

Imperial College London
Department of Civil and Environmental Engineering

Integrated Numerical Modeling and Data-Driven Techniques for Thermal Effluent Simulation in Coastal Waters

Nada Alsulaiman

Supervised by
Professor Matthew D. Piggott
Professor Maarten van Reeuwijk

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Civil and Environmental Engineering Research.

February 2025

Statement of Originality

I hereby declare that the work contained in this thesis is my own. Any external work or source material has been properly referenced and acknowledged.

Nada Alsulaiman,

1 February 2025.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

This thesis investigates the integration of numerical modelling and data-driven techniques to improve the simulation of thermal effluent discharged from coastal power plants into tidal waters. Uncertainties in key effluent characteristics, particularly the discharge flow rate and excess temperature relative to ambient water, pose significant challenges to accurate model representation, often resulting in errors in seawater temperature predictions. These challenges are further complicated by the dynamic nature of tidal systems, where strong advective and dispersive processes govern the transport and spread of discharged thermal plumes.

Three methodologies were integrated into the modelling framework to estimate discharge parameters and improve the accuracy of seawater temperature simulations. These methodologies include: (1) data assimilation (DA) using the Ensemble Kalman Filter (EnKF) for state estimation of seawater temperature; (2) Bayesian optimisation to calibrate hydrodynamic model parameters using temperature observations from in-situ profiles, thermal unmanned aerial vehicle (UAV) imagery, and satellite data; and (3) Large-Scale Particle Image Velocimetry (LSPIV) to integrate discharge rate estimates from optical UAV footage into the hydrodynamic model. Each approach was systematically evaluated for its effectiveness in reducing parameter uncertainties and improving model accuracy.

The data assimilation findings demonstrate that optimising data collection locations and their assimilation frequencies significantly reduces errors in simulated seawater temperature by maximising the spread and retention of DA adjustments across the modelled domain. Bayesian optimisation highlighted the critical role of spatial coverage in temperature observations on parameter estimation outcomes, demonstrating the importance of strategic data collection. LSPIV-derived discharge offers a viable, cost-effective method for estimating outflows from coastal power plants in tidal environments, with its integration into hydrodynamic models enhancing simulations and complementing calibration efforts.

Acknowledgements

I am deeply grateful to Professor Matthew Piggott and Professor Maarten van Reeuwijk for their unwavering support, guidance, and generosity in sharing their knowledge. Their mentorship allowed me the freedom to pursue ambitious ideas while providing the structure and insight needed to refine them into meaningful research. Their trust in my abilities has profoundly shaped my growth as a researcher, and I will always be grateful for their invaluable feedback and encouragement. I could not have asked for better supervisors.

I extend my sincere appreciation to my colleagues from the Coastal Management Program (CMP) at the Kuwait Institute for Scientific Research (KISR) for their support throughout this journey. In particular, I am deeply grateful to Dr. Yousef Alosairi, whose mentorship over the past decade has shaped my career into what it is today. Since joining the CMP, he has given me the opportunity to explore my research interests, treating me as a collaborator rather than just a junior researcher. His generosity in sharing his expertise and his unwavering encouragement have shaped my skills and confidence as a researcher, and for that, I will always be grateful.

I am also thankful to Mr. Ali Hammad, Mr. Khaled Alsayed, Mr. Fares Alamer, and Mr. Michael Ashok, whose dedication and patience made the extensive fieldwork for this research possible. Their support and willingness to be on board, even in the most challenging field conditions, made all the difference, and I truly appreciate their commitment.

Additionally, I would like to express my gratitude to Dr. Mohammed AlKhaldi, Ms. Dana Alhouti, and Dr. Alanoud Alragum for their unwavering support throughout my PhD. Their kindness and readiness to assist with anything I needed made a significant difference, and I am truly grateful for their help.

This research was funded through the Postgraduate Research Program at KISR, and I

sincerely appreciate the support that made this work possible.

I am fortunate to have friendships that have been a constant source of support throughout this journey, each marked by our weekly gatherings. James Tlhomole and Muyuan Liu made Fridays something to look forward to, with our coffee discussions sparking countless research ideas. Tom Burns, my office neighbour, always brought a smile to my face, and our interactions never failed to lift my spirits. Tuesdays were enriched by my meetings with the MvR Group, where engaging discussions continuously broadened my perspective and deepened my learning. And Sundays belonged to Eman Almudhaf, a friendship that has spanned years—from Kuwait to London, through different career paths, yet always anchored in unwavering support and companionship.

Lastly, I want to thank my friends and family in Kuwait for their continuous support throughout this journey. Their encouragement from afar has meant so much to me. Most importantly, I am profoundly grateful to my mother and father, whose love, sacrifices, and belief in me have made this achievement possible. More than mine, this accomplishment belongs to them.

Contents

1	Introduction	1
1.1	Motivation and Objectives	1
1.2	Thesis Structure	6
1.3	Publications	7
2	Data Assimilation for State Estimation	8
2.1	Introduction	10
2.2	Methods	14
2.2.1	2D Hydrodynamic Model	14
2.2.2	Data Assimilation	18
2.2.3	Observing System Simulation Experiments (OSSEs)	23
2.3	Results and Discussion	26
2.3.1	OSSE1: Assimilation Relative to Different Frequencies and Locations	26
2.3.2	OSSE2: Assimilation Based on the Ensemble-based Targeted Ob- servations Method	32
2.3.3	Limitations and Considerations	38
2.4	Conclusion	39
3	Bayesian Optimisation for Parameter Estimation	41
3.1	Introduction	43
3.2	Materials and Methods	45
3.2.1	Observational Data Sources	45
3.2.2	3D Hydrodynamic Model Setup	51
3.2.3	Bayesian Optimisation	56

3.2.4	Performance Metrics	62
3.2.5	Impact of Parameter Uncertainty and Data Selection on Calibration Results	62
3.2.6	Model Calibration Based on In Situ and Remotely Sensed Obser- vations	68
3.3	Critical Assessment of Estimated Discharge and Excess Temperature Values	79
3.3.1	Evaluation of Estimated Discharge	80
3.3.2	Evaluation of Estimated Excess Temperature	81
3.4	Limitations and Considerations	82
3.5	Conclusion	83
4	Large-Scale Particle Image Velocimetry for Discharge Estimation	86
4.1	Introduction	88
4.2	Materials and Methods	93
4.2.1	Observational Data	93
4.2.2	PIVlab Analysis	97
4.2.3	Discharge Estimation	101
4.2.4	Data Processing For LSPIV Validation	104
4.2.5	Workflow for LSPIV Investigation and Model Integration	105
4.3	Results and Discussion	109
4.3.1	Impact of Image Sampling Frequency and Interrogation Area on LSPIV Results	109
4.3.2	Integration of LSPIV-Derived Discharge into a Hydrodynamic Model	117
4.4	Conclusion	127
5	Conclusions	130
5.1	Summary of Thesis Achievements	130
5.2	Synthesis and Critical Evaluation of Results	132
5.3	Future Directions	134
5.3.1	Proposed Observing System for Monitoring Thermal Effluents and Bay Dynamics	134
5.3.2	Research Directions	135

Bibliography	139
Appendix A Hydrodynamic Model Overview	156
A.1 Governing Equations and Assumptions	156
A.1.1 Continuity Equation	156
A.1.2 Momentum Equations in Horizontal Direction	157
A.1.3 Vertical Velocities	158
A.1.4 Hydrostatic Pressure Assumption	158
A.1.5 Equations of State for the Density	159
A.1.6 Transport of Matter	159
A.1.7 Heat Balance	160
A.2 Kuwait Bay (KB) Model Setup	161
A.2.1 Overview of Model Configuration	161
A.2.2 Boundary and Atmospheric Forcing Inputs	161
A.2.3 Model Initialisation and Spin-Up Procedure	163
A.2.4 Representation of Surface Heat Fluxes in 2D Simulations	164
A.3 Sulai bekhat Bay (SB) Model	166
A.3.1 2D Model Setup	166
A.3.2 3D Model Setup	169
Appendix B Processing of Landsat TIR Scenes	174
Appendix C Selection of Parameters for Optimisation	176
Appendix D Estimation of Cooling Water Discharge	185
D.1 Thermal Efficiency and Waste Heat Calculation	185
D.2 Estimating Cooling Water Discharge	186
D.3 Application to the Doha West Power Plant	187

List of Tables

3.1	Acquisition times and details for datasets used for parameter estimation. .	46
3.2	Optimal discharge rate (Q) and excess temperature (ΔT) from calibration using three data sources.	70
3.3	RMSE values for baseline and calibrated models evaluated against in situ measurements.	74
4.1	Details of UAV footage for LSPIV analysis.	93
4.2	Summary of ADCP datasets for LSPIV analysis.	95
4.3	Summary of LSPIV processing parameters and their impact on displace- ment, data points, and computation time.	110
4.4	Impact of LSPIV sampling configurations on velocity and discharge estimates.	118
4.5	Model results summary:(Q), (ΔT), parameter sources, and RMSE values. .	123
5.1	Comparison of methods for estimating discharge rate (Q) and excess efflu- ent temperature (ΔT). The strengths and limitations of each method are also highlighted.	133
A.1	Kuwait Bay (KB) Model Setup Parameters	161
A.2	Sulaibekhat Bay (KB) 2D Model Setup Parameters	166
A.3	Sulaibekhat Bay (KB) 3D Model Setup Parameters	169
C.1	Summary of parameters and bounds used for sensitivity analysis	178

List of Figures

1.1	Study area showing Kuwait Bay, Sulaibikhat Bay, and seasonal thermal plume intensities near Doha West Station.	4
2.1	Two-dimensional modelling setup showing bathymetry, stations, and computational grids for KB and SB.	16
2.2	Validation of SB truth model using tidal and temperature comparisons. . .	18
2.3	Twin experiment setup showing the average temperature for truth and baseline models, along with variability in discharge and excess temperature. . .	23
2.4	Thermal plume interaction with a semidiurnal tidal cycle over 25 hours. . .	27
2.5	Synthetic temperature observations from the truth model for P2 and corresponding tidal phases.	28
2.6	OSSE1 results showing error reduction, observation correlations, and baseline RMSE.	29
2.7	Observation site ranking and their effectiveness in reducing the trace of $\mathbf{P}^a$. . .	34
2.8	Variance at PT1, PT2, and PT3 compared to observation error variance. . .	35
2.9	Comparison of error reduction, variance, and covariance in OSSE1 and OSSE2.	37
3.1	Spatial coverage of satellite, vertical profile, and UAV data in the study area. . .	47
3.2	Vertical temperature structure along transects from the discharge point. . .	49
3.3	Temperature offset corrections for UAV and satellite-derived temperatures. . .	51
3.4	Numerical grid of the KB model with key features and bathymetry.	52
3.5	Water level validation of the 3D KB model against KISR01 observations. . .	52
3.6	Grid resolution of the SB model with refinement near the discharge location. . .	54
3.7	Measurements for boundary conditions: temperature and salinity at KISR01. . .	55

3.8	Ensemble standard deviation for calibration parameters.	65
3.9	GP posterior surfaces for calibration parameters.	67
3.10	GP objective functions for calibrating discharge rate and excess temperature using three data sources.	70
3.11	Observed vs. simulated vertical temperature profiles across datasets and calibration models.	72
3.12	Thermal and optical satellite scenes comparison for datasets 1 and 3. . . .	75
3.13	Temperature differences between satellite-derived and simulated SST for different models with RMSE values.	76
3.14	Temperature differences between UAV-derived and simulated SST for different models with RMSE values.	78
3.15	RMSE vs Discharge (Q) for calibrated models across datasets.	80
4.1	UAV image of the outfall with ADCP-derived velocity vectors and calibration reference.	95
4.2	Interpolated bathymetry of the outfall region relative to high tide.	96
4.3	Estimating α in QRev from ADCP velocity profile.	103
4.4	Histograms and box plots of velocity magnitudes and directions for Dataset A.	105
4.5	Workflow for LSPIV analysis and integration with the hydrodynamic model.	108
4.6	LSPIV velocity fields at varying sampling frequencies and interrogation areas compared to ADCP measurements.	110
4.7	Comparison of ADCP-measured and LSPIV-derived velocities under varying processing parameters.	111
4.8	Cross-correlation matrices and optical displacement results for varying LSPIV sampling frequencies.	112
4.9	Velocity magnitude differences for varying sampling frequencies and interrogation area configurations during LSPIV analysis.	115
4.10	Velocity magnitude differences between IA configurations.	116
4.11	Thermal effluent analysis at the outfall with temperature, velocity, and incremental discharge along a transect.	119

4.12 Observed vs. simulated vertical temperature profiles across LSPIV configurations.	122
4.13 Observed vs. simulated vertical temperature profiles across datasets and models.	125
A.1 Forcing data at KB model's open boundary.	162
A.2 Forcing data for the heat exchange module.	162
A.3 Spin-up evaluation of the KB model	163
A.4 Effect of surface heat fluxes in 2D simulations	165
A.5 Forcing at the open boundary of 2D SB	167
A.6 SB model spin-up at station P2	168
A.7 Effect of surface heat flux in 2D SB model	169
A.8 Water level forcing for the SB model in 3D.	170
A.9 Temperature and salinity forcing for the SB model in 3D.	171
A.10 Validation of vertical temperature profiles using two boundary forcing approaches	172
A.11 Spin-up procedure and stability of the 3D SB model	173
C.1 Temperature sensitivity to Q and ΔT	180
C.2 Temperature sensitivity to horizontal and vertical eddy diffusivity	181
C.3 Temperature sensitivity to Dalton and Stanton numbers	182
C.4 Temperature sensitivity to bathymetry variations	183
C.5 Temperature sensitivity to Chezy roughness values	184

List of Acronyms

ADCP	Acoustic Doppler Current Profiler
ASV	Autonomous Surface Vehicle
DA	Data Assimilation
DCC	Discrete Cross-Correlation
DFT	Discrete Fourier Transform
DW	Doha West
EnKF	Ensemble Kalman Filter
FPS	Frames Per Second
GP	Gaussian Process
IA	Interrogation Area
KB	Kuwait Bay
KF	Kalman Filter
LHS	Latin Hypercube Sampling
LSPIV	Large-Scale Particle Image Velocimetry
OSSE	Observational System Simulation Experiment
PIV	Particle Image Velocimetry
RMSE	Root Mean Square Error
SSE	Sum of Squared Errors

SST	Sea Surface Temperature
TIR	Thermal Infrared
TIRS	Thermal Infrared Sensor
SB	Sulaibikhat Bay
UAV	Unmanned Aerial Vehicle

Chapter 1

Introduction

1.1 Motivation and Objectives

The rapid growth of global urbanisation has driven the development of power plants to meet the increasing demand for electricity. As this sectors continue to expand, the need for cooling water to absorb heat during operation also grows. In coastal regions, cooling water is typically drawn from the sea and returned at an elevated temperature to the same source, as water bodies are considered as an infinite heat sink (Morelissen et al., 2016; Punetha, 2018). However, the discharge of this heated water, known as thermal effluent, raises significant concerns about its environmental and operational impacts.

From an environmental standpoint, improper selection of discharge locations can pose significant threats to marine ecosystems, particularly in poorly flushed areas such as semi-enclosed water bodies, where heat can easily accumulate (Roberts et al., 2010; Kong et al., 2022). Thermal pollution, defined as the addition of excess heat to natural waters, can disrupt natural seasonal temperature cycles, alter the living conditions for species, and affect their survival (Hinrichs and Kleinbach, 2012; Ma et al., 2017). Moreover, elevated temperatures can reduce dissolved oxygen levels, hinder nutrient transport within the water column, promote algal growth, and increase the release of nitrogen and phosphorus from sediments (Shawky et al., 2013; Ma et al., 2017; Moossa et al., 2022).

From an operational perspective, selecting appropriate discharge locations is critical to minimising the risk of cooling water recirculation from the outfall back to the intake. If the

discharged water is drawn back into the system before it has had sufficient time to mix with the ambient water body, it can lead to elevated intake temperatures. This recirculation not only reduces plant efficiency but also increases the likelihood of operational issues, such as the proliferation of jellyfish or harmful algal blooms, which can clog intakes and disrupt production (Morelissen et al., 2016; Salehi, 2017; Lincoln et al., 2021).

Given the significant impacts of thermal effluents, models designed to simulate their dispersion and transport are valuable tools for understanding and managing their impacts. Building these models requires an understanding of several key factors, including discharge characteristics (e.g., flow rate and effluent excess temperature), receiving water properties, discharge-receiving water interactions (e.g., mixing driven by buoyancy, turbulence, and jet entrainment), and model characteristics, such as the dispersion field (e.g., near or far field), key approximations, and parametrisation used to represent physical processes (Punetha, 2018).

Accurately defining these components can be challenging due to two primary issues: the inherent variability in discharge characteristics and significant limitations in the availability of data related to thermal coastal effluents. Variability in discharge characteristics is influenced by several factors, particularly the operational capacities of power plants. Studies have shown a positive correlation between the increase in temperature due to cooling water discharge and the installed capacity of power plants (Lin et al., 2021; Zhang et al., 2022; Wang et al., 2024a). Furthermore, different types of power plants (e.g., coal, gas-fired, nuclear) exhibit varying cooling water requirements, reflecting their specific operational processes (Morelissen et al., 2016). Seasonal fluctuations in electricity generation add another layer of complexity. For example, in 2022, the Doha West (DW) station in Kuwait, a co-generation power and desalination plant, reported a peak electricity generation of 1,271.8 GWh in August and a low of 749.4 GWh in February, reflecting a 69.7% increase during the summer months due to higher consumption demands (Ministry of Electricity and Water, 2023). These seasonal variations are reflected in thermal satellite imagery of DW station, which shows notable differences in plume intensity between winter (Figure 1.1c) and summer (Figure 1.1d). This seasonal variability is primarily driven by Kuwait's harsh summer climate and the increased use of air conditioning, with electricity demand being two to three times higher than in winter (Alotaibi, 2011). This

causes the power plants to operate at their maximum load with a low thermal efficiency due to the high outdoor temperature, which in turn results in higher fuel consumption during the summer period (Alotaibi, 2011). Consequently, the additional heat generated at maximum load likely increases the demand for cooling water to dissipate this heat.

Adding to these challenges is the scarcity of reliable and comprehensive datasets on coastal effluents, which are critical for both modelling and monitoring efforts. Limited access to such datasets significantly hinders the ability to conduct long-term environmental impact assessments (Elhakeem and Elshorbagy, 2015). Additionally, much of the existing literature provides minimal quantitative data and often lacks sufficient detail in study design and statistical analyses, further complicating the interpretation of results (Roberts et al., 2010). Publicly available temperature data, such as source water temperatures and the difference between inlet and outlet cooling water, further restricts the ability to conduct estimates of cooling water discharge and develop effective management strategies (Ko and Kim, 2024).

To address these challenges, this study explores methods to improve the accuracy of thermal effluent modelling by addressing the variability in discharge characteristics and the limitations posed by the scarcity of data on these effluents. DW station, with its distinctive environmental and operational context, provides a valuable case study to investigate these issues and develop strategies applicable to similar shallow, tidal environments.

DW is situated in the Arabian/Persian Gulf (hereafter referred to as the Gulf), a hotspot for desalination activities (Lattemann and Höpner, 2008). While this study does not address the excess salinity of brine, it is important to acknowledge the elevated temperature of the brine relative to background levels in receiving waters, an aspect often overlooked in studies on desalination processes (Roberts et al., 2010). Multi-stage Flash (MSF) distillation, as used at DW, and other thermal distillation methods are the most common in the Gulf, producing discharges with temperatures ranging from 5 to 15°C above ambient water temperatures (Lattemann and Höpner, 2008).

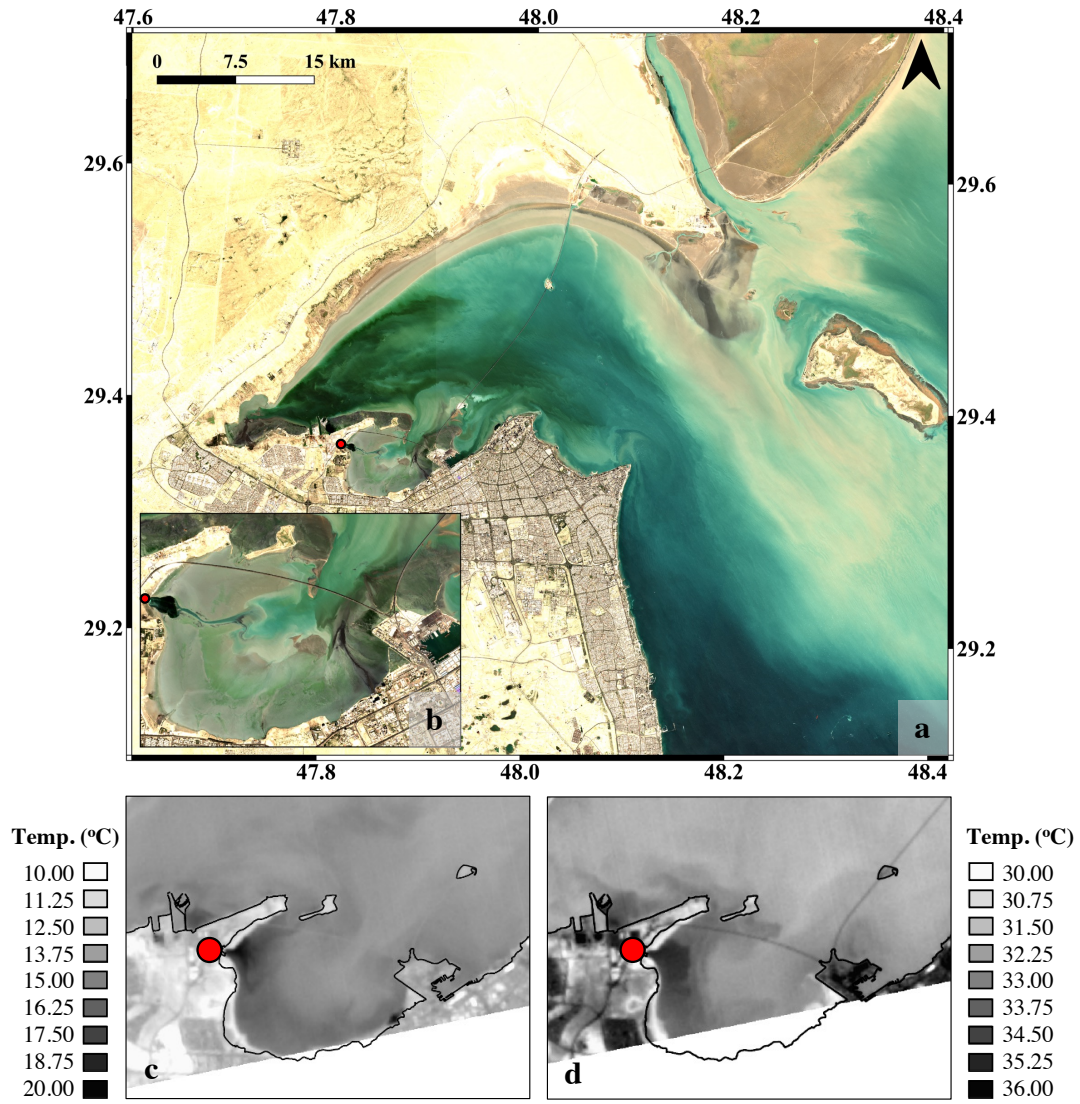


Figure 1.1: Study area showing Kuwait Bay (a) and a zoomed-in view of Sulaibikhat Bay (b), where the red circle indicates the location of Doha West Station. Panel (c) depicts the thermal plume intensity near Doha West station during high tide on 16th February 2022, representing the winter season with minimum power consumption. Panel (d) shows the thermal plume intensity during high tide on 24th June 2022, representing the summer season with maximum power consumption. (Images processed from Landsat 8 satellite – Band 10).

DW is strategically positioned in Sulaibikhat Bay (SB) (Figure 1.1b), a small embayment in the southern part of Kuwait Bay (KB) (Figure 1.1a), located in the northwest corner of the Gulf. SB spans approximately 46 km^2 and features a gradually sloping bottom with extensive intertidal mudflats (Wright et al., 1996). Both KB and SB are of immense biological importance, serving as critical habitats for shrimp, fish, and various migratory

bird species (Yamani et al., 2004; Al-Nafisi et al., 2009).

This region also faces significant environmental challenges, including rising sea surface temperatures. In July 2019, KB recorded an unprecedented sea surface temperature of 37.6°C, which led to significant fish kill incidents (Alosairi et al., 2020). Additionally, KB hosts four major power and desalination plants, including DW, which continuously discharges into SB (Rakha et al., 2009).

Combined with SB’s dynamic tidal regime, DW’s continuous discharge makes the bay an ideal location for evaluating methods to address and mitigate uncertainties in modelling industrial thermal effluents in tidal waters. Although this study is applied to DW, the findings are intended to serve as a general guideline and practical framework for modellers working with similar shallow, tidal environments or addressing comparable challenges.

In recent decades, advancements in computational resources and the expansion of global observational networks have significantly advanced the field of geophysical fluid modelling. However, the nonlinear nature of these systems, combined with the uniqueness of each geophysical event, continues to pose significant challenges to the modelling process. Particularly in operational forecast models, complexities arise from limited data, including the initial state of the system, numerical inaccuracies, and uncertainties in physical representations and parametrisations. As a result, the integration of observational data into numerical models has become essential (Brasseur, 2006; Blum et al., 2009; Hoteit et al., 2018).

Motivated by the unprecedented availability of observational data and advancements in computational resources, this PhD dissertation aims to integrate numerical models and data-driven approaches to address and mitigate inaccuracies arising from uncertainties in modelling thermal effluents in coastal waters. By leveraging the complementary strengths of numerical and data-driven methods, this research evaluates their practicality, applicability, and limitations, providing clear guidance on selecting the most effective method based on available resources.

To achieve this goal, the objectives of the thesis are as follows:

- Explore the integration of data-driven methods with traditional numerical modelling to address uncertainties in thermal effluent simulations, with a focus on utilising

diverse observational data sources such as in situ measurements, unmanned aerial vehicle (UAV) observations, and satellite imagery.

- Evaluate the strengths, limitations, and feasibility of data-driven approaches when combined with numerical models to improve thermal plume predictions under varying data availability.
- Develop transferable methodologies for integrating observational data into thermal effluent modelling, with findings applicable to similar shallow, tidal environments globally.

1.2 Thesis Structure

To achieve these objectives, the thesis is structured to systematically address the outlined goals, as follows:

- Chapter 2 explores the use of data assimilation (DA) techniques for state estimation of seawater temperature, assessing their ability to mitigate simulation discrepancies caused by uncertainties in discharge descriptions within the model. It also investigates the optimal assimilation frequency and observation placement to enhance DA performance.
- Chapter 3 implements Bayesian optimisation to calibrate discharge parameters in the model using three multi-sourced temperature observations: in situ vertical profile measurements, thermal UAV imagery, and thermal satellite imagery. An extensive field campaign, conducted to collect data simultaneously, facilitated the evaluation of the strengths and limitations of each data source in their effectiveness for parameter estimation.
- Chapter 4 applies Large-Scale Particle Image Velocimetry (LSPIV) as a cost-effective technique for deriving surface velocity and discharge rates from UAV footage. This chapter examines how variations in LSPIV analysis parameters influence the derived results. It also evaluates the potential improvements gained by integrating LSPIV-derived discharge estimations into the model.
- Chapter 5 concludes the thesis by summarising the key findings and contributions,

as well as providing recommendations for future work.

In each chapter, data-driven methods are integrated with a numerical hydrodynamic model of SB to demonstrate their effectiveness in mitigating uncertainties in modelling the effluents of the DW station. To the best of our knowledge, this research represents the first integration of data assimilation, Bayesian optimisation, and LSPIV techniques for modelling thermal discharges from coastal plants. We hope the innovations in this work advance the use of data-driven approaches in thermal effluent modelling and provide a foundation for future studies in similar contexts.

1.3 Publications

Parts of this thesis consist of works that have been submitted for publication or are in preparation for submission.

Chapter 2

- **N. Alsulaiman**, M. Van Reeuwijk, and M. D. Piggott. *Impact of Data Assimilation Frequency and Observation Location in Thermal Effluent Modeling for Coastal Waters*. Under review at *Earth and Space Science*, 2024.

Chapter 3

- **N. Alsulaiman**, Y. Alosairi, J. Tlhomole, M. Liu, M. Van Reeuwijk, and M. D. Piggott. *Bayesian Parameter Estimation of Coastal Thermal Effluents Using In Situ, UAV, and Satellite Observations*. In preparation for submission to *Environmental Modelling & Software*, 2025.

Chapter 4

- **N. Alsulaiman**, J. Tlhomole, Y. Alosairi, M. Van Reeuwijk, and M. D. Piggott. *UAV-Based Large-Scale Particle Image Velocimetry for Optimizing Coastal Thermal Effluent Modeling*. In preparation for submission to *Water Resources Research*, 2025.

Chapter 2

Data Assimilation for State Estimation

Highlights

- The Ensemble Kalman Filter was applied to address the uncertainties associated with the modelling of thermal effluents in coastal waters.
- Optimal frequency is characterised by its ability to overcome tidal advection and dispersion processes, without causing filter divergence.
- Optimal observation sites have both strong covariances with other state elements and high variance that exceeds prescribed observation error.

Summary

This chapter investigates the application of data assimilation (DA) using the Ensemble Kalman Filter (EnKF) to address the uncertainties associated with modelling the thermal effluents discharged from power plants into a shallow, tidal bay. A two-dimensional hydrodynamic model of Sulaibikhat Bay (SB), Kuwait, was developed using the Delft3D Flexible Mesh Suite to simulate the transport and dispersion of a thermal plume under the forces of predominantly semidiurnal tides. A series of observing system simulation experiments (OSSEs) were conducted using the OpenDA toolbox to identify the optimal

observation location and DA frequency. Significant reductions in temperature prediction errors were achieved using the EnKF for state estimation. The optimal DA frequency was characterised by its ability to retain the analysis adjustments in the system while maintaining computational efficiency. Relative to the dynamics of SB, the optimal frequency was found to be hourly. An excessive rate of DA was found to cause filter divergence, where forecast error is falsely underestimated due to diminishing ensemble variance. This led the filter to ignore the information provided by the observations. In contrast, a sparse rate of DA was found to cause the model to revert to its pre-assimilative state. The optimal station locations were identified using the ensemble-based targeted observation method, based on their ability to maximise the reduction of the analysis error variance. The optimal locations were attributed to having a balance between exhibiting strong covariances with the other state elements while experiencing local variance exceeding that of the prescribed observation error variance.

Publications and Disclaimer

The content within this chapter is based on:

N. Alsulaiman, M. Van Reeuwijk, and M. D. Piggott. *Impact of Data Assimilation Frequency and Observation Location in Thermal Effluent Modeling for Coastal Waters*. Under review at *Earth and Space Science*, 2024.

2.1 Introduction

In weather and oceanic forecasting, data assimilation (DA) is a widely used tool that effectively combines numerical model simulations with observational data to reach the best possible description of a dynamic system and its uncertainties. Now widely used across different research fields, it has evolved into a variety of methods that help improve the estimate of state variables, parameters, or both (Evensen et al., 2022). DA techniques can generally be placed under two categories based on how the observational data are blended into the system. Methods can either be classified as sequential or variational.

Sequential methods, often referred to as filters or statistical methods, can be described in two main steps: a propagation step, where the model’s previous state estimation and uncertainties are used to create an initial forecast, and an update step (i.e., analysis), where observations are recursively incorporated to adjust the state estimate up to the time of the analysis. The update step (also known as the analysis step) combines the model forecast with observations to produce an improved estimate of the system’s state (McLaughlin, 2002; Jamet and Loisel, 2009; Alvarado-Montero et al., 2022). These methods work within a probabilistic framework that aims to minimise the variance and maximise the posterior probability density function (PDF) whenever a new observation is made available (Khanarmuei, 2021). The Ensemble Kalman Filter (EnKF), the technique used in this chapter, is a popular example of sequential DA. It represents error statistics with an ensemble of perturbed model states to capture variability and quantify uncertainty in the system (Evensen, 2003).

In variational methods, commonly known as smoothers or nonsequential methods, the assimilation of the observations occurs over a predefined window of time. In the analysis step, a cost function that quantifies the difference between the predicted state and the observed state is needed to find the “optimal” state estimate, which should achieve the smallest difference between what is predicted and what is observed. The minimisation of the cost function requires the use of an adjoint model to compute the gradient of the cost function with respect to the state vector. Examples of variational methods include 3D-Var and 4D-Var methods (Rabier and Liu, 2003; Khanarmuei, 2021; Alvarado-Montero et al., 2022; Wang et al., 2023a).

To effectively utilise DA techniques, careful consideration must be given to the design of the observational network, as the success of assimilation depends not only on the method used but also on the timing and placement of observations. During the design process of an observational network for use in DA, the primary objective of monitoring needs to be clearly stated prior to deployment, as an observational network that is intended to monitor one aspect of ocean circulation might not be fit to monitor another (Oke and O’Kane, 2011). To fulfil the objectives of monitoring, the question of where and when to deploy and assimilate observations must be addressed.

Targeted (or adapted) observations are a class of monitoring systems that are intended to guide the design process through minimising the uncertainties in forecast models (Majumdar, 2016; Wu et al., 2020). A common approach to designing targeted observations is the use of ensemble-based methods (Bishop et al., 2001). These methods work by analysing the ensemble-based covariances of the system to identify the observations which can reduce the variance in the system’s analysis the most (Oke and O’Kane, 2011). In this chapter, the “objective array design method” of Sakov and Oke (2008) was selected due to its straightforward implementation and computational efficiency. The method is effective for designing stationary arrays aimed at long-term ocean deployments, as it uses an ensemble of system states from extended model runs or observation-based time series. Consequently, the reliability of the resulting design hinges greatly on how well the ensemble represents the system. The method has been successfully applied in a wide range of oceanic monitoring studies, including mooring array design (Liu et al., 2018) and the assessment of future and existing in situ stations (Geng et al., 2020). Notably, Wu et al. (2020) have extended this method to account for the flow-dependent variation in background error statistics and apply covariance localisation to suppress the influence of spurious correlations. Li et al. (2023) have introduced a projection vector to the method to develop a region-optional element to the design process. This enabled the identification of optimal locations within a specific area of interest. Rao et al. (2023) also introduced an algorithm that included a projection operator to enable a ‘cross-region’ and ‘cross-variable’ aspect to the design.

Identifying the optimal data collection frequency is considered to be another governing factor for the success of DA, yet, guidelines for the selection of the appropriate frequency

are still very limited in the literature (He et al., 2020). Numerical simulations show that update steps are temporarily retained in the system due to the gradual diminishing impact of the initial states over time (Berthet et al., 2009). In coastal environments, tide, wind, and buoyant forces create rapidly varying conditions that might hinder the retention time of the analysis, as a result, assimilation frequency must be selected with care. The variability of the observed parameters, like salinity and temperature, are often tied to the tidal cycle. As such, studies like Lai et al. (2021) and Liu et al. (2021b) suggest a dynamics-oriented assimilation attuned to local physics. While large assimilation intervals may overlook important hydrodynamic information, excessively short time steps lead to filter divergence and high computational costs (Khanarmuei et al., 2021). Filter divergence occurs when the ensemble variance lacks sufficient spread to adequately capture the uncertainty in the forecast error. This leads the filter to disregard the information provided by the observations in favour of the model’s own forecast (Whitaker and Hamill, 2012; Gottwald and Majda, 2013). An excessive rate of assimilation can cause the ensemble members to reduce their spread. The optimal DA cycle length must be aligned with the fundamental dynamical scales resolved by the numerical model and the observational system (Sandery, 2018).

This chapter will evaluate the performance of DA with a series of observing system simulation experiments (OSSEs) through the configuration of a numerical twin experiment. OSSEs are a class of numerical experiments in which synthetic (simulated) observations are assimilated within a DA framework to provide a cost-effective evaluation of observational network designs (Oke and O’Kane, 2011; Halliwell et al., 2014; Li et al., 2019). They build upon the concept of observing system experiments (OSEs), or data denial experiments, which assess the impact of observations by withholding them from assimilation, using synthetic observations from high-resolution simulations that serve as a proxy for the “true” ocean state (Halliwell et al., 2014). In a twin experiment setup, there are four key steps. First, a nature run (NR) is produced by integrating the model over a period of interest to simulate the “true” oceanic state. The NR is then used to generate synthetic observations that mimic the configuration of an observational network. Next, a control run is conducted using a model with a sufficiently different configuration from the NR, ensuring that the control run does not converge toward the NR during the in-

tegration period. In the third step, the synthetic observations are assimilated into the control run. Finally, the accuracy of the assimilative run is compared against the NR to assess the impact of DA on the simulation (Halliwell et al., 2014; Li et al., 2019; Yu et al., 2019). Throughout this chapter, the nature run is referred to as the “truth model”, and the control run is referred to as the “baseline model”. Twin experiments provide a testing ground without the need for real-world observations. They can be used to enhance our understanding of a wide range of spatiotemporal scales and allow us to conduct method comparisons, which are impractical to implement in real scenarios due to complex parameters and tuning needs (Loos et al., 2020).

The integration of observational data in coastal and ocean modelling to enhance state estimation has been well documented in the literature. DA has improved tidal simulations by adjusting open boundary conditions through the assimilation of harmonic constants derived from observed tides and currents (Shi et al., 2020), as well as by estimating bathymetry and bottom friction parameters (Wang et al., 2022). Operational forecasting models have also achieved improvements from the assimilation of observations from tidal gauge stations (Sørensen et al., 2004; Byrne et al., 2023). The enhancement introduced to the system by DA can extend beyond the assimilated variable. Liu et al. (2021b) reported that the assimilation of temperature and salinity profiles in a dynamic coastal environment reduced biases in water exchange simulations.

Despite the extensive application of DA in coastal modelling, its application to address uncertainties in thermal effluent modelling from coastal plants remains largely unexplored. Fluctuations in effluent volumes further complicate discharge parametrisation, introducing variability into simulation results. While several studies have combined observational measurements with numerical models to simulate thermal plume behaviour under various environmental conditions, such as seasonal changes or extreme events (Gaeta et al., 2020; Laguna-Zarate et al., 2021; Wang et al., 2024b), assess discharge compliance with environmental regulations (Lowe et al., 2009), or evaluate the design of both existing and planned plants to minimise recirculation risks and reduce operational costs (Salehi, 2017; Shaghaghi et al., 2024), none have utilised these methods within a DA framework.

This chapter aims to address two fundamental questions: First, can DA effectively reduce the uncertainties associated with modelling thermal effluents in tidal waters? Second,

what are the optimal observation locations and assimilation frequencies for tidal environments? To answer these questions, we conduct a series of OSSEs. The first experiment evaluates the impact of different assimilation rates and locations on model performance, while the second employs ensemble-based targeted observations to determine the most effective observational network design.

2.2 Methods

2.2.1 2D Hydrodynamic Model

The Delft3D Flexible Mesh Suite (D-Flow FM) is an integrated modelling software tool capable of simulating the unsteady, depth-averaged (2DH) or three-dimensional (3D) flow and transport phenomena that are governed by tides, wind, and pressure gradients due to free surface or density variations (Deltares, 2023). The model also allows for the inclusion of source and sink terms representing water discharge and extraction processes. D-Flow FM is ideal for studies of shallow waters, coastal areas, and estuaries, where horizontal scales significantly surpass the vertical scales (Deltares, 2023). The model’s governing equations are derived from the Navier-Stokes equations for incompressible free surface flow, under the Boussinesq, and in the case of 2DH, the shallow water assumptions. The hydrostatic pressure assumption is applied for the vertical momentum terms. D-Flow FM solves this comprehensive set of partial differential equations on an unstructured finite volume grid. For detailed model equations, approximations, and assumptions, the reader is referred to the Delft3D Flexible Mesh Suite manual (Deltares, 2023) and Appendix A.1.

In D-Flow FM, the thermal effluents discharged from a coastal plant are incorporated as source terms, quantified by either a constant or time-varying discharge rate and excess temperature added to the ambient waters. The thermal discharge is introduced into the model as a localised change in water volume and heat content. The discharge magnitude is included as a source term in the continuity equation to represent the water that is being added to the system. Simultaneously, the excess temperature is handled by an advection-diffusion equation, ensuring that it is transported and dispersed through the water body in a physically realistic manner. The effect of the excess temperature is coupled to the momentum equation through the influence of temperature on density as described by the

equation of state.

The use of the Delft3D modelling system, both structured and flexible mesh versions, for environmental studies on brine and thermal discharge has been documented well in the literature (Sana, 2015; Morelissen et al., 2016; Laguna-Zarate et al., 2021). Relevant to the scope of our study, Aljohani et al. (2022) demonstrated the use of the model to assess impacts associated with combined thermal and brine discharge from desalination plants. Salehi (2017) used the model to improve the design of intake and outfall structures to minimise the occurrence of thermal recirculation of cooling water. Chow et al. (2019) and Dols (2019) investigated the impacts of brine discharge from desalination plants on the Gulf’s salinity.

Careful consideration was given to the design of the SB hydrodynamic model, which serves as the primary model for the OSSEs, as it needed to balance hydrodynamic accuracy with computational efficiency during the execution of a large volume of ensemble runs. The modelling process began with the construction of a two-dimensional hydrodynamic model covering KB, as shown in Figure 2.1. The KB model was forced at the open boundary using water levels, temperature, and salinity measurements recorded at 10-minute intervals from the Salmiya station. The model accounted for heat exchange at the surface by considering the effects of solar and atmospheric radiation, as well as heat loss due to back radiation, evaporation, and convection. Meteorological data measured at 10-minute intervals—including wind, relative humidity, air temperature, and solar radiation—were obtained from the KISR01 station and applied uniformly across the KB domain. Bathymetry data were sourced from the Navionics Marine Charts and supplemented with data from recent survey campaigns for areas near the outfall, where navigational chart data were unavailable. Wetting and drying processes were included in the model to account for the extensive tidal flats.

A smaller model of SB was then constructed with a compact grid of approximately 340-meter resolution. The boundary forcings for the SB model, outlined in Figure 2.1, were extracted from the KB model simulation output. This ensured that the SB model remained connected to the larger bay system while maintaining computational efficiency for the EnKF application. Details on the KB and SB model setups are provided in Appendices A.2 and A.3.1, respectively.

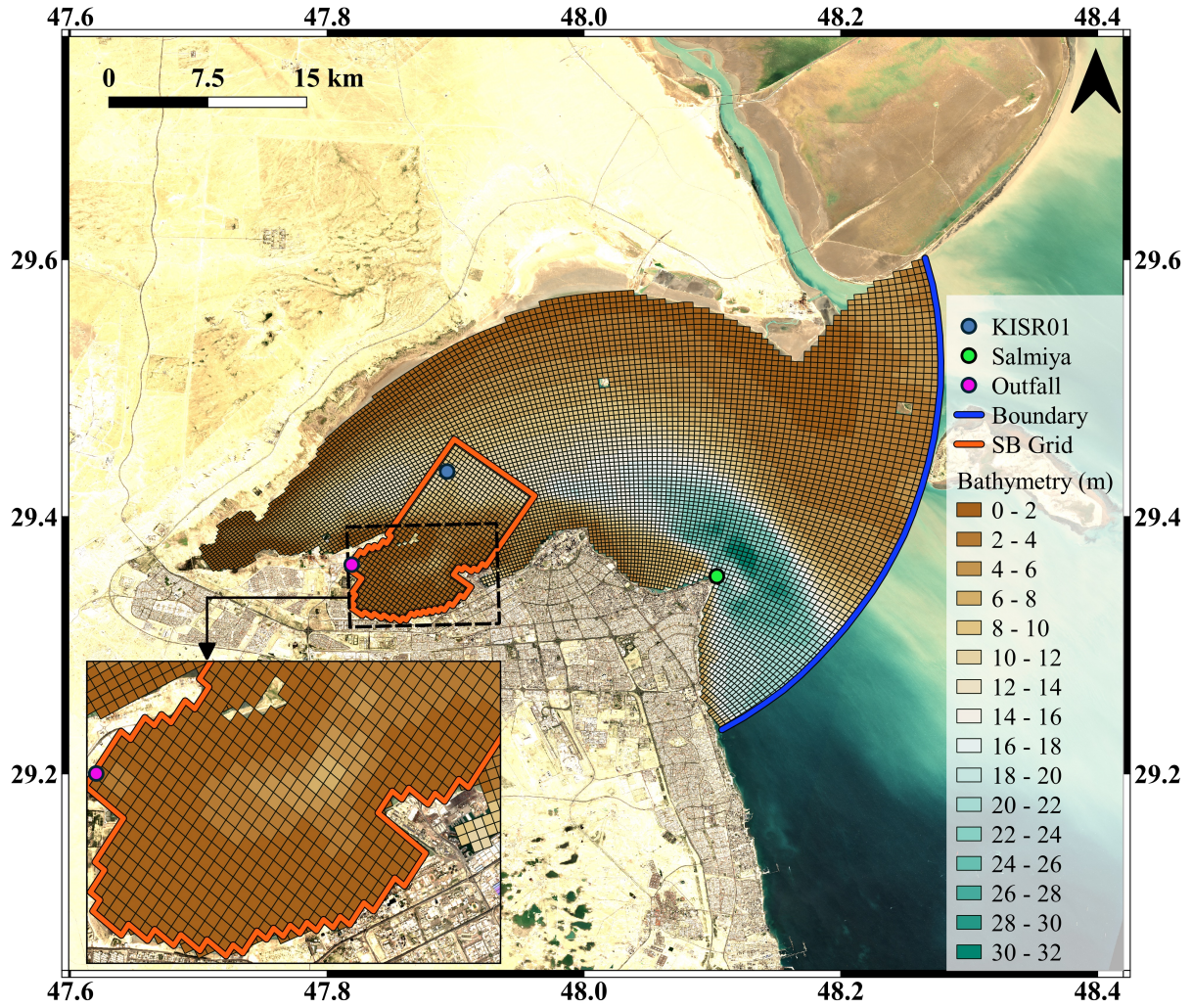


Figure 2.1: Two-dimensional modelling setup for this study. Bathymetry is indicated relative to mean sea level. Forcing and validation stations (Salmiya and KISR01), and Doha West (DW) power and desalination plant are marked on the map. Computational grids for Kuwait Bay (KB) and Sulaibekhat Bay (SB) are illustrated, with SB zoomed in at the bottom left corner.

While the truth model of SB in the twin experiment used synthetic data for discharge parametrisation (details are provided in 2.2.3.1), it still needed to be validated to ensure its credibility in capturing the observed dynamics. Measurements of water levels from the KISR01 station were compared against the simulated water levels from the SB model. Figure 2.2a shows good agreement between the phase and amplitude of the simulated and measured water levels, with residual differences within the ± 0.5 m range. This difference is deemed acceptable for this study’s objectives.

The horizontal spatial distribution of the depth-averaged simulated temperature in SB was compared against satellite scenes sourced from Landsat 8 Thermal Infrared Sensor (TIRS) through the United States Geological Survey (USGS) Earth Explorer platform (<https://earthexplorer.usgs.gov/>). The scenes were processed with the Semi-Automatic Classification Plugin (SCP), which converted TIR data into temperature values using the satellite scene parameters (Congedo, 2016). An overview of the procedure can be found in Appendix B. The satellite scenes, which overlapped with the simulation period, showed reasonable agreement on low (Figure 2.2b and c) and high tide states (Figure 2.2d and e), considering the model’s inherent limitations due to grid resolution and two-dimensional approximations.

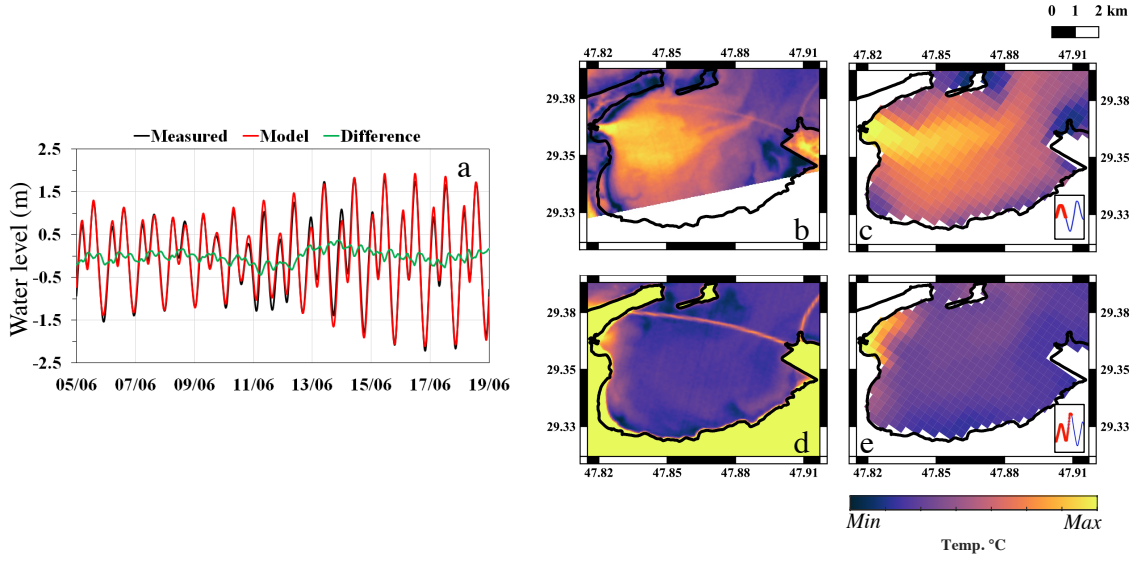


Figure 2.2: Validation of SB truth model. Tides were validated with (a) a comparison of measured and simulated water levels at the KISR01 station from 5th June to 19th June 2022. The horizontal distribution of temperature was validated by comparing simultaneous remotely sensed sea surface temperature and model predictions at different tidal states. Low tide state: (b) Landsat 8 TIR satellite observation on 8th June 2022 at 21:20. (c) SB model prediction for the same date and time. High tide state: (d) Landsat 8 TIR satellite observation on 12th June 2022 at 10:20. (e) SB model prediction for the same date and time. The water level displayed in panels c and e show the tidal phase, as indicated by the end of the red markers, at the time of validation.

2.2.2 Data Assimilation

2.2.2.1 The Ensemble Kalman Filter (EnKF)

The Ensemble Kalman Filter (EnKF) is a sequential DA method first introduced by Evensen (1994). It has gained a lot of attention due to its simple conceptual framework and ease of implementation. Using the Monte Carlo approach, the EnKF provides a practical alternative to the conventional Kalman Filter (KF) and the Extended Kalman Filter (EKF) by accounting for the inherent nonlinearity in the modelled system through representing the error statistics using an ensemble of the model state. The derivation presented in the following section for completeness is adapted from Evensen (2003).

The analysis step in the EnKF is based on the KF, which combines the model forecast

and observations to produce an improved estimate of the state as follows:

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{d} - \mathbf{H}\mathbf{x}^f), \quad (2.1)$$

where $\mathbf{x}^a$ is the n -dimensional vector of the ‘analysis’ (the updated state estimate), $\mathbf{x}^f$ is an n -dimensional vector of the model ‘forecast’ (the background state estimate), $\mathbf{d}$ is a p -dimensional vector of the ‘observations’ (i.e., measurements), $\mathbf{K}$ is the Kalman gain matrix of dimension $n \times p$, and $\mathbf{H}$ is the observation operator of dimension $p \times n$ which maps the model state space to the observation space. The innovation term, $(\mathbf{d} - \mathbf{H}\mathbf{x}^f)$, represents the difference between the observation and the model. The Kalman gain $\mathbf{K}$ is computed as:

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1}, \quad (2.2)$$

where $\mathbf{P}^f$ is the $n \times n$ forecast (i.e., background) error covariance matrix and $\mathbf{R}$ is the $p \times p$ diagonal observation error covariance matrix, as the errors are assumed to be uncorrelated in space. $\mathbf{P}^f$ plays an important role in projecting the innovation term onto the model state during assimilation (Oke and O’Kane, 2011).

As the true state is unknown, the ensemble covariance matrices in the EnKF are defined around the ensemble mean, which is interpreted to be the best estimate, and the spread of the ensemble around the mean represents the error. With m ensemble members, the ensemble mean $\hat{\mathbf{x}}$ is computed as:

$$\hat{\mathbf{x}} = \frac{1}{m} \sum_{i=1}^m \mathbf{x}_i. \quad (2.3)$$

The forecast error covariance matrix $\mathbf{P}^f$ is then given by:

$$\mathbf{P}^f = \frac{1}{m-1} \sum_{i=1}^m (\mathbf{x}_i - \hat{\mathbf{x}})(\mathbf{x}_i - \hat{\mathbf{x}})^T. \quad (2.4)$$

Using (2.1), the analysis for the EnKF updates each ensemble member according to the following:

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{d} + \boldsymbol{\epsilon}_i - \mathbf{H}\mathbf{x}_i^f), \quad i = 1, \dots, m, \quad (2.5)$$

where $\boldsymbol{\epsilon}_i$ is a vector randomly drawn from a normal distribution with zero mean and

covariance $\mathbf{R}$ to perturb the observations $\mathbf{d}$. The ensemble-based analysis error covariance becomes:

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{KH})\mathbf{P}^f. \quad (2.6)$$

2.2.2.2 Ensemble-based Targeted Observation Design

To determine the observational network design that most effectively minimises the uncertainties in the forecast model, we apply the Sakov and Oke (2008) method. This approach begins by considering an array consisting of p observations to be designed. The uncertainty is represented by a background error covariance matrix $\mathbf{P}^b$, constructed from (2.4). After assimilating the observations, $\mathbf{P}^b$ updates to the analysis error covariance $\mathbf{P}^a$ following (2.6), which can be rewritten as:

$$\mathbf{P}^a = \mathbf{P}^b - \mathbf{KHP}^b, \quad (2.7)$$

and from (2.2), we know that:

$$\mathbf{K} = \mathbf{P}^b \mathbf{H}^T (\mathbf{HP}^b \mathbf{H}^T + \mathbf{R})^{-1}. \quad (2.8)$$

The optimal set of observations, described by the observation matrix $\mathbf{H}^{opt}$, is the one which minimises the trace of the analysis error covariance matrix $\mathbf{P}^a$. As the diagonal elements of $\mathbf{P}^b$ indicate the variance of the ensemble, it can be stated that:

$$\mathbf{H}^{opt} = \arg \min \text{trace}(\mathbf{P}^a) = \arg \max \text{trace}(\mathbf{KHP}^b). \quad (2.9)$$

By substituting (2.7) and (2.8) into (2.9), and applying the property that $\text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})$, the solution can be rewritten as:

$$\mathbf{H}^{opt} = \arg \max \text{trace} \left(\frac{\mathbf{HP}^b (\mathbf{HP}^b)^T}{\mathbf{HP}^b \mathbf{H}^T + \mathbf{R}} \right). \quad (2.10)$$

The substantial computational expense of the EnKF arises primarily from the necessity of inverting the $p \times p$ matrix of $(\mathbf{HP}^b \mathbf{H}^T + \mathbf{R})$ (Bishop et al., 2001). Hence, it is a common practice to assimilate the observations in small batches or one at a time (Sakov and

Oke, 2008). Serial and parallel assimilation of uncorrelated observations yields identical analysis error covariance (Bierman, 1977). As such, the serial approach is often used to reduce the computational expenses when identifying targeted observations (Bishop et al., 2001). Proof of the equivalence of serial and parallel assimilation is provided by Sakov and Oke (2008).

When assimilating one measurement at a time (i.e., when $p = 1$), the observation matrix $\mathbf{H}$ reduces to a row vector with all elements set to zero except for a ‘1’ at the k th position, indicating the location of the observation. Consequently, $\mathbf{R}$ becomes a scalar r_{kk} , and the term $\mathbf{H}\mathbf{P}^b\mathbf{H}^T$ reduces to a scalar value, denoted as p_{kk} , indicating the variance of the k th element. The term $\mathbf{H}\mathbf{P}^b$ reduces to an n -dimensional vector indicating the covariance between the k th element and the other elements in the state vector.

After identifying the location of the first optimal observation using (2.10) and calculating $\mathbf{P}^a$ with (2.7), the serial process continues by using the analysis error covariance matrix associated with the first optimal station as the background error covariance matrix $\mathbf{P}^b$ to determine the location of the second observation by minimising its trace, and so on.

In this study, we directly used (2.7) to compute $\mathbf{P}^a$, as both the state variables and the ensemble were of a reasonable size. However, if $n \gg m$, direct manipulation of $\mathbf{P}$ becomes very challenging. A computationally efficient alternative is provided by the Ensemble Transform Kalman Filter (Bishop et al., 2001).

2.2.2.3 OpenDA Toolbox

OpenDA is an open-source toolbox, launched in 2010, for both data assimilation and calibration of dynamic models. With minimal coding intervention required, OpenDA offers the user different assimilation algorithms and uncertainty models for state and parameter estimation. While OpenDA is Java-based, it allows core calculations in other languages (Ridler et al., 2014).

OpenDA consists of several main blocks: method (assimilation or calibration algorithms), observations (stochastic observer for handling observational data), and the numerical model. OpenDA offers the user a great level of flexibility by allowing direct specification of algorithm types, ensemble sizes, variables, and grid locations. Additionally, it

lets the users define the uncertainties associated with the model, data, parameters, and observations (van Velzen et al., 2016).

In this study, we chose OpenDA for EnKF assimilation due to its simple “black box” integration with D-Flow FM, which facilitates handling the model’s input and output files without the need to access the model’s source code when performing the assimilation (van Velzen et al., 2016). The full source code of OpenDA can be accessed on their GitHub page (<https://github.com/OpenDA-Association/OpenDA>).

Several studies have reported successful implementation of the EnKF via the OpenDA framework. This includes the application of hydraulic head assimilation in Denmark’s Karup catchment (Ridler et al., 2014), water quality in South Korea’s Yeongsan River (Loos et al., 2020), and assimilation of in situ and satellite remote sensing temperature data in Lake Geneva’s Delft3D-FLOW model (Baracchini et al., 2020a). It has also improved velocity fields in a Delft3D-FM model of a predominantly semidiurnal micro-estuary (Khanarmuei et al., 2021; Mardani et al., 2021), wave simulation in the Gulf (Serpoushan et al., 2013), and sea surface height assimilation with an ocean model (van Velzen et al., 2016).

2.2.3 Observing System Simulation Experiments (OSSEs)

2.2.3.1 Twin Experiment Configuration

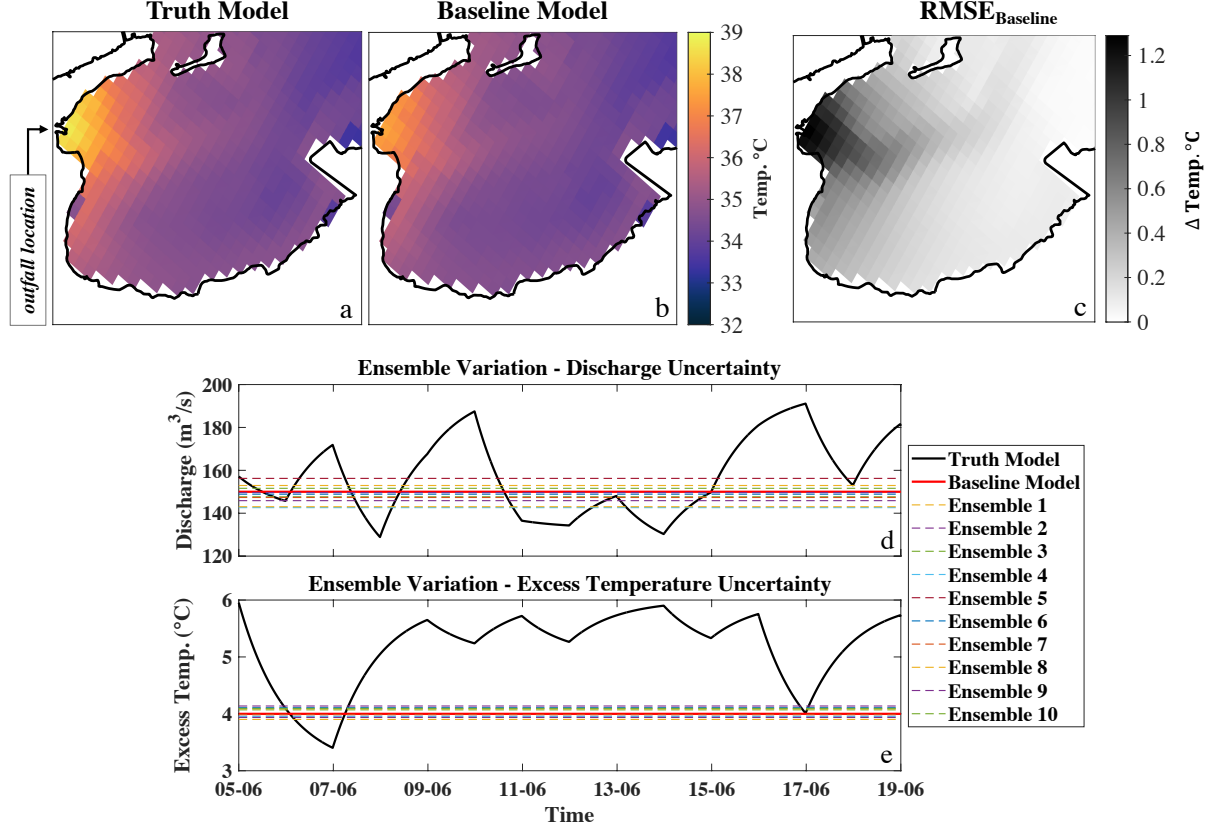


Figure 2.3: Summary of the setup of the twin experiment. The arrow points to the location of the Doha West (DW) outfall. (a) Average temperature from the truth model; (b) Average temperature from the baseline model; (c) Root-mean-squared error between the truth and baseline models. Variability for (d) the discharge rate and (e) the excess temperature for the ensemble generation used for data assimilation, with solid lines representing the truth (black) and baseline (red) values for each parameter.

The experimental setup is delineated as follows:

- Truth model: This model served as a proxy for the “true” state in SB. The model used a time-varying discharge rate and excess temperature that synthetically varied around $150 \text{ m}^3/\text{s}$ and 4°C , respectively (see time series in the solid black line in Figure 2.3d and e). These values are within a reasonable range of the reported values of $197 \text{ m}^3/\text{s}$ and 5°C for DW during the summer (Pokavanich and Alosairi, 2014). The average temperature of the truth model for the duration of the simulation

can be seen in Figure 2.3a.

- **Baseline model:** This model (seen in Figure 2.3b), used a constant discharge rate of $150 \text{ m}^3/\text{s}$ and excess temperature of 4°C (see time series in the solid red in Figure 2.3d and e) for the discharge definition. Open boundary conditions and grid configuration were intentionally kept identical for both truth and baseline models to allow the study to focus on the effectiveness of DA in adjusting the state estimate arising only from the uncertainties associated with the discharge source point. The difference in discharge parameters for both models drove a sufficient difference in the temperature fields (Figure 2.3c) to avoid the skill assessment issues associated with identical twin experiments.
- **Ensemble Generation:** We designed a 10-member ensemble with random fluctuations in discharge rate and temperature within $\pm 5\%$ of the baseline's values, as depicted in Figure 2.3d and e. Sensitivity analysis supported using a 10-member ensemble, as expanding beyond this, up to 80 ensembles, resulted in negligible differences in outcomes. The results of the ensemble size sensitivity analysis are likely attributed to the experiment's forced confinement of the uncertainties in the discharge definition only.
- **EnKF model:** This model represents the outcome of the EnKF DA runs. The 10-member ensembles which were generated based on the baseline model and the observations extracted from the truth model were assimilated using the EnKF algorithm to provide a better estimation of the seawater temperature. An appropriate spin-up period of one week was included in all assimilation models to allow the difference in the ensembles to grow while achieving a balanced initial state before the start of the assimilation process. As this study investigates the effectiveness of DA in tidal environments, we set the assimilation period to cover the duration of a full spring-neap tidal cycle (5th to 19th of June, 2022). We assigned the observation error covariance $\mathbf{R}$ a typical Gaussian observation error with a mean of 0 and a standard deviation of 0.05°C , to reflect the precision of the commonly available temperature sensors in the market.

Our OSSE design limitations should be addressed before proceeding with the result in-

terpretation. Since the truth and baseline models share identical boundary forcings, and the difference was solely confined to the model’s discharge definition, the outputs from both models converged in the regions close to SB’s boundary during high tides (i.e., when information from the open boundary prevails in the domain). Since the experiment operates under the assumption of a perfect model, meaning potential errors arising from inaccuracies in the model—such as those in bathymetry, open boundary conditions, and approximations within the numerical scheme—are not accounted for, results from our experiments should be regarded as representing an ideal scenario. We emphasise that the significance of any series of OSSEs heavily relies on the assumptions made in their configuration, which are often optimistic, and therefore the corresponding outputs should be considered indicative and primarily qualitative (Oke and O’Kane, 2011).

2.2.3.2 Performance Metrics

The Root Mean Square Error (RMSE) was chosen as the primary quantitative indicator for the performance of our OSSEs. This metric quantifies the average magnitude of errors between what is predicted and what is observed over the simulation period (T). We calculated two types of RMSE:

$$\text{RMSE}_{\text{Baseline}} = \sqrt{\frac{1}{T} \sum_{i=1}^T (x_i^{TM} - x_i^{BM})^2},$$

$$\text{RMSE}_{\text{EnKF}} = \sqrt{\frac{1}{T} \sum_{i=1}^T (x_i^{TM} - x_i^{EM})^2},$$

- $\text{RMSE}_{\text{Baseline}}$: measures the deviation of the truth model state x_i^{TM} from the baseline model prediction x_i^{BM} .
- $\text{RMSE}_{\text{EnKF}}$: measures the deviation of the truth model state x_i^{TM} from the EnKF model prediction x_i^{EM} .

The impact of the EnKF was assessed by calculating the percentage of RMSE reduction after performing DA on the baseline model. The error reduction percentage, defined as:

$$\text{Error Reduction (\%)} = \left(\frac{\text{RMSE}_{\text{Baseline}} - \text{RMSE}_{\text{EnKF}}}{\text{RMSE}_{\text{Baseline}}} \right) \times 100, \quad (2.11)$$

quantifies how much errors in the EnKF model have diminished relative to the baseline model. A higher percentage corresponds to a higher reduction in $\text{RMSE}_{\text{Baseline}}$.

2.3 Results and Discussion

2.3.1 OSSE1: Assimilation Relative to Different Frequencies and Locations

Three equally spaced points, P1, P2, and P3, were selected as the primary observation locations for OSSE1. Station P1 was placed directly at the discharge outfall, while P2 was stationed at an intermediate distance, and P3 was positioned the farthest distance from the outfall.

Relative to a typical predominantly semidiurnal tidal cycle, the stations experience the influence of the discharge to varying degrees. Due to P1's location, Figure 2.4 shows that it is continuously under the influence of the discharge regardless of the tidal phase. In contrast, P2 and P3 experience the influence of the plume only when it propagates outwards during low tide. The extent of the plume reaches P2 twice a day during low tide, and once a day for P3 during the lowest low tide. Due to the prescription of uncertainties in this experiment, the propagation of the ensemble variance $\mathbf{P}^b$ has a time-varying pattern that mirrors the propagation of the thermal plume under the influence of tides, similar to that seen in Figure 2.4.

We extracted observational data for the EnKF assimilation scenarios from the truth model at different frequencies. The extraction times included every high tide only, every low tide only, and every high and low tide combined. We then progressed to shorter assimilation intervals of every hour and every 10 minutes. Figure 2.5 shows an example of the extraction process at P2. The results of each assimilation scenario will be assessed based on their ability to reduce the RMSE of the baseline model (Figure 2.6x) using equation 2.11.

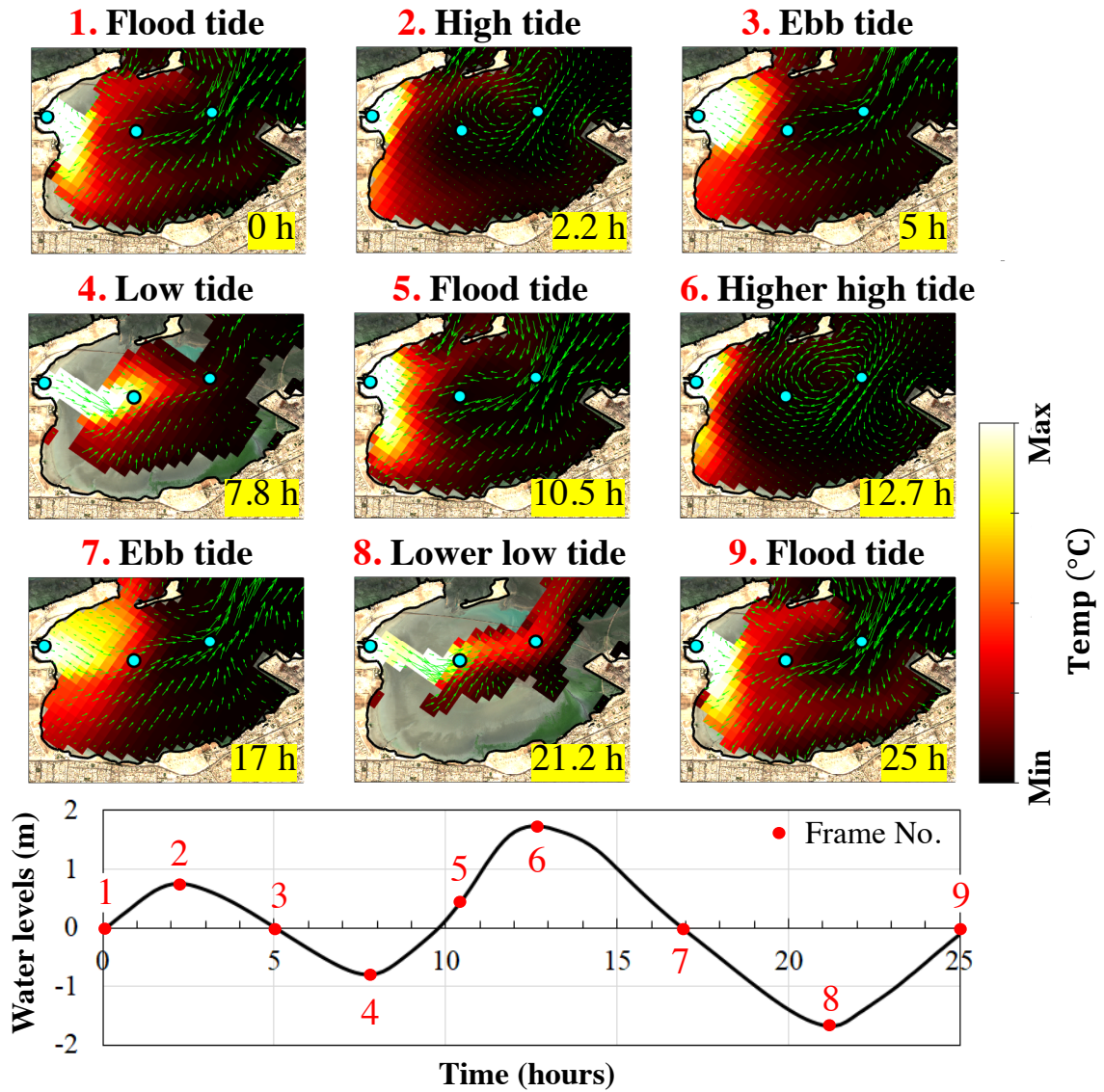


Figure 2.4: Thermal plume interaction with a predominantly semidiurnal tidal cycle during a full lunar day (25 hours). The series of snapshots (frames 1 through 9) depicts the thermal plume at different phases of the tidal cycle. The water levels (bottom) correlate the sequence of the tidal cycle with the snapshots above.

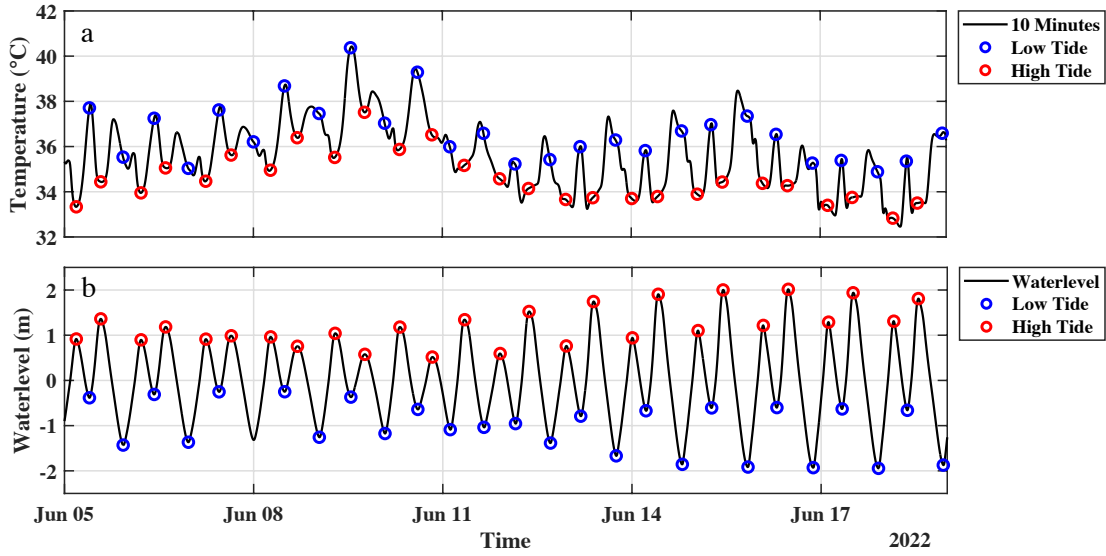


Figure 2.5: Observational data generated for P2 from the truth model. (a) Extraction of the synthetic temperature observations relative to (b) specific tidal phases for different assimilation scenarios.

This section investigates the performance of the filter relative to the assimilation frequency and observation location. Results are interpreted in relation to the propagation of the ensemble variance, the value of the observation's error covariance $\mathbf{R}$, the strength of correlation of the station to the other state elements, and the retention time of the analysis within the system.

When assimilating at a frequency of two times per day and targeting the process to coincide with high tides only, individual results from P3 and P2 showed negligible to no impact on the error reduction (Figure 2.6a and b respectively). Whereas, if we kept the same assimilation frequency, and thereby maintained the same computational cost, but synchronised the DA to low tides only, P3 and P2 start to show individual contributions to the error reduction (Figure 2.6e and f respectively), with P2 driving more corrections to the system than P3.

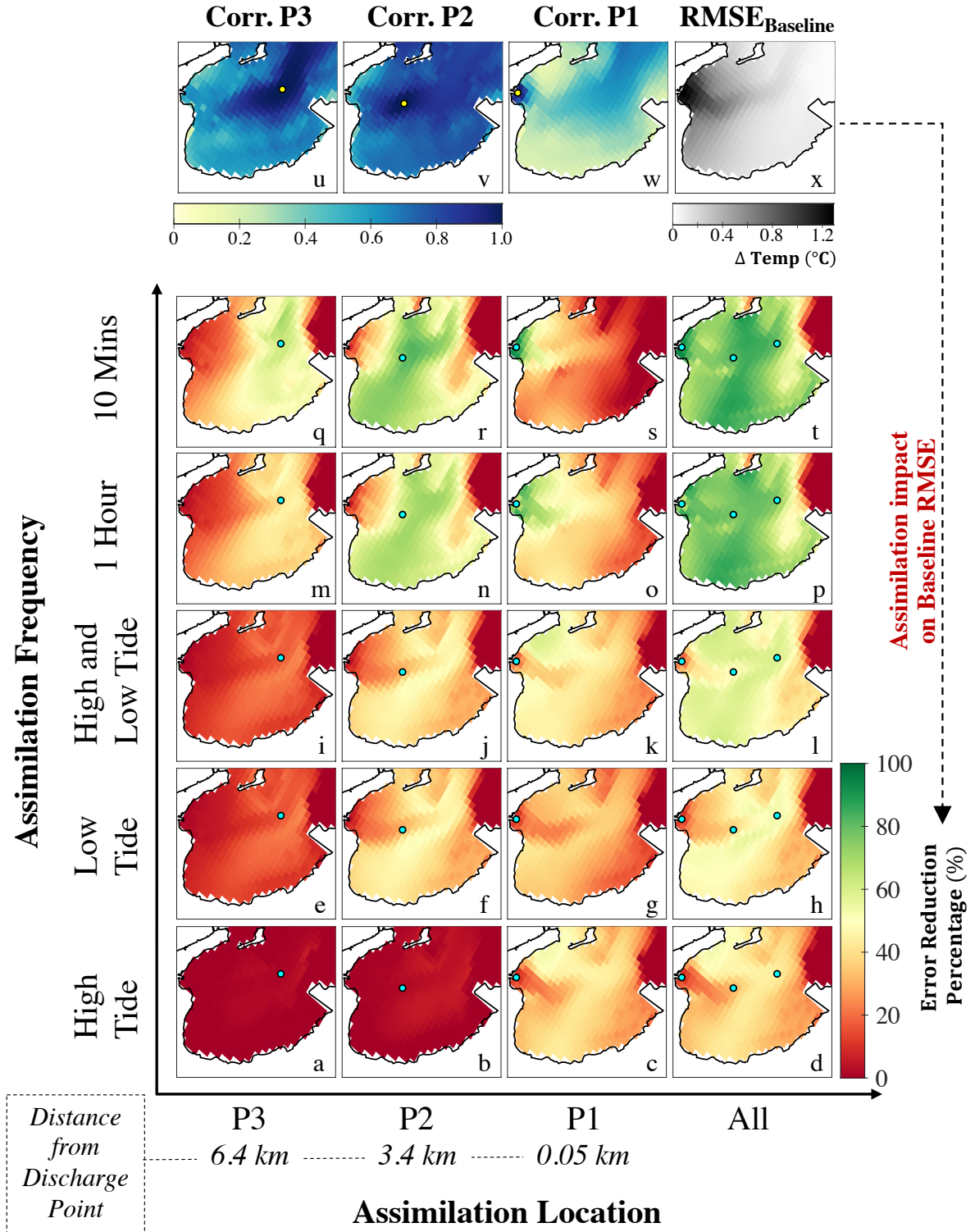


Figure 2.6: Summary of the results from OSSE1. (a–t) Percentage error reduction achieved for each assimilation scenario, with varying frequencies and observation locations. (u–w) Correlation strength of each observation location (P1, P2, P3) with other state elements. (x) RMSE of the baseline model, indicating areas with the highest prediction errors targeted for correction through data assimilation.

The difference in performance can be attributed to the plume’s propagation, and thereby the propagation of the ensemble’s variance, $\text{diag}(\mathbf{P}^b)$, under the influence of the tide. Referring to Figure 2.4, as both stations are momentarily placed outside of the plume’s direct influence during high tide, the local variance experienced at these locations is very small, smaller than the observation’s error variance $\mathbf{R}$, and as a result, the Kalman gain is too small to drive a significant correction to the state estimate. Targeting the assimilation to high tides only made the model’s pre- and post-assimilation performance almost identical. The opposite can be said for assimilating at low tides only. P2’s larger contribution to the error reduction than P3’s can be attributed to its position during low tide. P2 is always within the spread of the plume, and thereby under the influence of the source of uncertainty during both low tides experienced in a full lunar day, whereas the edge of the plume only reaches P3 during the lowest low tide, reducing its effect assimilation rate to one time per day.

As P1 is directly positioned at the outfall, it represents the location of the maximum variance in the ensemble, $\max \text{diag}(\mathbf{P}^b)$, and thereby offers a substantial innovation term for the EnKF regardless of the tidal phase. This is reflected by the significant error reduction attained at both high and low tides (Figure 2.6c and g respectively). The areas achieving the highest error reduction correlate with the plume’s position during the timing of the assimilation. For example, during high tides, the areas scoring the highest error reduction in Figure 2.6c are slightly north of the outfall location, aligning with the plume’s high tide position in Figure 2.4. The approximate computational time for assimilating twice per day was 1 hour and 19 minutes.

Doubling the assimilation frequency to four times per day, and thereby increasing the computational time to approximately 2 hours and 4 minutes, did not yield improvements for P2 and P3, as the outcomes seen in Figure 2.6i and j were identical to those obtained from assimilating during low tide only. The results do not justify the increased computational load. However, P1 benefited from the increased assimilation frequency (Figure 2.6k), as this location is consistently experiencing large variance that requires minimisation.

For all the scenarios, the simultaneous assimilation of all three stations provided better results than their individual contributions. For high tides only, P1 was the primary contributor to the reduction of errors observed in Figure 2.6d, whereas for low tides, the

reduction seen in Figure 2.6h was increased due to the additional contribution of P2. Assimilating every high and low tide provided the best results, where error reductions started to reach above 50% in Figure 2.6l, as a direct by-product of the enhancements achieved by P1 for this assimilation rate.

Yet for all cases, analysis adjustments were transient due to the tidal dynamics of SB. The fast, ebbing currents during low tides, seen in Figure 2.4, prohibited the retentions of any error reductions in the area overlapping with the plume’s ebbing trajectory (see Figure 2.6a through l). Further investigation on the impact of DA frequency was carried out by increasing the assimilation rate to an hourly basis, and then to an interval of 10 minutes, with the former costing approximately 9 hours and 35 minutes to complete, and the latter 56 hours in computational expenses.

For both assimilation frequencies, individual results for each case show that the area extending from the outfall location to P3 (Figure 2.6m and q) and P2 (Figure 2.6n and r) showed little error reduction. The lack of significant improvements in these zones, despite the increased assimilation rate, can be attributed to the following: for meaningful adjustment to the state estimate, the plume—carrying the embedded uncertainty—must contact the observation location to create a significant local variance, prompting a global analysis adjustment via the Kalman gain. Consequently, when the plume is outside the vicinity of the observation locations, no adjustments are made. As such, the performance results for these zones will generally be close to the model in its pre-assimilative state. However, when the plume does propagate to these locations, adjustments will take place, and the value of the innovation term will be appropriately scaled corresponding to the zones of strong correlations for each station (Figure 2.6u and v for P3 and P2, respectively). Since these blind zones of error reduction overlap with areas of high prediction error in the baseline model (Figure 2.6x), the high computational costs associated with the assimilation outcome were not justified.

Only for P1 was the analysis finally retained along the plume’s ebbing path when the assimilation rate was increased to hourly (Figure 2.6o) and 10-minute intervals (Figure 2.6s), with adjustments scaled proportional to the zones of strong correlations (Figure 2.6w). Yet we emphasise that an increased assimilation frequency does not inherently guarantee better results. Increasing the assimilation frequency to an excessive rate of 10

minutes caused major degradation in filter performance when compared to the results of hourly assimilation. The rapid assimilation cycle has caused the ensemble to progressively underestimate the background error covariance, resulting in the observation information being completely ignored by a large portion of the domain. This eventually led to filter divergence.

Consistent with the previous findings, simultaneous assimilation for all stations provided better results against individual performances for the hourly (Figure 2.6p) and the 10-minute (Figure 2.6t) scenarios. Notably, the output from the hourly DA provided a consistent error reduction across the bay, particularly in areas that overlap with high RMSE in Figure 2.6x, of which the 10-minute case was less successful in achieving. The difference in reduction consistency between Figure 2.6p and Figure 2.6t can be attributed to P1's deterioration in performance under the highest assimilation frequency. As such, an hourly rate of DA proved to be the most appropriate under SB's tidal dynamics.

From the findings of OSSE1 we conclude that, for an effective DA, the location of the observation needs to experience sufficient local variance at a level exceeding that of the prescribed observation variance, while simultaneously possessing strong correlations with other state elements. If the location was chosen appropriately, increasing the number of observations appears to be advantageous for error reduction. The frequency of assimilation needs to allow sufficient time for the system to adjust to the analysis updates. Assimilating too frequently can lead to filter divergence, and too sparingly can cause the model to revert to its pre-assimilative state.

2.3.2 OSSE2: Assimilation Based on the Ensemble-based Targeted Observations Method

The ensemble-based method of Sakov and Oke (2008), detailed in section 2.2.2.2, was used to identify an appropriate set of targeted observations suited for the stated monitoring objectives in SB. An ensemble size of 10 members, described in section 2.2.3.1, was integrated over the complete simulation period which covers a full spring-neap tidal cycle, to generate a time-varying background error covariance $\mathbf{P}^b$ from the model's output. $\mathbf{P}^b$ was then averaged over the simulation period to provide general statistics that are descriptive

of the uncertainties in the system. The method detailed in section 2.2.2.2 was applied on $\overline{\mathbf{P}^b}$ to identify the targeted observations.

Following the procedures detailed earlier, 10 locations were identified, as seen in Figure 2.7a. These sites were ranked from 1 to 10 based on their anticipated reduction of the total analysis error variance (i.e., $\text{trace}(\overline{\mathbf{P}^a})$) upon assimilation. Site no. 1 is expected to offer the most significant reduction, and so on. Figure 2.7b details the individual impact of each observation site on the reduction of variance, and Figure 2.7c demonstrates the cumulative percentage reduction achieved by sequentially adding the observations through the assimilation process. It suggests that assimilating all 10 sites could theoretically diminish the total variance across the domain by roughly 75%.

Denoting the observation location as ‘ k ’, Sakov and Oke (2008) argue that selection of the optimal location is not determined solely by the maximum variance p_{kk} associated with the observed state element x_k , but rather, by a balance between its own variance and the covariance it shares with other elements p_{ki} , where $i \neq k$. The authors additionally noted that the placement of the optimal observation is influenced by the magnitude of the prescribed observation error variance r_{kk} . Both Liu et al. (2018) and Geng et al. (2020) have reported that when using the ensemble-based method for the identification of targeted observations, the reduction in RMSE was sensitive to the magnitude of $\mathbf{R}$, with smaller observation errors yielding a larger percentage reduction in RMSE. Sakov and Oke (2008) concluded that when the selection criterion was based solely on the maximum background variance (i.e., $\max \text{diag}(\mathbf{P}^b)$), it resulted in suboptimal performance compared to their proposed criterion, which selects the locations based on the minimisation of the trace of $\mathbf{P}^a$. Their method benefits from including information about the background error covariances, rather than just the background error variances. Consequently, the results of OSSE2 have placed P1, the location of maximum variance, as 7th in assimilation effectiveness, as seen in Figure 2.7.

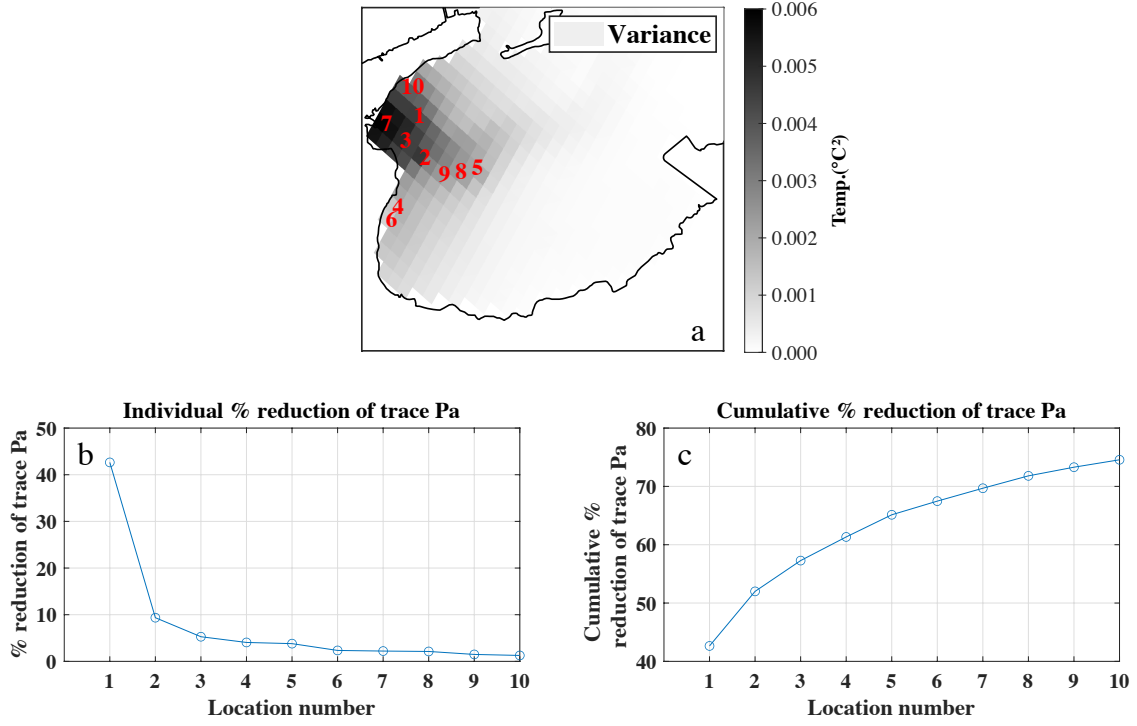


Figure 2.7: (a) The first 10 targeted observation locations, ranked by their effectiveness in reducing the trace of $\mathbf{P}^a$, with the numbering indicating each site's rank. The ranking is plotted against the trace of $\mathbf{P}^a$; (b) The individual percentage reduction of trace of $\mathbf{P}^a$ achieved using the n th observation site; (c) The cumulative percentage reduction of trace $\mathbf{P}^a$ achieved by the first 10 observation sites.

Since the observation locations were selected based on the system's averaged statistics, we propose the inclusion of an additional step in the assimilation process to incorporate the flow-dependent variation in the background error statistics which are influenced by the propagation of the thermal plume under the control of the tides. We confined the assimilation process to occur only when the local variance at the targeted observation p_{kk} surpasses that of the prescribed observation error variance r_{kk} . The temporal variation of the variance of the top three targeted observations, hereafter named PT1, PT2, and PT3, can be seen in Figure 2.8.

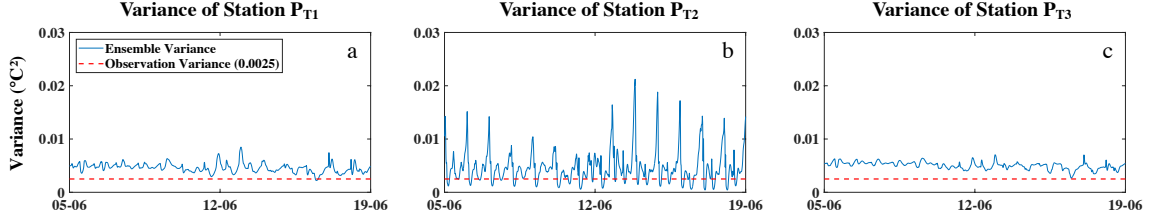


Figure 2.8: Variance p_{kk} of the stations (a) PT1; (b) PT2; and (c) PT3 compared with the observation error variance r_{kk} (dotted red line). These plots inform the determination of assimilation windows.

When p_{kk} is larger than r_{kk} , assimilation will commence at an hourly frequency. The rate of assimilation was chosen based on the findings of the previous section. Only the top three locations were assimilated, to facilitate a fair comparison with the results of OSSE1 to understand the impact targeted observations have on error reduction.

Figure 2.9a shows that while the error reduction achieved by the hourly assimilation from the equally spaced stations of OSSE1 was high, it was notably unevenly distributed across SB, a by-product of the uneven contribution each individual station has on the collective DA outcome. In contrast, Figure 2.9b shows that the relocation of the same number of observations allowed OSSE2 to achieve consistent and significant error reduction across the domain, notably in the western region of SB, where the prediction errors from the baseline model prevail (recall Figure 2.6x). Understandably, the area within a 2-grid cell vicinity of the outfall achieved less than 50% corrections, as the configuration of OSSE2 did not include an observation point there. However, from OSSE1 we concluded that while the inclusion of a point at the outfall location (P1) can benefit the immediate vicinity, the lack of strong correlations will limit its broader impact.

To understand the difference in performance between the two configurations further, Figure 2.9c through j provide an in-depth visualisation of the average background error covariance $\overline{\mathbf{P}}^b$ of the ensemble. Station variance $\overline{p_{kk}^b}$ and the covariances it shares with other elements $\overline{p_{kl}^b}$ relative to the prescribed observation variance r_{kk} are illustrated, with areas having values larger than r_{kk} highlighted using an orange-to-yellow gradient for ease of identification.

It is evident from Figure 2.9c and d, that regarding individual station variance, all targeted

observations of OSSE2 are all placed within the region of high variance, where $\overline{p_{kk}^b} > r_{kk}$. Whereas only P1 of OSSE1 is situated within that zone. This explains why the outcomes of individual assimilation for P2 and P3 were highly sensitive to the timing of assimilation relative to the propagation of variance under tidal forces.

As noted by Sakov and Oke (2008), variance is not the sole determinant of the DA efficiency. PT1, PT2 and PT3 share strong covariances with the other state elements (Figure 2.9f, h and j). Notably, these stations exhibited covariances larger than r_{kk} in the areas where the baseline model had the largest prediction errors. On the contrary, only P1 in OSSE1 had this property (Figure 2.9e), while P2 and P3 generally had weak covariances (Figure 2.9g and i). As such, we conclude that OSSE2 has outperformed OSSE1 due to its strategic placement of stations within the areas of both high variance and high covariance, effectively targeting the regions that required the most correction.

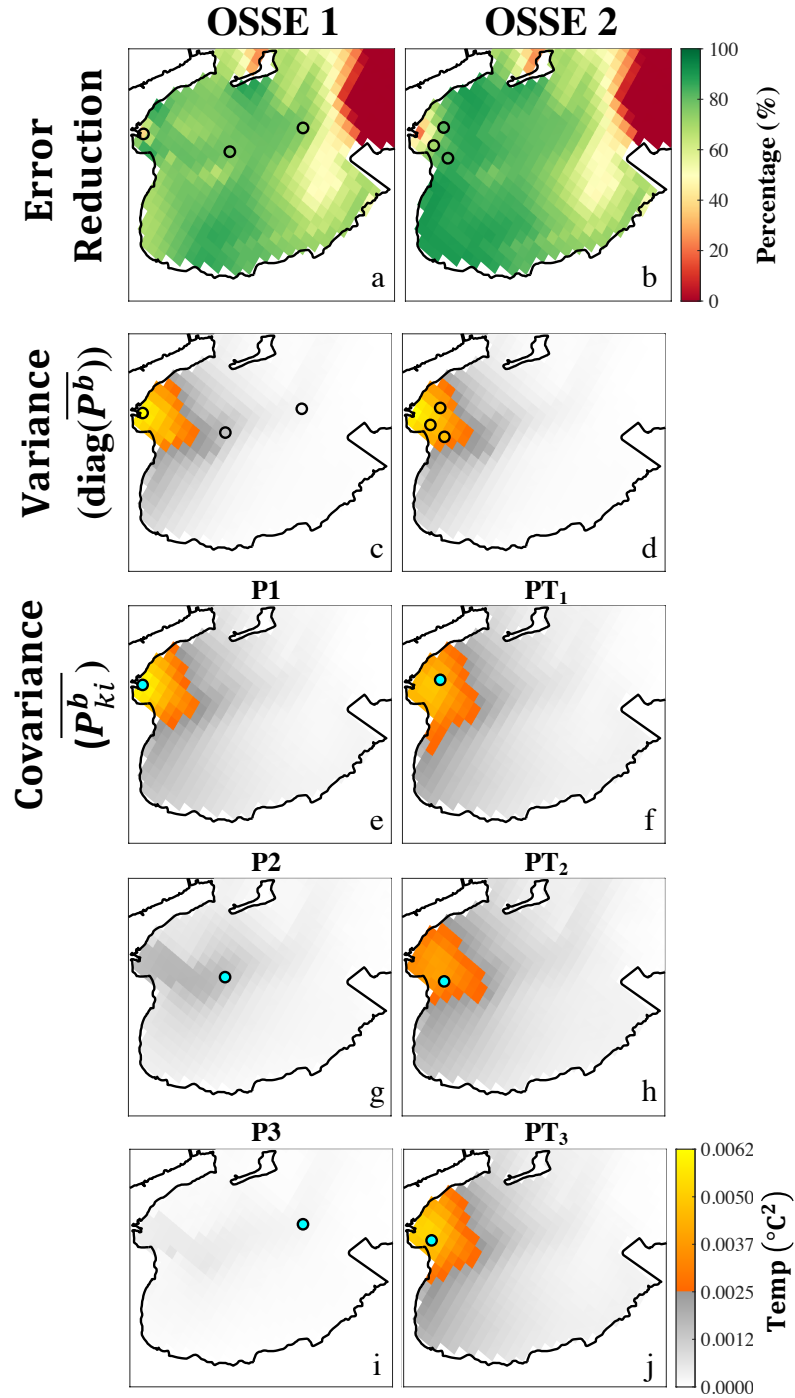


Figure 2.9: Comparison of error reduction, variance, and covariance in OSSE1 and OSSE2. (a) Error reduction percentage in OSSE1. (b) Error reduction percentage in OSSE2. (c-d) OSSE1 and OSSE2 stations marked against the time-averaged variance of the ensemble $\overline{\mathbf{P}^b}$. (e-j) Individual covariance maps for stations in OSSE1 (P1, P2, P3) and OSSE2 (PT1, PT2, PT3). The orange-to-yellow gradient corresponds to areas with values larger than the observation error variance ($> 0.0025^\circ\text{C}^2$).

2.3.3 Limitations and Considerations

The data assimilation framework presented in Chapter 2 demonstrates the utility of the Ensemble Kalman Filter (EnKF) for improving thermal effluent predictions in tidal waters. However, several limitations must be considered when evaluating the representativeness, generalisability, and modelling assumptions underlying the results.

- In this framework, the assimilated observations were synthetic, generated from a model spanning a full spring–neap cycle. However, the short summer simulation window does not capture seasonal changes in ambient seawater temperature. Additionally, episodic events such as Shamal (northwesterly) and Kous (southeasterly) winds, which are known to drive significant wind-induced circulation affecting temperature in the northwestern Gulf (Al Senafi and Anis, 2020), were not investigated.
- Time averaging the ensemble-based background error covariance matrix $\mathbf{P}^b$ may bias the optimal array design toward conditions that occur near mean sea level (i.e., water levels approaching 0 m). The propagation of ensemble variance closely follows the movement of the thermal plume under tidal influence, and this time dependence can be lost through averaging. A more balanced strategy could involve computing separate $\mathbf{P}^b$ averages for high tide, low tide, and slack conditions, then combining the top-ranked observation locations from each to form a composite design.
- The design of the optimal observation array was constrained by the model grid resolution. While adequate for capturing the general plume structure, a finer grid could enhance plume detail and improve the precision of optimal station placement.
- The SB model operates in a two-dimensional, depth-averaged mode. While suitable for the objectives of Chapter 2, applying the same methodology to a fully three-dimensional model could improve realism, particularly near the outfall or in vertically stratified conditions.
- Chapter 2 presents a proof-of-concept application of the EnKF in a controlled OSSE framework under idealised summer conditions. The insights regarding the relationship between assimilation frequency, observation location, and system dynamics are based on assumptions of a perfect model, a single uncertainty source (discharge),

fixed grid resolution, and a short-term simulation. Future studies should extend this approach to longer-term ensemble simulations, include seasonal variability, and employ more comprehensive model configurations to assess the robustness and generalisability of the conclusions.

2.4 Conclusion

The Ensemble Kalman Filter (EnKF) was successfully applied to address uncertainties in modelling thermal effluents from a coastal power plant in a shallow, tidally influenced bay. The filter generates enhanced initial conditions for the model whenever an observation is made available for assimilation. However, the complex dynamics of tidal environments, characterised by strong advection and dissipative processes, pose a challenge to the retention of data assimilation (DA) adjustments in the system.

To ensure a successful implementation of DA in tidal waters, we performed a series of observing system simulation experiments (OSSEs) through the configuration of a twin experiment, to identify the optimal assimilation frequency and observational location suited to our environment and study objectives. While the outcome of any OSSEs heavily depends on the assumptions made during their configuration, the results of this study can be taken to be an indicative characterisation of the optimal DA setup for shallow, tidal environments more generally. These findings can therefore be taken as a useful guide for similar coastal environments globally.

Optimal assimilation frequency was found to be that which allows the system sufficient time to integrate the analysis updates into the simulation. Longer assimilation cycles led the assimilative model to revert to its non-assimilative state, whereas rapid assimilation cycles led to filter divergence. We estimated the optimal frequency for our area to be on an hourly basis.

An ensemble-based targeted observation method was used to identify the optimal observation locations based on their ability to maximise the minimisation of the trace of the analysis error covariance of the ensemble. The observation stations with optimal locations were characterised by having both strong covariances with other state elements and experiencing large variances that exceed the value of the prescribed observation error variance.

Targeting observation locations which exhibit maximum variance only does not always guarantee optimal results. However, the reliability of this method depends on building an ensemble-based background error covariance that correctly captures the uncertainties found in the system.

While these findings demonstrated the potential of DA for improving simulation accuracy, the method's reliance on frequent adjustments to the system state highlighted a limitation in its ability to maintain accuracy over longer timescales. Given that the optimal assimilation frequency was determined to be hourly, it became clear that DA is most effective within this short time period. To address this limitation, the next chapter will shift from state estimation to parameter estimation using Bayesian optimisation.

Chapter 3

Bayesian Optimisation for Parameter Estimation

Highlights

- Bayesian optimisation was used to calibrate a hydrodynamic model by optimising two key parameters: discharge rate and excess effluent temperature, using multi-source data.
- Calibration outcomes varied across data sources, depending on how well each captured the thermal plume’s spatial extent, impacting the optimisation algorithm’s ability to estimate the parameters.
- Strategic data collection for calibration should target areas where model parameters exert the greatest influence on simulation outputs.

Summary

This chapter uses Bayesian optimisation to calibrate a three-dimensional hydrodynamic model of a shallow, tidal bay by estimating the discharge rate and excess temperature of thermal effluents from a coastal plant. Multi-source, synchronised temperature datasets—sourced from in situ profiles, unmanned aerial vehicles (UAVs), and satellites—are assessed for their ability to yield consistent calibration results independently.

Bayesian optimisation identifies the optimal parameters by minimising errors between observed and simulated temperatures. A Gaussian Process is constructed as a surrogate model for the objective function, while an acquisition function guides the parameter search through the solution space. The simulated thermal plume shows a high sensitivity to the calibration parameters. Consequently, the region affected by parameter uncertainty coincides with the plume, which is influenced by tidal movements, creating a time-varying zone of influence. Calibration results highlight the complex relationship between the spatial coverage of calibration data and the size of this zone. Datasets with limited spatial coverage, such as UAV data focused only on the immediate outfall area, are effective for calibrating one parameter but less reliable for the other. Conversely, datasets with broader spatial coverage—such as multiple in situ profiles and satellite observations—provide more comprehensive information, leading to more robust parameter estimation. However, when datasets with similar coverage but different resolutions are compared—such as in situ profiles and satellite observations—equifinality may arise, where different parameter combinations yield similar simulation outcomes. In this chapter, we conclude that aligning data collection with the parameter’s zone of influence is crucial for the successful application of Bayesian optimisation in calibrating thermal effluent models.

Publications and Disclaimer

The content within this chapter is based on:

N. Alsulaiman, Y. Alosairi, J. Tlhomole, M. Liu, M. Van Reeuwijk, and M. D. Piggott. *Bayesian Parameter Estimation of Coastal Thermal Effluents Using In Situ, UAV, and Satellite Observations*. In preparation for submission to *Environmental Modelling & Software*, 2025.

3.1 Introduction

Environmental models, despite continuous advancements, remain heavily parametrised approximations of real system dynamics, carrying inherent uncertainties in their simulation outputs. Often functioning as ‘black boxes’, these models can be highly computationally expensive to evaluate (Scyphers et al., 2024). Black box functions are characterised by having an unknown or effectively non-existent mathematical expression, which often restricts the use of conventional optimisation methods that typically require derivative information and a large number of function evaluations (Brochu et al., 2010; Diessner et al., 2022). Modellers frequently perform a calibration process, which is an inverse, multi-step problem aimed at aligning modelled and measured data by adjusting the model’s parameters, ideally within a physically reasonable range, to achieve optimal performance (Schwindt et al., 2023). However, there is arguably limited guidance available in the literature for calibration approaches, which are often conducted in a somewhat ad hoc manner and/or based on user experience, and with no standardised approach currently existing (Williams and Esteves, 2017; Schwindt et al., 2023). Autocalibration can be a powerful tool to reduce subjectivity and manual labour during the process; however, it often requires a large number of model runs, which increases with the number of parameters to be estimated (Williams and Esteves, 2017; Piccioni et al., 2022). To address this problem, a highly efficient calibration method is needed, particularly for complex hydro-ecological models, that requires fewer model runs while maintaining global optimal performance. One such method is Bayesian optimisation (Ma et al., 2022).

Bayesian optimisation is a statistical, machine-learning-based method used to optimise an objective function without a closed-form expression over a feasible search space, using observations obtained at sampled values. The calibration process yields a probability distribution of the parameters through a global search across the solution space and identifies the remaining (post-calibration) uncertainty of these parameters (Brochu et al., 2010; Frazier, 2018; Schwindt et al., 2023; Scyphers et al., 2024). It uses a Bayesian approach by establishing a prior distribution over the objective function, which is updated as new observations are made to form a posterior distribution. This process is often interpreted as estimating the objective function with a surrogate model, typically a Gaussian Process

(GP). An acquisition function is then constructed from the posterior GP to guide the selection of the next observation point by strategically balancing exploration (targeting areas with high uncertainty) and exploitation (focusing on areas that are likely to improve upon the current best result) (Brochu et al., 2010; Frazier, 2018; Scyphers et al., 2024). This dual consideration ensures efficient optimisation by continuously refining the model based on both new observations and an understanding of the current uncertainty.

Bayesian optimisation is particularly attractive for hydro-environmental models due to its effectiveness in optimising functions with high computational costs, as it supports parallel evaluations and adapts well to varying amounts of initial training data. Additionally, it can handle noisy objective functions as efficiently as deterministic ones and has been shown to excel in optimising functions with varying characteristics and numbers of optima (Diessner et al., 2022). Due to its advantages, Bayesian optimisation has gained popularity in recent years, extending to applications such as the automatic calibration of complex hydrological models (Ma et al., 2022), highly parametrised hydro-ecological models (Piccioni et al., 2022), three-dimensional hydrodynamic models (Schwindt et al., 2023), and promising applications in real-time urban flood modelling (Tanim et al., 2024). However, there remains a research gap regarding its effectiveness when applied to thermal effluent models.

This chapter explores the application of Bayesian optimisation in modelling thermal effluents by calibrating a three-dimensional hydrodynamic model of a shallow, tidally influenced bay, where the discharge rate and excess temperature of the effluent are initially unknown. The challenge lies in the strong dependence of these parameters on the observed thermal plume, which evolves with the tides (recall Figure 2.4). As these parameters govern the discharge characteristics, their sensitivity is reflected in the simulated thermal plume, creating a time-varying zone of uncertainty within the modelled domain. We assess how the spatial coverage and resolution of different data sources influence calibration outcomes, considering the plume’s evolving structure and its interaction with tidal dynamics. To address these questions, an extensive field campaign was conducted in June 2022, collecting simultaneous temperature measurements from in situ vertical profiles, a thermal unmanned aerial vehicle (UAV), and Landsat 9 satellite observations.

3.2 Materials and Methods

3.2.1 Observational Data Sources

Traditional in situ observations, such as from buoys and mooring lines, provide accurate temperature measurements across the water column through direct contact with the water. However, their limited spatial coverage, high operational costs, and their ability to only capture instantaneous and localised data make them insufficient for large-scale thermal assessments and tracking long-term variability (Lowe et al., 2009; Huang et al., 2023; Wang et al., 2024b).

Remote sensing technologies, such as satellites and UAVs, provide an effective alternative to in situ measurements, with the ability to capture surface properties over a wide range of spatial scales, without the need for direct contact (Alvir et al., 2024). Satellite-based sea surface temperature (SST) retrieval is a well-established method for monitoring thermal plume dynamics and their environmental impacts, offering a cost-effective solution for large spatial coverage (Durán-Colmenares et al., 2016; Huang et al., 2023; Mokarram et al., 2023; Wang et al., 2024b). However, satellite remote sensing is limited by atmospheric conditions, cloud cover, and relatively low spatial resolution, which can reduce accuracy in detecting nearshore thermal plumes, especially near land structures where mixed water and land pixels complicate temperature retrieval (Roy et al., 2022; Huang et al., 2023).

UAVs have become increasingly popular in environmental monitoring due to their affordability, portability, and ability to capture high-resolution images in otherwise inaccessible areas (Ferrara et al., 2017; Joyce et al., 2018; Vélez-Nicolás et al., 2021). Equipped with thermal infrared cameras and GPS-enabled autopilot systems, UAVs provide detailed, real-time data on local thermal discharges, bridging the gap between the coarse resolution of satellites and the point-specific measurements of in situ methods, while also reducing operational costs and errors (Ferrara et al., 2017; Xiang et al., 2017; Wang et al., 2023b, 2024b; Ardiny et al., 2024). Despite these advantages, UAVs are limited by short flight times, payload constraints, sensitivity to weather conditions, and regulatory restrictions (Ardiny et al., 2024).

A key limitation of remote sensing is that it measures radiant temperatures, while in

situ methods capture kinetic temperatures. Due to surface emissivity, radiant temperatures are typically lower, which necessitates validation and adjustment of remote sensing measurements using in situ data (Vanhellemont, 2020; Caldwell et al., 2019; Roy et al., 2022). Additionally, remote sensing is limited to surface-level observations, restricting comprehensive, three-dimensional monitoring of thermal discharges (Wang et al., 2024b). To overcome these limitations, recent studies have integrated in situ and remote sensing data with numerical models, combining the strengths of each to address their individual limitations for a more complete representation of thermal discharge dynamics (Durán-Colmenares et al., 2016; Laguna-Zarate et al., 2021; Kong et al., 2022; Wang et al., 2024b).

In this chapter, three overlapping multi-source datasets were collected to monitor the thermal plume emitted from DW at different tidal stages in SB. Dataset 1 was collected during the rising tide (i.e., one hour after low tide); Dataset 2 was collected at high tide; and Dataset 3 was collected at low tide. Table 3.1 provides details on each dataset and their acquisition times. Although the datasets overlapped temporally, they varied in spatial coverage and resolution, as illustrated in Figure 3.1.

Table 3.1: Acquisition times for each dataset used in the study. For vertical profiles, the number of profiles is indicated in parentheses. For UAV data, the altitude of the image is specified in parentheses. For satellite data, the source is noted in parentheses.

Acquisition Time			
Dataset	Vertical Profiles	UAV	Satellite
No. 1	04-06-2022 09:27 – 10:56 AM (11 profiles)	04-06-2022 10:00 AM (alt. 335 m)	04-06-2022 10:16 AM (Landsat 9)
No. 2	13-06-2022 09:00 – 10:56 AM (15 profiles)	13-06-2022 10:05 AM (alt. 350 m)	
No. 3	20-06-2022 09:30 – 10:15 AM (10 profiles)	20-06-2022 10:08 AM (alt. 344 m)	20-06-2022 10:16 AM (Landsat 9)

Sixteen vertical profiling stations, illustrated in Figure 3.1, were established at approximately equal intervals across SB, covering a 10 km survey route that extended to within 50 meters of the outfall. Tidal conditions influenced station coverage during data collection: Datasets 1 and 3 included stations P1 to P11 and P10, respectively, due to shallow water restricting access further upstream, while Dataset 2, collected at high tide, covered all 16 stations. Vertical profiles were measured using a water quality profiler (AAQ-RINKO, JFE Advantech), which recorded water temperature at 10 Hz with a resolution of $0.001,^{\circ}\text{C}$ and an accuracy of $\pm 0.01,^{\circ}\text{C}$. The data, post-processed to average readings at 0.1 m depth intervals, are considered the true water temperature, requiring no further corrections.



Figure 3.1: Spatial coverage of the data sources. Satellite data (red box) encompasses the entire region, vertical profile stations (scatter dots) are colour-coded by distance from the DW outfall, and UAV data (green box) focuses on the immediate vicinity of the outfall.

To examine the vertical structure and downstream evolution of the thermal plume, temperature profiles along the transect were extracted from Datasets 1, 2, and 3 and are presented as vertical sections in Figure 3.2. Each panel illustrates how the thermal signature varies with depth and distance from the discharge point. In Datasets 1 (panel a) and 3 (panel c), both collected during ebb tide, the plume extends approximately 7–7.5 km downstream. Beyond the 4 km mark, the thermal signal becomes increasingly confined to the upper water column as the plume is advected seaward. In contrast, Dataset 2 (panel b), acquired during flood tide, shows a much shorter plume extent of around 2 km, with heat retained closer to the discharge source as tidal reversal drives upstream transport.

It is also worth noting that distinguishing the outer extent of the plume becomes challenging when it occupies only the surface layers, particularly at greater distances from the outfall. Since the vertical profiles were collected between approximately 09:00 and 11:00 AM, surface heating due to solar radiation is expected during this period. This atmospheric warming can blend with the thermal signal of the effluent in the upper water column, making it challenging to precisely determine the plume boundary, especially in areas where the signal is already weakened by mixing and advection.

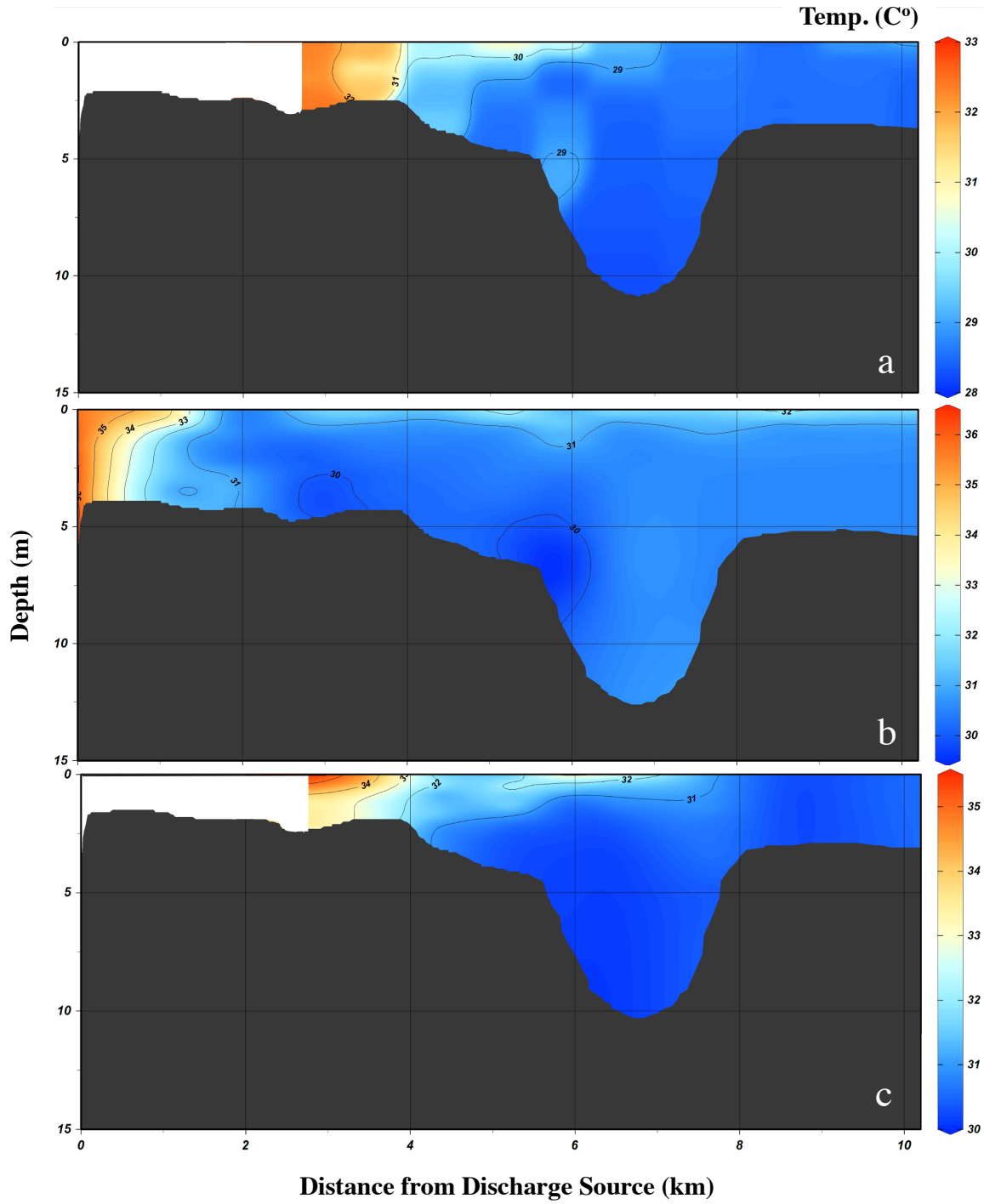


Figure 3.2: Vertical temperature cross-sections from Datasets 1 (a), 2 (b), and 3 (c) showing the thermal plume evolution with depth and distance from the discharge source.

Temperature data were collected using a DJI Matrice 210 V2 UAV equipped with a FLIR Tau 2 thermal imaging camera (640×520 resolution). Flying at an average altitude of 340 meters, the UAV covered the area shown in green in Figure 3.1. The camera's un-

cooled microbolometer thermal infrared (TIR) sensor is susceptible to temperature drift, requiring radiometric corrections for factors including flight distance, air temperature, humidity, emissivity, and reflective effects (Caldwell et al., 2019; Vélez-Nicolás et al., 2021). Additionally, daytime water scenes may produce false readings due to sun glint, surface illumination, shadows, or foam, necessitating ground-truthing with in situ measurements (Joyce et al., 2018; Vélez-Nicolás et al., 2021).

Ground-truthing of UAV data was performed using synchronised in situ measurements taken during the drone flights, following practices by Wang et al. (2023b) and Wang et al. (2024b) for monitoring power plant discharges using thermal UAVs. Raw TIR images were processed using ThermoViewer software (<https://thermalcapture.com/thermoviewer/>) to extract temperature readings. Due to limited access near the outfall during low tide, UAV data were validated using SST measurements from Dataset 2 at stations P14 and P15, which were acquired at the same time as the thermal image, within the UAV’s coverage area (Figure 3.1). Notably, while P16 is also within the UAV’s coverage area, it was acquired approximately an hour earlier than the thermal image and was therefore excluded from the validation. The UAV readings were found to underestimate surface temperatures by an average of 3.14 °C, as shown in Figure 3.3a. Since datasets 1 and 3 were collected under similar weather conditions and altitudes (Table 3.1), a constant offset of +3.14 °C was applied to all UAV temperature measurements.

Landsat 9 satellite imagery was acquired from the United States Geological Survey (USGS) Earth Explorer platform (<https://earthexplorer.usgs.gov/>) for the dates listed in Table 3.1. This study used the TIRS Band 10, with a wavelength range of 10.6–11.19 μm and a spatial resolution of 100 m (USGS, 2022). All scenes were cloud-free and processed with the Semi-Automatic Classification Plug-in (SCP), which converted TIR data into temperature values using the satellite scene parameters (Congedo, 2016) (an overview of the procedure can be found in Appendix B). Ground-truthing of the SST data derived from Landsat 9 was performed using profiling measurements taken within a ± 45 -minute window of the satellite overpass time, a period during which horizontal temperature distribution remains relatively stable. The temperature difference between the profiler’s measured SST and the satellite-derived SST was calculated (Figure 3.3b and c), and the average offset was applied to correct the satellite readings for each scene.

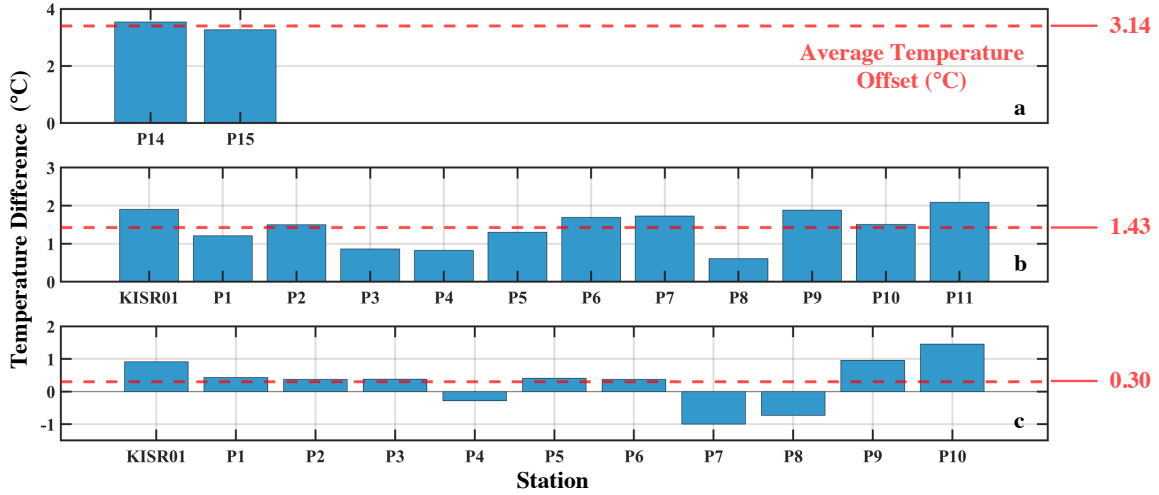


Figure 3.3: Temperature offset corrections for remotely sensed data. (a) Average temperature offset for UAV-derived temperatures based on overlapping in situ measurements from Dataset 2. (b) Temperature offset correction for satellite-derived temperatures in Dataset 1. (c) Temperature offset correction for satellite-derived temperatures in Dataset 3. The x -axis represents the profiling stations, and the offsets were calculated as the average difference between in situ measurements and remotely sensed temperatures.

3.2.2 3D Hydrodynamic Model Setup

The three-dimensional (3D) hydrodynamic model setup of KB and SB, shown in Figure 3.4, was developed using the Delft3D Flexible Mesh Suite (D-Flow FM), the integrated modelling software described in Section 2.2.1 and Appendix A.1. The 3D KB model setup closely follows the 2D configuration described in Section 2.2.1, with the addition of 10 σ -layers in the vertical direction, each representing 10% of the total water depth (see Appendix A.2 for details). The KB model provided water level forcing for the SB model and was validated against measurements at the KISR01 station (marked in blue in Figure 3.4), demonstrating agreement within an acceptable accuracy range for this chapter’s modelling objectives (Figure 3.5).

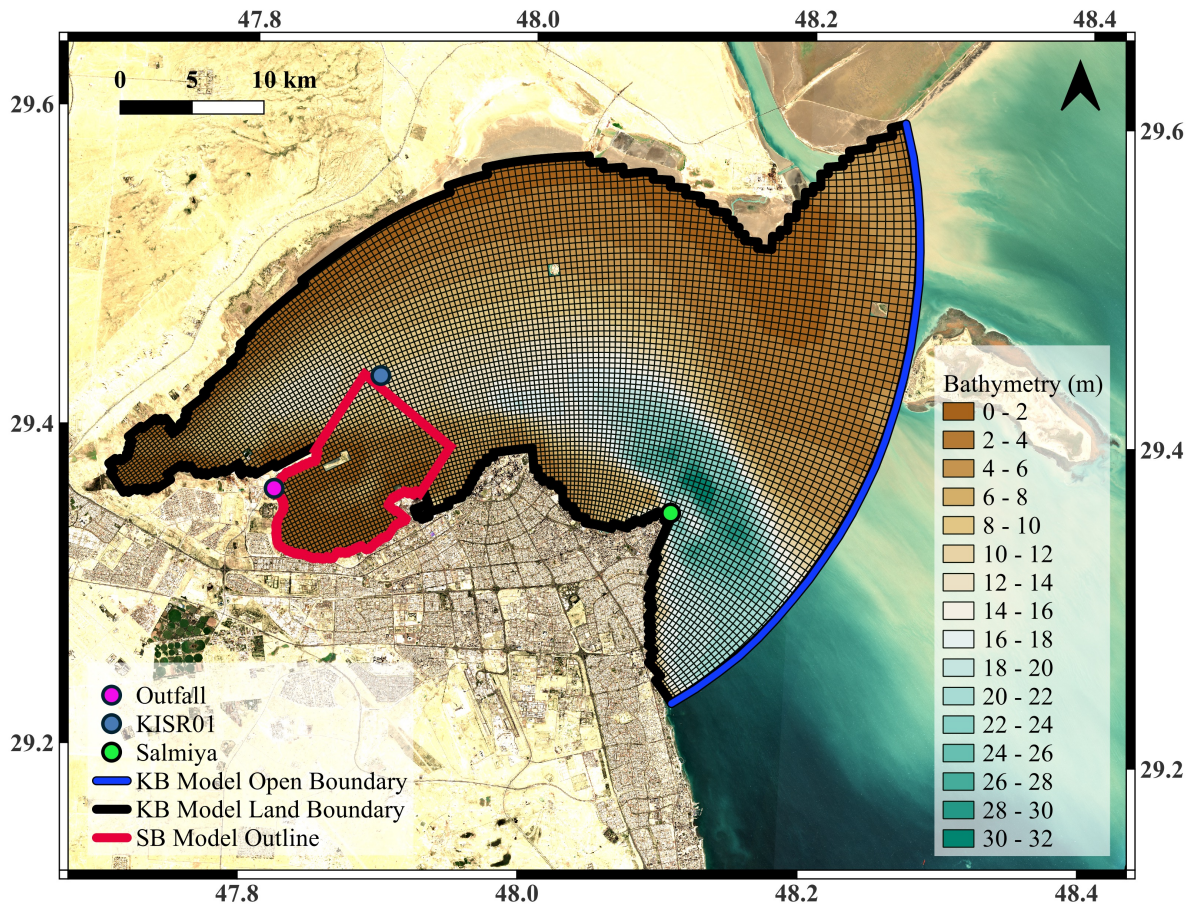


Figure 3.4: Numerical grid of the KB model, illustrating key features such as open and land boundaries, as well as the outline of the SB model domain. Reference points, including the DW outfall, KISR01 station, and Salmiya station, are marked. Bathymetry is shown relative to mean sea level.

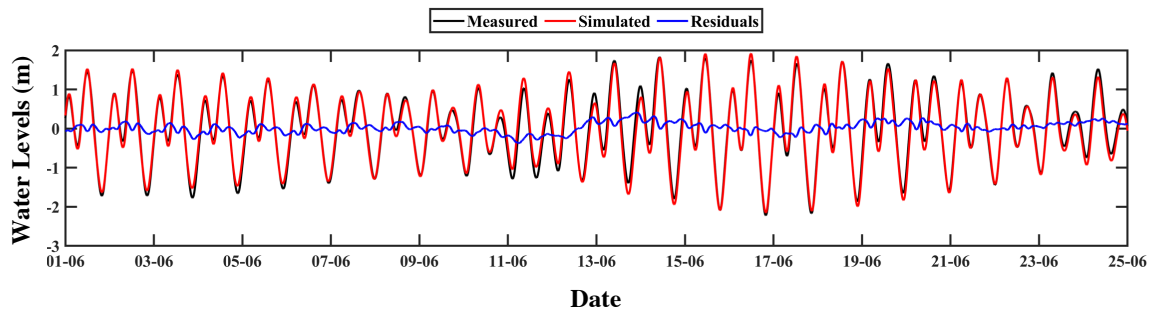


Figure 3.5: Water level validation of the KB model against measured water levels from KISR01 station. The residuals between the modelled and observed water levels indicate good model accuracy.

Building on the KB model, the SB 3D model features a high-resolution grid with progressive refinement near the outfall to capture localised hydrodynamic processes (Figure 3.6). The vertical grid matches the KB model configuration with 10 σ -layers. Open boundary conditions for water levels were sourced from the validated KB model. Vertically varying boundary conditions for temperature (Figure 3.7a) and salinity (Figure 3.7b) were applied using near-surface and near-bed observations from the KISR01 station. The upper 20% of the water column (from the surface) was forced with near-surface observations, while the lower 70% (from the seabed) was forced with near-bed observations, with linear interpolation applied in between. This approach is supported by preliminary analysis of observed profiles in the region, which indicate a characteristic mixed layer structure consistent with these depth-based divisions. Although the temperature and salinity forcing varied vertically, they were kept horizontally uniform along the boundary. The proximity of the KISR01 station to the SB boundary provided accurate state representation, especially during high tide, when the influence of the boundary on the modelled domain is at its strongest.

Heat exchange, meteorological forcing, and wetting and drying processes were applied in an identical manner to the KB model. The thermal discharge was initially modelled using a discharge rate (Q) of $197 \text{ m}^3/\text{s}$ and an excess temperature (ΔT) of 5°C , based on Pokavanich and Alosairi (2014)’s study on KB’s summer flushing characteristics. In this chapter, this configuration is referred to as the “baseline model”.

The simulation spanned the period from June 1 to 25, 2022, capturing interactions between tidal forces and the thermal plume over a full spring-neap tidal cycle. All models were allocated an appropriate spin-up period to achieve dynamic equilibrium and ensure stable initial conditions. All models were allocated an appropriate spin-up period to achieve dynamic equilibrium and ensure stable initial conditions. Further details on model setup, boundary prescriptions, forcing data, and spin-up procedures are provided in Appendix A.3.2.



Figure 3.6: Grid resolution of the SB model, with the blue line marking the open boundary. The grid is progressively refined near key features around the discharge location.

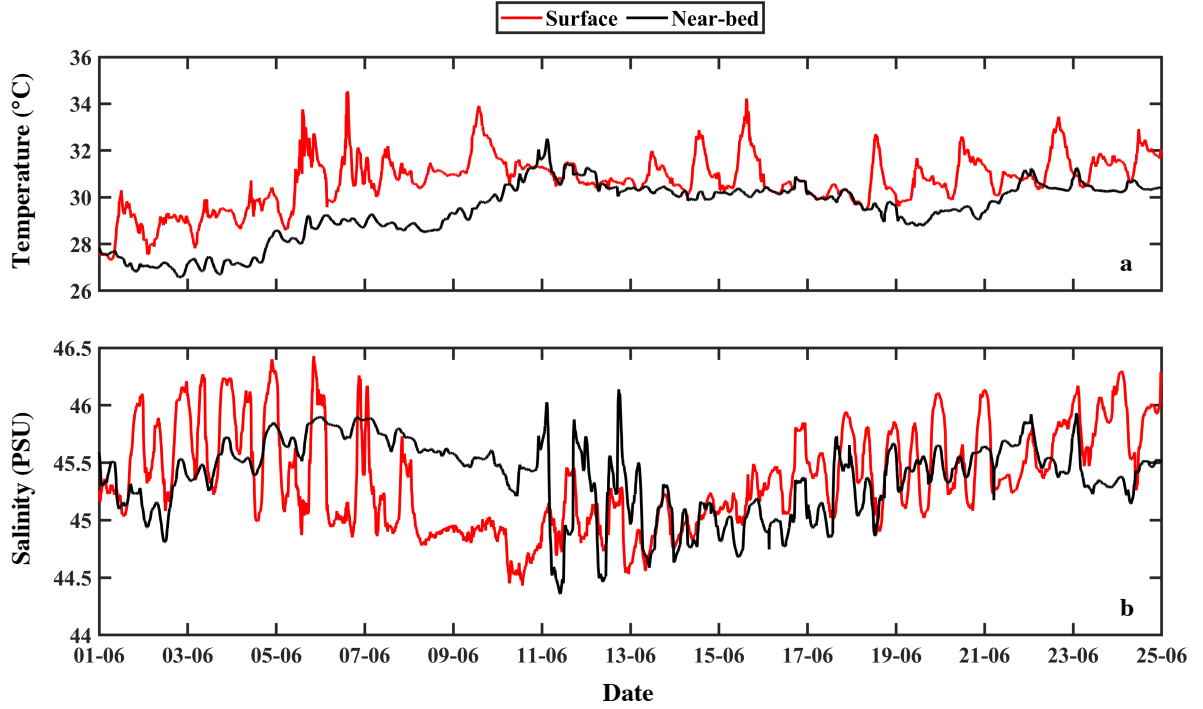


Figure 3.7: Measurements used to define the vertically varying boundary conditions in the SB model: (a) Surface and near-bed temperature, and (b) surface and near-bed salinity, both recorded at the KISR01 station.

The model setup adopted in this study is well suited for resolving intermediate- to far-field plume dynamics but has known limitations in simulating the near-field zone. Near-field processes are governed by rapid, momentum- and buoyancy-driven mixing immediately at the outfall, typically within 10 to 100 metres of the discharge point and over timescales of 100 to 1,000 seconds (Morelissen et al., 2013; Castro Faccetti, 2020). Due to the hydrostatic assumptions of models like D-Flow FM, which are typically used for far-field applications, the near-field region cannot be described accurately through these models alone. Additionally, the constraints on the grid resolution of these models introduce significant numerical diffusion, reducing the accuracy of effluent concentration estimates around the discharge point (Morelissen et al., 2013, 2016; Premathilake and Khangaonkar, 2019).

Although the SB model includes a high-resolution grid with refinement near the outfall, it does not explicitly resolve the discharge jet or initial momentum-driven mixing. The representation of the discharge as a point source, while computationally practical and consistent with typical far-field modelling practice, inherently simplifies the outfall geom-

etry and initial mixing characteristics, and introduces limitations in capturing fine-scale temperature gradients close to the discharge location. As a result, the model may underestimate near-field thermal maxima or over-smooth steep gradients due to numerical diffusion. These assumptions are reasonable for the study’s focus on plume evolution at the bay scale and its interaction with tides, but the simulation of the very near-field should be interpreted with awareness of these inherent model limitations.

3.2.3 Bayesian Optimisation

Consider the objective function $f(Q, \Delta T)$, which maps the uncertain model inputs to the sum of squared errors (SSE) between the measured and simulated temperatures. To minimise f using Bayesian optimisation, the workflow typically follows the steps outlined by Scyphers et al. (2024):

1. **Initial Evaluation:** The objective function f is first evaluated using a high-fidelity numerical model at n initial points $\{(Q_1, \Delta T_1), (Q_2, \Delta T_2), \dots, (Q_n, \Delta T_n)\}$ within a bounded search space. These points are typically generated using Latin Hypercube Sampling (LHS), a statistical technique that divides the parameter space into equally probable intervals to ensure uniform and efficient coverage.
2. **Surrogate Model:** A GP model is built as a surrogate for the objective function f by fitting a posterior distribution based on the initial evaluation points.
3. **Acquisition Function:** An acquisition function is computed over the posterior distribution to guide the selection of the next evaluation point, $(Q_{n+1}, \Delta T_{n+1})$.
4. **Evaluation:** The objective function $f(Q_{n+1}, \Delta T_{n+1})$ is evaluated using the high-fidelity numerical model.
5. **Update Surrogate Model:** The GP model is updated with the newly obtained data point $(Q_{n+1}, \Delta T_{n+1}, f(Q_{n+1}, \Delta T_{n+1}))$.
6. **Iteration:** Repeat steps 3–5 until the stopping criteria, such as convergence or a specified evaluation budget, are met.

The implementation of Bayesian optimisation in this study was performed using MATLAB’s `bayesopt` library (MathWorks, 2024b). An overview of the two main components of

Bayesian optimisation—the GP surrogate model and the acquisition function—as well as the chosen objective function to be minimised during the calibration process, is presented in the following sections.

3.2.3.1 Gaussian Process (GP)

Gaussian Processes (GP) are probabilistic tools widely used as surrogate models in Bayesian optimisation for representing objective functions as stochastic processes. A GP extends the Gaussian distribution to an infinite-dimensional domain, where any finite collection of function values follows a Gaussian distribution (Diessner et al., 2022). Similar to how a Gaussian distribution defines a probability distribution over a random variable using its mean and covariance, a GP defines a prior probability distribution over functions, fully characterised by its mean function and covariance function. The mean function represents the expected value of the function f at any point in the domain, while the covariance function, often referred to as the kernel, quantifies the similarity between input pairs in the search space (Brochu et al., 2010; Diessner et al., 2022; Scyphers et al., 2024). The kernel encodes prior assumptions about the function f , such as smoothness or periodicity (Snoek et al., 2012; Frazier, 2018; Scyphers et al., 2024). When data are added, the prior is updated to form a posterior distribution, allowing the GP to model complex, non-linear relationships effectively. The prior distribution of the GP is given by:

$$f|\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{K}),$$

where:

- $\mathbf{X} = [(Q_1, \Delta T_1), \dots, (Q_n, \Delta T_n)]^\top$: the input parameters, where n is the number of initial evaluations of the objective function.
- $\mathbf{m} = \mu_0(\mathbf{X})$: the prior mean vector of dimensions $n \times 1$. Typically, μ_0 is assumed to be zero.
- $\mathbf{K}$: the covariance matrix of dimensions $n \times n$, where each entry K_{ij} is computed using a kernel function k that quantifies the similarity between pairs of input parameters $(Q_i, \Delta T_i)$ and $(Q_j, \Delta T_j)$, such that $K_{ij} = k((Q_i, \Delta T_i), (Q_j, \Delta T_j))$.

The posterior predictive distribution of the GP for a new input point $\mathbf{x}_{n+1} = (Q_{n+1}, \Delta T_{n+1})$ is defined by the following mean and variance:

$$\begin{aligned}\mu_n(\mathbf{x}_{n+1}) &= \mu_0(\mathbf{x}_{n+1}) + \mathbf{k}_{n+1}^\top (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} (\mathbf{y} - \mathbf{m}), \\ \sigma_n^2(\mathbf{x}_{n+1}) &= k(\mathbf{x}_{n+1}, \mathbf{x}_{n+1}) - \mathbf{k}_{n+1}^\top (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{k}_{n+1},\end{aligned}$$

where

- $\mathbf{y}$: the vector of observed objective function values, with dimensions $n \times 1$.
- $\mathbf{k}_{n+1}$: a vector of covariances between the new input $\mathbf{x}_{n+1}$ and the observed points $\mathbf{X}$, with dimensions $n \times 1$. It is defined as:

$$\mathbf{k}_{n+1} = [k((Q_{n+1}, \Delta T_{n+1}), (Q_1, \Delta T_1)), \dots, k((Q_{n+1}, \Delta T_{n+1}), (Q_n, \Delta T_n))]^\top.$$

- $\sigma^2 \mathbf{I}$: the noise variance matrix, with dimensions $n \times n$, where σ^2 represents the variance of observational noise. For deterministic objective functions, $\sigma^2 = 0$.
- $k(\mathbf{x}_{n+1}, \mathbf{x}_{n+1})$: the self-covariance of the new input point, which is a scalar.

In this study, we implement the Automatic Relevance Determination (ARD) Matérn 5/2 kernel, which is the default kernel in MATLAB's 'bayesopt' function. The kernel models the covariance between the input parameters as:

$$k((Q_i, \Delta T_i), (Q_j, \Delta T_j) | \theta) = \sigma_f^2 \left(1 + \sqrt{5} r + \frac{5}{3} r^2 \right) \exp(-\sqrt{5} r),$$

where r is the scaled distance between the two input points $(Q_i, \Delta T_i)$ and $(Q_j, \Delta T_j)$, given by:

$$r = \sqrt{\left(\frac{(Q_i - Q_j)^2}{\ell_Q^2} \right) + \left(\frac{(\Delta T_i - \Delta T_j)^2}{\ell_{\Delta T}^2} \right)}.$$

Here, σ_f^2 represents the signal variance (or kernel amplitude), while ℓ_Q and $\ell_{\Delta T}$ are the length scales for Q and ΔT , respectively. The ARD mechanism allows the kernel to assign separate length scales to each input dimension, reflecting their relative importance in the model.

In MATLAB's GP implementation, the hyperparameters that characterise the covari-

ance function k are estimated by maximising the marginal log likelihood of the observed data, a process known as Maximum Likelihood Estimation (MLE) (MathWorks, 2024c). MATLAB’s GP optimises the hyperparameters $\theta = [\log \sigma_f, \log \ell_Q, \log \ell_{\Delta T}]$ to best fit the observed data. Further details on the Matérn kernel and hyperparameter optimisation can be found in the MATLAB documentation (MathWorks, 2024d).

3.2.3.2 Acquisition Function

An acquisition function defines the search strategy in Bayesian optimisation, utilising the posterior mean and variance from the GP to guide the search for the optimum. It balances exploration, which evaluates points in less-explored regions of the search space, and exploitation, which refines evaluations near the predicted optimum (Snoek et al., 2012; Frazier, 2018). Excessive exploration can lead to unnecessary iterations without significant improvement, while excessive exploitation risks convergence to local minima (Brochu et al., 2010).

In this study, we employed the Expected Improvement Plus (EI+) acquisition function provided by MATLAB’s Bayesian optimisation framework. The EI+ function builds upon the standard Expected Improvement (EI) function, which selects the next evaluation point based on the expected reduction in the objective function. The Expected Improvement function is defined as:

$$\text{EI}(x) = \mathbb{E}_P [\max(0, f(x_{\text{best}}) - f(x))],$$

where $f(x_{\text{best}})$ is the current minimum value of the objective function, evaluated using the high-fidelity model, and $f(x)$ represents the predicted value of the objective function at a candidate point x , as estimated by the GP model. The EI+ function modifies EI by incorporating an exploration term to avoid over-exploitation, particularly in regions where the algorithm detects excessive refinement near the predicted optimum. The EI+ function is expressed as:

$$\text{EI}^+(x) = \Delta(x)\Phi\left(\frac{\Delta(x)}{\sigma_P^+(x)}\right) + \sigma_P^+(x)\phi\left(\frac{\Delta(x)}{\sigma_P^+(x)}\right),$$

where $\Delta(x) = \mu_P(x_{\text{best}}) - \mu_P(x)$ represents the expected improvement, $\mu_P(x)$ is the

predicted posterior mean of the objective function at the candidate point, and $\sigma_P^+(x)$ is the inflated posterior standard deviation at the candidate point. The terms $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative distribution function and probability density function of the standard normal distribution, respectively.

In MATLAB’s implementation, the total variance at a candidate point x is expressed as:

$$\sigma_P^2(x) = \sigma_F^2(x) + \sigma^2,$$

where $\sigma_F^2(x)$ is the posterior variance of the objective function, and σ^2 is the posterior variance of the additive noise. To address over-exploitation, MATLAB’s EI+ function adjusts the kernel parameters by dynamically inflating the posterior variance, promoting exploration in less-sampled regions of the search space. This adjustment is controlled by the ‘ExplorationRatio’ parameter, denoted as t_σ , which was set to its default value of 0.5 in this study.

After each iteration, the algorithm evaluates whether the next candidate point x satisfies the condition:

$$\sigma_F(x) < t_\sigma \cdot \sigma.$$

If this criterion is met, the algorithm identifies the point x as over-exploited. If the algorithm detects over-exploitation, it modifies the kernel function by multiplying the hyperparameter θ by the number of iterations, as suggested by Bull (2011). This adjustment raises the posterior variance $\sigma_P^2(x)$ for points between observations, encouraging exploration. If the newly generated point is still over-exploited, the acquisition function further multiplies θ by an additional factor of 10 and tries again. This process is repeated up to five times, iteratively modifying the kernel parameters to generate a point x that is not over-exploited. Once a suitable point is identified, it is accepted as the next evaluation point. For more details on the implementation of the EI+ function, refer to the MathWorks documentation (MathWorks, 2024a).

3.2.3.3 Objective Function

The objective function in this chapter quantifies the discrepancy between observed and simulated water temperatures, forming the basis for model calibration. Multi-source

temperature datasets are used, and the sum of squared errors (SSE) is minimised under a transformation into a logarithmic scale to improve optimisation. This transformation reduces the impact of large errors and prevents the GP surface from becoming excessively flat, aiding in the identification of optimal solutions.

To ensure consistency with the numerical model’s vertical structure and address depth mismatches between observed and simulated data, vertical temperature profiles at each station were divided into ten layers, with each layer’s temperature represented as an average. The discrepancy between observed and simulated profiles was quantified using the SSE across all layers and stations. The objective function for the profiles dataset is given by:

$$\text{SSE}_{\text{profile}} = \log \left(1 + \sum_{s=1}^S \sum_{l=1}^L (T_{\text{obs},s,l} - T_{\text{sim},s,l})^2 \right), \quad (3.1)$$

where S is the total number of stations, L is the number of vertical layers per station (10 layers), $T_{\text{obs},s,l}$ is the observed temperature at station s and layer l , and $T_{\text{sim},s,l}$ is the simulated temperature at station s and layer l .

For satellite- and UAV-derived SST datasets, the numerical model grid was overlaid onto each dataset to extract temperature values overlapping with the centre of each grid cell, aligning with D-Flow FM’s representation of horizontal surface temperatures. Observed and simulated surface temperatures were then compared at these grid points to compute the SSE. The objective function for the satellite dataset is expressed as:

$$\text{SSE}_{\text{satellite}} = \log \left(1 + \sum_{g=1}^G (T_{\text{sat},g} - T_{\text{sim},g})^2 \right), \quad (3.2)$$

where G represents the total number of grid cells in the satellite domain, $T_{\text{sat},g}$ is the observed satellite-derived temperature, and $T_{\text{sim},g}$ is the simulated temperature at grid cell g .

Similarly, the UAV dataset was processed in the same manner, with the SSE for UAV-derived surface temperatures computed using:

$$\text{SSE}_{\text{UAV}} = \log \left(1 + \sum_{g=1}^G (T_{\text{UAV},g} - T_{\text{sim},g})^2 \right), \quad (3.3)$$

where G is the total number of grid cells in the UAV image domain, $T_{\text{UAV},g}$ is the observed UAV-derived temperature, and $T_{\text{sim},g}$ is the simulated temperature at grid cell g .

3.2.4 Performance Metrics

The Root Mean Square Error (RMSE) was used to evaluate the model's performance during calibration. For vertical profiles, the RMSE was calculated as:

$$\text{RMSE}_{\text{vertical}} = \sqrt{\frac{1}{S \times L} \sum_{s=1}^S \sum_{l=1}^L (T_{\text{obs},s,l} - T_{\text{sim},s,l})^2}. \quad (3.4)$$

For satellite-derived SST:

$$\text{RMSE}_{\text{satellite}} = \sqrt{\frac{1}{G} \sum_{g=1}^G (T_{\text{sat},g} - T_{\text{sim},g})^2}. \quad (3.5)$$

For UAV-derived SST:

$$\text{RMSE}_{\text{UAV}} = \sqrt{\frac{1}{G} \sum_{g=1}^G (T_{\text{UAV},g} - T_{\text{sim},g})^2}. \quad (3.6)$$

3.2.5 Impact of Parameter Uncertainty and Data Selection on Calibration Results

Model calibration is a critical step in numerical modelling, ensuring that simulations align with observed data. During this process, parameters are optimised to fit the calibration data; however, this optimisation does not guarantee that the estimated parameter values reflect their true physical counterparts (Ma et al., 2022). In some cases, model simplifications or physically unrealistic setups can force the calibration process to fit the data using implausible parameter combinations (Schwindt et al., 2023). When parameter estimates do not correspond to their real-world values, equifinality can occur, where multiple parameter sets produce similar model outputs due to complex interdependencies among calibration parameters within the numerical model (Piccioni et al., 2022). Recognising these challenges, the goal of calibration is not necessarily to recover exact physical values but rather to identify parameter sets that yield simulations closely matching observations

(Piccioni et al., 2022).

Studies by Piccioni et al. (2022) and Schwindt et al. (2023) highlight the importance of dataset selection in achieving accurate fits and physically meaningful calibration results. Piccioni et al. (2022) calibrated a hydro-ecological model using three target variables: total chlorophyll, cyanobacteria, and dissolved oxygen concentration. They found that calibrating with total chlorophyll data alone slightly improved chlorophyll predictions but resulted in highly inaccurate dynamics for oxygen and cyanobacteria. Similarly, Schwindt et al. (2023) calibrated a hydrodynamic reservoir model for eddy diffusivity, viscosity, and initial temperature using either flow velocity, water temperature, or both. When calibrating with a single variable, temperature, the model overemphasised diffusion, driving parameters for eddy diffusivity and viscosity to extreme, unrealistic values. Despite differing modelling contexts, both studies emphasise the importance of using diverse measured variables for calibration. This approach provides a more accurate representation of physical processes than relying solely on numerous measurements of a single variable.

While these studies focused on how calibration outcomes depend on the choice of different target variables, this section extends the discussion by examining how the selection of calibration parameters and the spatial characteristics of datasets measuring the same variable influence calibration results.

To systematically evaluate this, we employ two approaches. First, we conduct ensemble-based sensitivity analysis to assess how uncertainties in calibration parameters propagate through model outputs. Second, we perform synthetic data experiments to examine how different spatial datasets influence the shape of the GP posterior, providing insight into how dataset coverage affects the optimisation landscape.

Sensitivity Analysis of Calibration Parameters

Before applying Bayesian optimisation with real observations, it is crucial to assess the sensitivity of model outputs to the selected calibration parameters. Sensitivity analyses involve systematically varying input parameters within a physically realistic range and examining the corresponding changes in model outputs. This approach enhances our understanding of the relationships between input and output variables, enabling a streamlined calibration process by identifying parameters that significantly contribute to

model uncertainty while disregarding those with minimal or negligible impact (Williams and Esteves, 2017).

While multiple parameters, such as wind drag coefficients, heat flux coefficients, and eddy diffusivity, can influence simulated temperature fields (Baracchini et al., 2020b), as well as important model elements like bathymetry and bed roughness (Williams and Esteves, 2017), this study focuses on the discharge rate (Q) and the excess temperature (ΔT) of the effluents. These parameters are prioritised due to their dominant influence on simulated thermal plumes. Additionally, limiting the number of calibration parameters reduces the computational cost of Bayesian optimisation, as increasing dimensionality significantly expands the search space. A detailed justification for this choice, along with a sensitivity analysis of other potentially influential parameters, is provided in Appendix C.

To quantify the effect of Q and ΔT on model uncertainty, we generated a 10-member ensemble by perturbing each parameter by $\pm 20\%$ from the baseline model values ($197 \text{ m}^3/\text{s}$ and 5°C). The standard deviation of computed temperatures across the ensemble was analysed as a function of space to assess uncertainty propagation. Figure 3.8 illustrates the ensemble standard deviation at low and high tide, revealing a time-varying pattern that reflects tidal influences on the parameters' uncertainty region.

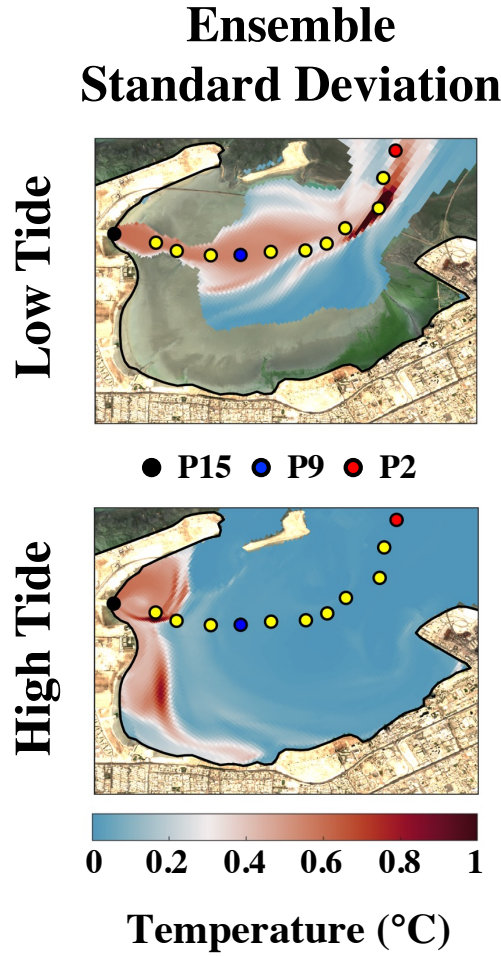


Figure 3.8: Ensemble standard deviation of simulated temperature, illustrating the influence of calibration parameters (Q and ΔT) on uncertainty propagation. The top panel corresponds to low tide conditions, while the bottom panel represents high tide. Stations P15 (closest to the outfall), P9 (intermediate distance), and P2 (farthest from the outfall) are highlighted in black, blue, and red, respectively.

If we define 0.3°C as the threshold of influence, Figure 3.8 shows that P15 (near the outfall) remains consistently affected by uncertainty in both parameters, regardless of tidal phase. In contrast, P9 and P2 are influenced only during low tide, when they fall within the thermal plume’s zone of influence. At high tide, these stations lie outside this region as inflow from the open boundary dominates. Understanding this tide-dependent, dynamic influence zone is essential for interpreting the calibration results presented in the following sections.

Sensitivity of GP Posterior to Dataset Coverage

To further assess the impact of data selection on calibration outcomes, a synthetic experiment was conducted under the assumption that the baseline model ($Q = 197 \text{ m}^3/\text{s}$, $\Delta T = 5^\circ\text{C}$) represents the true system state. The experiment examined how the spatial coverage of calibration data for the same variable (i.e. seawater temperature), along with the time-varying region of influence of the parameters, affects calibration results.

The search space for Q was defined between 10–500 m^3/s , while ΔT ranged from 1–10 $^\circ\text{C}$. Using Latin Hypercube Sampling (LHS), 20 parameter pairs were selected to ensure broad coverage, with four additional evaluations placed at the boundaries of the search space (e.g., (10,1), (500,1), (10,10), (500,10)).

To mimic in situ vertical profile collection, temperature values were extracted from the baseline model at 12 stations (locations marked in Figure 3.8) and treated as synthetic truth data. The objective function, defined as the sum of squared errors (SSE) in equation 3.1, was then computed relative to the baseline model. The GP posterior was evaluated for four cases: considering data from P15 only (closest to the discharge), P9 only (intermediate location), P2 only (farthest from the discharge), and all 12 stations combined.

Figure 3.9 demonstrates that station location significantly influences the shape of the objective function. During low tide, the region of influence—defined as the area where temperature deviations exceed 0.3°C (Figure 3.8)—extends across all stations along the survey line.

Panel (a) of Figure 3.9 shows that measurements from P15 provide a well-defined indication for ΔT but offer little constraint on Q . This is evident from the elongated region of low SSE along the horizontal axis, spanning almost the entire range of Q values, while narrowly constraining ΔT between approximately 4–6 $^\circ\text{C}$. This pattern suggests that calibration based solely on P15 would lead to high uncertainty in Q , necessitating extensive exploration by the search algorithm.

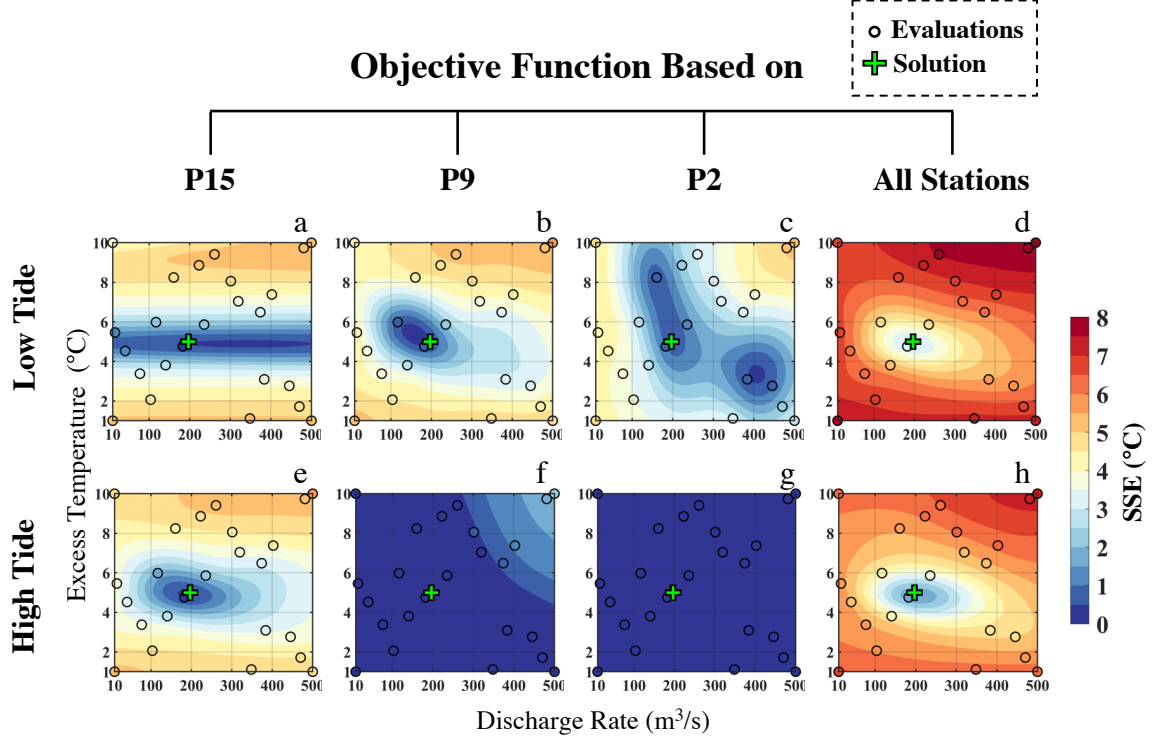


Figure 3.9: GP posterior surfaces for the objective function (SSE) in the synthetic case, illustrating the influence of calibration parameters (Q and ΔT) based on measurements from different stations. Panels (a–d) correspond to low tide, while panels (e–h) correspond to high tide. The objective function is computed using measurements from: (a, e) P15 (closest to the outfall), (b, f) P9 (intermediate distance), (c, g) P2 (farthest from the outfall), (d, h) all stations combined. The green marker represents the true parameter values, and open circles indicate evaluation points.

At P9 (Figure 3.9b), sensitivity shifts slightly toward Q , resulting in an oval-shaped contour elongated along the ΔT -axis. At P2 (Figure 3.9c), sensitivity to Q becomes slightly more dominant, with the objective function exhibiting vertically elongated contours for ΔT and broad regions of potential solutions. When all 12 stations are combined (Figure 3.9d), the GP posterior distinctly converges toward the true solution. This would enable the acquisition function to efficiently identify the optimal parameters during calibration.

However, in a realistic scenario, shallow waters during low tide may restrict access to stations near the outfall, limiting calibration data to intermediate stations (e.g., P2–P9). This limitation would result in vertically elongated contours and greater ambiguity for ΔT if the calibration were based solely on these datasets.

During high tide, the region of influence is confined to stations closest to the outfall. At

P15 (Figure 3.9e), the objective function remains ambiguous for Q , though its contours show slight improvement compared to low tide. At P9 and P2 (Figures 3.9f and 3.9g), now outside the influence zone, the objective function provides minimal guidance for calibration. When all stations are combined (Figure 3.9h), the optimal parameter solution region is further constrained, primarily influenced by data from stations within the high-tide zone of influence, shown earlier in Figure 3.8.

These results highlight the importance of considering both spatial coverage and tidal phases when selecting calibration data, as the evolving region where calibration parameters exert the greatest influence on model results directly impacts the clarity and accuracy of the solution space.

We recommend conducting a synthetic sensitivity analysis before applying Bayesian optimisation to real measurements. This allows modellers to assess the influence of each calibration parameter on model outputs and prioritise those with the strongest impact. Additionally, analysing the region of influence of calibration parameters can guide field data collection by identifying optimal sampling locations and acquisition times for calibration. Initial evaluations fitted to a GP can also help in selecting an appropriate acquisition function; for example, flat objective functions benefit from increased exploration rather than improvement-based functions (Diessner et al., 2022).

For large datasets, such as satellite imagery, this analysis can help identify areas contributing minimal information, allowing them to be masked to reduce computational burden. While this section mimicked vertical profile data as an example, the conclusions extend to other data types, as the key factor is measurement coverage rather than the specific data source. For instance, thermal images from satellites or UAVs limited to the outfall area would yield objective functions similar to those based on P15 data.

3.2.6 Model Calibration Based on In Situ and Remotely Sensed Observations

Building on the conclusions from the synthetic data analysis in Section 3.2.5, we apply these insights to real observations. Three multi-source temperature datasets, collected simultaneously (detailed in Section 3.2.1), are used to optimise the discharge parameters

of the DW plant in the SB model.

The Bayesian optimisation algorithm begins by evaluating the objective function (Section 3.2.3.3) using the high-fidelity hydrodynamic model for 35 parameter pairs $(Q, \Delta T)$. The search space is defined by discharge rates of 10–500 m³/s and excess temperatures of 1–10°C. These 35 points were selected using a combination of strategies: 20 points from LHS for broad coverage, 4 points at the search range corners, 1 point to represent the baseline model, and 10 points varying the baseline values by $\pm 20\%$ (the 10-ensemble data from the previous section).

While more initial evaluation points improve the GP’s representation of the objective function, Diessner et al. (2022) noted that this does not always enhance Bayesian optimisation performance, and the key challenge lies in dividing the predefined computational budget between initialising the GP and evaluations proposed by the optimisation algorithm. Given the computational demands of this study, an evaluation budget of 50—divided into 35 initial evaluations with a further 15 for optimisation—was deemed a reasonable compromise.

Figure 3.10 presents the GP approximation of the objective function for each data source. Black circles represent initial evaluations, blue crosses denote evaluations from the Bayesian optimisation algorithm, and the green triangle marks the optimal parameter pair with the lowest SSE. The corresponding optimal Q and ΔT values are listed in Table 3.2.

Although Datasets 1, 2, and 3 were collected at different tidal phases, the independent calibration results for vertical profiles and UAV measurements consistently converged to the same ΔT value. This convergence can be attributed to the high resolution of temperature measurements: vertical profiles provide fine vertical resolution, while UAV measurements offer high horizontal resolution, as evident in the contours of their objective functions in the first and second columns of Figure 3.10.

In contrast, calibration using satellite-derived SST data resulted in a coarser GP objective function, as seen in the third column of Figure 3.10, due to its lower spatial resolution (100 m) and the exclusion of areas near the coastline and outfall during preprocessing to mask regions affected by land temperature contamination. Despite these limitations, Bayesian optimisation successfully identified an optimal parameter pair, demonstrating

robustness even with lower-resolution data.

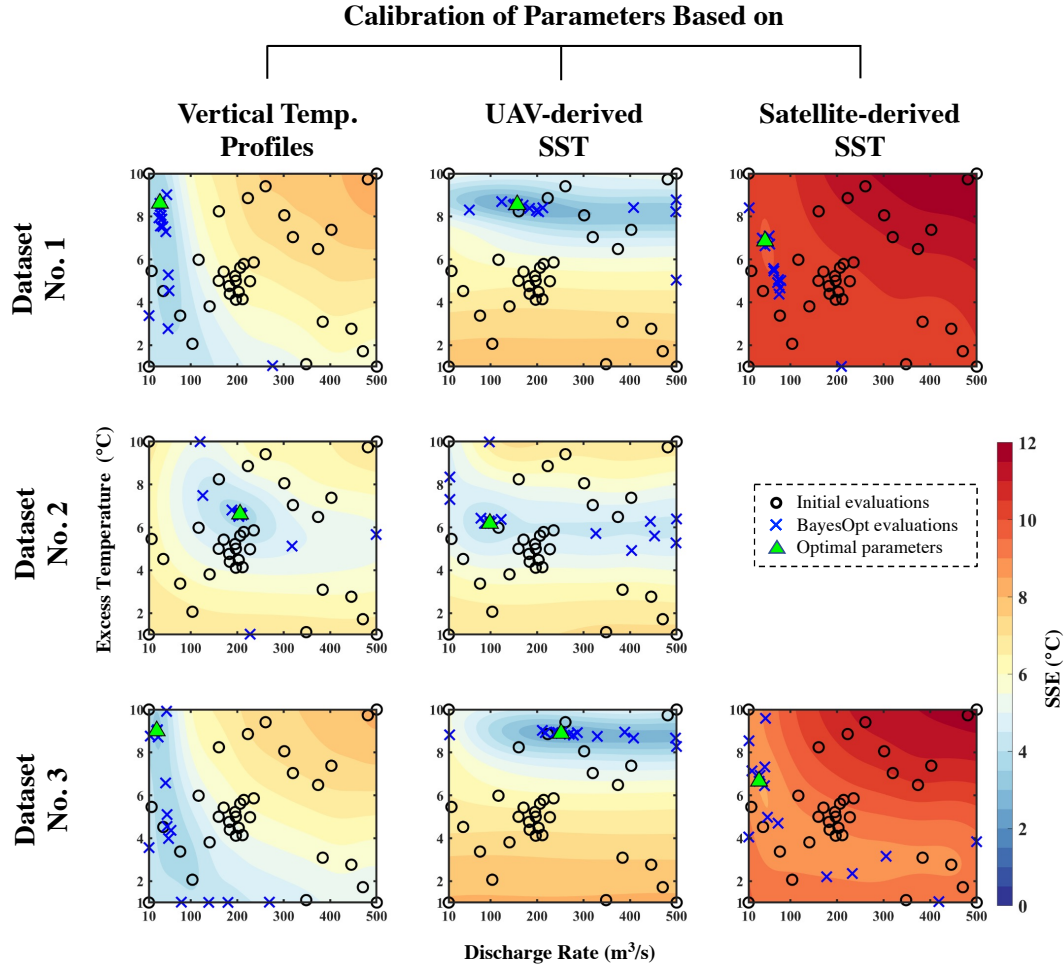


Figure 3.10: GP for the objective functions for the calibration of two parameters—discharge rate and excess temperature—based on three different data sources: vertical profiles of seawater temperature (first column), UAV-derived sea surface temperature (second column), and satellite-derived sea surface temperature (third column).

Table 3.2: Optimal parameter values for discharge rate (Q) and excess temperature (ΔT) obtained from calibration based on vertical profile measurements, UAV-derived SST, and satellite-derived SST for each dataset.

Dataset	Vertical Profiles	UAV	Satellite
No. 1	$Q = 33.67 \text{ (m}^3/\text{s)}$ $\Delta T = 8.62 \text{ }^\circ\text{C}$	$Q = 157.25 \text{ (m}^3/\text{s)}$ $\Delta T = 8.55 \text{ }^\circ\text{C}$	$Q = 45.35 \text{ (m}^3/\text{s)}$ $\Delta T = 6.86 \text{ }^\circ\text{C}$
No. 2	$Q = 205.95 \text{ (m}^3/\text{s)}$ $\Delta T = 6.62 \text{ }^\circ\text{C}$	$Q = 98.44 \text{ (m}^3/\text{s)}$ $\Delta T = 6.19 \text{ }^\circ\text{C}$	
No. 3	$Q = 26.75 \text{ (m}^3/\text{s)}$ $\Delta T = 9.00 \text{ }^\circ\text{C}$	$Q = 251.66 \text{ (m}^3/\text{s)}$ $\Delta T = 8.89 \text{ }^\circ\text{C}$	$Q = 32.39 \text{ (m}^3/\text{s)}$ $\Delta T = 6.66 \text{ }^\circ\text{C}$

As anticipated and consistent with the conclusions from the synthetic data in Section 3.2.5, the GP approximation of the objective function based on the low tide vertical profile data in Datasets 1 and 3 displayed an elongated shape along the ΔT axis, reflecting the spatial limitations of their coverage. In contrast, Dataset 2, collected during high tide, exhibited a narrower contour around the solution due to its expanded coverage. The GP for Datasets 1 and 3 prompted more exploratory behaviour from the Bayesian optimisation algorithm compared to Dataset 2, reflecting the increased uncertainty caused by limited proximity of the data to the outfall.

For UAV-derived SST data, restricted spatial coverage (recall Figure 3.1) resulted in an elongated zone of low SSE for Q , similar to what was observed with the synthetic data in the previous section. As a result, the optimisation algorithm exhausted the 15-evaluation budget without fully converging to a clear solution for Q .

When comparing the final parameters in Table 3.2, a consistent trend emerges: vertical profiles and UAV-derived SST data converge to similar ΔT values, but UAV data struggles with Q . Satellite-based Q values are consistently slightly higher than those from vertical profiles, while profile-based ΔT values are higher than satellite-based values.

The performance of each calibrated model will be assessed using the metrics in Section 3.2.4 across all data sources in each dataset. Each model will be evaluated both against the data used in its calibration as well as against synchronised data from other sources within the same dataset. This cross-validation approach will help determine whether the calibration is over-fitted to its specific data source or generalises effectively across multiple sources.

3.2.6.1 Performance of Calibrated Models Relative to Vertical Temperature Profiles

For each dataset, we evaluate the performance of four models: the baseline model, the model calibrated with vertical profiles, the model calibrated with UAV-derived SST, and the model calibrated with satellite-derived SST. Their performance is assessed against the vertical profile measurements from the respective dataset.

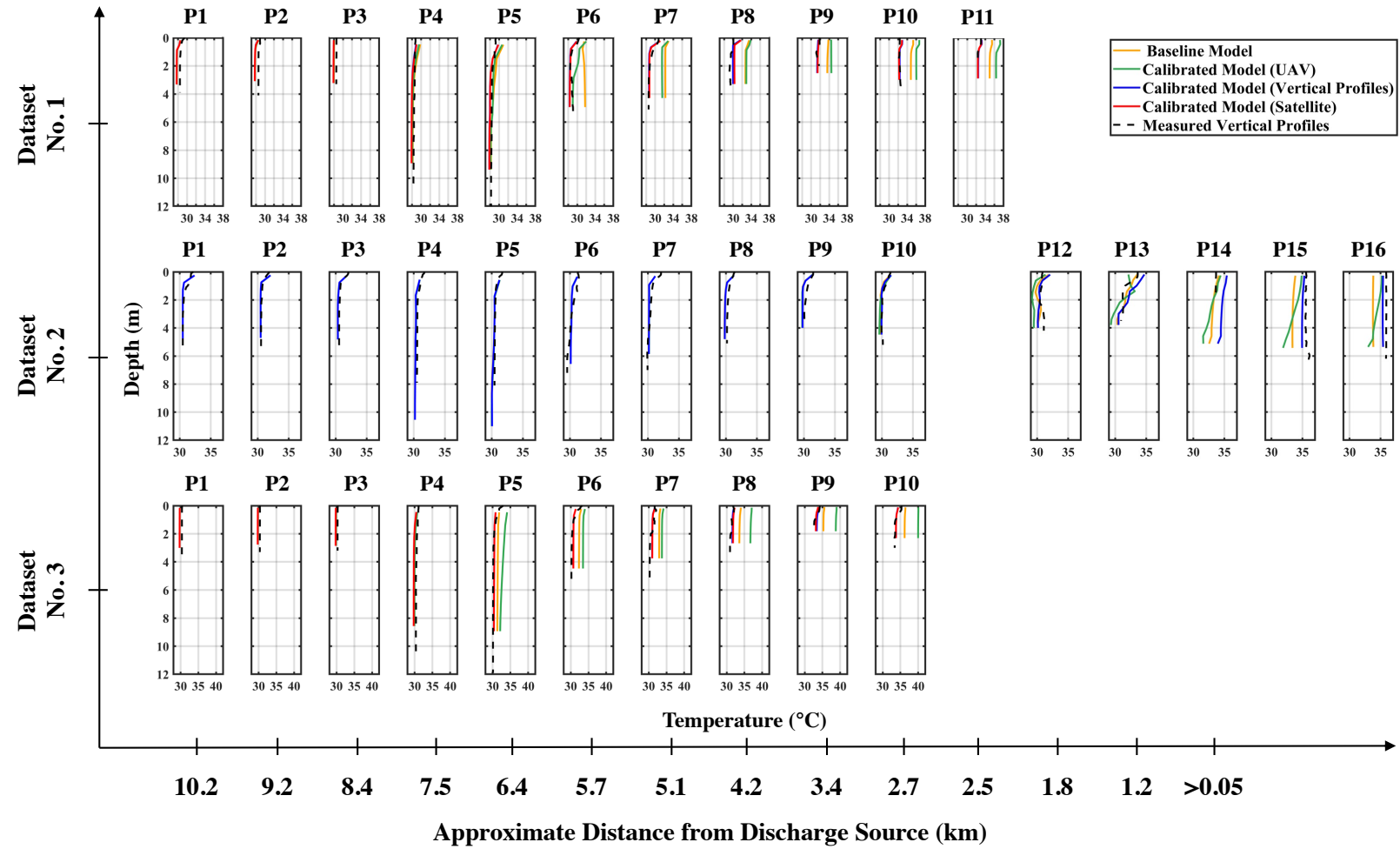


Figure 3.11: Observed (dashed line) vs. simulated (solid line) vertical temperature profiles for each dataset. Rows correspond to datasets, and colours represent different model configurations: baseline model (yellow), UAV-calibrated (green), profile-calibrated (blue), and satellite-calibrated (red). The x -axis indicates station distance from the discharge outfall, while the y -axis shows depth (m).

Figure 3.11 shows that in Dataset 1 (stations P1 to P5), Dataset 2 (stations P1 to P10), and Dataset 3 (stations P1 to P4), all models produce identical results. This occurs because these stations lie outside the thermal plume’s region of influence during the measurement acquisition time and are therefore unaffected by discharge parameters (recall Figure 3.8). Instead, their behaviour is governed by boundary conditions, leading to similar outputs across models despite variations in Q and ΔT . This observation is further supported by the vertical profile transects shown in Figure 3.2, which confirm the absence of elevated temperatures at these locations, indicating that the thermal plume did not reach them during sampling.

If used alone for calibration, these stations would provide little meaningful guidance for the Bayesian optimisation algorithm. Notably, all models closely match observed data at these locations, benefiting from vertically varying boundary forcings derived from KISR01 station.

To quantify performance, we calculated the RMSE for each model against observed profiles using equation 3.4, with results summarised in Table 3.3. Across all datasets, the vertical profile-calibrated model consistently achieved the lowest RMSE values. In Datasets 1 and 3, the satellite-calibrated model performed nearly as well, with substantial RMSE reductions relative to the baseline model of 75.54% and 67.70% for the profile-calibrated model, and 74.22% and 66.97% for the satellite-calibrated model. This is a case of equifinality, where different parameter sets produce comparable results, a limitation the Bayesian optimisation algorithm could not overcome. Despite differences in resolution, both data sources yielded comparable calibration outcomes.

The UAV-calibrated model, while achieving ΔT values similar to the profile-calibrated model (Table 3.2), performed worse due to inaccuracies in Q . This mismatch resulted in RMSE increases of 12.47%, 11.37%, and 114.62% relative to the baseline model for Datasets 1, 2, and 3, respectively. These results emphasise that accurate estimation of ΔT alone is insufficient; Q must also be accurately calibrated for optimal model performance.

Table 3.3: RMSE values ($^{\circ}\text{C}$) for the baseline model and models calibrated with vertical profiles, UAV-derived SST, and satellite-derived SST, evaluated against in situ vertical profile measurements.

Dataset	Baseline Model	Calibrated Model		
		Vertical Profiles	UAV	Satellite
No. 1	2.0972	0.5127	2.3589	0.5406
No. 2	0.8740	0.5401	0.9734	
No. 3	1.6437	0.5308	3.5280	0.5427

3.2.6.2 Performance of Calibrated Models Relative to Satellite-Derived Data

Before evaluating calibration performance, it is important to illustrate how the limitations of thermal satellite data (discussed in Section 3.2.1) are reflected in our datasets. As shown in Figure 3.12, the 100 m resolution of the satellite scenes is relatively coarse, exceeding the size of the outfall and resulting in a significant loss of temperature detail. This coarse resolution also leads to contamination from nearby land along the shoreline, requiring the cropping of these areas during preprocessing.

Additionally, satellite images were adjusted using vertical profile measurements (refer to Section 3.2.1) with an average offset correction. While this correction reduces discrepancies, it may leave residual errors. Satellite data also inherently include noise and may capture temperature patterns from processes that are not represented by the model, such as thermal exchange with tidal flats. Together, these factors introduce a baseline level of error in the RMSE that cannot be fully eliminated by the calibrated model.

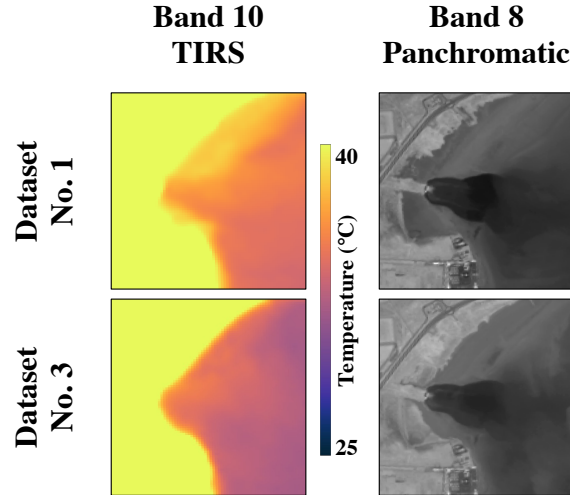


Figure 3.12: Comparison of thermal and optical satellite scenes for Datasets 1 and 3. The left column shows temperature data from the Landsat 9 Band 10 thermal infrared sensor (TIRS) with a 100 m resolution, while the right column presents high-resolution 15 m resolution panchromatic images from Band 8.

As shown in Figure 3.13, models calibrated with satellite and profile data displayed similar performance, achieving comparable RMSE values calculated using equation 3.5. The satellite-calibrated model performed slightly better, as it was evaluated against the same dataset used for its calibration. Relative to the baseline model, profile calibration reduced RMSE by 15.94% and 29.32% for Datasets 1 and 3, respectively, while satellite calibration achieved reductions of 15.46% and 30.83% for the same datasets.

When these calibrated models were compared against the vertical profile measurements, the data source which we consider to represent the true temperature of SB, both the satellite- and profile-calibrated models achieved greater RMSE reductions than when evaluated against the satellite data. This highlights the importance of data selection not only for influencing calibration outcomes (as discussed in Section 3.2.5) but also for affecting the values of the validation performance metrics.

Consistent with the findings from the previous section, the UAV-calibrated model performed worse due to the data source’s inability to provide information on Q . RMSE values increased relative to the baseline model by 22.22% and 136.84% for Datasets 1 and 3, respectively.

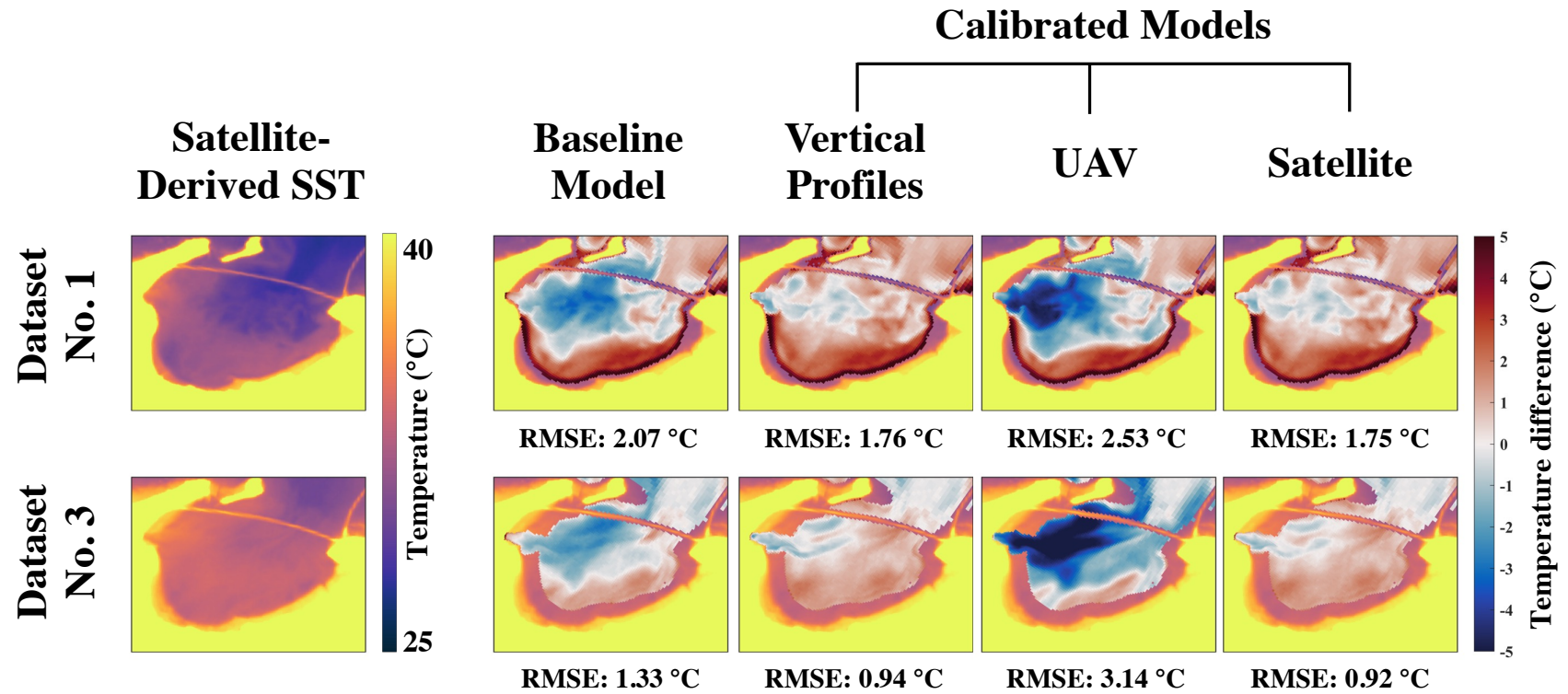


Figure 3.13: Comparison of temperature differences between satellite-derived and simulated SST for different models: baseline, vertical profile-calibrated, UAV-calibrated, and satellite-calibrated. The RMSE values for each model are displayed below the corresponding plots.

3.2.6.3 Performance of Calibrated Models Relative to UAV-Derived Data

Model validation using UAV thermal imagery requires careful consideration of its limitations. As shown in the first column of Figure 3.14, the high spatial resolution of UAV thermal images, while beneficial, creates a mismatch with the coarser model grid, posing challenges for capturing fine-scale near- and mid-field processes. For instance, cooler water pockets near the outfall, visible in Datasets 1 and 3, are absent in the model simulations. These differences are consistent with the model limitations discussed in Section 3.2.2, particularly the inability to fully resolve near-field dynamics. Additionally, the UAV images were adjusted using an average offset derived from vertical profile measurements (see Section 3.2.1), leading to inevitable residual errors when comparing model outputs to UAV-derived temperatures. .

To evaluate model performance against UAV thermal images, RMSE was calculated using equation 3.6, restricting the comparison to overlapping numerical grid cells within the UAV image extent. Temperature readings near the edges were masked to mitigate vignetting effects, a known sensor limitation that causes edges and corners to appear cooler than the centre (Yuan and Hua, 2022).

Among the calibrated models, the UAV-calibrated model achieved the lowest RMSE, followed by the profile- and satellite-calibrated models, with all calibrated models outperforming the baseline. As expected, the UAV-calibrated model performed best on this dataset, as it was validated using the same data employed for its calibration. However, its poorer performance when evaluated against other data sources highlights that the spatial extent of the data source significantly influences both calibration outcomes and apparent validation metrics.

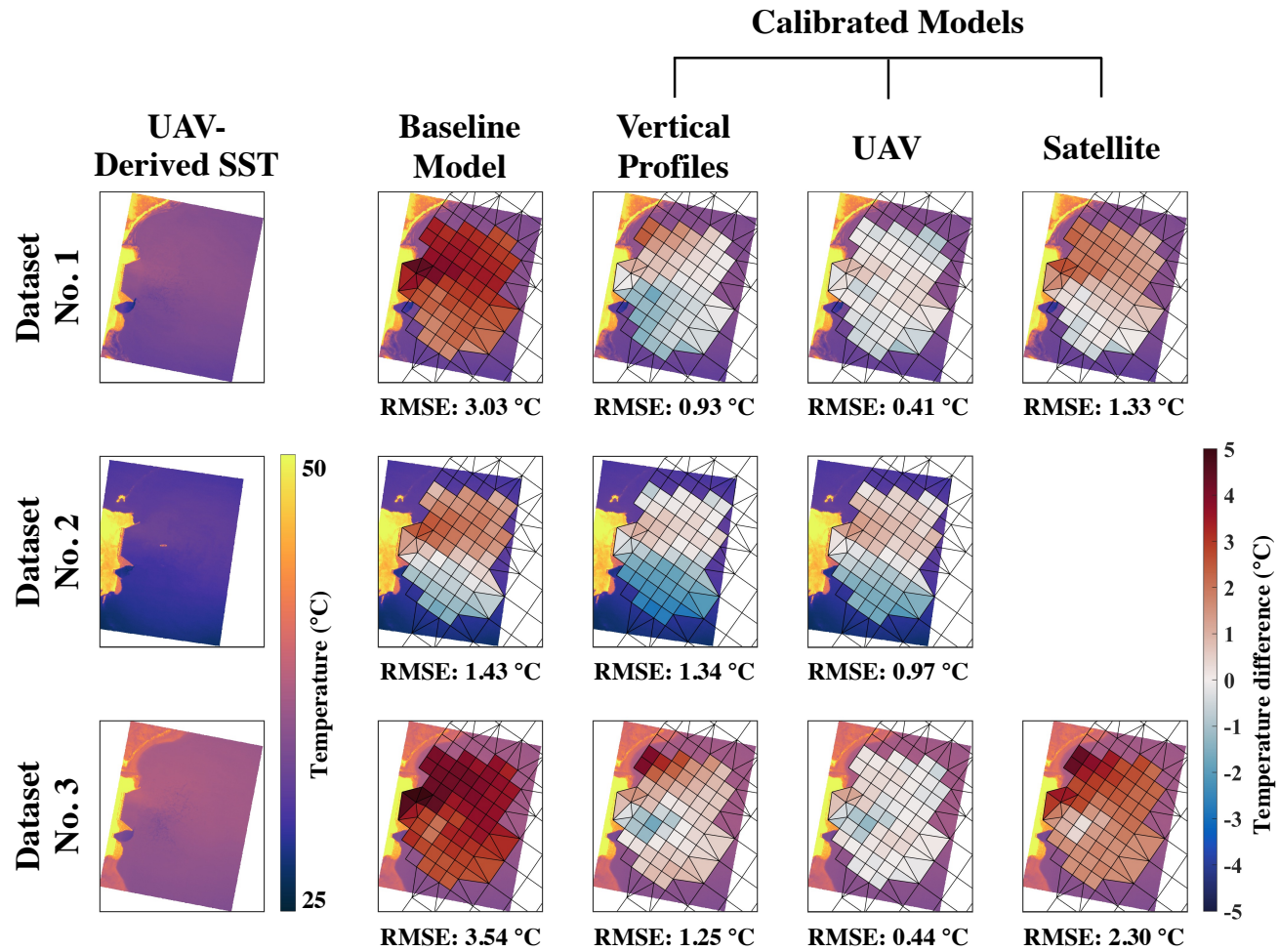


Figure 3.14: Comparison of temperature differences between UAV-derived and simulated SST for different models: baseline model, calibrated model using vertical profile data, calibrated model using UAV data, and calibrated model using satellite data. Each row corresponds to a different dataset, with RMSE values displayed below each model to indicate the accuracy of the simulated temperature relative to the measured values.

Table 3.2 highlights the discrepancies in Q values between UAV- and profile-based calibrations. In Dataset 1, the Q value from UAV calibration ($157.25 \text{ m}^3/\text{s}$) is nearly five times higher than that from profile-based calibration ($33.67 \text{ m}^3/\text{s}$), while in Dataset 3, the Q value from UAV calibration ($251.66 \text{ m}^3/\text{s}$) is approximately 9.4 times larger than the profile-based calibration ($26.75 \text{ m}^3/\text{s}$). Despite these significant differences in Q , the RMSE differences between the reduction achieved by both calibrated models are only 0.52°C and 0.81°C , respectively, illustrating the UAV’s data limitations not only in informing the calibration parameter (Q) but also in informing on the performance behaviour during validation.

This limitation also explains the higher RMSE of the satellite-calibrated model when evaluated against UAV data, despite performing similarly to the profile-calibrated model in Sections 3.2.6.1 and 3.2.6.2. As shown in Table 3.2, the Bayesian optimisation algorithm produced a different ΔT value for the satellite calibration, whereas the profile and UAV calibrations converged on identical ΔT values. This demonstrates that UAV data, due to its limited spatial extent, primarily provides information on a single calibration parameter, (ΔT), and offers limited utility for comprehensive calibration and validation.

3.3 Critical Assessment of Estimated Discharge and Excess Temperature Values

The calibration of the hydrodynamic model using Bayesian optimisation yielded estimates for the discharge rate (Q) and excess temperature (ΔT) that achieved low RMSE values against observed temperature data. However, low RMSE alone does not ensure that the estimated parameters are physically meaningful. Given that the calibration was performed using snapshot observations rather than continuous time series, there is a risk of overfitting, where the model optimises parameter values to fit isolated conditions. Additionally, parameter compensation may occur, where inaccuracies in one aspect of the model (e.g., bathymetry or boundary conditions) are absorbed by adjustments to Q or ΔT . To address these concerns, this section critically evaluates the validity of the estimated values, assessing their consistency with independent estimates and expected ranges.

3.3.1 Evaluation of Estimated Discharge

The discharge rate of the thermal effluent was initially estimated using simplified calculations based on plant efficiency and thermal characteristics, as detailed in Appendix D. These calculations yielded a plausible range of 169.8 to 238.68 m³/s, depending on the assumed thermal efficiency of the power plant. Additionally, Pokavanich and Alosairi (2014) reported a value of 197 m³/s, offering further context for the estimated Q values.

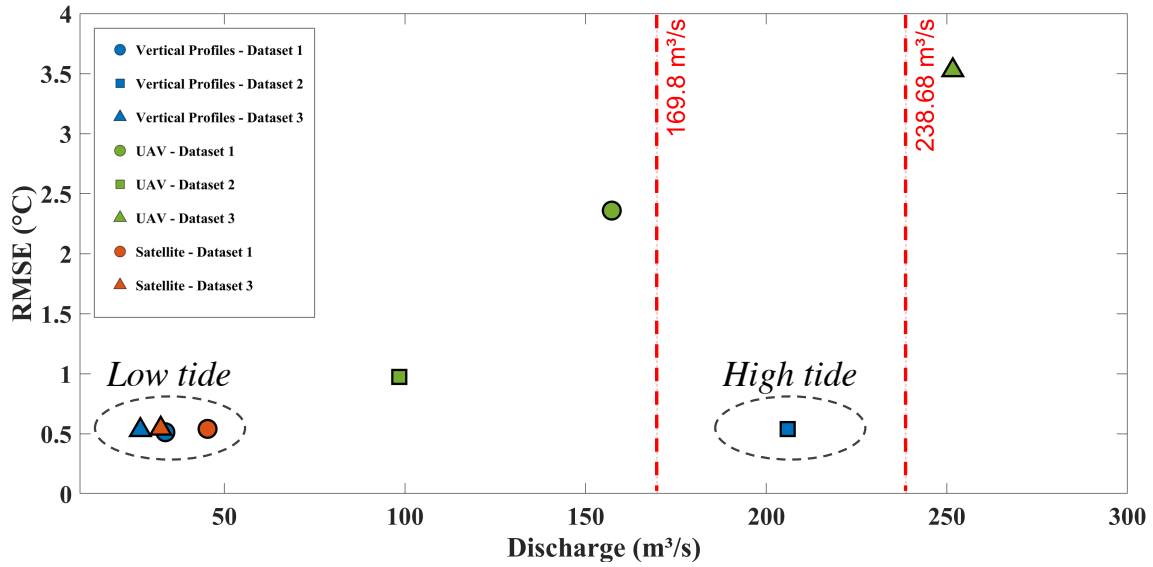


Figure 3.15: RMSE of calibrated models versus estimated discharge (Q). Triangles, squares, and circles represent Datasets 1, 2, and 3. Blue, green, and red markers indicate calibration using vertical profiles, UAV, and satellite data, respectively. Dashed red lines mark the plausible Q range (169.8 to 238.68 m³/s). Dashed ovals highlight low tide (left) and high tide (right) clusters.

Figure 3.15 shows the RMSE of the calibrated models compared to vertical profile measurements. Circles, squares, and triangles indicate Datasets 1, 2, and 3, respectively. Blue, red, and green markers correspond to calibration using vertical profiles, satellite, and UAV data. Key observations as follows:

- **UAV-Calibrated Values:** The green markers representing UAV-calibrated values consistently exceed the plausible range, resulting in high RMSE values. This indicates that the UAV's limited spatial coverage is insufficient for accurate Q estimation, as highlighted in Section 3.2.6.
- **Satellite and Profile-Calibrated Values:** Both data sources, with broader spa-

tial coverage, achieve low RMSE values. However, a clear pattern is observed:

- When observations were acquired under low tide conditions (Datasets 1 and 3), the calibrated Q values are significantly low ($< 50 \text{ m}^3/\text{s}$), falling well below the plausible range. This suggests possible overfitting.
- When observations were acquired under high tide conditions (Dataset 2), the estimated Q aligns with the plausible range, indicating a more realistic calibration.

These results suggest that during low tide, the calibration process may be compensating for inaccuracies in model configuration by artificially lowering Q to fit the observed temperature data. One potential explanation for the low Q values estimated during low tide is inaccuracies in the model's bathymetry. Specifically, if the bathymetry is too shallow, the thermal plume may extend further offshore than it should. In response, the calibration reduces Q , resulting in an artificially lower discharge value that confines the plume, better matching the observed temperature distribution.

This sensitivity analysis, shown in Appendix C (Figure C.4), demonstrates that even a uniform bathymetric adjustment of $\pm 0.5 \text{ m}$ can significantly alter the plume's extent. A shallower bathymetry (-0.5 m) allows the plume to spread further with the ebbing tide, while a deeper configuration ($+0.5 \text{ m}$) restricts it closer to the source. Although bathymetry appears to be a plausible factor, further investigation is needed, as other factors may also influence this behaviour.

3.3.2 Evaluation of Estimated Excess Temperature

Thermal effluent discharges from power and desalination plants in Kuwait Bay are known to elevate seawater temperatures. Lattemann and Höpner (2008) report that thermal desalination plants worldwide can increase seawater temperatures by 5°C to 15°C above ambient conditions, depending on plant design and discharge methods. In KB, Al-Rashidi (2009) specifically note that thermal plumes from power plants can raise seawater temperatures by up to 9°C above ambient. Pokavanich and Alosairi (2014) reported an excess temperature of 5°C for DW in their modelling study of KB.

The calibrated excess temperature (ΔT) values obtained in this chapter ranged from

6.19°C to 9.00°C (Table 3.2), which are consistent with these reported values.

3.4 Limitations and Considerations

This chapter is subject to several key limitations that may affect the reliability and generalisability of the calibrated model parameters. These limitations primarily relate to the following aspects:

- **Temporal Limitations:** The calibration was performed using short-duration observational datasets, each representing a brief snapshot of the thermal plume under specific tidal and atmospheric conditions. In contrast, the model simulates plume behaviour continuously over a 25-day period, capturing spring–neap tidal cycles, diurnal heat exchange, and boundary-driven variability. This mismatch raises questions about whether the calibrated parameters reflect persistent system characteristics or simply fit isolated conditions. For instance:
 - Daytime measurements may capture elevated surface temperatures due to atmospheric heating, particularly in shallow areas where background warming can mask the thermal signal of the effluent.
 - Data collected on windy days with strong surface mixing may capture atypical plume dynamics, deviating from normal conditions.

In both cases, the calibration may yield parameter estimates that reflect these transient conditions rather than the true behaviour of the thermal plume, resulting in potentially distorted or unrepresentative outcomes.

- **Limited Calibration Scope:** The calibration was focused solely on two parameters: discharge rate (Q) and excess temperature (ΔT), while other influential numerical parameters and model configuration aspects were excluded. This approach increases the risk of parameter compensation, where inaccuracies in these uncalibrated components are absorbed into the optimised values of Q and ΔT . As discussed in Appendix C and the previous section, if these unoptimised factors are not accounted for, they can compromise the robustness and physical validity of the calibrated parameters.

To improve the reliability and generalisability of calibration results, future work should incorporate extended, multi-temporal observations and consider expanding the calibration to include additional parameters. Long-term datasets are particularly important to evaluate whether the optimised values generalise across varying conditions or simply reflect a best fit to isolated snapshots.

3.5 Conclusion

Numerical models are crucial for understanding and predicting the impact of thermal effluents on coastal environments, but their accuracy relies on properly representing discharge characteristics, which are subject to uncertainty. Bayesian optimisation, a machine-learning-based approach, efficiently addresses this challenge by minimising errors between observed and simulated temperatures, using a Gaussian Process as a surrogate model and an acquisition function to balance exploration and exploitation of the solution space. In this chapter, we apply Bayesian optimisation to calibrate a thermal effluent model and explore how different data sources measuring the same variable affect calibration outcomes.

Three extensive field campaigns were conducted in Sulaibikhat Bay (SB), a shallow tidal bay receiving continuous discharge from the Doha West (DW) co-desalination and power plant. SB’s dynamic tidal environment makes it an ideal test site, providing findings which should be applicable to similar tidal environments. Temperature observations were collected simultaneously from three sources: in situ vertical profiles, thermal UAVs, and Landsat 9 satellite, with some overlapping spatial coverage. The campaigns spanned different tidal phases to monitor the thermal plume’s interaction with tidal dynamics. Bayesian optimisation, performed using MATLAB’s BayesOpt tool coupled with the Delft3D Flexible Mesh (D-Flow FM) model, calibrated two key parameters: discharge rate Q and excess effluent temperature ΔT , defining the discharge as a source point in the hydrodynamic model.

Each data source was used independently for calibration, producing optimal values for Q and ΔT . A cross-validation approach tested these values across all subsets within the same dataset, enabling assessment of model performance while reducing over-fitting risk,

and identifying performance variability due to the unique characteristics of each data source.

A key challenge in the calibration process was the sensitivity of model outputs to the calibration parameters, as changes in these parameters directly influenced the simulated thermal plume. The plume, representing the region of influence of the parameters in the modelled domain, is dynamically advected and dispersed by tidal forces, creating a time-varying region of impact. Calibration outcomes for discharge rate Q and excess temperature ΔT were significantly influenced by how effectively each data source captured this shifting region, as well as by the resolution of temperature measurements provided by each source.

Despite differences in plume coverage, UAV and vertical profile data consistently converged to similar ΔT values due to their high measurement resolutions. Vertical profiles provided detailed temperature readings along the water column, while UAVs offered high-resolution measurements across the horizontal plane, resulting in sharper gradients in their GP objective functions. In contrast, satellite data, with its lower spatial resolution of 100 m, produced a coarser GP approximation and arrived at different ΔT values.

However, sharper GP contours did not necessarily guarantee reliable calibration outcomes. The limited spatial coverage of UAV data, focused near the outfall, hindered its ability to inform the calibration of Q , effectively guiding only ΔT . Conversely, the broader spatial coverage of satellite data, despite its lower resolution, enabled it to inform the calibration of both Q and ΔT due to its more extensive coverage across the region of influence.

Vertical profile and satellite data, with comparable plume coverage but differing resolutions, produced distinct pairs of Q and ΔT . Despite these differences, both models achieved nearly identical validation performance, reducing errors significantly compared to the pre-calibrated model. This outcome highlights equifinality—a phenomenon where different parameter sets yield similar results in calibration.

Cross-validation further revealed that the apparent values of the validation performance metrics heavily depend on the dataset used. Assuming in situ measurements represent the true temperature of SB, models calibrated with profile and satellite data achieved over 60% error reduction when validated against these in situ measurements. However, when

the same high-performing calibrated models were validated against satellite data, their error reduction dropped to less than 30%. This discrepancy highlights the potential for misleading assessments of model performance and emphasise the importance of carefully selecting validation datasets to ensure accurate evaluations of model accuracy.

This chapter emphasises the importance of assessing the region of influence for the parameters of interest within the model domain prior to initiating calibration. By identifying this region, researchers can focus data collection on areas most affected by these parameters, thereby maximising the effectiveness of Bayesian optimisation. Equally important is the critical evaluation of calibrated values, as low RMSE alone does not guarantee their physical plausibility. This study revealed instances of parameter compensation, where the optimisation process adjusted parameter values to account for unrepresented or inaccurate model characteristics. These insights provide a basis for more reliable model calibration in environmental applications.

Given the limitations of UAV data in parameter estimation and model performance assessment, the following chapter shifts focus to a complementary method, large-scale particle image velocimetry (LSPIV), a cost-effective alternative technique for discharge estimation.

Chapter 4

Large-Scale Particle Image Velocimetry for Discharge Estimation

Highlights

- LSPIV represents a viable method for estimating uncertain outflows from coastal power plants in dynamic tidal environments, extending its application beyond riverine systems.
- Integrating LSPIV-derived discharge into a hydrodynamic model improved thermal plume simulation and complemented broader calibration efforts effectively.
- Discharge estimates varied with image sampling frequency and interrogation area size, though this variability had minimal impact on model performance.

Summary

This chapter examines the application of Large-Scale Particle Image Velocimetry (LSPIV) for measuring surface velocities and estimating discharge in the dynamic outflow of a coastal power plant. The study investigates the sensitivity of LSPIV-derived velocity and discharge estimates to two key processing parameters: image sampling frequency

and interrogation area size. These discharge estimates were incorporated into a three-dimensional hydrodynamic model to simulate the thermal effluent, allowing for an assessment of the influence of LSPIV parameter choices on model accuracy. The results highlights the practicality and cost-effectiveness of UAV-based LSPIV as a non-contact data collection method in challenging coastal environments. While the method demonstrated reliable performance, particularly when combined with other data sources and calibration techniques, the study acknowledges certain limitations and proposes recommendations for refinement in future applications. The findings validate LSPIV as a viable tool for discharge estimation in power plant outflows within coastal environments, marking a novel expansion of its use beyond riverine systems.

Publications and Disclaimer

The content within this chapter is based on:

N. Alsulaiman, J. Tlhomole, Y. Alosairi, M. Van Reeuwijk, and M. D. Piggott. *UAV-Based Large-Scale Particle Image Velocimetry for Optimizing Coastal Thermal Effluent Modeling*. In preparation for submission to *Water Resources Research*, 2025.

4.1 Introduction

Accurate characterisation of flow dynamics is essential for developing reliable hydrodynamic models. Velocity measurements, in particular, provide valuable insights into local flow behaviour, informing decisions on model setup, calibration, and validation. When combined with bathymetric data, these measurements facilitate the estimation of discharge flow rates. Discharges carrying heat, salinity, or other concentrations serve as source terms in coastal water models, shaping the simulation of key transport processes within these environments. Enhancing methods for collecting velocity and discharge data is therefore instrumental for improving model accuracy.

Traditionally, velocity measurements are commonly acquired using Acoustic Doppler Current Profilers (ADCPs). ADCPs have evolved over the past four decades into versatile tools capable of measuring velocity and discharge in both deep and shallow waters. ADCPs operate by transmitting high-frequency acoustic pulses, which reflect off suspended particles moving with the flow. Water velocity is calculated using the Doppler shift in frequency of the returned signals (Mueller et al., 2013; Teledyne Marine, 2021). This technology has significantly reduced the time required for discharge measurements and enabled spatial and temporal resolutions that were previously unattainable (Mueller et al., 2013; Lewis et al., 2018). However, ADCPs face notable limitations. They cannot measure velocities near the water surface, bed, or banks due to acoustic noise, transducer draft, and beam-spreading requirements, leaving these regions unmeasured and reliant on interpolation (Lewis et al., 2018). Furthermore, these methods remain labour-intensive, costly, and dependent on direct contact with the water body, making them impractical in high-flow conditions, inaccessible areas, or during extreme hydrological events such as floods (Pumo et al., 2021; Jolley et al., 2021; Bandini et al., 2021). Non-contact sensors for velocity measurements, such as surface velocity radars, have emerged as a viable alternative. However, their high deployment costs and operational limitations, such as reliance on surface roughness and susceptibility to environmental noise, have hindered their widespread adoption (Jolley et al., 2021).

The challenges associated with traditional methods have driven the adoption of image-based approaches for non-contact velocity measurement, as they offer significant advan-

tages such as reduced operational risks, lower costs, and applicability in inaccessible locations during extreme hydrological events (Dal Sasso et al., 2021; Bodart et al., 2024). Advances in UAV technology, coupled with increased affordability, have further enhanced the potential of these image-based approaches. UAVs provide greater spatial coverage and flexibility over traditional instruments, making them an increasingly attractive option for collecting observational data (Lewis et al., 2018; Bandini et al., 2021; Fairley et al., 2022).

The inference of velocity fields from the displacement of objects or patterns observed in consecutive images has been extensively studied in experimental fluid dynamics (Thomole et al., 2024). Particle Image Velocimetry (PIV), a widely adopted image-based technique originally developed for controlled laboratory settings, has been adapted for large-scale applications in natural environments. These adaptations, known as Large-Scale Particle Image Velocimetry (LSPIV), extend the principles of PIV to field-scale studies, enabling the measurement of environmental flows in complex settings (Fujita et al., 1998; Bandini et al., 2021; Pumo et al., 2021).

LSPIV operates from an Eulerian perspective, estimating velocity by analysing patterns or groups of particles within defined image regions. The surface velocity field is indirectly obtained by tracking the motion of floating tracer particles, which can either be naturally present or artificially introduced into the flow. These tracers are assumed to move with the local flow velocity, making them effective proxies for surface flow dynamics. This approach accommodates a wide variety of naturally occurring floating tracers with varying shapes and sizes, enhancing its applicability for environmental flow measurements (Fujita et al., 1998; Tauro et al., 2017; Bandini et al., 2021; Pumo et al., 2021).

The method divides an image into smaller subregions, known as interrogation areas (IAs), and calculates particle displacement using cross-correlation techniques. The maximum correlation point within each IA indicates the most probable displacement between consecutive frames. Velocity is then calculated by dividing this displacement by the time interval (Δt), resulting in one velocity vector per IA (Fujita et al., 1998; Stamhuis and Thielicke, 2014; Thielicke and Sonntag, 2021; Jolley et al., 2021).

Accurate LSPIV analysis depends on the presence of suitable tracers on the water surface, as these features enable meaningful cross-correlation analysis. Tracers must exhibit suf-

ficient contrast and distinctness to be detected. Their spatial and temporal distribution plays an important role in shaping the quality of LSPIV results (Jolley et al., 2021).

In addition to tracer characteristics, the success of LSPIV is heavily influenced by the processing parameters selected during the analysis stage, particularly the image sampling frequency (or feature extraction rate) and the IA size (Lewis et al., 2018; Pearce et al., 2020; Pumo et al., 2021; Bodart et al., 2024). Optimising these user-controlled parameters is crucial for maximising the accuracy of velocity estimation, particularly in dynamic environments where other conditions, like tracer density within the captured scene, may be difficult to control.

The choice of optimal image sampling frequency is closely tied to the local flow dynamics. High flow velocities require higher frame capture rates from the imaging device to ensure tracer displacement is adequately captured between consecutive frames. Conversely, lower flow conditions benefit from reduced frame rates, allowing sufficient tracer movement to occur between image pairs (Liu et al., 2021a; Pumo et al., 2021). Striking the right balance in image sampling frequency is essential for producing reliable velocity estimates. On one hand, sampling frequencies that are too low relative to the flow velocity can result in excessive tracer displacement, leading to pattern deformation, reduced correlation accuracy, and an increased likelihood of spurious velocity vectors (Lewis et al., 2018; Bodart et al., 2024). On the other hand, excessively high sampling frequencies in low-velocity conditions may result in minimal tracer displacement being captured, which undermines the matching algorithm’s ability to detect the displacement and infer velocity from it (Pumo et al., 2021).

Improper choices in sampling frequency can also exacerbate unintended interactions with frequency characteristics, such as camera movement caused by wind or wind-induced surface waves in the flow, introducing distortions or errors into the analysis (Lewis et al., 2018). The selected sampling frequency must ensure sufficient tracer displacement to distinguish true velocities from near-zero values while avoiding excessive displacements that exceed the IA size, which would compromise correlation accuracy (Pumo et al., 2021; Pearce et al., 2020).

Traditionally, laboratory PIV studies recommend that particle displacements between

frames should not exceed 25% of the IA size to maintain accurate vector recovery (Raffel et al., 2007). However, in outdoor LSPIV applications, stricter thresholds are recommended, limiting frame-to-frame displacements to 15% of the IA size to account for the dynamic and less controlled nature of such environments (Pumo et al., 2021).

Another important factor influencing the accuracy of LSPIV results is the IA size. Smaller IAs provide higher spatial resolution for analysis, enabling detailed velocity mapping, but they risk insufficient tracer density within the area, which can degrade cross-correlation performance and produce unreliable velocity vectors (Lewis et al., 2018; Liu et al., 2021a). In contrast, larger IAs are more likely to capture sufficient tracer information, reducing errors in velocity estimation. However, excessively large IAs can smooth out small-scale flow structures, leaving finer motion details unresolved when these structures are smaller than the IA size (Kim et al., 2013; Fujita et al., 1998). The optimal IA size must therefore be calibrated to balance these trade-offs: large enough to ensure adequate tracer density for cross-correlation, yet small enough to preserve finer flow details and avoid ambiguities (Pearce et al., 2020).

Importantly, IA size and image sampling frequency are directly related: if tracer displacement exceeds the IA size between frames, cross-correlation accuracy diminishes, resulting in unreliable velocity estimates (Pearce et al., 2020). This interdependence underscores the need to calibrate these parameters together, rather than in isolation.

Large-scale velocimetry methods have been widely applied in the literature, primarily in riverine environments, where they are commonly used for reconstructing velocity fields and estimating discharge (Le Coz et al., 2010; Lewis et al., 2018; Pearce et al., 2020; Bandini et al., 2021; Bodart et al., 2024; Torres et al., 2024). In addition to the use of optical imagery, alternative approaches such as thermal imagery (Kinzel and Legleiter, 2019; Legleiter et al., 2017; Schweitzer and Cowen, 2021) and satellite-based video remote sensing have also been explored for river monitoring applications (Masafu et al., 2023).

Applications outside riverine environments remain relatively limited and are largely in the exploratory stage. Tlhomole et al. (2023) explored the use of satellite-based velocimetry to infer glacier velocities from remote sensing imagery. In coastal settings, Dérian and Almar (2017) tracked foam patterns in the surf zone to estimate nearshore surface currents

and monitor flash rip events. Fairley et al. (2022) and McIlvenny et al. (2023) applied UAV-based LSPIV to assess tidal stream energy sites, finding that the method’s accuracy relied on favourable environmental conditions with sufficient surface texture for effective tracer tracking. Ruessmeier and Zielinski (2021) investigated the use of combined visual and infrared camera to increase the likelihood of recording usable data for tidal currents observations even in challenging ocean conditions with limited naturally present tracer textures.

In industrial applications, Wang et al. (2014) used LSPIV to measure surface velocities near a power plant intake, and to validate a numerical model to assess alternative designs for reducing ice transport and blockage risks. Matias et al. (2021) used LSPIV to characterise flow dynamics, including estimating discharge rates, in a wastewater treatment plant and proposed design alternatives to improve flow hydraulics in the contact chamber.

While LSPIV has emerged as a preferred method for estimating environmental flows, it still faces challenges that limit its broader application. Environmental factors such as lighting, tracer characteristics, wind, and surface wave interactions can significantly affect the accuracy of velocity estimates (Jolley et al., 2021; Manfreda et al., 2024). Furthermore, operator decisions during video acquisition and data processing introduce errors due to the method’s sensitivity to the analysis parameters (Pearce et al., 2020; Bodart et al., 2024). The lack of standardised protocols for image acquisition and analysis further complicates implementation, often requiring extensive training to ensure accurate parameter selection and result interpretation (Pearce et al., 2020; Pumo et al., 2021).

Given the advancements and limitations of LSPIV, this chapter addresses several key research gaps. It first investigates the feasibility of using LSPIV to estimate surface velocities in the dynamic and turbulent outflows of a coastal power plant, a novel application not previously explored in the literature. To achieve this, the study evaluates the sensitivity of LSPIV-derived velocity fields to two user-defined processing parameters—image sampling frequency and IA size. While these parameters have been extensively studied in riverine contexts, their impact on the complex and dynamic flow conditions of coastal power plant discharges remains largely unexplored, providing an opportunity to advance their application in challenging environments. Lastly, the study examines the potential of using LSPIV-derived velocities for estimating effluent discharge rates and integrating

them into a three-dimensional hydrodynamic model for thermal effluent simulations. This work serves as a proof of concept to extend LSPIV applications beyond riverine settings, laying the groundwork for its use in dynamic coastal environments and offering insights to inform future applications.

4.2 Materials and Methods

4.2.1 Observational Data

Multi-source observational data were collected to facilitate LSPIV analysis and validate its viability for surface velocity estimation. Video footage of the discharge location (recall Figure 3.1), captured at nadir to avoid the need for orthorectification, was recorded using a DJI Mavic Pro 2 at 60 frames per second (fps). The UAV was flown at an altitude of approximately 110 m to ensure that both the ADCP validation sites and the discharge outfall were within the frame (Figure 4.1). Full specifications of the video clip are provided in Table 4.1.

The video was stabilised using VISION, an open-source MATLAB software for video stabilisation that offers seven automatic feature detection algorithms capable of distinguishing corners, blobs, and regions with uniform intensity (Pizarro et al., 2022). For this study, the KAZE algorithm, recommended for natural environments, was applied to a defined region of interest (ROI) on the structure surrounding the discharge area, excluding the water surface to ensure robust feature detection. Aside from specifying the ROI and stabilisation algorithm type, all other settings were kept at default. For further details on the stabilisation algorithm, readers are referred to Pizarro et al. (2022).

Table 4.1: Details of UAV footage for LSPIV analysis.

Category	Value
Acquisition Time	13-06-2022, 09:00 AM
Approximate Altitude	110 m
Duration	3 minutes
Drone	DJI Mavic 2 Pro
Resolution	1920 × 1080 pixels
Frame Rate (fps)	60

For validation, velocity measurements were obtained using a RiverPro 1200 kHz ADCP (Teledyne RD Instruments), designed for accurate velocity profiling and bottom tracking in shallow waters. The instrument’s auto-adaptive sampling capability allowed for efficient data collection without extensive user configuration (Teledyne Marine, 2022). To ensure precise positioning, the ADCP was integrated with an A222 GNSS Smart Antenna (Hemisphere GNSS). Manufacturer-recommended calibration and measurement procedures were followed prior to each survey, as outlined by Teledyne Marine (2021, 2022).

For Datasets A, B and C (Figure 4.1a–c), the ADCP was mounted on an ARC-Boat, a 2 m remotely controlled Autonomous Surface Vehicle (ASV) developed by HR Wallingford. The ASV’s shallow draft and high manoeuvrability made it ideal for collecting hydro-metric and bathymetric data in shallow waters (Watchorn et al., 2021). For Dataset D (Figure 4.1d), the ADCP was deployed on a tethered RDI Riverboat, specifically designed for high-velocity and turbulent conditions. The Riverboat ensured stable deployment by maintaining the ADCP at a consistent depth with minimal flow disturbance (Teledyne Marine, 2022). In our setup, the boats positioned the ADCP approximately 9 cm below the water surface, with the blanking distance—the region directly beneath the transducer where velocity cannot be measured—typically minimal but occasionally reaching up to 17 cm in worst-case conditions.

All velocity data were collected in real time using Teledyne’s WinRiver II software (Teledyne Marine, 2021) and subsequently processed with QRev for quality assurance of the measurements (Mueller, 2016).

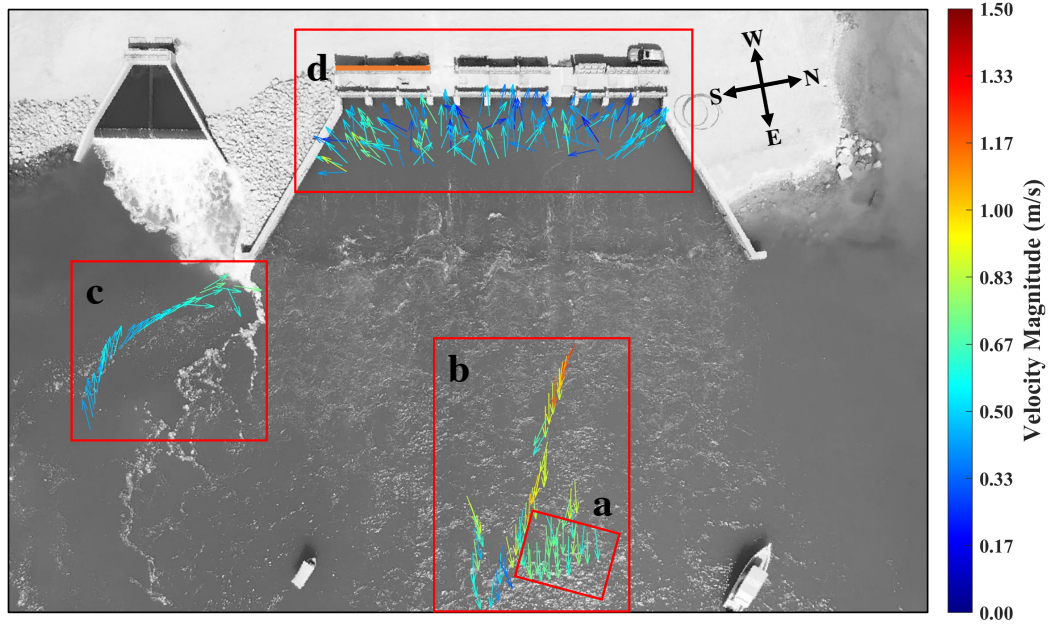


Figure 4.1: UAV aerial image of the outfall showing velocity vectors derived from ADCP measurements, averaged over a 1.5×1.5 m grid for improved readability. Datasets A, B, C, and D are highlighted within red boxes and indicated by their respective lowercase labels. Vector lengths are scaled to 4 m, and colours represent velocity magnitude (m/s). The orange line next to the letter d indicates a calibration reference length (13.260 m) based on the structure size, used to determine the pixel-to-meter ratio. The compass provides orientation for the scene.

In total, four ADCP datasets (Table 4.2) were collected to support LSPIV validation. Although the video for LSPIV analysis was recorded 50 minutes before the acquisition of Dataset A, this dataset was chosen for validation due to its longer duration (19 minutes) and semi-stationary collection, enabling spatial and temporal averaging. Datasets B, C, and D, collected using moving transects, provided instantaneous velocities unsuitable for direct validation but were used for qualitative visual comparisons. Notably, Datasets C and D, recorded on different days but during the same tidal state (high tide), offered additional contextual insights.

Table 4.2: Summary of ADCP datasets for LSPIV analysis.

Dataset	Acquisition Time	Duration (mins)	Collection Method	Relevance to LSPIV Analysis
A	13-06-2022, 09:49–10:08 AM	19	Semi-stationary (< 10 m movement)	Used for validation of LSPIV results.
B	13-06-2022, 09:46–09:49 AM	3	Moving transect	Used for visual comparison to LSPIV results, but not included in quantitative analysis of performance.
C	25-06-2022, 09:17–09:18 AM	1		
D	25-06-2023, 10:10–10:27 AM	17		

In addition to ADCP data, supplementary field data were required for LSPIV analysis and discharge calculations. The structure highlighted by the orange line (Figure 4.1) was measured at a length of 13.260 m and used as a calibration reference to convert pixel measurements into metres in the analysis results. Bathymetric data near the outfall, collected from scatter points and transects during previous surveys, were interpolated onto a 400×400 point grid (Figure 4.2) using MATLAB's "v4" `griddata` method, a 2-D interpolation technique suitable for creating a smooth and continuous representation of the bathymetric surface. Depth values in Figure 4.2 are referenced to high tide, as the data were collected during high tide due to the area's inaccessibility at lower tidal states. Similarly, the LSPIV footage was captured during high tide, ensuring that velocity estimates align with the high-tide bathymetry for accurate discharge estimation.

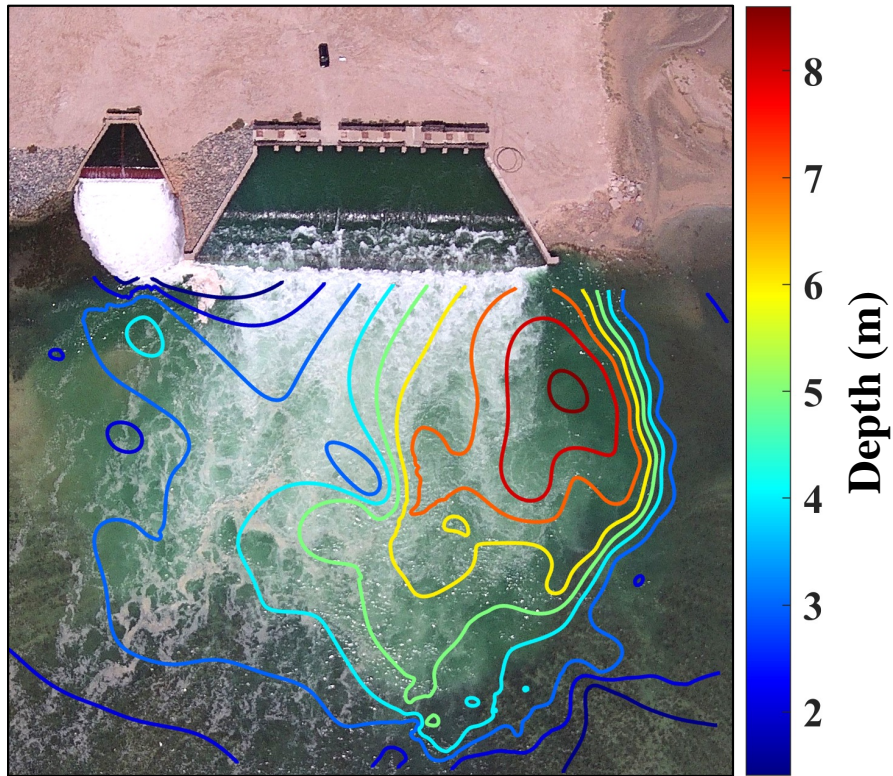


Figure 4.2: Interpolated bathymetry of the outfall region based on field surveys, showing depths relative to high-tide conditions.

4.2.2 PIVlab Analysis

To perform LSPIV analysis, PIVlab, a GUI-based open-source tool for Digital Particle Image Velocimetry (DPIV) analysis in MATLAB, was used. The analysis in PIVlab follows three main steps: image pre-processing, image evaluation, and post-processing. The images were imported in a sequential overlapping style (e.g., 1-2, 2-3) to ensure continuous displacement tracking across frames.

4.2.2.1 Image Pre-Processing

Image pre-processing is an important step in PIV analysis, as variations in image intensity can significantly influence the correlation signal. Non-uniform illumination introduces noise to the correlation plane, where brighter particles disproportionately dominate the correlation peak, reducing the contribution of weaker signals. To address these issues, several enhancement techniques need to be applied to the images before processing (Raffel et al., 2007).

Frames from the UAV video footage were imported into PIVlab, where various pre-processing tools were employed to enhance image quality. Among these:

- **Contrast-Limited Adaptive Histogram Equalisation (CLAHE)** was used to improve local contrast by redistributing the most frequent intensity values within 20-pixel tiles, while preserving details in areas with varying brightness levels (Stamhuis and Thielicke, 2014).
- **Intensity Capping** mitigated biases caused by pixel intensities exceeding a predefined threshold by capping these values to an upper intensity limit (Stamhuis and Thielicke, 2014).
- **Wiener Filter (Wiener2)** an adaptive noise-removal filter, was used to reduce local noise within a default value of 3×3 pixel neighbourhood (MathWorks, 2024).
- **Background Subtraction** was performed by calculating an average intensity image to minimise the influence of stationary features (Raffel et al., 2007).

4.2.2.2 Image Evaluation

The image evaluation stage forms the core of PIV analysis, where the cross-correlation algorithm is used to determine particle displacement between consecutive frames. This statistical pattern-matching technique compares small subregions, known as interrogation areas (IAs), within image pairs. The peak value in the cross-correlation matrix provides the most probable displacement of particle patterns (Raffel et al., 2007; Stamhuis and Thielicke, 2014).

The discrete cross-correlation (DCC) function, $C(m, n)$, calculates the displacement between IAs in images A and B . It is expressed as:

$$C(m, n) = \sum_i \sum_j A(i, j) B(i - m, j - n), \quad (4.1)$$

where:

- $A(i, j)$ and $B(i, j)$ are the intensity values in corresponding IAs from images A and B .
- m and n are the displacements in the x - and y -directions, respectively.
- The summation runs over all pixel pairs within the IAs.

The location of the intensity peak in the correlation matrix $C(m, n)$ indicates the most probable displacement of particles from image A to image B . This displacement is divided by the time interval (Δt) between frames to compute the velocity field. While the DCC method provides reliable results, it can be computationally intensive (Stamhuis and Thielicke, 2014).

To improve computational efficiency, equation 4.1 is often evaluated in the frequency domain using discrete Fourier transform (DFT). According to the correlation theorem, the cross-correlation $C(m, n)$ is equivalent to the inverse Fourier transform of the product of the Fourier transforms of A and B , with the latter being conjugated (Raffel et al., 2007):

$$C(m, n) \iff \mathcal{F}^{-1}\{\hat{A} \cdot \hat{B}^*\},$$

where $\hat{A}$ and $\hat{B}$ are the Fourier transforms of equal-sized IAs A and B , and $*$ denotes the complex conjugate. The frequency-domain approach reduces the computational cost from $O(N^4)$ in DCC to $O(N^2 \log_2 N)$, making it ideal for large datasets (Raffel et al., 2007).

Multi-Pass and Deformation Techniques

In DFT, to address the loss of information due to large particle displacements within identical-sized IAs, a multi-pass approach is employed (Stamhuis and Thielicke, 2014). The first pass provides a coarse displacement estimation at the IA centres. This information is then used to offset IA B in subsequent passes to align more closely with IA A , thereby improving the correlation signal. Typically, IAs overlap by a defined percentage (e.g., 50%), providing additional displacement information at their borders and corners. This information is used to calculate displacement information at every pixel of the IA via bilinear interpolation (Stamhuis and Thielicke, 2014). Additionally, the grid is refined with each pass, enhancing spatial resolution and signal-to-noise ratio.

In real flows, non-uniform particle motion, including shear and rotation, can broaden the intensity peak in the correlation matrix, and in extreme cases, it can split into multiple peaks (Raffel et al., 2007). To address this, deformation is applied to IA B based on displacement data from earlier passes. IA B is stretched, rotated, or warped to align more closely with IA A through bilinear (fast) or spline interpolation (slower, but high precision). At each pass, displacement information is accumulated to produce a refined velocity field (Stamhuis and Thielicke, 2014).

Circular and Linear Cross-correlation

During DFT analysis, PIVlab typically employs circular cross-correlation, which assumes periodicity in the signal. While computationally efficient, this assumption can introduce non-existent frequencies into the spectrum, leading to inaccuracies. To mitigate these issues, linear cross-correlation is used as an alternative. This method approximates non-periodic data by zero-padding the signal, reducing edge discontinuities and improving the accuracy of displacement estimates. However, this enhancement comes with increased computational costs due to the additional data processing required (Thielicke and Sonntag, 2021).

To further enhance robustness, PIVlab offers repeated correlation, a technique that shifts interrogation windows in multiple directions (e.g., left-up, right-up, etc.) by 25% of their size, performing five cross-correlation iterations. The resulting correlation matrices are then multiplied, suppressing noise and yielding a sharper, more distinct peak (Thielicke and Sonntag, 2021).

PIVlab combines these techniques into its ‘Correlation Robustness’ feature, offering three levels of robustness:

- **Standard:** Linear window deformation with circular cross-correlation.
- **High:** Spline window deformation with linear cross-correlation.
- **Extreme:** Spline window deformation with linear cross-correlation and repeated correlation.

In this study, the ‘Extreme’ setting was chosen to maximise the accuracy of displacement estimates, despite the additional computational cost.

Peak Finding

The accuracy of PIV largely depends on the choice of the peak finding technique, which determines particle displacement with sub-pixel precision. The standard method involves fitting a Gaussian function to the intensity distribution in the correlation matrix using adjacent vertical and horizontal pixels (3-point fit for each axis). For cases with shear, rotation, or motion blur, a two-dimensional Gaussian function (9-point fit) provides better performance by accounting for elliptical peak shapes (Stamhuis and Thielicke, 2014). Both methods are implemented in PIVlab, and in this study, the 2D (9-point) Gaussian peak fitting method was selected.

Ensemble Correlation

Ensemble correlation is a technique that averages correlation planes from a sequence of images before applying a peak-searching algorithm. By accumulating displacement information across multiple frames, this approach enhances the signal-to-noise ratio and yields a well-defined correlation peak that represents the mean displacement of the flow (Raffel et al., 2007; Thielicke and Sonntag, 2021).

In this study, ensemble correlation was applied to the three-minute UAV footage to generate a representative velocity field for discharge estimation. By smoothing out short-term variations, this approach is particularly suited for capturing the overall flow characteristics necessary for accurate discharge calculations.

4.2.2.3 Post-Processing

Post-processing functions in PIVlab were employed to improve the accuracy and reliability of the velocity estimates derived from the analysis. The following filters were applied in this study:

- A **standard deviation filter** with a threshold of 8 was used to remove outliers by comparing velocity components to the mean and standard deviation of the velocity field.
- A **local median filter** evaluated velocity fluctuations within a 3×3 neighbourhood to eliminate inconsistent vectors.
- A **correlation threshold** of 0.5 retained only vectors with high correlation values.

Missing data were not interpolated, as the focus was on preserving only validated vectors to ensure data integrity.

4.2.3 Discharge Estimation

Discharge estimation can be performed using the velocity-area method, which integrates the product of velocity and cross-sectional area along a transect. For LSPIV-derived surface velocity fields, this involves combining the calculated surface velocity field with bathymetric data and adopting simplified assumptions about vertical velocity profiles (Kinzel and Legleiter, 2019; Pumo et al., 2021).

Surface velocities obtained from PIVlab are converted into depth-averaged velocities using a scaling factor, also referred to as the velocity coefficient or calibration factor, denoted as α (Biggs et al., 2023). The coefficient is defined as:

$$\alpha = \frac{U}{u_s},$$

where U is the depth-averaged velocity, and u_s is the surface velocity.

Site-specific α can be estimated from measured velocity profiles using the extrapolation tool in QRev software (Biggs et al., 2023). In this study, ADCP velocity profiles from Dataset A were analysed to estimate the value of α . QRev displayed the profiles, as shown in Figure 4.3, with normalised elevation (z/H) on the y -axis and normalised velocities (u_s/U) on the x -axis. Using its automated method, which employs linear least squares regression and empirical criteria (Mueller, 2013, 2016), the software selected a power-law fit with an exponent of 0.4146. The normalised velocity where the extrapolation intersects the surface, indicated by the red cross in Figure 4.3, was used to calculate α based on the method outlined in (Biggs et al., 2023):

$$\alpha = \frac{1}{\text{Normalised velocity at surface}} = \frac{U}{u_s},$$

The value of α typically falls within the range of 0.7 to 0.9, and is commonly assigned a standard value of 0.85, derived from velocity profiles that follow a 1/6th power-law (Biggs et al., 2021, 2023). Based on Figure 4.3, α was estimated to be 0.82.

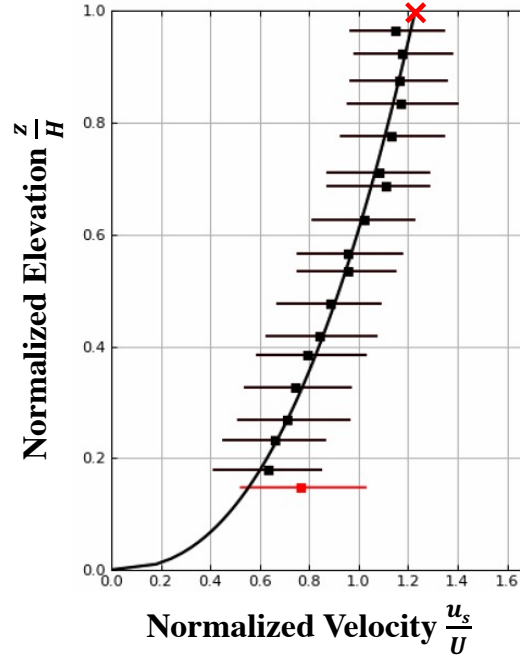


Figure 4.3: Estimation of α from ADCP Dataset A using QRev with power-law extrapolation. The red cross ($\times$) indicates the normalised velocity at the surface, $u_s/U = 1.22$, with α calculated as the reciprocal of the normalised surface velocity, yielding a value of $\alpha = 0.82$.

Depth-averaged velocities are combined with the bathymetric cross-section to compute discharge using the velocity-area method. The general equation for discharge is:

$$Q = \sum_{n=1}^N U_n \cdot \Delta A_n, \quad (4.2)$$

where U_n represents the depth-averaged velocity in a specific segment or “bin” of the transect, and ΔA_n is the corresponding cross-sectional area of that bin (Biggs et al., 2023).

Discharge estimation using this method is subject to several sources of uncertainty, with the conversion of surface velocity to depth-averaged velocity being a primary contributor. The estimation of α , often assumed constant across the transect, exhibits natural variability due to site geometry, flow conditions, and wind influence (Biggs et al., 2023). Additionally, extrapolating ADCP velocity profiles to the water surface introduces further uncertainty, as discrepancies may arise between the extrapolated velocities and actual surface velocities (Biggs et al., 2023).

Uncertainties in bathymetric data, including errors in depth measurements and interpolation, also significantly affect discharge calculations, as errors in the cross-sectional area propagate into flow rate estimates (Torres et al., 2024). Despite these challenges, and given the potential complete lack of any discharge data at some application sites, the methodology presented here offers a practical and effective framework for estimating discharge from velocimetry-derived data.

4.2.4 Data Processing For LSPIV Validation

One of the challenges in this study was comparing LSPIV and ADCP data due to the difference in temporal coverage between the two datasets. The LSPIV data are derived from a 3-minute UAV video, whereas Dataset A consists of a 19-minute ADCP recording. To evaluate the suitability of comparing these datasets, the velocity magnitudes and directions of the ADCP data were first visualised as histograms (Figures 4.4a and b). These histograms illustrate the variability contained within the 19-minute ADCP dataset, with the averages for velocity magnitude (0.71 m/s) and direction (101.3°) marked by black dashed lines.

To account for the differing temporal scales, the 19-minute ADCP dataset was randomly subsampled into continuous 3-minute intervals to match the duration of the UAV footage. This subsampling process was repeated 1,000 times to determine whether the averages of these shorter intervals significantly deviated from the 19-minute dataset. The results of this analysis are presented as box plots (Figures 4.4c and d), which summarise the distribution of subsampled averages.

The box plots demonstrate that the 3-minute subsampled averages align closely with the overall 19-minute average, showing minimal deviation. The narrow spread of the box plot edges reflects the limited variability in the subsampled data and supports the reliability of using the 19-minute ADCP dataset for comparison against the 3-minute LSPIV data.

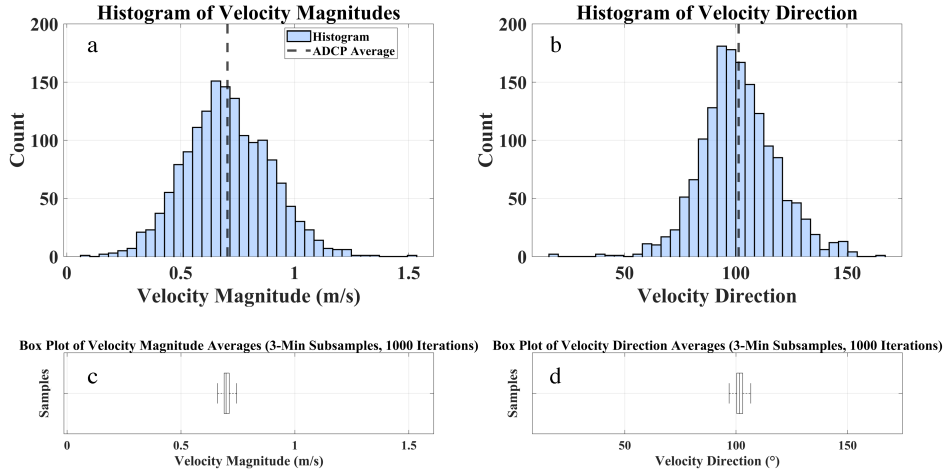


Figure 4.4: (a) Histogram of velocity magnitudes and (b) histogram of directions for Dataset A, with dashed vertical lines indicating the overall dataset averages (0.71 m/s for magnitude and 101.3° for direction). (c) Box plots of velocity magnitude and (d) direction averages calculated from 1,000 subsamples of continuous 3-minute intervals from the 19-minute dataset.

Another challenge in comparing LSPIV results with ADCP measurements stems from the ADCP’s inability to directly measure surface velocities, as discussed in Section 4.2.1. This limitation has been addressed in the literature through various methodologies. For instance, Pearce et al. (2020) applied characteristic velocity profile fitting to extrapolate surface velocities from subsurface ADCP data. Similarly, Kinzel and Legleiter (2019) used spatial averaging and grid matching techniques to reconcile differences between measured and velocimetry-derived datasets, facilitating performance comparisons.

In this study, we addressed this limitation by averaging (thinning) the velocities from Dataset A onto a 2×2 m grid, matching the spatial resolution of the smallest IA configuration (32 px^2 , equivalent to 2.55 m^2). LSPIV-derived velocity fields were then interpolated to the locations of the thinned ADCP grid using natural neighbour interpolation. This approach ensured spatial overlap between the datasets, enabling direct comparisons through differencing and improving the evaluation of LSPIV accuracy.

4.2.5 Workflow for LSPIV Investigation and Model Integration

To achieve the objectives of this chapter, the workflow outlined in Figure 4.5 was implemented. This approach was designed to validate LSPIV-derived velocities, assess their

suitability for discharge estimation, and integrate the results into a hydrodynamic model to simulate thermal effluents from a coastal power plant.

The process began with extracting frames from stabilised UAV footage, which were subsequently processed in PIVlab using the pre-processing steps outlined in Section 4.2.2.1. These pre-processed images were analysed using the ensemble multi-pass FFT technique with window deformation, as described in Section 4.2.2.2.

To investigate the influence of image sampling frequency and IA size on the analysis results, a series of experiments were conducted. The first set of experiments focused on sampling frequency. The original 60 fps UAV footage was systematically subsampled to 30, 15, and 8 fps. Each sampling configuration was processed with an IA setup starting at 128 px^2 (10.20 m^2) in the first pass and refining to 64 px^2 (5.10 m^2) in the second pass, with a 50% overlap in IAs maintained throughout. These tests aimed to evaluate the extent to which temporal resolution could be reduced while still accurately capturing tracer motion, particularly in the fast-paced and turbulent flow near the outfall.

Further analysis focused on the 8 fps footage, as it represented the lowest sampling rate and was suspected to be insufficient for capturing the turbulent discharge flow. This presented an opportunity to investigate whether adjustments to IA size and multi-pass configurations could mitigate the limitations of reduced temporal resolution. Three IA configurations were evaluated to explore how IA size and refinement influence velocity estimates:

1. **Standard Two-Pass Configuration:** The analysis began with an IA size of 128 px^2 (10.20 m^2) in the first pass, refined to 64 px^2 (5.10 m^2) in the second pass. This configuration, consistent with the setup used in the sampling frequency experiment, provided a baseline for comparing the effects of IA adjustments on velocity estimates.
2. **Reduced IA Two-Pass Configuration:** A smaller starting IA size of 64 px^2 was refined to 32 px^2 (2.55 m^2) in the second pass. This configuration introduced a denser grid of velocity measurements, enabling a closer examination of flow structures at finer scales.
3. **Three-Pass Configuration with Progressive Refinement:** This configuration introduced an additional refinement pass, starting with 128 px^2 , followed by 64 and

then 32 px². This setup aimed to evaluate the impact of additional refinement steps on velocity estimation accuracy.

Additionally, the comparison between the second and third configurations, both of which concluded with a final IA size of 32 px², allowed for an evaluation of how the initial IA size and the number of refinement passes impacted the ability to capture flow features accurately, particularly in regions with varying tracer densities. To ensure consistent and reliable comparisons, a 50% overlap was maintained across all configurations.

In all experiments, the PIVlab filtering procedures outlined in Section 4.2.2.3 were applied to remove invalid vectors. The post-processed velocity fields were subsequently compared against ADCP-measured velocities, following the methodology detailed in Section 4.2.4, to evaluate their accuracy and overall robustness.

Subsequently, discharge was calculated along a transect near the outfall using the LSPIV-derived velocity fields, following the methodology detailed in Section 4.2.3. These discharge estimates were integrated into the SB model (Section 3.2.2) to simulate the thermal plume generated by the DW power plant. Model performance was assessed by comparing simulated temperature profiles to in situ measurements (Dataset 2, Section 3.2.1) using the RMSE metric (Equation 3.4).

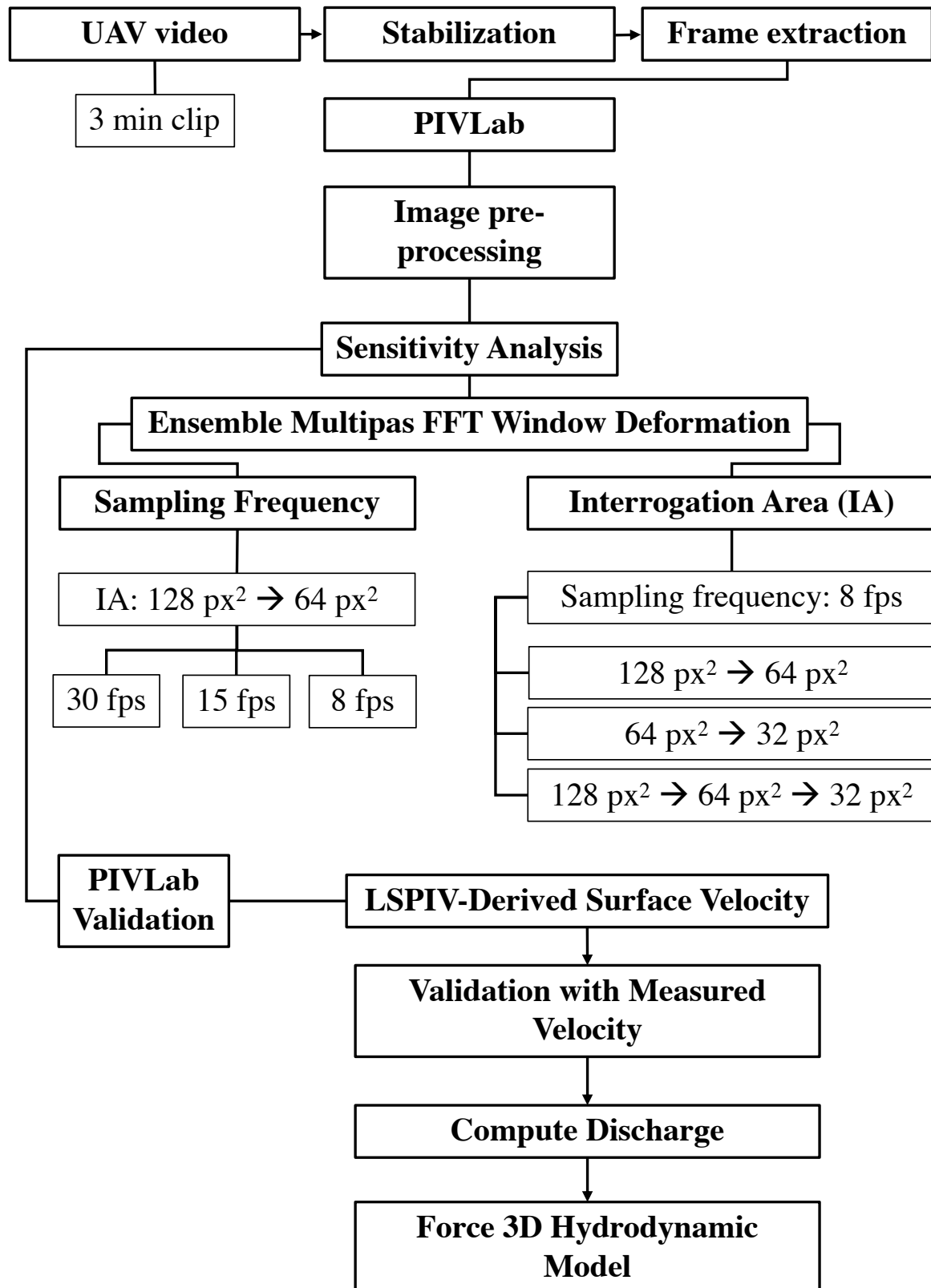


Figure 4.5: Workflow for LSPIV analysis and integration with the hydrodynamic model.

4.3 Results and Discussion

4.3.1 Impact of Image Sampling Frequency and Interrogation Area on LSPIV Results

In this study, naturally occurring surface textures—such as ripples, bubbles, and foam generated by the turbulent discharge from the outfall—served as tracers for cross-correlation analysis. While these features provided valuable information on the flow, their visibility and reliability were influenced by external factors. Wind-induced ripples often distorted or obscured the discharge-generated patterns, diminishing their clarity. Similarly, sunlight, while enhancing tracer visibility by illuminating surface textures, introduced complications such as sun glare, which created stationary high-intensity regions that biased the cross-correlation. This section investigates how image sampling frequency and IA size interact with the specific tracer characteristics of the study environment to influence LSPIV-derived velocity fields.

Initial visual inspection of the velocity fields derived from the LSPIV experiments with varying sampling frequencies and IA configurations in Figure 4.5a–e, compared to the measured velocity fields in panel (f), shows that LSPIV effectively captures the expected flow trajectories in and around the outfall, with some slight variation among the tested configurations. The downstream flow patterns from Datasets A and B are well-represented in the LSPIV results. Additionally, the clockwise rotational patterns visible on the left side of the outfall region align with Dataset C, while the radially divergent surface flow caused by the submerged jets matches the flow patterns in Dataset D. Notably, despite Datasets C and D not being temporally aligned with the UAV footage, the general agreement in flow patterns demonstrates that the algorithm was able to capture the primary flow features of the study area.

In the following sections, we will analyse the results of each sensitivity configuration in detail, drawing on comparisons with ADCP data (Figure 4.7) and key performance metrics from each experiment (Table 4.3) to provide a foundation for understanding the accuracy and implications of each configuration.

Table 4.3: Summary of LSPIV processing parameters and their impact on pixel displacement, valid data points and computation time.

Sampling Frequency (fps)	Interrogation Area (px ²)	Pixel Displacement per Frame (m/s)	Valid Data Points (%)	Computation Time (hr)
30	128 – 64	2.392	100	22.71
15	128 – 64	1.1958	97.2	12.31
8	128 – 64	0.6378	95.9	6.58
8	64 – 32	0.6378	97.7	5.73
8	128 – 64 – 32	0.6378	97.5	6.75

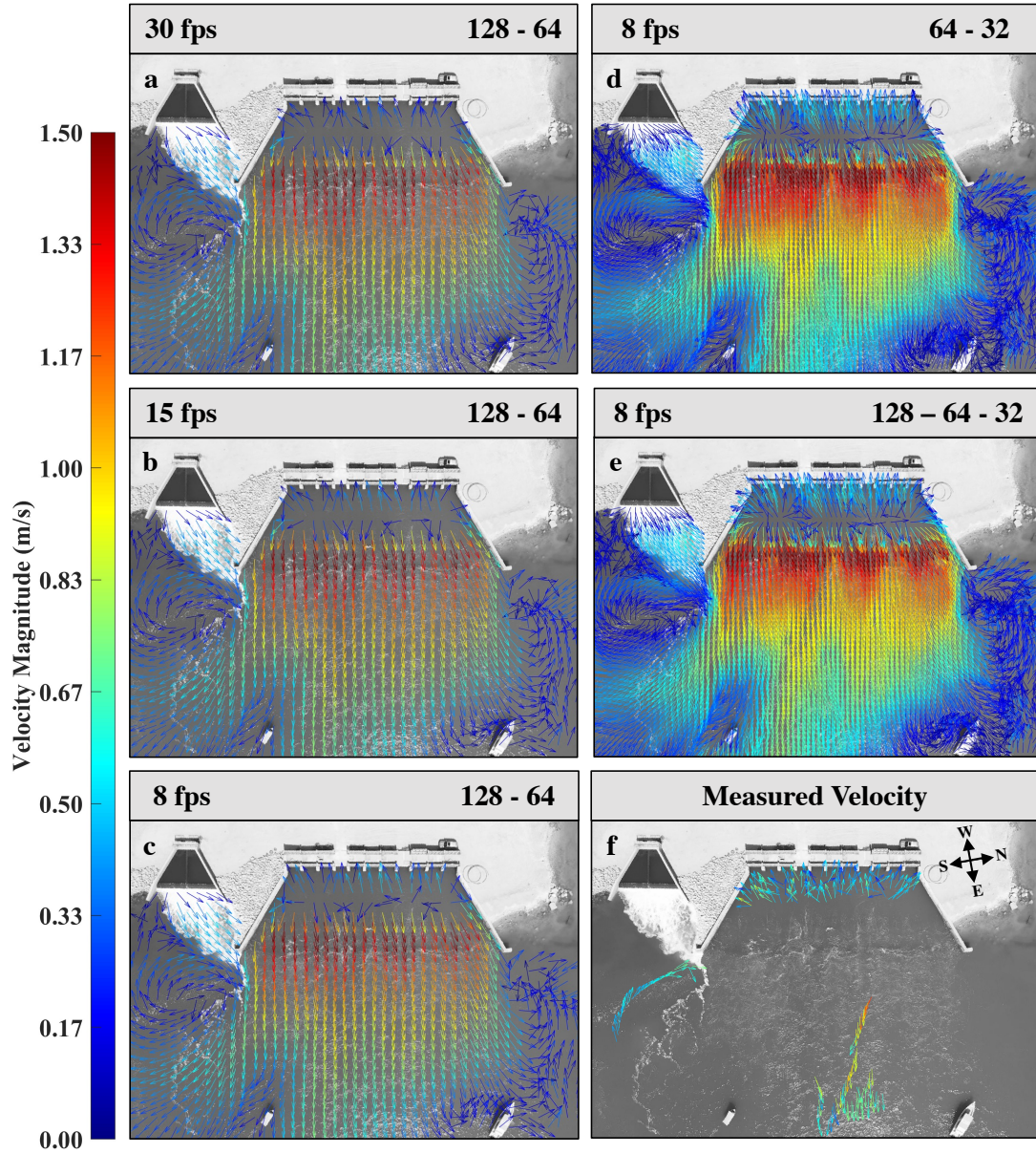


Figure 4.6: LSPIV results showing average velocity fields over a 3-minute clip with varying sampling frequencies and interrogation areas (IA). Panels (a–c) show results at 30 fps, 15 fps, and 8 fps, respectively, using an IA of 128–64 px². Panels (d–e) display LSPIV results at 8 fps with progressively refined IAs of 64–32 px² and 128–64–32 px². Vectors represent a distance of 5 m, and colours indicate velocity magnitude (m/s). Panel (f) shows the measured velocity field derived from the ADCP datasets, spatially averaged to a 1.5 × 1.5 m grid, with vector lengths scaled to 4 m for improved visibility.

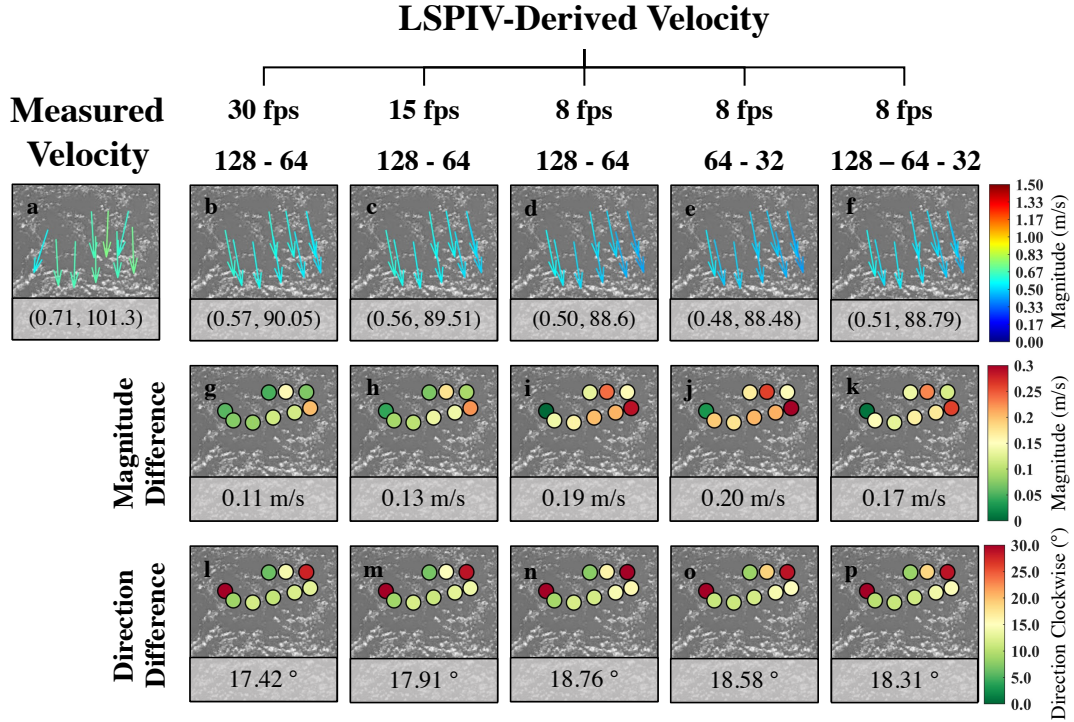


Figure 4.7: Comparison of measured ADCP velocity and LSPIV-derived velocity under varying processing parameters. Panel (a) shows the measured ADCP velocity from Dataset A. Panels (b–d) present LSPIV-derived velocity fields at 30 fps, 15 fps, and 8 fps, respectively, with an interrogation area (IA) of 128–64 px². Panels (e–f) display LSPIV results at 8 fps with IAs of 64–32 px² and 128–64–32 px². All vectors represent a 4 m distance, with colours indicating velocity magnitude. The text annotations below panels (a–f) show average velocity magnitude and direction. Panels (g–k) and (l–p) display the differences in velocity magnitude and direction, respectively, between ADCP and LSPIV, with RMSE values shown below each panel.

4.3.1.1 Impact of Image Sampling Frequency

The influence of image sampling frequency on LSPIV performance is evaluated in Figure 4.7. The average velocities derived from different subsampling configurations (Figures 4.7b–d) ranged from 0.57 m/s at 30 fps to 0.50 m/s at 8 fps, with corresponding average directions spanning 89.51° to 88.48°. While these differences are subtle—less than 0.1 m/s and 1.5°—they reveal how the cross-correlation algorithm estimates displacement as the time interval between image pairs increases. RMSE values for velocity magnitude and direction (Figures 4.7g–i and l–n) show that higher frame rates improve alignment with measured data, with 30 fps achieving the lowest RMSE. However, these results represent a relatively small portion of the overall domain.

Figure 4.8 provides additional context by showcasing snapshots of image pairs for each sampling frequency after applying preprocessing steps outlined in Section 4.2.2.1. These snapshots, though representing isolated instances in time, are illustrative of the challenges associated with tracer characteristics in this environment. Surface textures generated by the discharge rapidly decay, deform, or merge as they travel downstream to the location of the validation dataset. At 30 fps, these textures are mostly well-preserved in the image pairs, providing sufficient detail for cross-correlation. At 15 fps, partial decay of tracers is evident, while at 8 fps, notable degradation occurs, limiting the algorithm’s ability to reliably match tracer pairs between frames.

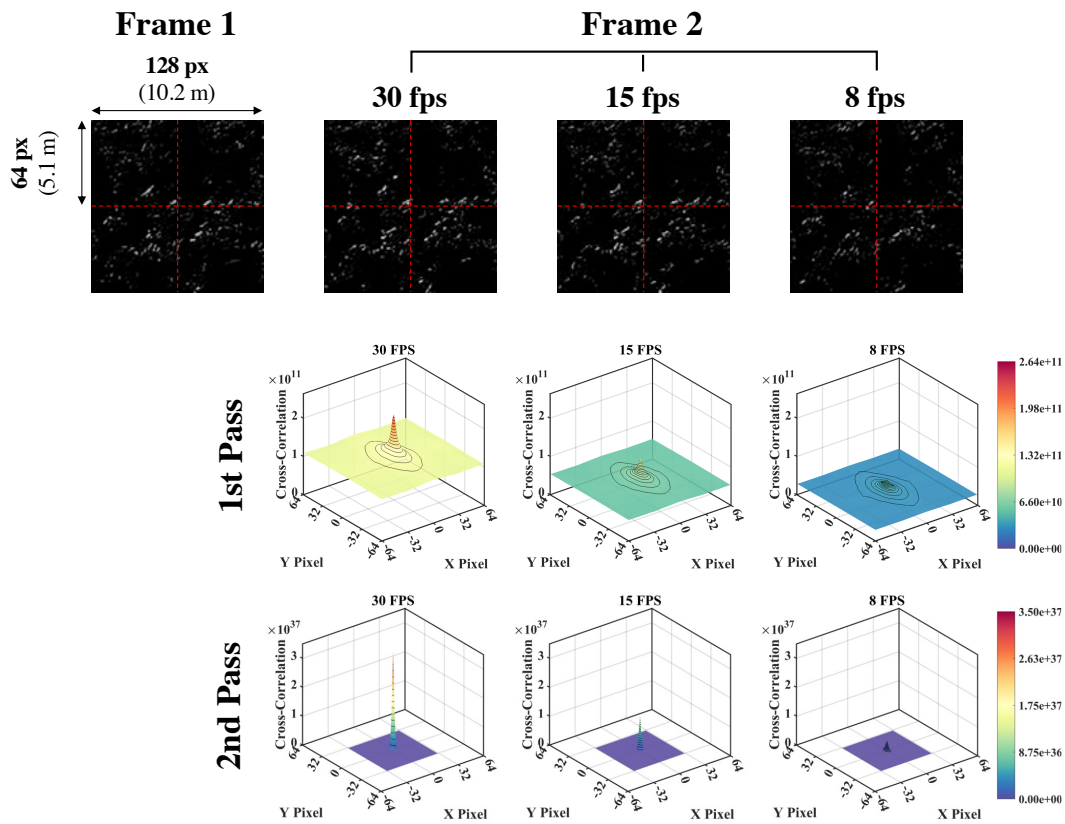


Figure 4.8: Optical displacement and cross-correlation matrices for LSPIV sampling frequencies (30 fps, 15 fps, and 8 fps). The first row shows snapshot frames after preprocessing, with grid overlays indicating pixel dimensions (px) and real-world equivalents (m). The second row displays first-pass cross-correlation matrices (128–64 px²), while the third row shows refined second-pass matrices with smaller interrogation windows.

The amplitude of the cross-correlation peak signal, illustrated in the second row of Figure 4.8 for the first pass and the third row for the second pass, reflects the impact of

sampling frequency. Peak intensity diminishes with decreasing frame rates, with 30 fps yielding the strongest and most distinct peaks, followed by 15 fps and 8 fps. This trend is consistent across both passes, with the second pass producing sharper peaks due to refinement from the first pass, where initial displacement estimates narrow the search region and improve particle matching accuracy.

The stronger peak intensity observed at 30 fps also aligns with the higher percentage of valid data points retained after PIVlab filtering. Table 4.3 shows a retention rate of 100% for 30 fps, compared to 97.2% for 15 fps and 95.9% for 8 fps. The reduced retention at lower frame rates highlights the challenges posed by degraded tracer visibility and weakened correlation signals, further emphasising the benefit of higher sampling frequencies for maintaining accuracy in this specific part of the domain.

The narrow width of the signal highlights another critical factor: tracer visibility is influenced by the ground sampling distance (GSD), which represents the physical area captured by a single pixel in the image. When the GSD is too large relative to the tracers' size, surface features become poorly resolved, limiting their effectiveness for image velocimetry (Jolley et al., 2021). In this study, the GSD was calculated in PIVlab using the structure length highlighted in orange in Figure 4.1, yielding a resolution of approximately 0.0797 m per pixel. This resolution, combined with the time between consecutive frames—0.0333 s for 30 fps, 0.0667 s for 15 fps, and 0.125 s for 8 fps—directly affects tracer displacement, resulting in pixel displacements of 2.4 m/s, 1.19 m/s, and 0.64 m/s per frame, respectively (Table 4.3). These displacement values, particularly at higher sampling frequency configurations, are relatively high when interpreted at the integer-pixel level in the cross-correlation grid, compared to the velocity range observed in the field (Figure 4.6f), meaning that significant information is condensed into single integer-pixel movements. This places greater emphasis on selecting an effective sub-pixel peak-finding technique, as many observed displacements are likely to lie between integer-pixel values.

To mitigate these challenges while maintaining a high image sampling frequency, reducing the altitude of image capture is a viable solution. For example, at 30 fps, capturing images at an altitude of approximately 110 m results in a GSD of about 8 cm/pixel and a pixel displacement of 2.4 m/s per frame. Lowering the altitude to 55 m decreases the GSD to 4 cm/pixel and the pixel displacement to 1.2 m/s, while an altitude of 27.5 m further

reduces the GSD to 2 cm/pixel and the displacement to 0.6 m/s per frame. As such, it is recommended to estimate the optimal altitude prior to data collection to ensure particle movements are adequately resolved at the integer-pixel level, aligning with the expected flow velocities in the field.

While our analysis demonstrated that 30 fps provided the best performance, and 8 fps exhibited the lowest accuracy in the downstream region, this conclusion cannot be generalised across the entire domain. The optimal sampling frequency must align with the characteristics of the tracers and the flow dynamics in the specific region of interest. Different areas within the domain may exhibit distinct tracer behaviours that necessitate tailored parameter settings.

For example, examining Dataset D in Figure 4.6f, which corresponds to the radially divergent surface currents generated by the submerged jets, reveals a different pattern. The jets intermittently produce vertical streams of bubbles that appear on the surface. At higher frame rates, such as 30 fps (Figure 4.6a), the shorter time interval between frames limits the observable displacement of these tracers, biasing the velocity estimates toward slower values. In contrast, 8 fps (Figure 4.6c) accommodates the intermittent nature of these tracers by allowing more time for displacement between frames. This results in velocity estimates that more accurately reflect the tracer dynamics, with performance further enhanced when finer IA configurations are used, as shown in Figures 4.6d and 4.6e. Additionally, the pixel displacement per frame at 8 fps (0.64 m/s) aligns more closely with the observed velocity range in Dataset D compared to the 2.4 m/s displacement at 30 fps.

Figure 4.9 provides additional insight into the impact of sampling frequency on velocity estimates by visualising the magnitude differences between configurations. The left panel shows an optical frame with enhanced contrast, improving tracer visibility, while the right panels depict the magnitude differences between velocity fields derived from 30 fps and 15 fps, and 30 fps and 8 fps, respectively.

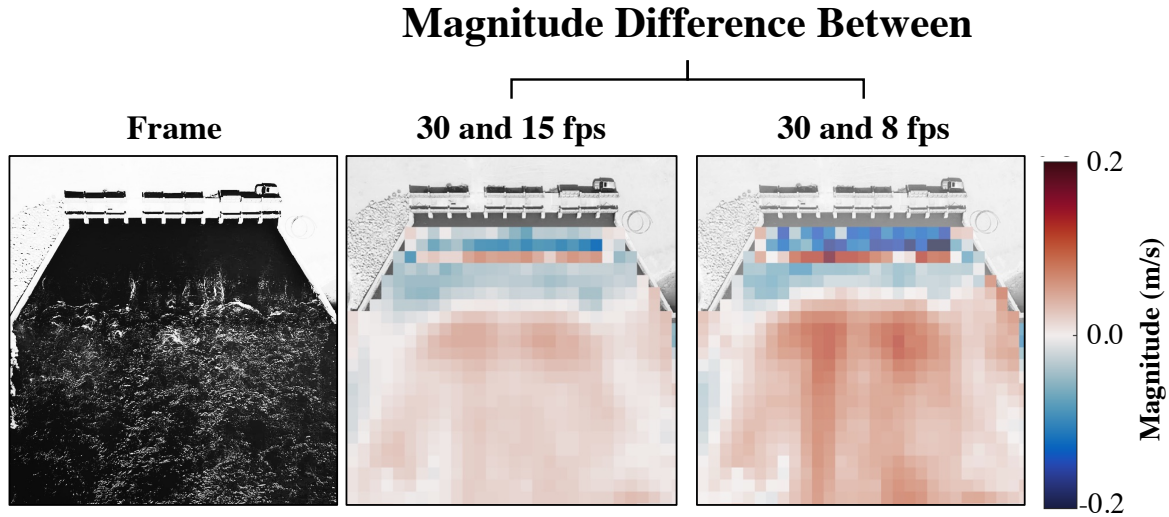


Figure 4.9: Magnitude differences in velocity fields between various sampling frequencies and interrogation area configurations during LSPIV analysis. The first column shows frames with increased contrast for better visibility of tracers. The second and third columns illustrate the magnitude difference in velocity fields between 30 fps and 15 fps, and 30 fps and 8 fps, respectively.

The difference maps reveal distinct patterns. Downstream of the outfall, 30 fps tends to produce higher velocity estimates compared to both 15 fps and 8 fps, as indicated by the red regions. This aligns with the earlier observation that 30 fps ensures better tracer preservation and displacement resolution, particularly in areas with fast-moving turbulent flow. Conversely, in regions closer to the submerged jets, characterised by radially divergent flow, 30 fps underestimates velocities relative to 8 fps, as shown by the blue areas.

Lastly, while aligning the frame rate to local flow dynamics is essential for accurate velocity estimation, an important consideration when selecting the appropriate sampling frequency is the computational time associated with higher frame rates. As shown in Table 4.3, the computational time for the 3-minute footage increases nearly 3.5-fold when moving from 8 fps to 30 fps, rising from 6.58 hours to 22.71 hours. This trade-off between accuracy and computational cost becomes critical when determining an acceptable error threshold for the intended application of the LSPIV data. If the differences in results are subtle, the significantly increased processing time at higher frame rates may not be justified for applications where such high precision is not strictly required.

4.3.1.2 Impact of Interrogation Area

To evaluate the influence of interrogation area (IA) size, we turn to Figure 4.10. The left panel shows the optical frame processed to enhance contrast and reveal surface textures acting as tracers, while the right panel illustrates the velocity differences between the 128–64–32 px² and 64–32 px² configurations for the 8 fps case. Although both configurations achieve the same final resolution, they differ in their initial IA sizes and number of passes, leading to distinct velocity estimates:

- **Red areas (up to +0.2 m/s):** Predominantly downstream of the discharge, indicating that the 128–64–32 configuration estimated higher velocities.
- **Blue areas (up to –0.2 m/s):** Concentrated upstream, directly underneath the submerged jets, where the 64–32 configuration estimated higher velocities.

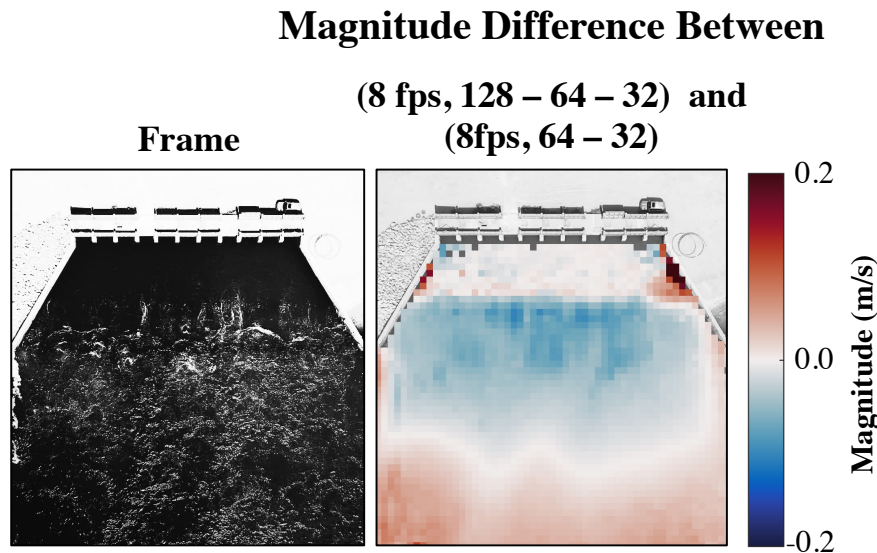


Figure 4.10: Velocity magnitude differences between two IA configurations, 128–64–32 px² and 64–32 px², for 8 fps. The left panel shows the optical frame with enhanced contrast to highlight surface textures. The right panel displays differences, with red areas (+0.2 m/s) where 128–64–32 estimated higher velocities, downstream with sparse tracers, and blue areas (–0.2 m/s) where 64–32 estimated higher velocities, upstream with dense tracers. The boundary reflects the transition in surface texture characteristics.

A clear boundary between these regions aligns with the transition in surface texture characteristics observed in the left panel. Downstream, surface textures are sparse, with visible gaps between tracer-rich areas. Conversely, upstream regions feature dense surface textures generated by turbulence and discharge activity.

The observed differences in velocity estimates arise from the mechanics of the multi-pass cross-correlation process. Larger IAs (e.g., 128 px²) in the initial pass average displacement estimates over a broader area, capturing overall flow direction even in regions with sparse tracers. Subsequent passes progressively narrow the search region, refining these initial estimates and improving displacement accuracy. However, starting with a smaller IA size (e.g., 64 px²) can introduce biases in regions with intermittent tracer gaps, where the algorithm struggles to establish reliable displacements due to insufficient surface textures.

This behaviour is evident in the validation dataset, which represents a small region of the downstream flow (Figure 4.7). The 8 fps 128–64 px² configuration, although underperforming in the sampling frequency analysis, demonstrated improvement when a third pass (128–64–32 px²) was introduced, reducing RMSE in both velocity magnitude and direction (Figures 4.6k and p). In contrast, the 64–32 px² configuration yielded slightly higher RMSE values for velocity magnitude but slightly lower RMSE values for direction (Figures 4.6j and o).

These findings emphasise the importance of tailoring IA sizes to the tracer density within the region of interest. Larger IAs are effective in capturing flow dynamics in areas with sparse or intermittent tracers, while smaller IAs are better suited for regions with dense tracer textures, preserving finer details and more accurate velocity estimates.

Lastly, the difference in computational times between IA configurations at 8 fps (Table 4.3) was less pronounced compared to the sampling frequency sensitivity tests, with variations of approximately one hour across configurations, reflecting the relatively smaller impact of IA adjustments on total processing time.

4.3.2 Integration of LSPIV-Derived Discharge into a Hydrodynamic Model

In this section, we evaluate the integration of LSPIV-derived discharge estimates into a three-dimensional hydrodynamic model to simulate thermal effluents in a dynamic coastal environment. Earlier analyses explored the sensitivity of LSPIV-derived velocity fields to image sampling frequency and IA size, revealing subtle differences in velocity estimates

(± 0.2 m/s) and more pronounced variations in computational time. These findings raise an important question: how do such variations in velocity fields translate into discharge estimates, and to what extent do these differences affect model performance?

To address this, the Sulaibikhat Bay (SB) model (Section 3.2.2) was employed. Model performance was assessed using the RMSE metric (equation 3.4) by comparing simulated temperature profiles to in situ measurements collected from 16 stations across the bay. These measurements, referred to as Dataset No. 2 in the previous chapter (Table 3.1), were acquired within a similar window of time as the UAV footage used in the LSPIV analysis. This alignment provided an opportunity to comprehensively evaluate the model's ability to replicate observed temperature profiles across multiple locations in the domain.

The workflow for computing discharge from LSPIV-derived velocities is summarised in Figure 4.11, which uses the 30 fps case as an example. This process was applied consistently across all configurations. A transect near the outfall, indicated by the black line in Figure 4.11a, was selected based on its proximity to the discharge region and its alignment with available bathymetric data.

Surface velocities along the transect, shown in Figure 4.11c, were extracted from the LSPIV results and scaled to depth-averaged velocities using the scaling factor detailed in Section 4.2.3. The transect was divided into bins, with bin widths defined by the spacing between consecutive velocity points. For each bin, the average velocity and depth were calculated based on values at the bin boundaries, using bathymetry data from Figure 4.2. Incremental discharge values were computed for each bin using equation 4.2 and summed to obtain the total discharge across the transect, as illustrated in Figure 4.11d.

Table 4.4: Summary of LSPIV sampling configurations and their impact on average transect velocity, computed discharge, and RMSE.

Sampling Frequency (fps)	Interrogation Area (px ²)	Average Transect Surface Velocity (m/s)	Discharge (m ³ /s)	RMSE (°C)
30	128-64	1.082	235.23	0.7181
15	128-64	1.059	229.99	0.7144
8	128-64	1.016	220.26	0.7157
8	64-32	1.000	215.17	0.7183
8	128-64-32	0.969	209.00	0.7224

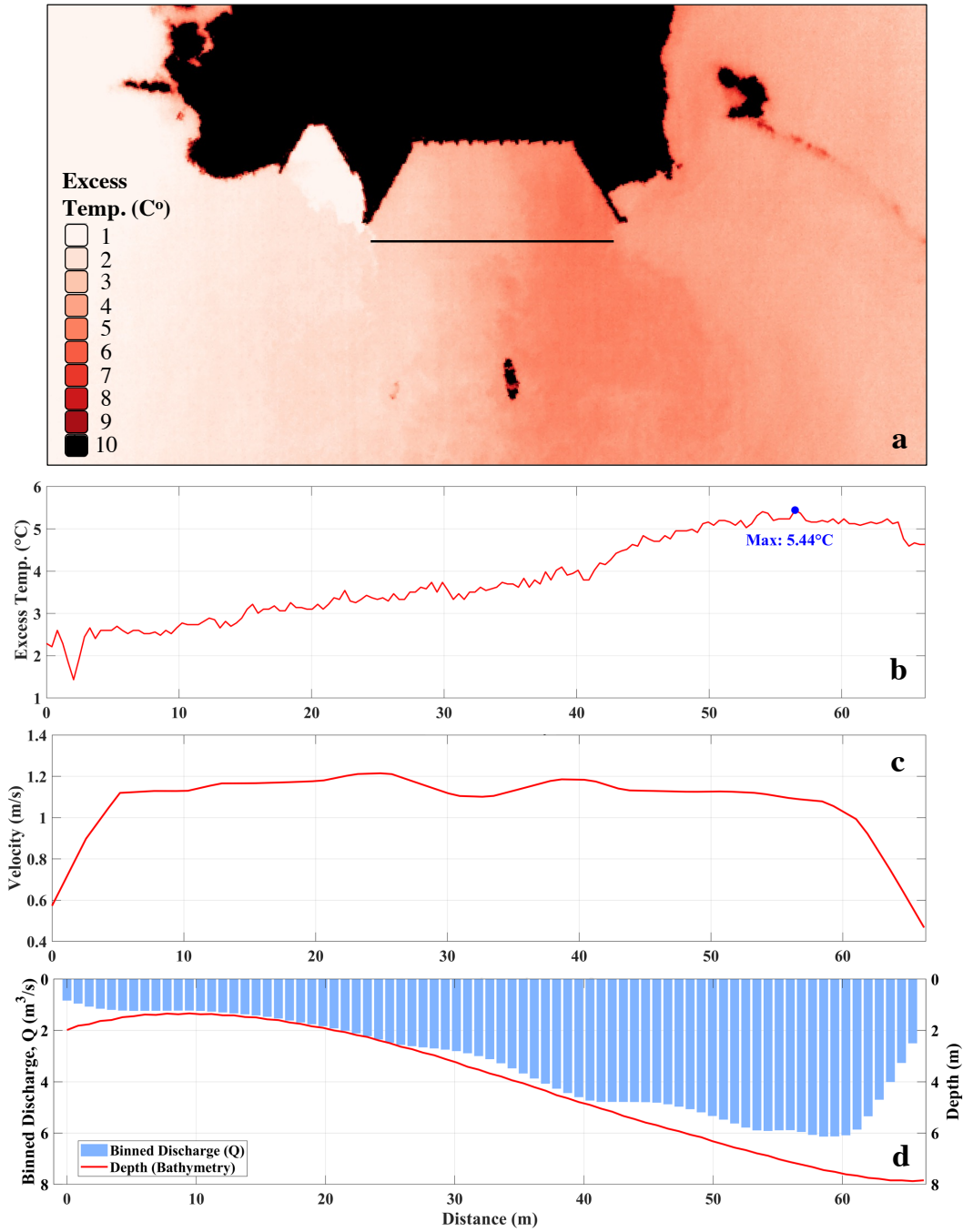


Figure 4.11: Thermal effluent analysis at the outfall location. (a) Thermal image of the discharge area showing excess temperature computed as the difference between the measured temperature in the scene and the boundary temperature. The black line indicates the transect along which measurements were extracted. (b) Excess temperature along the transect, with a maximum value of 5.44 °C. (c) Surface velocity along the transect derived from the 30 fps LSPIV case with a 128–64 interrogation area configuration. (d) Incremental discharge (binned by width) along the transect computed from the 30 fps LSPIV case, shown alongside the bathymetric depth profile. The blue bars represent discharge, and the red line depicts depth.

Table 4.4 presents the key results for each LSPIV configuration. The average surface velocities along the transect ranged from 0.969 m/s to 1.082 m/s, with a total variability of 0.113 m/s across all configurations. This variation translated into discharge estimates ranging from 209 m³/s to 235.23 m³/s, resulting in a total difference of approximately 26 m³/s across the five configurations. Notably, these values are within the range of discharge estimates derived using simplified energy calculations for the DW power plant (169.8 to 238.68 m³/s), as presented in Appendix D. They are also close to the value of 197 m³/s reported by Pokavanich and Alosairi (2014) and the profile-calibrated estimate of 205.95 m³/s for Dataset 2, which was collected simultaneously with the UAV footage. This agreement indicates that the LSPIV-derived values are within a physically reasonable range.

To evaluate the practical implications of this variability, it is essential to assess its impact within the hydrodynamic model, particularly in terms of its influence on the simulated thermal plume and the overall model performance. We explored a scenario where both discharge (Q) and excess temperature (ΔT) were derived entirely from UAV datasets. The thermal image of the discharge area from Dataset No. 2 was transformed into an excess temperature map (Figure 4.11a) by calculating the difference between the temperatures observed in the thermal image and the temperature at the model boundary (recall Figure 3.4) during the time of image acquisition. Excess temperature values were extracted along the transect (black line in Figure 4.11a), with the extracted temperatures shown in Figure 4.11b. The maximum temperature across the transect was selected as the ΔT input for model forcing, as this approach better represents the thermal anomaly introduced by the discharge. This decision accounts for the larger size of the computational grid, which cannot resolve the fine-scale gradients observed in the near field, making the use of an average temperature less appropriate. The timing of the thermal image acquisition aligns closely with the UAV video footage (Table 3.1), ensuring consistency between the datasets used for discharge and temperature estimation.

Using the ΔT value derived from the thermal image and the discharge (Q) values from each LSPIV configuration, the simulation results are presented in Figure 4.12. Stations P1 to P8, which are under the boundary's influence, were omitted from the visualisation to better highlight the regions dominated by effluent dynamics. Across the different

configurations, minimal variation was observed in the simulated temperature profiles, with the most noticeable differences occurring at station P12. However, these differences were small and insignificant relative to changes in the RMSE. The RMSE values, summarised in Table 4.4, were nearly identical across configurations, confirming that the variability in discharge estimates (within the observed range of $< 26 \text{ m}^3/\text{s}$) had a negligible impact on model performance.

These results demonstrate that, for far-field plume simulations with moderate numerical grid resolution, variability in Q (within the observed range of $< 26 \text{ m}^3/\text{s}$) has a negligible impact on the model's temperature predictions. Consequently, the additional computational effort outlined in Table 4.3 to achieve higher precision may not be justified. As such, modellers should evaluate their numerical configurations and study objectives to determine an acceptable level of sensitivity in results that aligns with the intended application.

Nevertheless, given the results, the model performance across all LSPIV configurations achieved an RMSE of approximately 0.72°C , which indicates good agreement with measured data, particularly at stations located 50 m or more from the outfall (Figure 4.12). Beyond this distance, the model closely aligned with observed temperature profiles, affirming the feasibility of using LSPIV as a practical method for deriving Q inputs in thermal effluent simulations.

At stations P15 and P16, located less than 50 m from the discharge, deviations were primarily influenced by the choice of ΔT rather than Q . This finding presents an opportunity to enhance the model by integrating insights from the previous chapter on Bayesian optimisation with the results of the LSPIV analysis. In the previous chapter, UAV-derived thermal imagery proved effective for calibrating excess temperature (ΔT) but fell short in accurately estimating discharge (Q). This limitation motivated the use of optical UAV footage as a cost-effective alternative for discharge estimation through LSPIV.

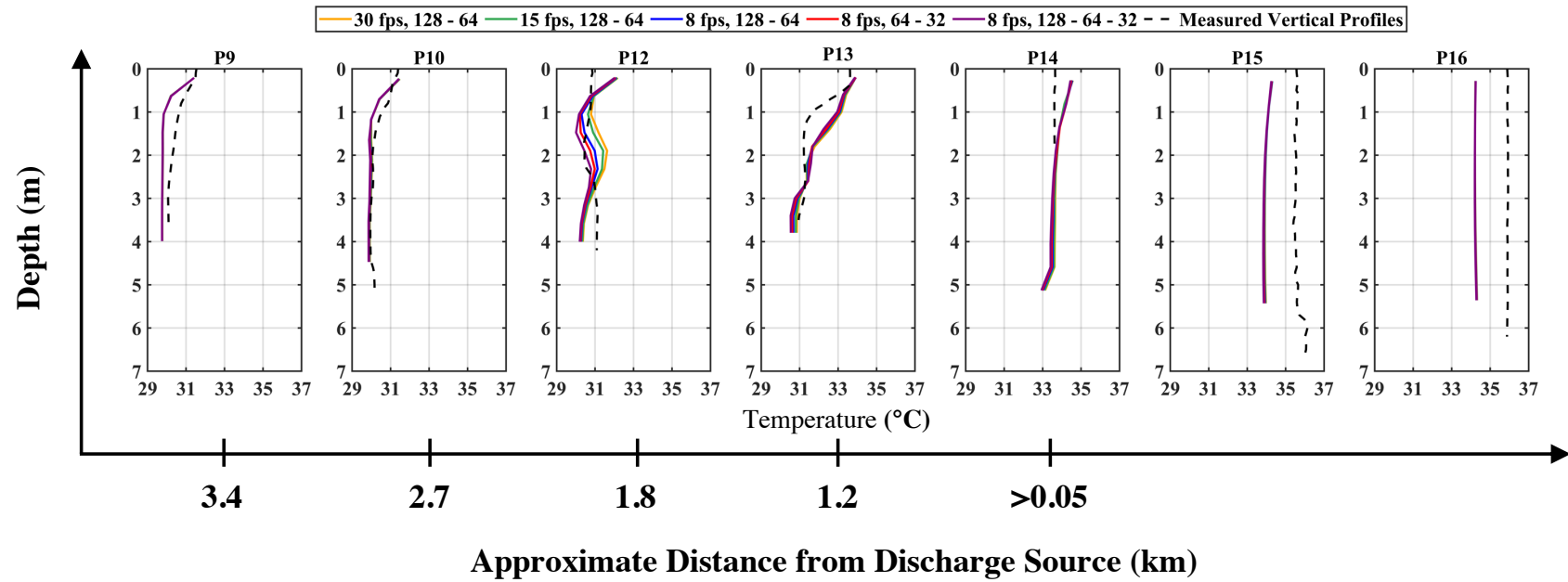


Figure 4.12: Observed (dashed lines) vs. simulated (solid lines) vertical temperature profiles for different LSPIV configurations. The profiles compare results for LSPIV setups with varying image sampling frequencies and interrogation area sizes. The x -axis indicates the approximate station distance from the discharge outfall, while the y -axis represents depth from the surface (m).

To evaluate whether LSPIV-derived discharge estimates can supplement the hydrodynamic model's Q values and combine effectively with ΔT values derived from Bayesian calibration, a comparative analysis was conducted using five model configurations. For this analysis, the 30 fps 128–64 px² case, which exhibited the best alignment with ADCP data during sensitivity analyses, was selected as the representative configuration for Q .

The model comparisons included three setups from the Bayesian optimisation chapter:

- **Baseline Model:** Utilising literature-based values for both Q and ΔT .
- **Profile-Calibrated Model:** Both Q and ΔT values were derived through Bayesian optimisation using in situ profile data.
- **UAV-Calibrated Model:** Both Q and ΔT values were derived through Bayesian optimisation using thermal UAV imagery.

Additionally, two new configurations were introduced to incorporate the LSPIV-derived discharge estimates:

- **UAV-Derived Model:** Both Q and ΔT values were extracted from UAV datasets. Specifically, Q was derived using LSPIV analysis, and ΔT was based on the excess temperature computed from the UAV-derived thermal image.
- **Hybrid Model:** Q was derived from LSPIV analysis, while ΔT was obtained through Bayesian optimisation applied to UAV thermal imagery.

Table 4.5: Summary of model results, including flow rate (Q), excess temperature (ΔT), parameter sources, and root mean square error (RMSE) between observed and simulated temperature profiles.

Model	Q (m ³ /s)	ΔT (°C)	Q Source	ΔT Source	RMSE (°C)
Baseline Model	197.00	5.00	Pokavanich and Alosairi (2014)	Pokavanich and Alosairi (2014)	0.8740
Profile-Calibrated	205.95	6.62	BayesOpt (Profile Data)	BayesOpt (Profile Data)	0.5401
UAV-Calibrated	98.44	6.19	BayesOpt (UAV Thermal Images)	BayesOpt (UAV Thermal Images)	0.9734
Hybrid Model	235.23	6.19	LSPIV	BayesOpt (UAV Thermal Images)	0.5918
UAV-Derived	235.23	5.44	LSPIV	Extracted Max (UAV Thermal Images)	0.7181

Table 4.5 summarises the key parameters and results for all configurations, including the sources of Q and ΔT , their respective values, and the RMSE achieved when compared with in situ temperature profile measurements.

Figure 4.13 and Table 4.5 summarise the model results. The baseline model, which relied on literature-based values for both Q and ΔT , demonstrated reasonable agreement with observed temperature profiles. Compared to the baseline model, the UAV-derived model achieved a modest reduction in RMSE, decreasing from 0.87 to 0.72 °C. While this improvement is not substantial, it is notable and highlights the potential of the UAV-derived approach as a viable alternative when literature-based data are unavailable. Additionally, it can serve as a practical starting point for further refinements, such as targeted calibration efforts.

The hybrid model, which combines LSPIV-derived Q with optimised ΔT from Bayesian calibration, demonstrated significant improvement over the UAV-calibrated model, reducing the RMSE from 0.97 to 0.59 °C. Remarkably, the hybrid model achieved RMSE values comparable to the profile-calibrated model (0.54 °C), which represents the most accurate fit against the validation dataset. These findings reinforce the potential of LSPIV-based discharge estimation to address gaps in traditional modelling approaches and complement existing calibration methods, ultimately enhancing overall model performance.

While the integration of LSPIV-derived discharge into the hydrodynamic model demonstrated promising results, several uncertainties and limitations must be acknowledged. The scaling factor (α) used to convert surface velocities into depth-averaged velocities was derived from Dataset A, representing conditions downstream of the discharge. This approach assumes a similar shape to the vertical velocity profile along the transect, which may not hold true in regions closer to the outfall where turbulence is more pronounced. Furthermore, tidal dynamics add complexity, as velocity profiles can vary significantly across different tidal phases. For instance, an α value derived during high tide may not accurately represent conditions during low or slack tide. Developing stage-specific α -rating curves that account for these variations could improve the applicability of LSPIV-based discharge estimation across broader tidal conditions, especially in dynamic coastal environments where flow variability is substantial.

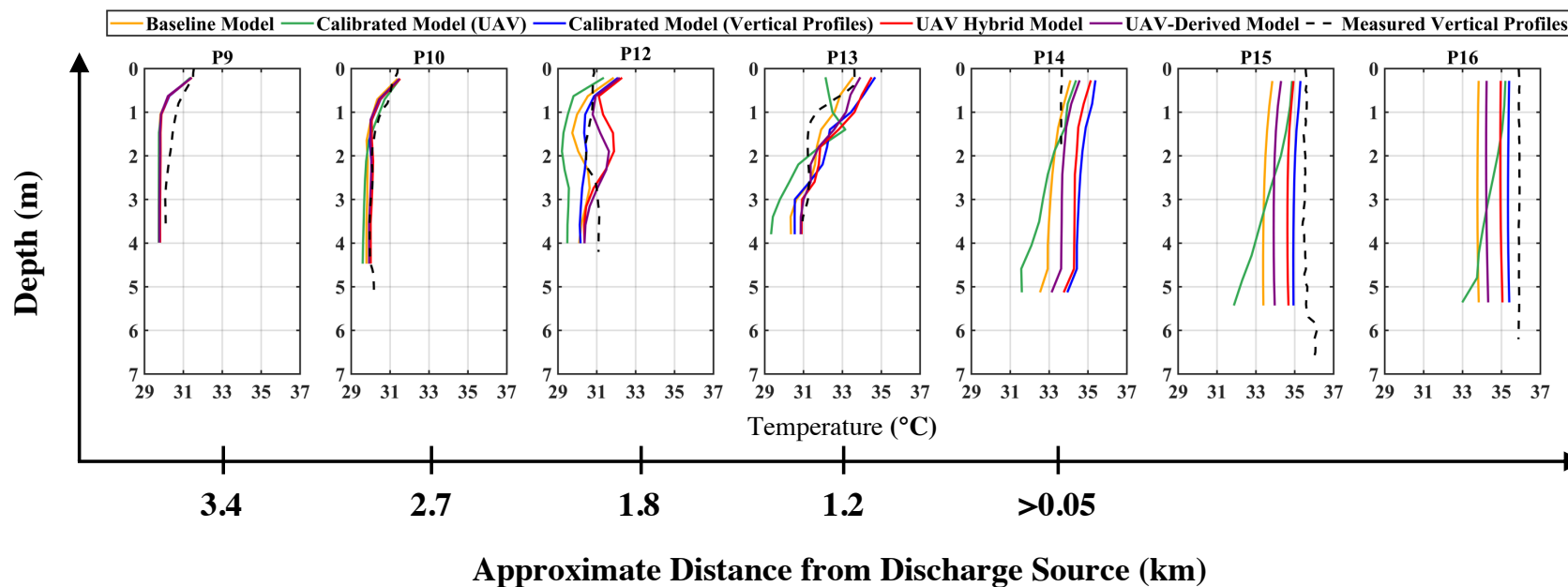


Figure 4.13: Observed (dashed lines) vs. simulated (solid lines) vertical temperature profiles across datasets. The profiles compare results from the baseline model, UAV-calibrated model, profile-calibrated model, Hybrid model, and UAV-derived model. The x -axis indicates the approximate station distance from the discharge outfall, while the y -axis represents depth from the surface (m).

The selection of the transect location may also influence discharge estimates. Shifting the transect closer to or further from the discharge source could introduce variability in the estimated discharge values. However, as discussed earlier, the impact of such variability on far-field model results is likely minimal. Nonetheless, it remains a potential source of uncertainty that must be acknowledged.

Additionally, this integration exercise was conducted during high tide, when boundary conditions dominate the domain and the thermal plume is compressed against the coastline, occupying a smaller area within the domain (see Figure 2.4). These high tide conditions may have diminished the apparent sensitivity of Q variations on simulated temperature profiles. To achieve a more comprehensive analysis, future experiments should be conducted during low tide, when the thermal plume is more dispersed across the domain. Such conditions could potentially increase the sensitivity of discharge variations on model performance, offering further insights into the robustness of the approach.

A key limitation of this study was the absence of direct discharge measurements along the transect for validation. Although LSPIV proved to be a viable method for estimating discharge, the lack of ground-truth data restricted the ability to directly compare LSPIV-derived Q values with actual measurements. Future studies should incorporate ADCP measurements along the transect, complemented by detailed bathymetric surveys, to validate the LSPIV-derived estimates. Such efforts would provide essential validation and increase confidence in the results.

Despite these limitations, this study provides compelling evidence supporting the viability of LSPIV for discharge estimation of coastal power plants in tidal environments, a novel application that has been previously unexplored. The UAV-derived model demonstrated that optical imagery can be used to reliably estimate Q , offering a practical alternative in scenarios where literature-based or in situ discharge measurements are unavailable. The hybrid model further demonstrated the potential of LSPIV-derived discharge to complement optimised ΔT values from Bayesian calibration, effectively addressing limitations in parameter estimation due to the calibration data's restricted spatial coverage of the plume.

4.4 Conclusion

This chapter investigated the application of Large-Scale Particle Image Velocimetry (LSPIV) to estimate discharge from a power plant outflow in a dynamic coastal environment. Traditional methods, including Acoustic Doppler Current Profilers (ADCPs), while accurate, are often labour-intensive, costly, and impractical to deploy in challenging conditions. LSPIV provides a contactless, image-based alternative by estimating surface velocities through the motion of tracer particles on the water surface. Combined with affordable UAV technology, LSPIV emerges as an efficient, practical and flexible solution for data acquisition in such environments.

The LSPIV method divides each image into smaller regions, known as interrogation areas (IAs), and estimates particle displacement between consecutive images using cross-correlation techniques. The velocity field is then calculated by dividing the displacement by the time interval between image pairs. In this study, LSPIV analysis was performed using the open-source PIVlab software, which employed ensemble multi-pass Fast Fourier Transform (FFT) techniques with window deformation to estimate particle displacement. LSPIV results are highly sensitive to analysis parameters, such as image sampling frequency and IA size, requiring careful consideration of tracer characteristics in the scene to ensure accurate outcomes.

The primary objectives of this study were twofold. First, it investigated the feasibility of using LSPIV to estimate surface velocities in turbulent coastal discharges. This included assessing the sensitivity of analysis parameters and evaluating their influence on velocity estimates and, consequently, on discharge estimates derived using the velocity-area method. Second, the study explored the integration of LSPIV-derived discharge estimates into a three-dimensional hydrodynamic model to simulate the transport and dispersion of thermal effluents. Accurate discharge rate (Q) serves as a fundamental input required for the model to accurately represent the continuous discharge from the plant.

Different analysis configurations, created by varying image sampling frequencies and IA sizes, revealed a nuanced interplay between these parameters and the surface textures generated by the turbulent discharge, which served as tracers at the study site. Surface textures rapidly decayed, deformed, or merged as they travelled downstream, with

higher sampling frequencies preserving these features more effectively for sharper correlation detection. Conversely, regions with slower, intermittent tracers benefited from lower sampling frequencies, allowing sufficient displacement to be captured between frames. Similarly, larger IAs were effective for capturing flow dynamics in areas with sparse or intermittent tracers, while smaller IAs preserved finer details and produced more accurate velocity estimates in regions with dense tracer textures.

These findings highlight the importance of calibrating the sampling frequency and IA size in tandem to align with the characteristics of the tracers and flow dynamics in the specific region of interest. If tracer displacement exceeds the IA size between frames, cross-correlation accuracy diminishes, as the IA cannot effectively track larger displacements. The sampling frequency determines how far tracers move between frames, while the IA size sets the limit for how large a tracer displacement can be tracked within the defined search area during the cross-correlation process.

The sensitivity of LSPIV-derived velocity fields to image sampling frequency and IA size revealed subtle differences in velocity estimates and, consequently, discharge estimates. These differences were accompanied by more pronounced variations in computational time. To evaluate the impact of the variation in discharge estimates (Q) and whether the additional computational time is justified, multiple hydrodynamic model simulations were conducted using the LSPIV-derived Q values from different configurations, while keeping the effluent excess temperature (ΔT), extracted directly from UAV thermal images, constant across all setups. The results showed that the variations in Q had minimal impact on the model's performance, as the simulation results were similar across configurations. As such, modellers should evaluate their numerical configurations and study objectives to determine an acceptable level of sensitivity in results that aligns with the intended application.

To further capitalise on the potential of LSPIV, a hybrid approach was developed by combining Q values derived from LSPIV with optimised ΔT values obtained through Bayesian optimisation against UAV-derived thermal images. This hybrid model demonstrated error reductions comparable to the model calibrated using in situ profile data, which represents the most accurate fit against the validation observations. Additionally, the hybrid model outperformed configurations relying solely on UAV-calibrated Q and

ΔT , as well as models using literature-based estimates for these parameters.

Despite limitations in processing assumptions, such as the use of a constant scaling factor (α) and the lack of ground-truth discharge measurements, this work provides a compelling proof of concept. It advances the applicability of LSPIV beyond riverine systems to dynamic coastal environments and offers valuable insights for future applications. These findings highlight the enhanced predictive capability achieved by integrating LSPIV-derived discharge with optimised temperature inputs, and the importance of combining complementary data sources to improve model performance.

Chapter 5

Conclusions

5.1 Summary of Thesis Achievements

Advancing the capabilities of thermal effluent models improves our ability to understand, predict, and manage their impacts, ensuring the protection of marine ecosystems and supporting sustainable and efficient plant operations in coastal environments. Building on this need, this thesis presents a comprehensive framework that integrates numerical modelling and data-driven techniques to simulate thermal effluents. By addressing challenges related to uncertainties in discharge parameters and the complexities of simulating plume dynamics under tidal forces, this work explores the application of three novel methodologies, implemented for the first time in the context of thermal effluent modelling, offering transferable knowledge and practical tools for studies in similar coastal environments.

Chapter 2 demonstrated the application of data assimilation (DA) using the Ensemble Kalman Filter (EnKF) to improve thermal effluent simulations. Using Observing System Simulation Experiments (OSSE), this chapter showcased how to optimise observational network design by determining the optimal assimilation frequency and employing an ensemble-based targeted observation method to identify the optimal observation locations which maximise the impact of analysis updates. These findings highlighted the potential of DA to enhance model accuracy while addressing limitations in retaining analysis updates in the modelled domain over longer timescales.

To address the limitation of DA's reliance on frequent observational updates to adjust

simulation results, Chapter 3 introduced Bayesian optimisation as a method for direct parameter estimation. Temperature datasets from in situ profiles, UAV thermal imagery, and satellite observations were acquired simultaneously during different tidal phases to calibrate two key model parameters: discharge rate and excess effluent temperature. We demonstrated the strengths and limitations of each dataset in parameter estimation by using them independently for calibration. Through an ensemble-based sensitivity analysis, we perturbed the discharge parameters to identify the region within the model where these parameters generate the greatest variance in simulated results. This region, corresponding to the simulated thermal plume, had a time-varying shape and distribution due to tidal dynamics. We demonstrated that the success of parameter optimisation relies on how well calibration data capture this region, while avoiding overfitting and preventing parameter compensation for unaccounted physical processes and model inaccuracies.

Chapter 4 addressed the limitation of UAV data in informing discharge rate estimation during Bayesian optimisation by applying Large-Scale Particle Image Velocimetry (LSPIV) to derive these values. We demonstrated that LSPIV represents a cost-effective and practical method for estimating discharge rates from coastal power plant outflows, extending its application beyond riverine systems. Optimal LSPIV analysis parameters, focusing on image sampling frequency and interrogation area size, were identified and evaluated for their influence on discharge estimates. By integrating discharge rates from different LSPIV configurations into a hydrodynamic model, we demonstrated that while variations in these discharge estimates—caused by changes in analysis parameters—significantly influenced the LSPIV algorithm’s run time, the simulation results remained consistent, with minimal differences across the model outputs. Additionally, the chapter highlighted that combining LSPIV-derived discharge rates with excess temperature estimates—obtained either through direct extraction from thermal UAV data or calibration against it—offered another viable data-driven approach to enhance the predictive capabilities of thermal effluent models.

5.2 Synthesis and Critical Evaluation of Results

To contextualise the findings, it is essential to establish the expected ranges for Q and ΔT based on literature and independent calculations. For discharge (Q), a plausible range of 169.8 to 238.68 m³/s was calculated based on thermal efficiency estimates for the Doha West (DW) power plant (Appendix D). This range aligns with the value of 197 m³/s reported by Pokavanich and Alosairi (2014) for DW. Regarding excess temperature (ΔT), thermal plumes in Kuwait Bay have been observed to increase seawater temperatures by up to 9°C (Al-Rashidi, 2009), while Pokavanich and Alosairi (2014) specifically reported an excess temperature of 5°C for DW. These ranges (Q : 169.8 - 238.68 m³/s, ΔT : 5 - 9°C) serve as benchmarks for assessing the consistency of the results obtained using the different methods, as detailed in Table 5.1.

Table 5.1: Comparison of methods for estimating discharge rate (Q) and excess effluent temperature (ΔT). The strengths and limitations of each method are also highlighted.

Method	Q Range (m ³ /s)	(ΔT) Range (°C)	Strengths	Limitations
DA	No direct estimation of parameters (state update only). Synthetic model (truth) varied Q and ΔT within the expected range.		- Dynamically adjusts model state in real-time, even with sparse observations, if location and assimilation frequency are optimised. - Accounts for uncertainty through ensemble spread.	- Does not directly estimate parameters ($Q, \Delta T$). - Performance depends on assimilation frequency and observation location. - Frequent updates required in dynamic tidal systems (hourly in this study).
Bayesian Optimisation	UAV: Unrealistic (outside range). Profile and Satellite: Plausible at high tide (205.95), unrealistic at low tide (< 50).	6.62 to 9.00 (within expected range).	- Directly estimates Q and ΔT. - Suitable for complex, non-linear parameter spaces and high-fidelity models. - Adaptable to multiple data sources (UAV, satellite, profiles).	- Highly sensitive to data coverage of the plume. - Prone to overfitting with temporally sparse data. - Risk of parameter compensation when optimising a limited set of parameters while neglecting others.
LSPIV	209 to 235.23 (within range).	Does not provide ΔT estimation.	- Direct, non-intrusive surface velocity measurement. - Cost-effective and can be performed frequently. - Provides spatial mapping of surface flow patterns.	- Measures only surface velocity; discharge is estimated using depth-averaged values, which are sensitive to assumed scaling factors. - Affected by surface conditions (e.g., glare, debris, wave action). - Results are sensitive to user choices in LSPIV analysis (e.g., frame rate, interrogation window size).

5.3 Future Directions

5.3.1 Proposed Observing System for Monitoring Thermal Effluents and Bay Dynamics

Based on the insights gained throughout this thesis regarding the complexities of modelling thermal plume dynamics and the critical role of observational data, a comprehensive, real-world observing system is proposed.

Accurate representation of boundary conditions is essential for reliable hydrodynamic simulations. To achieve this, a metocean buoy will be placed at the open boundary of the model domain. The buoy will be equipped with Conductivity-Temperature-Depth (CTD) sensors at surface, mid, and near-bottom layers to capture water column stratification (temperature and salinity), an upward-looking ADCP with a pressure sensor for current profiles and water levels, and a meteorological station measuring wind, solar radiation, air temperature, relative humidity, and precipitation. These measurements ensure accurate tidal forcing and the representation of temperature and salinity at the boundary. Data will be transmitted in real-time to a central server to provide continuous input for model forcing.

Given the model's sensitivity to bathymetry, especially in shallow areas, a detailed survey is needed for the intertidal zones and discharge area, where existing charts like Navionics lack coverage. This will enhance the model's ability to capture nearshore dynamics and accurately simulate plume behaviour.

Direct and continuous monitoring at the outfall is essential. Where feasible, in-pipe sensors will directly measure discharge (Q) and excess temperature (ΔT) from the power plant, transmitting data in real-time. If direct measurement is not feasible, an automated ADCP system will be deployed. This system consists of a downward-looking ADCP mounted on a motorized platform that traverses a fixed track across the outfall. At regular intervals (e.g., hourly), the ADCP will perform transects, measuring velocity profiles and discharge. The ADCP's integrated temperature sensor will also capture surface temperature, which can be assumed representative of the well-mixed outfall conditions due to turbulence.

To capture the spatio-temporal variability of the thermal plume and assess how well the model represents broader bay dynamics, a multi-component monitoring strategy is proposed. A buoy will be positioned in the mid-bay region (e.g., near site P9), where it can capture the interaction between boundary-driven flow and the thermal plume. This buoy will be equipped with surface and bottom CTD sensors and an ADCP with pressure sensor, providing continuous measurements of temperature, salinity, current profiles, and water levels. These data will be used to validate the model’s representation of tidal advection, mixing processes, and far-field plume behaviour.

Additionally, a vertical mooring line with temperature sensors (densely spaced near the surface and sparser with depth) will be deployed at the edge of the plume’s intermediate zone, marking the transition from near-field to far-field. This setup will capture vertical stratification and mixing, providing critical data for calibrating and validating the model’s vertical mixing schemes.

To supplement these fixed stations, regular profiling campaigns (e.g., weekly or bi-weekly) will be conducted along transects designed to capture the plume under different tidal phases (ebb and flood). These profiles will measure temperature, salinity, and depth, providing comprehensive spatial coverage, as highlighted in Chapter 3.

Large-scale plume mapping will be achieved using satellite-derived sea surface temperature data (e.g., Landsat). To address the limitations of satellite resolution at the outfall, high-resolution UAV-based thermal imagery will be captured over the outfall region concurrently with satellite overpasses. This dual remote sensing approach combines the wide coverage of satellite imagery with the fine detail of UAV sensor.

All data streams from this observing system will be transmitted in near-real-time to a central data management system, supporting continuous model calibration, validation, and adaptive monitoring. The system will also enable automated alerts if predefined environmental thresholds (e.g., ΔT) are exceeded.

5.3.2 Research Directions

Building on the findings of this research, several promising directions for future investigation are proposed:

- **Parameter and State Estimation via Data Assimilation:**

The limitations of Chapter 2 could be addressed by extending the data assimilation process to estimate both state variables and model parameters. This can be achieved through the concept of augmented state, which incorporates model parameters into the state vector. The data assimilation algorithm is then applied to simultaneously update both parameters and state variables whenever observations are available (Smith et al., 2008; Li et al., 2024). Successful implementation, however, requires sufficient knowledge of the relationships between parameters and state components to define cross-covariances. These cross-covariances describe the correlations between errors in state and parameter estimates, facilitating accurate updates when temperature observations are assimilated (Smith et al., 2008).

Future work could build on the methodologies presented in Chapter 2 by incorporating the augmented state approach to identify the optimal assimilation frequency and observation locations. This extension would enable an investigation into how the inclusion of parameter estimation influences the design of the observational network, offering deeper insights into the interactions between parameters and state variables.

Furthermore, integrating parameter and state estimation into data assimilation opens the door for real-time monitoring and decision support systems. By continuously updating model parameters and states as new observational data become available, adaptive models could be developed to provide real-time predictions of thermal plume behaviour. Such models would enhance the proactive management of thermal discharges, ensuring compliance with environmental regulations and mitigating their impacts on marine ecosystems.

- **Combining Multi-Source Data for Calibration:**

The Bayesian optimisation framework developed in Chapter 3 highlighted the critical role of data coverage in capturing the thermal plume and its dynamics during parameter calibration. A key limitation of individual data sources lies in their spatial constraints, often leading to incomplete plume representation. Future work could focus on the simultaneous integration of multiple data sources during calibration to

leverage their complementary strengths and address these limitations.

- **Exploration of Thermal vs. Optical LSPIV:**

The application of LSPIV in Chapter 4 utilised optical imagery to estimate surface velocities and discharge rates. A promising avenue for future research involves comparing the performance of optical LSPIV with thermal LSPIV, particularly leveraging the temperature elevation (ΔT) of the discharge. Thermal LSPIV, which utilises infrared images to track temperature-induced patterns at the water surface, offers distinct advantages. Unlike optical methods, it does not rely on natural or artificial tracers and is not restricted by daylight hours or subject to challenges such as illumination variability, thereby enhancing its robustness under certain conditions (Kinzel and Legleiter, 2019; Schweitzer and Cowen, 2021).

However, several open questions make this exploration compelling. First, the effectiveness of thermal LSPIV depends on the presence of sufficient temperature gradients in the captured field to enable feature tracking. This raises questions about its performance in scenarios where horizontal temperature differences are minimal, such as in well-mixed thermal conditions. Second, the resolution of thermal imagery is typically lower than that of optical cameras, potentially limiting its ability to detect fine-scale flow features. Future studies could systematically compare these techniques to evaluate their strengths and limitations under varying environmental conditions, providing insights into the optimal contexts for employing optical or thermal LSPIV.

- **Generating Time-Varying Parameter Estimates During Calibration:**

The parameter estimation methods employed in this thesis, including Bayesian optimisation and LSPIV, relied on snapshot observations captured at discrete moments in time. Consequently, the discharge rate and excess temperature were estimated as constant values applied over the entire simulation period. While effective within controlled conditions, this static approach does not capture the temporal variability in effluent characteristics that can arise from operational or environmental changes.

A potential future direction involves developing calibration processes capable of generating time-varying parameter estimates. By incorporating temporally rich cal-

ibration datasets—such as measurements spanning multiple days, tidal cycles, or seasons—these methods could produce dynamic parameterisations of effluent characteristics over the simulation period. This approach would enable models to more accurately capture the temporal variability of effluent dynamics, improving alignment with validation data collected over extended timescales and enhancing the model’s predictive capability for long-term simulations.

- **Scaling to Larger Domains with Multiple Discharge Sources:**

The methodologies presented in this thesis were applied to a single coastal power plant outfall. However, scaling these approaches to larger domains with multiple discharge sources, such as Kuwait Bay or the Gulf, presents both significant challenges and opportunities for future research.

Larger domains introduce additional complexities, including interactions between thermal plumes from different sources, varying operational schedules of multiple plants, and broader hydrodynamic influences. Applying data assimilation, Bayesian optimisation, and LSPIV techniques in such scenarios would require careful consideration of computational efficiency and data requirements to ensure successful scalability.

Bibliography

- Al-Nafisi, R.S., Al-Ghadban, A., Gharib, I., Bhat, N.R., 2009. Positive Impacts of Mangrove Plantations on Kuwait's Coastal Environment. *European Journal of Scientific Research* 26, 510–521.
- Al-Rashidi, T., 2009. An Analysis of Drivers of Seawater Temperature in Kuwait Bay. Arabian Gulf PhD Thesis, University of Southampton .
- Al Senafi, F., Anis, A., 2020. Wind-driven Flow Dynamics Off the Northwestern Arabian Gulf Coast. *Estuarine, Coastal and Shelf Science* 233, 106511.
- Aljohani, N.S., Kavil, Y.N., Shanas, P.R., Al-Farawati, R.K., Shabbaj, I.I., Aljohani, N.H., Turki, A.J., Abdel Salam, M., 2022. Environmental Impacts of Thermal and Brine Dispersion Using Hydrodynamic Modelling for Yanbu Desalination Plant, on the Eastern Coast of the Red Sea. *Sustainability* 14, 4389.
- Alosairi, Y., Alsulaiman, N., Rashed, A., Al-Houti, D., 2020. World Record Extreme Sea Surface Temperatures in the Northwestern Arabian/persian Gulf Verified by in Situ Measurements. *Marine Pollution Bulletin* 161, 111766.
- Alotaibi, S., 2011. Energy Consumption in Kuwait: Prospects and Future Approaches. *Energy Policy* 39, 637–643.
- Alvarado-Montero, R., Uysal, G., Collados-Lara, A.J., Şorman, A.A., Pulido-Velazquez, D., Şensoy, A., 2022. Comparison of Sequential and Variational Assimilation Methods to Improve Hydrological Predictions in Snow Dominated Mountainous Catchments. *Journal of Hydrology* 612, 127981.
- Alvir, M., Lučin, I., Kranjčević, L., 2024. Application of Remote Sensing Technologies

- on Industrial Outfalls. CoastLab 2024: Physical Modelling in Coastal Engineering and Science .
- Ardiny, H., Beigzadeh, A., Mahani, H., 2024. Applications of Unmanned Aerial Vehicles in Radiological Monitoring: a Review. *Nuclear Engineering and Design* 422, 113110.
- Bandini, F., Lüthi, B., Peña-Haro, S., Borst, C., Liu, J., Karagkiolidou, S., Hu, X., Lemaire, G.G., Bjerg, P.L., Bauer-Gottwein, P., 2021. A Drone-borne Method to Jointly Estimate Discharge and Manning’s Roughness of Natural Streams. *Water Resources Research* 57, e2020WR028266.
- Baracchini, T., Chu, P.Y., Šukys, J., Lieberherr, G., Wunderle, S., Wüest, A., Bouffard, D., 2020a. Data Assimilation of in Situ and Satellite Remote Sensing Data to 3d Hydrodynamic Lake Models: a Case Study Using Delft3d-flow V4. 03 and Openda V2. *4. Geoscientific Model Development* 13, 1267–1284.
- Baracchini, T., Hummel, S., Verlaan, M., Cimatoribus, A., Wüest, A., Bouffard, D., 2020b. An Automated Calibration Framework and Open Source Tools for 3d Lake Hydrodynamic Models. *Environmental Modelling & Software* 134, 104787.
- Berthet, L., Ramos, M.H., Perrin, C., Andréassian, V., Loumagne, C., et al., 2009. Can Discharge Assimilation Methods Be Used to Improve Flood Forecasting When Few Data Are Available? *IAHS publication* 333, 94.
- Bierman, G.J., 1977. Factorization Methods for Discrete Sequential Estimation. Academic Press.
- Biggs, H., Smart, G., Doyle, M., Eickelberg, N., Aberle, J., Randall, M., Detert, M., 2023. Surface Velocity to Depth-averaged Velocity—a Review of Methods to Estimate Alpha and Remaining Challenges. *Water* 15, 3711.
- Biggs, H., Smart, G., Holwerda, N., Doyle, M., McDonald, M., Ede, M., 2021. River Discharge From Surface Velocity Measurements-a Field Guide for Selecting Alpha. *Envirolink Advice Report*. Christchurch, New Zealand .
- Bishop, C.H., Etherton, B.J., Majumdar, S.J., 2001. Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly weather review* 129, 420–436.

- Blum, J., Le Dimet, F.X., Navon, I.M., 2009. Data Assimilation for Geophysical Fluids, in: Handbook of numerical analysis. Elsevier. volume 14, pp. 385–441.
- Bodart, G., Le Coz, J., Jodeau, M., Hauet, A., 2024. Quantifying and Reducing the Operator Effect in Lspiv Discharge Measurements. *Water Resources Research* 60, e2023WR034740.
- Brasseur, P., 2006. Ocean Data Assimilation Using Sequential Methods Based on the Kalman Filter: From Theory to Practical Implementations. *Ocean Weather Forecasting: An Integrated View of Oceanography* , 271–316.
- Brochu, E., Cora, V.M., De Freitas, N., 2010. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. *arXiv preprint arXiv:1012.2599* .
- Bull, A.D., 2011. Convergence Rates of Efficient Global Optimization Algorithms. *Journal of Machine Learning Research* 12.
- Byrne, D., Horsburgh, K., Williams, J., 2023. Variational Data Assimilation of Sea Surface Height Into a Regional Storm Surge Model: Benefits and Limitations. *Journal of Operational Oceanography* 16, 1–14.
- Caldwell, S., Kelleher, C., Baker, E., Lautz, L., 2019. Relative Information From Thermal Infrared Imagery via Unoccupied Aerial Vehicle Informs Simulations and Spatially-distributed Assessments of Stream Temperature. *Science of the Total Environment* 661, 364–374.
- Castro Faccetti, C.F., 2020. Modelling Dilution and Transport Processes From Marine Outfall Discharges. Ph.D. thesis. University of Leeds.
- Chow, A.C., Verbruggen, W., Morelissen, R., Al-Osairi, Y., Ponnumani, P., Lababidi, H.M., Al-Anzi, B., Adams, E.E., 2019. Numerical Prediction of Background Buildup of Salinity Due to Desalination Brine Discharges Into the Northern Arabian Gulf. *Water* 11, 2284.
- Congedo, L., 2016. Semi-automatic Classification Plugin Documentation. Release 4, 29.
- Dal Sasso, S.F., Pizarro, A., Pearce, S., Maddock, I., Manfreda, S., 2021. Increasing Lspiv

- Performances by Exploiting the Seeding Distribution Index at Different Spatial Scales. *Journal of Hydrology* 598, 126438.
- Darwish, M., 2001. On Electric Power and Desalted Water Production in Kuwait. *Desalination* 138, 183–190.
- Darwish, M.A., Al-Najem, N., 2005. The Water Problem in Kuwait. *Desalination* 177, 167–177.
- Delgado, A., Herzog, H.J., 2012. A Simple Model to Help Understand Water Use at Power Plants. Massachusetts Institute of Technology .
- Delgado Martín, A., 2012. Water Footprint of Electric Power Generation: Modeling Its Use and Analyzing Options for a Water-scarce Future. Ph.D. thesis. Massachusetts Institute of Technology.
- Deltares, 2023. D-flow Flexible Mesh User Manual.
- Dérian, P., Almar, R., 2017. Wavelet-based Optical Flow Estimation of Instant Surface Currents From Shore-based and Uav Videos. *IEEE Transactions on Geoscience and Remote Sensing* 55, 5790–5797.
- Diessner, M., O'Connor, J., Wynn, A., Laizet, S., Guan, Y., Wilson, K., Whalley, R.D., 2022. Investigating Bayesian Optimization for Expensive-to-evaluate Black Box Functions: Application in Fluid Dynamics. *Frontiers in Applied Mathematics and Statistics* 8, 1076296.
- Dols, F., 2019. The Impact of Desalination and Climate Change on Salinity in the Arabian Gulf. Master's thesis. University of Twente.
- Durán-Colmenares, A., Barrios-Piña, H., Ramírez-León, H., 2016. Numerical Modeling of Water Thermal Plumes Emitted by Thermal Power Plants. *Water* 8, 482.
- Elhakeem, A., Elshorbagy, W., 2015. Hydrodynamic Evaluation of Long Term Impacts of Climate Change and Coastal Effluents in the Arabian Gulf. *Marine Pollution Bulletin* 101, 667–685.
- Evensen, G., 1994. Sequential Data Assimilation with a Nonlinear Quasi-geostrophic

- Model Using Monte Carlo Methods to Forecast Error Statistics. *Journal of Geophysical Research: Oceans* 99, 10143–10162.
- Evensen, G., 2003. The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation. *Ocean dynamics* 53, 343–367.
- Evensen, G., Vossepoel, F.C., Van Leeuwen, P.J., 2022. *Data Assimilation Fundamentals: a Unified Formulation of the State and Parameter Estimation Problem*. Springer Nature.
- Fairley, I., Williamson, B.J., McIlvenny, J., King, N., Masters, I., Lewis, M., Neill, S., Glasby, D., Coles, D., Powell, B., et al., 2022. Drone-based Large-scale Particle Image Velocimetry Applied to Tidal Stream Energy Resource Assessment. *Renewable Energy* 196, 839–855.
- Ferrara, C., Lega, M., Fusco, G., Bishop, P., Endreny, T., 2017. Characterization of Terrestrial Discharges Into Coastal Waters with Thermal Imagery From a Hierarchical Monitoring Program. *Water* 9, 500.
- Frazier, P.I., 2018. A Tutorial on Bayesian Optimization. *arXiv preprint arXiv:1807.02811*.
- Fujita, I., Muste, M., Kruger, A., 1998. Large-scale Particle Image Velocimetry for Flow Analysis in Hydraulic Engineering Applications. *Journal of hydraulic Research* 36, 397–414.
- Gaeta, M.G., Samaras, A.G., Archetti, R., 2020. Numerical Investigation of Thermal Discharge to Coastal Areas: a Case Study in South Italy. *Environmental Modelling & Software* 124, 104596.
- Geng, W., Cheng, F., Xie, Q., Zou, X., He, W., Wang, Z., Shu, Y., Chen, G., Liu, D., Ye, D., et al., 2020. Observation System Simulation Experiments Using an Ensemble-based Method in the Northeastern South China Sea. *Journal of oceanology and limnology* 38, 1729–1745.
- Gottwald, G.A., Majda, A.J., 2013. A Mechanism for Catastrophic Filter Divergence in Data Assimilation for Sparse Observation Networks. *Nonlinear Processes in Geophysics* 20, 705–712.

- Halliwell, G., Srinivasan, A., Kourafalou, V., Yang, H., Willey, D., Le Hénaff, M., Atlas, R., 2014. Rigorous Evaluation of a Fraternal Twin Ocean Osse System for the Open Gulf of Mexico. *Journal of Atmospheric and Oceanic Technology* 31, 105–130.
- He, H., Lei, L., Whitaker, J.S., Tan, Z.M., 2020. Impacts of Assimilation Frequency on Ensemble Kalman Filter Data Assimilation and Imbalances. *Journal of Advances in Modeling Earth Systems* 12, e2020MS002187.
- Hinrichs, R., Kleinbach, M., 2012. *Energy: Its Use and the Environment*. Cengage Learning. URL: <https://books.google.co.uk/books?id=Q-9dza3I0CsC>.
- Hoteit, I., Luo, X., Bocquet, M., Kohl, A., Ait-El-Fquih, B., 2018. Data Assimilation in Oceanography: Current Status and New Directions. *New frontiers in operational oceanography* , 465–512.
- Huang, W., Jiao, J., Zhao, L., Hu, Z., Peng, X., Yang, L., Li, X., Chen, F., 2023. Thermal Discharge Temperature Retrieval and Monitoring of Npps Based on Sdgsat-1 Images. *Remote Sensing* 15, 2298.
- Hussain, H., Sebzali, M., Ameer, B., 2014. Efficiency Improvement of Steam Power Plants in Kuwait. *WIT Transactions on Ecology and the Environment* 186, 173–184.
- Jamet, C., Loisel, H., 2009. Data Assimilation Methods. *Surface Ocean–Lower Atmosphere Processes* 187, 303–317.
- Jolley, M.J., Russell, A.J., Quinn, P.F., Perks, M.T., 2021. Considerations When Applying Large-scale Piv and Ptv for Determining River Flow Velocity. *Frontiers in Water* 3, 709269.
- Joyce, K.E., Duce, S., Leahy, S.M., Leon, J., Maier, S., 2018. Principles and Practice of Acquiring Drone-based Image Data in Marine Environments. *Marine and Freshwater Research* 70, 952–963.
- Khanarmuei, M., 2021. Improving Flow Characterisations in Complex Estuary and Coastal Waterways Using Lagrangian Drifter Data. Ph.D. thesis. Queensland University of Technology.
- Khanarmuei, M., Mardani, N., Suara, K., Sumihar, J., McCue, S.W., Sidle, R.C., Mc-

- Callum, A., Brown, R.J., 2021. Assimilation of Gps-tracked Drifter Data to Improve the Eulerian Velocity Fields in an Estuary. *Estuarine, Coastal and Shelf Science* 262, 107575.
- Kim, S., Yu, K., Yoon, B., Park, G., 2013. Effects of Interrogation Area on the Accuracy of Surface Image Velocimetry, in: *Proceedings of the 35th IAHR World Congress*, pp. 1–9.
- Kinzel, P.J., Legleiter, C.J., 2019. Suas-based Remote Sensing of River Discharge Using Thermal Particle Image Velocimetry and Bathymetric Lidar. *Remote Sensing* 11, 2317.
- Ko, Y.H., Kim, T.W., 2024. Assessment of the Potential Effect of Thermal Effluents on Co2 Absorption in Coastal Waters. *Frontiers in Marine Science* 11, 1338832.
- Kong, G., Li, L., Guan, W., 2022. Influences of Tidal Flat and Thermal Discharge on Heat Dynamics in Xiangshan Bay. *Frontiers in Marine Science* 9, 850672.
- Laguna-Zarate, L., Barrios-Piña, H., Ramírez-León, H., García-Díaz, R., Becerril-Piña, R., 2021. Analysis of Thermal Plume Dispersion Into the Sea by Remote Sensing and Numerical Modeling. *Journal of Marine Science and Engineering* 9, 1437.
- Lai, W., Gan, J., Liu, Y., Liu, Z., Xie, J., Zhu, J., 2021. Assimilating in Situ and Remote Sensing Observations in a Highly Variable Estuary–shelf Model. *Journal of Atmospheric and Oceanic Technology* 38, 459–479.
- Lattemann, S., Höpner, T., 2008. Environmental Impact and Impact Assessment of Seawater Desalination. *Desalination* 220, 1–15.
- Le Coz, J., Hauet, A., Pierrefeu, G., Dramais, G., Camenen, B., 2010. Performance of Image-based Velocimetry (lspiv) Applied to Flash-flood Discharge Measurements in Mediterranean Rivers. *Journal of hydrology* 394, 42–52.
- Legleiter, C.J., Kinzel, P.J., Nelson, J.M., 2017. Remote Measurement of River Discharge Using Thermal Particle Image Velocimetry (piv) and Various Sources of Bathymetric Information. *Journal of Hydrology* 554, 490–506.
- Lewis, Q.W., Lindroth, E.M., Rhoads, B.L., 2018. Integrating Unmanned Aerial Systems

- and Lspiv for Rapid, Cost-effective Stream Gauging. *Journal of Hydrology* 560, 230–246.
- Li, X., Liu, F., Ma, C., Hou, J., Zheng, D., Ma, H., Bai, Y., Han, X., Vereecken, H., Yang, K., et al., 2024. Land Data Assimilation: Harmonizing Theory and Data in Land Surface Process Studies. *Reviews of Geophysics* 62, e2022RG000801.
- Li, X., Tang, Y., Shen, Z., Zhou, F., Song, X., Wu, Y., Li, Y., 2023. A Region-optional Targeted Observation Method and Its Application in the Sea Surface Temperature Prediction Associated with the Indian Ocean Dipole. *Journal of Geophysical Research: Oceans* 128, e2023JC019781.
- Li, Z., Wang, J., Fu, L.L., 2019. An Observing System Simulation Experiment for Ocean State Estimation to Assess the Performance of the Swot Mission: Part 1—a Twin Experiment. *Journal of Geophysical Research: Oceans* 124, 4838–4855.
- Lin, J., Zou, X., Huang, F., Yao, Y., 2021. Quantitative Estimation of Sea Surface Temperature Increases Resulting From the Thermal Discharge of Coastal Power Plants in China. *Marine Pollution Bulletin* 164, 112020.
- Lincoln, S., Buckley, P., Howes, E.L., Maltby, K.M., Pinnegar, J.K., Ali, T.S., Alosairi, Y., Al-Ragum, A., Baglee, A., Balmes, C.O., et al., 2021. A Regional Review of Marine and Coastal Impacts of Climate Change on the Ropme Sea Area. *Sustainability* 13, 13810.
- Liu, D., Zhu, J., Shu, Y., Wang, D., Wang, W., Yan, C., Zhou, W., 2018. Targeted Observation Analysis of a Northwestern Tropical Pacific Ocean Mooring Array Using an Ensemble-based Method. *Ocean Dynamics* 68, 1109–1119.
- Liu, W.C., Lu, C.H., Huang, W.C., 2021a. Large-scale Particle Image Velocimetry to Measure Streamflow From Videos Recorded From Unmanned Aerial Vehicle and Fixed Imaging System. *Remote Sensing* 13, 2661.
- Liu, Y., Xie, J., Liu, Z., Gan, J., Zhu, J., 2021b. The Assimilation of Temperature and Salinity Profile Observations for Forecasting the River–estuary–shelf Waters. *Journal of Geophysical Research: Oceans* 126, e2020JC017043.
- Loos, S., Shin, C.M., Sumihar, J., Kim, K., Cho, J., Weerts, A.H., 2020. Ensemble Data

- Assimilation Methods for Improving River Water Quality Forecasting Accuracy. *Water research* 171, 115343.
- Lowe, S.A., Schuepfer, F., Dunning, D.J., 2009. Case Study: Three-dimensional Hydrodynamic Model of a Power Plant Thermal Discharge. *Journal of Hydraulic Engineering* 135, 247–256.
- Ma, J., Zhang, J., Li, R., Zheng, H., Li, W., 2022. Using Bayesian Optimization to Automate the Calibration of Complex Hydrological Models: Framework and Application. *Environmental Modelling & Software* 147, 105235.
- Ma, P., Dai, X., Guo, Z., Wei, C., Ma, W., 2017. Detection of Thermal Pollution From Power Plants on China’s Eastern Coast Using Remote Sensing Data. *Stochastic environmental research and risk assessment* 31, 1957–1975.
- Majumdar, S.J., 2016. A Review of Targeted Observations. *Bulletin of the American Meteorological Society* 97, 2287–2303.
- Manfreda, S., Miglino, D., Saddi, K.C., Jomaa, S., Eltner, A., Perks, M., Peña-Haro, S., Bogaard, T., van Emmerik, T.H., Mariani, S., et al., 2024. Advancing River Monitoring Using Image-based Techniques: Challenges and Opportunities. *Hydrological Sciences Journal* 69, 657–677.
- Mardani, N., Khanarmuei, M., Suara, K., Brown, R., McCallum, A., Sidle, R.C., 2021. Lagrangian Data Assimilation for Improving Model Estimates of Velocity Fields and Residual Currents in a Tidal Estuary. *Applied Sciences* 11, 11006.
- Masafu, C., Williams, R., Hurst, M.D., 2023. Satellite Video Remote Sensing for Estimation of River Discharge. *Geophysical Research Letters* 50, e2023GL105839.
- MathWorks, 2024a. Bayesian Optimization Algorithm. URL: <https://mathworks.com/help/stats/bayesian-optimization-algorithm.html>. accessed: November 2024.
- MathWorks, 2024b. Bayesian Optimization with bayesopt. URL: <https://www.mathworks.com/help/stats/bayesopt.html>. accessed: November 2024.
- MathWorks, 2024c. Fitrgp: Gaussian Process Regression Models. URL: <https://mathworks.com/help/stats/fitrgp.html>. accessed: November 2024.

- MathWorks, 2024d. Kernel (Covariance) Function Options. URL: <https://mathworks.com/help/stats/kernel-covariance-function-options.html>. accessed: November 2024.
- MathWorks, 2024. Wiener2: 2-D Adaptive Noise-Removal Filtering. URL: <https://www.mathworks.com/help/images/ref/wiener2.html>. accessed: 22-Dec-2024.
- Matias, R.I., Gianina, L.R., Facundo, G., Melina, D.B., Gerardo, B.B., 2021. Combined Use of Ls-piv and Cfd for the Characterization of Turbulent Flow in the Contact Chamber of ‘costa Azul’ wastewater Treatment Plant, Carlos Paz. *Journal of Hydroinformatics* 23, 1083–1097.
- McIlvenny, J., Williamson, B., Fairley, I., Lewis, M., Neill, S., Masters, I., Reeve, D.E., 2023. Comparison of Dense Optical Flow and Piv Techniques for Mapping Surface Current Flow in Tidal Stream Energy Sites. *International Journal of Energy and Environmental Engineering* 14, 273–285.
- McLaughlin, D., 2002. An Integrated Approach to Hydrologic Data Assimilation: Interpolation, Smoothing, and Filtering. *Advances in Water Resources* 25, 1275–1286.
- Ministry of Electricity and Water, 2023. Statistical Yearly Book for 2022 - Electrical Energy. Ministry of Electricity and Water, Kuwait, Kuwait.
- Mokarram, M., Mokarram, M.J., Najafi, A., 2023. Thermal Power Plants Pollution Assessment Based on Deep Neural Networks, Remote Sensing, and Gis: a Real Case Study in Iran. *Marine Pollution Bulletin* 192, 115069.
- Moossa, B., Trivedi, P., Saleem, H., Zaidi, S.J., 2022. Desalination in the Gcc Countries-a Review. *Journal of Cleaner Production* 357, 131717.
- Morelissen, R., van der Kaaij, T., Bleninger, T., 2013. Dynamic Coupling of Near Field and Far Field Models for Simulating Effluent Discharges. *Water science and technology* 67, 2210–2220.
- Morelissen, R., Vlijm, R., Hwang, I., Doneker, R.L., Ramachandran, A.S., 2016. Hydrodynamic Modelling of Large-scale Cooling Water Outfalls with a Dynamically Coupled Near-field–far-field Modelling System. *Journal of Applied Water Engineering and Research* 4, 138–151.

- Mueller, D.S., 2013. Extrapol: Software to Assist the Selection of Extrapolation Methods for Moving-boat Adcp Streamflow Measurements. *Computers & Geosciences* 54, 211–218.
- Mueller, D.S., 2016. QRev—Software for Computation and Quality Assurance of Acoustic Doppler Current Profiler Moving-boat Streamflow Measurements—Technical Manual for Version 2.8. Technical Report. US Geological Survey.
- Mueller, D.S., Wagner, C.R., Rehmel, M.S., Oberg, K.A., Rainville, F., 2013. Measuring Discharge with Acoustic Doppler Current Profilers From a Moving Boat. Technical Report. US Geological Survey.
- Oke, P.R., O’Kane, T.J., 2011. Observing System Design and Assessment. *Operational Oceanography in the 21st Century* , 123–151.
- Pearce, S., Ljubičić, R., Peña-Haro, S., Perks, M., Tauro, F., Pizarro, A., Dal Sasso, S.F., Strelnikova, D., Grimaldi, S., Maddock, I., et al., 2020. An Evaluation of Image Velocimetry Techniques Under Low Flow Conditions and High Seeding Densities Using Unmanned Aerial Systems. *Remote Sensing* 12, 232.
- Piccioni, F., Casenave, C., Baragatti, M., Cloez, B., Vinçon-Leite, B., 2022. Calibration of a Complex Hydro-ecological Model Through Approximate Bayesian Computation and Random Forest Combined with Sensitivity Analysis. *Ecological Informatics* 71, 101764.
- Pizarro, A., Dal Sasso, S.F., Manfreda, S., 2022. Vision: Video Stabilisation Using Automatic Features Selection for Image Velocimetry Analysis in Rivers. *SoftwareX* 19, 101173.
- Pokavanich, T., Alosairi, Y., 2014. Summer Flushing Characteristics of Kuwait Bay. *Journal of Coastal Research* 30, 1066–1073.
- Premathilake, L., Khangaonkar, T., 2019. Fvcom-plume—a Three-dimensional Lagrangian Outfall Plume Dilution and Transport Model for Dynamic Tidal Environments: Model Development. *Marine pollution bulletin* 149, 110554.
- Pumo, D., Alongi, F., Ciraolo, G., Noto, L.V., 2021. Optical Methods for River Monitoring: a Simulation-based Approach to Explore Optimal Experimental Setup for Lspiv. *Water* 13, 247.

- Punetha, M., 2018. Thermal Pollution: Mathematical Modelling and Analysis. *Environmental Contaminants: Measurement, Modelling and Control* , 409–431.
- Rabier, F., Liu, Z., 2003. Variational Data Assimilation: Theory and Overview, in: *Proc. ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, Reading, UK, September 8–12, pp. 29–43.
- Raffel, M., Willert, C.E., Kompenhans, J., et al., 2007. *Particle Image Velocimetry: a Practical Guide*. Springer Science & Business Media.
- Rakha, K., Al-Banaa, K., Al-Rashidi, T., 2009. Modeling of Doha Power Plant Outfall in Kuwait Bay, in: *Proceedings of the 33rd Congress of IAHR*, pp. 4735 – 4742.
- Rao, W., Tang, Y., Wu, Y., Shen, Z., Song, X., Li, X., Lian, T., Chen, D., Zhou, F., 2023. A New Ensemble-based Targeted Observational Method and Its Application in the Tpos 2020. *National Science Review* 10, nwad231.
- Ridler, M.E., van Velzen, N., Hummel, S., Sandholt, I., Falk, A.K., Heemink, A., Madsen, H., 2014. Data Assimilation Framework: Linking an Open Data Assimilation Library (openda) to a Widely Adopted Model Interface (openmi). *Environmental Modelling & Software* 57, 76–89.
- Roberts, D.A., Johnston, E.L., Knott, N.A., 2010. Impacts of Desalination Plant Discharges on the Marine Environment: a Critical Review of Published Studies. *Water research* 44, 5117–5128.
- Roy, P., Rao, I.N., Martha, T.R., Kumar, K.V., 2022. Discharge Water Temperature Assessment of Thermal Power Plant Using Remote Sensing Techniques. *Energy Geoscience* 3, 172–181.
- Ruessmeier, N., Zielinski, O., 2021. Ocean Current Observations by Infrared and Visual Large Scale Particle Image Velocimetry (lspiv), in: *OCEANS 2021: San Diego–Porto*, IEEE. pp. 1–10.
- Rutberg, M.J., 2012. Modeling Water Use at Thermoelectric Power Plants. Ph.D. thesis. Massachusetts Institute of Technology.

- Sakov, P., Oke, P.R., 2008. Objective Array Design: Application to the Tropical Indian Ocean. *Journal of atmospheric and oceanic technology* 25, 794–807.
- Salehi, M., 2017. Thermal Recirculation Modeling for Power Plants in an Estuarine Environment. *Journal of Marine Science and Engineering* 5, 5.
- Sana, A., 2015. Hydrodynamic and Thermal Dispersion Modelling of the Effluent in a Coastal Channel. *Recent Progress in Desalination, Environmental and Marine Outfall Systems* , 269–283.
- Sandery, P., 2018. Data Assimilation Cycle Length and Observation Impact in Mesoscale Ocean Forecasting. *Geoscientific Model Development* 11, 4011–4019.
- Schweitzer, S.A., Cowen, E.A., 2021. Instantaneous River-wide Water Surface Velocity Field Measurements at Centimeter Scales Using Infrared Quantitative Image Velocimetry. *Water Resources Research* 57, e2020WR029279.
- Schwindt, S., Callau Medrano, S., Mouris, K., Beckers, F., Haun, S., Nowak, W., Wieprecht, S., Oladyshkin, S., 2023. Bayesian Calibration Points to Misconceptions in Three-dimensional Hydrodynamic Reservoir Modeling. *Water Resources Research* 59, e2022WR033660.
- Scyphers, M.E., Missik, J.E., Kujawa, H., Paulson, J.A., Bohrer, G., 2024. Bayesian Optimization for Anything (boa): an Open-source Framework for Accessible, User-friendly Bayesian Optimization. *Environmental Modelling & Software* 182, 106191.
- Serpoushan, N., Zeinoddini, M., Golestani, M., 2013. An Ensemble Kalman Filter Data Assimilation Scheme for Modeling the Wave Climate in Persian Gulf, in: *International Conference on Offshore Mechanics and Arctic Engineering*, American Society of Mechanical Engineers. p. V005T06A028.
- Shaghghi, G., Javid, A.H., Allahyaribeik, S., moradi, A.M., 2024. Determining the Optimization of Seawater Concentrate Discharge of Coastal Desalination Plants Into the Marine Environment, Based on Numerical Modeling. *Journal of Environmental Health Science and Engineering* , 1–13.
- Shawky, Y., Nada, A.M., Abdelhaleem, F.S., 2013. Environmental and Hydraulic Design

- of Thermal Power Plants Outfalls “case Study: Banha Thermal Power Plant, Egypt”. *Ain Shams Engineering Journal* 4, 333–342.
- Shi, L., Tang, L., Myers, E., 2020. Variational Data Assimilation of Tides. *Journal of Marine Science and Engineering* 8, 54.
- Smith, P., Baines, M., Dance, S., Nichols, N., Scott, T., 2008. Data Assimilation for Parameter Estimation with Application to a Simple Morphodynamic Model. *Mathematics Report* 2, 2008.
- Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in neural information processing systems* 25.
- Sørensen, J.V.T., Madsen, H., Madsen, H., 2004. Efficient Kalman Filter Techniques for the Assimilation of Tide Gauge Data in Three-dimensional Modeling of the North Sea and Baltic Sea System. *Journal of Geophysical Research: Oceans* 109.
- Stamhuis, E., Thielicke, W., 2014. Pivlab—towards User-friendly, Affordable and Accurate Digital Particle Image Velocimetry in Matlab. *Journal of open research software* 2, 30.
- Tanim, A.H., Smith-Lewis, C., Downey, A.R., Imran, J., Goharian, E., 2024. Bayes_opt-swmm: a Gaussian Process-based Bayesian Optimization Tool for Real-time Flood Modeling with Swmm. *Environmental Modelling & Software* , 106122.
- Tauro, F., Piscopia, R., Grimaldi, S., 2017. Streamflow Observations From Cameras: Large-scale Particle Image Velocimetry or Particle Tracking Velocimetry? *Water Resources Research* 53, 10374–10394.
- Teledyne Marine, 2021. WinRiver II Software User’s Guide. august 2021 ed. Teledyne RD Instruments.
- Teledyne Marine, 2022. RiverPro and RioPro ADCP Guide. february 2024 ed. Teledyne RD Instruments.
- Thielicke, W., Sonntag, R., 2021. Particle Image Velocimetry for Matlab: Accuracy and Enhanced Algorithms in Pivlab. *Journal of Open Research Software* .
- Thomole, J.B., Hughes, G.O., Zhang, M., Piggott, M.D., 2024. From Piv to Lspiv:

- Harnessing Deep Learning for Environmental Flow Velocimetry. *Journal of Hydrology*, 132446.
- Thomole, J.B., Piggott, M.D., Hughes, G.O., 2023. Deep Glacier Image Velocimetry: Mapping Glacier Velocities From Sentinel-2 Imagery with Deep Learning. *Methodology* 17, 18.
- Torres, W., Torres, A., Valencia, E., Pinchao, P., Escobar-Segovia, K., Cando, E., 2024. Experimental Validation of the Remote Sensing Method for River Velocity Measurement Using an Open-source Piv Scheme—case Study: Antisana River in the Ecuadorian Andes. *Water* 16, 3177.
- USGS, 2019. Landsat 8 (L8) Data Users Handbook. version 5.0 ed. Department of the Interior, U.S. Geological Survey.
- USGS, 2022. Landsat 9 Data Users Handbook. version 1.0 ed. Department of the Interior, U.S. Geological Survey.
- Vanhellemont, Q., 2020. Automated Water Surface Temperature Retrieval From Landsat 8/tirs. *Remote Sensing of Environment* 237, 111518.
- Vélez-Nicolás, M., García-López, S., Barbero, L., Ruiz-Ortiz, V., Sánchez-Bellón, Á., 2021. Applications of Unmanned Aerial Systems (uass) in Hydrology: a Review. *Remote Sensing* 13, 1359.
- van Velzen, N., Altaf, M.U., Verlaan, M., 2016. Openda-nemo Framework for Ocean Data Assimilation. *Ocean dynamics* 66, 691–702.
- Wang, B., Sun, Z., Jiang, X., Zeng, J., Liu, R., 2023a. Kalman Filter and Its Application in Data Assimilation. *Atmosphere* 14, 1319.
- Wang, H.Y., Fang, H.M., Chiang, Y.C., 2023b. Application of Unmanned Aerial Vehicle-based Infrared Images in Determining Characteristics of Sea Surface Temperature Distribution. *Journal of Marine Science and Technology* 31, 2.
- Wang, L., Li, G., Guo, X., Zhu, J., Sui, C., Dong, X., 2024a. Monitoring of Temperature Rise in Global Nuclear Power Plant Thermal Discharge From 2013 to 2022. *Journal of Environmental Management* 366, 121844.

- Wang, L., Li, G., Shi, H., Zhu, J., Zhan, C., Zhang, X., Wang, Q., 2024b. Integrated Monitoring and Prediction of Thermal Discharge From Nuclear Power Plants Using Satellite, Uav, and Numerical Simulation. *Environmental Monitoring and Assessment* 196, 736.
- Wang, X., Verlaan, M., Veenstra, J., Lin, H.X., 2022. Data-assimilation-based Parameter Estimation of Bathymetry and Bottom Friction Coefficient to Improve Coastal Accuracy in a Global Tide Model. *Ocean Science* 18, 881–904.
- Wang, Y., Politano, M., Ho, H.C., Muste, M., Michell, F., Stallings, J., 2014. Assessment of Ice Plugging of a Cooling Water Intake by a Numerical Model. *Journal of hydraulic research* 52, 81–92.
- Watchorn, R.P., Crossouard, A.N., Whitehouse, R.J.S., 2021. The Arc-boat: a Remote Controlled Rapid Response Scour Monitoring Platform, in: 10th International Conference on Scour and Erosion (ICSE-10), pp. 1–14.
- Whitaker, J.S., Hamill, T.M., 2012. Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. *Monthly Weather Review* 140, 3078–3089.
- Williams, J.J., Esteves, L.S., 2017. Guidance on Setup, Calibration, and Validation of Hydrodynamic, Wave, and Sediment Models for Shelf Seas and Estuaries. *Advances in civil engineering* 2017, 5251902.
- Wright, J.M., Abou-Seedo, F., Clayton, D., 1996. Long Term Changes in the Fish Assemblage of Sulaibikhat Bay, Kuwait. *Kuwait Journal of Science & Engineering* 23, 47–60.
- Wu, Y., Shen, Z., Tang, Y., 2020. A Flow-dependent Targeted Observation Method for Ensemble Kalman Filter Assimilation Systems. *Earth and Space Science* 7, e2020EA001149.
- Xiang, W., Xinxin, W., Jianhua, Z., Jianchao, F., Xiu, S., Dejun, Z., 2017. Monitoring the Thermal Discharge of Hongyanhe Nuclear Power Plant with Aerial Remote Sensing Technology Using a Uav Platform, in: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE. pp. 2958–2961.
- Yamani, F., al-Kuwayt lil-Abḥāth al ‘Ilmīyah, M., al-‘Āmmah lil Bī’ah, K.H., 2004.

Oceanographic Atlas of Kuwait's Waters. Kuwait Institute for Scientific Research.
URL: <https://books.google.co.uk/books?id=LxQytwAACAAJ>.

Yu, L., Fennel, K., Wang, B., Laurent, A., Thompson, K.R., Shay, L.K., 2019. Evaluation of Nonidentical Versus Identical Twin Approaches for Observation Impact Assessments: an Ensemble-kalman-filter-based Ocean Assimilation Application for the Gulf of Mexico. *Ocean science* 15, 1801–1814.

Yuan, W., Hua, W., 2022. A Case Study of Vignetting Nonuniformity in Uav-based Uncooled Thermal Cameras. *Drones* 6, 394.

Zhang, Z., Wang, D., Cheng, Y., Gong, F., 2022. Long-term Changes and Factors That Influence Changes in Thermal Discharge From Nuclear Power Plants in Daya Bay, China. *Remote Sensing* 14, 763.

Appendix A

Hydrodynamic Model Overview

A.1 Governing Equations and Assumptions

This section presents an overview of the governing equations and assumptions underlying the D-Flow FM hydrodynamic model. The content is derived from the D-Flow FM user manual and is intended to provide a general summary. Readers seeking more in-depth technical details are encouraged to refer directly to the manual (Deltares, 2023).

A.1.1 Continuity Equation

D-Flow FM solves the depth-averaged continuity equation, derived by integrating the continuity equation for incompressible fluids ($\nabla \cdot \mathbf{u} = 0$) over the total depth, taking into account the kinematic boundary conditions at water surface and bed level. The continuity equation is given by:

$$\frac{\partial h}{\partial t} + \frac{\partial(Uh)}{\partial x} + \frac{\partial(Vh)}{\partial y} = Q$$

where U and V are the depth-averaged velocities in the x - and y -directions, respectively, and h is the water depth. The term Q represents the contributions per unit area due to discharge or withdrawal of water, precipitation, and evaporation, as given by:

$$Q = \int_0^h (q_{in} - q_{out}) dz + P - E$$

where:

- q_{in} and q_{out} are the local sources and sinks of water per unit volume, respectively.
- P is the non-local source term for precipitation.
- E is the non-local sink term due to evaporation.

The term Q accounts for the net volume of water entering or exiting the system. For example, a power plant withdrawal of water is modelled as a sink (negative contribution), while precipitation contributes as a source. The free surface can also be a source or sink depending on whether precipitation or evaporation is considered.

A.1.2 Momentum Equations in Horizontal Direction

The momentum equations in the x - and y -directions in D-Flow FM are given by the following equations:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - f v = -\frac{1}{\rho_0} \frac{\partial P}{\partial x} + F_x + \frac{\partial}{\partial z} \left(\nu_V \frac{\partial u}{\partial z} \right) + M_x$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} + f u = -\frac{1}{\rho_0} \frac{\partial P}{\partial y} + F_y + \frac{\partial}{\partial z} \left(\nu_V \frac{\partial v}{\partial z} \right) + M_y$$

Where:

- u , v , and w are the velocity components in the x , y , and z directions, respectively.
- ν_V is the vertical eddy viscosity coefficient.
- ρ_0 is the reference density of the water.
- P is the pressure field.
- f is the Coriolis force.
- Forces F_x and F_y represent the unbalance of horizontal Reynolds stresses.

- M_x and M_y are additional momentum sources or sinks.

The Reynolds stresses are modelled using the eddy viscosity concept. The horizontal eddy viscosity ν_H is significantly larger than the vertical eddy viscosity ν_V and is typically expressed as a combination of three components:

- Sub-grid scale turbulence,
- 3D-turbulence,
- Dispersion for depth-averaged simulations.

In D-Flow FM, the horizontal eddy viscosity is determined using the Smagorinsky model, resulting in a space- and time-varying coefficient. Two turbulence closure models are used to determine the vertical eddy viscosity ν_V : the $k - \epsilon$ turbulence closure model and the $k - \tau$ turbulence closure model. In addition, D-Flow FM allows the specification of a constant background mixing coefficient for both horizontal and vertical eddy viscosity.

A.1.3 Vertical Velocities

The vertical velocity w is computed from the continuity equation in the adapting σ -coordinate system. The equation is given by:

$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} + \frac{\partial(vh)}{\partial y} + \frac{\partial w}{\partial z} = h(q_{\text{in}} - q_{\text{out}})$$

This vertical velocity is relative to the moving σ -plane and can be interpreted as the velocity associated with upward or downward motions of water.

A.1.4 Hydrostatic Pressure Assumption

Under the shallow water assumption, the vertical momentum equation is reduced to a hydrostatic pressure equation. Vertical accelerations due to buoyancy effects and sudden variations in bed topography are neglected. The equation is given by:

$$\frac{\partial P}{\partial z} = -\rho gh$$

where P is the pressure and g is the gravitational acceleration.

A.1.5 Equations of State for the Density

The density of water, denoted as ρ , is a function of both salinity (s) and temperature (T).

The density can be calculated using the UNESCO formulation. The formula is given by:

$$\rho = \rho_0 + As + Bs^{3/2} + Cs^2,$$

where:

$$\begin{aligned}\rho_0 &= 999.842594 + 6.793952 \times 10^{-2}T - 9.095290 \times 10^{-3}T^2 \\ &\quad + 1.001685 \times 10^{-4}T^3 - 1.120083 \times 10^{-6}T^4 + 6.536332 \times 10^{-9}T^5 \\ A &= 8.24493 \times 10^{-1} - 4.0899 \times 10^{-3}T + 7.6438 \times 10^{-5}T^2 \\ &\quad - 8.2467 \times 10^{-7}T^3 + 5.3875 \times 10^{-9}T^4 \\ B &= -5.72466 \times 10^{-3} + 1.0227 \times 10^{-4}T - 1.6546 \times 10^{-6}T^2 \\ C &= 4.8314 \times 10^{-4}\end{aligned}$$

A.1.6 Transport of Matter

In D-Flow FM, the transport of matter, such as salinity and temperature, is described by a general transport equation. The general form of the transport equation is given as:

$$\frac{d}{dt} \int_{V(t)} c dV + \int_{\partial V(t)} c(\mathbf{u} - \mathbf{v}) \cdot \mathbf{n} dS = \int_{\partial V(t)} (K \nabla c) \cdot \mathbf{n} dS + \int_{V(t)} s dV$$

where: c is the concentration (e.g., salinity, temperature), $\mathbf{u}$ is the flow velocity, $\mathbf{v}$ is the velocity of the vertically moving control volume, K is the diffusivity coefficient matrix, and s is a source term.

For two-dimensional (depth-averaged) flow, the transport equation is simplified as:

$$\frac{\partial hc}{\partial t} + \nabla \cdot (h\mathbf{u}c) = \nabla \cdot (hK\nabla c) + hs$$

where: h is the water depth, $\mathbf{u}$ is the horizontal velocity, K is the diffusivity matrix, and s is the source term.

A.1.7 Heat Balance

The total heat flux through the free surface is given by the following equation:

$$Q_{\text{tot}} = Q_{\text{sn}} + Q_{\text{an}} - Q_{\text{br}} - Q_{\text{ev}} - Q_{\text{co}} - Q_{\text{evfree}} - Q_{\text{cofree}}$$

Where the terms represent the following:

- Q_{sn} : net incident solar radiation (short wave),
- Q_{an} : net incident atmospheric radiation (long wave),
- Q_{br} : back radiation (long wave),
- Q_{ev} : evaporative heat flux (latent heat),
- Q_{co} : convective heat flux (sensible heat),
- Q_{evfree} : evaporative heat flux (free convection latent heat),
- Q_{cofree} : convective heat flux (free convection sensible heat).

A.2 Kuwait Bay (KB) Model Setup

A.2.1 Overview of Model Configuration

This section describes the numerical setup of the Kuwait Bay (KB) model used in Chapters 2 and 3. Both two-dimensional (2D) and three-dimensional (3D) configurations were implemented at different stages of the study, with core hydrodynamic settings kept largely consistent. Unless otherwise specified, the parameters listed in Table A.1 apply to both model versions.

Table A.1: Kuwait Bay (KB) Model Setup Parameters

Parameter	Value
Number of vertical layers (3D Model)	10 sigma layers. Uniform thickness of 10%
Maximum Courant number	0.8
Turbulence model (3D Model)	k-epsilon
Smagorinsky factor in horizontal turbulence	0.2
Threshold water depth for wet and dry cells (m)	0.1
Friction type	Chezy
Uniform friction coefficient	70
Uniform horizontal eddy viscosity	1
Uniform horizontal eddy diffusivity	10
Uniform vertical eddy viscosity	1E-06
Uniform vertical eddy diffusivity	1E-06
Heat flux model forcing	Refer to Figure A.2
Water level forcing	Refer to Figure A.1c
Temperature forcing	Refer to Figure A.1a
Salinity forcing	Refer to Figure A.1b
Excess salinity in DW discharge	2 PSU (Pokavanich and Alosairi, 2014)

A.2.2 Boundary and Atmospheric Forcing Inputs

KB model was forced at the open boundary using observed water levels, temperature, and salinity data recorded at 10 minute intervals from the Salmiya station (Figure A.1). These forcings were applied uniformly along the open boundary. In the 3D configuration, temperature and salinity were also applied uniformly across the vertical column. This was considered appropriate since the boundary is located approximately 13.2 kilometres offshore from the entrance to the bay, allowing the model to develop vertical structure within the target domain, as will be shown in the following section on spin up evalua-

tion. Meteorological inputs required for the surface heat exchange module, namely wind, relative humidity, air temperature, and solar radiation, were obtained from the KISR01 station (Figure A.2) and applied uniformly throughout the domain.

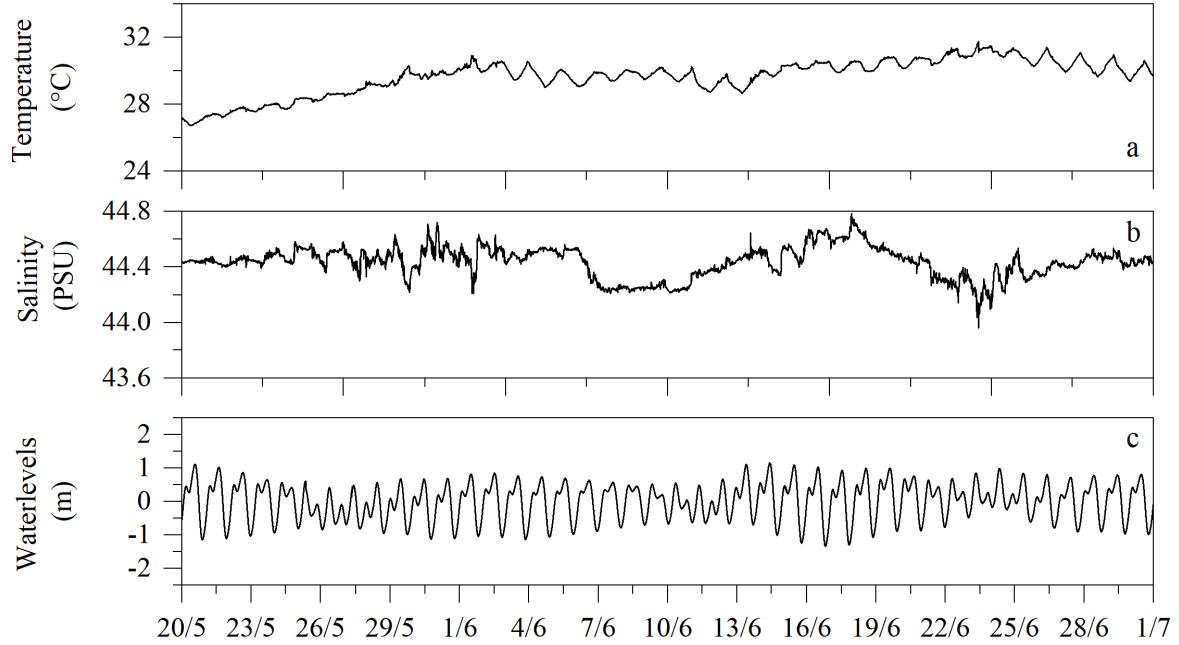


Figure A.1: Forcing data at the open boundary of the KB model: (a) Seawater temperature, (b) Salinity, and (c) Water levels, all sourced from the Salmiya station.

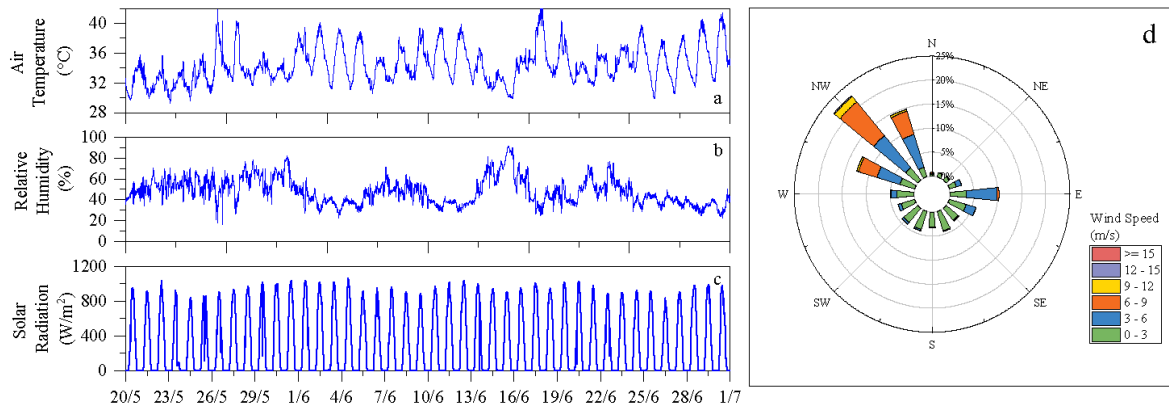


Figure A.2: Forcing data required for the heat exchange module in D-FLOW FM: (a) Air temperature, (b) Relative humidity, (c) Solar radiation, and (d) Wind measurements recorded at the KISR01 station.

A.2.3 Model Initialisation and Spin-Up Procedure

The KB model was initialised with horizontally and vertically uniform temperature and salinity fields set to 25.8°C and 44 PSU, respectively, based on field data. Initial water levels were set to mean sea level (0 m), and zero velocities were assumed throughout the domain. To ensure the model reached equilibrium, a staged spin-up procedure was employed. Due to the lack of long-term historical boundary data, a representative boundary forcing period was repeated prior to the main simulation period (June 2022). The first spin-up (26–31 May; 6 days) began from a cold start (labelled “Cold Start” in Figure A.3). The output from this run was used to initialise a second run (Spinup-1), which was followed by a third run (Spinup-2). The final run, using the output of Spinup-2, was extended through the full June simulation. As shown in Figure A.3, the final simulation reached a stable state and closely aligned with the measured surface and bottom temperatures at the KISR01 station, indicating that thermal equilibrium was achieved. Additionally, the model reproduced water levels with residuals within ± 0.5 m, as previously demonstrated in Figure 3.5.

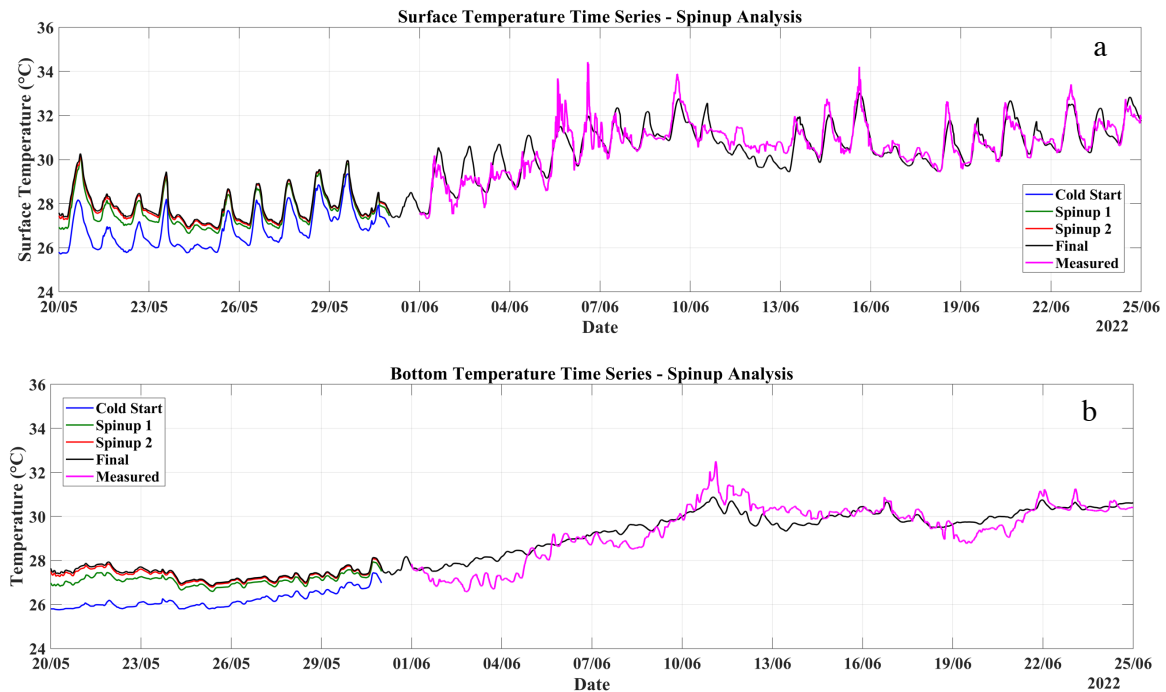


Figure A.3: Spin-up procedures of KB model showing the evolution of simulated temperature at the KISR01 station for different spin-up stages. Panel (a) shows the surface temperature, and panel (b) the bottom temperature.

A.2.4 Representation of Surface Heat Fluxes in 2D Simulations

When temperature is included in the simulation, D-FLOW FM allows users to choose how surface heat exchange is handled. The composite heat flux model, outlined in Section A.1.7, computes the net surface heat flux (Q_{tot}) from meteorological inputs. In a 3D model, the resulting temperature change is given by:

$$\frac{\partial T_s}{\partial t} = \frac{Q_{\text{tot}}}{\rho_w c_p \Delta z_s} \quad (\text{A.1})$$

where T_s is the temperature in the top layer, ρ_w is the water density, c_p is the specific heat capacity, and Δz_s is the thickness of the top layer. In this formulation, Q_{tot} is absorbed with exponential decay as a function of depth (Deltares, 2023). In a 2D depth-averaged model, Δz_s corresponds to the total water depth at each location, and Q_{tot} is absorbed entirely by the single model layer.

To assess the impact of activating surface heat fluxes in a 2D setup, an identical simulation was run using the *Transport Only* option, in which D-FLOW FM treats temperature as a passive tracer affected only by advection, diffusion, and boundary conditions. In this mode, no atmospheric heat exchange is applied (Deltares, 2023).

Figure A.4 compares both configurations against measured temperatures at KISR01. The simulation with heat fluxes enabled (green) captures the summertime warming and daily variability more accurately, aligning well with observed surface and bottom temperatures. In contrast, the simulation without heat exchange (red) shows minimal temperature evolution, driven solely by hydrodynamic processes. These results demonstrate that even in 2D mode, including surface heat fluxes is necessary for reproducing realistic temperature dynamics in the bay.

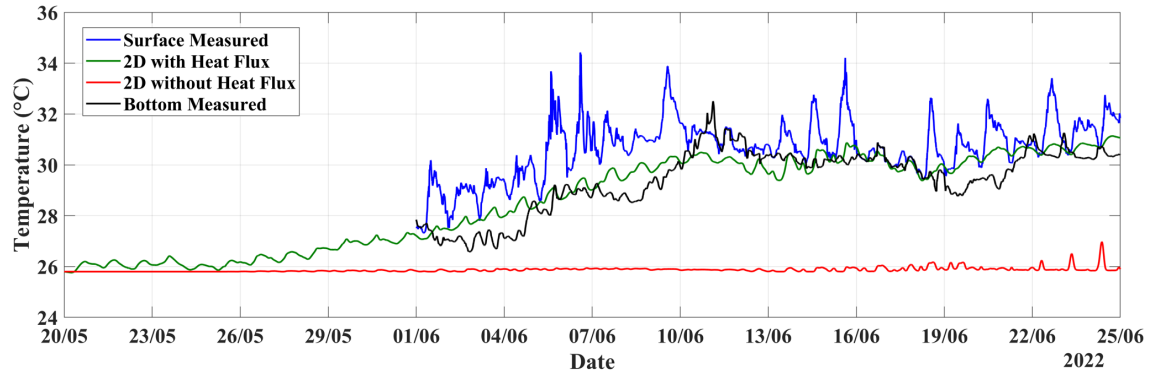


Figure A.4: Comparison of modelled and measured temperature at KISR01. The 2D simulation with heat fluxes (green) aligns more closely with observed surface (blue) and bottom (black) temperatures, while the simulation without heat fluxes (red) underestimates temporal variability and fails to match the observed warming trend.

A.3 Sulaibekhat Bay (SB) Model

A.3.1 2D Model Setup

A.3.1.1 Overview of Model Configuration

This section outlines the two-dimensional (2D) model setup for Sulaibekhat Bay (SB), used in the data assimilation experiments presented in Chapter 2. The core model parameters, turbulence settings, and forcing inputs are summarised in Table A.2.

Table A.2: Sulaibekhat Bay (KB) 2D Model Setup Parameters

Parameter	Value
Maximum Courant number	0.8
Smagorinsky factor in horizontal turbulence	0.2
Threshold water depth for wet and dry cells (m)	0.1
Friction type	Chezy
Uniform friction coefficient	70
Uniform horizontal eddy viscosity	1
Uniform horizontal eddy diffusivity	10
Uniform vertical eddy viscosity	1E-06
Uniform vertical eddy diffusivity	1E-06
Heat flux model forcing	Refer to Figure A.2
Water level forcing	Refer to Figure A.5a.
Temperature forcing	Refer to Figure A.5b.
Salinity forcing	Refer to Figure A.5c.
Excess salinity in DW discharge	2 PSU (Pokavanich and Alosairi, 2014)

A.3.1.2 Boundary and Atmospheric Forcing Inputs

The SB model was forced at its open boundary using water level, temperature, and salinity fields extracted from the 2D simulation of the KB model. These inputs varied spatially along the boundary and were updated at 10-minute intervals. Figure A.5 shows a representative time series of these boundary forcings. Meteorological data required for the surface heat exchange module were sourced from the KISR01 station and applied uniformly across the model domain (Figure A.2).

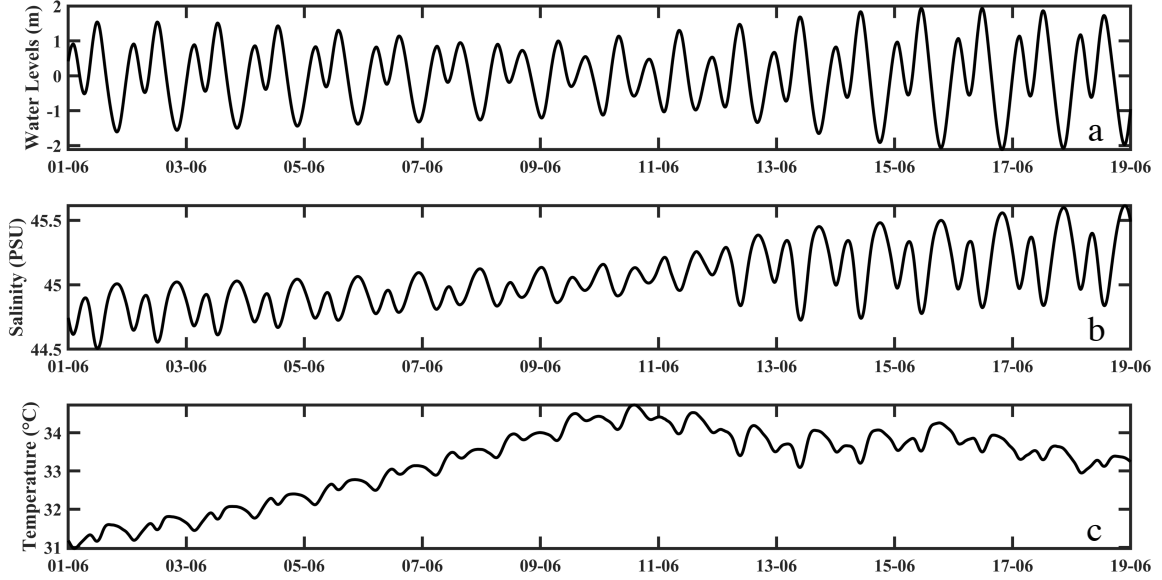


Figure A.5: Forcing data at the open boundary of the 2D SB model: (a) water levels, (b) salinity, and (c) temperature. All data were sourced from the 2D version of the KB model.

A.3.1.3 Model Initialisation and Spin-Up Procedure

Similar spin-up procedures were applied as in the KB model (Section A.2.3). Although no real measurements were available for validation due to the synthetic nature of the SB model used in the twin experiments, Figure A.6 shows that the depth-averaged temperature at station P2 in both the Truth and Baseline models (Chapter 2) remains stable and bounded, indicating that thermal equilibrium was achieved. The presence of diurnal variability and temperature fluctuation by tidal forcing further supports the establishment of realistic mixing behaviour. Specifically, during spring tides, stronger tidal currents enhance the advective transport of the plume, resulting in larger temperature oscillations. In contrast, during neap tides, weaker tidal forcing limits mixing, leading to more gradual warming and reduced variability.

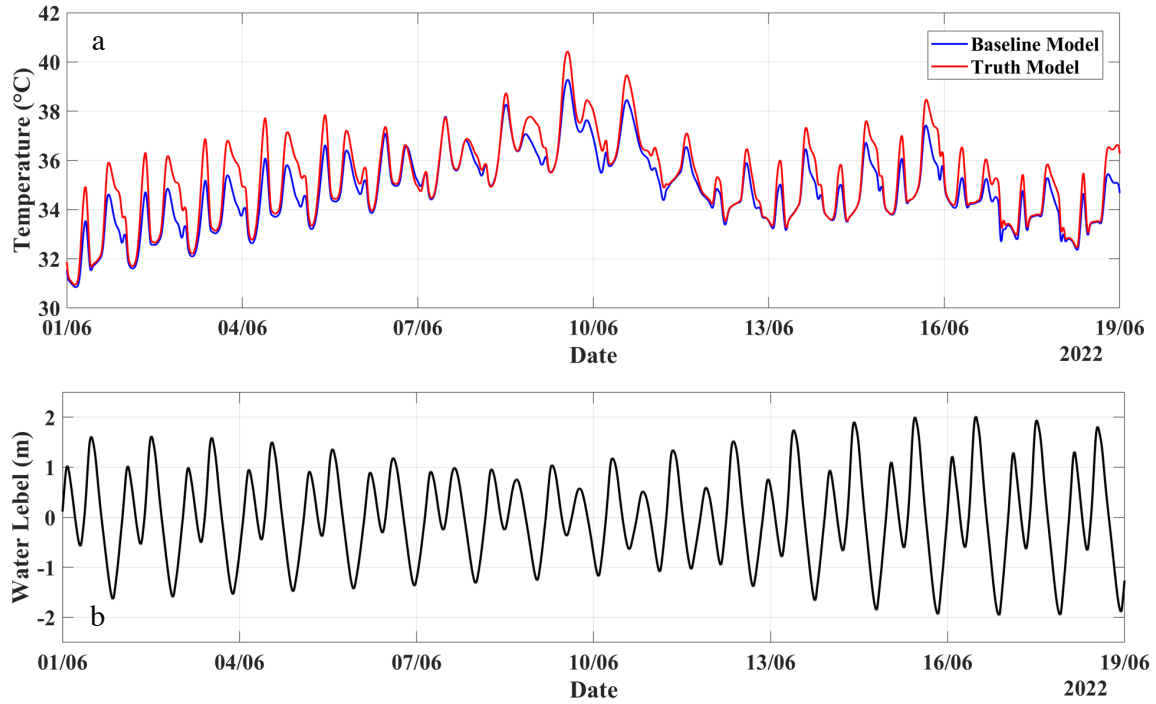


Figure A.6: Time series illustrating (a) depth-averaged temperature at station P2 from the Baseline and Truth models used in Chapter 2, and (b) the corresponding water level.

A.3.1.4 Representation of Surface Heat Fluxes in 2D Simulations

Building on the discussion in section A.2.4, a similar comparison was conducted for the 2D SB model. Figure A.7 shows the temperature time series at station P2, located near the centre of SB, for two identical model setups, with and without surface heat fluxes.

As in the KB case, including surface heat exchange leads to more realistic temperature evolution, capturing both tidal and diurnal variability. In contrast, the simulation without heat fluxes exhibits limited thermal dynamics, with temperature fluctuations driven solely by the plume being advected with the tides. This highlights that surface heat fluxes are essential for establishing the background temperature field in which the plume evolves. Without them, the model cannot reproduce the broader thermal context necessary for accurate plume simulation.

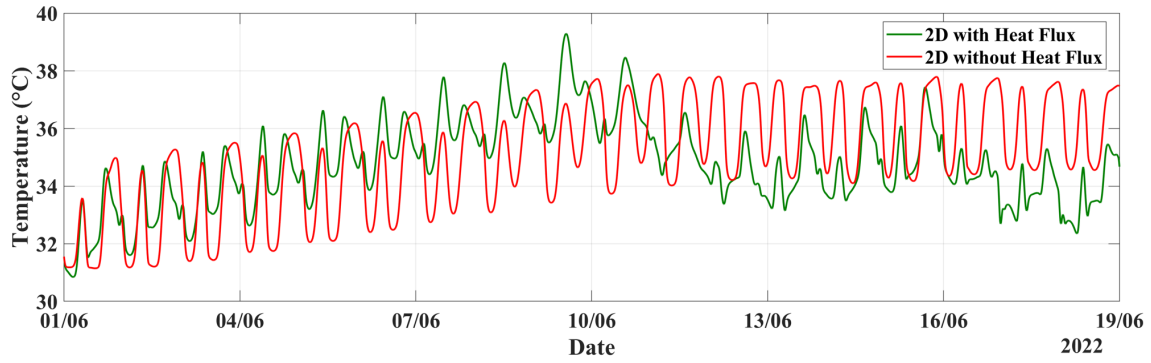


Figure A.7: Comparison of temperature time series at station P2 from two identical 2D simulations, one with surface heat fluxes (green) and one without (red).

A.3.2 3D Model Setup

A.3.2.1 Overview of Model Configuration

This section outlines the 3D SB model setup used in the parameter estimation experiments in Chapter 3. Key hydrodynamic parameters, turbulence settings, and forcing inputs are summarised in Table A.3.

Table A.3: Sulaibekhat Bay (KB) 3D Model Setup Parameters

Parameter	Value
Maximum Courant number	0.8
Smagorinsky factor in horizontal turbulence	0.2
Threshold water depth for wet and dry cells (m)	0.1
Friction type	Chezy
Uniform friction coefficient	70
Uniform horizontal eddy viscosity	0.001
Uniform horizontal eddy diffusivity	0.001
Uniform vertical eddy viscosity	1E-05
Uniform vertical eddy diffusivity	1E-05
Heat flux model forcing	Refer to Figure A.2
Water level forcing	Refer to Figure A.8
Temperature forcing	Refer to Figure A.9a
Salinity forcing	Refer to Figure A.9b
Excess salinity in DW discharge	2 PSU (Pokavanich and Alosairi, 2014)

A.3.2.2 Boundary and Atmospheric Forcing Inputs

The three dimensional SB model was forced at its open boundary using spatially varying water levels extracted from the validated KB model. Figure A.8 shows a representative time series from one location along the boundary.

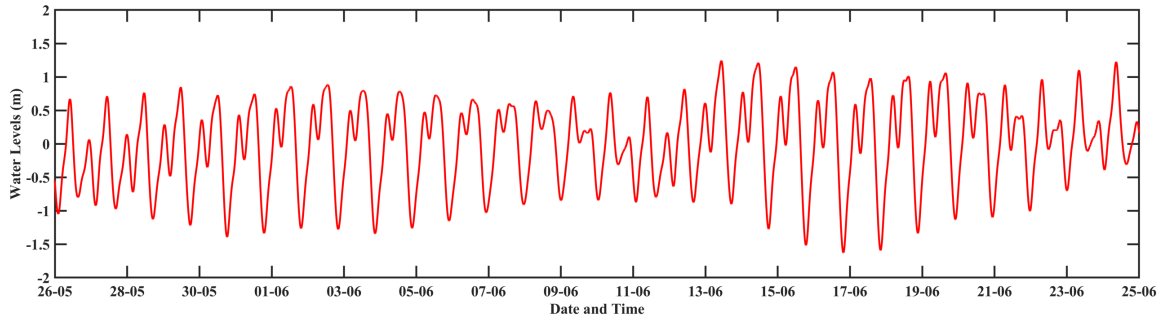


Figure A.8: Water level forcing data sourced from KB model for application in the 3D model of SB.

Vertically varying temperature and salinity boundary conditions were prescribed using near surface and near bed measurements from the KISR01 station (Figure A.9). Continuous vertical profiles were not available; only time series at the surface and bottom were recorded, along with occasional weekly profile snapshots. As a result, it was not feasible to impose complete vertical profiles at the boundary. Instead, a depth dependent structure was defined: the upper twenty percent of the water column was forced with surface observations, the lower seventy percent with bottom observations, and linear interpolation was applied across the intermediate ten percent. This twenty to seventy division was informed by the available snapshot profiles and aimed to approximate the typical vertical stratification.

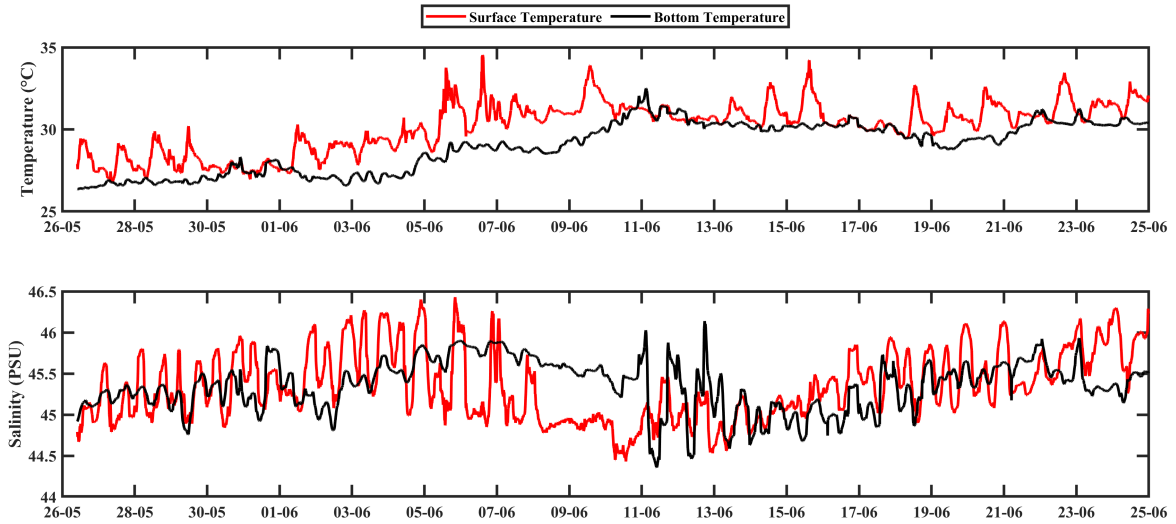


Figure A.9: Measurements used to define the vertically varying boundary conditions in the 3D model of SB: (A) surface and near-bed temperature, and (B) surface and near-bed salinity, both recorded at the KISR01 station.

Preliminary tests showed that a simpler approach, where surface and bottom values were applied only at the top and bottom layers and the model was allowed to interpolate in between, resulted in poor agreement with observed profiles (Figure A.10a). In contrast, the twenty to seventy forcing approach more accurately reproduced the vertical temperature structure observed in the field (Figure A.10a).

While vertical variability was prescribed, the same temperature and salinity profiles were applied uniformly along the horizontal extent of the open boundary. Meteorological data required for the surface heat exchange module were sourced from the KISR01 station and applied uniformly across the model domain (Figure A.2).

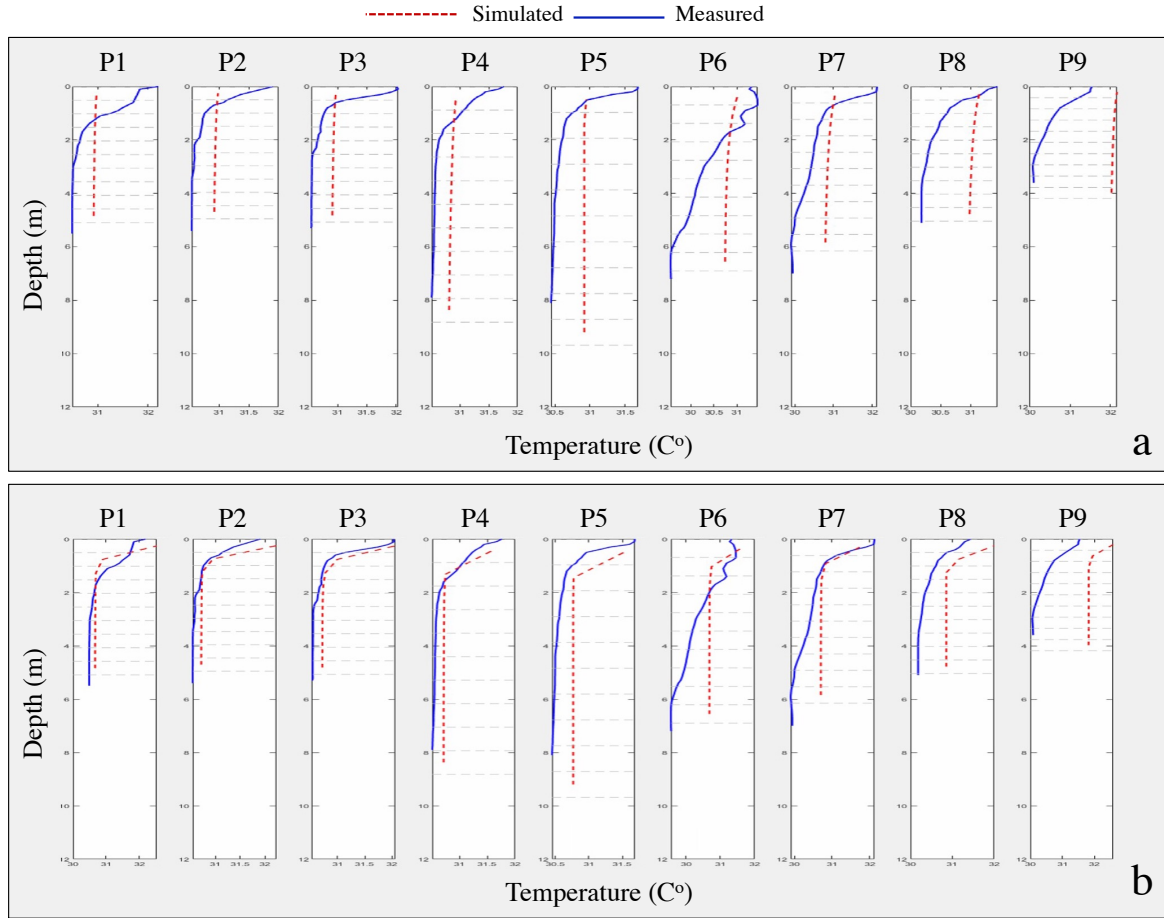


Figure A.10: Comparison of simulated and observed vertical temperature profiles at stations P1 to P9 using two boundary forcing methods. Panel (a) shows results using a simple surface and bottom layer assignment with internal interpolation. Panel (b) shows results using the depth-dependent twenty to seventy forcing scheme. Measured profiles (Dataset 2 from Chapter 3) are shown in blue, while simulated profiles are shown in dashed red.

A.3.2.3 Model Initialisation and Spin-Up Procedure

Similar spin-up procedures were applied as in the KB model (Section A.2.3). A representative boundary forcing period from 26 May to 3 June was repeated prior to the main simulation to allow the model to stabilise. Figure A.11a shows the depth-averaged temperature at station P9, located near the centre of the SB domain, extracted from the uncalibrated baseline model. Depth-averaged in situ profile measurements are overlaid as red scatter points. While some deviation is expected, the general agreement indicates that the model reached a stable thermal state suitable for calibration.

Temperature remains bounded within realistic limits, and the influence of tidal forcing is

clearly visible. Oscillations are stronger during spring tides and more subdued during neap periods (Figure A.11b). This behaviour reflects realistic mixing and exchange processes, confirming that steady state conditions were achieved prior to the calibration experiments.

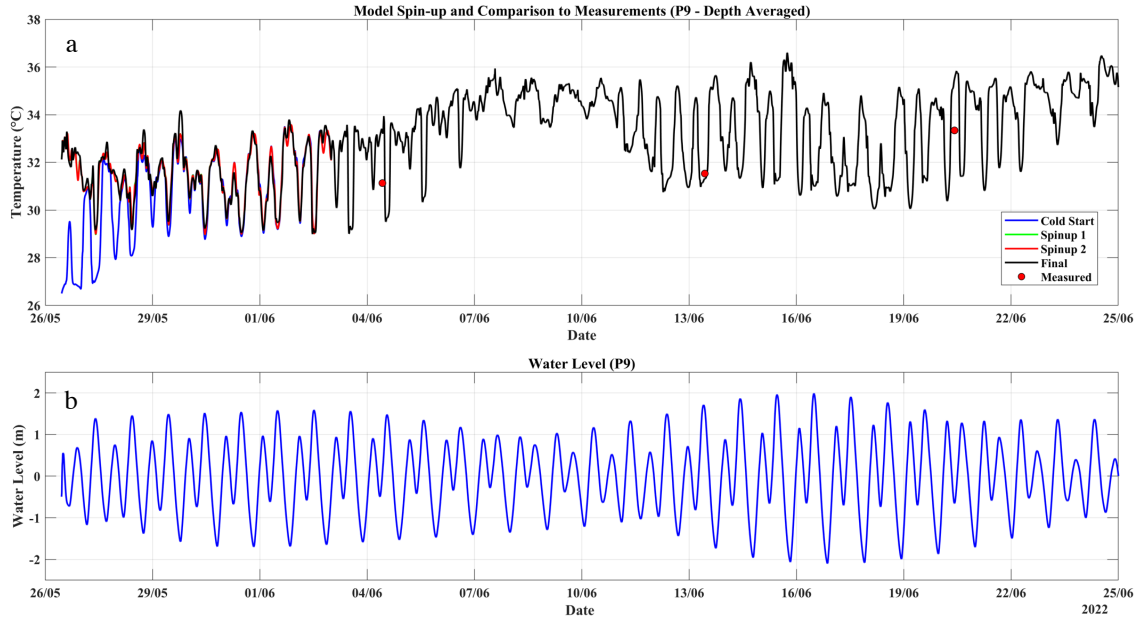


Figure A.11: Spin-up and stability analysis for the 3D SB model. (a) Depth-averaged temperature at station P9, showing the transition from cold start to final run, along with depth-averaged in situ measurements. (b) Corresponding water levels at P9.

Appendix B

Processing of Landsat TIR Scenes

This section provides an overview of the post-processing procedures used to compute sea surface temperature (SST) from satellite scenes obtained from Landsat 8 and 9. These procedures follow the guidelines outlined in the Landsat User's Handbooks (USGS, 2019, 2022).

Landsat Level-1 products consist of calibrated, scaled Digital Numbers (DNs) that represent the multispectral image data. These Level-1 DNs are delivered in an unsigned 16-bit integer format and can be converted to Top of Atmosphere (TOA) reflectance or radiance using the radiometric scaling factors provided in the scene metadata file. The following rescaling equation is applied:

$$L_{\lambda} = M_L \cdot Q_{\text{cal}} + A_L \quad (\text{B.1})$$

Spectral radiance (L_{λ}) is then converted to Top of Atmosphere (TOA) brightness temperature under the assumption of unity emissivity, using the following equation:

$$T = \frac{K_2}{\ln\left(\frac{K_1}{L_{\lambda}} + 1\right)} \quad (\text{B.2})$$

Equation parameters in B.1 and B.2 were obtained from the metadata accompanying the downloaded scenes. The parameters used in the equations are defined as follows:

- L_λ : Spectral radiance ($\text{W}/(\text{m}^2 * \text{sr} * \mu\text{m})$),
- M_L : Radiance multiplicative scaling factor,
- A_L : Radiance additive scaling factor,
- Q_{cal} : Level 1 pixel value in DN,
- T : Top of Atmosphere (TOA) brightness temperature (Kelvin),
- K_1 and K_2 : Band-specific thermal conversion constants.

Appendix C

Selection of Parameters for Optimisation

The Bayesian optimisation approach adopted in this study focused on calibrating two key parameters: the discharge rate (Q) and excess temperature (ΔT) of the thermal effluent. This section justifies the choice of these parameters, discusses their implications and limitations, and proposes a potential alternative to mitigate those limitations.

In addition to Q and ΔT , other model elements such as bathymetry and bed roughness significantly influence flow patterns, mixing, and dispersion within the model domain (Williams and Esteves, 2017). Parameters like eddy diffusivity, heat flux coefficients, and wind drag coefficients also affect the simulated temperature field (Baracchini et al., 2020b). However, the decision to focus the optimisation on Q and ΔT was based on their dominant influence over the simulated thermal plume, as they directly define the source characteristics of the effluent.

To illustrate this, a structured sensitivity analysis was conducted, in which eight relevant parameters were grouped into four conceptual categories, each representing a distinct physical process:

- **Discharge and Source Properties:** This group includes the discharge rate (Q) and excess effluent temperature (ΔT), which together define the magnitude and thermal intensity of the effluents.

- **Dispersion and Mixing:** This group includes horizontal and vertical eddy diffusivity. Horizontal diffusivity governs lateral spreading of temperature, while vertical diffusivity influences stratification and vertical mixing throughout the water column.
- **Air–Water Heat Exchange:** This group includes the Dalton number, which governs evaporative heat loss, and the Stanton number, which controls convective (sensible) heat exchange between the water surface and the atmosphere. Both are non-dimensional empirical coefficients used to parameterise latent and sensible heat fluxes, respectively (Deltares, 2023).
- **Bathymetry and Bed Roughness:** This group includes bathymetry, which defines the depth and slope of the model domain, and bed roughness, which controls the interaction between the flow and the seabed. Roughness is represented by the Chezy coefficient.

The bounds for Q and ΔT were set to match those used in Chapter 3 during the Bayesian optimisation experiments. For the eddy diffusivity and heat exchange parameters, upper and lower bounds were selected based on documented ranges in Baracchini et al. (2020b) and Deltares (2023). For bathymetry, depth was adjusted by applying spatially uniform offsets of ± 1.0 m to the baseline model. For bed roughness, the Chezy coefficient was varied by $\pm 50\%$ from the baseline value. Table C.1 summarises the parameter groupings, associated baseline model values, and the bounds used in the analysis.

Table C.1: Summary of parameters and bounds used for sensitivity analysis

Group	Parameters	Baseline Model Values	Lower Bound	Upper Bound
Discharge Properties	Discharge rate (Q)	197 m ³ /s	10 m ³ /s	500 m ³ /s
	Excess effluent temperature (ΔT)	5°C	1°C	10°C
Dispersion & Mixing	Horizontal eddy diffusivity (HD)	0.001 m ² /s	4.5×10^{-10} m ² /s	100 m ² /s
	Vertical eddy diffusivity (VD)	0.001 m ² /s	4.5×10^{-10} m ² /s	100 m ² /s
Heat Exchange (Air–Water)	Dalton number (D)	0.002	0	0.05
	Stanton number (S)	0.002	0	0.05
Bathymetry and Roughness	Bathymetry depth offset (B)	Spatially varying depth field	-1.0 m	+1.0 m
	Chezy bed roughness (R)	70	35 (50% decrease)	105 (50% increase)

Figures C.1 show that varying Q and ΔT relative to the baseline model leads to pronounced changes in the thermal plume's extent, intensity, and shape, both at the surface and near the bottom. These two parameters directly influence the thermal structure by defining the source characteristics of the effluent. In contrast, variations in horizontal and vertical eddy diffusivity (Figures C.2) and heat exchange coefficients (Figures C.3) act across the entire domain, affecting the background temperature field as well as the vertical stratification and lateral spreading of temperature.

The sensitivity analysis of bathymetry (Figure C.4) shows that altering seabed elevation affects the plume's extent and distribution. A deeper bathymetry (positive offset) reduces plume expansion, pushing it closer to the source region due to increased water depth (Figure C.4c–f). Conversely, a shallower bathymetry (negative offset) allows the plume to extend further offshore, where reduced depth enhances horizontal advection (Figure C.4g–j). The roughness sensitivity analysis (Figure C.5) indicates that higher roughness values (lower Chezy coefficient) restrict plume spread, while lower roughness values (higher Chezy coefficient) promote greater lateral dispersion.

While these parameters are important for accurately reproducing the overall thermal environment, including aspects of plume development, the decision to limit calibration to Q and ΔT was driven by the need to reduce the dimensionality of the optimisation problem. This allowed for a more efficient exploration of the parameter space with fewer model evaluations, a necessary trade-off given the high computational cost of running the 3D hydrodynamic model.

While the focus on calibrating Q and ΔT was justified, this approach introduces several limitations. A primary concern is parameter compensation, where errors in other parts of the model may be absorbed into the optimised values of Q and ΔT . Calibrating only two parameters implicitly assumes that all other model components are correctly specified, which may not hold in practice. For instance, an overly diffusive model setup might lead to higher calibrated values of Q and ΔT to match observations. Similarly, boundary condition uncertainties and modelling choices in forcing representation, such as those shown in Figure A.10, can significantly influence model outputs. This may lead the calibration process to unintentionally adjust Q and ΔT to compensate for these external effects.

An alternative strategy could involve staged calibration, beginning with the most influential parameters (Q and ΔT) and progressively incorporating others. Nonetheless, the two-parameter approach adopted in this study provides a practical starting point that balances computational efficiency with improved model accuracy. Although the calibrated values of Q and ΔT may be influenced by other model uncertainties, they offer valuable insights into effluent characteristics and significantly enhance the simulation of plume dynamics.

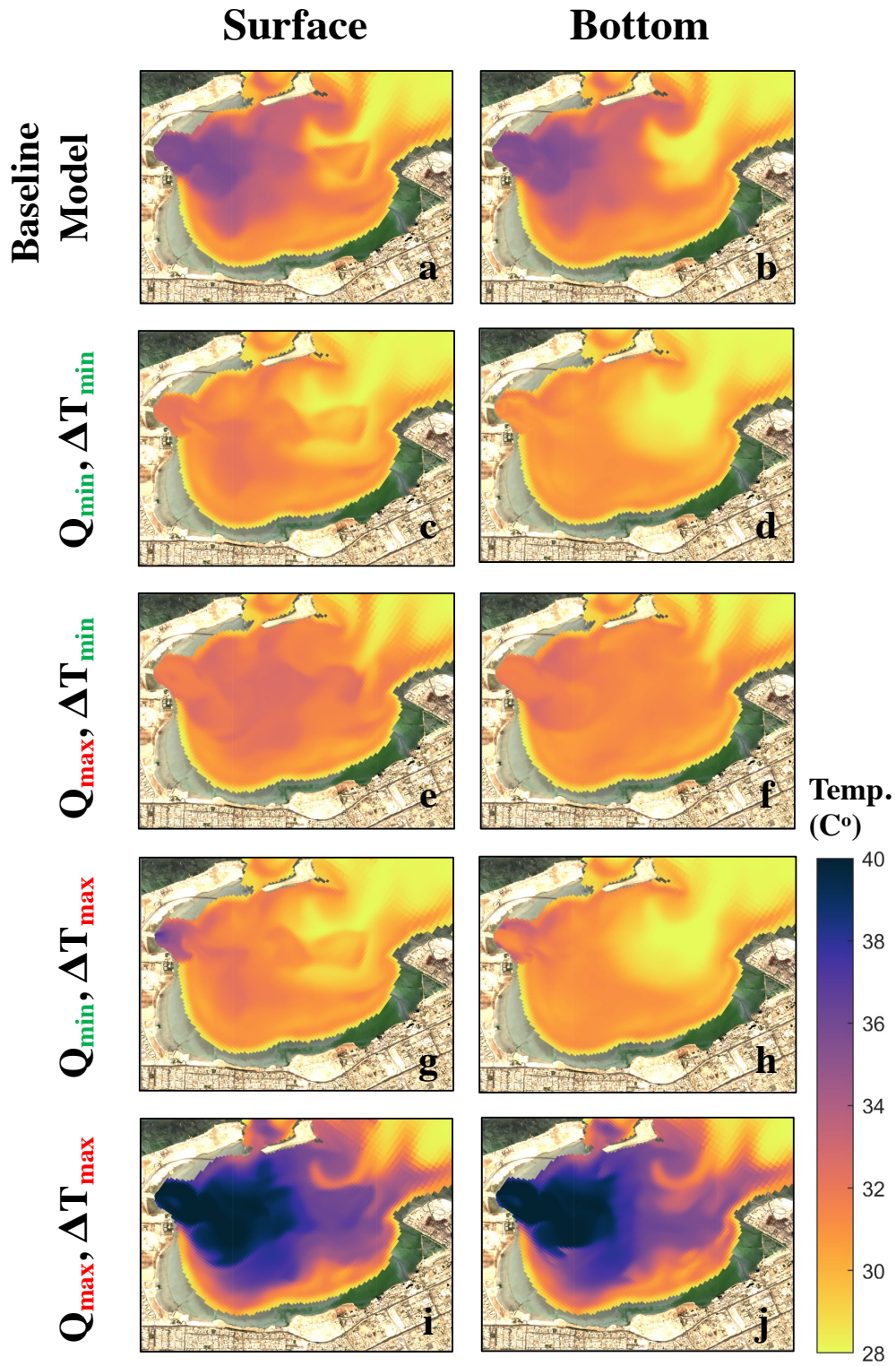


Figure C.1: Sensitivity of the simulated temperature field to variations in discharge rate (Q) and excess temperature (ΔT). Panels (a–b) show the surface and bottom temperatures from the baseline model. Panels (c–j) present results from scenarios using combinations of minimum and maximum bounds for Q and ΔT . Each row represents a distinct forcing combination, as indicated by the labels on the left.

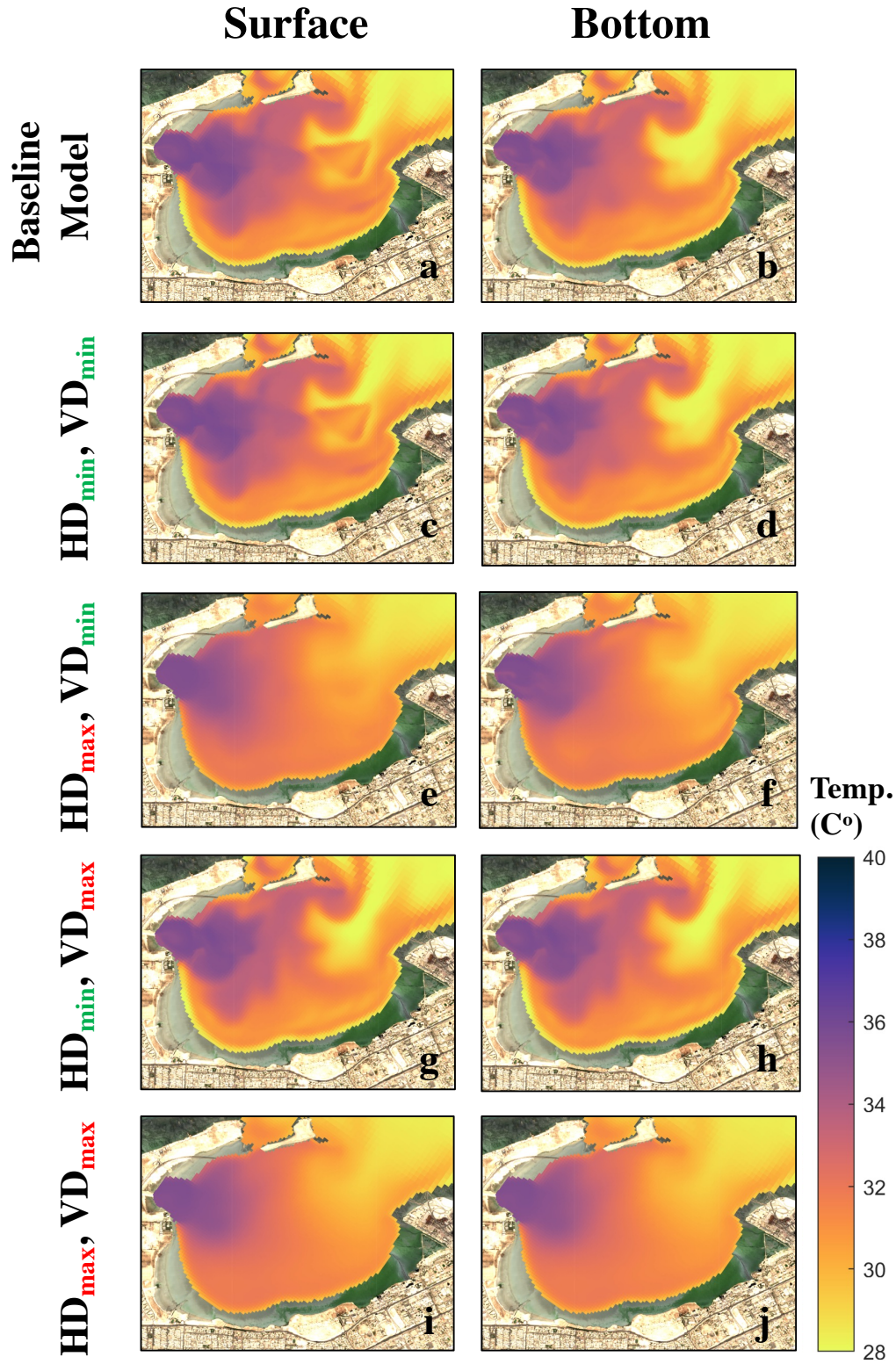


Figure C.2: Sensitivity of the simulated temperature field to variations in horizontal (HD) and vertical (VD) eddy diffusivity. Panels (a–b) show the surface and bottom temperatures from the baseline model. Panels (c–j) present results from scenarios using combinations of minimum and maximum bounds for HD and VD. Each row represents a distinct diffusivity combination, as indicated by the labels on the left.

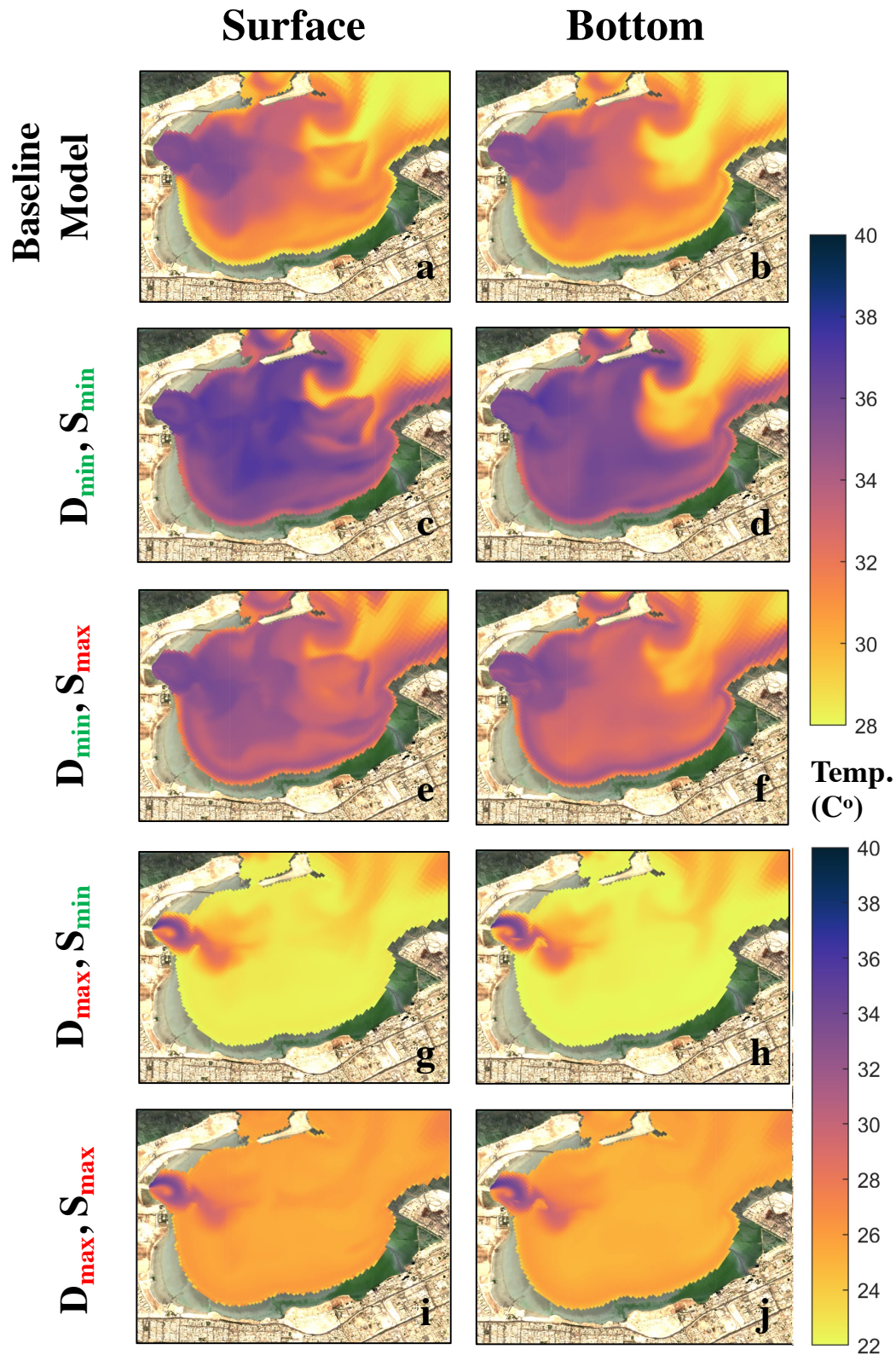


Figure C.3: Sensitivity of the simulated temperature field to variations in the Dalton (D) and Stanton (S) numbers, which govern evaporative and convective heat exchange, respectively. Panels (a–b) show the surface and bottom temperatures from the baseline model. Panels (c–j) present results from scenarios using combinations of minimum and maximum bounds for D and S. Each row represents a distinct heat exchange configuration, as indicated by the labels on the left. Note that panels (a–f) and (g–j) use different colour scales for improved readability.

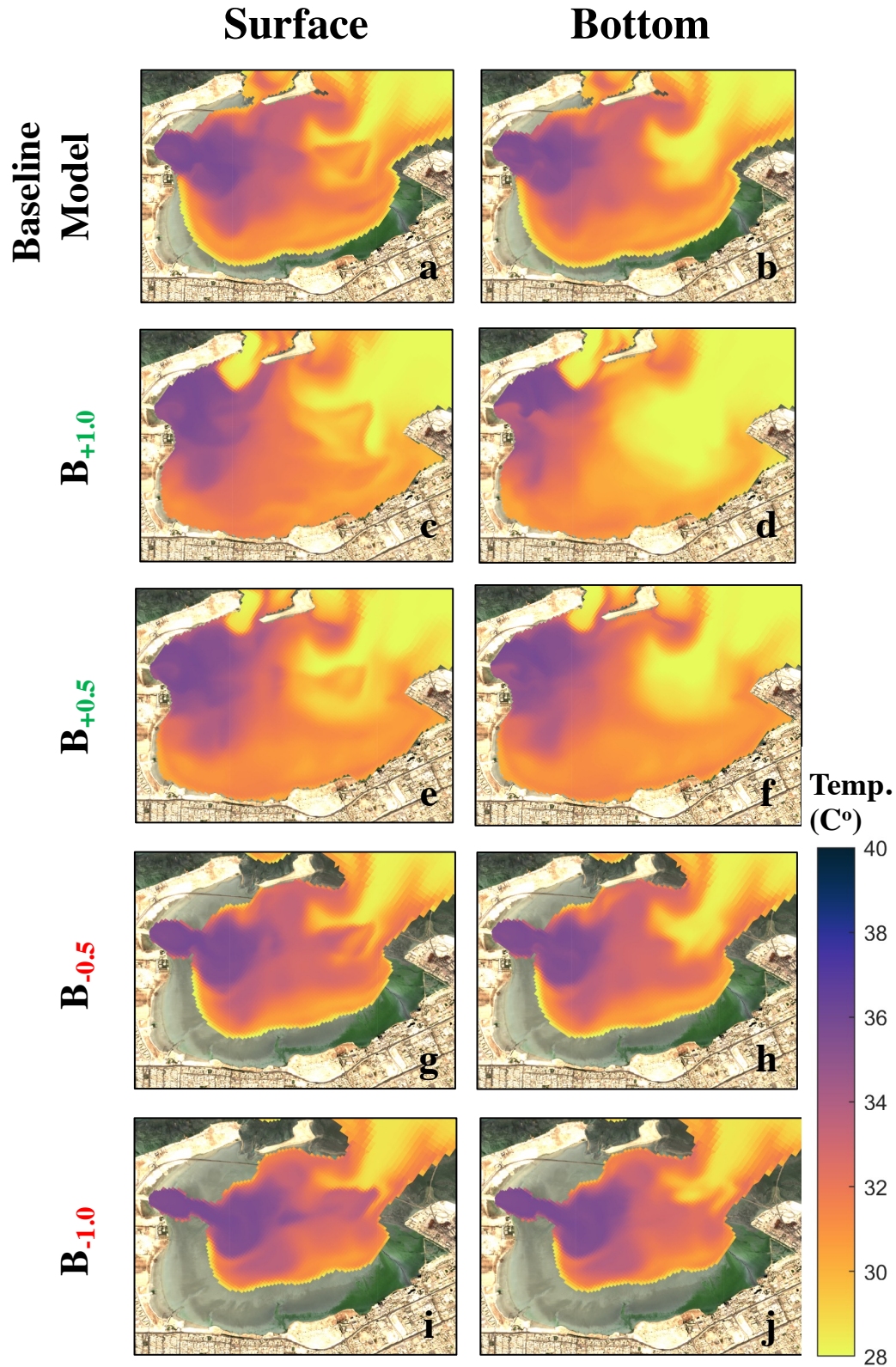


Figure C.4: Sensitivity of the simulated temperature field to bathymetry variations. Panels (a–b) show the surface and bottom temperatures from the baseline model. Panels (c–j) present results from scenarios using bathymetry adjustments of +1.0 m, +0.5 m, -0.5 m, and -1.0 m. Each row represents a distinct bathymetry configuration, as indicated by the labels on the left.

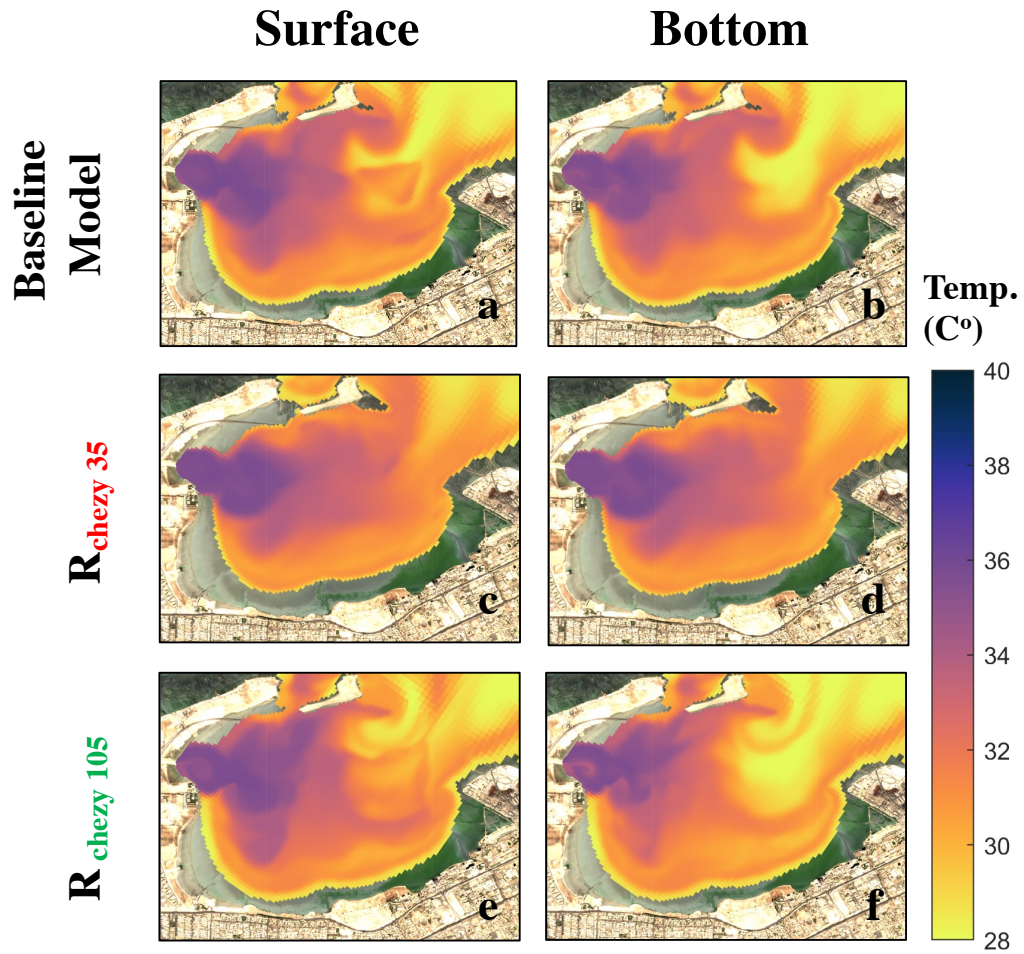


Figure C.5: Sensitivity of the simulated temperature field to variations in Chezy roughness values. Panels (a–b) show the surface and bottom temperatures from the baseline model (Chezy 70). Panels (c–d) show results for a rougher configuration (Chezy 35), while panels (e–f) show a smoother configuration (Chezy 105).

Appendix D

Estimation of Cooling Water Discharge

This section provides a systematic approach for estimating the cooling water discharge (Q) of thermal power plants, specifically for the Doha West Power Plant. The objective is to establish a plausible range for Q , providing a practical basis for comparing data-driven estimates with realistic operational conditions.

D.1 Thermal Efficiency and Waste Heat Calculation

Thermal power plants generate electricity by converting heat from various energy sources, such as coal, natural gas, uranium, solar energy, and geothermal energy (Delgado and Herzog, 2012). However, not all of this input energy is transformed into electricity. The unconverted portion becomes waste heat, which must be dissipated, typically through cooling systems that use water as the heat transfer medium (Delgado Martín, 2012). One of the most common cooling methods is once-through cooling, where water is drawn from a natural source (e.g., sea), passed through a heat exchanger to absorb waste heat, and then returned slightly warmer to the source (Rutberg, 2012).

The efficiency of a thermal power plant is defined as the ratio of useful electrical power output (P_{net}) to the total energy input (E_{in}):

$$\eta = \frac{P_{net}}{E_{in}} \quad (D.1)$$

This efficiency indicates the fraction of input energy that is converted into electricity, with the remainder classified as waste heat. Rearranging the efficiency equation allows for calculating the total energy input based on the plant's net output:

$$E_{in} = \frac{P_{net}}{\eta} \quad (D.2)$$

The waste heat (Q_{waste}) is then determined as the difference between the total energy input and the useful output energy:

$$Q_{waste} = E_{in} - P_{net} = \left(\frac{1}{\eta} - 1 \right) \times P_{net} \quad (D.3)$$

For context, a typical Combined Power and Desalination Plant (CPDP) in Kuwait, consuming 696.235 MW of fuel energy and producing 225 MW of electrical power and 45.043 MW of mechanical energy, has an equivalent efficiency calculated as:

$$\eta = \frac{225 + 43.295}{696.235} \approx 0.385 \quad (D.4)$$

This means that 38.5% of the input energy is converted into useful work, while the remaining is dissipated as waste heat (Darwish, 2001).

D.2 Estimating Cooling Water Discharge

The cooling water required to dissipate waste heat is calculated using the concept of water withdrawal per unit of waste heat (A), defined as (Delgado and Herzog, 2012):

$$A = \frac{1}{C_p \Delta T \rho_w} \quad (D.5)$$

where:

- A = Water withdrawal per unit of waste heat (L/kJ)
- C_p = Specific heat capacity of water (kJ/(kg · K))
- ΔT = The temperature increase of the water (°C)
- ρ_w = Density of water (kg/m³)

The total cooling water discharge (Q) is calculated as the product of water withdrawal per unit heat (A) and the total waste heat (Q_{waste}):

$$Q = A \times Q_{waste} = \frac{Q_{waste}}{C_p \Delta T \rho_w} \quad (D.6)$$

Substituting the expression for Q_{waste} :

$$Q = \frac{\left(\frac{1}{\eta} - 1\right) \times P_{net}}{C_p \Delta T \rho_w} \quad (D.7)$$

D.3 Application to the Doha West Power Plant

For the Doha West Power Plant, the following parameters are used:

- **Net Power Output (P_{net}):** 2541 MW (equivalent to 2541×10^3 kJ/s) (Ministry of Electricity and Water, 2023).
- **Specific Heat Capacity of Water (C_p):** 4.184 kJ/(kg · K) (Delgado and Herzog, 2012).
- **Temperature Difference (ΔT):** 4.8°C, estimated as the difference between average temperature of P15 and P1 in Dataset 2 (see Figure XX).
- **Thermal Efficiency (η):** Reported thermal efficiencies for CPDP in Kuwait, including Doha West, range between 0.34 and 0.42 (Darwish, 2001; Darwish and Al-Najem, 2005; Hussain et al., 2014). This range is used to account for variability.
- **Water Density (ρ_w):** 1029 kg/m³, sourced from the measured average density at P15.

Considering a range of thermal efficiency (η) between 0.34 and 0.42:

- For $\eta = 0.34$:

$$Q = \frac{\left(\frac{1}{0.34} - 1\right) \times 2541 \times 10^3}{1029 \times 4.184 \times 4.8} \approx 238.68 \text{ m}^3/\text{s} \quad (\text{D.8})$$

- For $\eta = 0.42$:

$$Q = \frac{\left(\frac{1}{0.42} - 1\right) \times 2541 \times 10^3}{1029 \times 4.184 \times 4.8} \approx 169.80 \text{ m}^3/\text{s} \quad (\text{D.9})$$

These calculations provide a practical estimate of the cooling water discharge for the Doha West Power Plant.