine Learning Methods for Prediction with a Small Dataset: A Case Study for Predicting the Consequences of Oil Spills on a Ground Environment. *Appl. Sci.* **2022**, *12*, 8252. [[CrossRef](#)]
34. MIKE, E. MIKE 21 & MIKE 3 Flow Model FM. Available online: https://scholar.google.com.hk/scholar?hl=zh-CN&as_sdt=0,5&q=37.%09MIKE,+E.+MIKE+21+&+MIKE+3+Flow+Model+FM&btnG= (accessed on 15 April 2024).
35. Pang, T.; Wang, X.; Nawaz, R.A.; Keefe, G.; Adekanmbi, T. Coastal erosion and climate change: A review on coastal-change process and modeling. *Ambio* **2023**, *52*, 2034–2052. [[CrossRef](#)] [[PubMed](#)]
36. Roe, P.L. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.* **1981**, *43*, 357–372. [[CrossRef](#)]
37. Jawahar, P.; Kamath, H. A high-resolution procedure for Euler and Navier–Stokes computations on unstructured grids. *J. Comput. Phys.* **2000**, *164*, 165–203. [[CrossRef](#)]
38. Darwish, M.S.; Moukalled, F. TVD schemes for unstructured grids. *Int. J. Heat Mass Transf.* **2003**, *46*, 599–611. [[CrossRef](#)]
39. Mackay, D.; Paterson, S.; Trudel, K. *A Mathematical Model of Oil Spill Behaviour: Environment Canada*; Environmental Protection Service, Environmental Impact Control Directorate, Environmental Emergency Branch, Research and Development Division: Beijing, China, 1980; Volume 38.
40. Reed, M. The physical fates component of the natural resource damage assessment model system. *Oil Chem. Pollut.* **1989**, *5*, 99–123. [[CrossRef](#)]
41. Xie, H.; Yapa, P.D.; Nakata, K. Modeling emulsification after an oil spill in the sea. *J. Mar. Syst.* **2007**, *68*, 489–506. [[CrossRef](#)]
42. Liu, G.; Zhang, W.; Wu, D.; Wang, X.; Feng, S.; Zhang, X.; Zhao, Y. Numerical simulation of crude oil leakage dispersion in the Lancang River crossing section of China-Myanmar crude oil pipeline. *Oil Gas Storage Transp.* **2021**, *40*, 96–106.
43. Kim, S.; Choi, Y.; Lee, M. Deep learning with support vector data description. *Neurocomputing* **2015**, *165*, 111–117. [[CrossRef](#)]
44. Ao, Y.; Li, H.; Zhu, L.; Ali, S.; Yang, Z. Identifying channel sand-body from multiple seismic attributes with an improved random forest algorithm. *J. Pet. Sci. Eng.* **2019**, *173*, 781–792. [[CrossRef](#)]
45. Wang, L.; Zeng, Y.; Chen, T. Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Syst. Appl.* **2015**, *42*, 855–863. [[CrossRef](#)]
46. Du, S.; Wang, J.; Wang, M.; Yang, J.; Zhang, C.; Zhao, Y.; Song, H. A systematic data-driven approach for production forecasting of coalbed methane incorporating deep learning and ensemble learning adapted to complex production patterns. *Energy* **2023**, *263*, 126121. [[CrossRef](#)]
47. Wang, Q.; Chen, D.; Li, M.; Li, S.; Wang, F.; Yang, Z.; Zhang, W.; Chen, S.; Yao, D. A novel method for petroleum and natural gas resource potential evaluation and prediction by support vector machines (SVM). *Appl. Energy* **2023**, *351*, 121836. [[CrossRef](#)]
48. Zhang, Y.; Pan, Z.; Wang, H.; Wang, J.; Zhao, Z.; Wang, F. Achieving wind power and photovoltaic power prediction: An intelligent prediction system based on a deep learning approach. *Energy* **2023**, *283*, 129005. [[CrossRef](#)]
49. Zhang, P.; Yin, Z.-Y.; Jin, Y.-F.; Chan, T.H.; Gao, F.-P. Intelligent modelling of clay compressibility using hybrid meta-heuristic and machine learning algorithms. *Geosci. Front.* **2021**, *12*, 441–452. [[CrossRef](#)]
50. Pant, M.; Zaheer, H.; Garcia-Hernandez, L.; Abraham, A. Differential Evolution: A review of more than two decades of research. *Eng. Appl. Artif. Intell.* **2020**, *90*, 103479.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.