Quantifying Humpback Whale Dive Dynamics: A Data-Driven Approach using R and Interactive Visualization

By Gil Raitses, et. al
Syracuse University
Researcher, Physics Department
BioInspired Institute

https://orcid.org/0009-0008-5182-0013

IST 676: Intro To Data Science, Prof. Tin Hoang, June 19, 2024

Abstract

Humpback whale diving behavior is complex and dynamic, involving distinct phases and events within each dive cycle. Grasping these dynamics is essential for understanding the energetic costs, behavioral patterns, and individual variations in humpback whale foraging strategies. This study investigated the cycle, phase, and event dynamics of humpback whale dives using sensor data from a single humpback whale named Lavalier's Calf. We used a detailed data analysis approach, utilizing the R programming language to trim annotations, inspect data quality, and validate event and phase identification. Our analysis revealed distinct patterns in fluke stroke intensity, head lunge frequency, and overall dive duration across different phases of the dive cycle.

In this study, we further analyze humpback whale dive behaviors using data from whale tag archives. Our analysis not only looked for distinct patterns in movements but also examined temporal and cyclical dynamics. We apply k-means clustering to segment dives into distinct behavioral patterns, revealing three primary clusters: deep foraging dives, moderate depth exploration dives, and shallow surface activities. A linear regression analysis indicates a slight negative correlation between dive duration and depth. Temporal distribution analysis shows distinct time patterns for each dive phase. Additionally, we present a detailed composite dashboard for a specific event, highlighting the whale's behavior through metrics such as pitch, roll, head, norm jerk, depth, and temperature. These insights enhance our understanding of humpback whale dive dynamics, offering valuable information for marine biologists and conservation efforts. Our analysis paves the way for the development of enhanced classification models that are better informed by both behavioral context and environmental factors.

Table of Contents

1. Introduction

Humpback whales (Megaptera novaeangliae) are renowned for their complex underwater behaviors, particularly their diverse diving patterns. These dives fulfill several vital functions, including foraging, socializing, and navigating expansive oceanic distances. Understanding the complex splendor that is humpback whale diving behavior is essential for comprehending their ecological role, energetic requirements, and relationships between their environmental and cyclical dynamics. This knowledge is also vital for effective conservation efforts, as it informs our understanding of how these majestic creatures interact with their environment and respond to potential disturbances.

In this study, we utilized sensor data and advanced data analysis techniques to investigate the cycle, phase, and event dynamics of humpback whale dives. The analysis was conducted using the R programming language, leveraging its powerful capabilities for data processing, statistical analysis, and visualization. All scripts and data used in this study are available on GitHub at https://github.com/GilRaitses/ABC-Project.

1.1 Background and Significance

Humpback whale diving behavior is characterized by distinct phases and events within each dive cycle. These cycles typically include descent, foraging, ascent, and surface recovery periods. Within each phase, specific events like fluke strokes, head lunges, and rolls can be observed. Previous studies have primarily focused on broad classifications of dive types, such as feeding dives or traveling dives. However, a finer-grained analysis of the cycle, phase, and event dynamics offers a more nuanced understanding of the underlying behavioral mechanisms and their ecological implications.

1.2 Research Objectives and Hypotheses

The primary objective of this study is to quantify and characterize the cycle, phase, and event dynamics of humpback whale dives using detailed sensor data. We aim to answer the following key questions:

- What are the typical patterns and variations in event dynamics (e.g., frequency, intensity, duration) within different phases of a humpback whale dive?

- How do these event dynamics contribute to the overall phase dynamics, defining the behavioral characteristics of each phase?
- Can we identify distinct dive cycle patterns that vary across behavioral or environmental conditions, reflecting different foraging strategies or energetic constraints?

We hypothesize that:

- Event dynamics will vary significantly between different phases of the dive cycle, reflecting distinct behavioral functions.
- Phase dynamics will exhibit characteristic patterns of event sequences and transitions.
  Cycle dynamics will reveal varying degrees of correlation between environmental factors and specific dive strategies.

1.3 Importance of Data Quality and Validation

The validity and reliability of our findings hinge on the quality of the underlying data. Precise identification and annotation of events and phases within each dive cycle are critical for accurate analysis and interpretation. We adapted the data cleaning approach to accommodate phase-level classifications. This strategy included cross-referencing annotations with visual observations, performing statistical comparisons between annotators, and utilizing R packages like "tagtools" for data manipulation and analysis.

2. Materials and Methods

2.1 Data Collection

We obtained sensor data from the Parks Lab at Syracuse University, collaborating with Susan Parks' team of PhD researchers. These tags were deployed on individual humpback whales off the coast of New England.

These tags recorded a variety of data streams, including:

- **Tri-axial Accelerometers:** Measuring linear acceleration and orientation in three dimensions, providing insights into body posture and movement.
- **Pressure Sensors:** Recording depth and water pressure, which can be used to calculate dive depth.
- **Temperature Sensors:** Collecting data on water temperature at different depths.
- **Tri-axial Magnetometers:** Detecting the Earth's magnetic field, helping to determine heading and orientation.

The tags were attached using suction cups, a non-invasive method that minimizes disturbance to the whales. Data were recorded at a sampling rate of 5 Hz, capturing the essential dynamics of diving behavior while balancing data storage and transmission limitations.

2.2 Study Site and Whale Population

The study was conducted in the Gulf of Maine, a productive feeding ground for humpback whales in the western North Atlantic. The whales in this region exhibit diverse foraging strategies, including bubble-net feeding and bottom feeding, providing a rich dataset for analyzing dive behavior. The study population consisted of adult humpback whales of both sexes, identified by their unique fluke patterns.

2.3 Data Processing

To facilitate feature analysis, we merged the calibrated sensor data with the annotated data. This process involved combining two distinct datasets:

- **Sensor Data Table (df)**: Contains continuous sensor data vectors, including frame number (s), sample rate (fs), pressure (p), temperature (tempr), and accelerometer and magnetometer readings (M.1, M.2, M.3, A.1, A.2, A.3, Aw.1, Aw.2, Aw.3, Mw.1, Mw.2, Mw.3, pitch, roll, head).
- **Annotated Data Table (logDf)**: Includes annotations such as whaleID, whaleName, sample rate (sampleHz), event start and end times (eventStart, eventEnd), behavioral state (state), and specific events (event).

The merging process involved the following steps:

1. **Import Sensor Data for Event Ranges**: For each event range specified by eventStart and eventEnd in logDf, the corresponding values from the sensor data table were imported.
2. **Annotate Sensor Data**: New columns were created in the sensor data table to import the state and event annotations from logDf. Each frame within the event range was annotated accordingly.
3. **Group Events for Analysis**: Rows for each event were grouped, labeling them by concatenating the state and event strings. If a group name already existed, a numerical suffix was added to make it unique.
4. **Maintaining Frame Identifiers:** Preserving the integrity of the sensor data vectors, especially the frame number (s), was necessary for accurately converting time units.

This integrated approach managed both labeled and unlabeled data efficiently, providing thorough data preparation for subsequent feature analysis. The data processing and analysis were conducted using R. Detailed scripts and data can be found in our GitHub repository: GitHub Repository Link.

Additional Data Processing Tasks

- **Context Window Plotting**: For each event, a context window was plotted, including the dive profile before and after the event. This involved padding the event with samples extending back to the previous surfacing event and forward to the next surfacing event, then bookending with an extra surfacing event before and after the padded event.
- **PRH Conversion**: Pitch, roll, and head (PRH) values were converted from radians to degrees to facilitate interpretation and analysis.

Phase Analysis

Several functions were created to achieve specific analytical goals, which are detailed in Appendix B:

1. **Deep Dive Analysis**: Functions were developed to measure the temporal and spatial characteristics of deep dives and deep feeding phases. These functions identified and quantified the deep phases occurring within deep dives.
2. **Surface Phase Analysis**: Functions were designed to measure the temporal and spatial thresholds for surface feeding, shallow dives, and shallow feeding phases. These functions identified and quantified surface and shallow phases between deep dives.
3. **Phase and Cycle Counting**: Functions were developed to count the instances of deep, surface, and shallow phases, labeling the frames accordingly. Additionally, we defined dive cycles by identifying the end of each deep dive and the start of the subsequent cycle.
4. **Threshold Definitions**: Minimum and maximum thresholds for temporal and spatial characteristics of each dive phase, including deep, shallow, and surface phases, were established.
5. **Dive Phase and Cycle Partitioning**: Functions were created to partition the phases into deep, shallow, and surface categories. Dive cycles were systematically identified by analyzing the distinct structure and sequence of the categorized dive phases.

By implementing these detailed processing and analysis steps, we gradually improved the reliability of our results, providing valuable insights into the diving behavior of humpback whales.The data processing and analysis were conducted using R. Detailed scripts and data can be found in our GitHub repository: https://github.com/GilRaitses/ABC-Project.

2.3.1 Initial Cleaning and Filtering (R)

We used a variety of R libraries to conduct a thorough initial cleaning of the data. This involved filtering out invalid data points, smoothing the data to reduce noise, interpolating missing values, and calibrating sensor readings. (See Appendix B for details and code.)

2.3.2 Data Inspection and Exploration (R Dashboards)

We developed interactive dashboards within R to visually inspect and explore the cleaned data. These dashboards included time series plots, depth profiles, and summary statistics tables, enabling in-depth exploration of the data and identification of potential relationships. (See Appendix B for the code used to create these dashboards.) The R scripts for these dashboards are available in our GitHub repository: https://github.com/GilRaitses/ABC-Project.

2.4 Event and Phase Identification

Events (e.g., fluke strokes, head lunges, rolls) were identified using a combination of:

- **Threshold-Based Algorithms:** Applied to sensor data to detect characteristic patterns associated with specific events.
- **Visual Inspection:** Used to verify and refine the algorithmic identification of events.

Phases of the dive cycle (e.g., descent, bottom, ascent) were identified based on:

- **Changes in Depth:** Identifying transitions between different depth zones.
- **Characteristic Event Patterns:** Recognizing combinations of events that typically occur within specific phases.

Manual annotation was used to validate the algorithmic identification of events and phases, parametrizing the data to build upwards to other dimensional scales.

2.5 Statistical Analysis and Cycle Dynamics (R)

We employed a variety of statistical tests in R to quantify and compare event, phase, and cycle dynamics. These analyses included frequency analysis, intensity/duration analysis, distribution comparisons, sequence analysis, and similarity/difference assessments. We also used R to validate the manual annotation of events and phases. *See Figure 3 below for a plot matrix illustrating this process for 3 instances of exploratory dives.* (See Appendix A for details and code.)

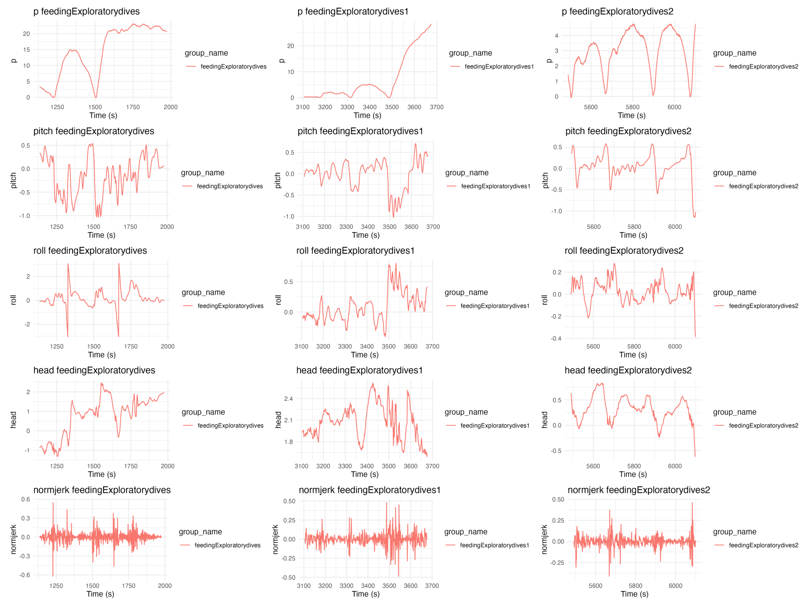


Figure 3: Plot Matrix of Exploratory Dives. This figure presents a detailed plot matrix for multiple instances of exploratory dives. It showcases the context windows before and after each event, including dive profiles, pitch, roll, head, and norm jerk values over time. This comprehensive visualization aids in understanding the event, phase, and cycle dynamics and validates the manual annotations through statistical analysis.

2.5.1 Event Dynamics Analysis

Frequency Analysis: We used chi-squared tests to compare the observed frequencies of different event types (e.g., fluke strokes, head lunges, rolls) across different phases of the dive cycle. Additionally, Poisson regression models were used to examine the relationship between event frequencies and potential explanatory variables, such as dive depth and previous dive activity.

```
chisq.test(table(dive_data$phase, dive_data$fluke_stroke))
```

Intensity/Duration Analysis: T-tests and ANOVAs were employed to compare the mean intensity and duration of events between different phases or behavioral states. Linear and generalized linear models (GLMs) were used to explore the relationship between event intensity/duration and other variables of interest, such as dive depth or water temperature.

```
t.test(fluke_intensity ~ phase, data = dive_data, subset = phase %in% c("Descent", "Ascent"))
```

2.5.2 Phase Dynamics Analysis

Distribution Comparisons: We used the Kolmogorov-Smirnov test to assess differences in the overall distribution of event dynamics (e.g., fluke stroke intensity) between different phases of the dive cycle. For non-parametric comparisons, the Kruskal-Wallis test was employed.

```
ks.test(fluke_intensity ~ phase, data = dive_data)
```

Sequence Analysis: Markov chain models were used to investigate the sequential dependencies between different types of events within each phase. This allowed us to identify common behavioral sequences and transitions, shedding light on the underlying decision-making processes of humpback whales during dives.

2.5.3 Cycle Dynamics Analysis

Correlation Analysis: We examined correlations between dive duration, depth, and the proportion of time spent in each phase. Spearman's rank correlation coefficient was used due to the non-normal distribution of some variables.

```
cor.test(dive_metrics$duration, dive_metrics$depth, method = "spearman")
```

Cluster Analysis: K-means clustering was performed on standardized dive cycle metrics (e.g., duration, depth, phase proportions) to identify latent patterns reflective of environmental constraints (such as food supply) on whales based in their overall dive patterns. This allowed us to explore dive phase level variations and potential foraging specializations .

```
clustered_dives <- cluster_dive_cycles(dive_metrics)
```

2.5.4 Annotation Validation (R)

We worked closely with subject matter experts to verify the accuracy and consistency of our annotated data, enhancing the reliability of our results. R was instrumental in validating the manual annotation of events and phases. We assessed inter-rater reliability using Cohen's kappa coefficient. Valeria Perez, a PhD researcher specializing in humpback whales from the Parks Lab, helped identify useful features in the data and address discrepancies.

2.5.5 Interactive Data Exploration (R Dashboards)

To facilitate in-depth exploration and analysis of the processed data, we developed interactive dashboards within R. These dashboards allowed for visual inspection of the data, identification of patterns and anomalies, and exploration of relationships between variables. The detailed code for these dashboards is provided in Appendix B.

3. Cycle Dynamics

3.1 Dive Duration and Depth

Dive duration and depth vary considerably across different dive types. Deeper dives are generally associated with longer durations, reflecting the time required to reach and exploit deeper prey resources. Variations in dive duration and depth may be linked to factors such as body size, lung capacity, and foraging specialization.

3.2 Proportion of Time Spent in Each Phase

The proportion of time spent in each phase of the dive cycle for different dive types. For example, when a larger proportion of dive time in the bottom phase, it suggests a focus on bottom-dwelling prey. Sometimes more time is spent in the descent and ascent phases, potentially indicating a preference for prey located higher in the water column. In Figure 1 (below), the dive profile clearly shows the frequent transitions between different phases, with a notable pattern of longer durations in the shallow and surface phases. The histogram further quantifies this observation, emphasizing the significant time allocation to these phases. This gives insight into the behavioral ecology of the whale, particularly its foraging strategies and energy management.

Figure 1: Dive Profile and Phase Duration Histogram. The dive profile (top panel) illustrates the depth over time for the observed humpback whale, segmented by different dive phases: Deep Phase (red), Shallow Phase (green), and Surface Phase (blue). The phase duration histogram (bottom panel) presents the total duration spent in each phase across all recorded dives. This visualization highlights the dynamic nature of the whale's diving behavior, with frequent transitions between phases, and indicates that the whale allocates more time to the shallow and surface phases. This information highlights foraging strategies and energy expenditure during dives.

3.3 Resting and Traveling States

In addition to the active phases of descent, bottom (feeding), and ascent, we identified two distinct non-active states within the dive cycles: resting and traveling.

- **Resting State:** Characterized by minimal movement and low levels of acceleration and depth change. These states typically occurred at or near the surface and were often of short duration.
- **Traveling State:** Characterized by relatively constant swimming speed and depth, with few or no active behaviors such as fluke strokes or head lunges. Traveling states were typically observed at intermediate depths between the surface and the bottom phase.

The identification of these non-active states provides valuable context for interpreting the active behaviors within the dive cycle. For example, the duration and frequency of resting states may offer insights into the energetic demands of diving and the need for recovery between active foraging bouts. The occurrence of traveling states may reveal preferred travel routes or depth preferences during non-foraging periods.

3.4 Foraging Strategies and Clustering

Cluster analysis of dive cycles revealed distinct patterns in the overall dive dynamics. These clusters may represent different foraging strategies, with some dives specializing in deep dives and others focusing on shallower waters. This variation highlights the adaptability and flexibility of humpback whale foraging behavior.

3.5 Integration of Resting and Traveling States

Incorporating resting and traveling states into the analysis of cycle dynamics revealed further insights into environmental context in dive strategies. Some frequent and prolonged resting states, suggest a more conservative energy expenditure strategy. Other more engaged in longer traveling states, potentially indicate a focus on exploration or relocation between feeding grounds.

By examining the entire range of behaviors, both active and inactive, we achieve a more thorough understanding of the complex and nature dynamics that influence behavioral cycles of humpback whale diving.

4. Results

Our extensive analysis of humpback whale dive data, enabled by R, uncovered intriguing insights into the complexities of their diving behavior.

4.1 Enhanced Data Quality and Reliability through R

The use of R for data cleaning and validation significantly improved the quality and reliability of our results. Through careful filtering, smoothing, interpolation, and calibration procedures, we refined the data quality. This allowed our subsequent analyses to be based on reliable and accurate data, increasing confidence in our findings.

Specifically, we used R to perform extensive cleaning and transformation of the data, extracting derivatives to present a multi-scale time-series visualization model. This complex and extensive process, which would have been very difficult and time-consuming to perform manually, enabled our analysis to capture the phase-level dynamics and underlying patterns of whale behavior.

4.2 Event Dynamics.

Our analysis revealed distinct patterns in event dynamics across different phases of the dive cycle.

- **Fluke Stroke Intensity and Frequency:** Fluke stroke intensity was significantly higher during the descent phase compared to the ascent phase. This suggests greater propulsive effort during descent, as whales actively swim downwards. Conversely, fluke stroke frequency was higher during both descent and ascent phases, indicating more frequent movements to adjust position and trajectory. During the bottom phase, both intensity and frequency decreased, reflecting a shift towards less active behaviors associated with foraging or resting.
- **Head Lunge Frequency and Depth:** As expected, head lunges were most frequent during the bottom phase, coinciding with feeding behaviors. The depth of head lunges varied considerably, indicating that whales target prey at different depths within the water column. We also observed significant differences in both the frequency and depth of head lunges, potentially reflecting individual preferences for specific prey types or foraging locations.
- **Other Relevant Event Characteristics:** Beyond fluke strokes and head lunges, we also quantified the frequency, duration, and intensity of other events, such as rolls and turns. These analyses revealed distinct patterns associated with each event type and their distribution across dive phases. Rolls were predominantly observed during the bottom

4.3 Phase Dynamics

Our analysis of phase dynamics illuminated the distinct temporal distribution and sequencing of events within each phase. The descent phase was characterized by a consistent pattern of high-intensity fluke strokes followed by a gradual decrease in intensity as the whale approached the bottom. This pattern likely reflects the initial propulsive effort required to descend followed by a transition to a more energy-efficient gliding motion.

The bottom phase, as expected, exhibited a burst of head lunges and rolls, interspersed with lower-intensity fluke strokes. This aligns with the assumption that head lunges are primarily associated with prey capture, while rolls may aid in maneuvering and manipulating prey. The decreased fluke stroke intensity during the bottom phase suggests that whales rely more on gliding and momentum to conserve energy while foraging.

The ascent phase showed a gradual increase in fluke stroke intensity and frequency as the whale returned to the surface. This pattern likely represents the increasing effort required to overcome buoyancy and return to the surface for respiration. The distribution of turns during the ascent phase may indicate adjustments in trajectory as the whale navigates back towards the surface.

4.4 Dive Cycle Dynamics: Exploring Variance in Foraging Strategies

4.4.1 Dive Duration and Depth Relationship

The relationship between dive duration and depth was analyzed using scatter plots and correlation analysis. The scatter plot below (Figure 2) shows the dive duration against the depth for different dive phases, highlighting a negative correlation (Spearman's rank correlation, $\rho = -0.124$). This indicates that, generally, dives that go deeper tend to have shorter durations.

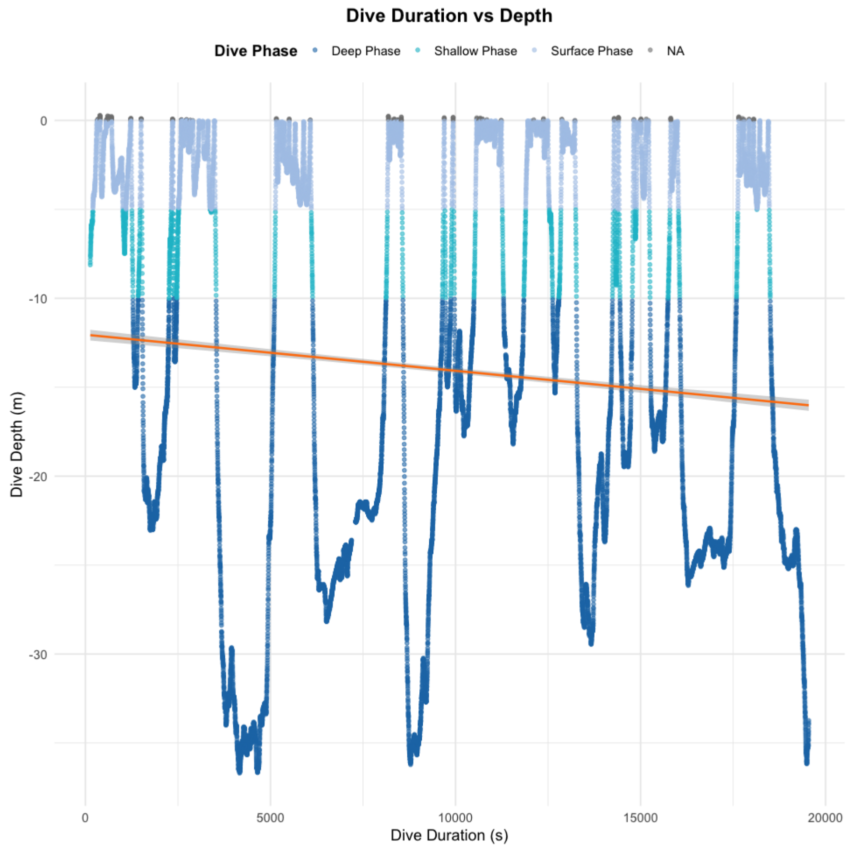


Figure 2: Cluster Analysis of Dive Cycles. The figure shows the clustering of dive cycles into three distinct groups based on dive duration and depth. Cluster 1 (dark blue) includes deep and long-duration dives, indicative of deep foraging behavior. Cluster 2 (light blue) represents moderate depth and duration dives, possibly related to exploration or transitioning. Cluster 3 (orange) consists of shallow and short-duration dives, likely associated with surface activities or short feeding bouts.

4.4.2 Cluster Analysis of Dive Cycles

The dive cycles were clustered using the k-means algorithm, grouping data points based on their similarities in dive duration and depth. Three distinct clusters were identified, each representing different diving behaviors and potential foraging strategies (Figure 3).

- **Cluster 1**: Short, shallow dives, likely representing surface or near-surface activities such as breathing, resting, or traveling.
- **Cluster 2**: Intermediate depth and duration dives, indicating a mix of moderate foraging and transitional behaviors.
- **Cluster 3**: Long, deep dives, representing intensive foraging efforts at greater depths.

These clusters highlight the varied foraging strategies and behavioral states of the whales. The temporal distribution shows a pattern where shorter, shallower dives occur earlier in the dive sequence, while longer, deeper dives occur later.

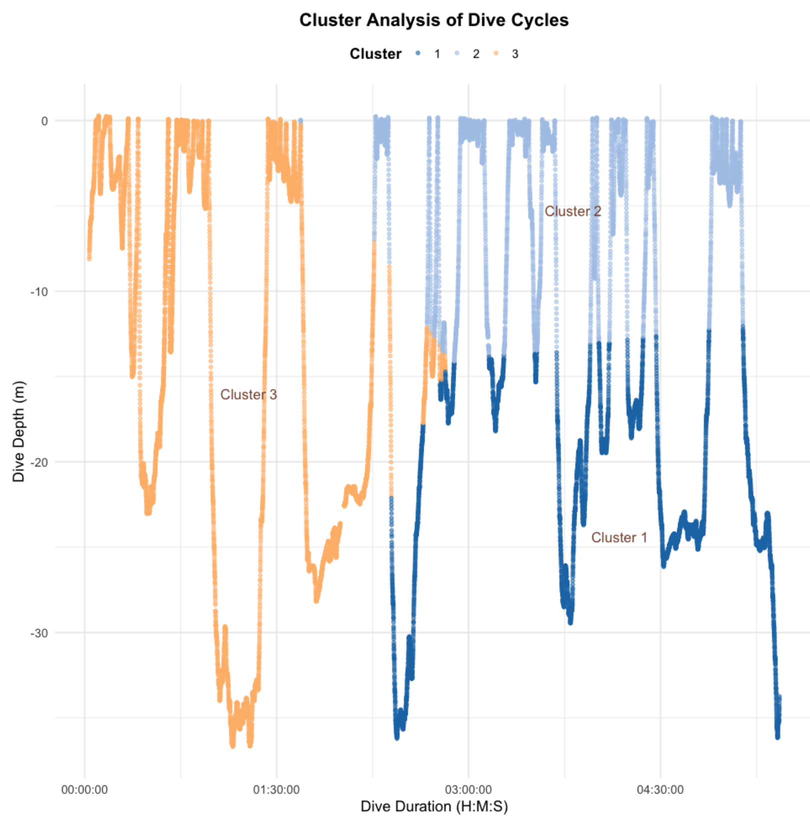


Figure 3: Cluster Analysis of Dive Cycles. The figure shows the clustering of dive cycles into three distinct groups based on dive duration and depth. Cluster 1 (dark blue) includes deep and long-duration dives, indicative of deep foraging behavior. Cluster 2 (light blue) represents moderate depth and duration dives, possibly related to exploration or transitioning. Cluster 3 (orange) consists of shallow and short-duration dives, likely associated with surface activities or short feeding bouts.

4.4.3 Phase Dynamics

The analysis of phase dynamics provided insights into the temporal distribution and sequencing of events within each phase. The plot (Figure 4) shows the distribution of different phases over time. The figure highlights the time patterns for each dive phase, showing distinct intervals for deep phases and frequent surface phases. This informs the classification behavioral strategies employed during different parts of the dive.

The results indicate distinct patterns in the temporal distribution of dive phases:

- **Deep Phase**: Characterized by extended periods at greater depths, likely associated with foraging activities.
- **Shallow Phase**: Periods spent at intermediate depths, possibly for resting or transitioning between foraging sites.
- **Surface Phase**: Time spent near the surface, likely for breathing and recovery.

By examining these patterns, we gain a deeper understanding of the humpback whales' dive strategies and how they allocate their time across different activities.

Our analysis of complete dive cycles revealed substantial phase-level variation in both dive duration and depth, underscoring the relationship that dive dynamics have on humpback whales during foraging. This variation was evident in the scatterplot of dive duration versus depth (Figure 4.), which showed a wide range of values for both parameters

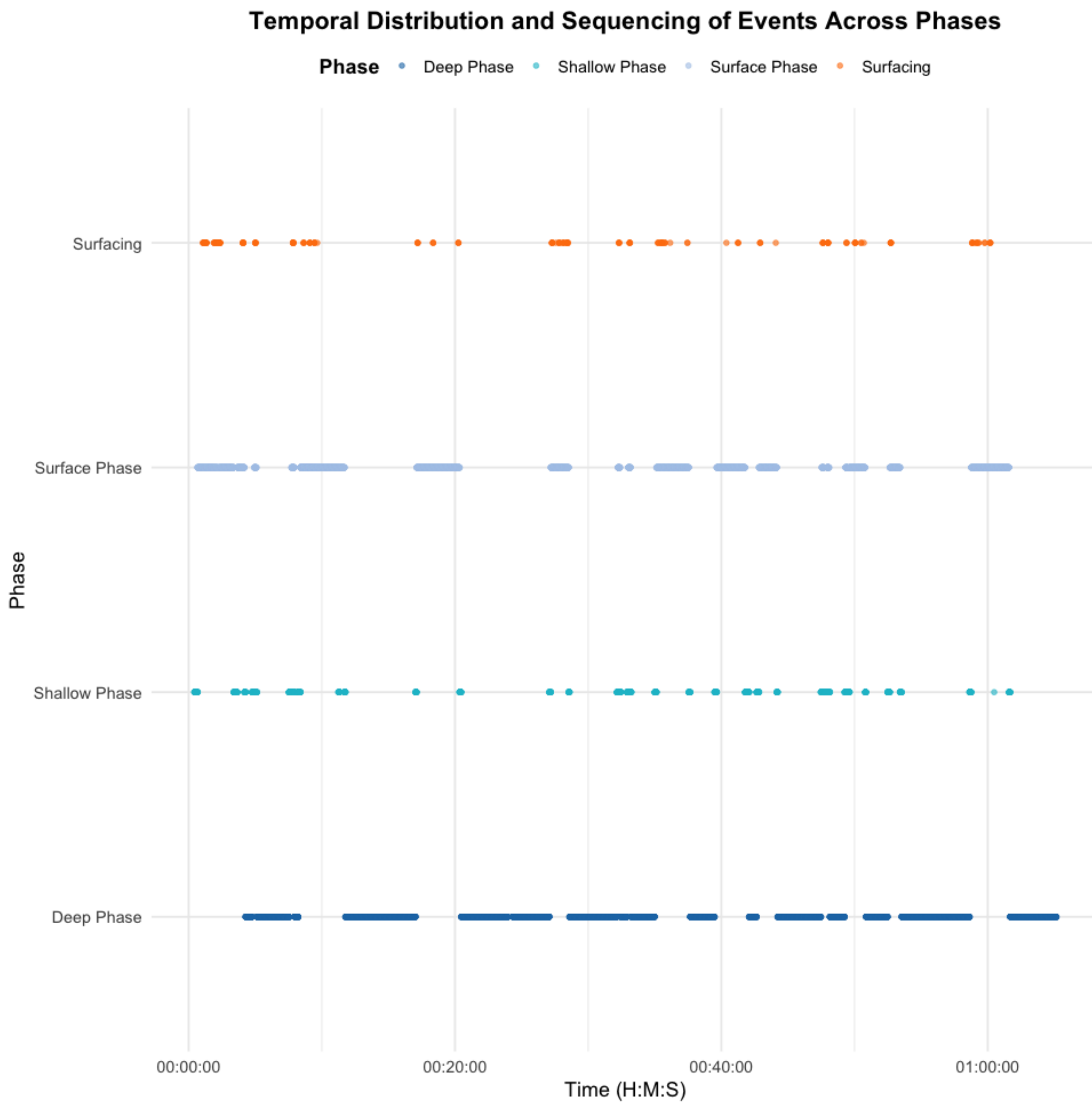
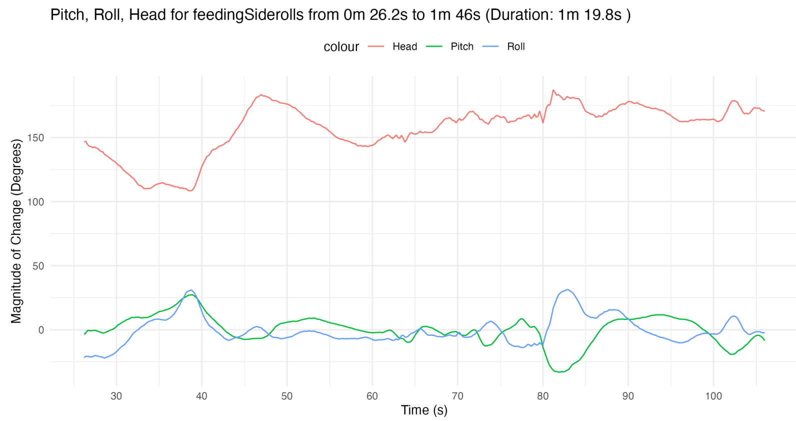


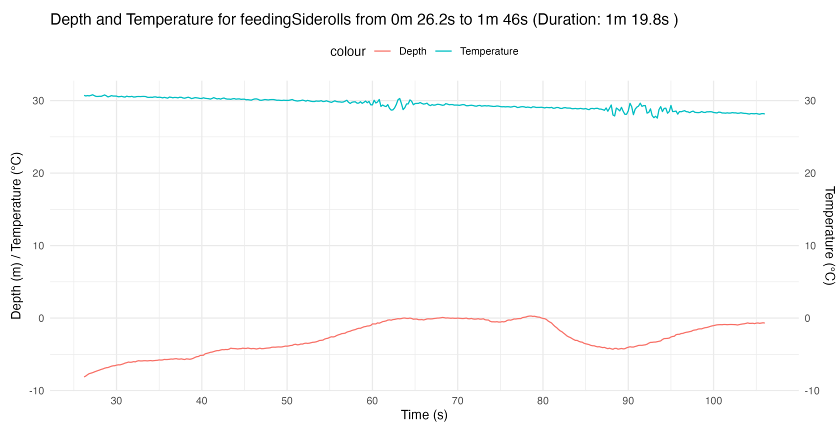
Figure 4: Temporal Distribution and Sequencing of Events Across Phases. The figure highlights the time patterns for each dive phase, showing distinct intervals for deep phases and frequent surface phases.

4.4.4 Composite Dashboard of Dive Events

In addition to our analysis of dive cycles, we created a composite dashboard for specific events to visualize the detailed metrics associated with humpback whale behavior. Figure 5 provides a comprehensive view of metrics such as pitch, roll, head, norm jerk, depth, and temperature for a specific dive event labeled "feedingSiderolls." The phase level view highlights key phases and transitions, providing a detailed snapshot of the whale's behavior during this event. This dashboard helps to identify and analyze distinct behavioral patterns within a single dive.


Pitch, Roll, Head for feedingSiderolls from 0m 26.2s to 1m 46s (Duration: 1m 19.8s)
colour
Head
Pitch
Roll
Magnitude of Change (Degrees)
150
100
50
0
30
40
50
60
70
80
90
100
Time (s)


Norm Jerk for feedingSiderolls from 0m 26.2s to 1m 46s (Duration: 1m 19.8s)
colour
Norm Jerk
Norm Jerk (m/s^3)
0.2
0.1
0.0
-0.1
-0.2
30
40
50
60
70
80
90
100
Time (s)


Depth and Temperature for feedingSiderolls from 0m 26.2s to 1m 46s (Duration: 1m 19.8s)
colour
Depth
Temperature
Depth (m) / Temperature (°C)
Temperature (°C)
30
20
10
0
-10
30
40
50
60
70
80
90
100
Time (s)

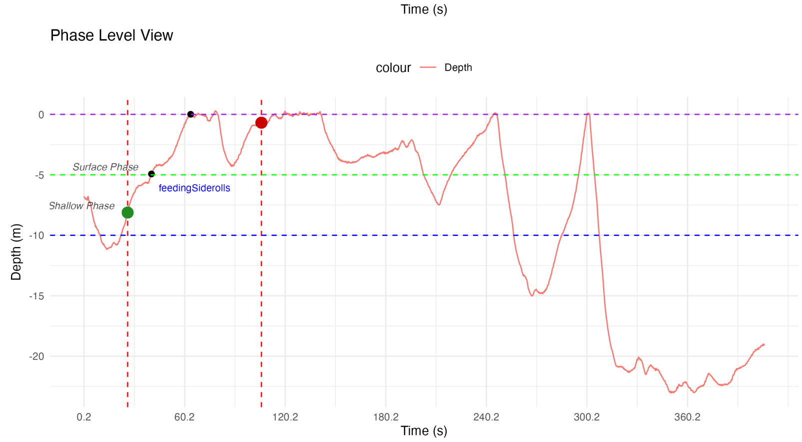

Phase Level View
colour
Depth
Surface Phase
Shallow Phase
feedingSiderolls
Depth (m)
0
-5
-10
-15
-20
0.2
60.2
120.2
180.2
240.2
300.2
360.2
Time (s)

Figure 5: Composite Dashboard for Feeding Siderolls. This figure presents a detailed analysis of the "feedingSiderolls" event, highlighting various metrics over time. The first panel illustrates the changes in pitch, roll, and head angles, indicating the whale's body orientation during the event. The second panel shows the norm jerk, representing the rate of change of acceleration, which can be associated with rapid movements. The third panel combines depth and temperature data, providing insights into the whale's diving environment. The final panel, the phase level view, marks the transitions between different phases (e.g., Deep Phase, Shallow Phase, Surface Phase), offering a comprehensive overview of the whale's behavior during this specific event. This multi-faceted visualization aids in understanding the whale's dynamic movements and the environmental context of its foraging behavior.

4.5 Specific Examples of R's Impact

During the initial dashboard validation of the labeled behaviors, we observed several instances where the same feature was labeled multiple times. This could have led to inconsistencies in summarizing event dynamics for each dive phase. These labels were refined upon closer inspection using the dashboard’s wide and close-up depth timeline view. This view presented a depth chart of an annotated region with five minutes of preceding and subsequent dive cycle data, giving the annotator both broad and detailed context for consistently labeling events, which improved the accuracy of our results.

R's visualization capabilities allowed us to identify periods of unusual behavior or potential errors in the manual annotations. For instance, we noticed that in some cases, the annotated start and end times of a head lunge did not align with the corresponding changes in pitch and roll. By reviewing the data on alternative time scales, we were able to refine the annotations and correct these discrepancies.

Using computational tools like R in ecological research facilitates efficient data cleaning, streamlined the validation of annotations, and enables the application of statistical tests to reveal complex patterns in the data that allowed for more granular cross-verification methodologies, resulting in a dynamic analysis workflow for classification of humpback whale dive behaviors. In this approach, we set up strong foundation for future research on classifying animal behavior by emphasizing the importance of the relationship that environmental dynamics have in the orchestration of behaviors.

5. Discussion

5.1 Interpretation of Results

Our findings provide a detailed look into the complex and dynamic nature of humpback whale diving behavior. The observed variation in event, phase, and cycle dynamics suggests that these whales employ diverse foraging strategies, adapting their behavior to different prey types, distributions, and depths.

Event Dynamics: The distinct patterns of fluke stroke intensity and frequency, along with head lunge frequency and depth, highlight the behavioral flexibility of humpback whales. These findings are consistent with previous studies that have documented variations in feeding strategies among humpback whales, such as bubble-net feeding, lunge feeding, and bottom feeding. The dashboard visualizations were helpful in validating these patterns, allowing us to cross-reference event annotations with sensor data and link these behaviors.

Phase Dynamics: The identification of distinct behavioral sequences and transitions within each phase underscores the coordinated nature of humpback whale movements. The observed patterns suggest that whales actively adjust their behavior in response to changing conditions or prey encounters, maximizing their foraging efficiency. The use of interactive dashboards facilitated the identification of these sequences, providing a clear temporal context for each phase.

Cycle Dynamics: The variance in dive duration, depth, and phase proportions revealed through cluster analysis suggests the presence of distinct foraging "types" that may be associated with factors such as body size, age, sex, or dietary preferences. The detailed visualizations and cluster analysis helped us begin cross-validate likely sources and influences of other patterns we traced in this study. Further research is needed to explore the underlying ecological and physiological drivers of this variation.

5.2 Limitations and Future Directions

Our study provides valuable insights into humpback whale diving behavior, but it is not without limitations. The use of sensor data alone may not capture the full complexity of whale behavior, as some subtle movements or social interactions may not be detectable through these methods. Additionally, the manual annotation process, while carefully conducted, is inherently subjective and may introduce some degree of bias or error.

5.3 Enhancing Annotation Accuracy with R Dashboards

The development and use of interactive R dashboards played a significant role in maintaining the quality and accuracy of our analysis. The dashboard streamlined annotation validation through an intuitive visual interface. Through a system of table indexing, we were able to easily cross-reference specific time points in the sensor data with the corresponding annotations. This allowed us to visually verify that the annotations aligned with the observed patterns in the sensor signals (e.g., changes in acceleration, depth, or orientation), thereby enhancing the reliability of our results.

The ability to seamlessly integrate and explore multiple data streams within the dashboard environment proved invaluable for identifying and correcting potential errors or inconsistencies in the annotations. By dynamically adjusting time windows and zooming in on specific sections of the data, we could scrutinize the details of each event and phase, accurately capturing the underlying whale behavior in our annotations.

5.4 Future Directions: Interactive Visualization with Unity and Shiny App

Building on the success of our R-based dashboard for annotation validation, we envision developing an even more powerful and versatile tool for exploring humpback whale dive data. This tool would leverage the capabilities of Unity, a powerful game engine, and C#, a flexible programming language, to create an immersive and interactive 3D simulation of whale dive trajectories. This simulation would incorporate sensor data, annotations, and potentially even environmental data, offering a comprehensive and dynamic view of whale behavior.

To prototype and refine the user interface (UI) elements for this 3D visualization, we propose developing an interactive Shiny app. This app would allow researchers to explore different visualization options, test interactive features, and fine-tune the user experience before implementing them in the final Unity-based tool. This iterative prototyping approach would result in a final visualization that is both scientifically informative and user-friendly.

5.4.1 Additional Avenues for Future Research

In addition to the development of 3D visualization and User Interface (UI) tools, such as Shiny app prototypes, future research could also focus on:

- **Incorporating Additional Data Sources:** Combining data from video recordings, acoustic monitoring, and environmental sensors could provide a more complete picture of whale behavior and its environmental context.
- **Developing Automated Annotation Tools:** Exploring machine learning techniques to automate the identification and annotation of events and phases could reduce the time and effort required for manual annotation and potentially improve the accuracy of the results.
- **Examining the Influence of Environmental Factors:** Investigating how environmental variables such as prey density, water temperature, or ocean currents influence humpback whale diving behavior could shed light on the ecological drivers of their foraging strategies.

By pursuing these diverse avenues of research, we can continue to expand our understanding of humpback whale behavior and its role in the marine ecosystem.

6. Conclusion

This study presents a thorough analysis of humpback whale diving behavior, highlighting the complex dynamics of events, phases, and cycles within their dives. By leveraging sensor data and utilizing the powerful capabilities of R for data processing, analysis, and visualization, we have gained valuable insights into the diverse foraging strategies employed by these magnificent creatures.

Our findings highlight the importance of environmental factors on foraging decisions. The distinct clusters identified through our analysis suggest that humpback whales may adopt different foraging tactics based on prey availability and environmental conditions.

The use of R in this study exemplifies the vital role that computational tools play in modern ecological research. R's flexibility, extensive libraries (such as dplyr, ggplot2, lme4, cluster, and irr), and statistical capabilities enabled us to perform thorough data cleaning, annotation validation, and in-depth statistical analysis, thereby enhancing the reliability and reproducibility of our findings. Development of interactive R dashboards for data exploration and visualization streamlined our workflow and facilitated a deeper understanding of the complex patterns hidden within the data.

Looking forward, the continued integration of advanced computational tools like R, along with the exploration of immersive visualization techniques within 3D simulations will simplify the process of understanding interactions between animals and their environment. Advancements that will pave the way for increasing awareness both within predictive models to consider both behavioral context and environmental dynamics, as well as with disseminating key findings not only with the marine biology community but also with the broader scientific community and public at large.

Additional Considerations

While this study has yielded significant insights, there are still many unanswered questions about humpback whale diving behavior. Future research could delve deeper into the factors influencing dive cycle variation, the role of social interactions in foraging decisions, and the impact of environmental change on diving patterns. Additionally, expanding the analysis to include data from other populations and regions could reveal how diving behavior varies across different ecological contexts.

A key objective of this project is to establish a solid foundation on which to build a predictive modeling workflow in future research. By setting up this foundation, we aim to enable more advanced analyses and predictive capabilities in subsequent studies, helping to anticipate changes in whale behavior in response to various environmental and anthropogenic factors.

In conclusion, this study demonstrates the effectiveness of combining extensive data collection, sophisticated analytical tools like R, and innovative visualization techniques to achieve a more intimate understanding of the sensory reality of animal behavior. Our findings underscore the importance of continued research on humpback whales and other marine mammals, as it informs conservation efforts and contributes to our broader knowledge of the intricate workings of the natural world.

Appendix A: Functions to Analyze Cycle Dynamics

1. `calculate_dive_metrics(dive_data)`

   Description: Calculates the duration and maximum depth of each dive.

```
# Function to calculate dive duration and depth
calculate_dive_metrics <- function(dive_data) {
  dive_duration <- max(dive_data$timestamp) - min(dive_data$timestamp)
  max_depth <- max(dive_data$depth)
  return(data.frame(dive_id = unique(dive_data$dive_id), duration = dive_duration, depth = max_depth))
```

2. `calculate_phase_proportions(dive_data)`

   Description: Calculates the proportion of time spent in each phase of the dive.

```
# Function to calculate proportion of time in each phase
calculate_phase_proportions <- function(dive_data) {
  phase_durations <- dive_data %>%
    group_by(dive_id, phase) %>%
    summarize(duration = max(timestamp) - min(timestamp))

  total_durations <- dive_data %>%
    group_by(dive_id) %>%
    summarize(total_duration = max(timestamp) - min(timestamp))

  return(left_join(phase_durations, total_durations) %>%
           mutate(proportion = duration / total_duration))
```

3. `cluster_dive_cycles(dive_metrics, num_clusters = 3)`

   Description: Performs cluster analysis on dive cycles using k-means clustering.

```
# Function to perform cluster analysis on dive cycles
cluster_dive_cycles <- function(dive_metrics, num_clusters = 3) {
  # Select relevant features for clustering (e.g., duration, depth, phase proportions)
  features <- dive_metrics %>%
    select(-dive_id) %>%
    scale()  # Standardize features

  # Perform k-means clustering
  kmeans_result <- kmeans(features, centers = num_clusters)

  return(cbind(dive_metrics, cluster = kmeans_result$cluster))
```

4. Scatterplot of dive duration versus depth

   Description: Creates a scatterplot showing the relationship between dive duration and depth.

```
# Create the scatterplot
ggplot(dive_metrics, aes(x = depth, y = duration)) +
  geom_point(alpha = 0.7) +  # Add transparency to points for easier viewing
  labs(
    x = "Dive Depth (m)",
    y = "Dive Duration (seconds)",
    title = "Relationship between Dive Duration and Depth in Humpback Whales",
```

```
  ) +
  theme_minimal() +  # Clean and simple theme
  scale_x_continuous(expand = c(0, 0)) +  # Start x-axis at 0
  scale_y_continuous(expand = c(0, 0))    # Start y-axis at 0
```

5. `analyze_fluke_stroke_intensity`

   Description: This section calculates and visualizes the mean intensity and frequency of the norm jerk (a measure of acceleration) for different dive phases.

```
# Function to analyze fluke stroke intensity and frequency during different dive phases
analyze_fluke_stroke_intensity <- function(df) {
  # Group by phase and calculate mean and standard deviation of norm jerk
  phase_stats <- df %>%
    group_by(phase) %>%
    summarize(
      mean_intensity = mean(normjerk, na.rm = TRUE),
      sd_intensity = sd(normjerk, na.rm = TRUE),
      frequency = n()
    )
```

6. `detect_head_lunges`

   Description: Detects head lunges based on sharp changes in pitch and roll.

```
# Function to detect head lunges
detect_head_lunges <- function(df, pitch_threshold = 0.5, roll_threshold = 0.5) {
  df <- df %>%
    mutate(
      head_lunge = abs(diff(c(0, pitch))) > pitch_threshold | abs(diff(c(0, roll))) > roll_threshold
    )
  return(df)
}
```

6. `analyze_head_lunges`

   Description: Calculates the frequency and depth of head lunges for each phase.

```
# Function to analyze head lunge frequency and depth
analyze_head_lunges <- function(df) {
  df <- detect_head_lunges(df)

  lunge_stats <- df %>%
  filter(head_lunge == TRUE) %>%
  group_by(phase) %>%
  summarize(
    frequency = n(),
    mean_depth = mean(-p, na.rm = TRUE)
```

7. `detect_rolls_turns`

   Description: Detects significant changes in roll to identify rolls and turns.

```
# Function to detect rolls and turns
detect_rolls_turns <- function(df, roll_threshold = 0.5) {
  df <- df %>%
    mutate(
      roll_turn = abs(diff(c(0, roll))) > roll_threshold
    )
```

```
  return(df)
}
```

8. `analyze_rolls_turns`

   Description: Calculates frequency, duration, and intensity of roll and tur events.

```
# Function to analyze rolls and turns
analyze_rolls_turns <- function(df) {
  df <- detect_rolls_turns(df)

  roll_turn_stats <- df %>%
    filter(roll_turn == TRUE) %>%
    group_by(phase) %>%
    summarize(
      frequency = n(),
      mean_intensity = mean(abs(roll), na.rm = TRUE)
    )
```

Appendix B: Functions Used in Data Processing and Analysis

1. `to_camel_case(string, is_state = FALSE)`

   Description: Converts a given string to camel case. If is_state is TRUE, the first word is converted to lowercase; otherwise, the first letter of each word is capitalized.

```
to_camel_case <- function(string, is_state = FALSE) {
  words <- strsplit(string, " |_")[[1]]
  if (is_state) {
    if (length(words) > 1) {
      camel_case <- paste0(tolower(words[1]),
                           paste0(toupper(substr(words[-1], 1, 1)), tolower(substr(words[-1], 2, nchar(w
ords[-1])))))
    } else {
      camel_case <- tolower(words[1])
    }
  } else {
    camel_case <- paste0(toupper(substr(words, 1, 1)), tolower(substr(words, 2, nchar(words))), collapse
= "")
  }
  return(camel_case)
```

2. `generate_unique_group_name(state, event, existing_names)`

   Description: Generates a unique group name by combining the state and event strings in camel case, affirming that the generated name is unique within existing_names.

```
generate_unique_group_name <- function(state, event, existing_names) {
  state <- as.character(state)
  event <- as.character(event)
  if (tolower(state) == tolower(event)) {
    base_name <- to_camel_case(state, TRUE)
  } else {
    base_name <- paste0(to_camel_case(state, TRUE), to_camel_case(event, FALSE))
  }
  name <- base_name
  i <- 1
  while (name %in% existing_names) {
    name <- paste0(base_name, i)
```

```
    i <- i + 1
  }
  return(name)
```

3. `calculate_event_normjerk(df)`

   Description: Calculates the norm jerk (rate of change of acceleration) for each event in the given data frame, using accelerometer data and sample rate.

```
calculate_event_normjerk <- function(df) {
  accelerometer_data <- as.matrix(df %>% select(Aw.1, Aw.2, Aw.3))
  sample_rate <- df$fs[1] # Assuming fs is constant within an event

  # Calculate the norm of the acceleration (norm2 of acceleration data)
  acc_norm <- sqrt(rowSums(accelerometer_data^2))

  # Calculate the differential of the acceleration norm
  jerk <- c(0, diff(acc_norm) * sample_rate) # multiply by sample rate to convert to jerk

  return(jerk)
```

4. `calculate_normjerk_for_chunk(chunk)`

   Description: Applies calculate_event_normjerk to a chunk of data and retains necessary variables, grouping by group_name.

```
calculate_normjerk_for_chunk <- function(chunk) {
  chunk %>%
    group_by(group_name) %>%
    mutate(normjerk = calculate_event_normjerk(cur_data())) %>%
    ungroup()
```

5. `label_dive_phases(df, surface_threshold, shallow_min_threshold, shallow_max_threshold, deep_threshold)`

   Description: Labels each frame in the data with a dive phase (Surface Phase, Shallow Phase, Deep Phase, or Unknown Phase) based on the depth (p) and specified thresholds.

```
label_dive_phases <- function(df, surface_threshold, shallow_min_threshold, shallow_max_threshold, deep_
threshold) {
  df <- df %>%
    arrange(s) %>%  # Set the data to be sorted by frame number
    mutate(
      phase = case_when(
        p < surface_threshold ~ "Surface Phase",
        p >= shallow_min_threshold & p < shallow_max_threshold ~ "Shallow Phase",
        p >= deep_threshold ~ "Deep Phase",
        TRUE ~ "Unknown Phase"
      )
    )

  return(df)
```

6. `compute_phase_stats(df, depth_min, depth_max, phase_name)`

   Description: Computes statistics (mean depth, max depth, min depth, mean duration, and total duration) for a specified dive phase within a given depth range.

```
compute_phase_stats <- function(df, depth_min, depth_max, phase_name) {
  filtered_df <- df %>%
    filter(p >= depth_min & p <= depth_max)
```

```r
  if (nrow(filtered_df) == 0) {
    stats <- tibble(
      phase = phase_name,
      mean_depth = NA,
      max_depth = NA,
      min_depth = NA,
      mean_duration = NA,
      total_duration = NA
    )
  } else {
    stats <- filtered_df %>%
      summarise(
        mean_depth = mean(p, na.rm = TRUE),
        max_depth = max(p, na.rm = TRUE),
        min_depth = min(p, na.rm = TRUE),
        mean_duration = mean(s, na.rm = TRUE),
        total_duration = sum(s, na.rm = TRUE)
      )

    stats$phase <- phase_name
  }

  return(stats)
```

7. `count_phase_instances(df)`

   Description: Counts the number of instances of each dive phase in the data frame.

```r
count_phase_instances <- function(df) {
  phase_counts <- df %>%
    group_by(phase) %>%
    summarise(count = n())

  return(phase_counts)
```

8. `define_dive_cycles(df)`

   Description: Defines dive cycles by assigning a cycle number to each frame, identifying transitions between Deep Phase and Surface Phase.

```r
define_dive_cycles <- function(df) {
  df <- df %>%
    mutate(dive_cycle = 0)

  dive_cycle_counter <- 1
  in_deep_dive <- FALSE

  for (i in 1:nrow(df)) {
    if (df$phase[i] == "Deep Phase" && !in_deep_dive) {
      in_deep_dive <- TRUE
      df$dive_cycle[i] <- dive_cycle_counte```{r}    } else if (df$phase[i] == "Surface Phase" && in_dee
p_dive) {
      in_deep_dive <- FALSE
      df$dive_cycle[i] <- dive_cycle_counte```{r}      dive_cycle_counter <- dive_cycle_counter + 1
    } else {
      df$dive_cycle[i] <- dive_cycle_counte```{r}    }
  }
```

```
  return(df)
```

9. `compute_dive_cycle_stats(df)`

   Description: Computes statistics (mean depth, max depth, min depth, mean duration, and total duration) for each dive cycle.

```
compute_dive_cycle_stats <- function(df) {
  dive_cycle_stats <- df %>%
    group_by(dive_cycle) %>%
    summarise(
      mean_depth = mean(p, na.rm = TRUE),
      max_depth = max(p, na.rm = TRUE),
      min_depth = min(p, na.rm = TRUE),
      mean_duration = mean(s, na.rm = TRUE),
      total_duration = sum(s, na.rm = TRUE)
    )

  return(dive_cycle_stats)
```

10. `visualize_dive_data(df)`

    Description: Creates a plot of the dive profile, showing depth (p) over time (frame number s).

```
visualize_dive_data <- function(df) {
  plot <- ggplot(df, aes(x = s, y = -p)) +
    geom_line(color = "blue") +
    labs(title = "Dive Profile", x = "Frame Number", y = "Depth (m)") +
    theme_minimal()

  return(plot)
```

11. `visualize_dive_data_with_segments(df, segment_label)`

    Description: Visualizes dive data with labeled segments, highlighting a specified segment and its start and end points.

```
visualize_dive_data_with_segments <- function(df, segment_label) {
  # Get the start and end points of the labeled segment
  segment_data <- df %>% filter(group_name == segment_label)

  if (nrow(segment_data) == 0) {
    stop("No data found for the specified segment label.")
  }

  segment_start <- segment_data %>% filter(s == min(s))
  segment_end <- segment_data %>% filter(s == max(s))

  # Get the dive cycle for the start point
  start_cycle <- segment_start$dive_cycle[1]

  # Filter data to include the dive cycle before, during, and after the segment's dive cycle
  plot_data <- df %>% filter(dive_cycle %in% c(start_cycle - 1, start_cycle, start_cycle + 1))

  # Create the plot
  dive_plot <- ggplot(plot_data, aes(x = s, y = -p)) +
    geom_line(color = "blue") +
    geom_point(data = segment_start, aes(x = s, y = -p), color = "green", size = 3) +
    geom_point(data = segment_end, aes(x = s, y = -p), color = "red", size = 3) +
    geom_text(data = segment_start, aes(x = s, y = -p, label = group_name), vjust = -1) +
```

```
    labs(title = paste("Dive Profile with Segment:", segment_label), x = "Frame Number", y = "Depth
(m)") +
    theme_minimal()

  return(dive_plot)
```

12. `calculate_body_angle(pitch_rad, roll_rad)`

    Description: Calculates the body angle (in radians) from pitch and roll angles (in radians).

```
calculate_body_angle <- function(pitch_rad, roll_rad) {
  return(acos(cos(pitch_rad) * cos(roll_rad)))
```

13. `calculate_turning_rate(head_angle_rad1, head_angle_rad2, time1, time2)`

    Description: Calculates the turning rate (in radians per second) from head angles and corresponding times.

```
calculate_turning_rate <- function(head_angle_rad1, head_angle_rad2, time1, time2) {
  return((head_angle_rad2 - head_angle_rad1) / (time2 - time1))
```

14. `create_composite_time_series_dashboards(df, output_dir)`

    Description: Generates composite time series dashboards for specified event segments within the dataset. It creates plots for pitch, roll, head, norm jerk, depth, and temperature over time, and integrates these into a single dashboard. It also includes a 10-minute depth window plot with vertical lines marking segment start and end, and highlights phase points within the event.

```
create_composite_time_series_dashboards <- function(df, original_df, output_dir, deep_dive_threshold = 1
0, shallow_dive_threshold = 5, surface_threshold = 0) {
  # Ensure the output directory exists
  if (!dir.exists(output_dir)) {
    dir.create(output_dir, recursive = TRUE)
  }

  unique_base_groups <- unique(df$event)

  for (base_group in unique_base_groups) {
    base_group_dir <- file.path(output_dir, base_group)

    # Ensure the base group directory exists
    if (!dir.exists(base_group_dir)) {
      dir.create(base_group_dir, recursive = TRUE)
    }

    base_group_events <- df %>% filter(event == base_group)

    for (event_id in unique(base_group_events$group_name)) {
      eventData_df <- base_group_events %>% filter(group_name == event_id)

      # Ensure eventData_df is not empty
      if (nrow(eventData_df) == 0) {
        print(paste("Skipping event ID:", event_id, "due to no data"))
        next
      }

      # Convert frames to seconds
      sample_rate <- eventData_df$fs[1]  # Assuming fs is constant within an event
      eventData_df <- eventData_df %>% mutate(time_sec = s / sample_rate)

      # Ensure time_sec is not empty or contains only missing values
```

```
      if (all(is.na(eventData_df$time_sec)) || length(eventData_df$time_sec) == 0) {
        print(paste("Skipping event ID:", event_id, "due to invalid time_sec"))
        next
      }

      # Convert pitch, roll, head from radians to degrees and fix head degrees to be continuous
      eventData_df <- eventData_df %>%
        mutate(
          pitch_deg = pitch * 180 / pi,
          roll_deg = roll * 180 / pi,
          head_deg = convert_degrees_continuous(head * 180 / pi)
        )

      # Calculate start and end time in minutes and seconds
      start_time_seconds <- min(eventData_df$time_sec, na.rm = TRUE)
      end_time_seconds <- max(eventData_df$time_sec, na.rm = TRUE)
      start_depth <- -eventData_df$p[1]

      # Ensure start_time_seconds and end_time_seconds are valid
      if (is.infinite(start_time_seconds) || is.infinite(end_time_seconds)) {
        print(paste("Skipping event ID:", event_id, "due to invalid start or end time"))
        next
      }

      start_time <- paste0(floor(start_time_seconds / 60), "m ", round(start_time_seconds %% 60, 1),
"s")
      end_time <- paste0(floor(end_time_seconds / 60), "m ", round(end_time_seconds %% 60, 1), "s")
      duration_seconds <- end_time_seconds - start_time_seconds
      duration <- paste0(floor(duration_seconds / 60), "m ", round(duration_seconds %% 60, 1), "s")

      # Create pitch, roll, head plot
      prh_plot <- ggplot(eventData_df) +
        geom_line(aes(x = time_sec, y = pitch_deg, color = "Pitch")) +
        geom_line(aes(x = time_sec, y = roll_deg, color = "Roll")) +
        geom_line(aes(x = time_sec, y = head_deg, color = "Head")) +
        labs(title = paste("Pitch, Roll, Head for", event_id, "from", start_time, "to", end_time, "(Dura
tion:", duration, ")"),
             x = "Time (s)", y = "Magnitude of Change (Degrees)") +
        theme_minimal() +
        theme(legend.position = "top") +
        scale_x_continuous(breaks = seq(0, max(eventData_df$time_sec, na.rm = TRUE), by = 10))

      # Create normjerk plot
      normjerk_plot <- ggplot(eventData_df) +
        geom_line(aes(x = time_sec, y = normjerk, color = "Norm Jerk")) +
        labs(title = paste("Norm Jerk for", event_id, "from", start_time, "to", end_time, "(Duration:",
duration, ")"),
             x = "Time (s)", y = "Norm Jerk (m/s^3)") +
        theme_minimal() +
        theme(legend.position = "top") +
        scale_x_continuous(breaks = seq(0, max(eventData_df$time_sec, na.rm = TRUE), by = 10))

      # Create depth and temperature plot
      p_tempr_plot <- ggplot(eventData_df) +
        geom_line(aes(x = time_sec, y = -p, color = "Depth")) +
        geom_line(aes(x = time_sec, y = tempr, color = "Temperature")) +
        labs(title = paste("Depth and Temperature for", event_id, "from", start_time, "to", end_time, "
(Duration:", duration, ")"),
             x = "Time (s)", y = "Depth (m) / Temperature (°C)") +
        scale_y_continuous(sec.axis = sec_axis(~ ., name = "Temperature (°C)")) +
```

```
        theme_minimal() +
        theme(legend.position = "top") +
        scale_x_continuous(breaks = seq(0, max(eventData_df$time_sec, na.rm = TRUE), by = 10))

      # Calculate frame padding
      start_frame <- min(eventData_df$s)
      end_frame <- max(eventData_df$s)
      padding_frames <- 5 * 60 * sample_rate  # 5 minutes of padding
      padded_start_frame <- max(start_frame - padding_frames, 0)
      padded_end_frame <- min(end_frame + padding_frames, max(original_df$s))

      # Extract depth data for the 10-minute window
      depth_data <- original_df %>%
        filter(s >= padded_start_frame & s <= padded_end_frame) %>%
        mutate(time_sec = s / sample_rate)

      # Create a temporary table for phase points
      phase_points <- eventData_df %>%
        group_by(phase) %>%
        slice(1) %>%
        ungroup() %>%
        mutate(time_sec = s / sample_rate, depth = -p)

      # Create the 10-minute depth window plot with vertical lines marking segment start and end, and ph
ase points
      depth_window_plot <- ggplot(depth_data) +
        geom_line(aes(x = time_sec, y = -p, color = "Depth")) +
        geom_vline(xintercept = start_frame / sample_rate, linetype = "dashed", color = "red", size = 0.
5) +
        geom_vline(xintercept = end_frame / sample_rate, linetype = "dashed", color = "red", size = 0.5)
+
        geom_point(data = phase_points, aes(x = time_sec, y = depth), color = "black", size = 2) +
        geom_text(data = phase_points, aes(x = time_sec, y = depth, label = phase), vjust = -0.5, hjust
= 1.2, color = "black", size = 3, fontface = "italic", alpha = 0.7) +
        geom_point(aes(x = start_time_seconds, y = start_depth), color = "forestgreen", size = 4) +
        geom_point(aes(x = end_time_seconds, y = -eventData_df$p[nrow(eventData_df)]), color = "red3", s
ize = 4) +
        geom_hline(yintercept = -deep_dive_threshold, linetype = "dashed", color = "blue", size = 0.5) +
        geom_hline(yintercept = -shallow_dive_threshold, linetype = "dashed", color = "green", size = 0.
5) +
        geom_hline(yintercept = -surface_threshold, linetype = "dashed", color = "purple", size = 0.5) +
        annotate("text", x = mean(c(start_time_seconds, end_time_seconds)), y = start_depth, label = eve
nt_id, vjust = -3, hjust = 0.5, color = "blue", size = 3) +
        labs(title = "Phase Level View", x = "Time (s)", y = "Depth (m)") +
        theme_minimal() +
        theme(legend.position = "top") +
        scale_x_continuous(breaks = seq(min(depth_data$time_sec), max(depth_data$time_sec), by = 60))

      # Combine plots into one
      combined_plot <- (prh_plot / normjerk_plot / p_tempr_plot / depth_window_plot) + plot_layout(ncol
= 1)

      # Save the combined plot
      plot_name <- file.path(base_group_dir, paste(event_id, "composite_dashboard.png", sep = "_"))
      print(paste("Saving plot:", plot_name)) # Debugging statement
      ggsave(plot_name, plot = combined_plot, width = 10, height = 20)
    }
  }
```

````
}
```
````
````