Gefördert durch:

Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages

S C H L U S S B E R I C H T

DigitShipData
Modellierung, Erfassung und Analyse von Schiffsbetriebsdaten für den effizienten Schiffsbetrieb und die Entwurfsoptimierung

Förderkennzeichen: 03SX514B

Zuwendungsempfänger: Universität Rostock

Ausführende Stelle: Universität Rostock, Fakultät für Maschinenbau und Schiffstechnik, Lehrstuhl für Schiffbau, Albert-Einstein-Straße 2, 18059 Rostock

Vorhabenleitung: Prof. Dr.-Ing. Florian Sprenger
Dr.-Ing. Wisam Jabary

Projektlaufzeit: 01.09.2020 bis 30.04.2024

Inhaltsverzeichnis

I. Kurze Darstellung 4
I.1. Aufgabenstellung 4
I.2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde 4
I.3 Planung und Ablauf des Vorhabens 4
I.4 Wissenschaftlicher und technischer Stand zu Projektbeginn 5
I.5 Zusammenarbeit mit anderen Stellen 6

II. Eingehende Darstellung 7
II.1 Erzielte Ergebnisse 7
II.1.1 Detaillierte Beschreibung der betrachteten Schiffsbetriebsdaten (AP 1000) 7
II.1.2 Verarbeitungssystem für Schiffsbetriebsdaten (AP 2000) 31
II.1.3 Prognose der Schiffsleistung (AP 3000) 48
II.1.4 Schiffsbetriebsprofile (AP 6000) 52
II.2 Erläuterungen zum zahlenmäßigen Nachweis 57
II.3 Notwendigkeit und Angemessenheit der Arbeiten 58
II.4 Voraussichtlicher Nutzen – Fortschreibung des Verwertungsplans 58
II.5 Während der Durchführung des Vorhabens bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen 58
II.6 Erfolgte und geplante Veröffentlichungen 58
II.6.1 Erfolgte Veröffentlichungen 58
II.6.2 Geplante Veröffentlichungen 59

Literaturverzeichnis 60

I. Kurze Darstellung

I.1. Aufgabenstellung

Das Konsortium des Verbundprojekts DigitShip hatte sich die Entwicklung von Methoden und Werkzeugen zur Optimierung des Schiffsbetriebs durch die strukturierte Erfassung, Analyse und Verwertung von Betriebsdaten zum Ziel gesetzt.
Die Aufgaben der Universität Rostock innerhalb des Verbunds lagen auf der Entwicklung und Bereitstellung eines Datenmanagementsystems sowie auf der Validierung und Vorverarbeitung der gesammelten Betriebs- und Umweltdaten. Die Aufgaben umfassten die Entwicklung und Bereitstellung eines Informationsmodells, das die synchronisierte Einbindung von großen Mengen heterogener Daten ermöglicht, sowie die Entwicklung von Methoden zur Bewertung der Unsicherheiten in der Datenanalyse. Zudem hatte sich die Universität Rostock zum Ziel gesetzt, eine Methode zu entwickeln, um aus Betriebsdaten repräsentative Betriebsprofile ableiten zu können, die der Verbesserung von Spezifikationen für Schiffsneubauten dienen können.

I.2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Die fortschreitende Installation von Sensorik an Bord von Schiffen sowie die zunehmende Digitalisierung und Verbesserung von Telemetrie ermöglichen es, umfangreiche Betriebsdaten zu erfassen und digital zu verarbeiten. Diese Daten können wertvolle Einblicke in den Schiffs- und Maschinenbetrieb sowie in Umwelteinflüsse wie Seegang, Wind und Strömung liefern. Die Herausforderung besteht jedoch darin, die Qualität und Aussagekraft dieser Daten sicherzustellen. Hierzu ist die Entwicklung spezifischer Daten(vor-)verarbeitungsmethoden erforderlich, die Schiffseignern vor Beginn des Vorhabens nicht oder nur in sehr eingeschränktem Umfang zur Verfügung standen. Ein zentraler Aspekt war die Notwendigkeit der Entwicklung eines Informationsmodells, um eine integrierte Modellierung von Betriebs- und Umweltdaten zu ermöglichen. Die Universität Rostock verfügte bereits vorab über umfangreiche Erfahrungen im Datenmanagement für schiffstechnische Anwendungen und war grundsätzlich in der Lage, große und komplex strukturierte Datenmengen zu verarbeiten und zu analysieren. Um die Betriebsdaten verschiedenen Projektpartnern zugänglich zu machen, war die Umsetzung einer cloudbasierten Datenbereitstellung vorgesehen, die über eine sichere Schnittstelle Zugriff auf die Daten gewährt.
Ein weiteres Ziel war die Entwicklung neuer Methoden zur Datenvorverarbeitung, insbesondere zur Filterung, Imputation und Plausibilitätsprüfung. Hierzu konnte auf den bisherigen Erfahrungen mit Schiffsbetriebsdaten aufgebaut werden. Die gesammelten Daten sollten außerdem durch Algorithmen zur Identifizierung charakteristischer Betriebsprofile ergänzt werden, um eine belastbare Grundlage für langfristige Analysen und Designoptimierungen zu schaffen. Das vorhandene Vorwissen der Arbeitsgruppe am Lehrstuhl für Schiffbau auf dem Gebiet des Schiffsentwurfs und -betriebs war für die Erreichung dieser Zielstellung unabdingbar.

I.3 Planung und Ablauf des Vorhabens

Das Vorhaben der Universität Rostock im Verbundprojekt DigitShip war in mehrere aufeinander aufbauende Teilschritte gegliedert.

Entwicklung des Messkonzeptes und Erfassung der Daten: Zunächst wurde ein Messkonzept für die Datenerfassung an Bord eines Containerschiffes des assoziierten Partners CPO entwickelt. Hierbei wurden die vorhandenen und im Rahmen des Projektes neu installierten Sensoren und Messgeräte in

ein Gesamtkonzept integriert. Die gesammelten Messdaten wurden durch Wetter- und AIS-Daten externer Anbieter ergänzt. Das realisierte Messkonzept ermöglicht eine kontinuierliche Aufnahme und Übertragung relevanter Daten aus unterschiedlichen Quellen in das durch die Universität Rostock entwickelte Datenmanagementsystem.

Entwicklung von Methoden zur Datenvorverarbeitung: Die Universität Rostock entwickelte im nächsten Schritt Methoden zur Plausibilitätsprüfung und Filterung der aufgenommenen Daten. Dies war für den weiteren Projektverlauf von zentraler Bedeutung, da valide statistische Aussagen nur mit sauber vorverarbeiteten Daten möglich sind. Im Rahmen dieses Schrittes wurden Fehlerquellen identifiziert, Ausreißer und ungenaue Messungen eliminiert, und die Synchronizität der Daten sichergestellt, um eine hohe Datenqualität für nachfolgende Analysen zu gewährleisten.

Clustererkennung und Analyse von Betriebsprofilen: Mittels speziell angepasster Algorithmen des maschinellen Lernens wurde ein Verfahren zur Mustererkennung in Langzeitmessdaten aus dem Schiffsbetrieb entwickelt. Hieraus lassen sich charakteristische Betriebszustände eines Schiffs ermitteln, die als Basis für langfristige Leistungsanalysen genutzt werden können und Potenziale zur Reduzierung des Brennstoffverbrauchs und somit der CO_2-Emissionen aufzeigen.

Langzeit-Validierung und Leistungsprognosen: Durch kontinuierliche Validierung wurden die über viele Monate aufgenommenen Langzeitmessungen auf Konsistenz geprüft. Außerdem wurden auf maschinellem Lernen basierende Methoden entwickelt, um den Kraftstoffverbrauch des Containerschiffes des assoziierten Partners CPO unter verschiedenen Umwelt- und Betriebsbedingungen vorherzusagen und den Brennstoffverbrauch in Abhängigkeit von der gefahrenen Route und den vorherrschenden Wetterbedingungen realistisch einzuschätzen.

I.4 Wissenschaftlicher und technischer Stand zu Projektbeginn

Mit der zunehmenden Verfügbarkeit von Sensorik und Digitalisierung an Bord moderner Schiffe stehen immer umfangreichere Betriebsdaten zur Verfügung. Diese ermöglichen die momentane Erfassung von Parametern wie Geschwindigkeit, Ladezustand und Umwelteinflüssen (Seegang, Wind und Strömung), die für die Bewertung der benötigten Antriebsleistung genutzt werden können. Die Datenqualität und Aussagekraft dieser Messungen bleiben jedoch eine zentrale Herausforderung. Zahlreiche Studien belegen, dass trotz fortschrittlicher Erfassungssysteme viele der aufgezeichneten Betriebsdaten unzuverlässig oder physikalisch unplausibel sind, was Analysen erschwert und zu fehlerhaften Ergebnissen führen kann. Beispielsweise konnte im Rahmen von Auswertungen nach der internationalen Norm ISO 19030 zur Messung von Veränderungen in der Rumpf- und Propellerleistung nur ein Bruchteil der erfassten Daten verwendet werden, da häufig Messfehler und ungeeignete Datenqualität vorlagen. Ein besonderes Problem bei der Datenanalyse ist die mangelnde Synchronizität der Messdaten. Verschiedene Sensorsysteme, die oft von unterschiedlichen Zulieferern stammen und keine abgestimmten Datenerfassungsparameter wie Abtastrate und Signalverarbeitung aufweisen, führen zu Inkonsistenzen. So kann der Zeitversatz zwischen unterschiedlichen Messgrößen die Genauigkeit der Analysen erheblich beeinträchtigen, was wiederum den Aussagewert der ermittelten Leistungsparameter verringert. Auch bei der Analyse von Langzeitdaten, die für die Ableitung von Betriebsprofilen und energetischen Bewertungen entscheidend sind, werden die Unsicherheiten in den Daten häufig nicht berücksichtigt oder offengelegt. Es besteht daher ein großer Bedarf an wissenschaftlich fundierten und standardisierten Methoden zur Filterung und

Qualitätsbewertung der Betriebsdaten, die auch die spezifischen Unsicherheiten und Fehlerquellen der unterschiedlichen Messsysteme mit einbeziehen. Eine zusätzliche Herausforderung besteht in der Bereitstellung und Integration externer Umweltdaten (z. B. Seegang, Windbedingungen), da viele Schiffe keine eigenen Systeme zur genauen Erfassung dieser Umgebungsbedingungen besitzen. Stattdessen müssen oft externe Datenquellen herangezogen und auf die vorhandenen Betriebsdaten abgestimmt werden, was die Anforderung an die Datenverarbeitung erhöht.
Der Stand der Forschung zu Beginn des Vorhabens legte nahe, dass die Nutzung hochentwickelter Filter- und Synchronisationsmethoden in Kombination mit objekt-orientierten Datenmanagementsystemen vielversprechende Ansätze zur Sicherstellung der Datenqualität und zur Erstellung repräsentativer Betriebsprofile bietet.

I.5 Zusammenarbeit mit anderen Stellen

Neben der Zusammenarbeit mit den Projektpartnern erfolgte im Rahmen des Vorhabens eine enge und vertrauensvolle Zusammenarbeit mit der Reederei CPO, wodurch die Aufnahme von Langzeitmessdaten an Bord eines Containerschiffes möglich wurde. Weitere Kooperationen zu anderen Forschungsstellen gab es nicht.

II. Eingehende Darstellung

II.1 Erzielte Ergebnisse

Der Schwerpunkt der Arbeiten der Universität Rostock lag in der Aufbereitung und Bereitstellung von Daten aus dem Schiffsbetrieb und von externen Anbietern in einem projektweiten Datenmanagement System sowie der Ergänzung der Daten für Analysen. Der besseren Übersicht halber ist die Darstellung der im Rahmen des Vorhabens erzielten Ergebnisse und durchgeführten Arbeiten gemäß der Projektstruktur nach den jeweiligen Arbeitspaketen untergliedert.

II.1.1 Detaillierte Beschreibung der betrachteten Schiffsbetriebsdaten (AP 1000)

Für die Arbeiten im Projekt wurden reale Schiffsbetriebsdaten von einem als Versuchsträger dienenden Containerschiff des assoziierten Partners CPO aus zwei unterschiedlichen Quellen und mit unterschiedlichen Eigenschaften herangezogen:

a) Angaben und Messungen durch Sensoren und die Besatzung an Bord des Schiffes, zusammenfassend als ”Borddaten” bezeichnet

b) Daten aus externen Quellen, die in Bezug auf das Schiff von Dritten erhoben werden, wozu primär navigatorische Daten sowie die Umweltbedingungen zählen.

Unter dem Begriff ”Borddaten” sind alle Messgroßen und Angaben zu verstehen, die an Bord des Versuchsträgers ermittelt werden. Hierzu gehören neben den von der Besatzung manuell zu führenden Berichten (”Noon-Reports”) sowie den Daten zum jeweiligen Beladungszustand aus dem Ladungsrechner die automatisch über Sensoren aufgezeichneten Daten zu allen relevanten schiffstechnischen Systemen als auch zu dem aktuellen Bewegungsverhalten des Schiffes mittels einer inertialen Messeinheit (englisch: Inertial Measurement Unit (IMU)). Abgesehen von der IMU, die spezifisch für das Projekt an Bord des Versuchsträgers installiert wurde, gehörte die verwendete Sensorik zur Ausstattung des Schiffes. Die Messdaten wurden in unterschiedlichen Frequenzen aufgenommen. Hierbei ist grundsätzlich zwischen niederfrequenten Daten (low frequency data), bspw. für Reports und Ladungsrechnerdaten sowie hochfrequenten Daten (high frequency data), bspw. für Sensor- und Bewegungsdaten zu unterscheiden. Das Gesamtszenario ist in Abbildung 1 schematisch dargestellt.

Sowohl Reports als auch die Ladungsrechnerdaten wurden mindestens einmal täglich vom Schiff über eine Satellitenverbindung zur Reederei übertragen und dort über eine Internetverbindung mit den entsprechenden Webanwendungen der Forschungsstelle bereitgestellt.

Die aufgezeichneten Messgroßen des Data Loggers sowie einer inertialen Messeinheit wurden an Bord gespeichert und mit Hilfe einer transportablen Festplatte an Land gebracht, wenn das Schiff einen europäischen Hafen anläuft. Der Anbieter FleetMon, über den die navigatorischen Daten (AIS, Automatic Identification System) mittels eines Unterauftrags bezogen wurden, stellte diese der Forschungsstelle ebenfalls über eine Online-Verbindung zur Verfügung. FleetMon erhielt die schiffsbezogenen Daten entweder über terrestrische Verbindungen oder, wenn dies aufgrund des Abstands zur Küste nicht möglich ist, ebenfalls über Satelliten. Darüber stellte FleetMon der Universität Rostock orts- und zeit-synchronisierte Umweltbedingungen für die jeweilige Position des Schiffes zur Verfügung.

Die Universität Rostock hat in einem ersten Schritt zunächst eine systematische Übersicht über sämtliche relevante Daten erstellt. Diese wichtige Vorarbeit diente als Lastenheft für die Umsetzung des Datenmanagement Systems.

In den folgenden Abschnitten werden alle in das Datenmanagement System eingebundenen

Daten detailliert beschrieben.

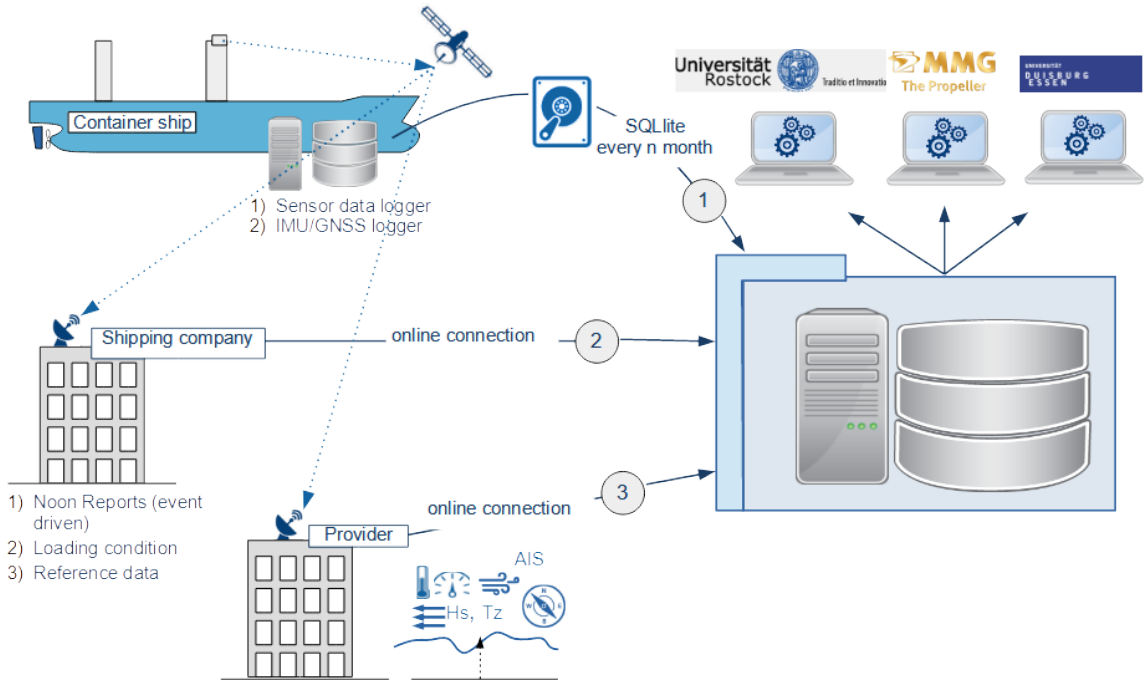


Abbildung 1: Schematische Darstellung der Erhebung von Schiffsbetriebsdaten aus unterschiedlichen Quellen sowie der nachfolgenden Verarbeitung in dem speziell entwickelten Datenmanagement-System.

II.1.1.1 Niederfrequente Daten

Zu den niederfrequenten Daten gehören zum einen die Berichte („Noon Reports“), die manuell an Bord von der Besatzung erstellt werden und zum anderen die Daten des Ladungsrechners, welche beim Auslaufen aus einem Hafen jeweils aktualisiert werden.

Berichte – Noon Reports

Noon Reports werden manuell von autorisierten Besatzungsmitgliedern an Bord gepflegt und stehen bei höher ausgerüsteten Schiffen - wie dem im Projekt zur Verfügung stehenden Versuchsträger - der technischen Inspektion der Reederei über eine online-Verbindung zur Verfügung. Für das Forschungsvorhaben hat die Reederei CPO der Forschungsstelle einen online Zugang eingerichtet, womit diese direkt und vollumfänglich auf die Daten des Versuchsträgers zugreifen konnte. Beim Versuchsträger handelte es sich um das 13.000 TEU Containerschiff „MSC Bari“ (siehe Abbildung 2), auf welchem das System „Bluetracker“ der Firma „navis“ [1] zum Einsatz kommt. Die Reports umfassen über 200 Angaben zum jeweils aktuellen Zustand des Schiffes, von Bordsystemen sowie z.B. auch zum Wetter sowie administrativen Informationen. Der traditionelle Begriff Noon-Report hat hier seine Gültigkeit verloren, da diese Berichte zusammengefasst ereignisgesteuert erfolgen, und nicht nur wie in der Vergangenheit regelmäßig um 12 Uhr mittags (noon) eines jeden Tages. Ereignisse (events), zu denen der Report jeweils aktualisiert wird, werden unterschieden in:

Abbildung 2: Der Versuchsträger MSC Bari des assoziierten Partners CPO ist ein 13.000 TEU Containerschiff, Baujahr 2011

Abkürzung	**Ereignis**
DEPA	Departure from Port
ARRI	Arrival in Port
BOSP	Beginning of Sea passage
EOSP	End of Sea passage
NOON (Sea)	Noon at Sea
NOON (Man)	Noon while Manoeuvring
NOON (Anc)	Noon at Anchor
NOON (Drf)	Noon while Drifting
NOON (Prt)	Noon in Port

Hieraus lassen sich die folgenden fünf Betriebszustände (states) ableiten, in denen sich ein Schiff zu einem bestimmten Zeitpunkt befinden kann:

- At Sea - Sea passage
- Drifting
- At Anchor
- Maneuvering
- In Port

Die durch die Besatzung manuell gepflegten Borddaten werden ergänzt durch Messergebnisse, die von Sensoren an Bord geliefert werden, und über entsprechende Software in dem betreffenden Zeitraum aggregiert bzw. gemittelt werden. Diese Werte unterscheiden sich nicht erheblich von den manuell erhobenen, dienen aber der Überprüfung der Plausibilität der von der Crew gemachten Angaben. Letztere sind Stand heute rechtsverbindlich in Auseinandersetzungen zwischen Reeder und Charterer. Die Sensordaten selbst werden gesondert in höheren Frequenzen protokolliert, siehe dazu den nachfolgenden Abschnitt II.1.1.2.

In den Berichten lassen sich die Angaben bzw. die entsprechenden Mess- und Berechnungsgroßen in die übergeordneten Kategorien

- Ship: übergeordneten Informationen zur Reise

- Navigation: navigatorische Informationen
- Loading Condition: grundlegende Informationen zum Beladungszustand des Schiffes
- Machinery: Informationen zur Hauptmaschine, den Hilfsmaschinen, den Abgasnachbehandlungsanlagen, zum Brennstoff einschließlich Verbrauch und Tankfüllungsständen, zur Propulsion, Performance-Indikatoren (KPIs)
- Environment: zusammenfassende Umgebungsbedingungen auf dem Routenabschnitt

einteilen, wobei alle Angaben für den jeweiligen Berichtszeitraum zusammengefasst bzw. gemittelt sind.
In den folgenden Abschnitten werden die zu den einzelnen Kategorien zugehörigen Angaben, entsprechenden Mess- und Berechnungsgroßen mit den verwendeten, nicht immer SI-konformen, Einheiten aufgelistet.

- **Ship**

In dieser Kategorie sind alle übergeordneten Informationen zur Reise zusammengefasst, die zum Zeitpunkt der Berichterstellung relevant sind. Hierzu gehören:

Parameter	**Format**
name Captain	[Zeichenkette]
name Chief Engineer	[Zeichenkette]
report timestamp	[UTC Zeit]
reporting period	[h]
sea time duration	[h]
status	[ENUM]
event	[ENUM]
ETA UTC	[timestamp]
ETA local	[timestamp]

Aus Datenschutzgründen werden die Namen des Kapitäns und des Chief Engineers beim Import in die DigitShip Projekt-Datenverwaltung anonymisiert. „status“ und „event“ sind Angaben aus einer der oben angegebenen Aufzählungen.

- **Navigation**

In dieser Kategorie sind alle Informationen zusammengefasst, die zum Zeitpunkt der Berichterstellung navigatorische Angaben betreffen. Hierzu gehören:

Parameter	**Format**
position latitude	[degree]
position longitude	[degree]
under keel clearance (UKC)	[m]
water depth	[m]
heading	[degree]
course made good	[degree]
speed over ground (SOG)	[kn]
averaged speed over ground	[kn]
speed through water (STW)	[kn]

distance over ground	[nautical miles]
distance through water	[nautical miles]

- **Loading Condion**

In dieser Kategorie sind alle übergeordneten Informationen zusammengefasst, die zum Zeitpunkt der Berichterstellung die Schwimmlage und den Beladungszustand des Schiffs betreffen. Hierzu gehören:

Parameter	Format
displacement	[t]
draught forward	[m]
draught aft	[m]
mean draught	[m]
trim	[m]
cargo load	[t]
container full	[TEU]
container empty	[TEU]
container reefer	[TEU]
water ballast	[t]

Weitere Details sind in den Daten des Ladungsrechners zu finden, siehe nächster Abschnitt.

- **Machinery**

In dieser Kategorie sind ausgewählte Informationen zusammengefasst, die zum Zeitpunkt der Berichterstellung die Maschinenanlage sowie alle Angaben zum Brennstoff und den Emissionen für den Berichtszeitraum betreffen. Ferner sind ausgewählte übergeordnete Leistungskennzahlen (KPI) enthalten. Diese Informationen können in die folgenden Unterkategorien weiter untergliedert werden: allgemeine Daten, Gesamtbrennstoffverbräuche, Daten für die Hauptmaschine(n), die Hilfsmaschinen, für den Hilfskessel, Umweltdaten und Kennzahlen.

Parameter	Format
Allgemeine Daten (general data)	
engine distance	[nautical miles]
total running hours ME	[h]
average shaft RPM	[1/min]
average slip through water	[%]
avg slip over ground	[%]
Gesamtbrennstoffverbräuche (fuel oil consumption)	
total fuel oil consumption (FOC)	[t]
total fuel oil consumption (ISO FOC)	[t]
daily fuel oil consumption (DFOC)	[t/d]
total fuel oil consumption heavy fuel oil (HFO)	[t]
total fuel oil consumption high sulfur (HFO HS)	[t]
sulfur content	[%]
total fuel oil consumption low sulfur (HFO LS)	[t]
total fuel oil consumption marine diesel oil (MDO)	[t]

total fuel oil consumption marine gas oil (MGO)	[t]
Hauptmaschine (main engine)	
average engine power	[kW]
average utilization	[%]
total energy	[kWh]
average shaft power	[kW]
total shaft energy	[kWh]
total fuel oil consumption (ME FOC)	[t]
total fuel oil consumption ISO conditions (ME FOC ISO)	[t]
daily fuel oil consumption (ME DFOC)	[t/d]
specific fuel oil consumption (ME SFOC)	[g/(kWh)]
specific fuel oil consumption ISO conditions (ME SFOC ISO)	[g/(kWh)]
Hilfsmaschinen (auxiliary engines)	
average engine power	[kW]
average utilization	[%]
total energy	[kWh]
total electrical energy	[kWh]
average electrical power	[kW]
average generators utilization	[%]
total fuel oil consumption (AE FOC)	[t]
total fuel oil consumption ISO conditions (AE FOC ISO)	[t]
daily fuel oil consumption (AE DFOC)	[t/d]
specific fuel oil consumption (AE SFOC)	[g/(kWh)]
specific fuel oil consumption ISO conditions (AE SFOC ISO)	[g/(kWh)]
Hilfskessel (auxiliary boiler)	
total fuel oil consumption	[t]
total fuel oil consumption ISO conditions (AB FOC ISO)	[t]
daily fuel oil consumption (AB DFOC)	[t/d]
Emissionen (emissions)	
total CO2 emissions	[t]
ME CO2 emissions	[t]
AE CO2 emissions	[t]
AB CO2 emissions	[t]
total SOx emissions	[t]
ME SOx emissions	[t]
AE SOx emissions	[t]
AB SOx emissions	[t]
total NOx emissions	[t]
ME NOx emissions	[t]
AE NOx emissions	[t]

AB NOx emissions	[t]
Scrubber running	[yes | no]
Umweltbezogenen Kennzahlen (KPI)	
EEOI per TEU	[gCO2/(TEU nautical mile)]
EEOI per load	[gCO2/(TEU nautical mile)]
transport efficiency	[gCO2/(cargo t nautical mile)]

- **Environment**

In dieser Kategorie sind alle Informationen zusammengefasst, die zum Zeitpunkt der Berichterstellung die vorherrschenden Umweltbedingungen betreffen. Hierzu gehören die folgenden Unterkategorien: Bathymetrie, Wasser- und Luftdaten, Wind, Seegang (Windwellen und Dünung). Zu beachten ist hierbei, dass einige Daten nicht an Bord erhoben werden, sondern von einem Wetterdienst bereitgestellt werden. Diese sind im Bericht mit „weather provider" gekennzeichnet.

Parameter	**Format**
Wasser und Luftparameter (water and air conditions)	
seawater temperature measured	[degree Celsius]
seawater temperature weather provider	[degree Celsius]
air temperature measured	[degree Celsius]
air temperature weather provider	[degree Celsius]
air pressure	[hPa]
air pressure weather provider	[hPa]
visibility range weather provider	[km]
Wind (wind)	
Beaufort	[Number: 0...12]
true wind speed	[m/s]
true wind weather provider	[m/s]
true wind direction	[degree]
true wind direction weather provider	[degree]
apparent wind speed	[m/s]
apparent wind direction	[degree]
Seegang (waves)	
sea state	[Number: 0...9]
wave height	[m]
wave height weather provider	[m]
wave direction	[degree]
swell height	[m]
swell height weather provider	[m]
swell direction	[degree]
swell direction weather provider	[degree]
swell period weather provider	[s]

Hinweis: In den Berichten, die der Universität Rostock im Rahmen des Vorhabens vorlagen, sind

nicht immer alle Angaben zu finden. Dies gilt auch für solche, die prinzipiell zur Verfügung stehen sollten.

Aktuelle Beladungssituation – Ladungsrechner

Die Daten des Ladungsrechners der MSC Bari konnten mit Hilfe eines Web-Interfaces des installierten MACS3 Connected Systems der Firma navis heruntergeladen werden. Die Aktualisierungen erfolgten immer zum Zeitpunkt des Verlassens eines Hafens (Ereignis DEPA). Zustände aus der Vergangenheit waren nur sehr eingeschränkt einsehbar. Die zur Verfügung stehenden Daten können weiter untergliedert werden in:

- Hydrostatik: Schwimmlage (Hydrostatics): Angaben zur Schwimmlage mit Tiefgängen und Schwerpunkten für das ganze Schiff
- Hydrostatik Stabilität (Stability): Parameterwerte der überprüften Stablitätsanforderungen
- Tanks (Tanks): Angaben zu den Füllständen für alle Tanks mit Masse und Schwerpunkt sowie Füllgrad
- Containerladung (Container Cargo): detaillierte Angaben zu jedem Container an Bord
- Beladungszustand (Loading Condition Summary): Massen und Schwerpunkte für das Schiff mit allen Massengruppen

Im Folgenden werden die zu den einzelnen Kategorien zugehörigen Angaben mit den jeweils verwendeten, nicht immer SI-konformen, Einheiten aufgelistet.

- **Hydrostatik: Schwimmlage (hydrostatics)**

Diese Kategorie beinhaltet alle Angaben zur aktuellen Schwimmlage und den relevanten Schwerpunkten für das gesamte Schiff.

Parameter	**Format**
Seawater density	[t/m3]
displacement	[t]
draft aft	[m]
draft midships	[m]
draft forward	[m]
list angle	[degree]
trim	[m]
FS Correction	[m]
LCF im Schiffskoordinatensystem	[m]
LCF relative to midship (Lpp/2)	[m]
VCG	[m]
VCG (corrected)	[m]

- **Hydrostatik: Stabilität (stability)**

In dieser Kategorie sind sämtliche Angaben zur Überprüfung der Stabilitätsanforderungen nach dem Intact Stability Code der IMO [2] zusammengefasst.

Parameter	Format
GM (corrected)	[m]
angle due to transverse moment	[degree]
angle due to wind + transverse moment	[degree]
max. lever GZ at angle >= 30 degrees	[m]
angle of max righting lever GZ	[degree]
area up to 30 degrees	[mrad]
area up to 40 degrees	[mrad]
area between 30 and 40 degrees	[mrad]
amplitude of rolling	[degree]
period of rolling	[s]
weather criterion	[Zahl]
min. draft fore	[m]

- **Tanks (tank fillings)**

Diese Kategorie enthält Angaben zu Füllständen sämtlicher Tanks an Bord des Schiffes sowie zu den daraus resultierenden Schwerpunkten in allen drei Schiffshauptkoordinatenrichtungen.

Parameter	Format
tank purpose (acc. contents)	[Zeichenkette]
tank ID	[Zeichenkette]
aktual filling	[m3]
filling ratio	[%]
maximum tank volume	[m3]
density of fluid	[t/m3]
mass	[t]
LCG	[m]
TCG	[m]
VCG	[m]

Hinweis: Die Angaben zu den absoluten und relativen Füllvolumina und die weiteren Angaben zum maximalen Tankvolumen sind redundant, gleiches gilt für die Angaben zur Dichte und Masse.

- **Ladung: Container (load, container)**

Diese Kategorie enthält für alle Container neben kennzeichnenden Angaben und solchen zur Beladungslogistik Angaben zur Masse und dem Stauplatz, hierzu gehören im Einzelnen:

Parameter	Format
Type	
{standard | dangerous goods | reefer}	[Zeichenkette]
position	[6-stellige Zahl]
mass	[t]
port of loading (POL)	[UN/LOCODE: Buchstabenkette]
port of discharge (POD)	[UN/LOCODE: Buchstabenkette]
serial number	[alphanum. Zeichenkette]

ISO Typ Code	[4-st. alphanum. Zeichenkette]
Length	[m]
Breadth	[m]
Height	[m]

Aus dem Stauplatz (position) lassen sich die Schwerpunkte im Schiffshauptkoordinatensystem bestimmen, wenn hierzu Annahmen zu den Schwerpunkten der Container selbst getroffen werden. Über die Abmessungen lässt sich der Typ schlussfolgern, bspw. TEU, FEU, high cube.

- **Zusammenfassung Beladungszustand (summary)**

Diese Kategorie enthält für Hauptgruppen der Massenverteilung Angaben zu (Gruppen-)Massen und Schwerpunkten in allen drei Schiffshauptkoordinatenrichtungen.

Parameter	**Format**
Tanks	Masse [t], LCG [m], VCG [m], TCG [m]
Container	Masse [t], LCG [m], VCG [m], TCG [m]
Deadweight	Masse [t], LCG [m], VCG [m], TCG [m]
Leeres Schiff	Masse [t], LCG [m], VCG [m], TCG [m]
Deplacement (Summe)	Masse [t], LCG [m], VCG [m], TCG [m]

II.1.1.2 Hochfrequente Daten

Zu den hochfrequenten Daten, die im Rahmen des Vorhabens auch der MSC Bari aufgenommen wurden, zählen die Sensordaten aus dem bereit vorhandenen System Performance Monitoring und zum anderen die Messdaten der speziell für das Projekt installierten Inertial Measurement Unit (IMU).

Sensordaten

An Bord des Versuchsträgers MSC Bari wurden umfangreiche Daten über das installierte System Performance Monitoring [3] der Firma Hoppe Marine zur Erfassung zentral bereitgestellt. Hiervon wurden für das Projekt in einem ersten Schritt diejenigen Mess- und Berechnungsgroßen ausgewählt, die eine besonderer Relevanz für die geplanten Analysen haben. Diese lassen sich in die folgenden übergeordneten Kategorien einteilen:

- Navigation: navigatorische Informationen
- Loading Condition: grundlegende Informationen zum Beladungszustand des Schiffes
- Machinery: Informationen zur Hauptmaschine, den Hilfsmaschinen, zum Brennstoff einschließlich Verbrauch, zur Propulsion sowie zu Leistungskennzahlen (KPIs)
- Environment: aktuelle Umgebungsbedingungen

Diese Kategorisierung erfolgt analog zu der Einteilung, die für die Noon Reports in Abschnitt II.1.1. vorgenommen wurde. In den folgenden Abschnitten werden die zu den einzelnen Kategorien zugehörigen Angaben, entsprechenden Mess- und Berechnungsgrößen mit den verwendeten, nicht immer SI konformen, Einheiten aufgelistet. Ferner ist die Frequenz angegeben, mit der die jeweiligen Größen protokolliert wurden.

Alle Mess- bzw. Berechnungsgroßen wurden einer Zeit- und Ortsangabe zugeordnet, diese sind definiert durch:

Parameter	**Format**
date	UNIX timestamp
latitude	[deg]
longitude	[deg]

- **Navigation**

Sämtliche Messgrößen, die der Kategorie navigatorische Informationen zugeordnet werden können, wurden mit einer einheitlichen Frequenz von 5Hz aufgenommen.

Parameter	**Format**	**Frequenz**
course over ground (COG)	[deg]	5 Hz
heading	[deg]	5 Hz
rate of turn	[deg/min]	5 Hz
speed over ground (SOG)	[kn]	5 Hz
distance over ground	[nautical miles]	5 Hz
speed through water (STW)	[kn]	5 Hz
distance through water	[nautical miles]	5 Hz
rudder angle max	[deg]	5 Hz
rudder angle min	[deg]	5 Hz

rudder angle average	[deg]	5 Hz
rudder angle RMS	[deg]	5 Hz

- **Loading Condition**

In diese Kategorie gehören alle Messwerte, die Informationen bezüglich des auf einem bestimmten Routenabschnitt jeweils vorherrschenden Beladungszustands liefern. Diese umfassen neben der Schwimmlage insbesondere die momentane Füllung (Massen) sämtlicher Tanks mit flüssigen Betriebs- und Verbrauchsstoffen an Bord des Schiffes. Die Aktualisierungsfrequenz dieser Daten variiert zwischen 5 mal pro Sekunde und alle 15 Minuten.

Parameter	Format	Frequenz
Schwimmlage		
draft forward at center line	[m]	5 Hz
draft midship at port side	[m]	5 Hz
draft midship at starboard side	[m]	5 Hz
draft aft at center line	[m]	5 Hz
Ballastwassertanks		
Füllung	[t]	1/900 Hz
Brennstofftanks		
Füllung	[t]	1/60 Hz
Schmieröltanks		
Füllung	[t]	1/900 Hz
Frischwassertanks		
Füllung	[t]	1/900 Hz

- **Machinery**

Die Sensorik am Antriebsstrang des Schiffes liefert wichtige Messdaten aus dem Maschinenraum, von der Antriebswelle und zum Brennstoff. Daraus abgeleitet werden wichtige Leistungsindikatoren (KPI) zum Schiffsantrieb berechnet. Die Aktualisierungsfrequenz dieser Daten variiert zwischen 5 mal pro Sekunde und alle 15 Minuten.

Parameter	Format	Frequenz
Hauptmaschine		
power	[kW]	5Hz
tourque	[kNm]	5Hz
load	[%]	5Hz
light run	[%]	5Hz
running hours	[h]	1/900Hz
total energy	[GWh]	1/900Hz
thrust bearing temperature	[°C]	1/60Hz
engine room temperature	[°C]	1/60Hz
Propulsion		
shaft speed	[rpm]	5Hz
shaft revolution	[rotations]	5Hz
slip through water	[%]	5Hz

slip over ground	[%]	5Hz
Brennstoff		
ME fuel mass supply	[t/h]	5Hz
ME fuel volume supply	[m3/h]	5Hz
ME total fuel mass supply	[t]	5Hz
ME total fuel mass supply HFO	[t]	5Hz
ME total fuel mass supply HFO LS	[t]	5Hz
ME total fuel mass supply MGO	[t]	5Hz
ME total fuel mass supply MGO LS	[t]	5Hz
ME total fuel mass supply MDO	[t]	5Hz
ME on HFO	[-]	5Hz
ME on HFO LS	[-]	5Hz
ME on MGO	[-]	5Hz
ME on MGO LS	[-]	5Hz
ME on MDO	[-]	5Hz
Density	[t/m3]	5Hz
temperature	[°C]	5Hz
viscosity	[mm2/s]	5Hz
lower calorific value	[MJ/kg]	5Hz
KPI		
main engine SFOC be	[g/(kWh)]	5Hz
main engine SFOC ISO	[g/(kWh)]	5Hz
main engine SFOC through water	[kg/(nautical mile)]	5Hz
main engine SFOC over ground	[kg/(nautical mile)]	5Hz
main engine efficiency	[%]	5Hz
auxiliary engines SFOC be	[g/(kWh)]	5Hz
auxiliary engines SFOC through water	[kg/(nautical mile)]	5Hz
auxiliary engines SFOC over ground	[kg/(nautical mile)]	5Hz
SFOC through water	[kg/(nautical mile)]	5Hz
SFOC over ground	[kg/(nautical mile)]	5Hz

- **Environment**

An Bord des Schiffes befindet sich Sensorik zur Messung des Windes (Anemometer) sowie zu Bestimmung der Wassertiefe unter dem Kiel –(under keel clearance, UKC). Diese mit einer einheitlichen Frequenz von 5Hz aufgenommen.

Parameter	**Format**	**Frequenz**
wind direction true North	[°]	5Hz
wind direction relative	[°]	5Hz
wind speed true	[m/s]	5Hz
wind speed relative	[m/s]	5Hz
under keel clearance UKC	[m]	5Hz

Daten der IMU - Schiffsbewegungen

An Bord des Versuchsträgers MSC Bari wurde speziell für das Projekt eine IMU installiert, die die Bewegungen des Schiffes im Seegang mit einer Abtastfrequenz von 100Hz aufzeichnen kann.

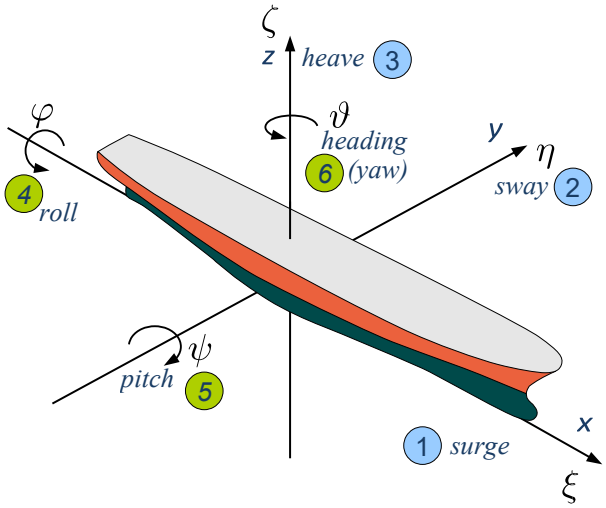


Abbildung 3: Freiheitsgrade (6DOF) und Messgrößen bzgl. der Schiffsbewegungen

Parameter	Format	Frequenz
date	UTC timestamp	100Hz
high precision timestamp	[hour, minute, second, millisecond]	100Hz
heading	[°]	100Hz
roll angle	[°]	100Hz
pitch angle	[°]	100Hz
heave	[m]	100Hz
heave without lever arm	[m]	100Hz
surge	[m]	100Hz
sway	[m]	100Hz
heave speed	[m/s]	100Hz
surge speed	[m/s]	100Hz
sway speed	[m/s]	100Hz
heading rate	[°/s]	100Hz
validity indicator	[binary]	100Hz

Während sich der Kurswinkel (heading) auf das globale Weltkoordinatensystem bezieht (Norden entspricht 0°, Osten 90°, Süden 180° und Westen 270°), erfolgen die Winkelangaben bzgl. der Bewegungsfreiheitsgrade Rollen (roll) und Stampfen (pitch) bezogen auf das schiffsfeste Koordinatensystem.

II.1.1.3 Daten aus externen Quellen

Neben den an Bord des Versuchsträgers aufgenommenen Daten wurden im Rahmen des Vorhabens weitere Daten aus externen Quellen herangezogen. Hierzu gehören zum einen die fortlaufenden AIS-Informationen und zum anderen Daten zur Beschreibung der Umweltbedingungen an den jeweiligen Orten, an denen sich das Schiff zu den bestimmten

Zeitpunkten aufhielt. Letztere werden häufig verkürzt auch als ”Wetterdaten” bezeichnet. Diese Ergänzung der Daten ist unabdinglich, um eine Zuordnung von Ursachen zum Energieverbrauch an Bord des Schiffes vornehmen zu können.
Die AIS-Informationen wurden für den Versuchsträger redundant erfasst. Dies war zum einen notwendig, um die Synchronisierung mit den Umweltdaten zu ermöglichen, zum anderen konnte damit bewertet werden, wie verlässlich diese Angaben aus einer externen Quelle bezogen werden können, was besonders für die Bearbeitung des Arbeitspaketes AP6000 eine wichtige Rolle spielte.

AIS

Für die im Forschungsvorhaben durchgeführten Analysen wurden nicht nur AIS-Daten des Containerschiffs MSC Bari, sondern auch Langzeitdaten von 25 weiteren Schiffen beschafft. In den nachfolgenden Tabellen sind die Informationselemente des AIS-Signals in vier Kategorien untergliedert, für Details siehe [4]:

1. Statische Daten (static), Tabelle 1
2. dynamische Daten (dynamic), Tabelle 2
3. Reiserelevante Daten (voyage related), Tabelle 3
4. sicherheitsrelevante Daten (safety related). Sicherheitsbezogene Kurznachrichten sind Textnachrichten im festen oder freien Format, die entweder an ein bestimmtes Ziel (MMSI) oder an alle Schiffe in diesem Gebiet gerichtet sind. Ihr Inhalt sollte für die Sicherheit der Schifffahrt relevant sein, z.B. ein gesichteter Eisberg oder eine Boje, die nicht auf Station ist. Die Meldungen sollten so kurz wie möglich gehalten¨ werden. Das System erlaubt bis zu 158 Zeichen pro Meldung, aber je kurzer die¨ Meldung ist, desto eher wird sie übertragen. Um alle Möglichkeiten zukünftig offen zu halten sind diese Meldungen zurzeit nicht weiter geregelt.

Tabelle 1: statische AIS-Informationselemente

Informationselement	**Einheit/Kurzbeschreibung**
IMO-Nummer	siebenstellige Nummer
Schiffsname	alphanumerische Zeichenkette
Rufzeichen	bis zu 7 alphanumerische Zeichenkette
MMSI-Nummer	neunstellige Nummer
Schiffstyp	Zeichenkette: aus einer vordefinierten Liste
Länge	Zahl [m]
Breite	Zahl [m]
Position der Antenne des elektronischen Positionsbestimmungssystems (EPFS)	vier Parameter [m]: - Distanz von Bug Distanz von Heck Distanz von Backboard Distanz von Steuerbord

Tabelle 2: dynamische AIS-Informationselemente

Informationselement	**Einheit/Kurzbeschreibung**
Navigationsstatus	aus einer vordefinierten Liste wie ”at anchor”, ”moored”, siehe [4]
Schiffsposition	Longitude [Zahl]

	Latitude [Zahl]
Zeit der Schiffsposition	UTC Zeitstempel
Kurs über Grund (COG)	relativ zum wahren Norden [°]
Geschwindigkeit über Grund (SOG)	Zahl [kn]
Vorausrichtung (HDG)	relativ zum wahren Norden [°]
Kursänderungsrate (ROT)	Grad pro Minute [°/min]

Tabelle 3: Reiserelevante AIS-Informationselemente

Informationselement	**Einheit/Kurzbeschreibung**
Tiefgang	Zahl [m]
Gefahrgutklasse der Ladung (IMO)	Code, Beispiele: DG: Dangerous Goods HS: Harmful substances MP: Marine pollutants
Reiseziel	entsprechend UN/LOCODE, siehe [5]
geschätzte Ankunftszeit (ETA)	UTC Zeitstempel
Routenplan (Wegpunkte)	für jeden Wegpunkt: Longitude und Latitude (manuell zu Beginn der Reise eingegeben)

Die Soll-Frequenzen der Aufnahme und somit der Aktualisierungsrate der Informationen und deren Übermittlung bestimmen sich nach der jeweiligen Informationskategorie:

1. statische und Reise relevante Daten: alle 6 Minuten,
2. dynamische Daten: abhängig von Geschwindigkeit und Kursänderung, siehe Tabelle 4 für die definierten Bereiche,
3. sicherheitsrelevante Daten: in der Regel nach Bedarf.

Tabelle 4: Sendeintervalle der dynamischen Daten nach [4]

Schiffsstatus	**Sendeintervall**
Schiffe, die SOLAS unterliegen	
vor Anker, oder festgemacht und bewegt sich nicht schneller als 3 kn	3 min
vor Anker, oder festgemacht und bewegt sich schneller als 3 kn	10 s
bewegt sich zwischen 0 und 14 kn	10 s
bewegt sich zwischen 0 und 14 kn bei Kursänderung¨	3 1/3 s
bewegt sich zwischen 14 und 23 kn	6 s
bewegt sich zwischen 14 und 23 kn bei Kursänderung	2 s
bewegt sich mit mehr als 23 kn	2 s
bewegt sich mit mehr als 23 kn bei Kursänderung¨	2 s
Schiffe, die nicht SOLAS unterliegen	
bewegt sich mit weniger als 2 kn	3 min
bewegt sich zwischen 2 und 14 kn	30 s
bewegt sich zwischen 14 und 23 kn	15 s
bewegt sich mit mehr als 23 kn	5 s

Da die betrachteten Schiffe ohne Ausnahme SOLAS unterliegen, sind nur die Angaben im oberen Abschnitt der Tabelle für die weiteren Betrachtungen relevant.

Umweltdaten

Die umfassende Bewertung der Performance eines Schiffes im Betrieb ist nur dann möglich, wenn die zu den betrachteten Zeitpunkten vorhandenen Umweltbedingungen an den jeweiligen Orten bekannt sind, häufig verkürzt als Wetterbedingungen bezeichnet.

Die geografischen Positionen, an denen die umweltbezogenen Daten vorliegen müssen, leiten sich aus dem Routenverlauf der betrachteten Schiffe ab. Die Aktualisierungsfrequenz der Umweltdaten bestimmt sich nach dem Operationsstatus des betrachteten Schiffes (bspw. in port, at sea, manoeuvring, at anchor, etc.) sowie der statistischen Änderungsrate von Wetterbedingungen an einem Ort. Es wird allgemein angenommen, dass sich Wetterbedingungen innerhalb von 30 Minuten nicht signifikant ändern und somit innerhalb dieses Intervalls als konstant angesehen werden können.

Für die Bewertung der Effizienz des Schiffsbetriebs ist die Kenntnis der in Tabelle 5 gelisteten orts- und zeitsynchronisierten Parameter erforderlich.

Tabelle 5: Notwendige Umweltdaten für die Bewertung der Schiffsperformance

Kategorie	**Einheit**
Wind waves	
significant wave heigth (Hs)	[m]
wave peak period (Tp)	[m]
mean wave direction	relative to North [°]
Swell	
significant wave heigth (Hs)	[m]
wave peak period (Tp)	[m]
mean wave direction	[°] relative to North
Current	
current speed	[m/s]
mean current direction	[°] relative to North
Wind	
wind speed	[m/s]
mean wind direction	[°] relative to North
strength of wind gusts	[m/s]
Water	
temperature	[°]
salinity	[%]
Air	
temperature	[°]
pressure	[hPa]
relative humidity	[%]

Referenzdaten des Schiffes

Unter Referenzdaten werden die Informationen des betrachteten Schiffes verstanden, die das Schiff und seine hier relevanten Komponenten im werftneuen Zustand beschreiben bzw. die Daten, die in Modellversuchen und Probefahrten ermittelt wurden. Diese können in die folgenden Kategorien unterteilt werden:

- Schiff: Hauptdaten, Generalplan, Schiffsform, Formkurven, Loading Manual, Ruder
- Maschinenanlage Anlagenparameter: Ergebnisse der Shop Tests
- Modellversuche: gemessene hydrodynamische Kenngroßen und daraus abgeleitete Prognosen für die Großausführung
- Probefahrtergebnisse: gemessene Maschinenleistung in Abhängigkeit vom Beladungszustand und der Schiffsgeschwindigkeit

- **Schiff**

Zu den Referenzdaten eines Schiffes zählen die Hauptdaten. Im Kontext des Vorhabens zählen hierzu die folgenden übergeordneten, eher informatorischen statischen allgemeinen Angaben, Hauptabmessungen, Ladungsinformationen und Dienstgeschwindigkeit(en).

Parameter	**Format**
Allgemeine Angaben	
IMO Number	[Zahl]
Name	[Zeichenkette]
type	[Zeichenkette]
owner	[Zeichenkette]
builder	[Zeichenkette]
yard number	[Zahl]
date of delivery	[Datum]
class notation hull	[Zeichenkette]
class notation machinery	[Zeichenkette]
Hauptabmessungen	
Length between perpendiculars	[m]
Length over all	[m]
Bredth	[m]
Design Draft	[m]
Scantling Draft	[m]
Depth	[m]
Freeboard Depth	[m]
Ladung	
nominal cargo capacity	[TEU]
cargo capacity homogeneous loading 14t	[TEU]
cargo capacity in holds	[TEU]
cargo capacity dangerous goods	[TEU]
number of reefer plugs	[Zahl]
Geschwindigkeiten	
service speed	[kn]
service speed draft	[m]

Der Generalplan liegt in Form einer Zeichnung (in der Regel im pdf-Format) vor, und hat im Kontext des Vorhabens lediglich informatorischen Charakter.

Die Schiffsform konnte aus den vorhandenen Ablieferungsunterlagen in einer für die vorgesehenen Analysen erforderlichen Genauigkeit rekonstruiert werden. Sie liegt in

verschiedenen Repräsentationen (Dateiformaten) vor.
Unter diesem traditionellen Begriff werden tabellarische Daten verstanden, die die geometrischen Parameter der Schiffsform als Funktion des Tiefgangs ausweisen. Es wird immer davon ausgegangen, dass das Schiff auf ebenem Kiel schwimmt - weder Krängung noch Trimm sind vorhanden. Zu den relevanten Daten der Schiffsform gehören:

Parameter	**Format**
Tiefgangsabhängige Daten	
draft	[m]
moulded volume	[m3]
displacement in fresh water	[t]
Unter Berücksichtigung der Dicke der Außenhaut	
displacement in sea water	[t]
LCF from AP	[m]
LCB from AP	[m]
VCB above BL	[m]
WPA - waterplan area	[m2]
moment to alter unit trim	[tm/1cm]
KM	[m]
WPA longitudinal moment of inertia	[m4]
WPA transverse moment of inertia	[m4]
wetted surface	[m2]
block coefficient cb	[-]
waterplane coefficient cw	[-]
midship section area coefficient cm	[-]

Das sehr umfangreiche *Ladungshandbuch* liegt in den Ablieferungsunterlagen der MSC Bari vor. Aus dieser Quelle waren die folgenden Daten im Kontext des Vorhabens relevant:

Parameter	**Format**
light ship weight	[t]
LCG	[m]
TCG	[m]
VCG	[m]
light ship weight distribution, function of LC	[t/m]
draft aft	[m]
draft forward	[m]
lateral area forward	[m2]
lateral area side	[m2]

Zu den relevanten übergeordneten Angaben den/die Propeller betreffend gehören:

Parameter	**Format**
Durchmesser	[m]
Nabendurchmesser	[m]

Anzahl Flügel	[Zahl]
Flächenverhältnis	[-]
Effizienzsteigernde Anbauten	[Zeichenkette]

Für das *Ruder* lag eine Zeichnung in den vorhandenen Ablieferungsunterlagen in einer für die im Rahmen des Vorhabens durchgeführten Analysen erforderlichen Genauigkeit vor. Die daraus abgeleitete geometrische Beschreibung stand in verschiedenen Repräsentationen (Dateiformaten) zur Verfügung.

- **Maschinenanlage**

Diese Kategorie fasst alle wichtigen Angaben zur Hauptmaschine und den Hilfsmaschinen zusammen. Dem vorliegenden Shop Test Report konnten für die Hauptmaschine die folgenden projektrelevanten Daten für die weiteren Analysen entnommen werden:

Parameter	**Format**
Allgemeine Angaben	
main engine type	[Zeichenkette]
main engine serial number	[Zeichenkette]
main engine maximum power MCR	[kW]
main maximum engine speed MCR	[RPM]
main engine nominal power NCR	[kW]
main nominal engine speed NCR	[RPM]
number of auxiliary engines	[Zahl]
auxiliary engine type(s)	[Zeichenkette]
total auxiliary engine power	[kW]
auxiliary engine speed	[RPM]
Shop Test	
date	[UTC]
lower calorific value LCV	[kcal/kg]
Abhängig von der maximalen Leistung	
power % MCR	[%]
power	[kW]
speed	[RPM]
specific fuel oil consumption	[g/kWh]
specific fuel oil consumption ISO conditions	[g/kWh]
Kennfeldbegrenzung, abhängig von der Drehzahl	
speed	[RPM]
maximum power relative to MCR	[kW]

- **Modellversuchsergebnisse**

Modellversuche in Versuchsanstalten spielen eine wesentliche Rolle in der Bewertung der erwarteten Geschwindigkeiten bei bestimmten Kombinationen von Ladezuständen/Tiefgängen und Leistungseinstellungen der Hauptmaschine. Die Daten zu den entsprechenden Versuchen der MSC Bari lagen für das Vorhaben vor. Sie lassen sich weiter untergliedern in allgemeine Daten zu den Versuchen, Modellversuchsergebnisse sowie Prognosen für die Großausführung. Zu letzteren

gehören Widerstands- und Propulsionskoeffizienten, Leistung und Propellerumdrehungen, Wirkungsgrade, Prognosedaten für die Probefahrt.

Im Folgenden sind zunächst die relevanten *allgemeinen Daten*, die im Vorhaben verarbeitet wurden, aufgelistet:

Parameter	**Format**
Administrative Angaben	
Versuchsanstalt	[Zeichenkette]
Berichtsnummer	[Zeichenkette]
Datum des Berichts	[Datum]
beauftragendes Unternehmen	[Zeichenkette]
Modell	
Modellnummer	[Zeichenkette]
Modellmaßstab (λ)	[Zahl]
Ruder Referenz	[Zeichenkette]
Formfaktor (k)	[Zahl]
Nachstromkorrektur für Großausführung	[Zahl]
Propeller	
Propeller Referenz	[Zeichenkette]
Durchmesser	[m]
Steigungsverhältnis (P/D)	[-]
Nabendurchmesser	[m]
Flächenverhältnis (AE/A0)	[-]
Anzahl Flügel	[natürliche Zahl]
Schlepptankbedingungen	
Temperatur	[°C]
kinematische Zähigkeit	[m2/s]
Dichte	[kg/m3]
Angenommene Seewasserbedingungen	
Temperatur	[°C]
kinematische Zähigkeit	[m2/s]
Dichte	[kg/m3]

Zu den zusammenfassenden Aussagen zählen die aus den Modellversuchen abgeleiteten zusammengefasste Prognosen für die Großausführung, d.h. für jede untersuchte Schwimmlage den prognostizierten Leistungsbedarf bei einer spezifizierten Geschwindigkeit.

Parameter	**Format**
Konditionen	
Schiffszustand und Umgebung	[Zeichenkette]
für jede untersuchte Schwimmlage	
Tiefgang achtern	[m]
Tiefgang vorn	[m]

Leistung am Propeller	[kW]
Geschwindigkeit	[kn]

Die konkreten Ergebnisse für die durchgeführten Versuche sowie Ergebnisse im Modellmaßstab und Prognosen für die Großausführung sind Teil der Daten. Diese liegen für jede Kombination aus untersuchtem Beladungszustand/Schwimmlage und Geschwindigkeit mit folgenden Parametern vor:

Parameter	**Format**
Schwimmlage	
Großausführung Tiefgang achtern	[m]
Großausführung Tiefgang vorn	[m]
Großausführung Verdrängung	[m3]
Modellzustand	
(Turbulenzerzeuger, Propeller, Ruder)	[Zeichenkette]
benetzte Lange Großausführung (Los)	[m]
benetzte Oberfläche Großausführung	[m2]
Korrelationskoeffizient Ca	[Zahl]
Propulsionsversuch (Modellmaßstab)	
Modellgeschwindigkeit	[m/s]
Froudezahl	[-]
Reynoldszahl	[-]
Gesamtwiderstand Modell	[N]
Propeller Drehzahl	[1/s]
Propeller Schub	[N]
Propeller Drehmoment	[Nm]
Reibungsabzug	[N]
Koeffizienten (Großausführung)*	
Schiffsgeschwindigkeit	[kn]
Froudezahl (Fn)	[-]
Reynoldszahl (Rn)	[-]
Restwiderstandsbeiwert CR	[-]
Reibungswiderstandsbeiwert CF	[-]
Gesamtwiderstandsbeiwert CT	[-]
unter Berücksichtigung Korrelationskoeffizient	[-]
Großausführung Propellerfortschrittsgrad (Jv)	[-]
Schubbeiwert (KT)	[-]
Momentenbeiwert (KQ)	[-]
Leistung und Propeller (Großausführung)*	
Schiffsgeschwindigkeit	[kn]
Schleppleistung PE	[kW]
Propellerleistung PD	[kW]
Propellerdrehzahl	[RPM]
Verhältnis Reibungs- zu Gesamtwiderstand	[-]

Widerstands-Verdrängungskoeffizient CTVOL	[-]
Leistungs-Verdrängungskoeffizient CDVOL	[-]
Propulsionswirkungsgrad (ηD)	[-]
Wirkungsgrade (Großausführung)*	
Schiffsgeschwindigkeit	[kn]
Propellerfortschrittsgrad	[-]
Sogziffer (t)	[-]
Nachstromziffer (w)	[-]
Schiffseinflussgrad (ηH)	[-]
Propellerfreifahrtwirkungsgrad (η0)	[-]
Schiffseinflussgrad (ηR)	[-]
Propulsionswirkungsgrad ηD	[-]
Schubbelastungsgrad (CTH)	[-]
Probefahrt (Großausführung)**	
Schiffsgeschwindigkeit	[kn]
Gesamtwiderstand (RT)	[kN]
Propellerschub (T)	[kN]
Schleppleistung (PE)	[kW]
Leistung am Propeller (PD)	[kW]
Propulsionswirkungsgrad ηD	[-]
Propellerdrehzahl	[RPM]
Wirkungsgrade Probefahrt (Großausführung)**	
Schiffsgeschwindigkeit	[kn]
Fortschrittsgrad	[-]
Sogziffer (t)	[-]
Nachstromziffer (w)	[-]
Schiffseinflussgrad (ηH)	[-]
Propellerfreifahrtwirkungsgrad (η0)	[-]
Schiffseinflussgrad (ηR)	[-]
Propulsionswirkungsgrad ηD	[-]
Schubbelastungsgrad (CTH)	[-]
Betriebsbedingungen (Großausführung)***	
Schiffsgeschwindigkeit	[kn]
Gesamtwiderstand (RT)	[kN]
Propellerschub (T)	[kN]
Schleppleistung (PE)	[kW]
Leistung am Propeller (PD)	[kW]
Propulsionswirkungsgrad ηD	[-]
Propellerdrehzahl	[RPM]
Wirkungsgrade Betriebsbedinungen (Großausführung)***	
Schiffsgeschwindigkeit	[kn]
Fortschrittsgrad	[-]
Sogziffer (t)	[-]
Nachstromziffer (w)	[-]

Schiffseinflussgrad (ηH)	[-]
Propellerfreifahrtwirkungsgrad (η0)	[-]
Schiffseinflussgrad (ηR)	[-]
Propulsionswirkungsgrad ηD	[-]
Schubbelastungsgrad (CTH)	[-]
Freifahrtversuch[****]	
Fortschrittsziffer (J)	[-]
Schubbeiwert (KT)	[-]
Momentenbeiwert (KQ)	[-]
Freifahrtwirkungsgrad (η0)	[-]

[*]ohne Korrekturen für Nachstrom und Propellerwirkungsgrad

[**]kein Wind und kein Seegang mit Korrekturen für Nachstrom und Propellerwirkungsgrad

[***]Zuschlag auf Probefahrtbedingungen: +15% auf PD (trial) mit Korrekturen für Nachstrom und Propellerwirkungsgrad

[****]jeweils für die drei Bedingungen: Freifahrtversuch, Freifahrtversuch mit Korrekturen für Reynoldszahl im Propulsionsversuch, Freifahrtversuch mit Korrekturen für Reynoldszahl der Großausführung, abhängig von der Fortschrittsziffer J.

- **Probefahrtergebnisse**

Das Protokoll der Geschwindigkeit-/Leistungsmessungen liegt für den Versuchsträger vor. Daten relevant für die Auswertungen im Vorhaben sind:

Parameter	**Format**
Allgemeine Angaben	
Berichtnummer	[Zeichenkette]
Datum der Probefahrtmessungen	[Datum]
Tiefgang achtern	[m]
Tiefgang vorn	[m]
Deplacement	[m3]
benetze Oberfläche	[m2]
Windlateralfläche von vorn	[m2]
Windlateralfläche von der Seite	[m2]
Windwiderstandskoeffizienten CAA als Funktion der scheinbaren Richtung	[°], [-]
Wetterbedingungen	[Zeichenkette]
Wassertiefe	[m]
Wassertemperatur	[°C]
Lufttemperatur	[°C]
Messergebnisse[*]	
Kennzeichen Lauf	[Zeichenkette]
Geschwindigkeit	[kn]
Maschinenleistung	[kW]
Drehzahl	[RPM]
gemittelte Geschwindigkeit	[kn]
gemittelte Maschinenleistung	[kW]

gemittelte Drehzahl [RPM]

*für jeden Messlauf einschließlich der Korrekturen nach ISO 15016 [6]

II.1.2 Verarbeitungssystem für Schiffsbetriebsdaten (AP 2000)

Beschreibung der Systemarchitektur

Die Implementierung des Informationsmodells zur Schiffsbetriebsdatenauswertung ermöglicht den fortlaufenden Import von Schiffsbetriebsdaten und deren Auswertung mit dem Ziel der Optimierung ds Schiffsbetriebs. Hierzu wurde von der Universität Rostock eine Systemarchitektur mit vier klar voneinander abgegrenzten Komponenten entwickelt und umgesetzt:

1. die Informationsmodellierung mit der Konfiguration der Datenverwaltung auf der Grundlage eines Datenbankmanagementsystems,
2. das Datenbankmanagementsystem in Form einer cloud-basierten Datenverwaltung mit spezifischen Datenmanipulationsmethoden für die Anwendungsprogrammierung,
3. die Parser für Daten aus verschiedenen Quellen zur Vorfilterung und für den Import in die integrierte Datenverwaltung sowie
4. die vielfaltigen Anwendungsprogramme zur Datenauswertung mit jeweils spezifischen Zielstellungen.

Abbildung 4 zeigt die vier Komponenten, die im Folgenden detailliert beschrieben werden, und deren Verbindungen mit einem höheren Abstraktionsgrad.

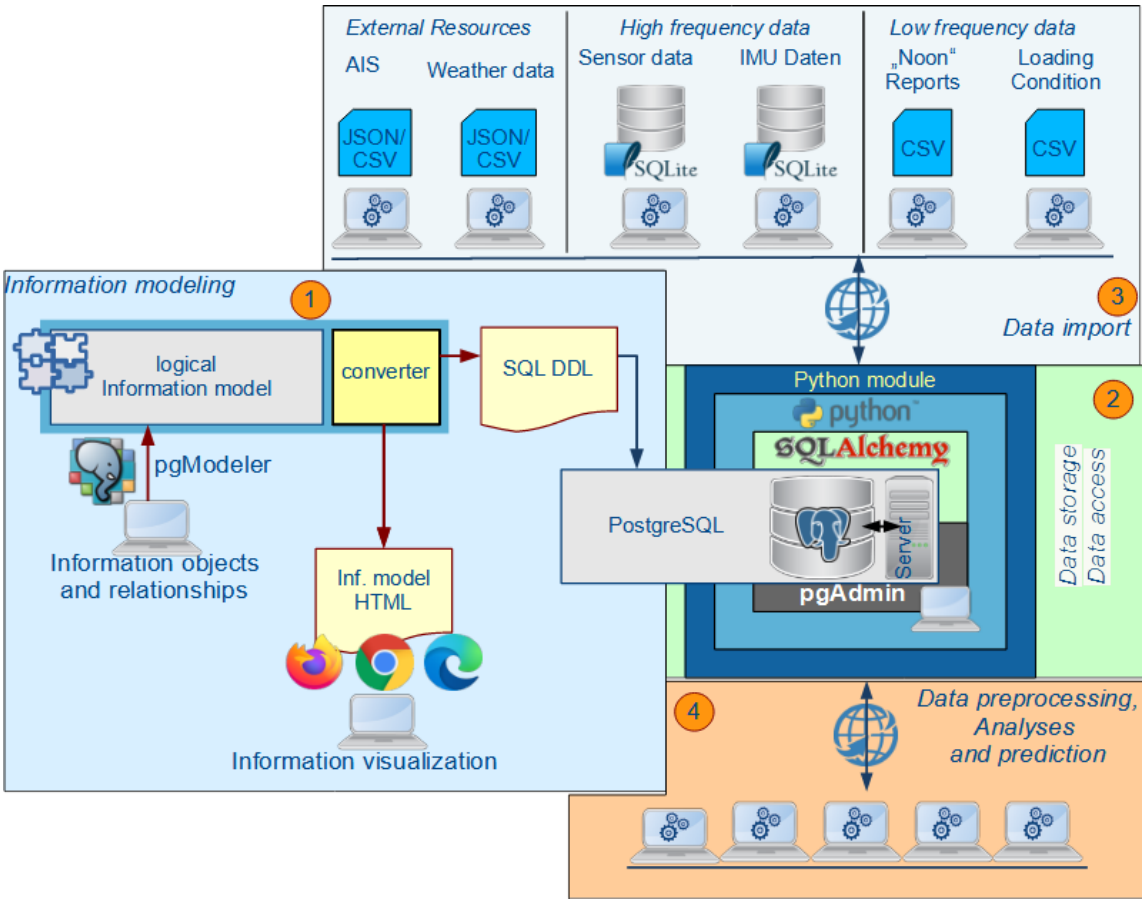


Abbildung 4: DigitShip Systemarchitektur mit Kennzeichnung der eingesetzten Softwarewerkzeuge

(1) Informationsmodellierung

Die Informationsmodellierung ist der erste Prozessschritt, mit dem die Konfiguration der Datenverwaltung für die applikationsspezifische Entwicklungsumgebung vor allen weiteren Aktivitäten erfolgt. Die Modellierung im Rahmen des Vorhabens wurde auf Basis der Softwareumgebung pgModeler [7] durchgeführt. Diese erlaubt die Formulierung eines logischen Informationsmodells, welches alle Objekte und deren Beziehungen zueinander formal beschreibt. Im Gegensatz zur klassischen Notation nach Chen kommt hierbei die Crow's Foot Notation [8] zum Einsatz, die ebenfalls dazu dient, ein ERM formal zu beschreiben und zu visualisieren. Diese Notation unterscheidet sich von der erstgenannten im Wesentlichen dadurch, dass Beziehungen nur durch benannte Verbindungslinien dargestellt werden, und die Kardinalitäten nicht durch Zahlen, sondern durch Symbole dargestellt werden.
Darüber hinaus ist die pgModeler Software in der Lage, das formulierte logische Modell in verschiedenen Ausgabeformaten für die weitere Verwendung zu exportieren. Im Rahmen des Vorhabens war insbesondere die Ausgabe in das SQL DDL Format relevant, um damit in einem weiteren Schritt das Datenbank-Managementsystem zu konfigurieren, und damit das physikalische Modell zu definieren. Ein weiteres Exportformat ist HTML, womit das Modell mit Webbrowsern visualisiert und in diesem navigiert werden kann. Dies war besonders für die Kommunikation über die Systemarchitektur während der Entwicklung sowie der Anwendungsprogrammierung hilfreich, da dadurch keine weiteren Spezialprogramme notwendig waren.
Das wichtigste Ergebnis der Informationsmodellierung stellt somit die Datendefinition "DigitShip SQL DDL" dar, mit dessen Hilfe die Datenverwaltung konfiguriert wurde.

(2) Datenbankmanagementsystem

Als Datenbankmanagementsystem wurde im Rahmen des Vorhabens PostgreSQL [9] verwendet. Die Software pgAdmin [10] diente der User Administration und der Konfiguration wichtiger Parameter der Datenbank. Darüber hinaus können damit interaktive Abfragen den Datenbestand betreffend erfolgen, um diesen z.B. in Bezug auf definierte Fragestellungen überprüfen zu können. Aufgrund der damit verbundenen umfassenden Änderungsfunktionen sind hierfür besondere Rechte erforderlich, die über das Rollensystem des Datenbankmanagementsystems kontrolliert werden.
Bei dem physikalischen Modell der Datenbank handelt es sich um ein relationales Modell mit Ergänzungen in Richtung eines objektorientierten Modells. Alle Informationen werden hierbei in Form von Tabellen gespeichert, deshalb auch als objektrelational bezeichnet. Die Überführung des ER-Modells in eine Tabellenrepräsentation erfolgt automatisch durch das vorgeschaltete Programm pgModeler, wobei auch die formulierten Konsistenzbedingungen an den Datenbestand in Form von diesen entsprechenden SQLProgrammsequenzen mitberücksichtigt, und in die Datenbank übertragen werden.
Der Zugriff auf den Datenbestand aus den Anwendungsprogrammen erfolgt hingegen objektorientiert, was zu einer deutlichen Effizienzsteigerung in der Programmierung führt. Hierfür wurde ein Object-Relational-Mapping (ORM) eingesetzt, welches in diesem Fall mit Hilfe der Software SQLAlchemy [11] für die Anwendungsprogrammiersprache Python realisiert wurde. Um weitere Methoden ergänzt stehen damit Python Module zur Verfügung, die die Grundlage für die Programmierung aller Anwendungen bildeten: sowohl für den Import der Schiffsbetriebsdaten

aus den verschiedenen Quellen als auch für die Analysen der Daten unter beliebigen Fragestellungen. In der *DigitShip-Python-Module-Schicht* werden die im Informationsmodell definierten Konsistenzbedingungen an den Datenbestand überprüft, und bei Verstoß eine Eintragung der Daten in die Datenbank verweigert. Dies bewirkt eine erhebliche Steigerung der Qualität und ist als Grundlage zum sicheren Umgang mit den Daten anzusehen. Die Überprüfung der Konsistenzbedingungen in diesen Fällen ist nicht zu verwechseln mit den vielfaltigen Algorithmen, die die Qualität, Vollständigkeit und Sinnhaftigkeit der Schiffsbetriebsdaten an sich bewerten.

(3) Datenimport

Der Import der Schiffsbetriebsdaten sowie der Referenzdaten des Schiffes als auch weiterer statischer Daten erfolgt durch hierfür jeweils speziell angepasste Schnittstellenprogramme. Der Zugriff auf die Datenbank wird dabei ebenfalls über die *DigitShip-Python-Module-Schicht* realisiert. Hierbei werden Daten aus den folgenden Quellen importiert, und in das Informationsmodell integriert:

- Referenzdaten
- Sensordaten von Bord über SQLite [12] Datenbanken
- Schiffsbewegungsdaten von Bord über SQLite Datenbanken
- Berichtsdaten aus der navis-Cloud über csv-formatierte Dateien
- Ladungsrechnerdaten aus der navis-Cloud über csv-formatierte Dateien
- von externen Anbietern bezogene AIS Daten über csv-formatierte Dateien
- von externen Anbietern bezogene Umweltdaten über csv-formatierte Dateien

(4) Datenvorverarbeitung, Analyse und Vorhersage

Die Vorbearbeitung, Analyse, Auswertung und Vorhersage der in der Datenbank gespeicherten Daten mit Hilfe der Programmiersprache Python erfolgt ebenfalls über die *DigitShip-Python-Module-Schicht*. Die Funktionalitäten dieser Komponente werden in den folgenden Abschnitten näher erläutert.

II.1.2.1 Schiffsbetriebsinformationsmodell (SOIM)

Die Informationsmodellierung und die darauf aufbauende Informationsverarbeitung erfolgt in klar definierten Prozessschritten, siehe Abbildung 5.
Zu diesem Zweck wurden im Vorhaben standardisierte Softwarekonzepte verwendet. Zu diesen Konzepten gehören zum einen die Modellierung mittels einen von Chen [13] entwickelten und speziell erweiterten Entity-Relationship-Modells mit Hilfe des von Silva [7] entwickelten Werkzeugs pgModeler und zum anderen die Verwendung einer objektrelationalen Datenbank, in diesem Fall des Datenbankmanagementsystems PostgreSQL.

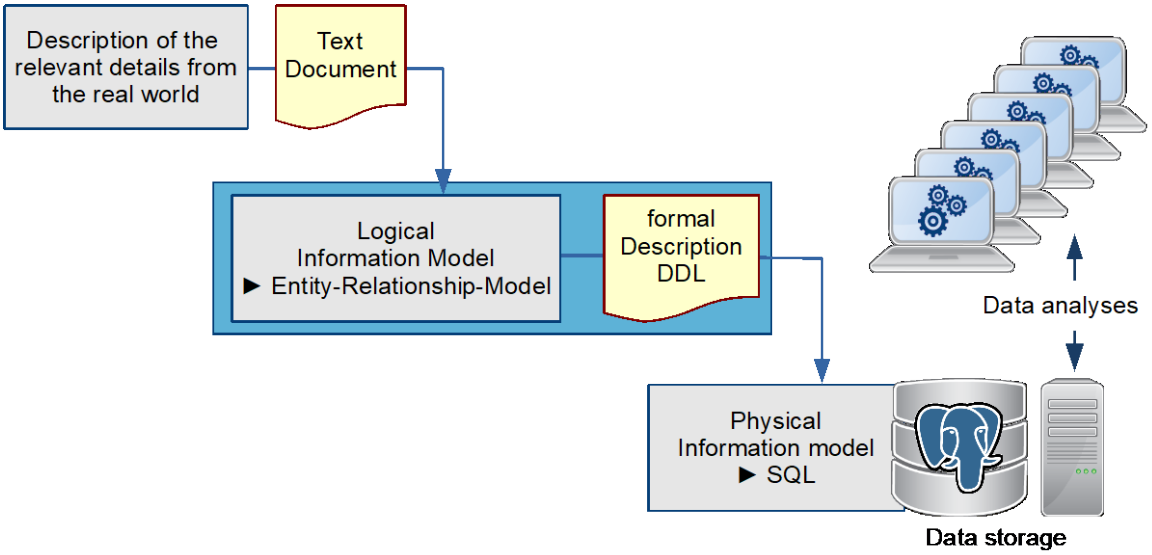


Abbildung 5: Schematische Darstellung der Schritte der Informationsmodellierung und anschließenden Datennutzung

Das logische Schiffsbetriebsinformationsmodell wurde aus der zuvor erwähnten textuellen Beschreibung der Schiffsbetriebsdaten abgeleitet. Die Informationen sind verschiedenen Untermodellen (Schemata) zugeordnet, die unterschiedliche Datenkategorien widerspiegeln, aus denen sich die Struktur des logischen Informationsmodells der obersten Ebene ableiten lässt. Die Organisation des Schiffsbetriebsmodells in einzelne Schemata, wie sie im Folgenden aufgeführt sind (siehe Abbildung 6) und in denen das *ship_operation_core_model* eine zentrale Rolle spielt, hat viele Vorteile auf der Prozessebene der Verwaltung der verschiedenen Datentypen.

- ship_operation_core_model: Enthält Daten bezogen auf spezifische Schiffsreisen mit Informationen über die jeweiligen Reiseabschnitte, deren zugeordnete Betriebsperioden und -zustände und besonderen Ereignissen sowie Positions- und Zeitangaben.
- board_data_low_frequency: Enthält Daten zu ereignisgesteuerten Meldungen sowie zu den jeweiligen Ladezuständen.
- board_data_high_frequency: Enthält Daten aus Messungen durch Bordsensorik, sowie die Schiffsbewegungsdaten der zusätzlich installierten IMU.
- external_resources: Enthält Daten aus externen Quellen – im Vorhaben waren dies AIS und Umweltdaten.
- general_resources: Enthält allgemeine Informationen, die mehrfach verwendet werden können, wie z. B. Verwaltungsdaten.
- Reference_data: Enthält Schiffsreferenzdaten, die alle relevanten Informationen über das Schiff und seine Systeme enthalten, insbesondere Maschinendaten, Modellversuchs- und Probefahrtergebnisse.

Da die den Schemata Reference_data und general_resources zugeordneten Daten statischer Natur sind und sich während einer Schiffsreise nicht ändern, werden im Folgenden nur die vier dynamischen Schemata ausführlich beschrieben.

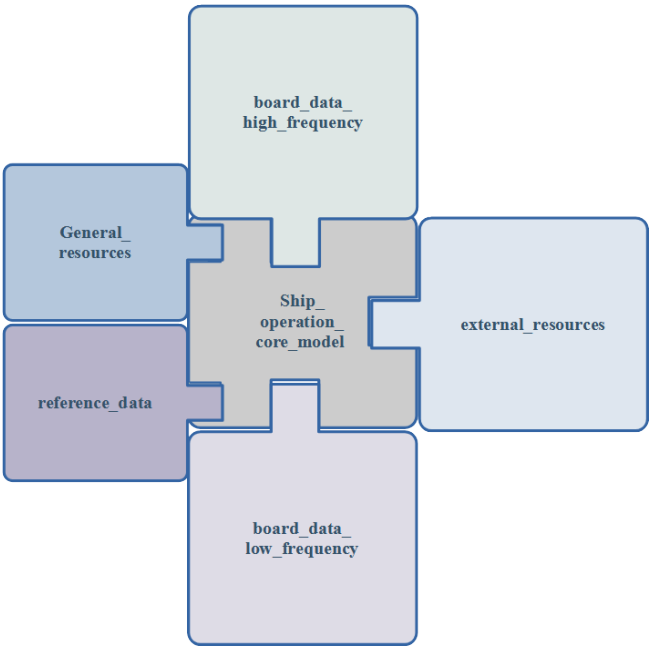


Abbildung 6: Übersicht der Schemata des DigitShip Schiffsbetriebsinformationsmodells

Kernmodell (ship_operation_core_model)

Unter *voyage logistic* werden im Rahmen des Vorhabens alle Informationen verstanden, die den Schiffsbetrieb als Funktion der Zeit abbilden. Diese Daten stellen die Basis für alle weiteren Informationen dar. Die Einführung des Begriffs *fleet* (Flotte) dient dazu, eine Anzahl von *ships* (Schiffe) nach einem beliebigen, vom Benutzer festgelegten Ordnungsprinzip zu gruppieren. Bei Bedarf können auch mehrere Flotten gleichzeitig verwaltet werden. Das Schiff ist das zentrale Informationsobjekt, dem alle gültigen Informationen zugeordnet sind. Dazu gehören neben einigen Verwaltungsdaten zum einen die bereits erwähnten statischen Referenzdaten und zum anderen Daten aus der Liste der Reisen, die das jeweilige Schiff unternommen hat. Während die Referenzdaten als zeitunabhängig angenommen werden, sind die Reisedaten, d.h. die Zeitspanne vom Verlassen des Starthafens bis zur Ankunft im Zielhafen, zeitabhängig.

Innerhalb des Kernmodells wird der Zeitstrahl einer Schiffsreise in einzelne Abschnitte unterteilt, die sequenziell definiert sind. Abbildung 7 verdeutlicht dieses Prinzip anhand eines Beispiels. Auf der obersten Ebene sind *voyages* (Reisen) nahtlos sequenziell angeordnet, diese sind wiederum in nahtlose Abschnitte, sog. *legs* unterteilt. Jeder dieser Abschnitte stellt den Teil einer Reise zwischen zwei aufeinanderfolgenden Abfahren aus Häfen dar. Auf der dritten Ebene wird ein *leg* basierend auf den Betriebszuständen in aufeinanderfolgenden Zeitspannen (*operation periods*) unterteilt. Abbildung 7 zeigt ein Beispiel für die Betriebszustände der Etappe 3 der Reise n+1. Diese Etappe beginnt mit einer kurzen Manövrierphase nach dem Auslaufen aus dem Hafen und geht dann in eine längere Seepassage über. Die Seepassage wird durch eine kurze Driftphase unterbrochen, nach der das Schiff seine Reise zum Zielhafen fortsetzt. Nach einer weiteren Driftphase ankert das Schiff und manövriert nach einer Wartezeit in den Zielhafen. Im Zielhafen verbringt das Schiff die letzte Zeit der Etappe, bevor es wieder zu einer neuen Etappe aufbricht. Dies zeigt, dass eine Strecke immer mit der Abfahrt beginnt und mit der nächsten Abfahrt aus dem Zielhafen endet. Die Hafenzeit stellt also stets die letzte Phase einer Etappe dar.

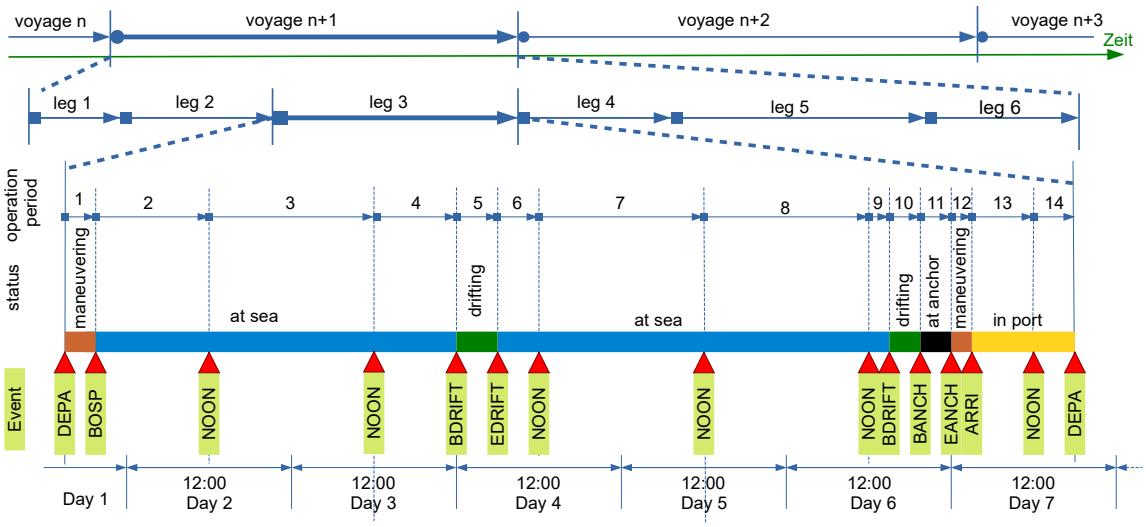


Abbildung 7: Prinzip der datentechnischen Gliederung einer Schiffsreise

Die Gesamtdauer einer Reise kann durch einen Fahrplan vorgegeben sein, wie es im transkontinentalen Containerliniendienst üblich ist. Die Dauer der Teilstrecken wird durch die Entfernungen und die Schiffsgeschwindigkeit zwischen den angelaufenen Häfen bestimmt. Die verschiedenen Bedingungen und ihre jeweilige Dauer (Betriebszeiten) hängen von den tatsächlich vorherrschenden Bedingungen ab, insbesondere von den Wetterbedingungen, die eine Änderung der Route gegenüber dem ursprünglichen Plan erforderlich machen können.

Zusätzlich zu den oben erwähnten Informationen über Zeitphasen und Betriebszustände sind die im vorherigen Abschnitt erwähnten Ereignisse wichtig, da sie zusätzliche Informationen auf der Zeitachse darstellen. Dazu gehören die beiden Ereignisse, die einen zeitlichen Abschnitt eines bestimmten Betriebszustandes abgrenzen: das (Start-Ereignis) und das (End-Ereignis). Neben der Zeitangabe enthalten die Ereignisse auch Informationen über den Ort, an dem sich das Schiff zum Zeitpunkt des Ereignisses befand.

Die oben skizzierten Informationsobjekte zur Beschreibung des Schiffsbetriebs im Rahmen des Vorhabens sind in einem Entity-Relationship-Modell (ERM)-Diagramm nach Chen [13] in Abbildung 8 dargestellt.

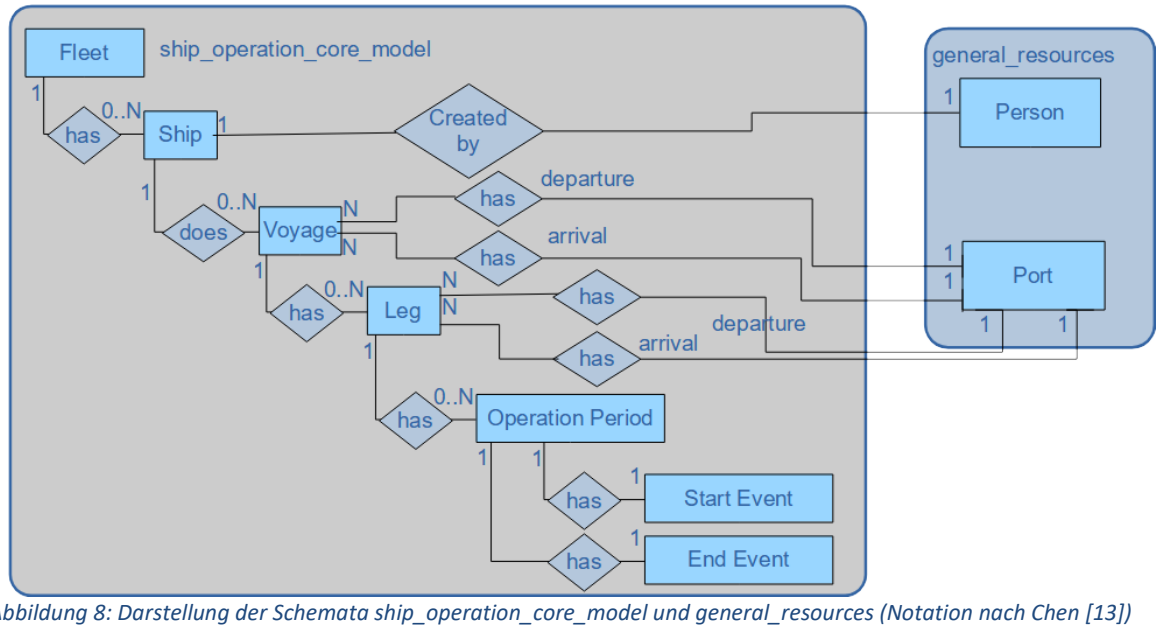


Abbildung 8: Darstellung der Schemata ship_operation_core_model und general_resources (Notation nach Chen [13])

Die Beziehung zwischen einer Flotte und den zugehörigen Schiffen ist ein einfaches Beispiel dafür, wie die Notation mit den Richtungen der Beziehung zu interpretieren ist: eine Flotte besteht aus einer beliebigen Anzahl von Schiffen (0..N), ein Schiff muss mit genau einer (und nicht mit keiner oder mehreren) Flotte verbunden sein (1). Die Implementierung in der Datenbank ist so gestaltet, dass viele der Beziehungen und damit ihre Objekte optional sind. Im Allgemeinen können Objekte, die auf einer niedrigeren Ebene in der modellierten Baumstruktur existieren, nicht existieren, ohne dass ihre zugehörigen Objekte auf den höheren Ebenen existieren; umgekehrt gilt dies nicht. Beispiel: eine bestimmte Strecke kann nur existieren, wenn die zugehörige Reise existiert, diese Reise kann nur existieren, wenn das zugehörige Schiff existiert, und dieses Schiff muss mit einer Flotte verbunden sein. Andererseits können z.B. mehrere Reisen definiert werden, ohne sie weiter in Unterreisen (Strecken) zu unterteilen (0..N).

Allgemeine Informationen (general_resources)

Hierzu zählen die statische Informationen, die für die Schiffsbetriebsoptimierung relevant sind. Darunter fallen administrative Angaben zur Informationsspeicherung und zu Häfen. Diese beiden Konzepte, die im Folgenden ausführlich beschrieben werden, wurden in den beiden Objekten "Person" und "Port" modelliert und verfügbar gemacht (siehe Abbildung 8).

Administration

Zu den im Rahmen des Vorhabens genutzten administrativen Daten gehören Angaben zur Person (Name, mit dem der Zugang zum System erfolgte), die die entsprechend gekennzeichneten Daten in die Datenbank einfugt hat, sowie der Zeitpunkt, an dem dies erfolgt ist (Hierbei handelt es sich nicht um den Zeitpunkt, an dem die Daten entstanden sind).

Damit wird nachvollziehbar, wer welche Daten wann für die weitere Nutzung freigegeben hat. Das Rollen-System sowie die Zeitfunktionen des Datenbankmanagementsystems kommen hierfür zum Einsatz, wodurch die Angaben automatisiert in die Datenbank eingetragen werden konnten. Ferner wurde damit sichergestellt, dass personenbezogene Zugriffsrechte definiert werden konnten, was in erster Linie dem unbeabsichtigten Löschen von Daten vorbeugen sollte.

UN/LOCODE

Der Ortscode der Vereinten Nationen für Handel und Transport (UN/LOCODE, United Nations Code for Trade and Transport Locations) ist ein standardisierter Code für die eindeutige Kennzeichnung geographische Orte. Alle Seehäfen, zumindest die, die von Fracht- oder Passagierschiffen angelaufen werden, können über eine fünfstellige Buchstabenkombination beschrieben werden [5]. Dem Code werden weitere Daten den Hafen betreffend zugewiesen. Hierzu gehören neben Informationen zur Infrastruktur immer die beiden Angaben LAT, LON im geographischen Koordinatensystem. Der Hafen von Rostock ist z.B. nach diesem Schema mit ”DERSK” eindeutig referenziert.

Häfen werden in der Seefahrt üblicher Weise mit Hilfe des LOCODE angegeben. Im Rahmen des Vorhabens wird dieser Methode gefolgt. Der vollständige Name und die Koordinaten LAT und LON Ort können bei Bedarf in Anwendungsprogrammen von einer in der Datenbank gespeicherten Referenztabelle abgefragt werden.

Referenzdaten (reference_data)

Unter Referenzdaten werden alle relevanten Informationen verstanden, die ein Schiff und seine Systeme Komponenten im Ablieferungszustand beschreiben bzw. die Daten, die in

Modellversuchen und der Probefahrt ermittelt wurden. Diese Daten haben sind statisch, d.h. sie sind unabhängig vom betrachteten Zeitpunkt und können in die folgenden Kategorien untergliedert werden:

- Hauptdaten (general characteristics) des Schiffes
- Maschinenanlage (machinery): Ergebnisse der Shop Tests
- Modellversuche (model tests): in Modellversuchen gemessene hydrodynamische Kenngrößen und daraus abgeleitete Prognosen für die Performance der Großausführung
- Probefahrtergebnisse (sea trials): gemessene Ergebnisse am werftneuen realen Schiff nach Fertigstellung, beinhaltet die erforderliche Maschinenleistung in Abhängigkeit von der Schwimmlage und der Schiffsgeschwindigkeit

Die Objektstruktur mit der Einbindung in die Grundstruktur des Informationsmodells kann in Abbildung 9 nachvollzogen werden. Die Notation folgt hier ebenfalls der für ein ERM Diagramm nach Chen [13]. Die Implementierung in die Datenbank erfolgte so, dass Referenzdaten verwaltet werden können, aber nicht müssen – sie sind somit optional. Die Verknüpfung von Referenzdaten erfolgt direkt mit dem jeweiligen Schiff.

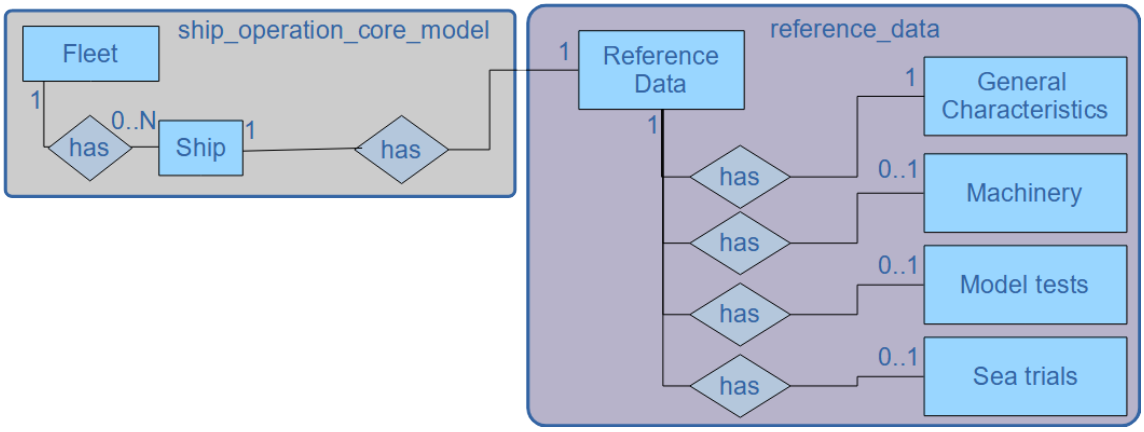


Abbildung 9: Schematische Darstellung des reference_data Modells (Notation nach Chen [13])

Modell der niederfrequenten Daten (board_data_low_frequency)

Die in Abschnitt II.1.1.1 beschriebenen niederfrequenten Daten wurden in einem eigenen Schema modelliert (siehe Abbildung 10). Zu diesen Daten gehören die von der Besatzung an Bord manuell erstellten Berichte und die automatisch aus aggregierten Sensoraufzeichnungen gewonnenen Messdaten sowie die Daten über den Beladungszustand, die in der Regel bei jedem Auslaufen des Schiffes aus einem Hafen aktualisiert werden.

Da Berichte zu bestimmten Ereignissen an Bord erstellt werden, beziehen sie sich auf den Zeitraum zwischen dem vorangegangenen (Start-)Ereignis und dem Ereignis, das den Bericht auslöst (Endereignis). Berichte werden daher über das zugehörige Endereignis genau einem Berichtszeitraum (Betriebszeitraum) zugeordnet. Einem Berichtszeitraum kann genau ein manuell (gemeldeter) und zusätzlich ein automatisch (gemessener) generierter Bericht zugeordnet werden. Abbildung 7 zeigt am bereits erwähnten Beispiel eines Fahrtabschnitts die aufeinanderfolgenden Betriebszustände (hier 14), für die jeweils ein Bericht erzeugt wird. So wird beispielsweise der vierte Bericht am dritten Tag der Reise zu dem Zeitpunkt erstellt, an dem das Schiff zu driften beginnt, so dass der Berichtszeitraum vom letzten Noon Report bis zu diesem Zeitpunkt reicht. Daraus ergeben sich unterschiedlich lange Berichtszeiträume, die jedoch 24 Stunden (Noon to Noon) nicht überschreiten. Die Beladungszustandsdaten hingegen werden

aufgrund ihrer Zuordnung zum Ereignis immer genau einer Teilstrecke der Reise zugeordnet: DEPA, wie bereits erwähnt. Wie in Abbildung 8 zu sehen ist, können jedem *leg* (1) mehrere Belastungszustände (0..N) zugeordnet werden.
Sämtliche Daten sind mit einem UTC-Zeitstempel (Universal Time Coordinated) sowie der jeweiligen Ortszeit versehen. Dies ermöglicht die Synchronisation unterschiedlicher Daten zu einem Schiff in der Datenbank, was wiederum erhebliche Vorteile bei der Auswertung von Schiffsbetriebsdaten mit sich bringt.

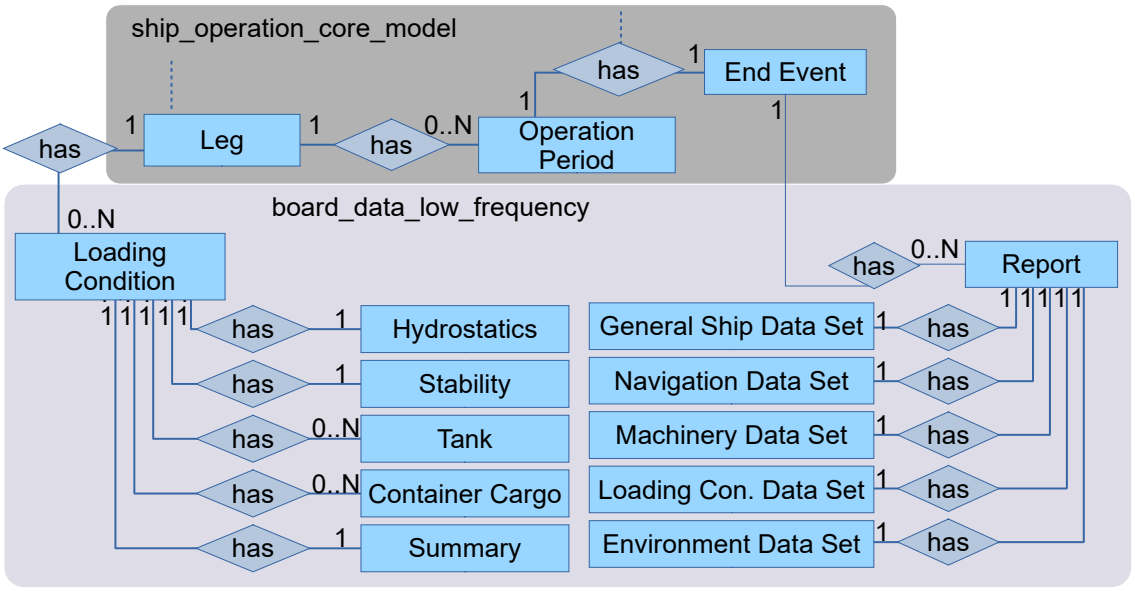


Abbildung 10: Schematische Darstellung des board_data_low_frequency Modells (Notation nach Chen [13])

Modell der hochfrequenten Daten (board_data_high_frequency) und externen Daten (external_resources)

Die schematische Darstellung in Abbildung 11 verdeutlicht die Modellstruktur für hochfrequente Bordmessdaten und Daten, die von externen Anbietern bezogen werden sowie deren Zuordnung zum jeweiligen Betriebszeitraum:
Wie in Abschnitt II.1.1.2 beschrieben, umfassen die hochfrequenten Daten zum einen die von den Sensoren an Bord gemessenen und für das Projekt relevanten Daten und zum anderen die Bewegungsdaten des Schiffes, die von einer eigens für das Forschungsprojekt installierten Inertial Measurement Unit (IMU) aufgezeichnet wurden. Die hochfrequenten Daten wurden entsprechend ihrer Messintervalle (15 min, 1 min und 200 Millisekunden) modelliert und gespeichert.
Zu den externen Daten zählen wie in Abschnitt II.1.1.3 beschrieben, insbesondere AIS- und Umweltdaten.

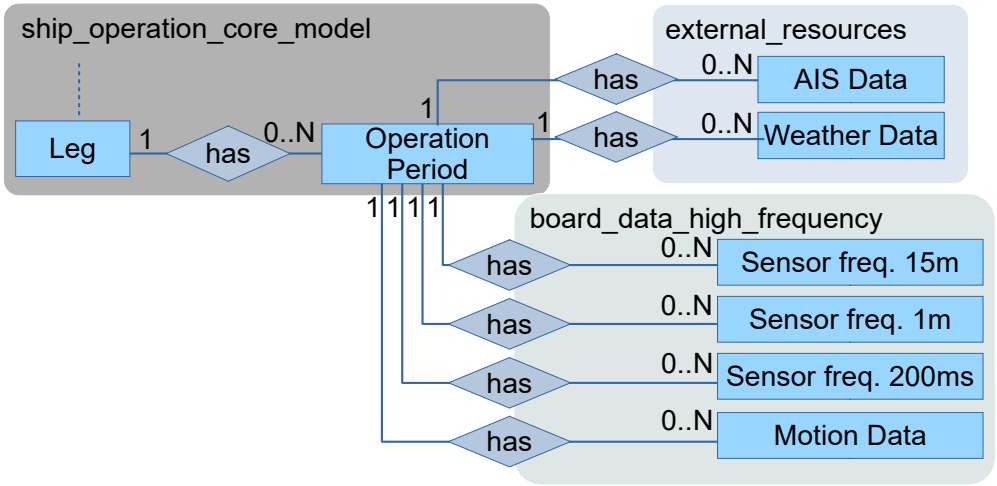


Abbildung 11: Schematische Darstellung der Modelle board_data_high_frequency und external_resources (Notation nach Chen [13])

II.1.2.2 Vorverarbeitung, Analyse und Bewertung der Schiffsbetriebsdaten

Vorverarbeitung der Daten

Um die verfügbaren gespeicherten Rohdaten aus dem Schiffsbetrieb für weiterführende Analysen verwenden zu können, müssen sie Vorverarbeitungsprozessen unterzogen werden, um ihre Qualität zu erhöhen und damit die Vertrauenswürdigkeit der darauf aufbauenden Analysen zu steigern.

Die Vorverarbeitung von Rohdaten aus dem Schiffsbetrieb ist jedoch mit einigen Herausforderungen verbunden. Diese lassen sich grob in zwei Hauptgruppen unterteilen:

1. Datenform: Da sich Betriebsdaten aus unterschiedlichen Quellen hinsichtlich ihres Formats und ihrer Frequenz (niedrig, mittel und hoch) unterscheiden, erfolgt zunächst eine Synchronisierung und anschließende Umwandlung in ein einheitliches Datenformat, um eine holistische und integrierte Verarbeitung dieser Daten zu ermöglichen.
2. Dateninhalt: Schiffsbetriebsdaten eignen sich in ihrer Rohform aufgrund einiger Probleme nicht zur Datenanalyse. Dazu zählen insbesondere
    - Fehlende Daten: Dazu zählen zu sog. „Null“-Daten auf vollständig fehlende Einträge, die unter anderem durch Sensorausfälle, Datenübertragungsprobleme und menschliches Versagen verursacht werden können.
    - Ausreißer: Ein Ausreißer ist ein Datenpunkt, der aus nicht-physikalischen Gründen erheblich vom Rest der übrigen Daten abweicht.

Diesen Herausforderungen wird innerhalb der DigitShip-Systemarchitektur in Schritt 4 (siehe Abbildung 4) durch spezielle Algorithmen zur Filterung und Imputation von Daten begegnet.

Datenfilterung

Die Filterung ist der erste Schritt bei der Vorverarbeitung der Daten. In diesem Schritt werden die verfügbaren Rohdaten aus dem Schiffsbetrieb überprüft, um die zuvor besprochenen inhaltlichen Probleme zu identifizieren, einschließlich fehlender Daten, „Nullen“ und Ausreißer. Diese Probleme können durch ungewollte Einflüsse aus der Umgebung oder ein fehlerhaftes Messgerät verursacht werden; in diesem Fall handelt es sich um fehlerhafte Datenpunkte, die bearbeitet werden sollten.

Andererseits kann ein Ausreißer auch durch natürliche Schwankungen entstehen. Unter der

Annahme, dass die beobachteten Daten einer Normalverteilung folgen, besteht die Möglichkeit, dass einige Datenpunkte stark von dieser Verteilung abweichen. Bei der nachtäglichen Betrachtung von Schiffsbetriebsdaten ist es unwahrscheinlich, dass man die Quelle dieser Abweichung identifizieren kann. Daher werden Ausreißer, „Nullen“ fehlende Daten entfernt und als fehlende Werte behandelt, die dann anschließend wie im folgenden Abschnitt beschrieben aufgefüllt (imputiert) werden. Es wird davon ausgegangen, dass Ausreißer aus Fehlmessungen resultieren. Da die fehlerhaften Messungen unabhängig von den einzelnen Messgeräten sind, sollten die Ausreißer univariat sein. Daher wird für jede Messgröße eine univariate Ausreißererkennung (wie von Tukey [14] beschrieben) durchgeführt.

Datenimputation

Wie bereits erwähnt, sollten nach der Filterung der operativen Rohdaten und vor Beginn der Datenanalyse die fehlenden Werte behandelt werden. Die einfachste Lösung zur Überwindung dieses Problems besteht darin, die Datenpunkte mit fehlenden Werten vollständig zu ignorieren und die Analyse mit einer Teilmenge des Datensatzes durchzuführen, was als listenweise Löschung oder vollständige Fallanalyse bezeichnet wird. Raaijmakers et al. [15] stellten jedoch fest, dass eine listenweise Löschung von 10% der Daten zu einer Reduzierung der statistischen Aussagekraft um 35 % führt und sogar um bis zu 98% wenn 30% der Daten fehlen sollten. Moderne Imputationsmethoden (bspw. Mean Imputation, Linear Regression, KNN) erhalten die statistischen Eigenschaften der Daten, sofern nicht mehr als 40% der Daten fehlen, wie Lodder [16] feststellte.

In Rahmen dieses Vorhabens wurden verschiedene Imputationsmethoden implementiert, um die abgeleiteten Ergebnisse (imputierte Daten) zu vergleichen und die am besten geeignete Methode für den betrachteten Datensatz zu ermitteln. Im Folgenden werden die angewandten Methoden kurz beschrieben:

- Mean Imputation: Sie wird als univariate Imputationsmethode eingestuft. Bei der Mean Imputation wird ein fehlender Wert mit dem Mittelwert aus dem Datensatz oder einer definierten Teilmenge der Daten aufgefüllt, wie es von Mohamed et al. [17] beschrieben wird.
- Lineare Regression: Hierbei handelt es sich ebenfalls um eine univariate Imputationsmethode (vgl. Wang et al. [18]). Um einen fehlenden Wert aufzufüllen, werden vorhandenen Variablen oder der letzte und der nächste beobachtete Wert der selben Variable verwendet, um eine Vorhersage über den fehlenden Wert zu treffen. Dies setzt jedoch lineare Zusammenhänge zwischen den betrachteten Variablen bzw. zeitlich lineares Verhalten der selben Variable voraus.
- KNN-Imputation: Bei dieser Methode handelt es sich um eine multivariate Imputation, die ursprünglich eingeführt wurde, um die Qualität von Trainingsdatensätzen für maschinelles Lernen zu verbessern, wie von Batista et al. [19] erläutert. Bei der KNN-Imputation wird der gesamte Datensatz als Trainingsdaten verwendet. Um einen fehlenden Wert auf der Metrik M in einem Datenpunkt P zu ersetzen, sucht die KNN-Imputation nach den k (ein ganzzahliger Parameter, der angegeben werden kann) nächstgelegenen Nachbarschaften von P, in denen die Werte auf M beobachtet werden. Dann wird der fehlende Wert mit dem Mittelwert oder dem gewichteten Durchschnitt der Nachbarn aufgefüllt.

Beispiel

Die Funktionalität der für die Vorverarbeitungsphase implementierten Algorithmen wird im Folgenden anhand eines konkreten Beispiels demonstriert. Die Rohdaten des Kraftstoffverbrauchs der Hauptmaschine (FOC ME) für eine Reise zwischen Europa und China der „MSC Bari“ werden aus der im Vorhaben entwickelten Datenbank gezogen. Diese Daten stammen ursprünglich aus den zuvor beschriebenen „Noon Reports“, die von der Besatzung manuell eingegeben wurden.

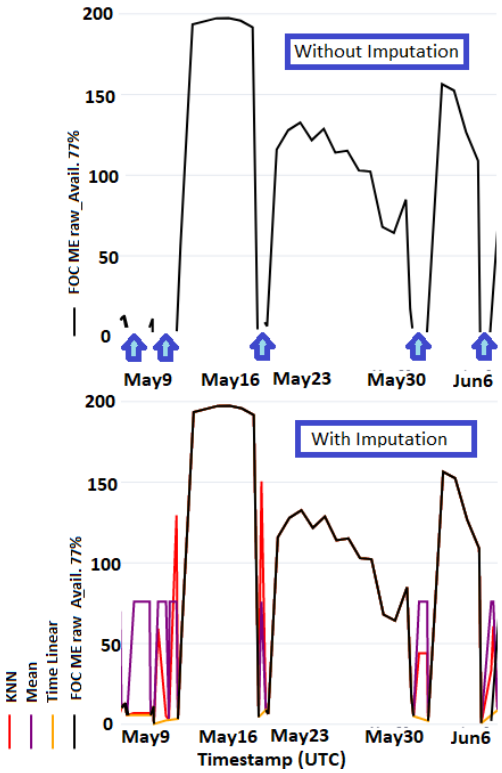


Abbildung 12: Manuell abgelesener Kraftstoffverbrauch der Hauptmaschine vor (oben) und nach der Datenimputation mit unterschiedlichen Verfahren (unten)

Abbildung 12 (oben) zeigt diese Daten, nachdem sie dem oben beschriebenen Filterprozess unterzogen wurden. Wie man sieht, beträgt die Verfügbarkeit dieser Daten nach der Filterung 77%, was bedeutet, dass 23% der Daten fehlen und aufgefüllt werden sollten (die Stellen, an denen Daten fehlen, sind durch Pfeile gekennzeichnet). Die oben beschriebenen Imputationsmethoden wurden verwendet, um die Datenlücken zu schließen. Die Ergebnisse unterscheiden sich je nach der verwendeten Methode, wie im unteren Teil von Abbildung 12 zu sehen ist. Dies wirft die wichtige Frage auf, welche Methode zum Auffüllen von Daten für einen bestimmten Parameter (in diesem Fall FOC ME) am besten geeignet ist. Hieraus ergibt sich die Notwendigkeit, klare Kriterien für den Vergleich verschiedener Methoden aufzustellen. Es ist jedoch möglich, dass eine Methode für einen bestimmten Parameter besser ist als eine andere, während sie für einen anderen Parameter schlechter ist.

Eines der einfachsten Kriterien, das verwendet werden kann, ist der Vergleich der betreffenden Daten mit ähnlichen Daten aus einer anderen Quelle unter Berücksichtigung der Zeitstempel. In diesem Beispiel sind die gesuchten Daten aus einer anderen Quelle verfügbar, nämlich dem automatisch gemessenen Noon Report (siehe Abbildung 13). Vergleicht man die gemessenen mit den aufgefüllten manuell abgelesenen Daten, so wird deutlich, dass die ideale Methode zum Ausfüllen in diesem Beispiel die Lineare Regression („Time Linear“) ist.

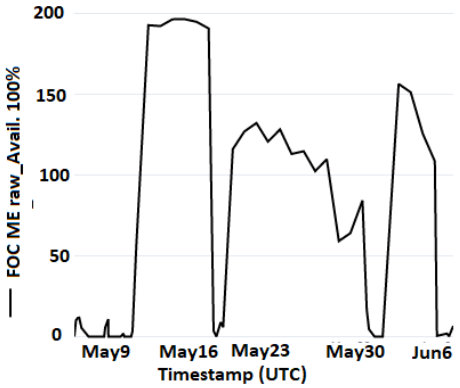


Abbildung 13: Kraftstoffverbrauch der Hauptmaschine (gemessen)

Ein weiteres zuverlässiges (allerdings komplexeres) Kriterium für die Akzeptanz oder den Ausschluss der Ergebnisse verschiedener Imputationsmethoden basiert auf der Untersuchung der Korrelationen und Abhängigkeiten des betreffenden Parameters von anderen operativen (oder mit operativen Parametern verbundenen) Parametern vor und nach der Imputation. Die Ergebnisse, die die Korrelation drastisch beeinträchtigen, sollten ausgeschlossen werden. Darüber hinaus wurde ein Imputationsverfahren entwickelt, bei dem die Daten eines oder mehrerer spezifischer Parameter, mit denen der interessierende Parameter stark korreliert oder abhängig ist (vor der Imputation), berücksichtigt werden. Auf Korrelationen und Abhängigkeiten wird im nächsten Abschnitt näher eingegangen.

Datenanalyse

Nachdem die Rohdaten aus dem Schiffsbetrieb mit geeigneten Algorithmen vorverarbeitet wurden, ist die Basis für weitere Analyseschritte geschaffen, die hauptsächlich auf ein besseres Verständnis der Schiffsperformance unter realen Betriebsbedingungen abzielen. Hier sind bspw. alle Faktoren von Interesse, die den Energieverbrauch eines Schiffes während des Betriebs beeinflussen. Das Hauptziel solcher Analysen besteht darin, den Schiffsbetrieb im Hinblick auf kritische Leistungsindikatoren wie Kraftstoffverbrauch und Emissionen zu optimieren, was zu ökologischen und wirtschaftlichen Vorteilen für Eigner führt.

Korrelationen von Betriebsdatenparametern

Korrelationen und Abhängigkeiten von Parametern aus dem Schiffsbetrieb werden im Folgenden durch einige konkrete Beispiele veranschaulicht, die verschiedene Arten von Datenbeziehungen für zwei wichtige Betriebsdatenparameter zeigen. Hierzu wird die Beziehung zwischen dem Gesamtkraftstoffverbrauch des Schiffes („Total FOC“) bzw. dem Gesamtkraftstoffverbrauch der Hauptmaschine („FOC ME“) mit unterschiedlichen Navigations- und Umweltparametern aus verschiedenen Quellen und mit unterschiedlicher Datenfrequenz für das untersuchte Schiff „MSC Bari“ auf einer Reise von Europa nach Asien untersucht.

Zunächst müssen jedoch die in der Studie verwendeten Methoden erläutert werden:

- Pearson Korrelation: Der Pearson‘sche Korrelationskoeffizient misst die lineare Korrelation zwischen zwei Parametern und hierbei ignoriert andere Arten der Korrelation (bspw. quadratisch oder exponentiell, siehe Sedgwick [20]). Der Koeffizient kann Werte zwischen [-1, 1] einnehmen. Wenn der Koeffizient gleich „-1“ oder „+1“ ist, liegen die Datenpunkte zweier Parameter genau auf einer Linie. Ist der Koeffizient nahezu null, liegt keine lineare Korrelation zwischen den betrachteten Parametern vor.

- Spearman Korrelation: Der Rangkorrelationskoeffizient nach Spearman gibt die monotone Beziehung zwischen zwei gemessenen Variablen an (siehe Zar [21]). Der Koeffizient kann Werte zwischen [-1, 1] annehmen. Bei zwei Parametern x und y ist der Korrelationskoeffizient „+1“, wenn x immer zunimmt, während y zunimmt (unabhängig vom Gradienten). Andernfalls, wenn x immer abnimmt, während y zunimmt, ist der Koeffizient „-1“. Es ist wichtig zu erwähnen, dass der grundlegende Unterschied zwischen diesem Ansatz und dem Ansatz nach Pearson darin besteht, dass mit Hilfe des Rangkorrelationskoeffizienten auch mit monotonen Beziehungen zwischen zwei Variablen abgedeckt werden.
- Abstandskorrelation: Mit diesem Verfahren können sowohl lineare als auch nichtlineare Beziehungen zwischen zwei Variablen gemessen werden (siehe Russell [22]). Der Korrelationskoeffizient kann Werte zwischen [0, 1] annehmen. Der Wert „0“ bedeutet, dass die beiden Metriken unabhängig sind, und der Wert „1“ zeigt an, dass die Unterräume der angesprochenen Parameter gleichwertig sind.

Diese Methoden ermöglichen die Erkennung und Analyse verschiedener Arten von Korrelationen und Abhängigkeiten zwischen den verschiedenen Schiffsbetriebsparametern, ob linear oder nichtlinear. Sie können bspw. dazu verwendet werden, die Ergebnisse der verschiedenen oben erwähnten Datenimputationsmethoden zu akzeptieren oder zu verwerfen. Darüber hinaus können sie verwendet werden, um ein ein tieferes Verständnis des Schiffsbetriebsprozesses zu erlangen, um daraus Empfehlungen für einen effizienteren Betrieb unter wirtschaftlichen und ökologischen Gesichtspunkten abzuleiten.

Beispiele

Anhand der folgenden drei Beispiele werden die Korrelationen der Betriebsdatenparameter „Gesamt-FOC“ und „FOC ME“ mit den übrigen Variablen anhand von drei (ausgewählten) Datensätzen aus drei unterschiedlichen Quellen überprüft, die mit drei unterschiedlichen Frequenzen aufgezeichnet wurden, und zwar in der Reihenfolge 1. niedrige Frequenz, 2. mittlere Frequenz, 3. hohe Frequenz. In einem vorgelagerten Schritt wurden die Daten dem zuvor beschriebenen Vorverarbeitungsprozess unterzogen. Für jeden der drei Datensätze wurde eine Matrix erstellt, die die Beziehungen zwischen den beiden genannten Parametern und den anderen Betriebsparametern mit den verschiedenen zuvor beschriebenen Methoden beschreiben - Pearson-Korrelation (1, untere Zeile), Spearman Korrelation (2, mittlere Zeile) und Abstandskorrelation (3, obere Zeile).

Beispiel 1 – niederfrequente Daten

Zunächst wird die Korrelation der Verbrauchsparameter „Gesamt-FOC“ und „FOC ME“ mit ausgewählten Wetter- und Navigationsparametern, die in den Noon Reports mit geringer Frequenz von einigen Minuten bis zu 24 Stunden enthalten sind (je nach den Ereignissen an Bord des Schiffes, wie in Jabary et. al. [25] beschrieben). An dieser Stelle ist anzumerken, dass einige in den Noon Reports enthaltenen Wetterdaten aus zwei Quellen stammen: manuell von der Besatzung abgelesen bzw. visuell bestimmt oder über einen externen Anbieter bezogene Daten (in den Matrizen mit „prov.“ Gekennzeichnet). Aus der in Abbildung 14 dargestellten Korrelationsmatrix für niederfrequente Daten lassen sich folgende Erkenntnisse ableiten:

Abbildung 14: Korrelationsmatrix für niederfrequente Daten unter Anwendung der Pearson Korrelation (1, untere Zeilen), Spearman Korrelation (2, mittlere Zeilen) und Abstandskorrelation (3, obere Zeilen)

- Je nach Verfahren ist eine starke Korrelation von 0,82 bis 0,99 zwischen Kraftstoffverbrauch und der zurückgelegten Strecke bzw. der Schiffsgeschwindigkeit durch das Wasser und über Grund (eine Erhöhung der Geschwindigkeit und eine längere Fahrstrecke führt zu einem Anstieg des Verbrauchs) zu beobachten, was logisch und erwartbar ist.
- Gemäß der manuell durch die Schiffsbesatzung eingegebenen Daten ist die Auswirkung der wahren und relativen Windgeschwindigkeit und -richtung auf den Kraftstoffverbrauch des Schiffes gering und reicht von weniger als 0,01 bis 0,29 für die drei Verfahren, während sie gemäß der durch den externen Anbieter gelieferten Daten bis zu 0,35 erreicht.
- Bei den Wind-, und Seegangsdaten sind signifikante Unterschiede in der Korrelation zwischen den beiden Datenquellen (manuell durch Besatzung eingetragen und durch externe Anbieter geliefert) zu beobachten. Während eine Zunahme der Wellenhöhe gemäß der durch die Besatzung eingetragenen Daten den Kraftstoffverbrauch erhöht, ist zu beobachten, dass der Verbrauch mit zunehmender Wellenhöhe sinkt, wenn man die Daten des externen Anbieters zu Grunde legt. Dies wirft die grundlegenden Fragen auf:
    - Wie genau sind die durch die Besatzung eingegebenen Daten bezüglich ihres Wertes und ihres Zeitstempels, zumal es sich um Daten mit niedriger Frequenz handelt, die einen Zeitraum von 24 Stunden repräsentieren sollen? Dies führt zur Notwendigkeit, Daten mit höherer Frequenz zu überprüfen, was im folgenden Beispiel erfolgt.
    - Inwieweit können die Daten aus den Noon Reports in ihrer jetzigen Form verwendet werden, um zuverlässige Analysen zu erstellen, die zur Verbesserung des Schiffsbetriebs genutzt werden können?

Beispiel 2 – Daten im mittleren Frequenzbereich

Im zweiten Beispiel wird die Korrelation der Verbrauchsparameter „Gesamt-FOC“ und „FOC ME“ mit ausgewählten Wetter- und Navigationsparametern, die in mittlerer Frequenz (15-Minuten-Intervalle) vorliegen analysiert. Diese Daten wurden durch einen weiteren externen Anbieter im Rahmen des Vorhabens bezogen. Es handelt sich hierbei nicht um Messdaten an Bord des Schiffes, sondern um Ergebnisse aus Simulationsmodellen. Aus der in Abbildung 15 dargestellten

Korrelationsmatrix für Daten im mittleren Frequenzbereich lassen sich folgende Erkenntnisse ableiten:

- Der Einfluss der Windgeschwindigkeit auf den Kraftstoffverbrauch scheint auch im mittleren Frequenzbereich eher schwach zu sein, während die Windrichtung Korrelationen von bis zu 0,39 aufweist.
- Der Einfluss des Seegangs (Periode, Höhe und Richtung der Wellen) ist hier im Vergleich zum niederfrequenten Beispiel stärker ausgeprägt: So weißt der Kraftstoffverbrauch eine Korrelation von bis zu 0,4 mit der Seegangsrichtung auf. Dieses Beispiel unterstreicht die Notwendigkeit, die Frequenz der erhobenen Daten mit in die Interpretation von Korrelationsergebnissen einzubeziehen.
- Diese Daten enthalten im Vergleich zu den niederfrequenten Daten in Abbildung 14 Informationen zu Strömungsverhältnissen am Ort des Schiffes. Danach ist der Kraftstoffverbrauch zwischen 0,19 und 0,31 mit Richtung und Geschwindigkeit der Strömung korreliert.

Es ist festzuhalten, dass sowohl der Windwiderstand nichtlinear mit der Windgeschwindigkeit ansteigt als auch der wellenbedingte Zusatzwiderstand nichtlinear mit der Wellenhöhe zunimmt. Dies wird dadurch verdeutlicht, dass die Abstandskorrelation (3) stets höhere Korrelationskoeffizienten als die Pearson (1) oder Spearman Korrelation (2) liefern.

Abbildung 15: Korrelationsmatrix für Daten im mittleren Frequenzbereich unter Anwendung der Pearson Korrelation (1, untere Zeilen), Spearman Korrelation (2, mittlere Zeilen) und Abstandskorrelation (3, obere Zeilen)

Beispiel 3 – Hochfrequente Daten

Im dritten Beispiel wird die Korrelation der Verbrauchsparameter „Gesamt-FOC“ und „FOC ME“ mit ausgewählten Wetter- und Navigationsparametern, die mit hohen Frequenzen (eine Sekunde) durch an Bord des Schiffes vorhandene Sensoren aufgezeichnet wurden, verglichen. Direkte Messdaten sind die genauste Form der Abbildung von Betriebs- und Umweltparametern. Die Verfügbarkeit dieser Daten ist jedoch direkt durch die an Bord eines Schiffes vorhandenen Sensoren eingeschränkt. In Abstimmung mit dem Betreiber eines sind ggf. zusätzliche, über die Standardausrüstung hinausgehende Installationen sinnvoll, um die Genauigkeit der Abbildung des Betriebsverhaltens eines Schiffes durch zusätzliche hochfrequente Informationen zu erhöhen. Aus der in Abbildung 16 dargestellten Korrelationsmatrix für hochfrequente Daten lassen sich

folgende Erkenntnisse ableiten:

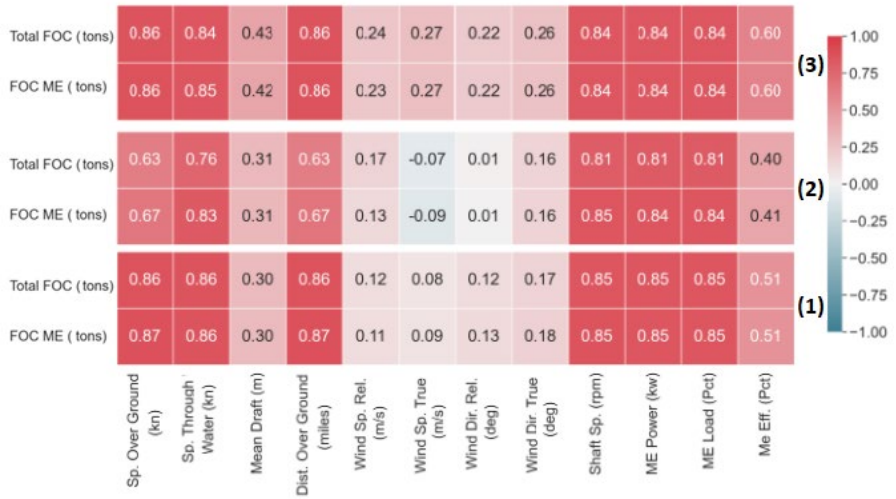


Abbildung 16: Korrelationsmatrix für hochfrequente Daten unter Anwendung der Pearson Korrelation (1, untere Zeilen), Spearman Korrelation (2, mittlere Zeilen) und Abstandskorrelation (3, obere Zeilen)

- Neben der Bestätigung der in den beiden vorangegangenen Beispielen bereits beobachteten Ergebnisse hinsichtlich der erwartbaren Korrelation zwischen dem Kraftstoffverbrauch und der Schiffsgeschwindigkeit sowie der zurückgelegte Strecke (bis zu 0,87) sowie der leicht ausgeprägten Korrelation mit dem Tiefgang (zwischen 0,3 und 0,43) sind in diesem Beispiel die Korrelationen der Verbrauchsparameter mit den Parametern der Hauptmaschine (Wellendrehzahl, Leistung, Last und Wirkungsgrad) hervorzuheben. Diese sind Schlüsselparameter für die Leistungsbewertung und Optimierung des Schiffsbetriebs.
- Da an Bord des Schiffes nur Sensorik zur Messung des Windes vorhanden war, lassen sich keine Betrachtungen zum Seegang und zu Strömungseinflüssen ableiten. Die hochfrequent gemessenen Winddaten an Bord der MSC Bari liefern folgende Erkenntnisse:
    - Der Kraftstoffverbrauch weißt eine moderate Korrelation (0,22 bis 0,27) mit der wahren und der relativen Windgeschwindigkeit auf, sofern man die Abstandskorrelation (3) verwendet. Aufgrund des nichtlinearen Zusammenhangs ist die Korrelation für die anderen beiden Verfahren nahezu null. Vergleichbare Tendenzen sind für die Korrelation mit der wahren und relativen Windrichtung erkennbar.

Zur Einordnung dieser Ergebnisse sollte berücksichtigt werden, dass die Daten eine Reise der MSC Bari von Europa nach Asien im Zeitraum Mai bis Juni repräsentieren. Im Frühling und Frühsommer sind die auf dieser Route vorherrschenden Wetterbedingungen schwächer als in der Herbst- und Wintersaison. Um ein vollständiges Bild zu erhalten, ist es daher notwendig, Daten über einen längeren Zeitraum zu betrachten.

Ableitung von Empfehlungen zur Anwendung von ISO 19030

Aus den Erfahrungen in der Vorverarbeitung und Analyse von Schiffsbetriebsdaten im Rahmen des Vorhabens lassen sich folgenden Empfehlungen ableiten, die zu einer Verbesserung der

Ergebnisqualität bei der Anwendung von Methoden gemäß ISO 19030 führen können.

1. Die Anwendung der ISO-Methoden auf Betriebsdaten aus unterschiedlichen Quellen und deren integrierte und synchronisierte Bereitstellung trägt zu einer Verbesserung der Ergebnisqualität bei.
2. Das in der ISO 19030 beschriebene Filterkonzepts kann durch Hinzufügen weiterer Kriterien verbessert werden (siehe zuvor erwähnte Filterungsmethoden). Die Filterung der Daten würde sich damit nicht mehr auf den Ausschluss von fehlenden Daten und von Daten, die nicht innerhalb eines in der ISO definierten Wertebereiches liegen, beschränken.
3. Die Imputation gefilterter Daten bietet einen guten Mechanismus zur Verbesserung der statistischen Aussagekraft eines Datensatzes, indem die Anzahl der Daten beibehalten wird und keine fehlenden oder gefilterten Daten verloren gehen.
4. Die Berücksichtigung des Navigationsstatus (z.B. at Sea, at Anchor, ...) bei der Anwendung der ISO-Methoden auf Schiffsbetriebsdaten trägt zur Verbesserung der Ergebnisse für die Leistungsvorhersage bei.

II.1.3 Prognose der Schiffsleistung (AP 3000)

Die in der Datenbank gespeicherten Rohdaten zu Navigations- und Umgebungsinformationen werden unterschiedlichen Vorverarbeitungsprozessen unterzogen, um ihre Qualität zu verbessern und sie für Vorhersagezwecke nutzen zu können. Dieser im Rahmen des Vorhabens entwickelte Prozess und Datenfluss ist in Abbildung 17 schematisch illustriert.

Vorverarbeitung der Rohdaten

Die Vorverarbeitung der Daten umfasst die folgenden beiden Prozessschritte:

- Filterung: In diesem Schritt werden die aus der Datenbank abgerufenen Rohdaten auf folgende Probleme geprüft: fehlende Daten, Nullen und nicht-physikalische Ausreißer. Jeder Datenpunkt, der einer der vorgenannten Kategorien zugeordnet werden kann, wird als „fehlend“ behandelt. Weitere Einzelheiten über diesen Prozessschritt finden sich in Jabary et. al. [26].
- Imputation: Die im Filterungsschritt ermittelten fehlenden Daten werden in Abhängigkeit von der Verfügbarkeit der übrigen gefilterten Daten imputiert (aufgefüllt). Zu diesem Zweck wurden verschiedene Imputationsmethoden implementiert: Mittelwert-Imputation, lineare Zeitregression und KNN-Imputation, weitere Einzelheiten siehe Jabary et. al. [26].

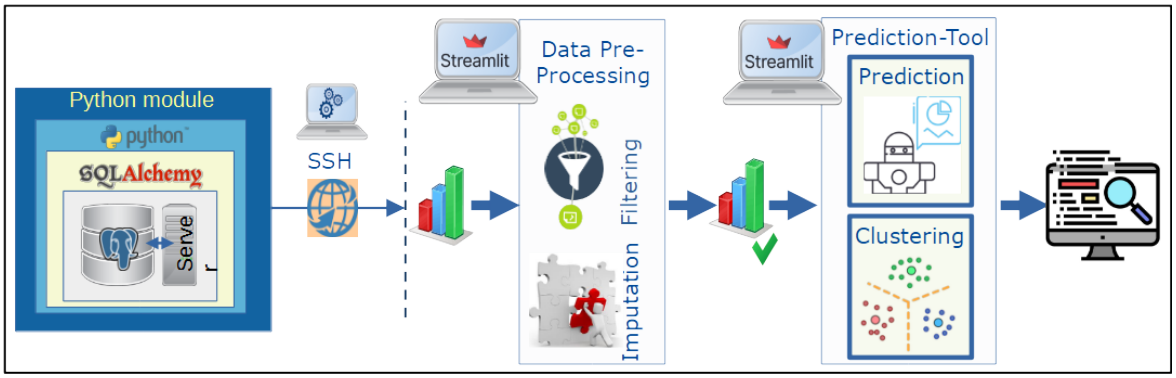


Abbildung 17: Schematische Darstellung der Vorverarbeitung von Betriebsdaten für die anschließende Nutzung zur Vorhersage der Schiffsperformance

Vorhersage-Tool

Um Schiffseignern und der Crew eine Vorhersage des Verhaltens bestimmter Betriebsparameter unter zu erwartenden Bedingungen zu ermöglichen, wurde im Rahmen des Vorhabens ein Vorhersagetool auf Basis von Algorithmen des maschinellen Lernens entwickelt. Die Eingabedaten für dieses Tool sind vorverarbeitete Betriebsdaten, die entsprechend des in Abbildung 17 dargestellten Datenflusses genutzt werden. Das in der Programmiersprache Python entwickelte Vorhersagetool bietet die eine Vielzahl von Funktionen. Mit Hilfe der Streamlit-Technik (Streamlit [23]) konnte eine interaktive Auswahl von Bedingungen realisiert werden.

Das entwickelte Tool bietet die folgenden Funktionalitäten:

- Prognose eines gewünschten Betriebsparameters für einen definierten Betriebszustand durch Aufteilung des adressierten Datensatzes (z.B. für eine Reise) nach dem Navigationsstatus, der „Moored“, „At Anchor“ und „Underway using engine“ umfasst
- Auswahl der in den Vorhersageprozess einzubeziehenden Betriebsparameter, so dass der Vorhersageprozess entweder nur auf Grundlage eines bestimmten Bereichs ausgewählter Parameter, z. B. Wellenhöhe oder Windgeschwindigkeit, oder auf der Grundlage aller verfügbaren Parameter durchgeführt werden kann
- Vorhersage sowohl für normalisierte (Anpassung der Datenwerte an eine gemeinsame Skala, um die Vorhersagegenauigkeit zu erhöhen) als auch für nicht normalisierte Datensätze
- Durchführung von Prognosen mit verschiedenen Methoden des maschinellen Lernens (ML) auf Basis von sklearn (Scikit-learn [24]). Um die genausten Vorhersagen erzielen zu können, werden die Ergebnisse der implementierten Algorithmen systematisch verglichen und die mit dem Vorhersageprozess verbundene Genauigkeit wird ermittelt
- Clustering-Funktion um den angesprochenen Datensatz in Abhängigkeit von bestimmten Betriebsparametern (Leitparameter) in Cluster (Gruppen) einteilen und die Vorhersage für diese Gruppen mit ML-Methoden durchführen zu können.

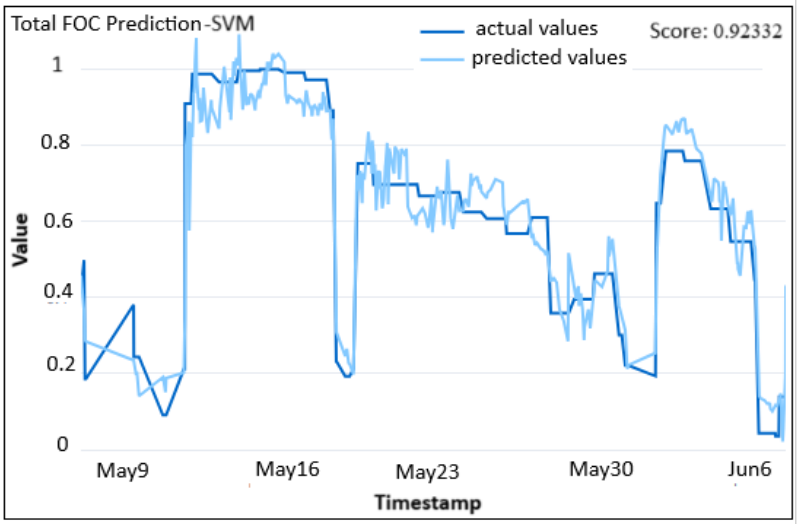
Beispiel

Anhand des folgenden Beispiels soll die Effizienz des Vorhersagetools verdeutlicht werden. Zunächst werden Navigations- und Umweltdaten (von Drittanbietern bezogen) und die Kraftstoffverbrauchsdaten (FOC) aus dem „Noon Report“ in Form von 2.450 Rohdatenpunkten für eine Reise zwischen Europa und Asien des Referenzschiffes MSc Bari aus der entwickelten Datenbank importiert und – wie oben beschrieben – vorverarbeitet und als Eingabe für das Vorhersagetool verwendet. Die Prognose des Gesamt-Kraftstoffverbrauchs wird nun unter Verwendung der zuvor beschriebenen Funktionalitäten des Vorhersagetools durchgeführt;

- Navigationsstatus: „Underway using engine“, Datensatz mit 1.877 Datenpunkten während der Reise
- Normalisierung der Daten
- Einbeziehung der Betriebsparameter: mittlerer Tiefgang, Geschwindigkeit über Grund (SOG), Wind (Richtung, Geschwindigkeit), Seegang (Höhe, Periode, Richtung), Strömung (Geschwindigkeit, Richtung).
- Verwendete Machine Learning Methoden: lineare Regression, Multi-Layer Perceptron (MLP) und Support Vector Machine (SVM)

Ein Vergleich der Vorhersagegenauigkeiten der genutzten Algorithmen, die in der Tabelle auf der

rechten Seite von Abbildung 18 aufgeführt sind, zeigt, dass alle drei Methoden grundsätzlich gute Ergebnisse liefern (Prognosegenauigkeit 90-92 %), wobei die SVM-Methode mit einem Wert von 0,92332 die beste Vorhersagegenauigkeit aufweist. Abbildung 17 (links) zeigt die tatsächlich gemessenen (dunkelblau) und die mittels SVM vorhergesagten Daten (hellblau) als Funktion der Zeit.



Methods comparision	
Method	Score
Linear regression	0.89797
SVM	0.92332
MLP	0.90118

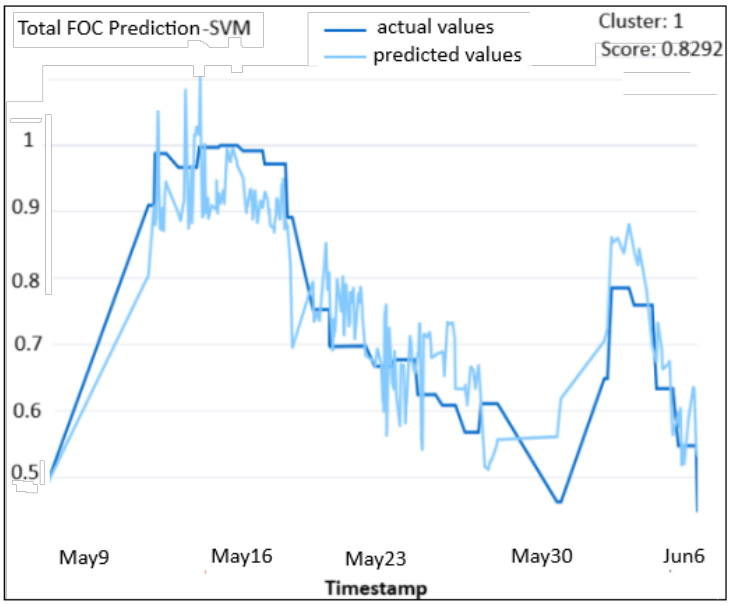
Abbildung 18: Prognosegüte für den Kraftstoffverbrauch (FOC) bei Anwendung verschiedener Algorithmen des maschinellen Lernens

Es sei erwähnt, dass die tatsächlich gemessenen Werte normalisiert sind und nur Werte im Bereich von [0-1] annehmen, während die vorhergesagten Daten auf den normalisierten Daten basieren und daher in einigen Punkten Wert > 1 annehmen, da die Grenzen der Normalisierung bei Anwendung des SVM Algorithmus nicht berücksichtigt werden.

Im nächsten Beispiel wird die Clustering-Funktion auf denselben Datensatz angewendet, wobei die folgenden Bedingungen gelten:

- Die Anzahl der Cluster unter Berücksichtigung der „Elbow Method“ (Scikit-learn, [25]) ist zwei
- der führende Parameter für das Clustering ist in diesem Beispiel die Geschwindigkeit über Grund (SOG). Das bedeutet, dass die beiden resultierenden Cluster auf der Grundlage der gemessenen SOG-Werte erstellt werden, die Vorhersage jedoch unter Berücksichtigung aller zuvor genannten Betriebsparameter erfolgt.

Abbildung 19 zeigt die Prognoseergebnisse für Cluster 1 (mittlere und hohe Geschwindigkeiten). Auf der rechten Seite der Abbildung sind die Größe, die Genauigkeit und die Mitte des Clusters (in Bezug auf SOG) dargestellt. Wie bereits erwähnt, verringert das Clustern von Daten die Genauigkeit der Vorhersageergebnisse in solch kleinen Datensätzen.



Clusters data		
	cluster1	cluster2
Size	1403	474
Score	0.8292	0.7003
SOG-Center	0.7382	0.253

Abbildung 19: Prognose des Kraftstoffverbrauchs (FOC) unter Anwendung des Cluster-Funktion

Das Angebot verschiedener Vorhersageansätze innerhalb des entwickelten Tools (mit/ohne Clustering-Funktion) trägt dazu bei, die besten Ergebnisse für die Vorhersage bestimmter Parameter in Abhängigkeit von einem bestimmten Satz von Parametern auswählen zu können, was viele Vorteile für der Erstellung von Betriebsprofilen mit sich bringt, die eine Kombination einer Reihe von Betriebsparametern unter bestimmten Betriebsbedingungen darstellen.

II.1.4 Schiffsbetriebsprofile (AP 6000)

Im Rahmen dieses Arbeitspaket wurde ein Verfahren erarbeitet, um aus Langzeitdaten für AIS und Wetter Betriebsprofile für Schiffe abzuleiten.

Datenvorverarbeitung

Über einen externen Anbieter wurden im Vorhaben entsprechende Daten für 25 Schiffe über einen Zeitraum von 5 Jahren bezogen. Die Rohdaten wurden im CSV-Format geliefert. Bei den 25 Schiffen, die in Abbildung 20 aufgelistet sind, handelt es sich um Containerschiffe in einem Kapazitätsbereich zwischen 3.000 und 24.000 TEU.

Nr.	IMO-Nr.	TEU	Name
1.	9461415	14000	MSC Deila
2.	9461374	14000	MSC Alexandra
3.	9447847	12552	Madrid Express
4.	9447885	12552	Southampton Express
5.	9456757	13630	Maersk Edinburgh
6.	9458030	13092	Maersk Edmonton
7.	9238789	4250	MSC Manu
8.	9227027	4250	Seaspan Dalian
9.	9224312	4250	Seaspan Chiwan
10.	9351579	4250	MSC Tamara
11.	9351608	4250	Seaspan Fraser
12.	9437385	4250	Seaspan Loncomilla
13.	9839179	23112	CMA CGM Jacques Saade
14.	9839181	23112	CMA CGM Palais Royal
15.	9839260	23656	MSC Mina
16.	9839272	23656	MSC Isabella
17.	9839430	23656	MSC Gulsun
18.	9839442	23656	MSC Samar
19.	9252735	3100	Independent Vision
20.	9334375	3108	Independent Horizon
21.	9298648	3091	Daphne
22.	9311880	3091	Leto
23.	9219379	3091	Polonia
24.	8913411	3005	MSC Santhya
25.	8913423	3007	MSC Shannon

Abbildung 20: Übersicht der 25 Containerschiffe, für die eine Analyse der Langzeit-Betriebsdaten durchgeführt wurde.

Die Rohdaten wurden gemäß dem in Abschnitt II.1.2.1 beschriebenen Schiffsbetriebsinformationsmodell gespeichert. Da für diese 25 Schiffe neben den AIS- und Wetterdaten keine weiteren Datenquellen wie z.B. Noon Reports zur Verfügung standen, war eine Untergliederung der Reisen wie in Abbildung 7 dargestellt jedoch nicht möglich. Um die Daten dennoch im System speichern zu können, wurde die Daten für jeden Schiff jeweils einer einzigen Betriebsphase zugeordnet:

Ship → one voyage → one leg → one operation period

Das entsprechende Speicherkonzept ist in Abbildung 21 dargestellt. Der Zugriff auf die Langzeitdaten für ein Schiff erfolgt über die IMO-Nummer als eindeutigen Identifikator.

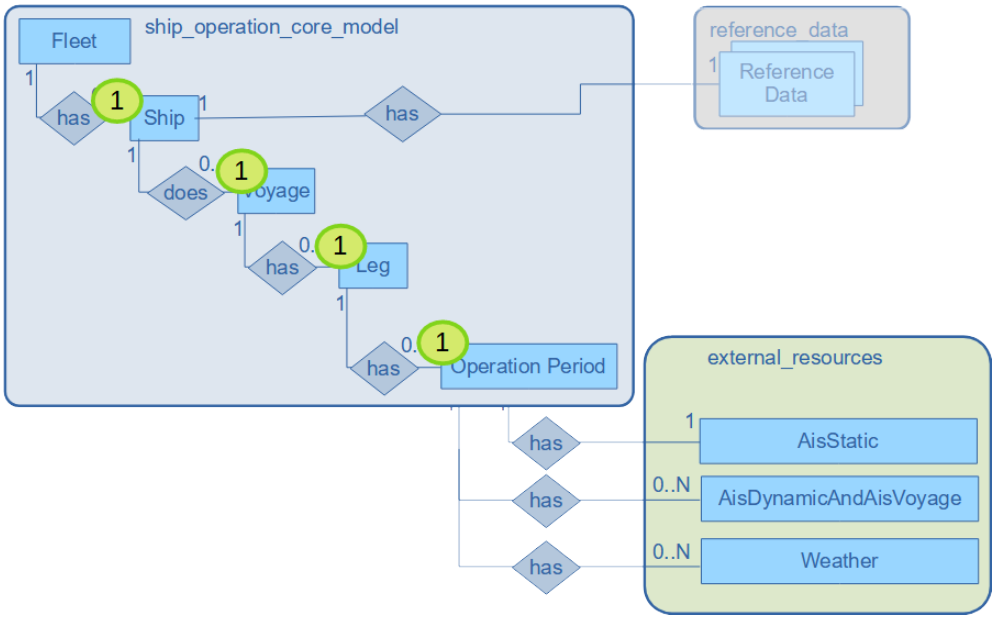


Abbildung 21: Speicherkonzept für die Langzeitdaten der 25 Schiffe

Um das Ziel dieses Arbeitspaketes – die Generierung von Betriebsprofilen aus Langzeitdaten – erreichen zu können, werden auch Informationen über die Leistungsdaten der Schiffe benötigt. Da im Rahmen des Vorhabens jedoch kein Anbieter gefunden wurde, der diese Daten zur Verfügung stellen konnte, wurden lediglich AIS- und Wetterdaten beschafft. Die nötigen Leistungsdaten wurden synthetisch mittels eines vom Projektpartner MMG zur Verfügung gestellten Leistungsprognose-Tools erzeugt und in der Datenbank abgelegt.

Die Rohdaten werden grundsätzlich im Originalzustand in der Datenbank gespeichert. Um sie für die weiterführenden Analysen aufzubereiten, durchlaufen sie bei Abruf aus der Datenbank zunächst die in Abschnitt II.1.2.2 beschriebene Vorverarbeitungsprozedur. Da sich jedoch herausgestellt hat, dass die Langzeitdaten zusätzliche Fehlertypen aufweisen, für die im bisherigen Vorverarbeitungsprozess noch nicht berücksichtigt wurden, war eine Erweiterung der Datenvorverarbeitungsprozedur um weitere Filterverfahren notwendig. Ein Beispiel für einen neuen Fehlertyp ist in Abbildung 22 dargestellt: Hier liegen eindeutig fehlerhafte Einträge des Aufenthaltsortes eines Schiffes vor, die als Outlier behandelt werden.

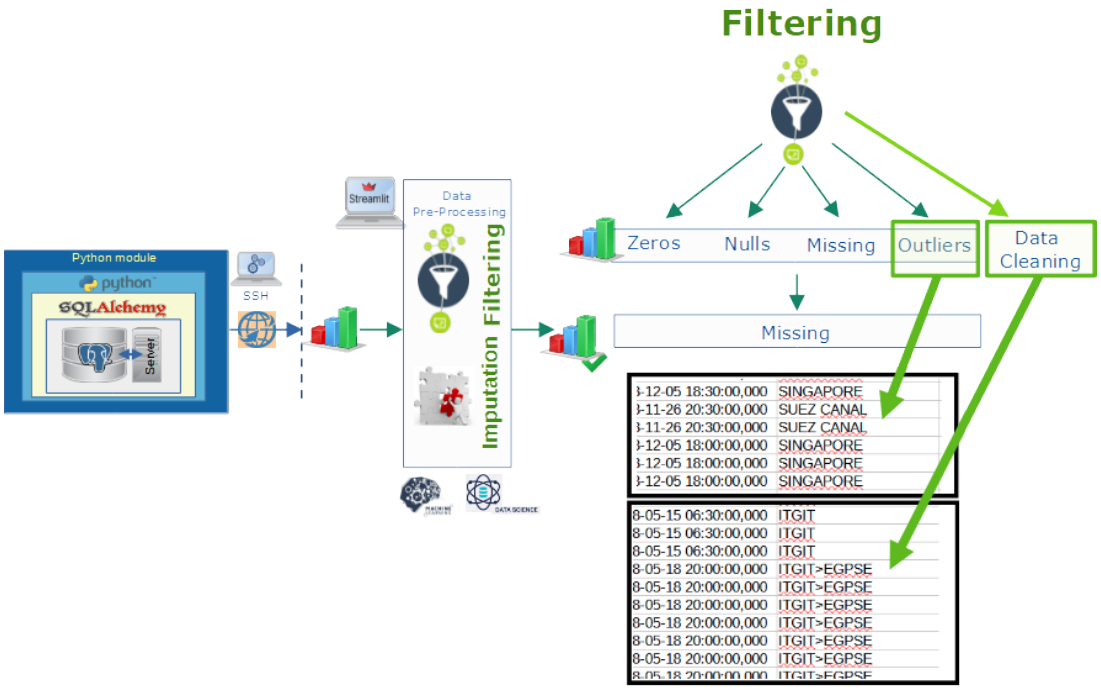


Abbildung 22: Schematischer Ablauf der Vorverarbeitung der Langzeitdaten mit beispielhaften Fehlern

Nach der Filterung der Rohdaten erfolgt auch hier die Imputation mit den in Abschnitt II.1.2.2 beschriebenen Verfahren, um eine möglichst hohe Datenverfügbarkeit und somit eine hohe statistische Aussagekraft der nachfolgenden Analysen zu erreichen.

In einem ersten Vorschritt für die Erstellung der Betriebsprofile wurden die Daten entsprechend der Fahrtgebiete der 25 Schiffe in die vier in Abbildung 23 visualisierten Gruppen unterteilt:

- Gruppe 1: Europa-Asien-Westküste USA
- Gruppe 2: Europa-Asien
- Gruppe 3: Europa-Ostküste USA
- Gruppe 4: Europa-Asien-Panamakanal-USA

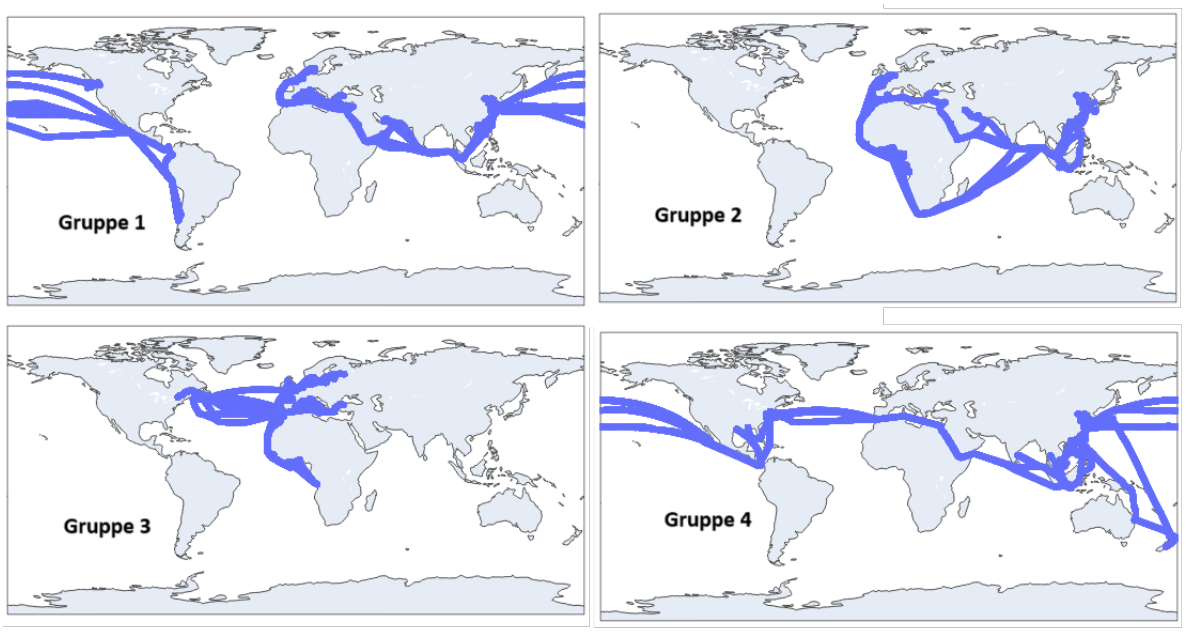


Abbildung 23: Visuelle Darstellung der Unterteilung der Langzeitdaten in vier Gruppen entsprechend ihrer Fahrtgebiete

Bildung der Schiffsbetriebsprofile

Zur Bildung der Betriebsprofile wurde eine eigenständige browserbasierte App mit Nutzeroberfläche entwickelt.

Die Erstellung von Betriebsprofilen basiert grundsätzlich auf der Clusterung ausgewählter Datenbereiche. Der grundsätzliche Ablauf ist in Abbildung 24 visualisiert, die entwickelte Benutzeroberfläche ist in Screenshots beispielhaft in Abbildung 25 dargestellt.

Grundsätzlich umfasst die Erstellung eines Betriebsprofils folgende Schritte:

1. Auswahl der Gruppe (Fahrtgebiet)
2. Auswahl eines Schiffes
3. Auswahl des/der gewünschten Betriebszustände (navigation status)
4. Auswahl des/der Jahreszeit
5. Auswahl der Cluster-Methode (Implementiert sind Algorithmen für K-Means und Mean Shift)
6. Auswahl der Anzahl der Cluster
7. Auswahl der Leitparameter (z.B. Geschwindigkeit, Leistung etc.)

Aufbauend auf diesen Eingaben wird durch die App ein Betriebsprofil-Bericht erstellt, der Informationen zu den Schwerpunktwerten der Leitparametern in den Clustern und den weiteren Parametern, die von Interesse sind enthält. Zusätzlich werden die Daten auf einer Weltkarte visualisiert, um eine einfache geographische Zuordnung der betrachteten Daten zur ermöglichen. Mit Hilfe dieser App ist es bspw. möglich, die Antriebsleistung, den Kraftstoffverbrauch oder die CO2-Emissionen eines Schiffes unter bestimmten Betriebsbedingungen zu analysieren. Die so gewonnenen Erkenntnisse können sowohl durch Reeder zur Optimierung des Schiffsbetriebs als auch durch Designer zur Verbesserung von Schiffsentwürfen der nächsten Generation genutzt werden.

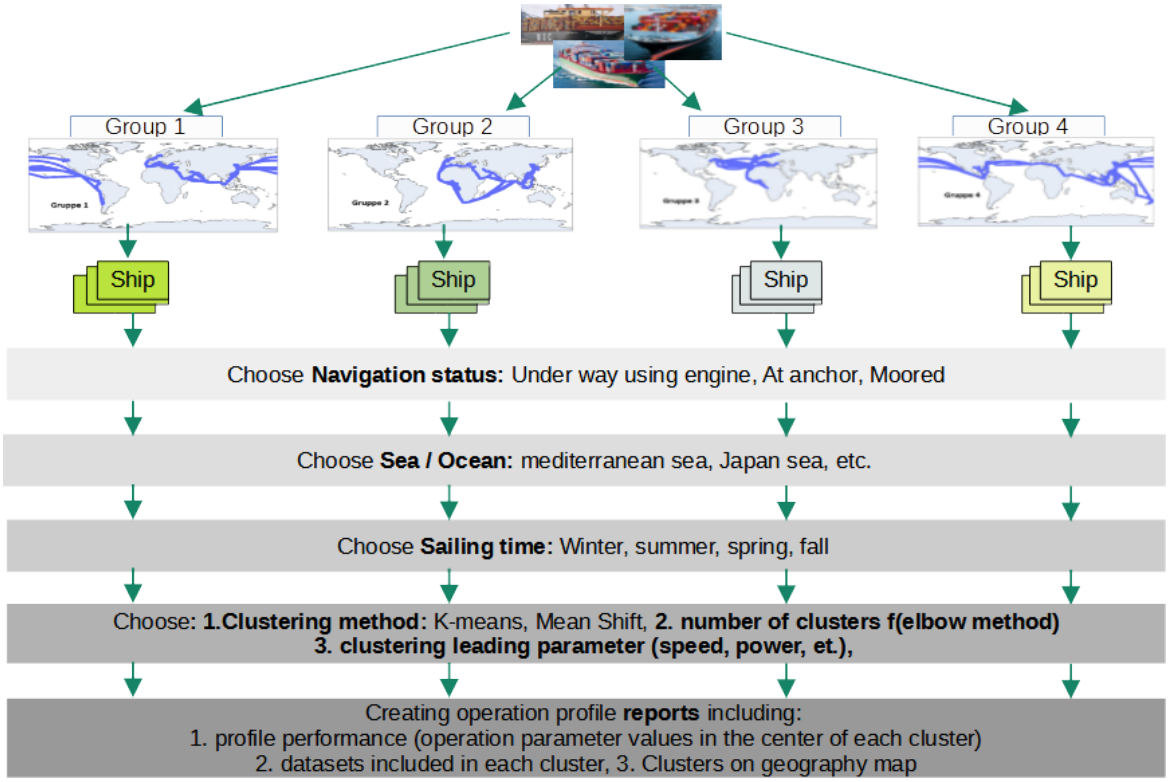


Abbildung 24: Schematische Darstellung des Ablaufs der Erstellung von Betriebsprofilen in der entwickelten App

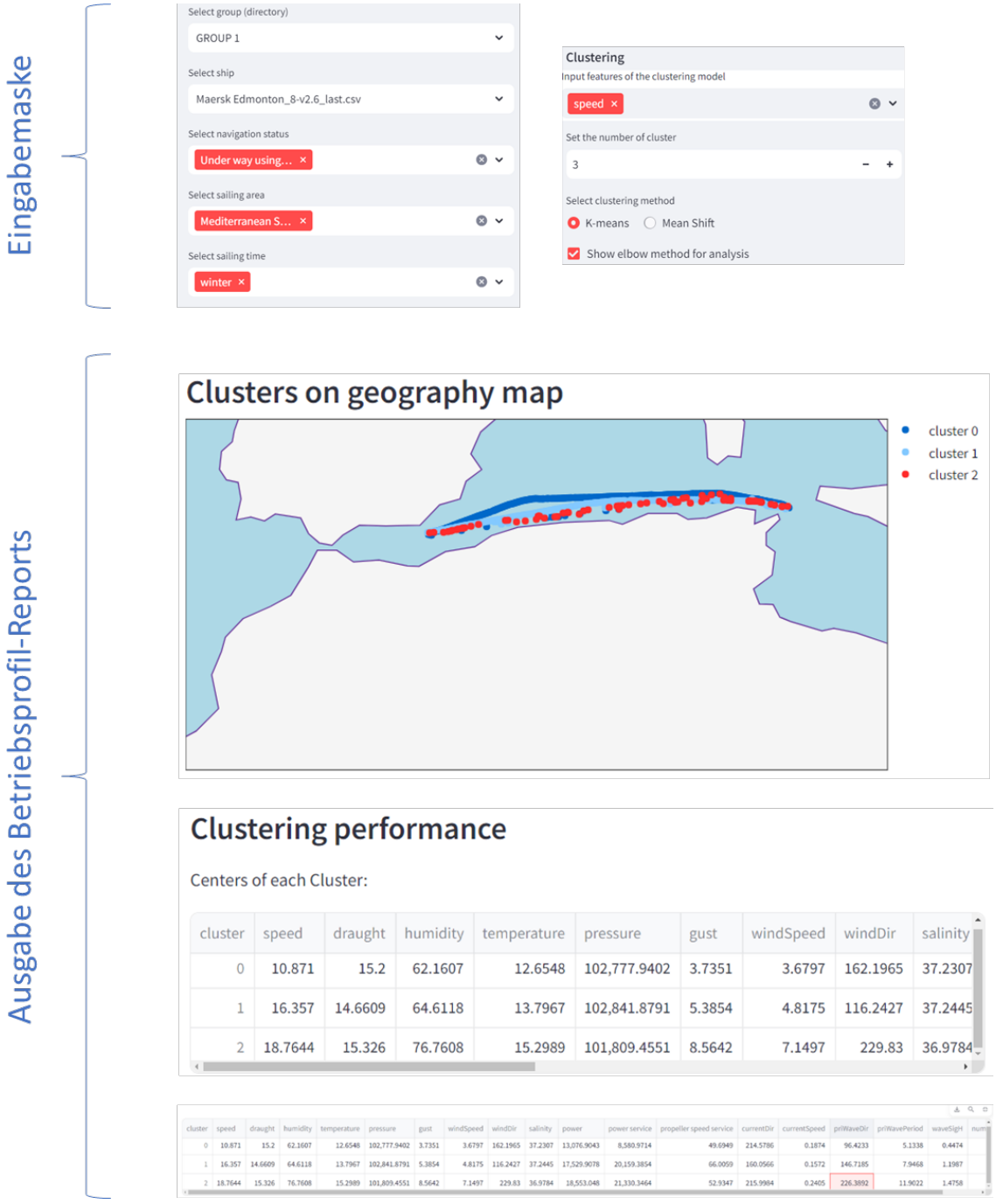


Abbildung 25: Screenshots der entwickelten browserbasierten App zur Generierung von Betriebsprofilen

Die App ist grundsätzlich modular programmiert und kann jederzeit funktional erweitert werden.

II.2 Erläuterungen zum zahlenmäßigen Nachweis

Pos. 0812

Auf dem Vorhaben wurde der wissenschaftliche Mitarbeiter Dr.-Ing. Wisam Jabary für insgesamt 40 PM beschäftigt. Mit Bescheid vom erfolgte eine kostenneutrale Laufzeitverlängerung um 8 Monate. Da Herr Jabarys Einstellung jedoch aus geschilderten Umständen verzögert erfolgte, war er insgesamt 40 PM auf dem Vorhaben beschäftigt. Die dadurch entstandenen Ausgaben beliefen sich auf 229.000,44€. Somit wurde diese Position fast vollständig verausgabt (232.673,16€ standen gemäß dem letzten gültigen Gesamtfinanzierungplan zur Verfügung).

Pos. 0822

Um die zwischenzeitlich entstandenen zeitlichen Verzögerungen aufholen zu können, waren Mit Herrn Jonas Musil, Herrn Chien Liu und Herr Ba Le auf dem Vorhaben mehr studentische Hilfskräfte als ursprünglich geplant beschäftigt. Die dadurch entstandenen Ausgaben beliefen sich auf 27.272,09€ gegenüber 16.694€ gemäß dem letzten gültigen Gesamtfinanzierungplan. Die Position wurde entsprechend überzogen.

Pos. 0834

Gemäß dem letzten gültigen Gesamtfinanzierungplan standen 179,00€ für Softwarelizenzen zur Verfügung. Aufgrund einer strategischen Entscheidung wurde jedoch vom ursprünglichen Beschaffungsplan abgesehen und Open-Source-Software eingesetzt.

Pos. 0835

Um die benötigten AIS- und Wetterdaten von externen Anbietern zu beschaffen, waren im letzten gültigen Gesamtfinanzierungplan 59.082,06€ vorgesehen. Entgegen des bei Antragstellung eingeholten Angebotes fielen die Kosten des Anbieters Jakota Cruise Systems (FleetMon) jedoch deutlich geringer aus. Insgesamt wurde nur 34.664,20€ verausgabt.

Pos. 0843

Die sonstigen Verwaltungsausgaben fielen mit 139,60€ gegenüber 1.600€ geringer als geplant aus. Herr Dr. Jabary belegte eine Fortbildung im Projektmanagement und die Produktionskosten für ein Poster des Vorhabens zur Ausstellung auf dem Forschungscamp der Uni Rostock wurde hier angesetzt.

Pos 0846.

Die Anzahl der tatsächlich durchgeführten Dienstreisen lag deutlich unter der ursprünglichen Planung. Dies liegt vor allem an der zunehmenden Zahl an Videokonferenzen seit der Corona-Pandemie. Im Konsortium fanden mit Ausnahme der Abschlussveranstaltung, an der Herr Dr. Jabary nicht teilnehmen konnte, tatsächlich keine Projekttreffen in Präsenz statt. Die einzigen Reisekosten fielen für Herrn Dr. Jabary an, der die Ergebnisse des Vorhabens auf den internationalen Fachkonferenzen COMPIT'22 in Siena, Italien und ISOPE2023 in Ottawa, Kanada präsentierte. Somit wurden nur 5.746,78€ von 9.821,00€ verausgabt.

Pos. 0850

Die Beschaffung des Servers für das Datenmanagementsystem war aufgrund von

Preissteigerungen teurer als ursprünglich angenommen. So wurden 9.220,12€ anstelle von 8.308.70€ verausgabt.

Insgesamt wurden somit 306.043,23€ anstelle der ursprünglich geplanten 328357,98€ verausgabt. Dies liegt insbesondere an den unerwarteten Minderausgaben für Daten in Pos. 0835.

II.3 Notwendigkeit und Angemessenheit der Arbeiten

Die durchgeführten Arbeiten und die dafür eingesetzten Ressourcen waren für die Erreichung der Zielstellung notwendig und angemessen. Die durch die Universität Rostock durchgeführten Arbeiten hätten ohne die bereitgestellte Förderung nicht durchgeführt werden können. Darüber hinaus mussten keine zusätzlichen Ressourcen zur Durchführung des Vorhabens aufgewandt werden.

II.4 Voraussichtlicher Nutzen – Fortschreibung des Verwertungsplans

Die im Vorhaben erarbeiteten Ergebnisse wurden in Form von Veröffentlichungen und Vorträgen auf internationalen Konferenzen einem breiten Fachpublikum zugänglich gemacht. Dies erfolgte bereits während der Laufzeit des Vorhabens. Zudem wurde durch zwei Masterarbeiten mit direktem Bezug zum Vorhaben ein Beitrag zur wissenschaftlichen Qualifikation von Studierenden geleistet. Erkenntnisse aus dem Projekt werden außerdem bereits in Vorlesungen und Übungen des Lehrstuhlinhabers integriert und so den Studierenden vermittelt. Dieser Know-How Zuwachs zukünftiger Absolventen dient einer Stärkung der Innovationskraft in der maritimen Wirtschaft.
Über die bereits getätigten Verwertungsaktivitäten hinaus plant die die Universität Rostock die Veröffentlichung eines weiteren Journal-Papers mit den Ergebnissen des Vorhabens. Zudem wird Anfang 2025 eine weitere Masterarbeit, die in direktem Zusammenhang mit dem Vorhaben steht, abgeschlossen. Auf Grundlage der Erkenntnisse des Vorhabens sind mittelfristig aufbauende und weiterführende Forschungsarbeiten mit Unternehmen aus der maritimen Branche geplant.

II.5 Während der Durchführung des Vorhabens bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Am Lehrstuhl für Schiffbau der Universität Rostock werden regelmäßig einschlägige Fachpublikationen in Form von Zeitschriften, Veröffentlichungen und Berichten ausgewertet. Sowohl die wissenschaftlichen Mitarbeiter als auch der Lehrstuhlinhaber sind damit über relevante Entwicklungen im Fachgebiet informiert. Zudem nimmt jeder Mitarbeiter mindestens einmal im Jahr an einer internationalen Fachkonferenz teil und übernimmt zusätzliche Reviewer-Aufgaben für Konferenzen und Journals. Im Rahmen dieser Tätigkeiten wurden keine Erkenntnisse dahingehend gewonnen, dass andere Stellen mit der Entwicklung vergleichbarer Methoden und Werkzeuge befasst waren.

II.6 Erfolgte und geplante Veröffentlichungen

II.6.1 Erfolgte Veröffentlichungen

Jabary, W. et al., 2022. *Development of a Unified Data Model to Improve Ship Operational Performance Analyses.* Proc 21st Conference on Computer Applications and Information Technology in the Maritime Industries, Pontignano, Italy

Jabary, W. et al., 2023. *Development of Machine Learning Approaches to Enhance Ship Operational Performance Evaluation Based on an Integrated Data Model.* Ottawa, Canada, The 33rd International Ocean and Polar Engineering Conference (ISOPE)

II.6.2 Geplante Veröffentlichungen

Derzeit befindet sich ein Journal-Paper mit den Ergebnissen des Vorhabens in der Vorbereitung.

Literaturverzeichnis

[1] Navis (2024), Fleet performance management. https://kaleris.com/solutions/terminal-operations/

[2] International Maritime Organisation (2008), International Code on Intact Stability. IMO IS Code

[3] Hoppe Marine (2024), System Performance Monitoring. https://www.hoppe-marine.com/

[4] International Maritime Organisation (2015), Revised Guidelines for the Onboard Operational use of Shipborne Automatic Identifcation Systems (AIS). IMO-Resolution A.1106(29)

[5] United Nations Economic Commission for Europe (1998), UN/LOCODE Codes for Ports and other Locations. recommendation no. 16, adopted by the centre for the facilitation of procedures practices for administration, commerce and transport, 3rd edition

[6] International Standards Organisation (2015), Ships and marine technology, Guidelines for the assessment of speed and power performance by analysis of speed trial data. ISO 15016:2015

[7] Raphael Arau´jo e Silva (2024), pgmodeler - PostgreSQL Database Modeler. https://pgmodeler.io

[8] Gordon C. Everest (1976), Basic Data Structure Models Explained With A Common Example. In Proceedings Fifth Texas Conference on Computing Systems, pages 39–46, Austin, TX

[9] The PostgreSQL Global Development Group (2024), PostgreSQL: The World's Most Advanced Open Source Relational Database. https://www.postgresql.org/

[10] Dave Page et al. (2024), pdAdmin. https://www.pgadmin.org/

[11] Michael Bayer (2024), The Python SQL Toolkit and Object Relational Mapper. https://www.sqlalchemy.org/

[12] The SQLite Team (2024), About SQLite. https://sqlite.org/about.html

[13] Chen, P. (1976), The entity-relationship model toward a unified view of data. ACM Transactions on Database Systems.

[14] Tukey, JW. (1977), Exploratory data analysis. Addison-Wesely.

[15] Raaijmakers, Q. A. (1999), Effectiveness of different missing data treatments in surveys with likert-type data: Introducing the relative mean substitution approach. Educational and Psychological Measurement, 59 (5), 725–748.

[16] Lodder, P. (2013), To Impute or not Impute, That's the Question. Advising on research methods: Selected topics: 1 - 7.

[17] Mohamed, C., Sedory, S. A., & Singh, S. (2018), Improved mean methods of imputation. Statistics, Optimization & Information Computing, 6(4), 526-535.

[18] Wang, Q., & Rao, J. N. K. (2001), Empirical likelihood for linear regression models under imputation for missing responses. Canadian Journal of Statistics, 29(4), 597-608.

[19] Batista, G.E., & Monard, M.C. (2002). A Study of K-Nearest Neighbour as an Imputation Method.

International Conference on Health Information Science.

[20]Sedgwick, P. (2012), "Pearson’s correlation coefficient." Bmj 345.

[21]Zar, Jerrold H. (2005), "Spearman rank correlation." Encyclopedia of biostatistics 7.

[22]RUSSELL, et. al. Similarity analyses of chromatographic herbal fingerprints: a review. Analytica Chimica Acta, 2013, 804. Jg., S. 16-28.

[23]Streamlit (2024), Streamlit’s open-source app framework, https://streamlit.io/

[24]Scikit-learn (2024), machine learning library for the Python programming language, https://scikit-learn.org/

[25]Jabary, W. et al. (2022), Development of a Unified Data Model to Improve Ship Operational Performance Analyses. Proc 21st Conference on Computer Applications and Information Technology in the Maritime Industries, Pontignano, Italy

[26]Jabary, W. et al. (2023), Development of Machine Learning Approaches to Enhance Ship Operational Performance Evaluation Based on an Integrated Data Model. Ottawa, Canada, The 33rd International Ocean and Polar Engineering Conference (ISOPE)