The National Academies of
SCIENCES · ENGINEERING · MEDICINE

TRANSPORTATION RESEARCH BOARD

How to Process AIS Data to Maximize Utility and Overcome Data Challenges

Tuesday, September 18, 2018
1:00-2:30 PM ET

The Transportation Research Board has met the standards and requirements of the Registered Continuing Education Providers Program. Credit earned on completion of this program will be reported to RCEP. A certificate of completion will be issued to participants that have registered and attended the entire session. As such, it does not include content that may be deemed or construed to be an approval or endorsement by RCEP.

RCEP

REGISTERED CONTINUING EDUCATION PROGRAM

Purpose

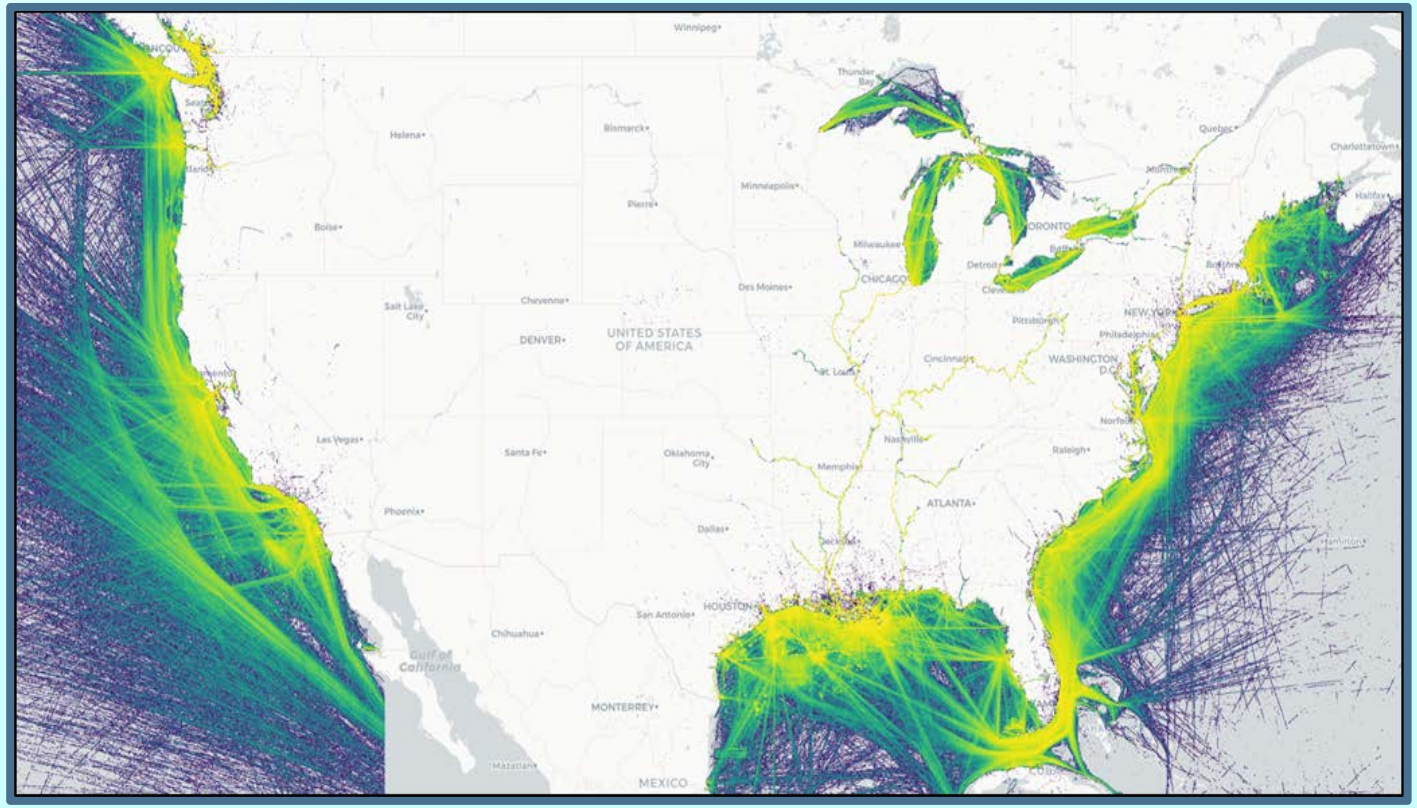
Examine how to process large marine vessel automatic identification system (AIS) datasets in order to maximize their usage.

Learning Objectives

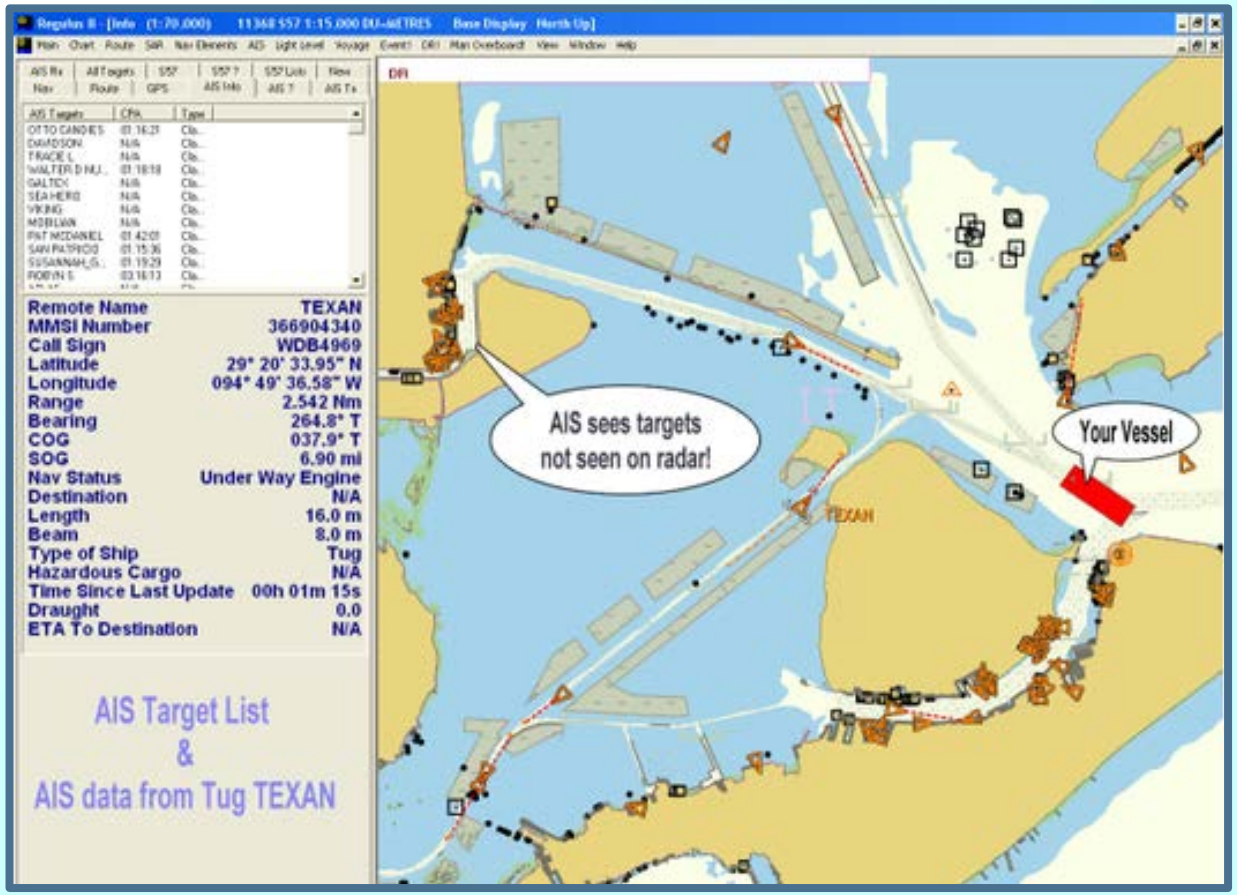
At the end of this webinar, you will be able to:

- List the steps that should be taken to process AIS data into a useful format
- Identify issues with AIS data and possible approaches that help mitigate these issues
- Describe methods to link AIS vessels to vessel characteristics
- Apply GIS data visualization tools to identify and correct data quality concerns

How to Process AIS Data to Maximize its Utility and Overcome Data Challenges



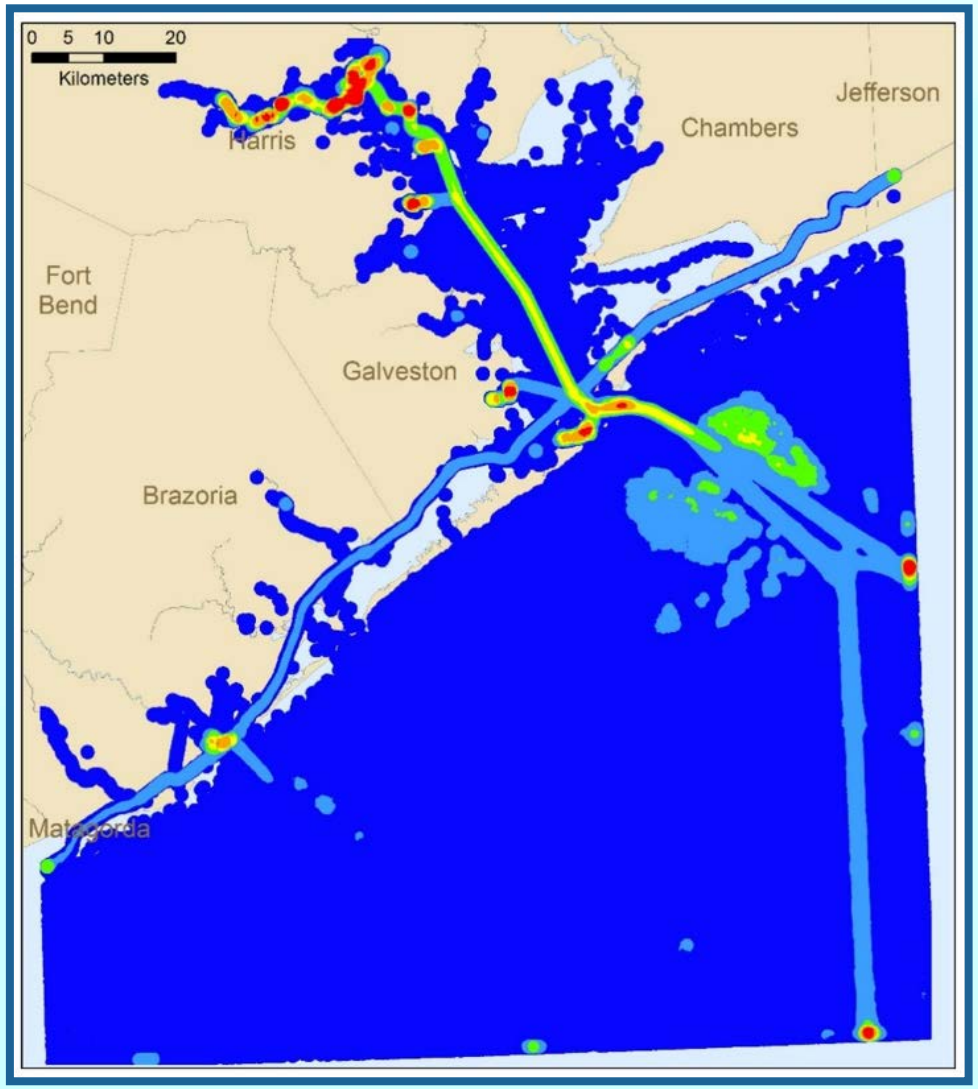
Introduction – Richard Billings, ERG



Presenters:

- Chris Leslie, Oceaneering/PortVision
- Heather Perez, Eastern Research Group
- Ed Carr, Ph.D., Environment and Energy Research Associates
- Tim Sturtz, Ph.D. Ramboll

Overview of the Flow of the Presentations

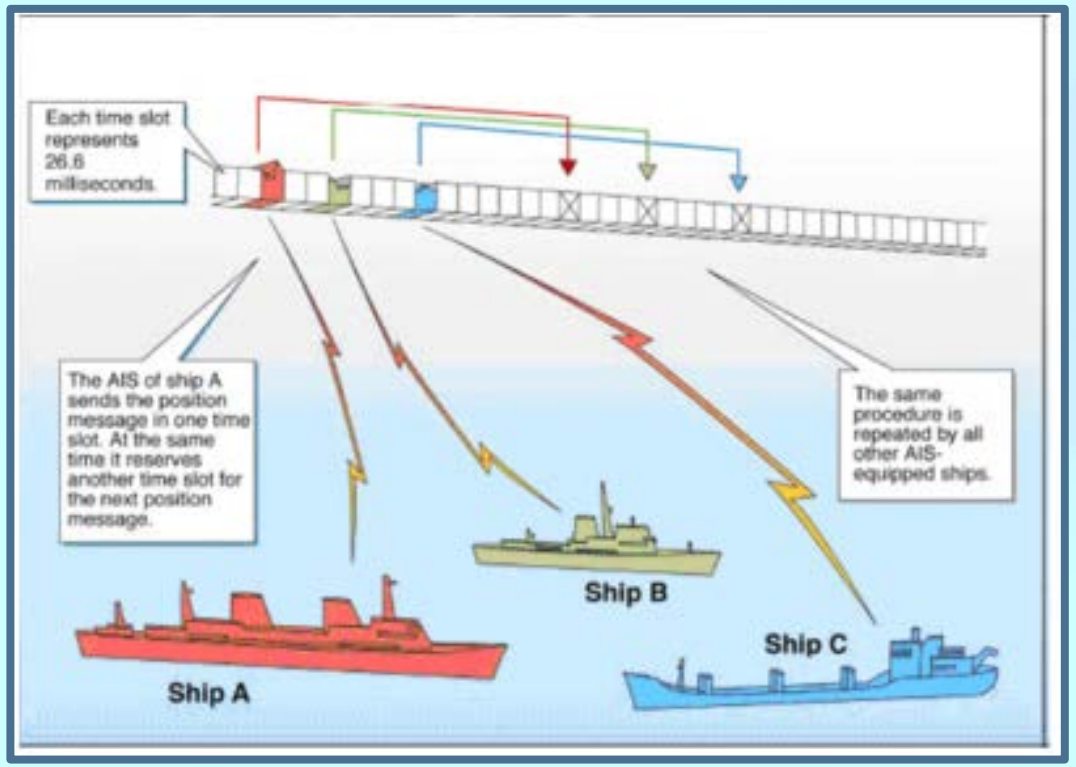


- Compiling raw AIS data
- AIS data processing
- Integrating AIS activity with vessel characteristics
- What to do about auxiliary engines and boilers
- Computing spatially resolved results from AIS and integrated data
- Overview of AIS-data costing, staffing and timelines

Compiling Raw AIS Data – Chris Leslie, PortVision

How AIS Works

- Autonomous and continuous Broadcasts via VHF transmitter
- 2 radio channels (87B, 88B)
- AIS stations sync to each other
- Data Tracking system (SOTDMA)
- Connected to ships GPS or GNSS
- Static info is manually programmed
- 20nm Range



Broadcast Information

Dynamic (positional) Data		Static & Voyage Data	
MMSI	Latitude/Longitude	IMO	Destination
Navigation Status	Time/UTC Seconds	Call Sign	ETA
Rate of Turn	True Heading	Vessel Name	Draught
Speed Over Ground	Course Over Ground	Vessel Type/Cargo Type	Dimensions (Length, Beam)

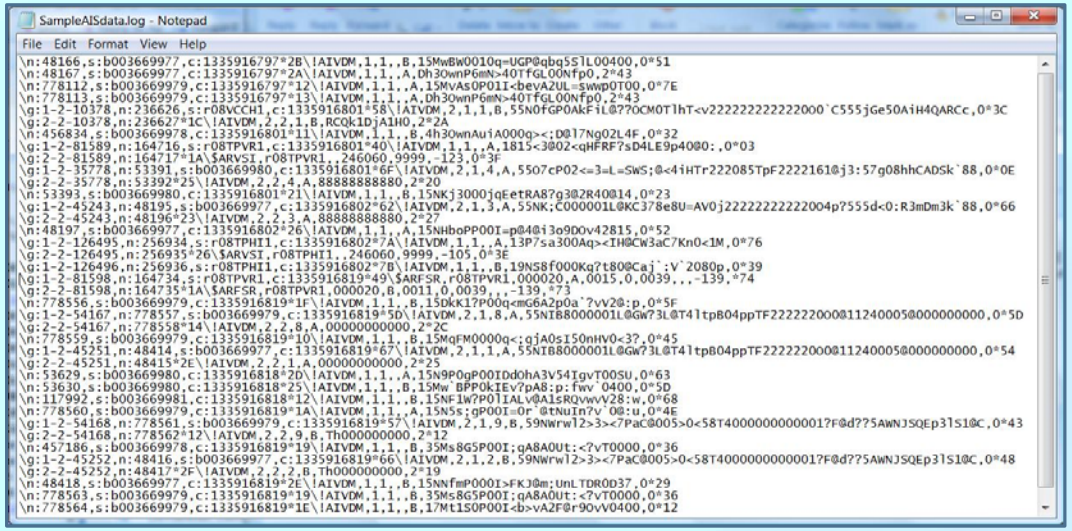
- Dynamic data broadcasts every 2-10 seconds while underway, every 3 minutes while at anchor (Class A)
- Static data broadcasts every 6 minutes (Class A)

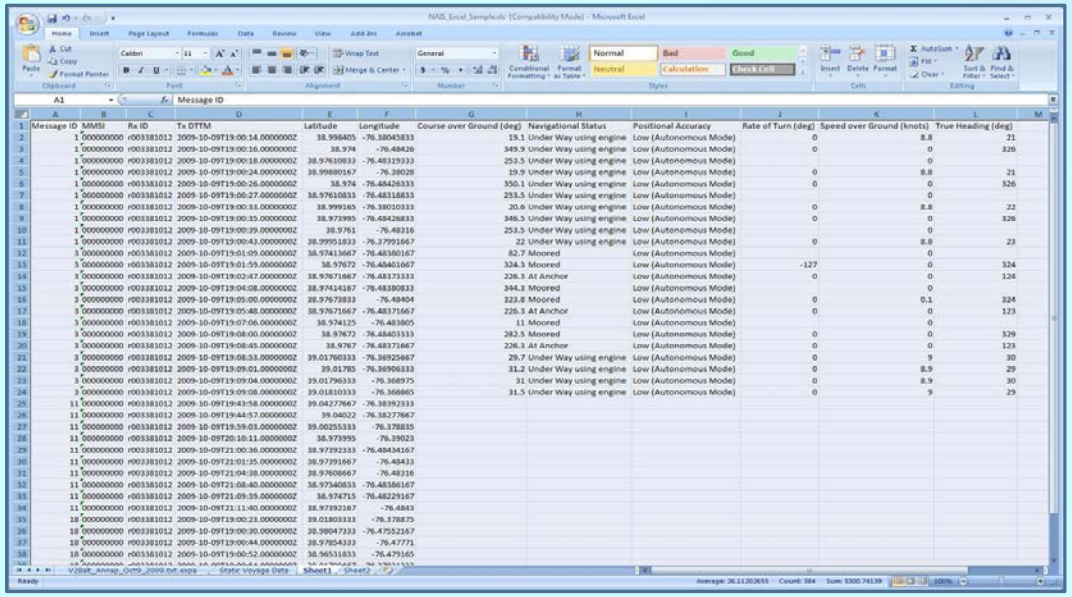
AIS Systems

- **Class A Transceiver**
  - **IMO Standard, Commercial Vessels, SOTDMA, 12.5 W, all AIS Messages**
- **Class B Transceiver**
  - **Lighter Commercial & Leisure Vessels, SOTDMA/CSTDMA, 2 W, all AIS Messages**
- Base Station
  - USCG operated, application specific messages, weather info, safety info
- AIS Aid to Navigation (ATON)
  - Application specific messages, every 3+ minutes
- AIS Search and Rescue Transmitter (SART)
  - Emergency distress beacon, maximum 5 miles

Combining AIS Data from Different Sources

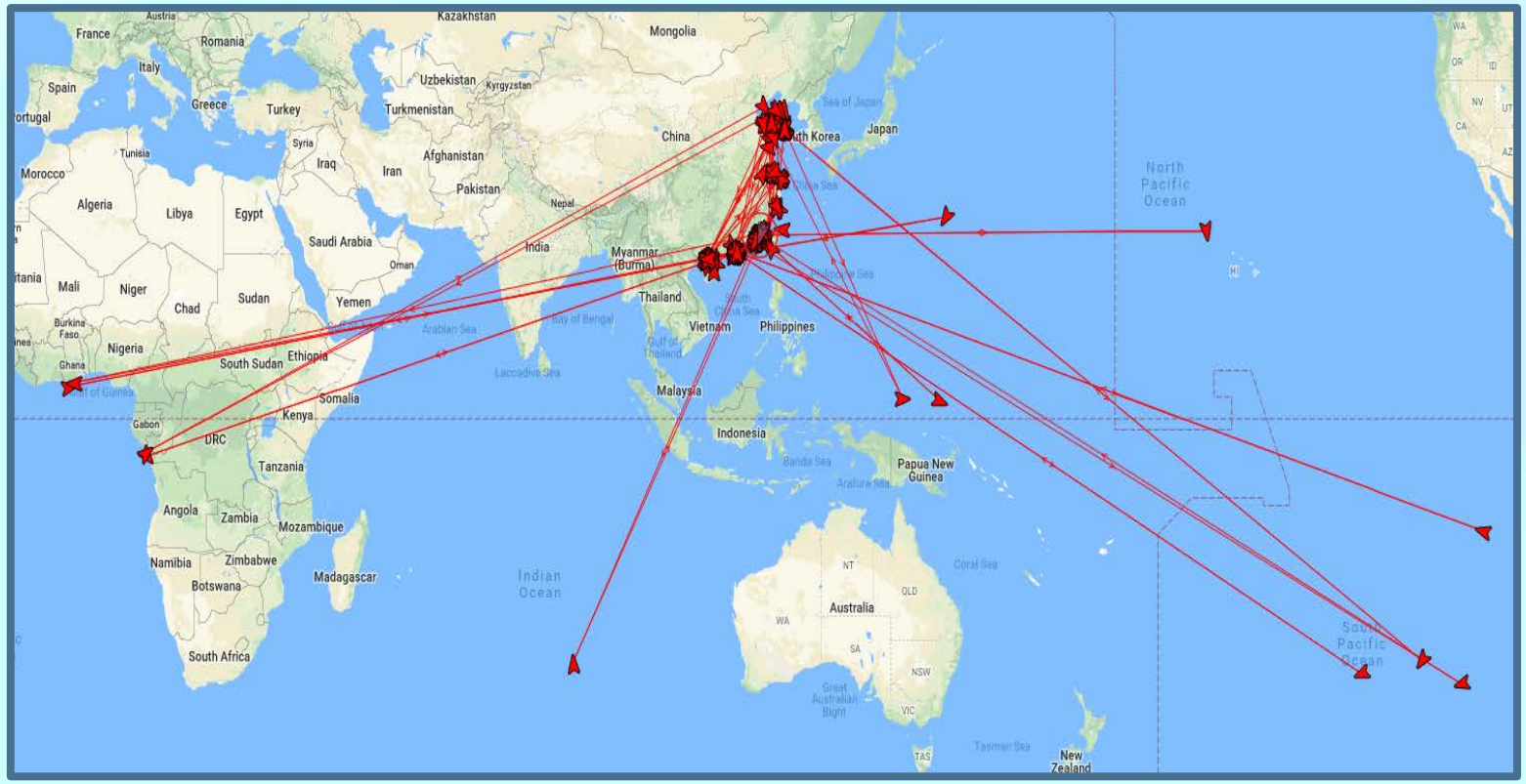
- Archive all AIS data from all sources
- Data source grouping
- Duplicate messages
- Throttling down data volume
- Ingest approach
- Record all invalid AIS messages





AIS Data Troubleshooting

- Latency of Satellite feeds
- Spoofing
- Typos
- “Gate Keeper Logic”
  - Geographic Feasibility
  - Incorrect IMOs
- 3rd party bulk data feeds



AIS Data for Air Quality Studies – Heather Perez, ERG

Temporal Resolution

- Capture activity
- Address data management challenges

Study Area

- Generous boundary
- Minimize edge effects

AIS Data Elements

- Vessel identifiers
- Date/Timestamp
- Geocoordinates
- Speed over ground

Stage 1: Basic AIS Data Cleansing

Identification and resolution of duplicate records

Missing/erroneous data elements

Remove “Ghost vessels”

Geographic Information Systems

- Validate spatial distribution of activity
- Identify data gaps
- Remove observations on land, etc.

MMSI	VESSEL NAME	IMO NUMBER	CALL SIGN	VESSEL TYPE	RECORD COUNT
205472642	MMSID: 205472642				2
205474019	MMSID: 205474019				2
205474035	MMSID: 205474035				2
205589584	MMSID: 205589584				3
205589608	MMSID: 205589608				3
205612010	MMSID: 205612010				1

Sample Data Overview

Total Observations: 2,435,032

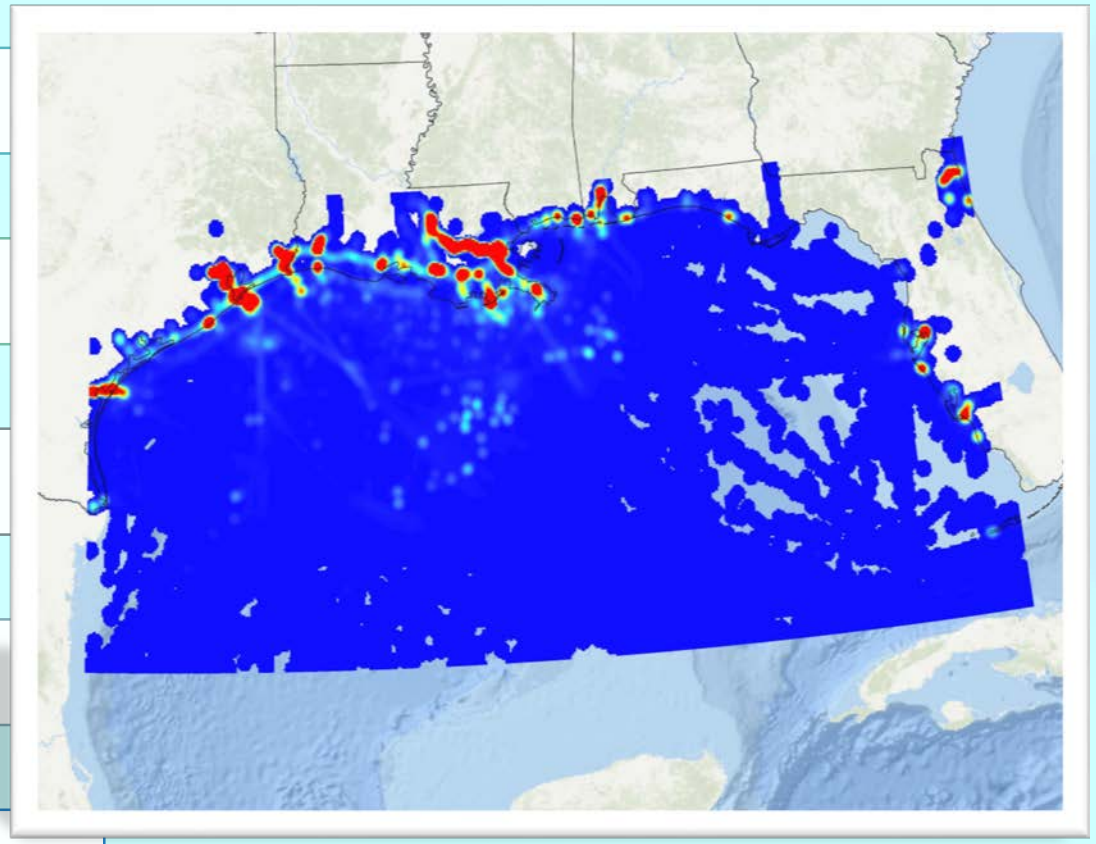
Unique MMSIs: 9,435

“Ghosts”: 1,800 (19%)

Max record count per MMSI: 744 (perfect)

Spatial Distribution: Good

Other issues: As Anticipated



Stage 2: Data Processing/Cleaning for Emissions Inventories

Order by Date/Timestamp

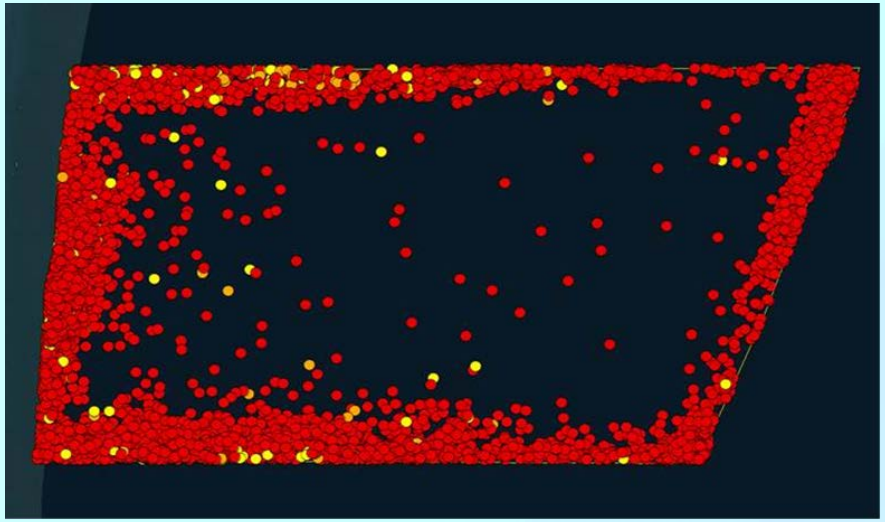
Calculate duration between observations*

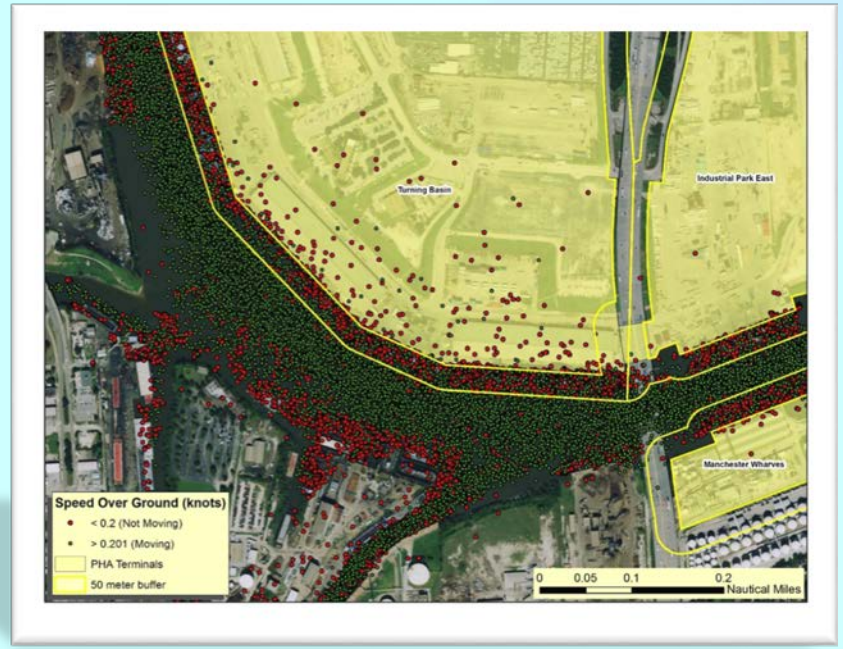
Investigate and adjust for:

- Long durations
- Unrealistic speeds
- Unrealistic speed/duration combinations

Extra attention to observations on land and near port areas

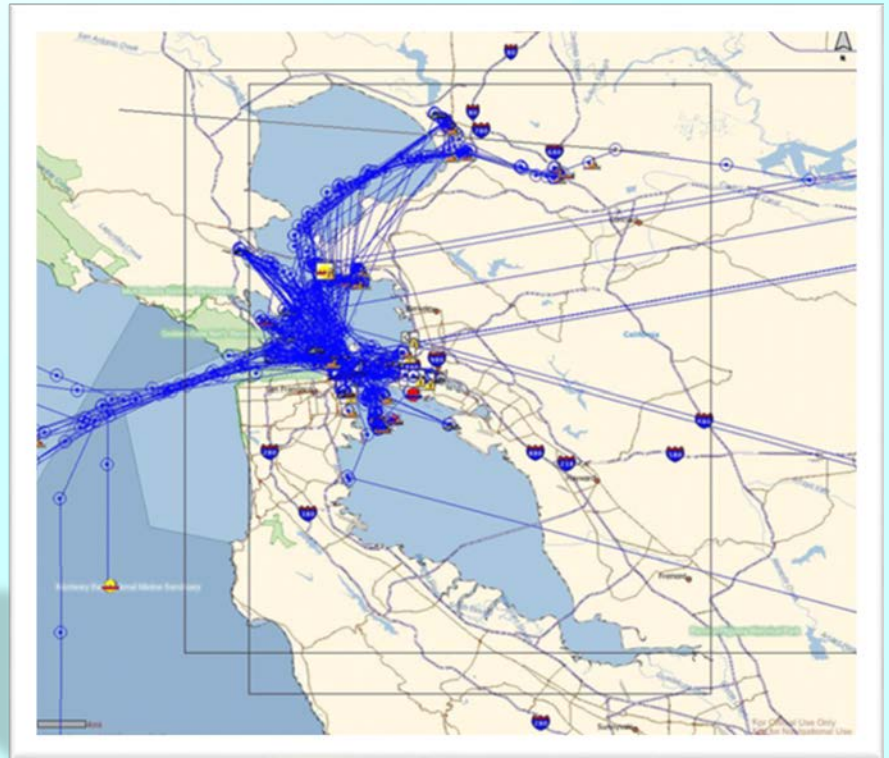
* Also allows for data consolidation

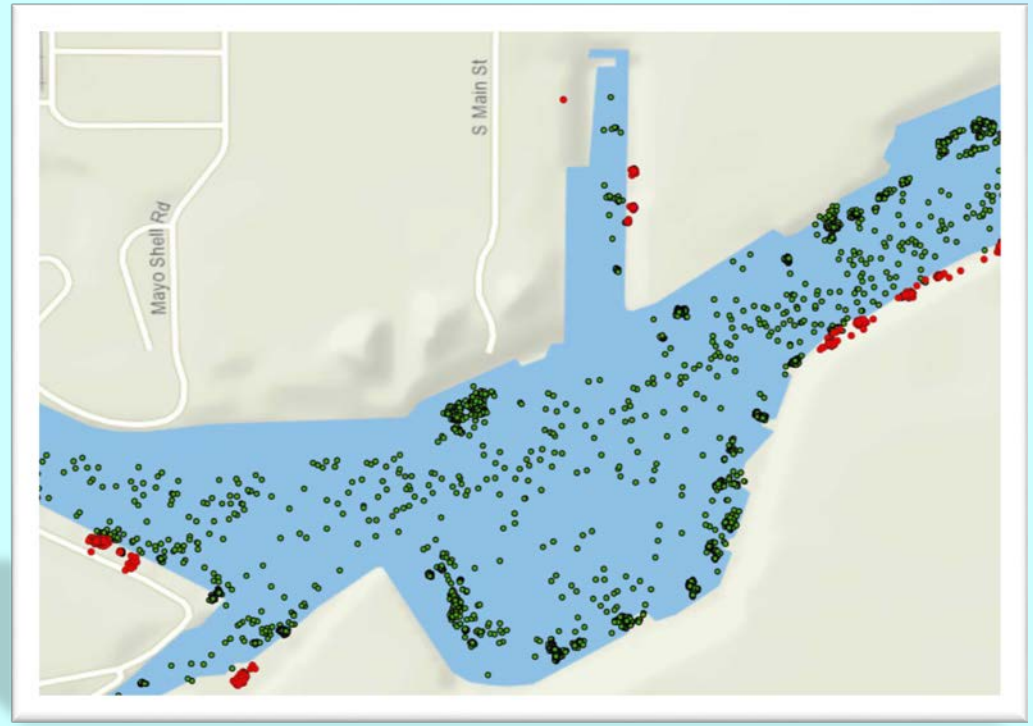




Special Case: Observations on Land

- Missing transmissions
- Deliberate tampering
- Horizontal accuracy variances
- Processing techniques





Take-Aways on AIS Data Cleansing

- Right resolution for inventory needs
- QA is an extensive, iterative process
- Geospatial and database skills required
- Prioritize CLEAN data over ALL data
- You’re only halfway there (if that)!

Data Cleansing for Emissions Inventories

AIS Data	Classification Society Data
• Duration from previous AIS observation (hours of operation) • Speed over ground • Spatial distribution of activity	• Match vessel IDs to characteristics databases • Gap-fill engine kW ratings and maximum speed • Refine vessel type, category, and build year

Integrating Activity and Vessel Characteristics Information – Ed Carr, EERA

- Why merge AIS activity data with vessel characteristics data?
  Data fields not present in AIS
  - Installed power
  - Design vessel speed
  - Cylinder diameter/stroke length (EPA engine Category determination)
  - Build year
  - Country of registration
  - Tonnage
  - Cargo capacity etc.
  - Passengers

- Classification Society
  - Establish and maintain a register of technical standards and vessel characteristics
  - International Association of Classification Societies (IACS)

Integrating AIS and Vessel Characteristics

Which fields to merge on?

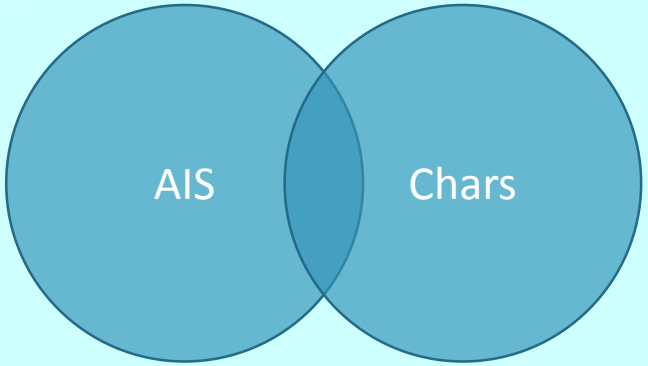
1. IMO Number
   - Permanent and unique number assigned to each vessel
   - Does not change on transfer of vessel
2. MMSI
   - Maritime Mobile Service Identity (MMSI)
   - Uniquely identifies a vessel
   - Can change on transfer of a vessel
3. Vessel Name
   - Non-unique vessel identifier
   - Can change over the life of the vessel
4. Call Sign
   - Uniquely identifies a vessel
   - Intermittent availability in AIS data

DateTimeUTC	MMSI	IMO	ShipName	CallSign
2013-10-01 00:00:00.000	215467000	6821080	LOUIS AURA	9HA2738
2013-10-01 00:00:00.000	309964000	7218395	BOUDICCA	C6VA3

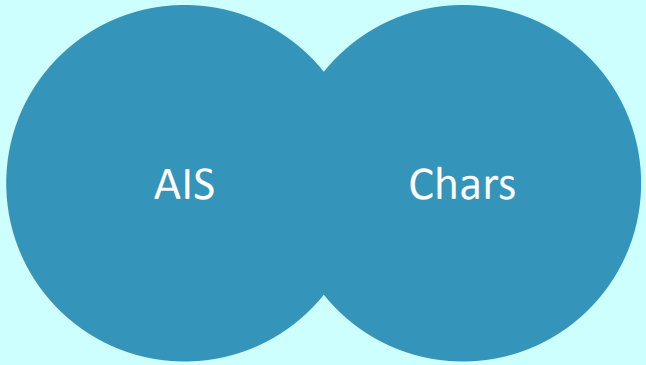
Integrating AIS and Vessel Characteristics

- Database Operations
  - Both AIS data and classification data are tabular with one or more fields in common
    - Join types vary depending on use case
    - Inner yields no Null values, but trims both AIS and Chars data
    - Full outer maintains all data, but yields Null values where AIS and Chars data don't match
    - Left outer keeps all AIS data and drops Chars data that doesn't match

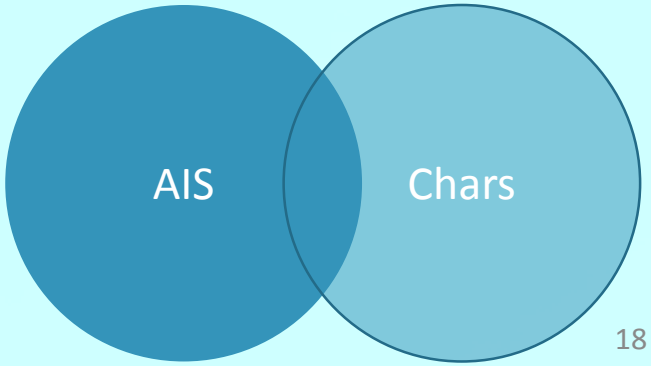
Inner



Full Outer



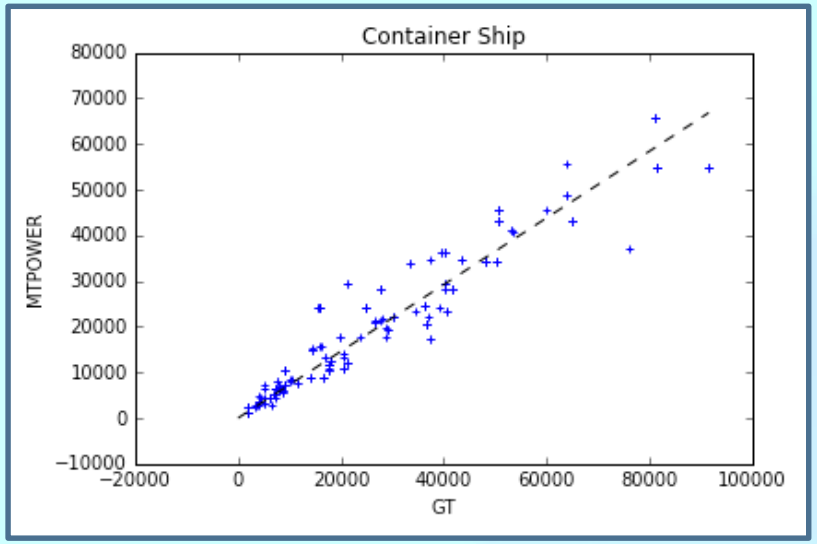
Left Outer

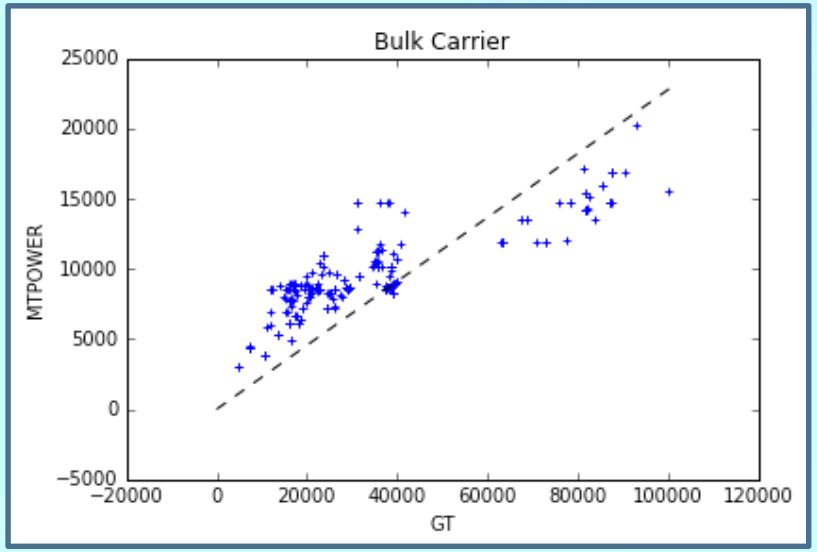


Integrating AIS and Vessel Characteristics

- Handling Missing Data
  - Very common when working with messy, real-world data

1. ~~Drop rows with missing data~~
   - Non-optimal approach as it crops the dataset
2. Gap fill with fleet averages
3. Gap fill using equations informed by vessel type, size, tonnage parameters
4. Obtain characteristics data from a secondary source

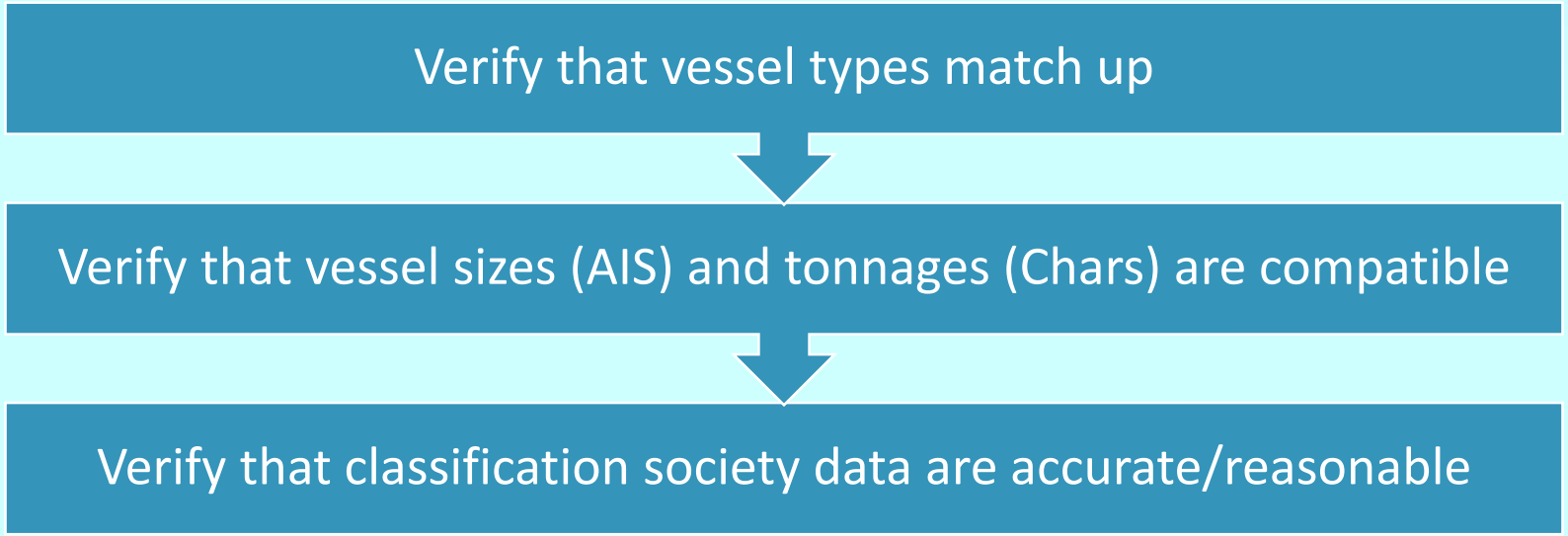




Integrating AIS and Vessel Characteristics

- Data Cleaning and Verification
  - Iterative process
  - Visually inspect
  - Use descriptive statistics

Example Flow Diagram



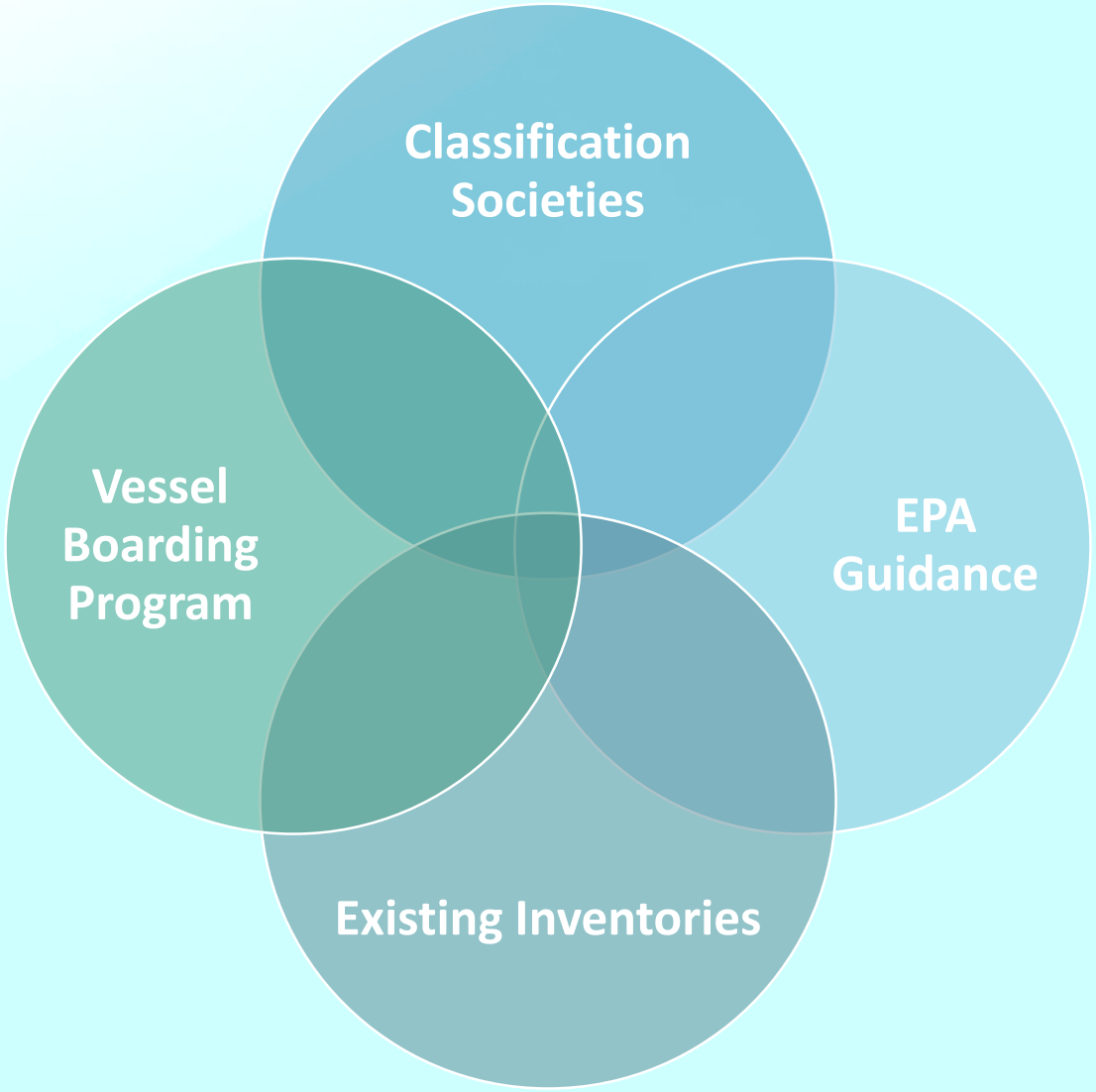
Auxiliary Engines and Boilers – Ed Carr, EERA

- Auxiliary engines
  - Provide power to vessel in addition to main engine
    - Lights
    - HVAC
  - Provide power at berth, unless shore power is used

- Auxiliary boilers
  - Steam generator, generates low pressure steam
  - Typically used when loads are < 20% MCR
    - Heat drawn from main engines in transit
  - Used to power various machinery, heat fuel, and produce hot water
  - Cruise ships and tankers often have high boiler loads in order to meet passenger and cargo demands

Auxiliary Engines and Boilers

- Data Sources



Auxiliary Engines and Boilers

- Importance
  - Auxiliary engines and boilers are important to consider in developing inventories
  - 2017 Port of Los Angeles Inventory
    - Auxiliary engines contributed ~35% of SO_X and 32% of CO_2e emissions
    - Auxiliary boilers contributed ~40% of SO_X and 41% of CO_2e emissions
  - Data can be scarce
    - Installed kW
    - Loads
  - Use best judgement for data sources

Developing Emissions – Tim Sturtz, Ramboll

Emission Factors

- Ocean-Going Vessel v. Harbor Craft
- Engine Type (e.g., medium speed diesel)
- Fuel Type

Vessel Specifications

- Vessel Type (e.g., cruise ship)
- Propulsion, Auxiliary, and Boiler Power
- Cruise Speed

AIS Records

- Vessel Identifier
- Position
- Segment Speed, Duration, and Distance

OGV By-Engine Emissions Development

Engine	Engine Type	Vessel Type*	Operating Mode	Fuel Type	Fuel Sulfur	Load Adj.**	ECA Adj.
Propulsion	✓	✓	✓	✓	✓	✓	✓
Auxiliary		✓	✓	✓	✓		✓
Boiler		✓	✓				

* For propulsion engines, vessel type is only needed if using EPA-provided cruise speeds
** Load adjustments not used for diesel electric or gas turbine electric engines

Harbor Craft Emissions Development

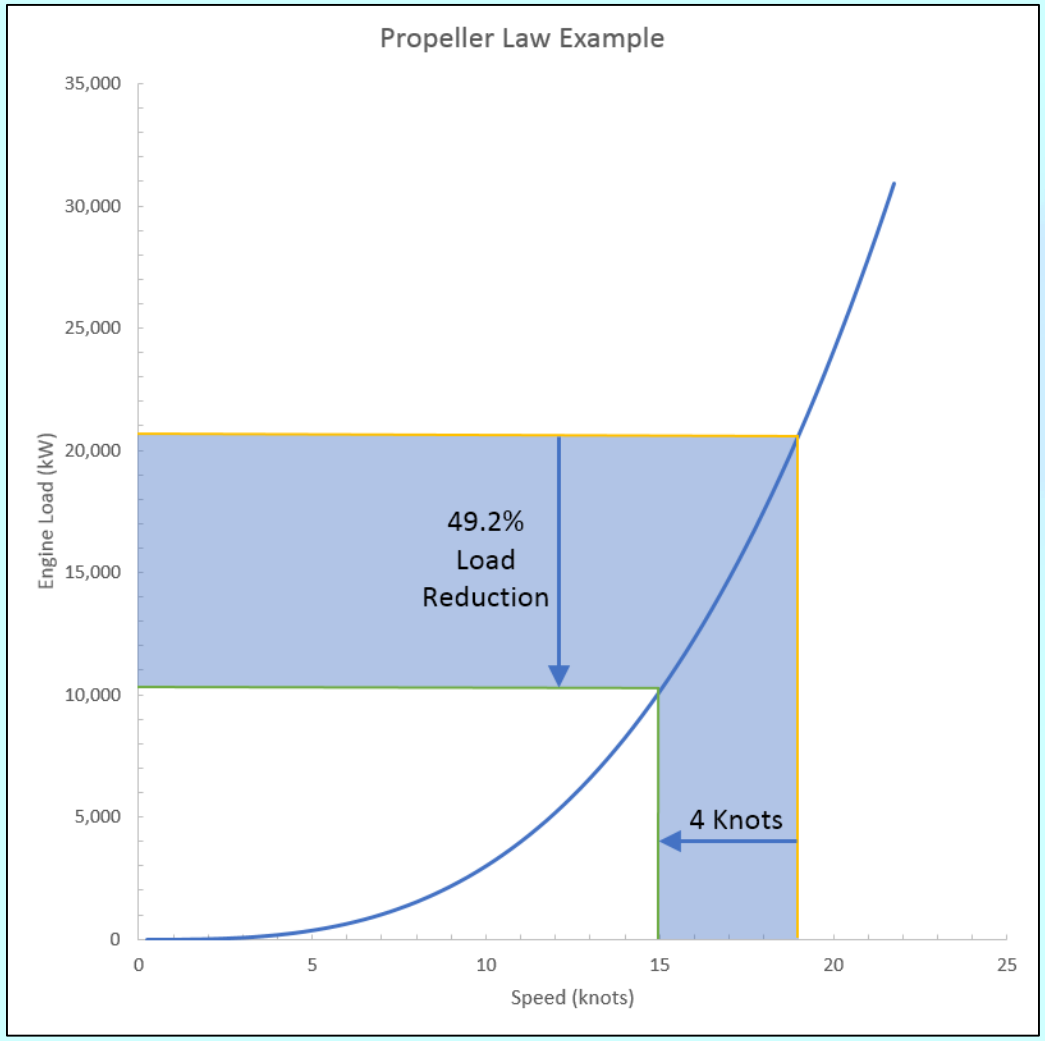
Engine	Engine Type*	Vessel Type	Fuel Type	Fuel Sulfur	Tier Adj.**
Propulsion	✓	✓	✓	✓	✓
Auxiliary		✓	✓	✓	✓

* Category 1 or 2

** EPA cautions that control factors should not be applied fleet wide, rather only to specific vessels

Propeller Law and Propulsion Engine Load

$$\text{Load} = \left(\frac{\text{Actual Speed}}{\text{Maximum Speed}}\right)^3$$



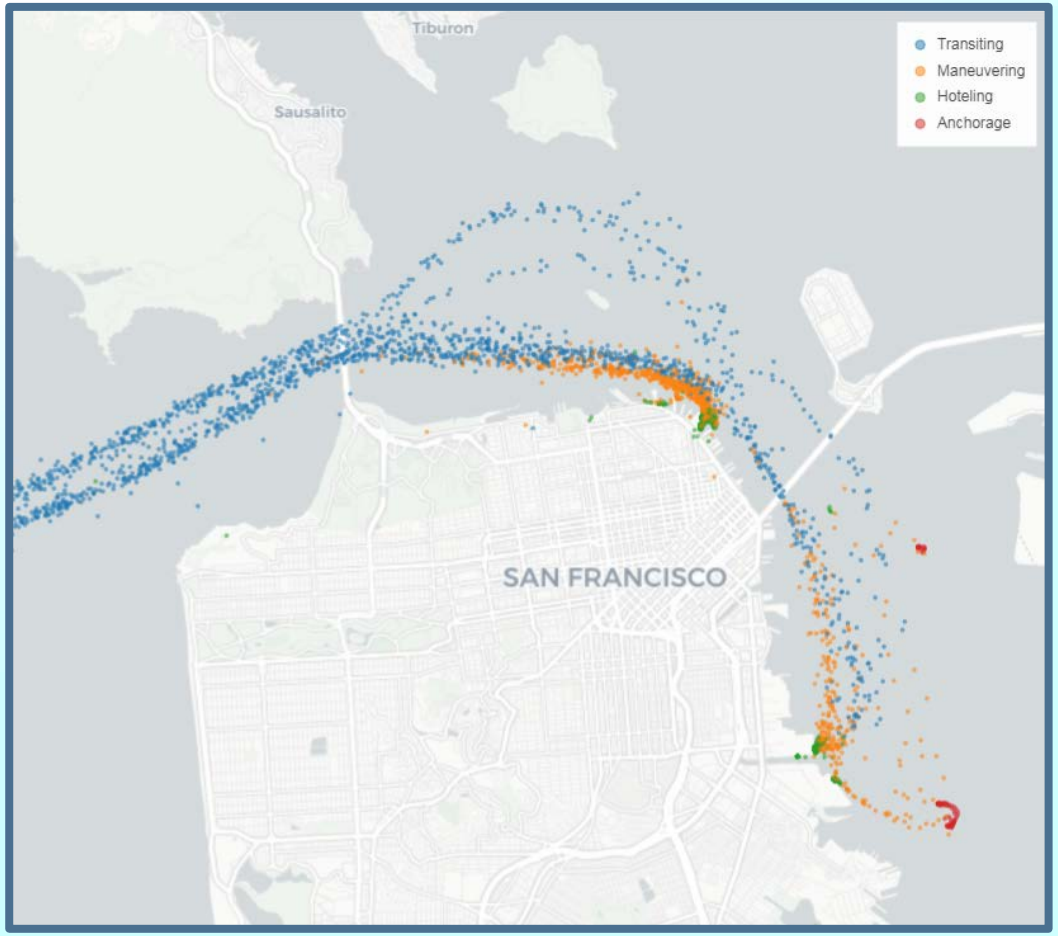
Calculating Emissions

- Identify EFs for each AIS record
- Multiply EFs by Durations
- Aggregate Emissions
- Review and QA

Example of Line-by-Line Emission Tabulations

NAME	vType	PERIOD	Dur	NOx	PM25	CO2
SILVER WHISPER	Cruise Ship	2017-01-06 12:40:00	00:05:00	0.004236	0.000065	0.233047
ZENITH LEADER	RORO	2017-01-13 11:10:00	00:05:00	0.003334	0.000054	0.163806
ARCADIA	Cruise Ship	2017-01-25 19:45:00	00:05:00	0.029252	0.000457	1.794465
AZALEA ACE	RORO	2017-01-26 09:50:00	00:05:00	0.002936	0.000042	0.146728
QUEEN ELIZABETH	Cruise Ship	2017-02-03 12:15:00	00:05:00	0.035661	0.000560	2.185734
GLOVIS COURAGE	RORO	2017-02-10 10:30:00	00:05:00	0.006126	0.000078	0.274735
GLOVIS CHALLENGE	RORO	2017-02-25 09:20:00	00:05:00	0.003377	0.000049	0.168573
ISLAND PRINCESS	Cruise Ship	2017-03-15 11:35:00	00:05:00	0.021326	0.000355	1.316388
RT HON PAUL E MARTIN	Bulk Carrier	2017-03-31 07:00:00	00:05:00	0.004653	0.000060	0.212981
GLOVIS CARAVEL	RORO	2017-04-14 08:25:00	00:05:00	0.003118	0.000047	0.157905
OPAL ACE	RORO	2017-04-29 07:10:00	00:05:00	0.002967	0.000046	0.149075
NORWEGIAN JEWEL	Cruise Ship	2017-05-02 11:20:00	00:05:00	0.023872	0.000407	1.450728
CRYSTAL SERENITY	Cruise Ship	2017-05-05 17:35:00	00:05:00	0.020379	0.000325	1.265354
EUROPA	Cruise Ship	2017-05-06 20:10:00	00:10:00	0.014628	0.000246	0.897321
WISDOM ACE	RORO	2017-05-09 09:20:00	00:05:00	0.004084	0.000071	0.202847

Visualizing AIS Data and Results



- Many visualization tools
  - → ESRI ArcGIS
  - → QGIS
  - → Python, R, Java and other languages
- Allows for validation of calculations
  - → Vessels properly spatially located?
  - → Operating mode identified correctly?
  - → Vessel density as expected?
  - → Strange hot-spots of emissions?
  - → Vessel type patterns appropriate?
- Interactive displays allow for further validation

Visualizing AIS Data and Results

- Many visualization tools
  - → ESRI ArcGIS
  - → QGIS
  - → Python, R, Java and other languages
- Allows for validation of calculations
  - → Vessels properly spatially located?
  - → Operating mode identified correctly?
  - → Vessel density as expected?
  - → Strange hot-spots of emissions?
  - → Vessel type patterns appropriate?
- Interactive displays allow for further validation

Applying Results

Emission Inventory

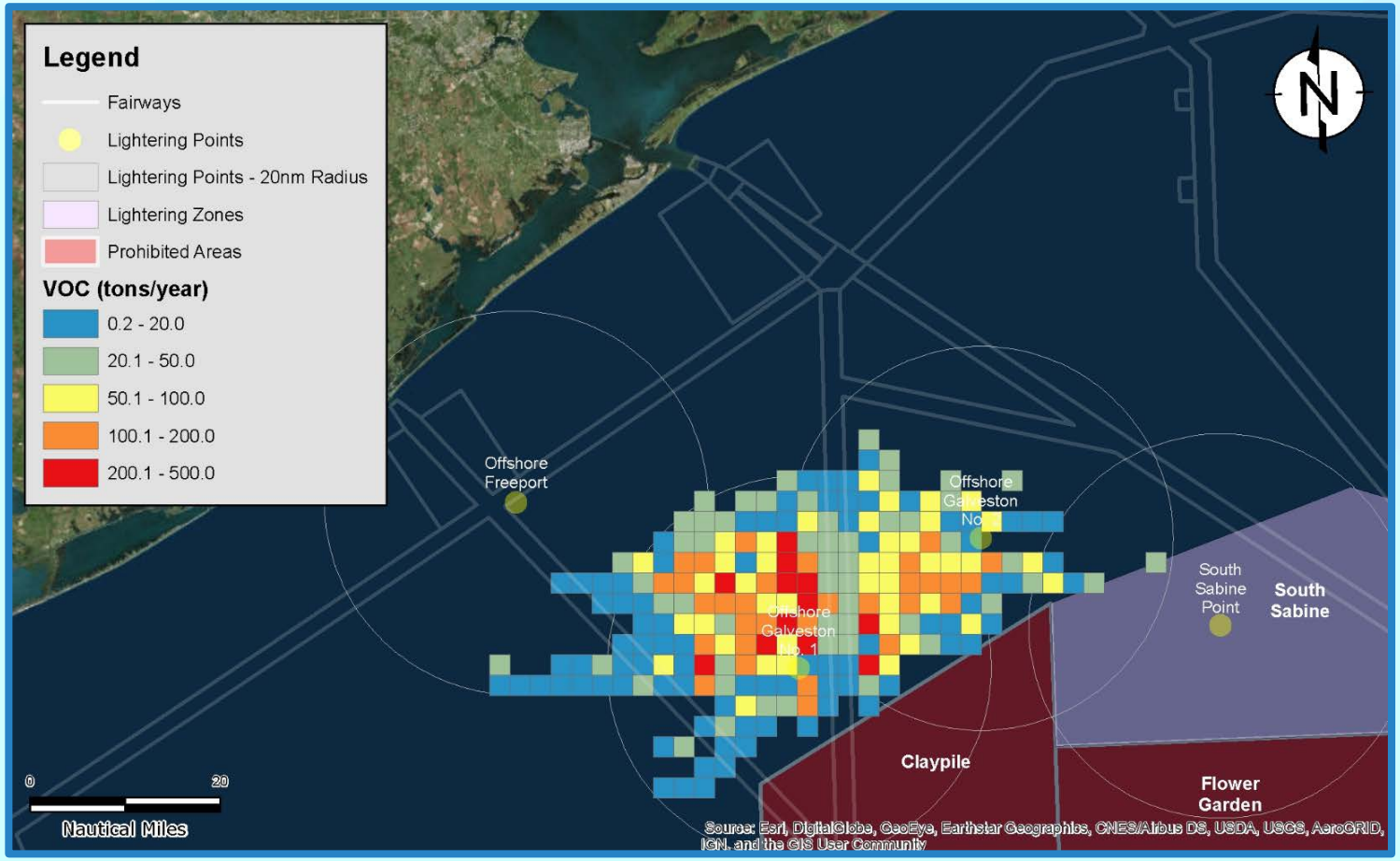
- Emission Summaries
- Temporal Trends
- State Implementation Plans

Modeling

- Impact Assessments
- Photochemical Modeling

Control Strategies

- Vessel Speed Reduction Programs



Overview of AIS-Data Costing, Staffing and Timelines – Richard Billings, ERG

- AIS data
  - Vendor compiled data
  - U.S. Coast Public data set
  - NOAA AIS data

- Vessel characterization data
  - Need to check on CBI issues

- Labor
  - Data prep, vessel matching, and gap filling
  - QA of compiled data and results
  - Documentation of process

- Schedule
  - Data cleaning
  - Vessel matching
  - QA activities

Today's Participants

- Richard Billings, *Eastern Research Group,* richard.billings@erg.com
- Chris Leslie, *Oceaneering,* cleslie@oceaneering.com
- Heather Perez, *Eastern Research Group,* heather.perez@erg.com
- Edward Carr, *Energy and Environmental Research Associates,* ewcarr@gmail.com
- Timothy Sturtz, *Ramboll,* tsturtz@ramboll.com

Panelists Presentations

http://onlinepubs.trb.org/onlinepubs/webinars/180918.pdf

After the webinar, you will receive a follow-up email containing a link to the recording

Get Involved with TRB

- Getting involved is free!
- Join a Standing Committee (http://bit.ly/2jYRrF6)
  - Search for AW030 (Standing Committee on Marine Environment)
- Become a Friend of a Committee (http://bit.ly/TRBcommittees)
  - Networking opportunities
  - May provide a path to become a Standing Committee member
- For more information: www.mytrb.org
  - Create your account
  - Update your profile

Receiving PDH credits

- Must register as an individual to receive credits (no group credits)
- Credits will be reported two to three business days after the webinar
- You will be able to retrieve your certificate from RCEP within one week of the webinar

TRB turns 100 on November 11, 2020

100 YEARS 2020

TRB

Help TRB:

- Promote the value of transportation research;
- Recognize, honor, and celebrate the TRB community; and
- Highlight 100 years of accomplishments.

Learn more at

www.TRB.org/Centennial

MOVING IDEAS: ADVANCING SOCIETY—100 YEARS OF TRANSPORTATION RESEARCH

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

TRB
TRANSPORTATION RESEARCH BOARD