



4 ALBERT EMBANKMENT
LONDON SE1 7SR
Tel: +44 (0)20 3970 1060 www.imso.org

Circular Letter No. 57
27 April 2026

To: All IMSO Parties
IMSO Observers

Subject: **Outcome of IMSO study on the operational impact of full exercise of Coastal State entitlement on the LRIT system**

1 The Director General presents his compliments and hereby informs IMSO Parties and Observers of the outcome of an analysis conducted by the IMSO Directorate on the operational impact of a full exercise of Coastal State entitlements on the Long-Range Identification and Tracking of Ships (LRIT) system.

2 The aim of this study is to provide an analytical response to questions raised on this matter at the fifty-second session of the Advisory Committee, which was held on 16 and 17 March 2026, in the context of the upcoming discussion at the 111th session of the International Maritime Organization's (IMO's) Maritime Safety Committee (MSC) on the financial architecture of the LRIT system.

3 The analysis enclosed to this Circular Letter reflects a technically complex exercise, requiring the mobilisation of significant internal resources, including in-house modelling, large-scale data processing as well as a dedicated software development. Owing to the breadth of the dataset and the depth of analysis required, the study could not be produced within a shorter timeframe.

4 The Directorate nevertheless undertook this work on an accelerated basis in order to provide IMSO Parties and Observers with a comprehensive and reliable assessment of the operational implications under consideration.

IMSO study on the operational impact of full exercise of Coastal State entitlement on the LRIT system

Executive Summary

1 The study examines the operational implications for the LRIT system in the event that all Contracting Governments were to fully exercise their coastal State entitlement under SOLAS regulation V/19-1. It is based on data collected from all 71 Data Centres, covering 135 flag States and a total of 39,517 ships, using 30-day audit samples distributed across the full calendar year. The dataset was reviewed to ensure that it provides a reliable basis for annual estimates, with no indication of seasonal bias affecting the results.

2 On this basis, the analysis indicates that the volume of messages exchanged through the IDE would reach approximately 214 million per year under current operating conditions, and up to 334 million in a scenario of full transmission compliance. This represents an average processing load of between 6.8 and 10.6 messages per second, remaining well within the current system capacity of 100 messages per second. Occasional short-lived peaks above this threshold were identified, typically linked to external factors, but these remain limited in duration and can be addressed through standard system configuration.

3 While an increase in data storage would be required, the volumes involved remain limited and do not give rise to any significant cost implications. Overall, the analysis indicates that the LRIT system is capable of supporting full use of coastal State entitlements within the existing technical framework, subject to appropriate configuration and scaling where necessary.

Study methodology

4 The study is set out in the two annexes: **Annex 1** presents the technical findings regarding the operational impact on the International Data Exchange (IDE) and Data Centres (DCs), and **Annex 2** documents the statistical validity of the data sample used.

Sample data

5 The sample data was drawn from the 2025 LRIT audit schedule and covers all 71 DCs, 135 flag States and 39,517 LRIT-compliant ships. Each DC sample consists of 30 days of operational traffic data; the 71 DC samples are distributed across all 12 calendar months of the year, spanning from July 2024 to August 2025. The aggregated 30-day sample comprises 3,042,131 standard six-hourly position reports.

6 This approach reflects the established IMO LRIT audit process, under which each Data Centre provides a representative 30-day sample of operational data for performance assessment. While individual Data Centres do not provide continuous 12-month datasets, the aggregation of one-month samples across all Data Centres, distributed over the full calendar year provides a representative and statistically robust basis for extrapolation to annual estimates (a process known as 'annualization').

Statistical validation and annualization

7 Prior to annualisation, three non-parametric diagnostic tests were applied to confirm that the sample is suitable for extrapolation to annual estimates: a Kruskal-Wallis test for seasonal assignment bias ($H = 9.65$, $p = 0.56$), a Spearman rank correlation for monotonic seasonal trend ($r_s = -0.20$, $p = 0.09$), and a jackknife resampling procedure for estimating the bias and variance of the sample. All three tests confirm the absence of any seasonal distortion of practical significance.

8 The sample contains 3,042,131 position reports against a theoretical maximum of 4,742,040 thus indicating a fleet-wide transmission completeness rate (TCR) of 64.2%. This shortfall is consistent with the finding in document NCSR 12/INF.3 that approximately 36% of ships experienced at least one reporting gap of 24 hours or more within a 30-day period, reflecting authorised transmission reductions, equipment issues and non-compliant interruptions. Two annualised estimates are therefore established: a theoretical upper bound (57,694,820 transmissions per year; 95% CI: $\pm 3.3\%$) used for infrastructure dimensioning, and an empirical estimate (37,043,672 transmissions per year; 95% CI: $\pm 3.3\%$) representing expected operational load.

Message routing analysis

9 For each position report in the sample data, a 'point-in-polygon' test was applied against the LRIT Data Distribution Plan (DDP) in force at the moment of receipt to identify all CGs from other DCs within whose 1,000 NM coastal zone the transmitting vessel was located. One database record was created for each such entitled CG, representing one message that would be routed through the IDE. The analysis covers only standard six-hourly periodic transmissions; on-demand polled requests are excluded.

Key findings

10 The analysis indicates that the LRIT system architecture is capable of accommodating a full exercise of coastal State entitlements under the current technical framework, subject to some adjustments further discussed below.

Processing capacity

Average load

11 Under a scenario based on an analysis of the actual data, henceforward referred to as the 'operational scenario', the IDE would process an average of 6.8 messages per second.

12 Under an ideal scenario in which every ship transmits four times a day without gaps (i.e. TCR = 100%), henceforward referred to as the 'theoretical maximum', the average rises to 10.6 messages per second.

13 In both cases, the average load remains well within the current processing capacity requirements of 100 messages per second.

Peak load

14 The analysis identified short-term burst events reaching up to 292 messages per second which were associated with connection disruptions in the LRIT or communication service provider infrastructure, causing buffered messages to be released simultaneously.

15 Such peak events represent 0.01% of operational time and are therefore limited in duration. The analysis indicates that these peaks can be effectively managed through the appropriate configuration of connection pool sizes and application-layer timeout parameters.

Storage capacity

International Data Exchange

16 The annual raw XML data received by the IDE is estimated to increase from approximately 33 GB to 500 GB under the operational scenario, representing an approximately 15-fold increase in the annual storage requirement for periodic LRIT position reports.

17 This increase does not constitute a limitation of the LRIT system architecture but requires appropriate scaling of storage capacity.

Data Centre outgoing traffic

18 Each standard position report received by a DC would, on average, generate 7 outgoing messages to other CGs' DCs. As an illustration, for every 100 ships a DC manages, a provision for approximately 1.5 GB of additional annual raw data storage for outgoing traffic under full coastal entitlement exercise.

19 A linear calculation based on this number will yield a rough estimate that supports impact assessment.

Potential upgrade

20 The study suggests that DCs would require an additional storage capacity of 1.5GB per 100 ships per year for the outgoing messages and the IDE around 500GB per year. These data storage capacity may already be available considering that it was assumed when the LRIT system was designed that Contracting Governments would actively monitor their coastal waters however even if more storage is needed the cost is negligible (around \$50 per year per 1000GB for an on-premises solution).

Quantitative findings

21 The principal quantitative findings are summarised in **table 1** below.

Table 1 - Summary of key findings

Metric	Value
Ships in scope (complete LRIT transmitting population)	39,517
LRIT position reports in the aggregated sample	3,042,131
Average coastal State matches per position report	5.7
Estimated annual IDE message volume (operational)	214,191,806
Estimated annual IDE message volume (theoretical maximum)	334,190,386
Average IDE throughput (operational scenario)	6.8 messages/sec
Average IDE throughput (theoretical maximum)	10.6 messages/sec
Observed peak throughput (sample data)	292 messages/sec
Percentage of time exceeding IDE capacity (100 message/sec)	0.01%
Current IDE processing capacity requirement	100 messages/sec

Metric	Value
Estimated increase in annual IDE storage (raw)	~15 × current annual increment
Contracting Governments receiving data as coastal States	124

Redelivery mechanisms

22 The current LRIT technical documentation does not require DCs to implement message redelivery upon failure. Earlier specifications applicable to the International LRIT Data Centre (IDC) included this requirement, but it was not retained when those provisions were removed. The presence or absence of DC-level redelivery mechanisms affects the resilience of message delivery to the IDE during peak periods.

23 Clarification or further guidance on these aspects would enhance system resilience, particularly in managing short-duration peak loads and ensuring continuity of message delivery.

Conclusions

24 Based on the outcome of the study, the existing technical requirements and the current ship reporting performance, it is concluded that the LRIT system architecture is capable of accommodating a full exercise of coastal State entitlements.

25 The study highlights that:

- .1 with regard to message processing capacity, the existing requirements for the IDE are sufficient to accommodate the increased message flow if all Contracting Governments decides to exercise their coastal State entitlements.
- .2 the peak of transmissions, which are usually caused by a shoreside component such as DC or CSPs, are short and infrequent, and can be managed through appropriate configuration of system parameters.
- .3 the projected increase in data volume may require appropriate scaling of storage capacity at both IDE and Data Centre level if needed. The financial impact would be negligible.

26 The study confirms the importance of a redelivery mechanism for the IDE and the DCs, especially when there is a peak in the number of messages sent or received.

27 The study outcome includes data on an estimated increase in incoming and outgoing message flow for individual DCs which is not published in this document. However, the DCs and Contracting Governments may consult the IMSO Directorate for detailed information regarding their DCs.

ANNEX 1

TECHNICAL REPORT

Operational impact of maximum LRIT periodic information exchange under the coastal state entitlement

Contents

List of Acronyms	3
1 Executive Summary	4
2 Background and Regulatory Context	4
2.1 <i>LRIT Entitlement Under SOLAS Regulation V/19-1</i>	4
2.2 <i>The LRIT Data Distribution Plan (DDP)</i>	5
2.3 <i>Current operational status of the LRIT system</i>	5
2.4 <i>LRIT Capacity Implications and Mitigation Mechanisms for Increased Traffic</i>	6
Pool of Connections and Processing Capacity in Distributed Communication Systems..	6
Session Lifecycle at the DC-IDE Interface	6
LRIT Governance	7
3 Questions to be answered	7
4 Methodology	7
4.1 <i>Data Sample</i>	7
4.2 <i>Limitations and Caveats</i>	8
4.3 <i>Treatment of non-standard-rate transmissions</i>	8
4.4 <i>Analysis Procedure</i>	8
4.5 <i>Operational and theoretical load under full coastal entitlement exercise</i>	9
5 Aggregated data overview	9
6 Results	11
6.1 <i>Average Load Estimation</i>	11
6.2 <i>Temporal Analysis — Transmission Peak Times</i>	11
6.3 <i>Increase in storage for IDE</i>	14
6.4 <i>Operational impact for Data Centres</i>	15
7 Conclusions	15
List of Acronyms	3
1 Executive Summary	4
2 Description of the Sample	4
3 Statistical Assessment of Sample Quality	5
3.1 <i>Nature of the Dataset</i>	5
3.2 <i>Scope of the Dataset</i>	5
3.3 <i>Distribution of Fleet Size Across Data Centre</i>	5
3.4 <i>Check for Seasonal Assignment Bias: Kruskal-Wallis Test</i>	6

- 3.5 *Check for Monotonic Seasonal Trend: Spearman Rank Correlation*..... 6
- 3.6 *Temporal Representativeness: Jackknife Estimate* 7
- 3.7 *Annual Position Transmissions, Transmission Completeness and Confidence Interval*7
- 4 Summary of Statistical Quality Metrics9**
- Appendix 1 — Step-by-Step Statistical Calculations 12**
 - 1 Distributional Analysis 12**
 - 2 Kruskal-Wallis Test Calculation 12**
 - 3 Spearman Rank Correlation Calculation 13**
 - 4 Jackknife Resampling Calculation 13**
 - 5 Annual Position Transmissions and Confidence Interval 14**
- Appendix 2 — Sample dataset 15**

List of Acronyms

DC	LRIT Data Centre
DDP	Data Distribution Plan
IDE	International Data Exchange
IMO	International Maritime Organization
LRIT	Long-Range Identification and Tracking
NM	Nautical Mile
SOLAS	International Convention for the Safety of Life at Sea
UTC	Coordinated Universal Time

1 Executive Summary

1.1 This report presents the findings of an analysis conducted by the LRIT Coordinator to estimate the volume and geographic distribution of LRIT messages that would flow through the International LRIT Data Exchange (IDE) if all Contracting Governments were to exercise, in full, their entitlement to receive LRIT information under SOLAS regulation V/19-1.

1.2 The analysis is based on sample data drawn from the 2025 audit schedule, covering 71 Data Centres (DCs). For each LRIT position report received in the sample, the geographic coordinates were cross-referenced against the LRIT Data Distribution Plan (DDP) in force at the moment of Timestamp 4, to determine all Contracting Governments (CG) within whose 1,000-nautical-mile coastal zone the transmitting vessel was located at that time. A database record was created for each such entitled CG, representing one message that, under a fully exercised entitlement regime, would be routed through the IDE to the LRIT Data Centre (DC) associated with the CG. The key findings are displayed in table 1:

Table 1 - Analysis key findings

Metric	Value
Total LRIT positions analysed (all DCs)	3,042,131
Total database records generated	17,604,806
Average records per position	5.7
Estimated annual IDE message volume (extrapolated)	214,191,806
Peak transmission hour (message/sec)	292 message/sec
Number of Contracting Governments as data recipients	124 SOLAS Contracting Governments ¹
Data Centres covered in the sample	71

1.3 The results offer a quantitative foundation for a comprehensive evaluation of the operational impact, concerning IDE throughput, DC ingestion load, and routing infrastructure, of a scenario wherein all Contracting Governments actively receive the LRIT information to which they are entitled under SOLAS.

2 Background and Regulatory Context

2.1 LRIT Entitlement Under SOLAS Regulation V/19-1

2.1.1 SOLAS regulation V/19-1 establishes the obligations of ships to transmit LRIT information and the rights of CGs to receive it. CGs are entitled to receive LRIT information in three distinct capacities: as a flag State, in respect of ships entitled to fly their flag; as a port State, in respect of ships proceeding to a port or place under their jurisdiction; and as a coastal State, in respect of ships navigating within 1,000 nautical miles of their coast, regardless of their flag.

2.1.2 Coastal State entitlement is the primary factor driving message volume in a fully operational LRIT system. Because these entitlement zones extend up to 1,000 NM from the coast of each CG, they overlap extensively across the world's oceans. As a result, a single ship's position can simultaneously fall within the zones of multiple CGs, each independently entitled to receive that LRIT information, directly increasing the volume of messages flowing through the system.

¹ 10 non-metropolitan territories and two special administrative regions are part of the LRIT, and they may be associated with LRIT DCs other than those used by their Contracting Governments.

2.1.3 To receive LRIT information in a given area, a CG must establish a Coastal State Standing Order (CSSO) in the IMO GISIS system for its LRIT polygons representing the geographic area it wishes to monitor. When processing position reports from their associated ships, DCs check each ship's location against the existing CSSOs to determine which LRIT Data Users (LDUs) are entitled to receive that information. This entire process runs automatically, without human intervention, on the DC servers.

2.2 The LRIT Data Distribution Plan (DDP)

2.2.1 CGs are required to upload information on their Internal Waters, Territorial Waters, and 1000 NM zones to the IMO GISIS database. This information is represented as polygons in GML format and distributed to all DCs via the DDP file and its subsequent updates. For the purposes of this analysis, the version of the DDP in force at the time of each Timestamp 4 event was used to determine whether a given ship position fell within a particular government's entitlement zone. This approach ensures that the analysis accurately reflects the state of the DDP at the moment each message was received by the DC, thereby accounting for any updates made to the plan throughout the sample period.

2.3 Current operational status of the LRIT system

2.3.1 In practice, only a very small number of CGs have configured CSSOs covering the full extent of their 1,000 NM entitlement zone. Most CGs instead limit data receipt to ships flying their own flag, making no use of their coastal entitlement. As a result, the LRIT traffic volume currently observed on the IDE represents only a fraction of what would be generated if all CGs fully exercised their entitlement.

2.3.2 The purpose of this analysis is to quantify that gap, providing the LRIT Coordinator and Contracting Governments with a data-driven estimate of the operational implications of full entitlement exercise.

2.3.3 As a reference, in 2025, the IDE received and processed 13,479,083 periodic messages under CSSO, including those addressing the IDF. For this study, the IDF was included because it significantly affects message flow.

2.3.4 The flow of messages in 2025 indicates that 88.2% of all messages in the system are LRIT information provided under CSSO, 0.7% are provided in response to LRIT requests, such as port requests or SAR requests, and 10.1% are other messages, such as LRIT Requests, system status, and receipt messages.

2.3.5 The requirements for the IDE related to timely message processing are: to use a store-and-forward buffer to ensure LRIT information is received (MSC.263(84)/Rev.1, paragraph 10.3.3); to process and handle any input within 30 seconds of receipt and to provide the appropriate output; and to be capable of receiving and processing at least 100 reports per second (MSC.1/Circ.1259/Rev.9, paragraph 4.1.1). In addition, MSC.1/Circ.1259/Rev.9, paragraph 3.3.1.2, states that designers and implementers are advised to adopt asynchronous patterns and reliable, non-blocking mechanisms to receive the HTTP response message, to ensure optimal performance.

2.3.6 The requirement to process and handle any input within 30 seconds of receipt and to provide the appropriate output is verified in every IDE audit, with no findings in the last audits.

2.3.7 The IDE's capability to receive and process at least 100 reports per second was reportedly tested twice in the Developmental Testing environment in 2013 and 2016, during testing of IDE software versions 2.2.0 and 3.0.0, respectively. A new test is being conducted by

the IDE at the time this report is being written.

2.4 LRIT Capacity Implications and Mitigation Mechanisms for Increased Traffic

2.4.1 The study aimed to assess the impact of the increased message traffic on the LRIT system. To perform the analysis, it was necessary to verify the system's current capacity against the existing requirements. The risk to consider is that surges in transmission could reach a point where the IDE would be unable to accept new connections while still processing existing ones. Thus, it is not sufficient that the average number of incoming messages is below 100 messages per second, because the average metric does not reflect surges or idle periods.

2.4.2 The LRIT architecture is a resilient distributed system based on an asynchronous communications protocol using SOAP over TLS 1.2 that provides the data interface between components. Every component offers and consumes services from other components to operate within the system. Each new communication creates a session that should end with the message being sent and received. To receive these incoming communications, each component provides a pool of connections to consumers. The pool of communications is standard in network systems; every server would have a configured pool to handle incoming traffic.

Pool of Connections and Processing Capacity in Distributed Communication Systems

2.4.3 The capacity to maintain a pool of incoming communication sessions and the capacity to process the messages delivered within those sessions are independent parameters in a server. For example, in theory, the IDE may maintain a pool of open transport connections from transmitting LRIT Data Centres, but be unable to process the messages carried over those connections due to resource saturation or internal queue congestion. Conversely, the IDE may have available processing capacity while being unable to accept new connections because the connection pool has reached its configured limit. The latter scenario can occur during a surge of incoming communications or as a result of a malicious denial-of-service (DoS) attack.

Session Lifecycle at the DC-IDE Interface

2.4.4 In SOAP over TLS 1.2 (typically over HTTP), a communication exchange is bound to the lifecycle of the underlying transport connection and the request–response pattern. A connection may be closed upon: completion of the HTTP/SOAP request–response exchange; expiry of the transmitting DC's application-level response timeout without receiving a reply; network- or transport-layer closure by either party; or release by the IDE of an idle connection exceeding its configured idle timeout.

2.4.5 When a timeout occurs at the transmitting DC, the connection may be closed or abandoned without notification to the IDE at the application layer. In such cases, any processing completed subsequently by the IDE cannot be delivered to the originating Data Centre over that connection.

2.4.6 Timeout parameters exist independently at each communication layer, from the TCP transport layer through the TLS and HTTP layers to the application layer. The connection pool is an application-level resource management mechanism that governs the number of simultaneous transport connections that the receiving component maintains. These parameters are independent of the component's processing capacity and must be considered together to ensure reliable handling of incoming communications.

LRIT Governance

2.4.7 Timeout configuration and connection pool limits operate independently of the processing capacity of a component; however, all three parameters must be considered together to ensure that incoming connections and messages are handled reliably under varying load conditions. The current LRIT technical documentation establishes requirements for processing capacity and processing time, but provides no guidance on timeout configuration or connection pool sizing.

2.4.8 A further mitigation mechanism widely implemented in distributed systems is message redelivery upon failure. In the LRIT system, the Technical Specifications (MSC.1/Circ.1259/Rev.9, paragraph 3.3.3.2) require the IDE to attempt redelivery of a message up to three times within a 12-minute window when connectivity issues occur with a destination LRIT Data Centre. No equivalent redelivery requirement is currently prescribed for LRIT Data Centres themselves.

2.4.9 Earlier versions of the technical specifications included redelivery requirements applicable to the International LRIT Data Centre (IDC). However, when the IDC provisions were subsequently removed, this requirement was not carried forward. As a result, it is reasonable to assume that some LRIT Data Centre implementations may still incorporate redelivery mechanisms based on those earlier specifications, despite the absence of a current formal requirement; however, this is a significant issue that should be confirmed, since it would provide a resilience layer to ensure message delivery from DCs to IDE.

3 Questions to be answered

3.1 The analysis aims to answer the following questions:

- .1 What would be the increase in the message flow through the IDE if all Contracting Governments decide to exercise their coastal State entitlements?
- .2 Under the existing requirements, would the IDE be able to accommodate the increased message flow?
- .3 What would be the estimated number of incoming and outgoing messages DCs should expect if all Contracting Governments decide to exercise their coastal State entitlements?

4 Methodology

4.1 Data Sample

4.1.1 The analysis was conducted on sample data extracted from the 2025 LRIT audit schedule. Therefore, the experimental sample is formed by aggregating the audit sample data, and it covers:

- All 71 Data Centres;
- All LRIT information transmitted by ships indicated by the associated Administrations to transmit to each DC during its 30-day audit sample period;
- 39,517 ships registered to the DCs in their audit sample period;
- A total of 3,042,131 individual LRIT position standard reports (6/6 hrs) across

all DCs;

- The data sample period spans from July 2024 to August 2025.

4.1.2 The data sample was assessed for its suitability for generalisation over a one-year period, taking into account each LRIT DC, its respective sample period, and the number of registered ships. Several statistical non-parametric tests were performed to verify that the data are representative and that no bias was inadvertently introduced.

4.2 *Limitations and Caveats*

4.2.1 The following limitations should be noted when interpreting the results of this analysis:

- .1 The DDP as used in this analysis reflects the geographic entitlement zones as submitted by Contracting Governments, and incorrections in the polygons may affect the results; and
- .2 The analysis does not account for on-demand (polled) transmissions requested by individual governments or overreporting from individual ships. Only periodic messages in the standard rate of 6/6 hours were included.

4.3 *Treatment of non-standard-rate transmissions*

4.3.1 The analysis considered only the 6/6 hour regular transmission of LRIT information, since this is the concern this study aims to address.

4.3.2 Transmissions in rates faster than 6/6 hours or responses to polling requests have been excluded from the sample programmatically at runtime while loading the samples.

4.4 *Analysis Procedure*

4.4.1 Each of the 71 DCs 30-day samples was initially processed individually, and the results were logged in a database.

4.4.2 For each LRIT position in the 30-day sample data of a DC, the following procedure was applied:

- .1 Ship identification and associated Administration, DC, geographic coordinates (latitude, longitude) and Timestamp 4 were extracted from the record.
- .2 The version of the DDP in force at the moment of Timestamp 4 was retrieved from the DDP archive.
- .3 A point-in-polygon test was performed to determine all CG in whose 1,000 NM coastal zone polygon the ship's position fell at that instant.
- .4 CG associated with the same DC of the sample data are excluded since this information is routed internally, not using the IDE²
- .5 One database record was created for each such Contracting Government, representing one LRIT message that the DC would be expected to route to that

² Cooperative DC, Regional DC, and National DC providing services to other Administrations may route LRIT information internally, and only the journal is periodically sent to the IDE.

government's Data Centre.

- .6 The new registry contains originator ship IMO number, data user (Administration), originator DC, destination data user (coastal CG), destination DC, position, timestamp4;
- .7 Timestamp 4 was recorded for each generated record to enable temporal analysis of transmission peaks.

4.4.3 After all records were processed, they were adjusted for the same 30-day period, forming an aggregated 30-day sample that represents a period in which all ships are transmitting. Such data provide the basis for the yearly generalisation and facilitate the calculations performed during the data analysis.

4.4.4 The annualization further accounts for the current number of ships actively transmitting to each DC, as recorded in the DC's ship registry at the time of the audit, to allow scaling of results to the full active fleet.

4.5 *Operational and theoretical load under full coastal entitlement exercise*

4.5.1 While assessing the sample data, it was noted that 39,517 ships in the sample data, accounting for 3,042,131 pieces of standard 6/6 hours LRIT information sent to the DCs indicated by their Administrations. This is the operational data used in the analysis.

4.5.2 However, it was observed that for such a number of ships, a total of 4,742,040 individual pieces of LRIT information should be generated in the same period, if all ships produced one message every 6 hours as per the requirement. The sample analysis report describes the causes for this Fleet-wide Transmission Completeness Rate of 64.2%, noting this is a known phenomenon in the system.

4.5.3 In order to assess the load under full coastal entitlement exercise, this matter was considered, and the results are discussed for the operational scenario, but also for the theoretical one, where the Fleet-wide Transmission Completeness Rate is 100%.

5 **Aggregated data overview**

5.1 The total estimated messages under the full CSSO entitlement exercise for a 30-day period, based on the sample analysis and the current operational status of the LRIT, is 17,604,806. If considering the full theoretical capacity, with all registered ships transmitting 4 messages per day, the expected load is 27,467,703 messages per month.

5.2 To produce an estimate of the year-long IDE message flow, the total number of database records generated for the sample period was scaled to a full calendar year using the following formula:

$$\text{Annual Estimate}(\text{operational}) = (\text{Total Records}/\text{Sample Days}) \times 365 = 214,191,806 \text{ MT1}$$

5.3 Considering a scenario where all ships transmit every 6 hours and do not miss transmissions (theoretical capacity):

$$\text{Annual Estimate}(\text{max_theoretical}) = (\text{Total Records}/\text{Sample Days}) \times 365 = 334,190,386 \text{ MT1}$$

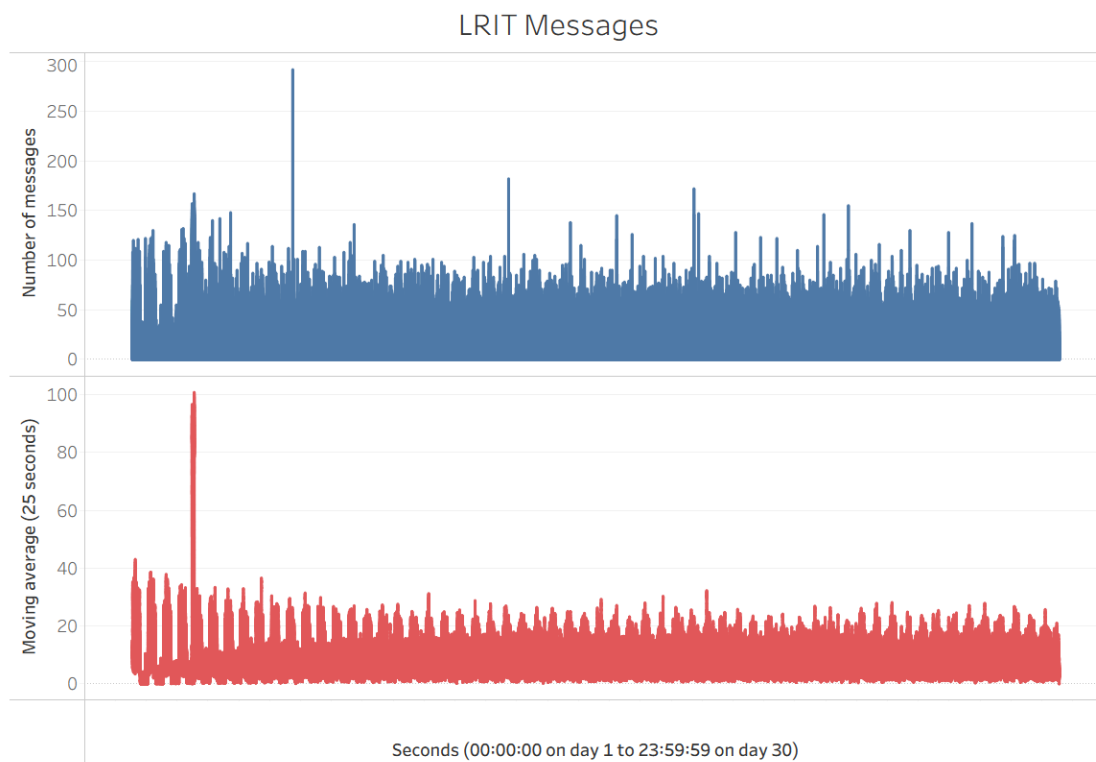
5.4 **Figure 1** presents two graphs, the first with the distribution of Timestamp 4 of periodic LRIT position report to be sent every second, for all DCs combined in a unique 30-period.

5.5 The second graph shows the moving average for 25 seconds in future, which gives an insight into the consequences of a 25-second timeout in the application layer. The moving average is calculated as follows:

$$MA(i) = (1/25) * \sum_{t=i}^{i+24} x(t)$$

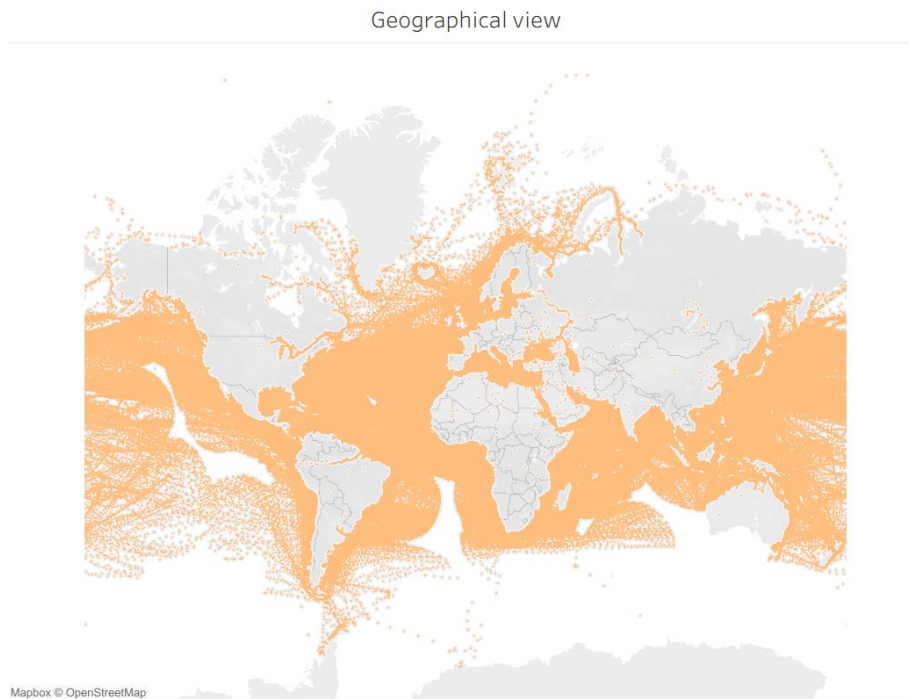
5.6 The peak observed in the moving average (nearly 200 messages) is not attributed to an outline in the ship's transmission. This phenomenon is historically associated with connection issues in the LRIT or CSP infrastructure (e.g., a CSP issue or a DC outage). When an issue occurs, the participant components may buffer the outgoing traffic and release it when possible, causing communication peaks.

Figure 1 - Estimated message flow through IDE per second.



5.7 A geo-referenced visualisation of estimated annual message traffic (**Figure 2**) was generated by aggregating LRIT data received by DCs from ships instructed by their Administrations to report to them. The resulting plot illustrates the volume and distribution of SOLAS ships' transmissions to the LRIT system over a full year, allowing coastal States to assess the information they would receive if they were to exercise their coastal entitlement.

Figure 2 - Geo-referenced LRIT positions from sample data.



6 Results

6.1 Average Load Estimation

6.1.1 The analysis of the operational data indicates that 214,191,806 messages are expected to be sent through the IDE in a one-year period if the CG exercises their coastal entitlement in full. This total represents an average of 6.8 messages per second. If the maximum theoretical capacity is reached, the average messages per second is estimated in 10.6. Considering the distributed load over time, in both scenarios the IDE will be operating well below its processing capacity of 100 messages per second.

6.1.2 The average increase in outgoing traffic of messages for the IDE is the same as the income, since the IDE receives the LRIT position reports and forwards them to the destination DCs.

6.2 Temporal Analysis — Transmission Peak Times

6.2.1 **Figure 3** presents the distribution of LRIT transmissions across one month, aggregated across all Data Centres in the sample. This analysis identifies peak and off-peak periods of IDE activity, which are relevant to dimensioning routing capacity and scheduling maintenance windows. A peak of messages is a second in which the IDE receives more than its processing capacity per second, 100 messages per second.

6.2.2 Considering distributed systems which communicate with asynchronous messages, such as the LRIT system, it is expected that in some situations communications will arrive faster than the system can process them. However, the system should be able to keep such communications, or sessions, alive while they have not been processed. As mentioned before, communication resilience may benefit from tuning the server connection pool and timeout mechanisms, and from implementing redelivery mechanisms.

6.2.3 Considering redelivering mechanism is not a requirement for DCs and should be further discussed, the analysis here intends to estimate the minimum connection pool and session timeout for the IDE to be able to couple with the peaks indicated by the sample data, considering its nominal capacity of processing 100 messages per second.

6.2.4 The analysis considers the composition of the DCs' data sample and the messages that would be generated by each piece of LRIT information. The 30-day aggregated sample data contain 3,042,131 periodic LRIT information sent by the ships to the DC indicated by their Administration. These messages cause 17,604,806 periodic messages to be sent to coastal states registered in other DCs, based on their 1000 NM coastal area. In short, due to the overlap of 1000 NM polygons, each piece of LRIT information from a ship would, on average, be sent to 5.7 other CG under their CSSO through the IDE if they exercise their coastal entitlement in full. Inter DC traffic is excluded because it does not affect the IDE.

6.2.5 Over the 30-day aggregated sample period, the message rate received by the IDE shows a clear central tendency of around 6.8 messages per second, representing the typical level of activity. Values up to 10 messages per second were registered in 75% of the time, indicating that the system will operate within this range under normal conditions. This suggests a stable and consistent workload for most of the time.

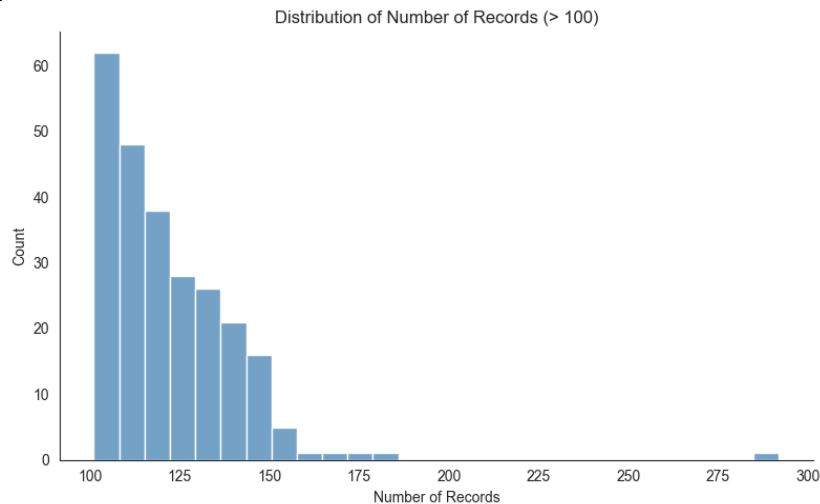
6.2.6 Statistically, outliers are identified using the interquartile range (IQR) method ($IQR = Q3 - Q1$; upper threshold = $Q3 + 1.5 \times IQR$). For the operational scenario, any value above 25 messages per second is classified as an outlier in the aggregated sample data. When extrapolated to the theoretical scenario, outliers correspond to values exceeding 40 messages per second. In the 30-day aggregated dataset, a total of 249 seconds recorded traffic above 100 messages per second, representing approximately 0.01% of the total operating time.

6.2.7 For the aggregate data sample, the maximum number of messages observed in a single second was 292, with the distribution showing a clear positive skew, as illustrated in the histogram. Only 10 seconds of recorded traffic exceeded 150 messages per second (Figure 3).

6.2.8 In a theoretical scenario, assuming continuous transmissions from all ships on a strict 6/6-hour cycle without interruption, traffic levels would increase by approximately 50%. However, analysis of the data does not identify any temporal pattern indicating concentration of messages at specific time intervals, nor any alignment with the 6/6-hour transmission cycle. No cyclic behaviour was observed.

6.2.9 Instead, the data exhibits short-term clustering. Further analysis indicates that the principal burst of messages within the sample originated from a single DC within a short interval of a few seconds. This behaviour is consistent with historical observations by the LRIT Coordinator during audits, which indicate that significant disturbances in message flow are typically associated with issues or outages affecting CSPs or LRIT system components, such as the IDE or DCs.

Figure 3 - Message peak histogram (Message per sec > 100).



6.2.10 To assess the nominal ability of the IDE to handle these peak times, the data was further analysed, calculating the minimum connection pool required for every second in the monthly aggregated data:

$$CPmin(t_0) = M(t_0)$$

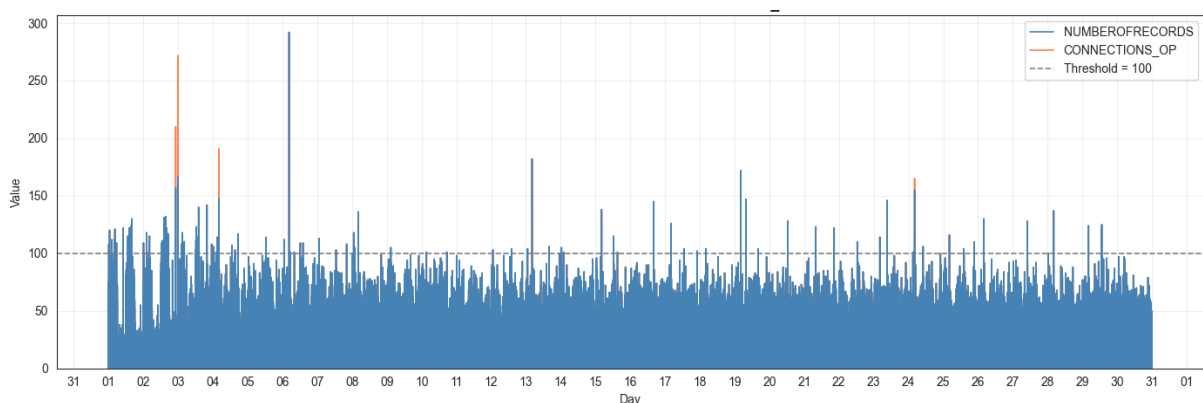
$$CPmin(t) = \max(0, CPmin(t-1) - 100) + M(t)$$

M(t): number of messages received at time t (per second)
CPmin(t): minimum required connection pool size

6.2.11 This metric indicates the size of the connection pool required to handle the incoming traffic, considering the IDE capacity of handling 100 messages per second.

6.2.12 **Figure 4** shows that during a burst period, considered an outlier and non-related to the ship's behaviour, the number of connections needed in the pool reaches 292 on day 6.

Figure 4 - Message peak and connection pool per second (aggregated sample data)

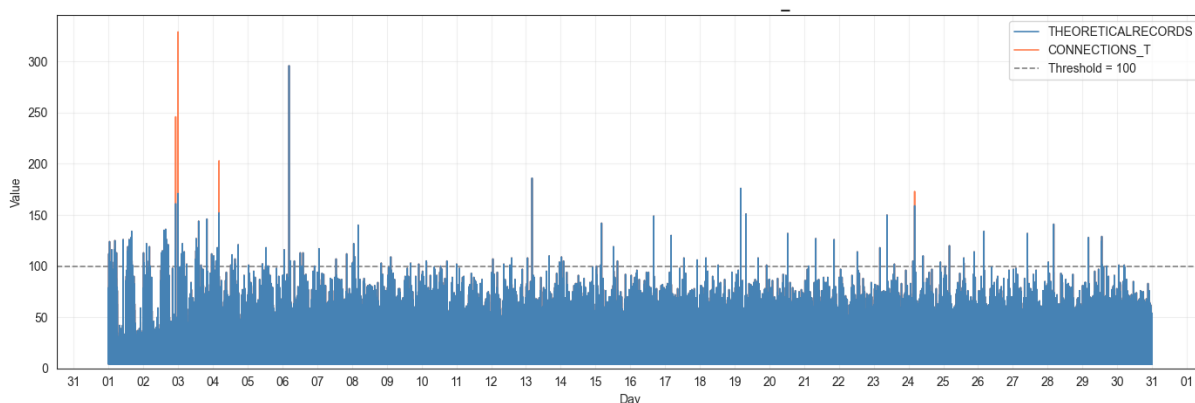


6.2.13 The data was further analysed to identify the maximum time interval between seconds that did not require an extra buffer (i.e. when the sum of the remaining connections and the incoming ones is lower than 100 messages). The resulting maximum time interval was 19 seconds, providing indication of the minimum timeout value that should be used when configuring the LRIT system components.

6.2.14 Special consideration was given to the full theoretical scenario where all ships are able to transmit their periodic information with no gaps. Given that it is unrealistic to assume these new transmissions would occur at the exact seconds we observe communications in the actual sample, a model was developed to distribute the new load evenly across the data.

6.2.15 **Figure 5** presents the simulation of such traffic. Considering the 100 messages per second capacity of the IDE, the pending connections reached 329, indicating the minimum pool size and the biggest gap between the two values of 100 messages in the pool was 20 seconds, suggesting a potential number for the timeout at the application level.

Figure 5 - Message peak and connection pool per second (theoretical scenario).



6.2.16 The foregoing analysis is based on the assumption that DCs do not implement re-delivery mechanisms, which could otherwise mitigate pressure on the IDE. From an architectural standpoint, the implementation of such mechanisms is recommended, even in the current scenario where only limited periodic reports are exchanged via the IDE, as it would enhance resilience to short-duration surges and temporary outages lasting a few minutes.

6.3 Increase in storage for IDE

6.3.1 Taking into account a binary XML representation of the periodic message, removing parts not stored by the IDE, such as latitude and longitude, a single message reduces to nearly 1.3KB uncompressed.

6.3.2 The Message flow is expected to increase from 15M to 214M messages. Thus, in theory, the raw XML data received will increase from approximately 33GB to 500 GB. When representing XML in a relational database, there is overhead due to indexes and internal database metadata. This overhead will vary depending on the implementation approach and the database management system. On the other hand, if the data is parsed and only the values of the files are stored, the storage size can be a fraction of the original data.

6.3.3 As presented, it is not feasible to calculate the actual increment in disk space for the IDE databases to couple with the new influx of information. However, it is a fair estimation to consider that the annual increment in size for the data will increase 15 times when compared to the current scenario, since LRIT position reports correspond to 90% of the current traffic

6.3.4 If considering the theoretical maximum capacity, where all ships transmit every 6 hours with no gap, the increase in messages will be from 14M to 334M, and the data may increase 23 times from the current annual increment.

6.4 Operational impact for Data Centres

Outgoing traffic:

6.4.1 The aggregated sample data show the receipt of 3,042,131 position reports by CG as flag States, and that these reports may imply 17,604,806 messages being sent to CG in other DCs if they exercise full coastal State entitlement. This suggests that, on average, every message received by the DC will cause 5.7 messages to be sent to other DCs, on average, and these messages must be stored in the DC.

6.4.2 As an example, if a DC has 100 ships, this implies 146,000 annual standard reports (100x4x365) that the DC should be receiving. In a scenario with full exercise of coastal entitlement, these DC, on average, will send 832,200 (146,000 x 5.7) messages to other CGs.

6.4.3 Considering that the full MT1 has all timestamps and the position of the ship. We may consider 1.6KB per message for estimation. One year will require around 1.5GB of raw data space for 100 ships, and this may be reflected in a database volume several times larger (e.g., 3x), due to metadata and DBMS overhead if stored as XML, or it can be compacted if saved as values in database tables.

6.4.4 For instance, a DC with 5,000 ships may need to provision an extra 75GB of raw data and consider how much this will impact the database.

Incoming traffic

6.4.5 It is not possible to provide a generalised rule to estimate incoming traffic, because this is highly dependent on the traffic density in the 1000NM from the coastline of the CG associated with the DCs.

6.4.6 However, the aggregated simulation developed in this experiment allows the LRIT Coordinator to inform each DC individually of their estimated traffic if they exercise their full 1000 NM coastal entitlement. Since the size of MT1 is the same as described in the paragraph above, these DCs can estimate incoming traffic and the need for storage resources.

6.4.7 The LRIT Coordinator notes that there is no requirement enforcing CG to receive information in the 1000 NM region; thus, it is sensible to assume that CG will only activate their polygons when their DCs are able to confirm they have the capacity to handle the incoming traffic.

7 Conclusions

7.1 The present study used a statistical experiment based on sample data that considered one month of traffic from each DC and aggregated these months to estimate annual traffic on the LRIT in a scenario in which all CGs exercise the full 1000 NM coastal entitlement. The suitability of the sample for a 12-month generalisation was assured through tests in a pre-processing phase.

7.2 The statistical analysis outcome, considered together with software engineering concepts, standards and specifications, allowed for the estimation of the impact on the IDE and the DCs in terms of message flow and storage needs.

7.3 The LRIT Coordinator notes that the information provided is generic and does not consider the specifics of each implementation, which may cause variation in the results. The

aim was not to provide precise data, which would be impossible by the very nature of a statistical experiment, but to look at the dimensions of potential impacts in all CG requires information in their 1000 NM, assisting the ongoing discussion on this matter.

7.4 Based on the outcome of the study, the existing technical requirements and the current ship reporting performance, it is concluded that:

- .1 with regard to message processing capacity, the existing requirements for the IDE are sufficient to accommodate the increased message flow if all Contracting Governments decide to exercise their coastal State entitlements.
- .2 the peak of transmissions, which are usually caused by a shoreside component such as DC or CSPs, are short and infrequent, and can be managed through appropriate configuration of system parameters.
- .3 the projected increase in data volume may require appropriate scaling of storage capacity at both IDE and Data Centre level if needed. The study suggests that DCs would require a storage capacity of 1.5GB per 100 ships per year for the outgoing messages and the IDE around 500GB per year (when the LRIT messages are saved in raw XML format). These data storage capacity may already be available considering that it was assumed when the LRIT system was designed that Contracting Governments would actively monitor their coastal waters.

7.5 The study outcome includes data on the estimated increase in incoming and outgoing message flow for individual DCs, which is not published in this document. However, the DCs and Contracting Governments may consult the LRIT Coordinator for detailed information regarding their DCs.

7.6 The study confirms the importance of a redelivery mechanism for the IDE and the DCs, especially when there is a peak in the number of messages sent or received.

ANNEX 2

STATISTICAL BASIS FOR THE GENERALISATION OF LRIT SAMPLE TRAFFIC DATA TO ANNUAL ESTIMATES

Contents

List of Acronyms	3
1 Executive Summary	4
2 Background and Regulatory Context	4
2.1 <i>LRIT Entitlement Under SOLAS Regulation V/19-1</i>	4
2.2 <i>The LRIT Data Distribution Plan (DDP)</i>	5
2.3 <i>Current operational status of the LRIT system</i>	5
2.4 <i>LRIT Capacity Implications and Mitigation Mechanisms for Increased Traffic</i>	6
Pool of Connections and Processing Capacity in Distributed Communication Systems..	6
Session Lifecycle at the DC-IDE Interface	6
LRIT Governance	7
3 Questions to be answered	7
4 Methodology	7
4.1 <i>Data Sample</i>	7
4.2 <i>Limitations and Caveats</i>	8
4.3 <i>Treatment of non-standard-rate transmissions</i>	8
4.4 <i>Analysis Procedure</i>	8
4.5 <i>Operational and theoretical load under full coastal entitlement exercise</i>	9
5 Aggregated data overview	9
6 Results	11
6.1 <i>Average Load Estimation</i>	11
6.2 <i>Temporal Analysis — Transmission Peak Times</i>	11
6.3 <i>Increase in storage for IDE</i>	14
6.4 <i>Operational impact for Data Centres</i>	15
7 Conclusions	15
List of Acronyms	3
1 Executive Summary	4
2 Description of the Sample	4
3 Statistical Assessment of Sample Quality	5
3.1 <i>Nature of the Dataset</i>	5
3.2 <i>Scope of the Dataset</i>	5
3.3 <i>Distribution of Fleet Size Across Data Centre</i>	5

- 3.4 *Check for Seasonal Assignment Bias: Kruskal-Wallis Test* 6
- 3.5 *Check for Monotonic Seasonal Trend: Spearman Rank Correlation*..... 6
- 3.6 *Temporal Representativeness: Jackknife Estimate* 7
- 3.7 *Annual Position Transmissions, Transmission Completeness and Confidence Interval*7
- 4 Summary of Statistical Quality Metrics****9**
- Appendix 1 — Step-by-Step Statistical Calculations** **12**
- 1 Distributional Analysis**12
- 2 Kruskal-Wallis Test Calculation**.....12
- 3 Spearman Rank Correlation Calculation**.....13
- 4 Jackknife Resampling Calculation**13
- 5 Annual Position Transmissions and Confidence Interval**.....14
- Appendix 2 — Sample dataset** **15**

List of Acronyms

CI	Confidence Interval
CV	Coefficient of Variation
DC	LRIT Data Centre
EU	European Union
HHI	Herfindahl-Hirschman Index
IDE	International Data Exchange
LRIT	Long-Range Identification and Tracking
NCP	Non-Centrality Parameter
NS	Not Statistically significant (at $\alpha = 0.05$)
R ²	Coefficient of Determination
SE	Standard Error
SOLAS	International Convention for the Safety of Life at Sea

1 Executive Summary

1.1 This annex documents the statistical validity of the one-month-per-DC sample used to extrapolate LRIT message traffic to annual estimates, and to assess IDE capacity requirements under full coastal State monitoring conditions. The scope of the analysis is limited to ships transmitting LRIT information to a registered Data Centre; no external fleet reference is used.

2 Description of the Sample

2.1 The dataset used in this experiment comprises LRIT traffic data drawn from 71 Data Centres (DCs) registered in the LRIT system, collectively representing 135 flag States. Each DC contributed one calendar month of operational LRIT traffic data, with the 71 DCs distributed across all 12 calendar months of the year, spanning the period July 2024 to July 2025.

2.2 Of the 71 DCs, 65 recorded at least one transmitting ship during their respective sample periods, accounting for a total of 39,517 SOLAS-compliant ships. The remaining 6 DCs recorded zero transmitting ships during their sample month. This outcome is operationally valid and consistent with historical LRIT data from previous years: it reflects periods during which no ships registered under those flag States transmitted LRIT information to the respective DC, rather than any deficiency in the sample design. These six DCs are therefore retained as valid observations in all statistical analyses, with a ship count of zero. The distribution of the sample data across the DCs is presented in table 1 of appendix 2.

2.3 The distribution of fleet size across DCs is highly heterogeneous, reflecting that some DCs have associated CGs with which Administrations have a large number of ships registered; in contrast, there are DCs receiving information from a limited number or no ship at all. The ten largest DCs account for approximately 81% of all ships in the dataset. The Herfindahl-Hirschman Index (HHI) was computed on the fleet-share distribution across all 71 DCs. The outcome was 0.0992, indicating a non-concentrated distribution of ships across DCs.

Table 2 - Distribution of Data Centres and Ships by Sample Month

Calendar Month	DCs Sampled	Ships Represented	Flag States
January	9	6,210	23
February	6	5,987	8
March	6	471	6
April	4	7,816	4
May	7	615	7
June	6	1,170	9
July	5	5,463	5
August	5	1,549	7
September	7	1,019	7
October	7	7,791	43
November	3	181	6
December	6	1,245	10
Total	71	39,517	135

2.4 All 12 calendar months are represented in the sample (Table 1), ensuring temporal completeness across the full annual cycle. The variability in ship counts across months reflects the unequal distribution of fleet sizes among the DCs assigned to each month, rather than any systematic seasonal pattern in vessel activity (a distinction that is confirmed statistically in Section 3).

3 Statistical Assessment of Sample Quality

3.1 Nature of the Dataset

3.1.1 The 71 DCs in this dataset constitute the entire population of LRIT Data Centres registered in the system; it is a complete census, not a random sample. The aggregated ship count of 39,517 is the directly observed total for the full DC population.

3.1.2 The purpose of the statistical assessment of the data sample is to verify that the one-month-per-DC audit sample for the 2025 audit schedule is suitable for extrapolating LRIT message traffic to annual estimates and for assessing IDE capacity under full coastal State monitoring conditions. Specifically, the sample assessment addresses two conditions that must hold for the extrapolation to be valid:

- (a) the single month assigned to each DC must not be systematically biased: that is, months with disproportionately large or small fleets must not cluster together in ways that would skew the annual projection; and
- (b) fleet composition must be sufficiently stable across the year that the observed ship counts can be treated as representative of the full 12-month period.

3.1.3 The tests that follow serve as diagnostic checks against these two conditions, and their results are interpreted in terms of whether any bias of practical significance, i.e. large enough to materially affect the annual message load estimate or the IDE capacity assessment, is present in the data.

3.2 Scope of the Dataset

3.2.1 The scope of this experiment is explicitly defined as all ships transmitting LRIT information to a registered DC in the 2025 audit schedule. A ship falls within scope if and only if it meets this criterion during the sample period; ships outside this definition are excluded by design. The fleet total of 39,517 ships, therefore, constitutes a complete count rather than an estimate drawn from a wider population, and no coverage ratio against an external reference is applicable or required.

3.3 Distribution of Fleet Size Across Data Centre

3.3.1 The fleet size distribution across the 71 DCs is strongly right-skewed (skewness = 3.456; excess kurtosis = 11.435), reflecting the concentration of ships registered under a small number of Administrations. A Shapiro-Wilk normality test confirms a significant departure from normality ($W = 0.4414$, $p < 0.001$). Applying a $\log(1+x)$ transformation, which accommodates the six DCs with zero-ship, produces a near-normal distribution ($W = 0.9744$, $p = 0.1557$; skewness = 0.003; excess kurtosis = -0.733). The $\log(1+x)$ -transformed fleet sizes are consistent with normality; accordingly, the original scale is highly skewed and non-parametric methods were used. The distributional statistics of fleet size across DC are summarised in Table 2.

Table 2 - Distributional Statistics for DC Fleet Sizes (n = 71)

Mean fleet size per DC	556.6 ships
Geometric mean	60.6 ships
Standard deviation	1,378.2 ships
Coefficient of variation	247.6%
Skewness (raw)	3.456
Excess kurtosis (raw)	11.435
Shapiro-Wilk W (raw)	0.4414 p < 0.001
Shapiro-Wilk W (log(1+x))	0.9744 p = 0.1557

3.4 Check for Seasonal Assignment Bias: Kruskal-Wallis Test

3.4.1 For the monthly-to-annual extrapolation to be valid, the fleet sizes of DCs must not vary systematically across the calendar months to which they were assigned. If, for example, months sampled early in the year consistently contained DCs with larger fleets than months sampled later, the resulting annual projection would be distorted. A Kruskal-Wallis test was applied as a diagnostic check to verify that no such grouping effect is present in the data.

3.4.2 The test yields $H = 9.6456$ (degrees of freedom = 11), $p = 0.5625$. The observed statistic is only 0.4902 of the critical value $\chi^2(11, \alpha = 0.05) = 19.6751$, well below the threshold at which a grouping effect would be considered detectable. The epsilon-squared effect size is $\epsilon^2 \approx 0$, indicating that calendar month assignment accounts for essentially none of the variance in DC fleet sizes. The minimum effect size that would be detectable at 80% sensitivity is $\epsilon^2 = 0.241$, but no effect approaching this magnitude was found. No statistically detectable month-group effect on fleet size was observed. The summary of the outcomes of the Kruskal-Wallis test is presented in Table 3.

Table 3 - Kruskal-Wallis Test Results

Test statistic H	9.6456
Degrees of freedom	11
Critical value $\chi^2(11, \alpha = 0.05)$	19.6751
H / χ^2_{crit} ratio	0.4902
p-value	0.5625
Effect size ϵ^2	≈ 0
Significance level α (Type I error)	0.05
Power $(1-\beta)$ at observed effect	0.05 (= α , as $\epsilon^2 \approx 0$)
Sensitivity $(1-\beta)$ at observed effect	0.05 (effect ≈ 0 ; see note)
Minimum detectable bias (ϵ^2) at 80% sensitivity	0.241

3.4.3 The sensitivity figure ($1-\beta = 0.05$) is a mathematical consequence of the negligible observed effect size: when the true effect is absent, no test can have power to detect it. The H statistic at less than half the critical value is a strong indication that no assignment bias of practical relevance to the annual extrapolation exists in this dataset.

3.5 Check for Monotonic Seasonal Trend: Spearman Rank Correlation

3.5.1 A complementary diagnostic check was applied to detect any monotonic trend across the calendar year; specifically, whether DC fleet sizes increase or decrease progressively from January to December. Such a trend, if present, would mean that the annual extrapolation is anchored to a systematically high or low portion of the year, introducing a directional bias in the estimated message load.

3.5.2 A Spearman rank correlation between the assigned calendar month number (1–12) and DC fleet size yields $r_s = -0.2039$, $p = 0.0881$ ($n = 71$). The observed correlation accounts for only $R^2 = 4.2\%$ of the variance in fleet sizes across DCs: a negligible proportion with no practical bearing on the annual estimate. The test has sufficient sensitivity to detect a medium-strength trend ($|r_s| = 0.327$, $R^2 = 10.7\%$) with 80% probability; no trend of this magnitude was found. Table 4 presents the Spearman Rank Correlation test results.

3.5.3 The p-value of 0.0881 is noted transparently. It falls short of the conventional 0.05 significance threshold and reflects a weak, practically insignificant negative tendency in the data. The Spearman test found only a weak, non-significant negative trend, with no practical effect on the annual estimate. Taken together, the Kruskal-Wallis and Spearman diagnostic checks provide convergent evidence that the monthly assignment of DCs introduces no seasonal distortion into the annual extrapolation.

Table 4 - Spearman Rank Correlation Test Results

Spearman r_s	-0.2039
p-value (two-tailed)	0.0881
R^2 (variance explained by month)	4.2%
Significance threshold α	0.05
Sensitivity ($1-\beta$) at observed r_s	40.0%
Practical significance threshold $ r_s $ at 80% sensitivity	0.327
Min. detectable $ r_s $ at 90% sensitivity	0.374
Min. detectable $ r_s $ at 95% sensitivity	0.411

3.6 Temporal Representativeness: Jackknife Estimate

3.6.1 A jackknife resampling procedure was applied to the 12 monthly ship totals. In each of the 12 iterations, one month was omitted, and the mean monthly total was recomputed from the remaining 11 observations. The jackknife standard error is 883 ships, yielding a 95% confidence interval for the mean monthly total of [1,350 – 5,236 ships], with an observed mean of 3,293 ships. The width of this interval reflects the substantial variation in fleet sizes across months, driven by the unequal distribution of ships among Data Centres assigned to each month, rather than any instability in the underlying census total of 39,517 ships.

3.7 Annual Position Transmissions, Transmission Completeness and Confidence Interval

3.7.1 This section quantifies the total number of LRIT position transmissions generated annually by the ships in scope. These transmissions are the input to the main experiment: for each position received, the experiment determines which coastal States have that position within their 1,000 nautical mile monitoring polygon and records each such occurrence as a potential message to be exchanged between the flag State Data Centre and the relevant coastal State Data Centre via the IDE. The results are stored in a dedicated database for subsequent calculation of average IDE message load and peak traffic in messages per second. The annual transmission count, therefore, represents the volume of position events entering the experiment, not the IDE message load itself, which is determined downstream by the coastal State polygon analysis.

3.7.2 Under SOLAS Regulation V/19-1, each ship transmits LRIT information at a default interval of six hours, equivalent to four transmissions per day. Given that the 39,517 ships in the dataset constitute the complete in-scope population, the calculated total number for annual transmissions entering the experiment is:

$$39,517 \text{ ships} \times 4 \text{ transmissions/day} \times 365 \text{ days} = 57,694,820 \text{ transmissions/year}$$

3.7.3 This figure represents the upper bound for capacity planning purposes, assuming full and continuous transmission by all ships throughout the year. It is used for IDE infrastructure dimensioning in Annex 2.

3.7.4 The aggregated 30-day sample across all 71 DCs contains 3,042,131 standard six-hourly position reports. The theoretical expectation for the same period is:

$$39,517 \text{ ships} \times 4 \text{ transmissions/day} \times 30 \text{ days} = 4,742,040 \text{ transmissions/month}$$

3.7.5 The observed count is therefore 35.8% below the theoretical maximum. This shortfall is not a deficiency of the sample design, the sample contains all and only the 6/6 periodic messages actually transmitted during the respective sample periods, with no filtering or omission. The gap reflects real-world LRIT transmission behaviour: ships undergoing repairs, modifications or conversions in port or dry dock may reduce their transmission frequency to one per 24 hours or suspend transmission entirely in accordance with paragraph 4.4.1 of resolution MSC.263(84)/Rev.1; equipment failures and connectivity interruptions account for a further portion; and a fraction represents non-compliant gaps. This interpretation is consistent with the finding reported in document NCSR 12/INF.3, paragraph 29, that approximately 36% of ships across all DCs and participating flags experienced at least one reporting gap of 24 hours or more within a 30-day period, a figure that is broadly consistent with the 35.8% shortfall observed here and provides independent corroboration that the sample is capturing the operational state of the system accurately.

3.7.6 The 35.8% shortfall does not affect the validity of the annualized estimates for the following reasons. First, the theoretical maximum is retained as the basis for IDE capacity planning in Annex 2: dimensioning infrastructure against the upper bound is the conservative and appropriate approach. Second, the actual transmission rate observed in the sample is itself a stable empirical characteristic of the fleet that, provided it is not distributed unevenly across the calendar year, is equally valid as a basis for annualization of the observed traffic.

3.7.7 The only remaining source of temporal uncertainty in applying this figure to a full calendar year is fleet volatility: i.e. ships being added to or removed from the LRIT system by Administrations during the year. The uncertainty is quantified empirically from the LRIT system's own historical annual fleet counts, as shown in Table 5, requiring no external reference. Since the experiment processes each position transmission individually and stores results in a segregated database, any change in fleet composition during the year is naturally captured in the position records as it occurs. The confidence interval below therefore characterises the uncertainty in the annual transmission count used for capacity planning purposes.

Table 5 - Historical LRIT Fleet Size (Ships Transmitting to the System)

Year	Ships Transmitting	Annual Change	% Change
2018	35,589	—	—
2019	36,243	+654	1.84%
2020	36,364	+121	0.33%
2021	36,444	+80	0.22%
2022	36,882	+438	1.20%
2023	38,860	+1,978	5.36%
2024	38,767	-93	-0.24%
2025	39,517	+750	1.93%

3.7.8 Over the 2018–2025 period, the mean absolute annual change is 588 ships, with a standard deviation of 671.61 ships. The fleet shows a statistically significant upward trend

(slope = +571.9 ships/year, $R^2 = 0.8967$, $p = 0.0004$), confirming that the 2025 figure of 39,517 ships highest recorded count and is appropriate as the basis for capacity planning. Applying the empirical standard deviation as the uncertainty measure and a 95% confidence interval ($z = 1.96$):

$\sigma = 671.61$ ships (calculated as the sample standard deviation of the absolute annual changes in LRIT transmitting fleet size over 2018–2025)

95% CI: [55,772,935 – 59,616,705] transmissions/year ($\pm 3.3\%$)

3.7.9 The corresponding 95% confidence interval for the empirical annual transmission estimate is:

95% CI: [35,779,000 – 38,308,000] transmissions/year ($\pm 3.3\%$)

3.7.10 Both intervals are grounded entirely in observed LRIT system data. The theoretical upper bound interval is used for infrastructure dimensioning; the empirical interval characterises the expected operational range.

4 Summary of Statistical Quality Metrics

4.1 The statistical analysis set out in this annex confirms that the one-month-per-DC sample is fit for the purpose of extrapolating LRIT message traffic to annual estimates and for assessing IDE and DC capacity requirements under full coastal State monitoring conditions.

4.2 The dataset constitutes a complete census of all 71 LRIT Data Centres registered in the system, covering 135 flag States and 39,517 SOLAS-compliant ships, representing the entirety of the in-scope transmitting population. All 12 calendar months of the year are represented. The distributional analysis confirms that DC fleet sizes follow a log-normal distribution, consistent with the known concentration of vessel registration, and that non-parametric diagnostic methods are appropriate throughout.

4.3 The Kruskal-Wallis test on fleet size ($H = 9.6456$, $p = 0.5625$; $\epsilon^2 \approx 0$) confirms the absence of any seasonal assignment bias of practical significance in DC fleet composition. The Spearman rank correlation ($r_s = -0.2039$, $p = 0.0881$; $R^2 = 4.2\%$) detects no monotonic seasonal trend in fleet size capable of introducing a directional distortion into the annual extrapolation. The jackknife resampling procedure confirms the stability of the monthly mean across all 12 calendar months.

4.4 The aggregated 30-day sample contains 3,042,131 standard six-hourly position reports against a theoretical maximum of 4,742,040, yielding a fleet-wide Transmission Completeness Rate (TCR) of 0.642. This 35.8% shortfall reflects the known operational state of the LRIT system, including authorised transmission reductions, equipment and connectivity gaps, and non-compliant interruptions. This interpretation is corroborated by the findings in document NCSR 12/INF.3 that approximately 36% of ships experienced at least one reporting gap of 24 hours or more within a 30-day period.

4.5 Two annualized transmission estimates are therefore established. The theoretical upper bound (57,694,820 transmissions/year; 95% CI: 55,772,935 – 59,616,705; $\pm 3.3\%$) assumes full and continuous transmission by all ships and is used as the basis for IDE infrastructure dimensioning in Annex 2, providing the conservative maximum load against which capacity must be assessed. The empirical estimate (37,043,672 transmissions/year; 95% CI: approximately 35,779,000 – 38,308,000; $\pm 3.3\%$) reflects observed fleet-wide transmission behaviour and represents the expected operational load under realistic conditions. Both intervals are grounded entirely in observed LRIT system data and require no external reference.

4.6 Taken together, these results confirm that the sample design meets the requirements for statistical validity. The theoretical upper bound provides a reliable and demonstrably conservative basis for infrastructure planning; the empirical estimate characterizes the realistic operational range. The use of the upper bound for capacity dimensioning in Annex 2 ensures that the conclusions of this study do not underestimate the load that the LRIT system would need to sustain.

4.7 The consolidated metrics supporting these conclusions are set out in Table 6 below.

Table 6 - Consolidated Statistical Quality Metrics

Metric	Value
Dataset Overview	
Total DCs in LRIT system	71
DCs with transmitting ships	65
DCs with zero ships (valid observations)	6
Total flag States represented	135
Total ships in scope (complete population)	39,517
Calendar months covered	12 / 12 (100%)
Fleet Size Distribution	
Distribution of DC fleet sizes	Log-normal ($W = 0.9744$, $p = 0.1557$)
Kruskal-Wallis H — fleet size (seasonal group test)	9.6456, $p = 0.5625$ (NS)
— Effect size ϵ^2	≈ 0
— Significance threshold α	0.05
— Sensitivity at observed effect ($\epsilon^2 \approx 0$)	5% ($= \alpha$; no detectable effect)
— Min. detectable bias ϵ^2 at 80% sensitivity	0.241
Spearman r_s — fleet size (monotonic trend test)	-0.2039, $p = 0.0881$ (NS)
— R^2 (variance explained by month)	4.2%
— Sensitivity ($1-\beta$) at observed r_s	40.0%
— Min. detectable $ r_s $ at 80% sensitivity	0.327
Jackknife standard error of monthly total	883 ships
Jackknife 95% CI for monthly mean	[1,350 – 5,236 ships]
Transmission Completeness	
Standard position reports in sample (actual)	3,042,131
Theoretical maximum for sample period	4,742,040
Fleet-wide Transmission Completeness Rate (TCR)	0.642 (64.2%)
Cross-validation: NCSR 12/INF.3 gap finding	$\sim 36\%$ of ships with ≥ 24 h gap (consistent with TCR)
Annual Transmission Estimates	
Theoretical Upper Bound (basis: full compliance assumed)	
Point estimate	57,694,820 transmissions/year
95% CI (empirical fleet variability)	[55,772,935 – 59,616,705]
Relative uncertainty	$\pm 3.3\%$ ($\sigma = 671.61$ ships)
Purpose	Infrastructure dimensioning — conservative maximum
Empirical Estimate (basis: observed fleet-wide TCR = 0.642)	
Point estimate	37,043,672 transmissions/year
95% CI	[$\sim 35,779,000$ – $\sim 38,308,000$]
Relative uncertainty	$\pm 3.3\%$

Metric	Value
Purpose	Expected operational load under realistic conditions

NS = not statistically significant at $\alpha = 0.05$.

Appendix 1 - Step-by-Step Statistical Calculations

1 Distributional Analysis

The following calculations characterise the distribution of ship counts across the 71 DCs.

N = 71 Data Centres
Sum of ships = 39,517
Mean (μ) = 39,517 / 71 = 556.6
Standard deviation (s) = 1378.2
CV = $s / \mu = 1378.2 / 556.6 = 247.6\%$
Skewness = 3.456 (strongly right-skewed; 0 = symmetric)
Excess kurtosis = 11.435 (heavy-tailed; 0 = normal)

Shapiro-Wilk test for normality (raw values):

W = 0.4414, $p < 0.001 \rightarrow$ Distribution is NOT normal

Log(1+x) transformation applied to accommodate zero values:

log(1 + x_i) applied to all 71 DC ship counts
Shapiro-Wilk on transformed values: W = 0.9744, $p = 0.1557$
Skewness (log) = 0.003 Kurtosis (log) = -0.733
Geometric mean = $\exp(\text{mean}(\log(1+x))) - 1 = 60.6$ ships

Conclusion: DC fleet sizes follow a log-normal distribution. Non-parametric tests are required for seasonal analysis.

2 Kruskal-Wallis Test Calculation

The Kruskal-Wallis H statistic checks whether DC fleet sizes differ systematically across the 12 calendar month groups, which would indicate a seasonal assignment bias that could distort the annual extrapolation.

$$H = [12 / (n(n+1))] \times \sum (R_j^2 / n_j) - 3(n+1)$$

where n = total observations, n_j = observations in group j, R_j = sum of ranks in group j.

n = 71 (total DCs across all months)
k = 12 (number of month groups)
H = 9.6456
Degrees of freedom = k - 1 = 11
Critical value: $\chi^2(\text{df}=11, \alpha=0.05) = 19.6751$
H / $\chi^2_{\text{crit}} = 9.6456 / 19.6751 = 0.4902$ (<< 1.0)
p-value = 0.5625 >> $\alpha = 0.05 \rightarrow$ No seasonal assignment bias detected

Effect size — epsilon-squared:

$\epsilon^2 = \max(0, (H - k + 1) / (n - k))$
 $\epsilon^2 = \max(0, (9.6456 - 12 + 1) / (71 - 12))$
 $\epsilon^2 \approx 0 \rightarrow$ Month assignment explains $\approx 0\%$ of variance in DC fleet sizes

Sensitivity analysis — minimum detectable bias at 80% sensitivity:

Using non-central chi-squared approximation: $H \sim \chi^2(k-1, NCP)$ where $NCP = \epsilon^2 \times (n-1)$

$$\begin{aligned} \text{Solve: } 1 - F_{\chi^2}(\chi^2_{\text{crit}}; df=11, NCP=\epsilon^2 \times (n-1)) &= 0.80 \\ \rightarrow \text{Minimum detectable bias } \epsilon^2 \text{ at 80\% sensitivity} &= 0.241 \end{aligned}$$

3 Spearman Rank Correlation Calculation

The Spearman rank correlation checks for a directional seasonal trend: whether DC fleet sizes increase or decrease monotonically across the calendar year in the order months were assigned. Such a trend would introduce a directional bias into the annual extrapolation.

$$r_s = 1 - [6 \times \sum d_i^2 / (n(n^2-1))]$$

where d_i = difference in ranks between month number and ship count for DC i.

$$\begin{aligned} n &= 71 \\ r_s &= -0.2039 \\ p &= 0.0881 \text{ (two-tailed)} > \alpha = 0.05 \rightarrow \text{No directional seasonal trend detected} \\ R^2 &= r_s^2 = (-0.2039)^2 = 0.0416 \text{ (4.2\% of variance explained)} \end{aligned}$$

Sensitivity analysis: minimum detectable trend at 80% sensitivity, using Fisher z-transformation (Cohen, 1988; Zar, 2010):

$$\begin{aligned} z &= \text{arctanh}(|r_s|) = \text{arctanh}(0.2039) = 0.2068 \\ SE(z) &= 1 / \sqrt{(n-3)} = 1 / \sqrt{(71-3)} = 0.1213 \\ NCP &= z / SE(z) = 1.7053 \\ z_{\alpha/2} \text{ (two-tailed, } \alpha=0.05) &= 1.9600 \\ \text{Sensitivity } (1-\beta) &= 1 - \Phi(z_{\alpha/2} - NCP) + \Phi(-z_{\alpha/2} - NCP) = 40.0\% \\ \beta \text{ (missed-detection rate)} &= 1 - \text{Sensitivity} = 60.0\% \end{aligned}$$

Minimum detectable trend $|r_s|$ at 80% sensitivity:

$$\begin{aligned} z_{\beta} \text{ (80\% power)} &= \Phi^{-1}(0.80) = 0.8416 \\ \text{min}_z &= (z_{\alpha/2} + z_{\beta}) \times SE(z) = (1.9600 + 0.8416) \times 0.1213 = 0.3397 \\ \text{min } |r_s| &= \tanh(\text{min}_z) = 0.327 \text{ (} R^2 = 10.7\%) \end{aligned}$$

4 Jackknife Resampling Calculation

The jackknife procedure assesses sensitivity of the mean monthly ship total to the omission of any single month. Table 1 presents the Jackknife mean calculated by excluding each month and the deviation from the grand mean.

$$\begin{aligned} \text{Grand mean of monthly totals} &= 3293 \\ \text{Jackknife SE} &= \sqrt{[(n-1)/n \times \sum(\theta_i - \theta)^2]} = 883 \text{ ships} \\ t_{\text{critical}} \text{ (} df=11, \alpha=0.05) &= 2.201 \\ 95\% \text{ CI} &= [3293 \pm 2.201 \times 883] \\ &= [1,350 - 5,236 \text{ ships}] \end{aligned}$$

Table 1 - Monthly Ship Totals and Jackknife Iterations

Month	Monthly Total (ships)	Jackknife Mean (-this month)	Deviation from Grand Mean
January	6,210	3028	-265
February	5,987	3048	-245
March	471	3550	+257
April	7,816	2882	-411
May	615	3537	+243
June	1,170	3486	+193
July	5,463	3096	-197
August	1,549	3452	+159
September	1,019	3500	+207
October	7,791	2884	-409
November	181	3576	+283
December	1,245	3479	+186

5 Annual Position Transmissions and Confidence Interval

Annual position transmissions from ships in scope: the input stream to the coastal State polygon analysis (deterministic calculation). IDE message load is derived downstream from the polygon matching results.

Transmission rate = 1 per 6 hours = 4 per day (SOLAS V/19-1)

Transmissions/ship/year = 4 × 365 = 1460

Annual transmissions = 39,517 × 1460 = 57,694,820

Confidence interval based on empirical LRIT fleet variability (2018–2025):

σ = std(absolute annual ship count changes, 2018-2025, ddof=1) = 671.61 ships

95% CI lower = (39,517 - 1.96 × 671.61) × 1460 = 55,772,935

95% CI upper = (39,517 + 1.96 × 671.61) × 1460 = 59,616,705

95% CI = [55,772,935 - 59,616,705] transmissions/year (+-3.3%)

Appendix 2 - Sample dataset

Table 1 lists all 71 Data Centres included in the analysis, together with their LRIT ID, assigned sample month, number of flag States, and ship count. Data Centres with zero ships are retained as valid observations.

Table 1 - Complete DC Dataset (n = 71)

Data Centre	Sample Month	Flag States	Ships
LRIT DC 01	September 2024	1	6
LRIT DC 02	May 2025	1	30
LRIT DC 03	November 2024	1	12
LRIT DC 04	October 2024	6	546
LRIT DC 05	October 2024	1	117
LRIT DC 06	July 2025	1	1,168
LRIT DC 07	January 2025	4	234
LRIT DC 08	September 2024	1	60
LRIT DC 09	December 2024	1	313
LRIT DC 10	September 2024	1	0
LRIT DC 11	May 2025	1	17
LRIT DC 12	August 2025	1	64
LRIT DC 13	January 2025	2	65
LRIT DC 14	January 2025	3	2,570
LRIT DC 15	December 2024	1	2
LRIT DC 16	September 2024	1	78
LRIT DC 17	September 2024	1	0
LRIT DC 18	May 2025	1	5
LRIT DC 19	July 2024	1	34
LRIT DC 20	March 2025	1	80
LRIT DC 21	December 2024	1	6
LRIT DC 22	July 2025	1	23
LRIT DC 23	February 2025	3	426
LRIT DC 24	June 2025	1	232
LRIT DC 25	March 2025	1	87
LRIT DC 26	August 2024	1	2
LRIT DC 27	January 2025	1	10
LRIT DC 28	September 2024	1	407
LRIT DC 29	February 2025	1	14
LRIT DC 30	December 2024	1	0
LRIT DC 31	May 2025	1	11
LRIT DC 32	February 2025	1	5,257
LRIT DC 33	March 2025	1	244
LRIT DC 34	July 2025	1	4,221
LRIT DC 35	July 2024	1	17
LRIT DC 36	March 2025	1	17
LRIT DC 37	April 2025	1	10
LRIT DC 38	February 2025	1	1
LRIT DC 39	October 2024	1	34
LRIT DC 40	May 2025	1	6
LRIT DC 41	April 2025	1	6,441
LRIT DC 42	October 2024	1	0
LRIT DC 43	August 2024	1	130

Data Centre	Sample Month	Flag States	Ships
LRIT DC 44	May 2025	1	50
LRIT DC 45	September 2024	1	468
LRIT DC 46	April 2025	1	1,192
LRIT DC 47	January 2025	1	312
LRIT DC 48	June 2025	1	4
LRIT DC 49	April 2025	1	173
LRIT DC 50	February 2025	1	214
LRIT DC 51	January 2025	1	2,322
LRIT DC 52	January 2025	7	13
LRIT DC 53	June 2025	1	12
LRIT DC 54	August 2024	1	177
LRIT DC 55	June 2025	1	1
LRIT DC 56	January 2025	1	300
LRIT DC 57	November 2024	4	159
LRIT DC 58	November 2024	1	10
LRIT DC 59	June 2025	4	438
LRIT DC 60	October 2024	1	60
LRIT DC 61	June 2025	1	483
LRIT DC 62	February 2025	1	75
LRIT DC 63	March 2025	1	0
LRIT DC 64	May 2025	1	496
LRIT DC 65	December 2024	1	0
LRIT DC 66	March 2025	1	43
LRIT DC 67	October 2024	1	59
LRIT DC 68	December 2024	5	924
LRIT DC 69	October 2024	32	6,975
LRIT DC 70	January 2025	3	384
LRIT DC 71	August 2024	3	1,176
