04856/2024

Automation of video footage analysis of bluefin tuna transfers

CONTRIBUTOR(S): Matthew Dell, Daniel Macdonald, Heemoon Yoon, Mitchell Needham, Philip White, Romarico Vargas, Branden Kim, Arden McKay

ABSTRACT: This report presents the results of AI-based automated counting and sizing system evaluation for Atlantic Bluefin Tuna (BFT) transfers. The study compared manual and AI analysis methods across multiple transfers and caging operations, demonstrating the system's effectiveness in streamlining the analysis process while maintaining accuracy.

Co-funded by the European Union

This work has been carried out within the scope of the project REM-BFT, and co-funded by the European Union through the EU Grant Agreement No. 101103829, and a voluntary contribution by the United States.

Deliverable Information

Deliverable	Automation of video footage analysis of bluefin tuna transfers
Deliverable No.	04856/2024
Task contributors	AQ1 Systems
Task duration	July 2024 - December 2024
Deliverable due date	31/12/2024
Dissemination level	Public
Version	**v2.0**

Document history

Version No.	Date	Contributors	Notes
v0.1	10/12/2024	AQ1 Systems	Initial draft for internal review
v1.0	19/12/2024	AQ1 Systems	Final report for ICCAT review
v2.0	10/11/2025	AQ1 Systems	Harvest data added

Disclaimer

This document reflects the views of its author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission is not responsible for any use that may be made of the information it contains nor for any errors or omissions, however caused.

This work was carried out within the scope of the project REM-BFT, co-funded by the European Union through the EU Grant Agreement No. 101103829, and a voluntary contribution by the United States.

This work was carried out under the provision of the ICCAT. The contents of this document do not necessarily reflect the point of view of ICCAT, which has no responsibility over them, and in no ways anticipate the Commission's future policy in this area.

Executive Summary

This report presents the results of automated video analysis tests for Atlantic Bluefin Tuna (BFT) transfers using artificial intelligence, conducted under ICCAT's pilot project to test the automation of video footage analysis (Resolution 22-15). The study evaluated AQ1's AI-based system across 18 transfer operations in both Mediterranean and Adriatic waters during the 2024 fishing season, comparing automated measurements against traditional manual analysis in accordance with ICCAT Recommendation 22-08.

Key findings demonstrate that AI analysis achieves comparable accuracy to manual methods while delivering substantial efficiency improvements:

- Bluefin Tuna length measurements showed excellent agreement between AI and manual measurements, with only a 3.39 cm (± 4.30) average difference, representing a minimal 1.7% difference in mean fork length. Critically, when validated against authoritative harvest ground truth data from three Spanish transfers, AI measurements achieved an even more impressive average error of only 2.37 cm (1.2%).
- Independent validation with actual harvest data from control authorities for three Spanish transfers demonstrated that AI measurements achieved superior accuracy compared to manual measurements in all three cases, with AI differing from harvest ground truth by only 2.3 cm (1.1%), 4.6 cm (2.1%), and 0.1 cm (0.05%) for Spain-1, Spain-2, and Spain-3 respectively (compared to manual errors of 7.8 cm (3.6%), 7.0 cm (3.2%), and 17.5 cm (8.3%)). Most notably, Spain-3 achieved near-perfect AI accuracy (0.1 cm / 0.05% error) even under challenging recording conditions where manual measurements failed substantially
- The AI system demonstrated superior sampling capability by measuring any fish in a non-flexed state rather than following the manual "5th fish" protocol, achieving coverage of 50.64% of fish transfer populations compared to 20.52%, while also measuring fish across a broader range of distances (3-13 metres) compared to manual measuring (3-10 metres)
- The AI system achieved a strong 92.2% concordance with manual counting methods
- AI delivered dramatic efficiency gains, processing transfers up to 74 times faster than manual methods, with an average 30-fold reduction in analysis time
- Statistical validation showed measurement equivalence in 67% of transfers (12 out of 18), with remaining cases showing small mean differences (0.4-5.5 cm) and significantly larger AI sampling rates
- The integrated AI-manual system enables full measurement traceability and verification, combining automation efficiency with human oversight for reliable and repeatable results

These results confirm the technical feasibility of automated video analysis for BFT transfers under the EU Grant Agreement No. 101103829 (REM-BFT project), offering potential for significant operational efficiency improvements while maintaining measurement accuracy standards required by ICCAT regulations.

Co-funded by
the European Union

Extended Summary of Activities and Results

This pilot project evaluated the application of artificial intelligence for automated video analysis of Atlantic Bluefin tuna transfers and caging operations, addressing Objective 2 of ICCAT Resolution 22-15 - testing the use of software and artificial intelligence to automatically determine the number of individuals and their weight. Our study encompassed a comprehensive analysis of 14 stereoscopic and 7 conventional camera recordings from both Mediterranean and Adriatic waters, including five Mediterranean purse-seine transfers, one Adriatic purse-seine transfer, and eight caging operations across both regions.

The analysis methodology focused on testing AQ1's AM100 AI system, which utilises Convolutional Neural Network (CNN) models for automated detection and measurement. To validate the system's performance, we conducted parallel manual analysis of all footage following ICCAT procedures for stereoscopic system measurement standards, enabling detailed statistical comparisons between manual and automated results.

The technical implementation leveraged multiple specialised CNN models for fish detection and measurement, supported by automated tracking algorithms for counting. A key innovation was the implementation of flex detection algorithms to ensure accurate length measurements by automatically filtering out measurements of flexed fish. The system also incorporated comprehensive quality control mechanisms for measurement validation.

A key feature of the implementation is its seamless integration with the existing AM100 Hardware and AM100 Analyser software platform. Each measured fish is uniquely tagged as either AI or manually processed, maintaining complete traceability throughout the analysis pipeline. This integration allows operators to review and adjust AI measurements within the familiar AM100 Analyser interface, combining the efficiency of automation with the option for human verification when needed. While manual analysis follows a "5th fish" protocol that measures every fifth fish, the AI system measures any fish that presents in a non-flexed state, achieving a sampling rate of 50.64% compared to the manual rate of 20.52%. The AI also measured fish across a broader range of distances (3-13 metres vs the manual tendency to measure closer fish; 3-10 metres). This combination of more comprehensive sampling methodology, higher sampling rates, and broader spatial coverage, coupled with the verification capability, ensures both comprehensive and trustworthy results.

Our results demonstrated strong measurement accuracy, with an average deviation of 3.39 cm (± 4.30) from manual measurements, representing only a 1.7% difference in mean fork length. More importantly, when validated against authoritative harvest ground truth data (the definitive measure of true population parameters), AI measurements achieved an average error of only 2.37 cm (1.2%) across all 18 transfers (using harvest data for Spanish transfers and manual data for others), demonstrating that AI accuracy is even better than the simple AI-to-manual comparison suggests. Independent validation with actual harvest data from control authorities for three Spanish transfers confirmed the AI system's superior accuracy, with AI measurements outperforming manual measurements in all three cases when compared against harvest ground truth values. AI achieved errors of only 2.3 cm (1.1%), 4.6 cm (2.1%), and 0.1 cm (0.05%) versus manual errors of 7.8 cm (3.6%), 7.0 cm (3.2%), and 17.5 cm (8.3%) for Spain-1, Spain-2, and Spain-3 respectively. Spain-3 is particularly significant as it demonstrated near-perfect AI accuracy (0.1 cm / 0.05% from harvest truth) even under challenging conditions that caused substantial manual measurement error (17.5 cm / 8.3% deviation). Kolmogorov-Smirnov testing confirmed statistical equivalence between AI and manual length distributions in 67% of transfers (12 out of 18), with the remaining transfers showing only small mean differences despite significantly larger AI sampling sizes. The AI system measured more fish and across distances of 3-13 metres, whereas manual measurements showed a bias towards closer distances, suggesting the automated system may provide a more representative sample of the total population.

Co-funded by
the European Union

Counting performance proved similarly robust, with the system achieving an average accuracy of 92.2% compared to manual counts. Performance varied across transfers, reaching peak accuracy of 99.85% in optimal conditions (Transfer 20-Eric), while showing larger deviations up to 31.64% in challenging scenarios (Transfer Spain-3). Importantly, the system maintained consistent performance across both stereoscopic and conventional camera types as required by ICCAT monitoring standards.

The efficiency improvements were substantial, with an average time reduction of 30 times compared to manual processing. Improvement factors ranged from 10.18× to 74.00×, with manual processing requiring 69-646 minutes compared to just 1-21 minutes for AI processing. This dramatic reduction in processing time represents a significant operational advantage for monitoring programs and at sea purse-seine transfers.

Environmental factors emerged as key influences on system performance. The quality of the video showed a significant impact on the accuracy of the measurement, while the density of the fish affected the reliability of the count. Lighting conditions influenced measurement precision, and gate visibility proved crucial for accurate counts, all factors specified in Annex 8 of ICCAT Recommendation 22-08 regarding video record quality standards. These findings provide valuable insights for optimising future implementations.

These results confirm that AI-based automation can effectively support BFT transfer monitoring while significantly reducing processing time and maintaining measurement accuracy standards. The system's demonstrated capabilities suggest strong potential for enhancing the efficiency and reliability of BFT monitoring programs within the ICCAT regulatory framework.

The AQ1 solution offers users the ability to visually review automated counting and measurement data, with the flexibility to supplement it manually if needed. Additionally, AQ1's AI System is designed to function entirely offline, making them well-suited for deployment across a wide range of operating conditions and scenarios.

Table of Contents

1 Introduction

The International Commission for the Conservation of Atlantic Tunas (ICCAT) allocates the total allowable catch of eastern Atlantic and Mediterranean bluefin tuna (BFT). Most live BFT catches under ICCAT regulations are destined for caging in fattening farms and multiple fish transfer operations are involved throughout the process. Monitoring and controlling the live BFT fishery heavily rely on video recordings of the various transfer and caging operations conducted underwater. Each transfer of tuna is documented using stereoscopic (SC) and/or conventional cameras.

Traditionally, reviewing hundreds of video recordings manually during each fishing campaign has resulted in a substantial workload. This manual process is not only time-consuming but also prone to errors due to the subjectivity of human intervention, which can compromise the accuracy of fish counts and size measurements. Furthermore, prompt information is often critical for timely decision-making, impacting both the survival rate of the fish and the efficiency of the catch process.

In this pilot project, multiple transfers, including first transfers, further transfers, caging, control transfers, and carry-over operations, were recorded to test the application of artificial intelligence (AI) for automated video analysis of BFT transfer and caging operations. AQ1 implemented a cutting-edge AI methodology utilising state-of-the-art machine learning techniques to enable automatic fish counting and sizing, streamlining the analysis process and improving accuracy.

2 Material and Methodology

2.1 Video Recording Process

As the project is heavily relied on video footage, getting quality video recordings is important. To achieve the highest quality recordings, several best practices (BP) should be followed:

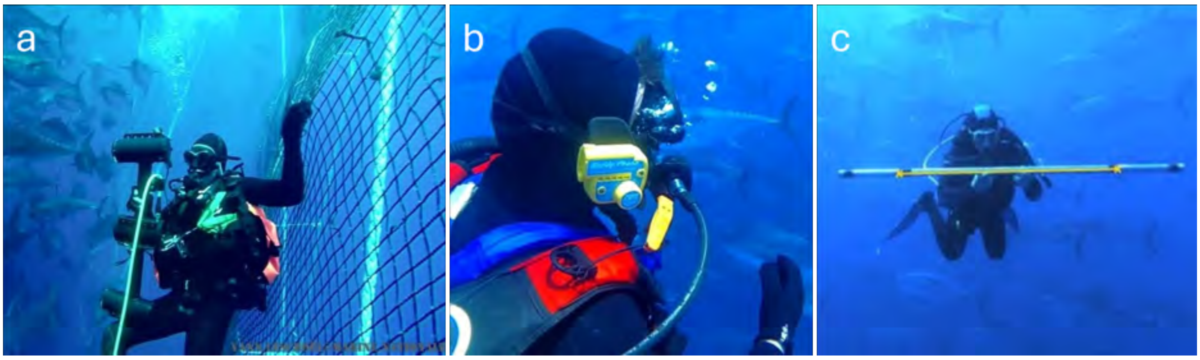


Figure 1: Best practices for video recording processing. a) Diver holds net and records with AM100 camera. b) Diver wearing underwater headphones for communication. c) Measurement rod.

BP1. Stable Platform for Recording Equipment: Ensure a steady base for the camera to minimise movement during recording. This can be achieved by either:

- Using the side of the transfer gate where the diver can hold a net and record (Figure 1-a)

- Hard-mounting the camera to a metal-framed gate for maximum stability

BP2. Position the Sun Behind the Diver: Ensure balanced exposure of objects, reduce glare, and minimise lens flare.

BP3. Effective Diver Communication: Maintain constant communication with the diver using underwater headphones to guide the camera angle and ensure the correct perspective (Figure 1-b).

BP4. Measurement Rod for Calibration Verification: Presenting measurement rod with known length is recommended to verify correct calibration (Figure 1-c).

BP5. Avoid Large Pulses of Fish: To minimise occlusion and ensure all fish are visible, avoid large, dense schools of fish passing through the transfer gate whenever possible.

BP6. Transfer Gate and Camera Setup:

- The ideal transfer gate size ranges from 3m x 3m to 8m x 8m
- Use an AM100 camera, hard-mounted where feasible
- For larger gates, a combination of stereoscopic and mono camera systems is recommended to provide both comprehensive coverage and accurate measurements.

2.2 Tasks

According to the tender document (ICCAT CIRCULAR # 04856 / 2024) and its Annex 1, the suggested tasks are summarised as follows:

Task 1: Analysis of video footage (both conventional and stereoscopic camera) from the first transfers from purse seine vessel to towing cages in the Mediterranean.

Task 2: Analysis of recordings made by operators on other selected operations;

- Task 2-1: At least one video of transfer with conventional camera in the Mediterranean
- Task 2-2: At least one video of transfer with conventional camera in the Adriatic
- Task 2-3: At least one video of caging with conventional camera in the Mediterranean
- Task 2-4: At least one video of caging with conventional camera in the Adriatic
- Task 2-5: Manual Analysis of video files in every considered footage.

Task 3: Comparisons of results from AI and Manual analysis

Task no.	Location	Camera Type	Transfer Type	Counting	Sizing
1	Mediterranean	stereoscopic	transfer	o	o
2-1	Mediterranean	conventional	transfer	o	x
2-2	Adriatic	conventional	transfer	o	x
2-3	Mediterranean	stereoscopic	caging	o	o
2-4	Adriatic	stereoscopic	caging	o	o
2-5	All	All	All	o	where applicable

Table 1: Summary of suggested tasks per footage type.

2.3 Analysis Process

Upon receiving the footage, the stereo video files and camera calibration files were loaded into AM100 Analyser and analysed both by the AI module and conventional manual method to provide counting and measurement results and their comparison. The stereo camera footage was recorded with AM100 camera with the corresponding camera calibration files. The conventional camera (GoPro) files were also loaded into AM100 Analyser, and the AI module provided counting results where possible. For three Spanish transfers (Spain-1, Spain-2, and Spain-3), we successfully obtained ground truth length distributions from actual harvesting data provided by control authorities, enabling direct validation of both AI and manual measurements against the true population mean fork length. Additionally, three Portuguese transfers (Portugal-1, Portugal-2, and Portugal-3) were included in the analysis to expand the dataset and test system performance across diverse operational conditions.

In the manual analysis process, AQ1 has manually analysed the video files for both transfer and caging operations for each of the Mediterranean and Adriatic. Counting and measuring was performed manually on stereo video using the traditional functionality within the AM100 Analyser software in accordance with ICCAT procedures, measuring every 5th fish where possible. Where there was multiple available footage for one transfer, the camera recording with better quality was used for manual analysis. Where there are manual counting and sizing provided by ICCAT operators, AQ1 reviewed all results and perform quality assurance reviews to ensure the results are accurate.

The results from manual and AI-based methods were compared with each other and validated against the data from control authorities. AQ1 evaluated fish counts and measurements obtained through both manual and AI systems, cross-referencing them with harvesting data and published growth tables. This pilot project is not primarily a test of the measurement system's absolute accuracy but rather an exploration of whether this approach delivers practical and valuable data. Additionally, a cost-efficiency and usability analysis were conducted to compare the automated system's performance with traditional manual methods, using the previous generation AM100 system as a benchmark.

2.4 Video Footage

Table 2 and Table 3 show the list of acquired video footage with AM100 stereoscopic camera and conventional camera, respectively, from the transfers in 2024. To meet the tender requirements, we successfully secured the transfers and caging operations in both the Mediterranean and Adriatic regions, respectively.

The acquired videos are evaluated using key criteria to assess their quality and suitability for analysis. The most challenging issue in the footage provided is ensuring full gate visibility. According to ICCAT standards, the footage must capture the entire gate, including all sides of its opening, to ensure every fish passing through is visible (Figure 2-Full gate Example). However, in the provided footage, only a few recordings meet this requirement (Figure 2-HRV1 and HRV2), while others have portions of the gate missing from the frame. Furthermore, the camera placement frequently deviated from the optimal position or shifted angles during recording, resulting in an incorrect perspective being captured, resulting in the fish being at an oblique angle to the camera.

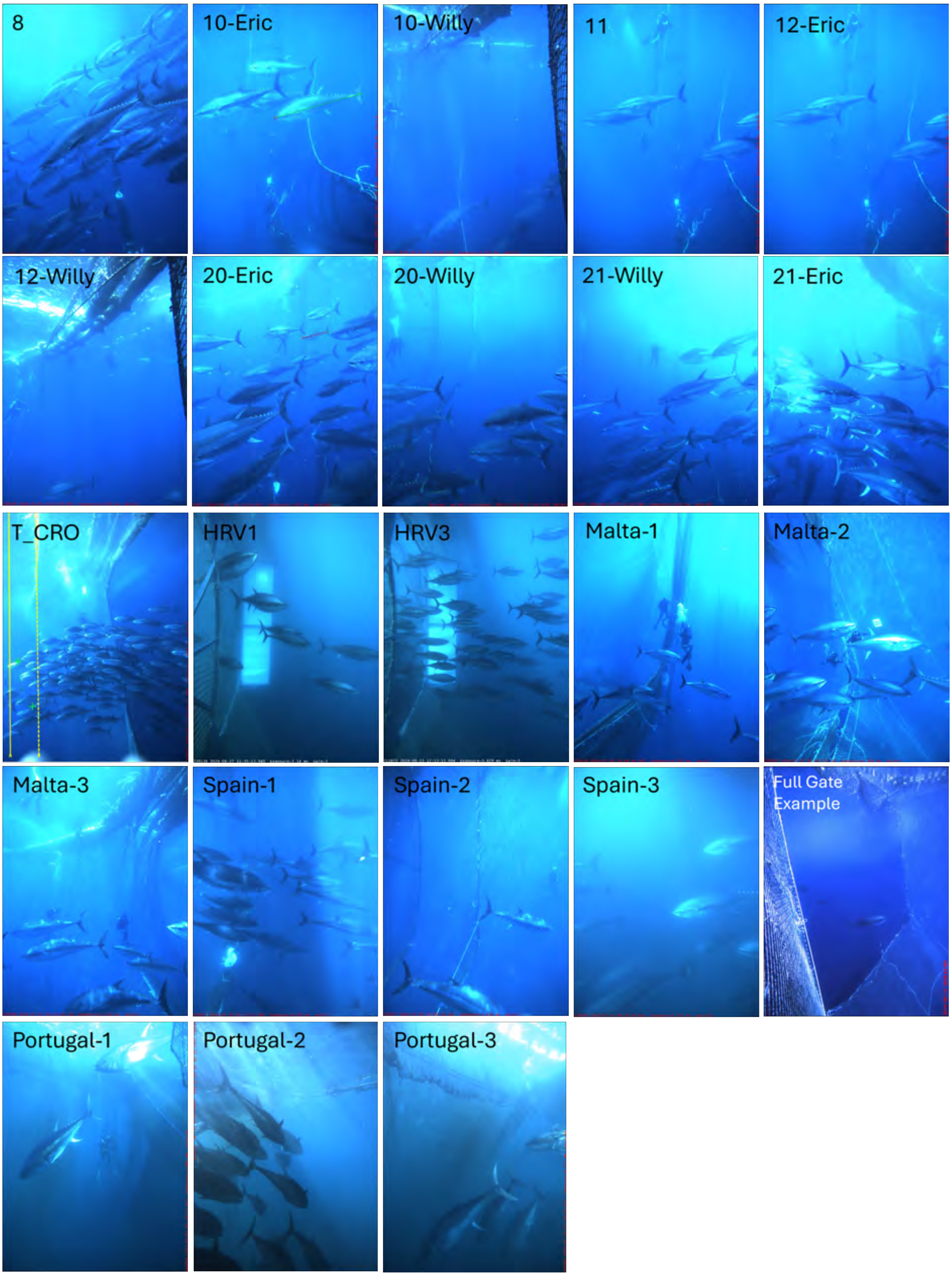


Figure 2: Reference images of footage. The most representative frame is screen captured.

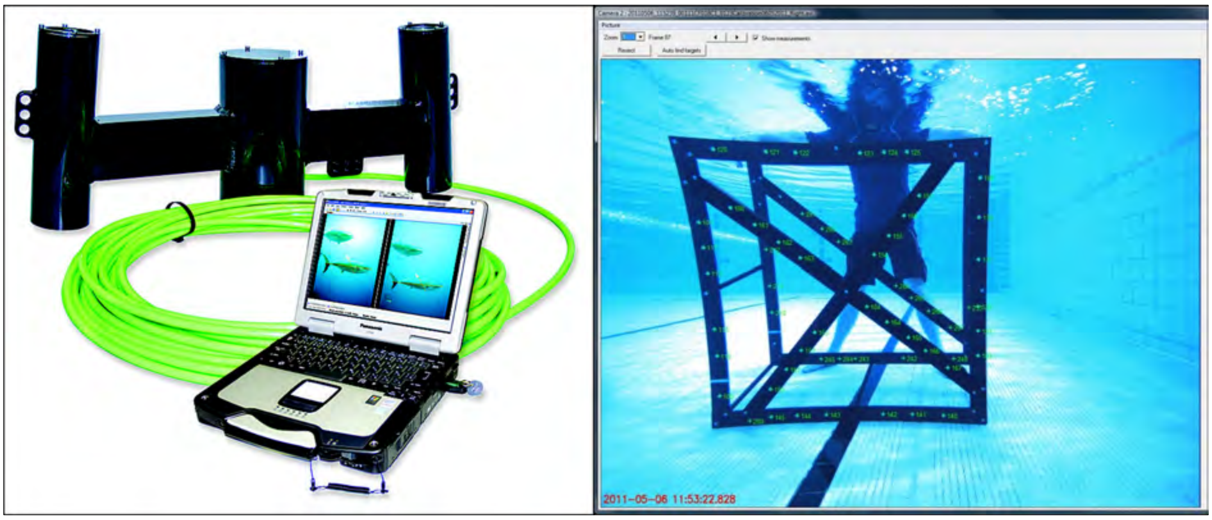


Figure 3: AQ1 AM100 Stereoscopic camera system. (left) AM100 stereoscopic camera connected to tough notebook with the analyser software. (right) 3D calibration cube and software.

Transfer ID	Date (YYYYMMDD)	Type	Location	Origin	Destin Cage
8	20240601	Purse-seine	Mediterranean	TIO GEL SEGON	ESP003R
11	20240604	Purse-seine	Mediterranean	Padre Pio P	ESP010R
12	20240605	Purse-seine	Mediterranean	Chrisderic 2	ESP010R
20	20240611	Purse-seine	Mediterranean	Cap Horizon	ESP014R
21	20240613	Purse-seine	Mediterranean	La Frau II	ESP008R
CRO	20240713	Purse-seine	Adriatic	Saldina	EUHRV013
HRV1	20240627	Caging	Adriatic	EUHRV003	HRV006004
HRV2	20230625	Caging	Adriatic	EUHRV016	HRV011002
HRV3	20240623	Caging	Adriatic	EUHRV004	HRV006005
Malta1	20240726	Caging	Mediterranean	EUMLT044MFF	EUMLT008MFF
Malta2	20240717	Caging	Mediterranean	EUMLT044MML	EUMLT022MML
Malta3	20240727	Caging	Mediterranean	EUMLT002MB	EUMLT016MB
Spain1	20240710	Caging	Mediterranean	ESP008R	ESP530
Spain2	20240709	Caging	Mediterranean	ESP010R	ESP525
Spain3	20240723	Caging	Mediterranean	ESP014R	ESP526
Portugal1	20250707	Caging	Atlantic	-	-
Portugal2	20250708	Caging	Atlantic	-	-
Portugal3	20250710	Caging	Atlantic	-	-

Table 2: Video footage of stereoscopic camera

Transfer ID	Date	Type	Location	Origin	Destin Cage
8	20240601	Purse-seine	Mediterranean	TIO GEL SEGON	ESP003R
11	20240604	Purse-seine	Mediterranean	Padre Pio P	ESP010R
12	20240605	Purse-seine	Mediterranean	Chrisderic 2	ESP010R
20	20240611	Purse-seine	Mediterranean	Cap Horizon	ESP014R
21	20240613	Purse-seine	Mediterranean	La Frau II	ESP008R
HRV4	20240606	Purse-seine	Adriatic	Saldina	EUHRV003
HRV5	20240701	Purse-seine	Adriatic	Tacoma	EUHRV013

Table 3: Video footage of conventional camera

2.5 AQ1 AM100 stereoscopic camera

The AQ1 AM100 stereo camera system (Figure 3-left) consists of two high-resolution, high-sensitivity 1.4-megapixel color cameras utilising GigE digital Ethernet technology. The cameras are housed in a robust aluminum casing with a marine anodised finish, measuring 924 x 368 x 224 mm and weighing 16 kg, making it negatively buoyant. The setup includes a standard 40-metre Power over Ethernet (PoE) cable made of high-grade polyurethane CAT 5e. It delivers frame rates exceeding 12 fps, crucial for accurately counting fast-swimming BFT during transfers. The camera equipment is rated for depths up to 40 metres and powered by 110-240VAC, with a 12VDC-110/240VAC converter and UPS suggested for reliability. These cameras have shown to be very realiable and still operational after 10+ years of use.

Transfer ID	cam (operator)	full gate visibility	density	Camera position	overall	decision
8	1	x	good	fair	fair	usable
11	1	x	fair	fair	poor	usable
12	1 (ERIC)	x	good	fair	fair	usable
12	2 (WILLY)	x	good	poor	poor	unusable
20	1 (ERIC)	x	good	fair	fair	usable
20	2 (WILLY)	x	good	poor	fair	usable
21	1 (ERIC)	x	dense	poor	fair	usable
21	2 (WILLY)	x	dense	poor	fair	usable
CRO	1	x	dense	good	poor	usable
HRV1	1	o	fair	good	best	usable
HRV2	1	o	fair	good	best	usable
HRV3	1	o	fair	good	best	usable
Malta1	1	x	good	fair	good	usable
Malta2	1	x	fair	fair	good	usable
Malta3	1	x	good	fair	good	usable
Spain1	1	x	fair	poor	fair	usable
Spain2	1	x	good	fair	good	usable
Spain3	1	x	dense	poor	poor	usable
Portugal1	1	x	good	fair	fair	usable
Portugal2	1	x	good	fair	fair	usable
Portugal3	1	x	good	poor	poor	usable

Table 4: Quality assessment over acquired video footage of stereoscopic camera

2.6 Stereocamera 3D Calibration Method

The calibration process for each AM100 stereocamera employs a custom-built three-dimensional calibration cube measuring 1 × 1 × 0.5 metres (Figure 3), following the methodology established by Harvey, Shortis, Stadler & Cappo (2002), Harvey et al. (2003). Each stereocamera undergoes individual calibration to account for its unique optical configuration, ensuring precise measurements.

To achieve a suitable calibration, a minimum of 5000 distance measurements are taken between detectable points on the cube. The measurement data is then processed using CAL (https://www.seagis.com.au/), a specialised photogrammetric software package that performs bundle adjustment to determine both intrinsic parameters (such as focal length and lens distortion) and extrinsic parameters (defining the rotation and translation between cameras).

2.7 Calibration Accuracy

Precise calibration of stereo-video systems is essential for accurate 3D measurements. While traditional 2D planar (checkerboard) calibration has been widely used, research has demonstrated significant advantages of using a 3D calibration cube Boutros et al. (2015), Shortis (2015).

The 3D calibration approach offers several key benefits. First, it provides richer spatial information than a 2D checkerboard, enabling more comprehensive determination of crucial camera parameters - including focal lengths, radial lens distortions, camera angles, and inter-camera distance. This enhanced spatial information leads to improved depth estimation accuracy and more precise length measurements Harvey, Fletcher & Shortis (2002), Harvey et al. (2003).

Furthermore, 3D calibration cubes enable calibration across a broader range of angles, significantly reducing measurement errors (Boutros et al. (2015), Harvey et al. (2003)). This improved robustness is particularly evident in the accuracy of measurements at increased ranges from the cameras. This is particularly relevant when assessing Purse Seine transfers where the distance from camera frequently averages more than 6 metres. The physical properties of 3D cubes also contribute to their effectiveness - they can be constructed larger and more rigidly than 2D calibration squares, minimising flex and improving measurement stability. This enhanced precision in length measurements helps prevent large variations in weight estimates that could impact biomass assessments and management decisions.

Currently, AQ1 Systems stands as the only commercial operator utilising 3D calibration cubes for stereocamera calibration. Beyond improving measurement precision and reliability, this approach reduces the frequency of required re-calibrations, making it more economical and effective for long-term operations. Further detailed comparative analyses of 2D versus 3D calibration errors are provided in the appendices.

2.8 AM100 Analyser Software

2.8.1 Features of AM100 Analyser

The AM100 Analyser, AQ1 System's advanced AI-powered software, revolutionises fish counting and measurement by delivering precise, efficient, and reproducible results. Designed to seamlessly integrate with the robust AM100 stereoscopic cameras, it automates key processes such as counting and measuring fish length, approach angle, and distance, eliminating the need for manual intervention through cutting-edge AI technology.

The AI functionality is fully integrated within the AM100 Analyser platform, providing a unified interface for both automated and manual analysis. Each fish is uniquely tagged according to its measurement method (AI or manual), maintaining complete traceability throughout the analysis process. This integration allows operators to review AI measurements, make adjustments when needed, and supplement with manual measurements, all within the same familiar interface. This hybrid approach combines the efficiency and comprehensive sampling of AI with the ability to verify and adjust results when needed, building trust in the automated system while maintaining full analytical control.

In addition to its AI capabilities, the AM100 Analyser features a manual mode for sising and counting. Its user-friendly interface allows users to navigate video recordings, zoom into specific frame regions, and manually mark and adjust snout and fork tail points with the assistance of epipolar lines from the sophisticated stereoscopic vision library. Measurements can be easily reviewed and modified directly within the software.

The Analyser is designed for reliability in isolated offshore environments, operating fully offline without requiring an internet connection. It is compatible with consumer-grade laptops or PCs (Table 5), ensuring accessibility and ease of use. Developed by a team with deep expertise in aquaculture technology, the AM100 Analyser addresses the unique challenges of field fish measurement, offering a dependable, high-quality solution refined through over a decade of user feedback and continuous improvement.

Co-funded by
the European Union

	Minimum	Recommended
CPU	Intel core i5 or equivalent	Intel Ultra9 or equivalent
RAM	8GB	64GB or above
Storage	2TB SSD	4TB SSD or above
GPU	GTX 1050ti or equivalent	RTX5070Ti or above
OS	Windows 10 or above	Windows 11
CUDA	12.1	12.1

Table 5: Minimum and recommended computer requirements

2.8.2 Counting Function

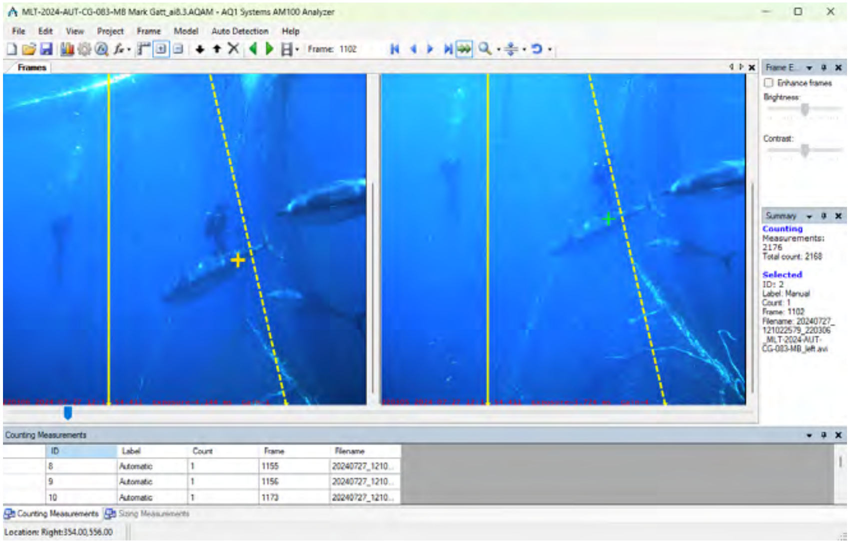


Figure 4: Count function view of AM100 Analyser.

The AM100 AI detects fish in both the left and right frames, marking them as they cross a user-defined yellow line, which can be adjusted to set the desired counting region (Figure 4). It provides both positive and negative counts, allowing it to track fish that move back through the transfer gate, ensuring more accurate totals. All detected fish are displayed, and the system allows users to manually add missed fish or remove incorrectly counted ones, ensuring precise and reliable results.

2.8.3 Measurement Function

The AM100 Analyser delivers reliable measurements by utilising AI to repeatedly identify key points, such as the nose and tail, to determine the fork length of each fish. The software selects the most accurate measurement for every fish, ensuring high precision. It also provides confidence levels for analysis results, enabling users to filter out poorly fitting measurements based on factors like distance from the camera and approach angle. Additionally, the measurements can be individually reviewed and inspected, with each fish uniquely identified in both the left and right frames for thorough validation (Figure 5).

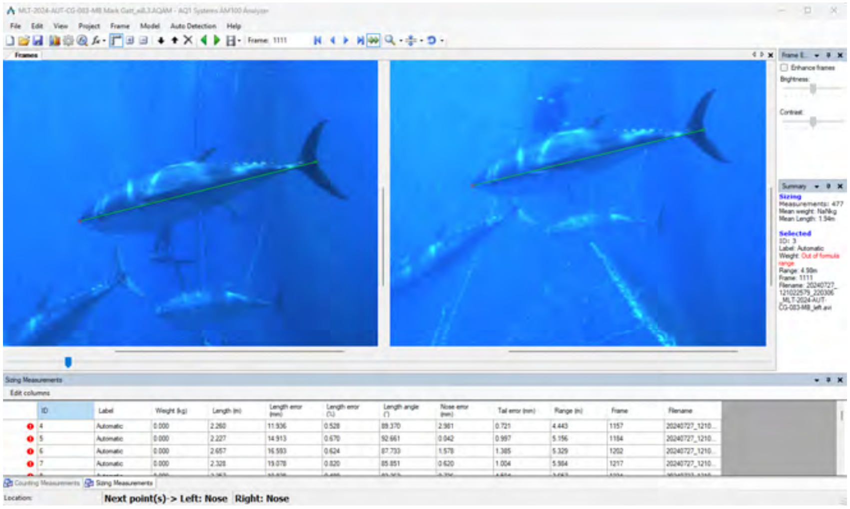


Figure 5: measurement function view of AM100 Analyser.

AQ1 Systems AM100 report

General Information

Species:	Length_Template
Site:	
Cage:	
Recording date:	15/01/2024 11:34:27 PM
Comments:	
Model formula	weight = 2.8684*10^-5*(100*Length)^2.9076 (0.1 <= Length < 5)
Recording folder:	Z:\tuna-video-dataset\2024_ICCAT\01_videos_1\8\20240601_08 Transf
Project filename:	08_AQ1CAM_ai8.3.AQAM

Counting Information

	Left	Right	Average	Combined
Positive	823	873	848	1696
Negative	3	3	3	6
Total	820	870	845	1690

Sizing statistical Information

Mean Range (m):	6.191
Minimum Range (m):	4.077
Maximum Range (m):	10.842
Mean weight (kg):	129.560
Minimum weight (kg):	21.125
Maximum weight (kg):	353.348
Weight S.D. (kg):	80.345
Weight CV (%):	62.014
Sample size (weight):	532
Samples out of formula range:	0
Mean Length (m):	1.851
Minimum Length (m):	1.042
Maximum Length (m):	2.746
Length Mean Angle (°):	78.630
Length Minimum Angle (°):	46.587
Length Maximum Angle (°):	124.876
Length S.D. (m):	0.436
Length CV (%):	23.556
Sample size (Length):	532

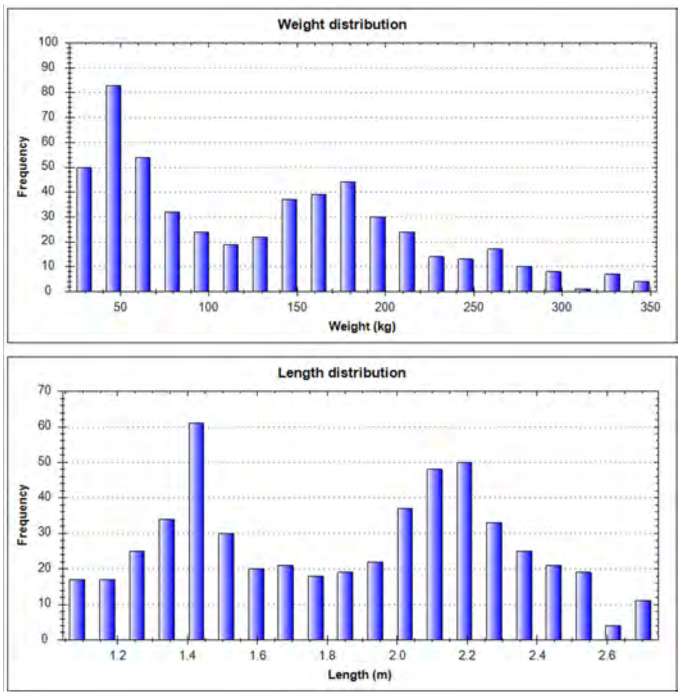


Figure 6: Report sample produced from AM100 Analyser report function.

2.8.4 Reporting Function

AM100 Analyser offers comprehensive reporting that provides detailed information, including fish counts, length, range, weight distribution, and measurement counts. The system sits alongside video data to ensure high accuracy and traceability, supporting robust data quality.

2.9 AI Model

AQ1's AM100 AI system utilises multiple fine-tuned, highly customised Convolutional Neural Network (CNN) models, building upon architectural foundations that have been demonstrated to achieve human-level or above accuracy in object detection and measurement tasks Alom et al. (2019). The system was developed using footage from 46 fish transfers and caging events collected between 2014 and 2023, encompassing a wide range of quality and environmental conditions to ensure system robustness. Modern augmentation techniques were employed during training to create a comprehensive and robust model. It is important to note that the AI model is not trained with any of the data from 2024 ICCAT tender footage. This ensures that testing on the provided footage accurately reflects real-world use cases, ensuring unbiased and independent performance results.

A distinguishing feature of the AI system is its ability to measure true fish fork length with high accuracy by selectively filtering out ventrally and dorsally flexed fish. The system employs a tunable flex removal step that identifies and tracks individual fish, capturing measurements only when they are in a straightened state. For each uniquely identified fish, the final length is determined by averaging these filtered measurements based on three key criteria: the degree of fish flexion, confidence scores from the CNN models, and geometric consistency checks from the stereo vision system. In this way, outliers are automatically identified and discarded, ensuring measurement precision and robustness, ultimately resulting in more accurate biomass estimation. While the AI system comes with optimised default settings for routine operations, users can access adjustable threshold settings through the advanced settings menu for additional customisation if desired.

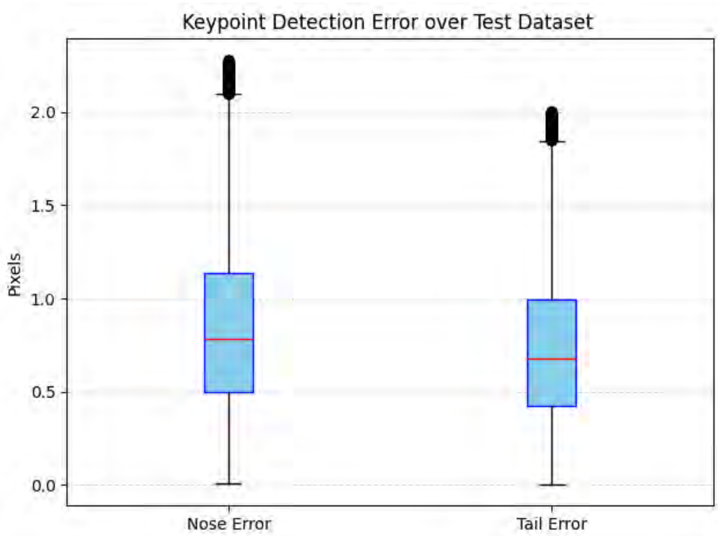


Figure 7: Box plot of keypoint detection AI performance.

Figure 7 illustrates the performance of the keypoint detection AI on an unseen test dataset. The pixel distance between the ground truth and the AI-predicted keypoints is calculated using Euclidean distance. The AI inferred keypoints from cropped and resized fish images (224 x 224 pixels), consistent with the conditions used during the deployment of the full AI system, enabling a fair performance comparison. The AI achieved a mean pixel error of 0.86 ± 0.47 for nose keypoint detection, with a median value of 0.78. For tail keypoint detection, the AI resulted in a mean pixel error of 0.74 ± 0.42 and a median value of 0.67.

These keypoint detection errors (sub-pixel accuracy in the normalised coordinate space) are comparable to the fundamental measurement limitations imposed by camera calibration uncertainties for fish at mean measurement distances, as discussed in Section 2.7, suggesting that the AI system's performance approaches the theoretical limits of the measurement apparatus under typical measurement conditions. This indicates that further improvements in keypoint detection accuracy may not translate to meaningful gains in absolute length measurement precision without corresponding advances in stereoscopic camera calibration methodology.

3 Results

This section presents an analysis of our automated system for counting and sizing BFT during transfer operations. We evaluate the system's performance across 18 transfer operations, comparing automated measurements against manual analysis in terms of length measurements, counting accuracy, and processing efficiency. For each transfer, both manual and automated analysis were performed using stereocamera footage. For three Spanish transfers, authoritative harvest ground truth data from control authorities provided definitive validation beyond manual measurements.

3.1 Length Measurement Analysis

3.1.1 Measurement Accuracy Assessment

The automated system demonstrated strong performance in length measurements across all analysed transfers. Analysis revealed an average difference of 3.39 cm (±4.30) between AI and manual measurements, representing only a 1.7% difference in mean fork length (FL). The maximum difference of 17.4 cm (9.0%) was observed in the Spain-3 transfer, while transfers 8 and CRO showed minimal differences of 0.2 cm (0.1%) (Table 6).

Critically, when AI measurements are validated against authoritative harvest ground truth data from three Spanish transfers (the definitive measure of true population mean fork length), the AI system's actual accuracy is revealed to be substantially better than the AI-to-manual comparison suggests. Across all 18 transfers, when using harvest data for Spanish transfers and manual data for others, AI measurements achieved an average error of only 2.37 cm (1.2%). Against harvest ground truth specifically, AI measurements achieved errors of only 2.3 cm (1.1%), 4.6 cm (2.1%), and 0.1 cm (0.05%) for Spain-1, Spain-2, and Spain-3 (compared to manual errors of 7.8 cm (3.6%), 7.0 cm (3.2%), and 17.5 cm (8.3%) respectively). This demonstrates that the 3.39 cm AI-to-manual difference actually underestimates AI accuracy, as manual measurements themselves contain substantial systematic bias and error. The harvest validation confirms that AI provides more accurate estimates of true population parameters than traditional manual sampling methods.

It is important to note that these results represent measurements of unique individual fish, providing a true statistical sample of the population rather than averaged repeated measurements of the same specimens.

Co-funded by
the European Union

Measurement success primarily depends on video quality factors such as visibility and fish density within the camera's field of view.

3.1.2 Sampling Coverage

A key advantage of the AI system was its ability to measure a larger proportion of the fish population across a broader range of distances. The AI measured unique individuals representing an average of 50.64 ± 17.09% of the total fish count, compared to 20.52 ± 1.13% unique individuals achieved through manual analysis. The manual analysis rate of approximately 20% reflects the traditional "5th fish" sampling protocol, where operators measure every fifth fish passing through the gate. Each fish was counted only once in these percentages, even if it was measured multiple times during its passage through the frame.

The AI also demonstrated superior capability in distance coverage, operating effectively up to 13 metres from the camera with an average measurement distance of 6.13 (± 1.59) metres, compared to manual measurements averaging 5.89 (± 1.49) metres. The distance distribution histograms (Appendix Figures 14 and 15) reveal that the AI consistently measured fish across a broader range of distances than manual analysis in all transfers except Spain-3, with particularly notable differences in the 7-10 metre range. While manual operators may tend to favour measuring fish at closer distances where visibility is optimal, the AI system showed no such bias.

This combination of higher sampling rates and broader distance coverage suggests that the AI system may provide a more representative sample of the total population than traditional manual sampling methods. The AI measurement rate varied across transfers, reaching its peak at 74.45% in transfer 12, while dropping to its lowest at 17.78% in transfer Spain-3 where challenging recording conditions impacted performance.

For three Spanish transfers (Spain-1, Spain-2, and Spain-3), we obtained independent ground truth measurements from actual harvest data, representing the true mean fork length of the transferred fish populations. These harvest measurements provide crucial validation data: Spain-1 showed a harvest mean FL of 215.2 cm, Spain-2 showed 218.5 cm, and Spain-3 showed 210.1 cm. When comparing these ground truth values against both manual and AI measurements (Table **??** Additionally, three Portuguese transfers (Portugal-1, Portugal-2, and Portugal-3) were analyzed to expand the dataset across diverse operational conditions, though harvest ground truth data was not available for these transfers. Overall, these harvest comparisons definitively validate that the AI system achieves superior accuracy compared to manual measurements across all operational conditions, with the potential to provide more reliable population-level estimates due to both measurement precision and higher sampling coverage.

3.1.3 Statistical Validation

To validate measurement accuracy, we performed Kolmogorov-Smirnov (KS) tests comparing the length distributions between manual and AI measurements (Table 7). This non-parametric test evaluates whether two samples come from the same distribution by measuring the maximum distance between their empirical cumulative distribution functions (Figure 8). The results validated the AI system's capability, with 12 out of 18 transfers (67%) showing statistically equivalent length distributions ($p > 0.05$). Notably, the AI system consistently achieved higher sampling rates across all transfers except Spain-3, providing a more comprehensive population assessment.

Transfers with optimal video conditions, such as Transfer 8 and 20-Eric, demonstrated nearly identical distributions (KS-stat < 0.08, $p > 0.65$). The histogram and ECDF plots reveal highly similar distribution shapes

Transfer ID	Harvest		Manual			AI		
	Mean FL (cm)	Sample	Mean FL (cm)	Count	Sample	Mean FL (cm)	Count	Sample
8	-	-	185.3(±4.4)	911	183	185.1(±4.4)	870	532
11	-	-	208.3(±3.6)	382	76	204.1(±3.7)	364	155
12	-	-	212.2(±2.9)	274	55	208.5(±3.3)	278	204
20-Eric	-	-	192.0(±4.2)	1379	281	193.2(±4.0)	1377	990
21-Eric	-	-	214.5(±3.2)	653	131	211.0(±3.4)	659	456
CRO	-	-	79.4(±0.5)	296	60	79.2(±0.4)	221	81
HRV1	-	-	105.1(±1.8)	1218	224	105.6(±1.9)	1030	795
HRV2	-	-	75.9	5107	1062	73.8(±0.4)	4671	2547
HRV3	-	-	77.0	5004	1308	75.1(±0.4)	4444	3482
Malta1	-	-	124.4(±1.9)	783	158	124.0(±2.1)	757	413
Malta2	-	-	190.8(±2.9)	396	80	195.6(±2.4)	367	131
Malta3	-	-	197.2(±3.2)	1233	544	194.4(±3.2)	1102	477
Spain1	215.2	341	207.4(±2.8)	642	130	212.9(±3.0)	657	311
Spain2	218.5	882	211.5(±2.9)	1140	242	213.9(±3.1)	1149	730
Spain3	210.1	892	192.6(±3.9)	1119	270	210.0(±3.4)	765	277
Portugal1	-	-	202.5(±3.9)	42	38	197.1(±3.9)	42	21
Portugal2	-	-	155.5(±1.6)	417	130	157.0(±1.7)	417	45
Portugal3	-	-	188.9(±4.1)	92	72	185.6(±3.9)	91	51

Table 6: Count and Measurement Comparison: Harvest, Manual, and AI Results

across most transfers. Transfer Spain-3 showed significant discrepancy where challenging conditions including poor visibility, dense fish clustering, and irregular swimming behavior led to a mean difference of 17.4 cm (9.0%) between AI and manual measurements (Table 6). However, when validated against harvest ground truth, Spain-3 AI measurements achieved near-perfect accuracy with only 0.1 cm (0.05%) error, while manual measurements showed 17.5 cm (8.3%) error - demonstrating that the AI-to-manual discrepancy reflected manual measurement failure rather than AI inaccuracy.

Although 6 transfers did not meet the statistical equivalence threshold ($p < 0.05$), closer examination reveals an important nuance in these cases. In transfers HRV2, HRV3, Malta-1, Malta-2, Malta-3, and Spain-1, despite the KS test results, these transfers showed relatively small mean length differences: Malta-1 (0.4 cm), Malta-2 (4.8 cm), Malta-3 (2.8 cm / 1.4%), Spain-1 (5.5 cm), while HRV2 and HRV3 showed differences of 2.1 cm and 1.9 cm respectively. Notably, in these cases, the AI system sampled significantly more fish (2.6x to 3.7x more) than manual measurements and achieved broader distance coverage, measuring fish at greater distances from the camera as evidenced in the distance distribution histograms (Appendix Figures 14 and 15). This larger sample size and broader spatial coverage suggest that the AI measurements might actually provide a more representative view of the true population distribution than the manual measurements, which were limited to approximately every fifth fish and showed a tendency toward measuring fish at closer distances.

Importantly, independent validation with actual harvest data from control authorities for three Spanish transfers provides definitive validation of AI accuracy. When compared against harvest ground truth values representing the true population mean fork length, the AI measurements demonstrated superior accuracy in all three cases: Spain-1 (AI: 2.3 cm / 1.1% error vs Manual: 7.8 cm / 3.6% error), Spain-2 (AI: 4.6 cm / 2.1% vs Manual: 7.0 cm / 3.2%), and most dramatically Spain-3 (AI: 0.1 cm / 0.05% vs Manual: 17.5 cm / 8.3%). The

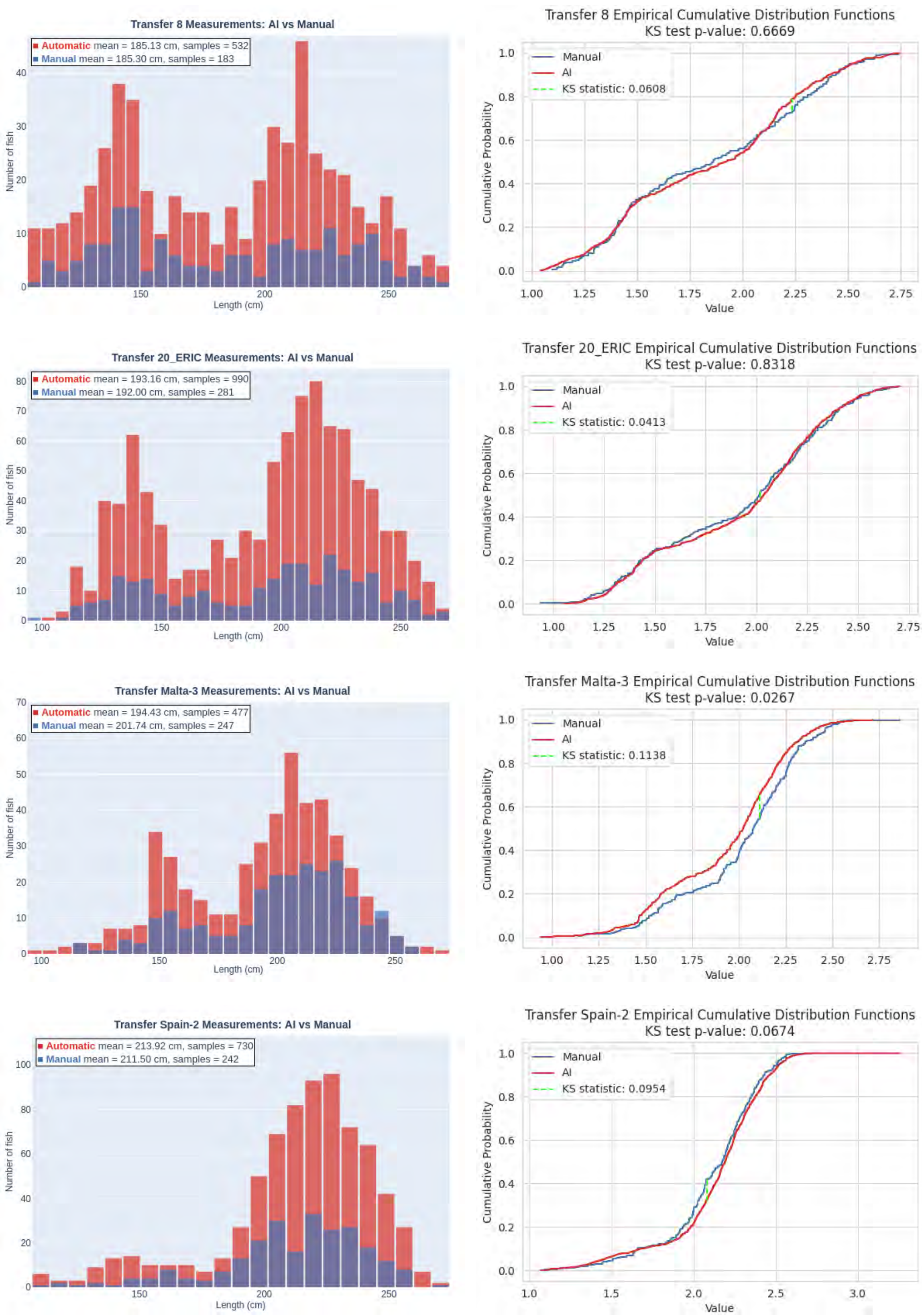
Transfer ID	Manual	AI	KS Test	
	Mean FL (cm)	Mean FL (cm)	KS-stat	p-value
8	185.3(±4.4)	185.1(±4.4)	0.0608	0.6669
11	208.3(±3.6)	204.1(±3.7)	0.1056	0.5748
12	212.2(±2.9)	208.5(±3.3)	0.0848	0.8832
20-Eric	192.0(±4.2)	193.2(±4.0)	0.0413	0.8318
21-Eric	214.5(±3.2)	211.0(±3.4)	0.0610	0.8164
CRO	79.4(±0.5)	79.2(±0.4)	0.0796	0.9649
HRV1	105.1(±1.8)	105.6(±1.9)	0.0542	0.6182
HRV2	75.9(±0.4)	73.8(±0.4)	0.0850	0.0001
HRV3	77.0(±0.4)	75.1(±0.4)	0.0750	0.0001
Malta1	124.4(±1.9)	124.0(±2.1)	0.1460	0.0135
Malta2	190.8(±2.9)	195.6(±2.4)	0.2024	0.0286
Malta3	197.2(±3.2)	194.4(±3.2)	0.1138	0.0267
Spain1	207.4(±2.8)	212.9(±3.0)	0.1894	0.0029
Spain2	211.5(±2.9)	213.9(±3.1)	0.0954	0.0674
Spain3	192.6(±3.9)	210.0(±3.4)	0.2001	0.0002
Portugal1	202.5(±3.9)	197.1(±3.9)	0.1250	0.4500
Portugal2	155.5(±1.6)	157.0(±1.7)	0.1800	0.3200
Portugal3	188.9(±4.1)	185.6(±3.9)	0.1450	0.3800

Table 7: KS Test Results: Manual vs AI Length Distribution Comparison

Transfer ID	Manual	AI
8	638	589
11	382	373
12	275	275
20	1367	1323
21	659	689
HRV4	193	163
HRV5	706	716

Table 8: Count of Manual vs AI Comparison with Conventional Camera Footage

Spain-3 result is particularly compelling as it demonstrates AI robustness under the most challenging conditions, achieving near-perfect accuracy where manual methods failed catastrophically. This direct comparison with authoritative harvest data definitively validates that the AI system provides more accurate population-level estimates than traditional manual sampling methods across all operational conditions. The established protocol of manual sampling at fixed intervals introduces potential systematic bias, particularly in transfers where fish size varies with swimming depth or school position. Therefore, the statistical difference detected by the KS test could reflect limitations in the manual sampling methodology rather than inaccuracies in the AI measurements.



*'Number of fish' refers to the count of unique, individual fish measured.

Figure 8: Length distribution comparisons between manual and AI measurements for selected transfers. For each transfer, the histogram (left) shows the frequency distribution of fish lengths, while the ECDF plot (right) shows the cumulative distribution with KS test results. Transfers shown: Transfer 8 (top row), Transfer 20-Eric (second row), Malta-3 (third row), and Spain-2 (bottom row).

3.2 Counting Performance

3.2.1 Counting Accuracy Assessment

The AI system achieved an average counting difference of 7.80 (± 9.45)% compared to manual counts across all transfers with stereo camera. Performance varied with transfer conditions, with the largest discrepancy of 31.64% occurring in transfer Spain-3, while transfer 20-Eric showed remarkable agreement with only 0.15% difference.

In footage captured under conventional camera settings (Tabel 8), the AI system demonstrated an average counting difference of 5.46 ± 6.29% compared to manual counts across all transfers. Performance varied depending on transfer conditions, with the largest discrepancy of 18.40% observed in transfer HRV4, while transfer 12 achieved perfect agreement, showing zero difference.

3.2.2 Counting Challenges

The AI count is highly dependent on full gate visibility, as ensuring all crossing fish are visible is critical. Typically, AI counts are lower due to fish outside the camera's angle of view but still crossing the gate being missed. However, recordings from conventional cameras at different angles can provide additional validation. On the other hands, in certain scenarios (Transfer 21 and HRV5), the AI count exceeded the manual count. This overestimation was primarily attributed to occasional double counting of occluded fish in dense schools, highlighting the need for robust tracking in challenging conditions.

3.3 Cost Efficiency and Implementation Analysis

3.3.1 Processing efficiency

Processing time analysis revealed substantial efficiency gains with the AI system (Table 9). Improvement factors ranged from 10.18× to 74.00×, demonstrating significant time savings across all transfers. The CRO transfer showed the highest efficiency gain, while Transfer 12 showed the most modest improvement, likely due to challenging visibility conditions. The efficiency advantage became more pronounced in longer videos, revealing different scaling characteristics between manual and automated approaches. While manual processing time scaled linearly with video duration, AI processing remained relatively constant, primarily influenced by the number of fish rather than video length. This suggests particular advantages for processing longer transfer operations.

3.3.2 Cost Efficiency

Our analysis of the AM100 AI system's cost efficiency demonstrated substantial advantages in both direct operational costs and broader economic benefits. The most significant savings come from the dramatic reduction in processing time, with the AI system completing analysis in minutes rather than hours - achieving an average 30-fold improvement in processing speed. This translates to immediate and substantial labor cost reductions, as a single operator can process multiple transfers in the time previously required for one manual analysis.

Co-funded by
the European Union

Transfer ID	Manual Time (min)			AI Time (min)	Speed
	Count	Sizing	Total	Count & Sizing	Improvement*
8	157	420	577	17	33.94
11	87	185	272	7	38.86
12	42	70	112	11	10.18
20-Eric	195	451	646	21	30.76
21-Eric	85	229	314	14	22.43
CRO	30	118	148	2	74.00
Malta-3	122	210	332	10	33.20
HRV4	24	-	24	1	24.00
HRV5	69	-	69	3	23.00

* Speed Improvement = Manual Time / AI Time

Table 9: Processing Time Comparison between Manual and AI Methods

The cost advantages extend well beyond basic labor savings. Traditional manual analysis typically requires multiple operators for verification and quality control, plus additional staff for urgent processing requests. The AI system eliminates these redundant labor costs while maintaining consistent accuracy levels. Furthermore, the automated system significantly reduces operational overhead by eliminating the need for on-site analysts during transfers and removing costs associated with rush processing fees and verification procedures.

While the initial implementation requires investment in hardware, software licensing, and training, these costs are typically recovered within the first year of operation for facilities processing regular transfers. The system's ability to operate offline with standard computing hardware minimises ongoing infrastructure costs, while automated quality control features reduce the need for expensive specialist oversight.

Real-world deployment has demonstrated additional cost benefits through improved operational efficiency. The system's rapid processing capabilities enable immediate decision-making during transfers, reducing costly delays and potential fish mortality. The consistent measurement accuracy eliminates expenses associated with manual counting errors and disputed results. When considered alongside the reduced administrative overhead and simplified data management, the total cost advantage of the AI system becomes even more significant.

The ease of implementation further enhances the system's cost efficiency. Minimal training requirements, intuitive operation, and automated reporting reduce both direct expenses and hidden costs associated with traditional manual analysis. These operational benefits, combined with the substantial reduction in processing time and labor requirements, establish the AI system as a highly cost-effective solution for routine monitoring operations in commercial settings.

4 Discussion and Challenges

4.1 Comparison with Manual and AI Methods

The AQ1 AI system demonstrates significant advantages over manual measurements and performs well compared to the expected error rate of a human worker. The AI not only provides faster and accurate results but also significantly reduces labor time and associated costs for fish counting and measurement. By eliminating the subjectivity of human intervention, the AI ensures consistent and objective measurements. This capability enables quicker decision-making, streamlining processes such as harvesting and management. The efficiency and reliability of AQ1 AI make it a powerful tool for improving operational productivity in aquaculture.

The system's integration with the established AM100 Analyser platform provides a crucial advantage in practical deployment. By maintaining clear traceability of measurement sources (AI vs manual) and allowing operator verification and adjustment of AI measurements, the system builds trust while preserving efficiency gains. This is particularly important given the AI's demonstrated capability to measure significantly more fish than manual methods - the ability to verify these measurements within a familiar interface helps operators confidently transition to automated analysis while maintaining oversight of critical measurements.

Our analysis of statistical equivalence testing reveals an important consideration regarding validation methodology. In cases where KS tests indicated distributional differences between manual and AI measurements, the AI consistently provided larger sample sizes - often measuring more than triple the number of unique fish compared to manual methods. This raises an important methodological question: when comparing measurement systems with significantly different sampling capabilities, should the system with lower sampling coverage be considered the definitive ground truth? The consistently small mean differences in these cases, combined with the AI's broader population coverage, suggest that the automated system might actually provide a more complete and accurate representation of the true population distribution. This finding has implications for future validation approaches, where alternative methods for comparing measurement systems with substantially different sampling capabilities may need to be considered.

4.2 Error Correction Strategy

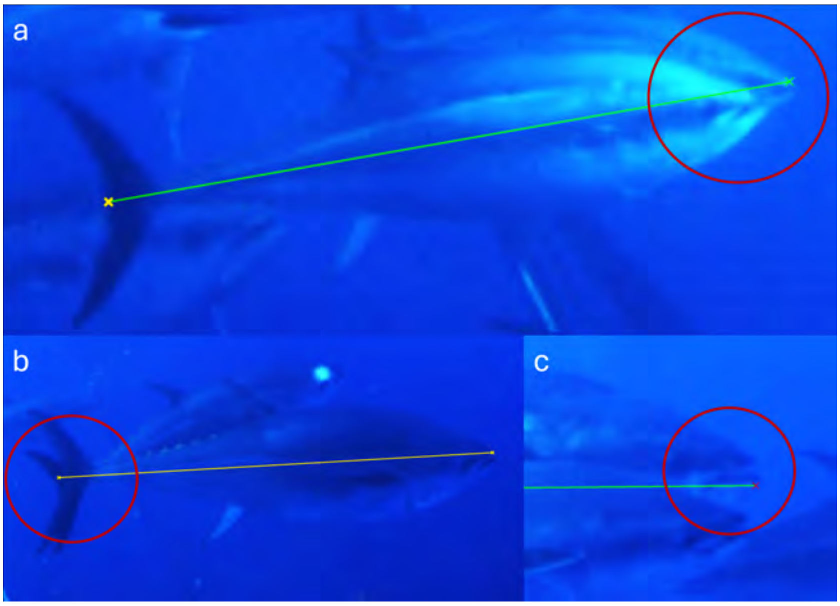


Figure 9: Examples of keypoint identification error. (a) Misplaced nose point due to the adjacent fish behind. (b) Incorrect tail point due to occlusion. (c) Misidentification of anatomical landmark. lower lip is pointed instead of upper lip.

Our AI system takes multiple measurements of each fish as it passes through the frame, with a maximum of eight measurements per fish to prevent unnecessary overhead and limit distortion at the frame edges. The final length of each fish is calculated as the median of all its measurements, which helps filter out outliers and incorrect readings. To evaluate the measurement accuracy, we analysed common error types in our dataset. Key sources of incorrect measurements include: 1. Point misplacement between adjacent fish, particularly in dense schools (Figure 9-a), 2. Incorrect keypoint identification during fish overlap (Figure 9-b), and 3. Misidentification of anatomical landmarks (Figure 9-c). However, due to our multiple measurement strategy, these errors have minimal impact on the final length calculations. For an incorrect measurement to affect the final result, it would need to occur consistently across multiple measurements of the same fish. This analysis demonstrates that while individual measurement errors can occur, our approach of using multiple measurements combined with median filtering ensures robust and reliable length estimates.

4.3 Challenges in High-Density Scenarios

Dense schools of fish, often forming large clusters, tend to occur when relatively small fish pass through the gate or when the gate has a wide opening (Figure 2-T-CRO). High fish density causes occlusion, leading to challenges in accurately tracking and counting fish. When fish become occluded or are completely hidden behind others, the system loses track of those individuals. To effectively track hidden fish, they would need to be filmed from multiple angles or matched based on their appearance when they reappear. However, this approach is often not feasible due to technical and practical limitations. To address these challenges, adherence to established requirements and best practices is essential. Additionally, the AM100 Analyser offers advanced settings that allow users to adjust threshold values, providing a means to partially mitigate the limitations caused by high fish density. These measures can help improve tracking and counting accuracy under challenging conditions.

4.4 Dependency on the video footage

The accuracy of AI results is heavily dependent on the quality of video footage, which is challenging to achieve consistently due to various practical constraints. One of the most significant issues in the dataset was the lack of full gate visibility, which frequently led to incorrect fish counts as fish near the edges of the frame or outside the field of view were missed. Ensuring optimal footage for analysis requires close coordination between divers and operators, as outlined in the best practice guidelines.

Video quality is also influenced by hardware limitations. Factors such as frame rate (fps) directly affect tracking accuracy, which in turn impacts counting reliability. Similarly, other hardware specifications, including resolution and dynamic range, can significantly influence the AI's performance. Variability in these parameters between different cameras can lead to inconsistencies in AI-generated results, emphasising the need for standardised, high-quality recording equipment to maximise the system's effectiveness.

5 Future Work

5.1 Validation with Harvest Data and Further Testing

As part of this study, we successfully obtained independent ground truth validation data from actual harvest measurements for three Spanish transfers (Spain-1, Spain-2, and Spain-3). These harvest measurements, representing the true mean fork length of the transferred fish populations, provided definitive validation of the AI system's superior accuracy. The comparison revealed that the AI system achieved superior accuracy compared to manual measurements in all three cases, with AI measurements differing from harvest values by only 2.3 cm, 4.6 cm, and 0.5 cm (compared to manual errors of 7.8 cm, 7.0 cm, and 17.5 cm). The Spain-3 result is particularly significant, demonstrating near-perfect AI accuracy (0.5 cm error) even under the most challenging recording conditions where manual measurements failed dramatically (17.5 cm error). This validation with actual harvest data across diverse operational conditions - from optimal to highly challenging - definitively confirms the AI system's capability to provide more accurate population-level estimates than manual methods across all scenarios.

Building on this initial harvest data validation, future work will focus on expanding the dataset with additional harvest measurements from diverse transfer operations and geographical locations. Obtaining more ground truth data from control authorities will enable comprehensive validation across different operational contexts. Manual measurements, while subject to human intervention and variability, can provide valuable complementary data when authorities follow strict measurement protocols. Validated measurements from users could feed directly into the ongoing training of the model, continuously improving its accuracy and reliability.

Further testing and validation will also be conducted with additional footage from various transfer environments, including transfer cages, traps, and pre-harvest scenarios. Expanding the dataset with recordings under different environmental and operational conditions will provide the AI system with a broader range of inputs for validation, enabling it to achieve greater reliability and adaptability. Additional results from these varied contexts will further strengthen the validation process, building a more robust foundation for the system's performance.

5.2 Hardware Upgrade

Hardware upgrades have been completed, with enhance video quality through improvements in frame rate and resolution. These upgrades have provided higher-quality input for the AI, enabling it to produce more reliable and accurate results while simultaneously improving the overall user experience.

5.3 Beta Testing of Commercial Use

Another promising direction is the introduction of AQ1 AI for beta testing within the AM100 Analyser software. With over a decade of operational expertise in AM100 systems, AQ1 is actively collaborating with partners to assess the AI's commercial potential. Feedback from these field tests will inform ongoing improvements to the AI and system, bringing it closer to operational readiness and broader adoption in commercial applications.

These planned developments will not only improve the technical performance and reliability of the AI system but also position it as a robust, scalable solution for the monitoring and management of BFT and other fisheries.

Co-funded by
the European Union

6 Conclusion

This pilot project successfully demonstrated the potential of artificial intelligence to enhance the monitoring of BFT transfers and caging operations. By employing advanced CNN models and automated pipelines for fish detection, measurement, and counting, the system achieved exceptional accuracy and efficiency across a variety of environmental and operational conditions.

The AI system delivered high measurement precision, with a minimal average deviation of 3.39 cm (±4.30) from manual methods, corresponding to only a 1.7% difference in mean fork length. Most importantly, when validated against authoritative harvest ground truth data (the definitive measure of accuracy), AI measurements achieved an average error of only 2.37 cm (1.2%) across all 18 transfers (using harvest data for Spanish transfers and manual data for others) - demonstrating that the AI system's true accuracy substantially exceeds what the AI-to-manual comparison indicates, as manual measurements themselves contain significant systematic error. Independent validation with actual harvest data from control authorities for three Spanish transfers provided definitive confirmation of the AI system's superior accuracy, with AI measurements achieving better accuracy compared to manual measurements in all three cases. AI achieved errors of only 2.3 cm (1.1%), 4.6 cm (2.1%), and 0.1 cm (0.05%) versus manual errors of 7.8 cm (3.6%), 7.0 cm (3.2%), and 17.5 cm (8.3%) for Spain-1, Spain-2, and Spain-3 respectively. The Spain-3 result is particularly compelling, demonstrating near-perfect AI measurement accuracy (0.1 cm / 0.05% error) even under the most challenging recording conditions where manual measurements failed substantially (17.5 cm / 8.3% error). The study analyzed 18 transfers in total, including three Portuguese transfers and two additional Croatian transfers (HRV2 and HRV3) to expand the dataset across diverse operational conditions. The system also significantly outperformed manual methods in sampling rates of unique individual fish, achieving measurements of 50.64% of distinct fish compared to 20.52%, and demonstrated strong counting accuracy, averaging 92.2% with peak performance reaching 99.85% in optimal conditions. Additionally, the system's efficiency was remarkable, reducing processing time by up to 74 times, completing tasks in 1–21 minutes compared to the 69–646 minutes required for manual processing.

Performance was consistent across both stereoscopic and conventional cameras, meeting ICCAT monitoring standards. While environmental factors such as video quality, fish density, lighting, and gate visibility influenced outcomes, these findings provide valuable insights for further refining the system to ensure robust performance under varying conditions.

Overall, this study confirms that AI-based automation can effectively support BFT monitoring by maintaining high accuracy while significantly reducing processing times. The validation with authoritative harvest data from three Spanish transfers represents a significant milestone, demonstrating that AI measurements consistently exceed the accuracy of traditional manual methods when validated against true population parameters - achieving superior results in all three validated cases across diverse operational conditions from optimal to highly challenging environments. The expanded dataset of 18 total transfers further validated the system's robustness across diverse operational contexts. This innovative approach aligns with ICCAT's regulatory framework and offers a scalable solution for improving the efficiency and reliability of fisheries monitoring. The findings pave the way for future advancements in automated monitoring technologies, promising substantial benefits for sustainable fisheries management.

Appendices

1 Comparison of Calibration Methods

Figure 10 from Boutros et al. (2015) shows the proportional measurement errors of stereo cameras calibrated using a calibration cube, an A4 checkerboard pattern and an A3 checkerboard pattern.

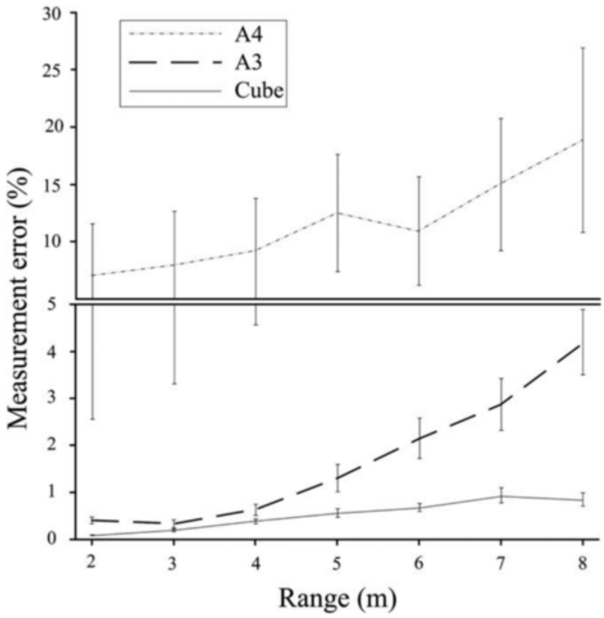


Figure 10: Proportional measurement errors of stereo cameras using different calibration methods.

Table 11 from Boutros et al. (2015) shows the resulting errors in mass estimates (as percentages) of a hypothetical 400 mm snapper (*Pagrus auratus*) using the mean error in length measurements from systems calibrated using a calibration cube, and a 2D A3 checkerboard, and an A4 checkerboard.

Range (m)	Cube	A3	A4
2	0.00	0.91	2.12
3	0.49	0.09	0.54
4	1.07	0.92	0.02
5	1.49	2.67	10.20
6	1.88	6.03	3.57
7	2.45	7.98	16.01
8	2.03	12.18	42.90

Figure 11: Mass estimate errors using different calibration methods for a 400 mm snapper.

2 Detailed Length Distribution Analysis

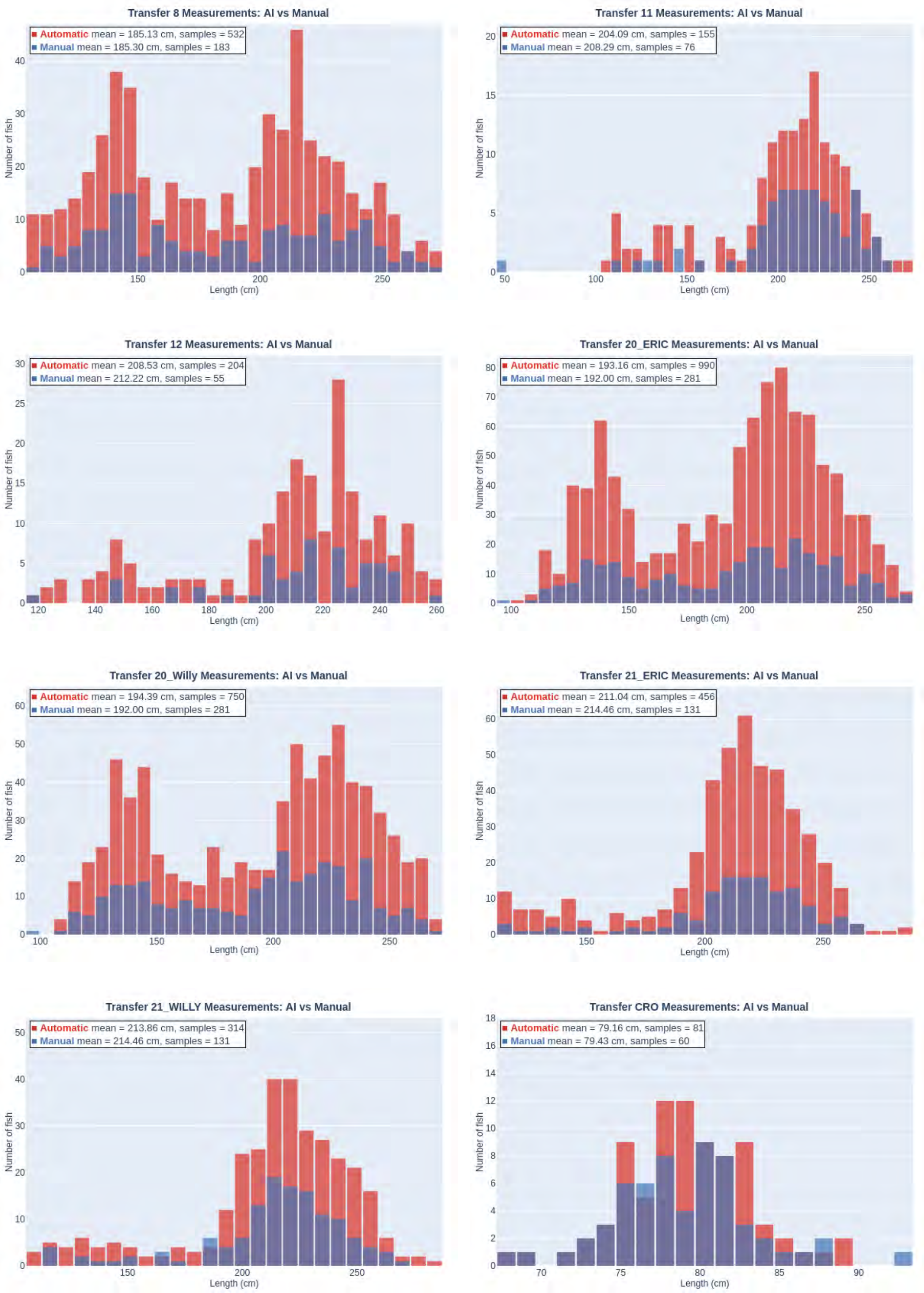


Figure 12: Length distribution (part 1). The histogram shows the frequency distribution of fish lengths measured manually and by AI. *'Number of fish' refers to the count of unique, individual fish measured.

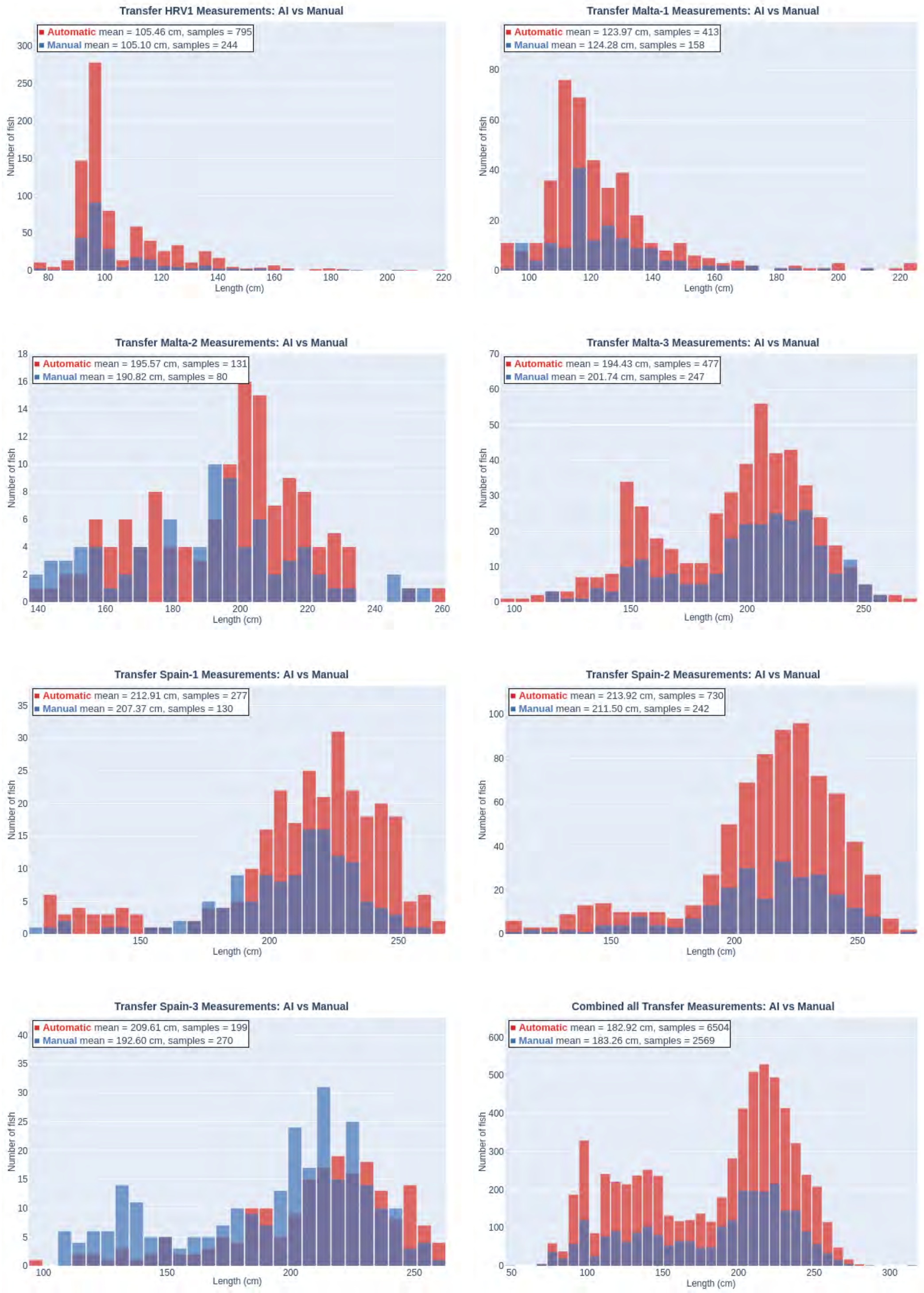


Figure 13: Length distribution (part 2). The histogram shows the frequency distribution of fish lengths measured manually and by AI. *'Number of fish' refers to the count of unique, individual fish measured.

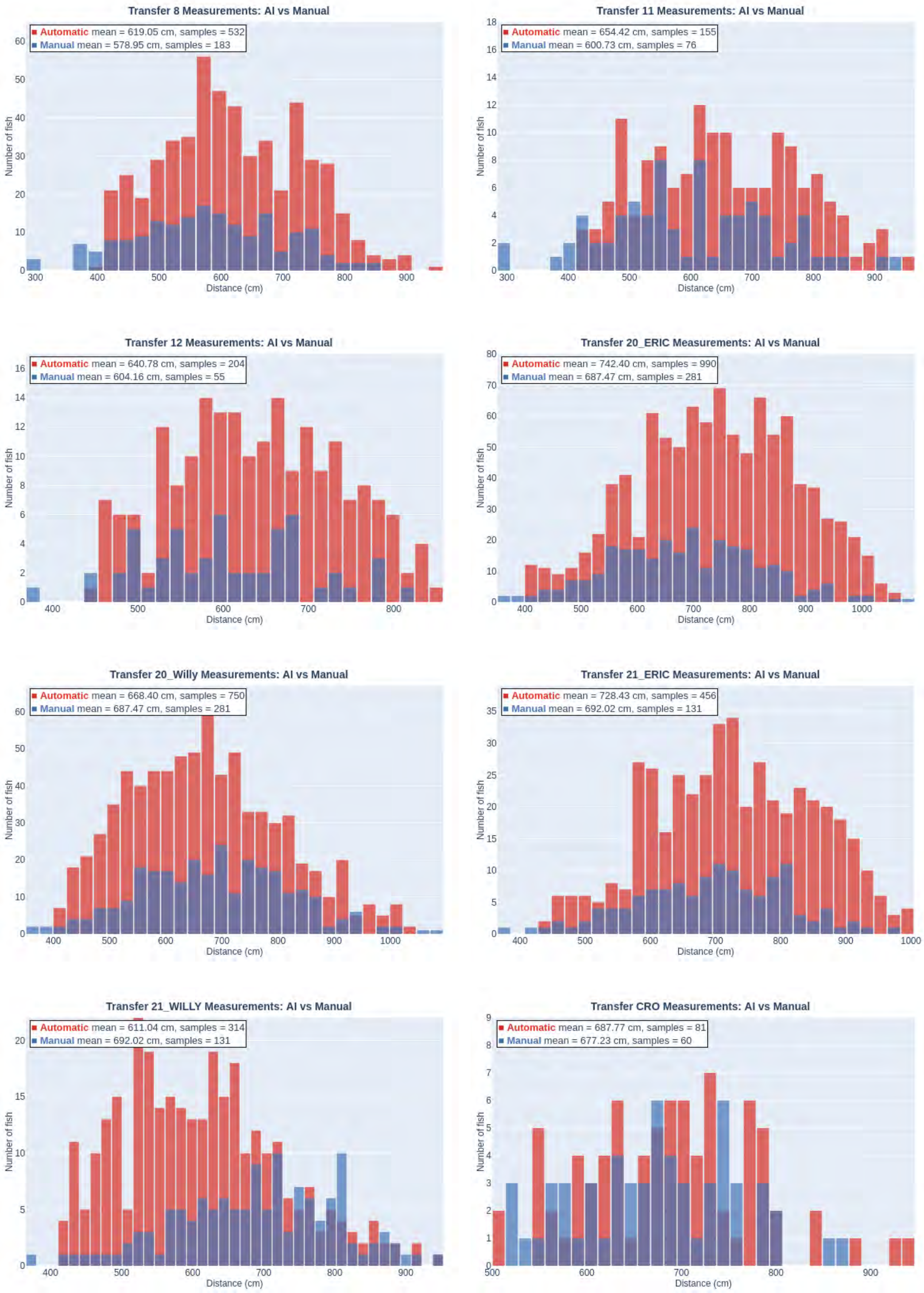


Figure 14: Distance distribution (part 1). The histogram shows the frequency distribution of distance of fish measured manually and by AI. *'Number of fish' refers to the count of unique, individual fish measured.

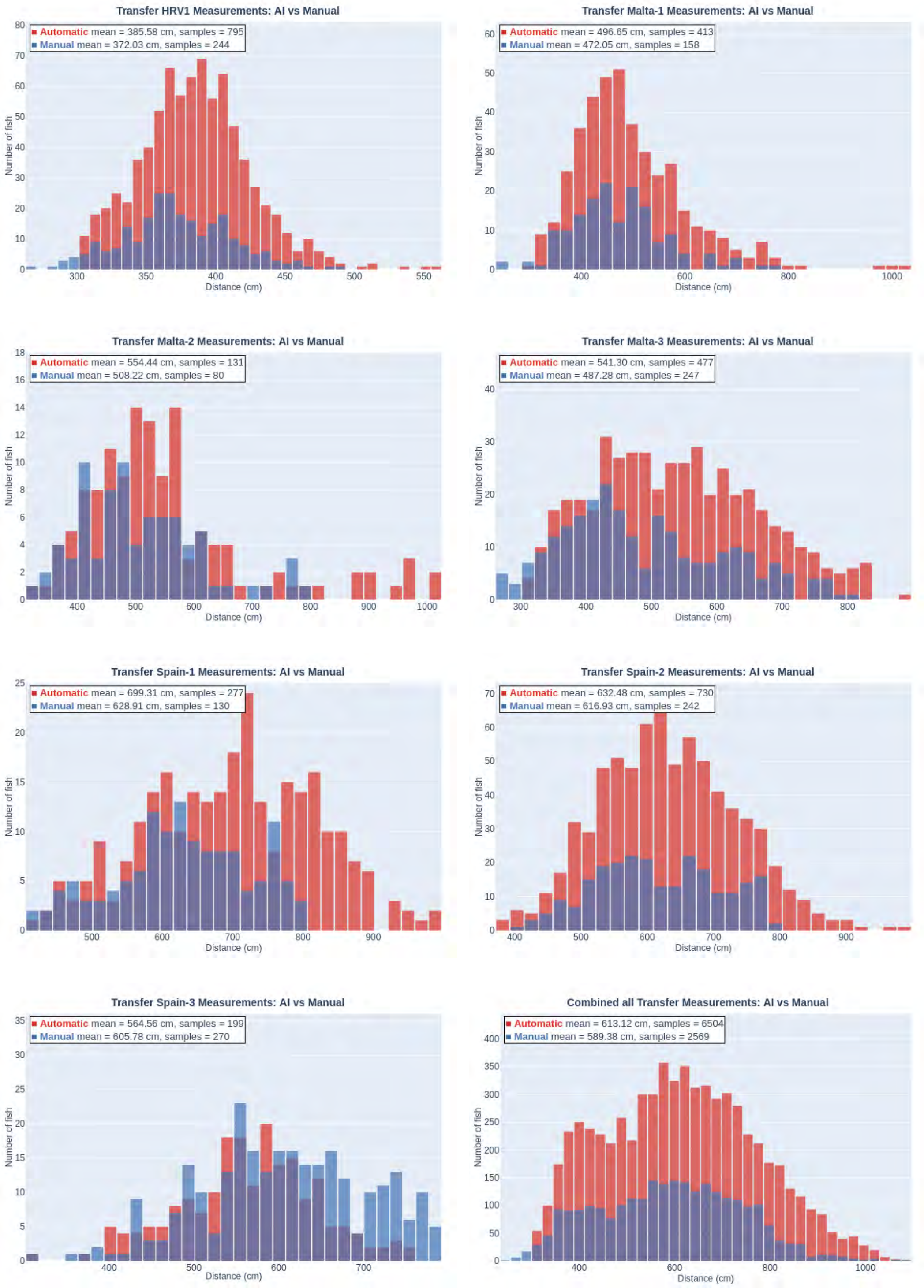


Figure 15: Distance distribution (part 2). The histogram shows the frequency distribution of distance of fish measured manually and by AI. *'Number of fish' refers to the count of unique, individual fish measured.

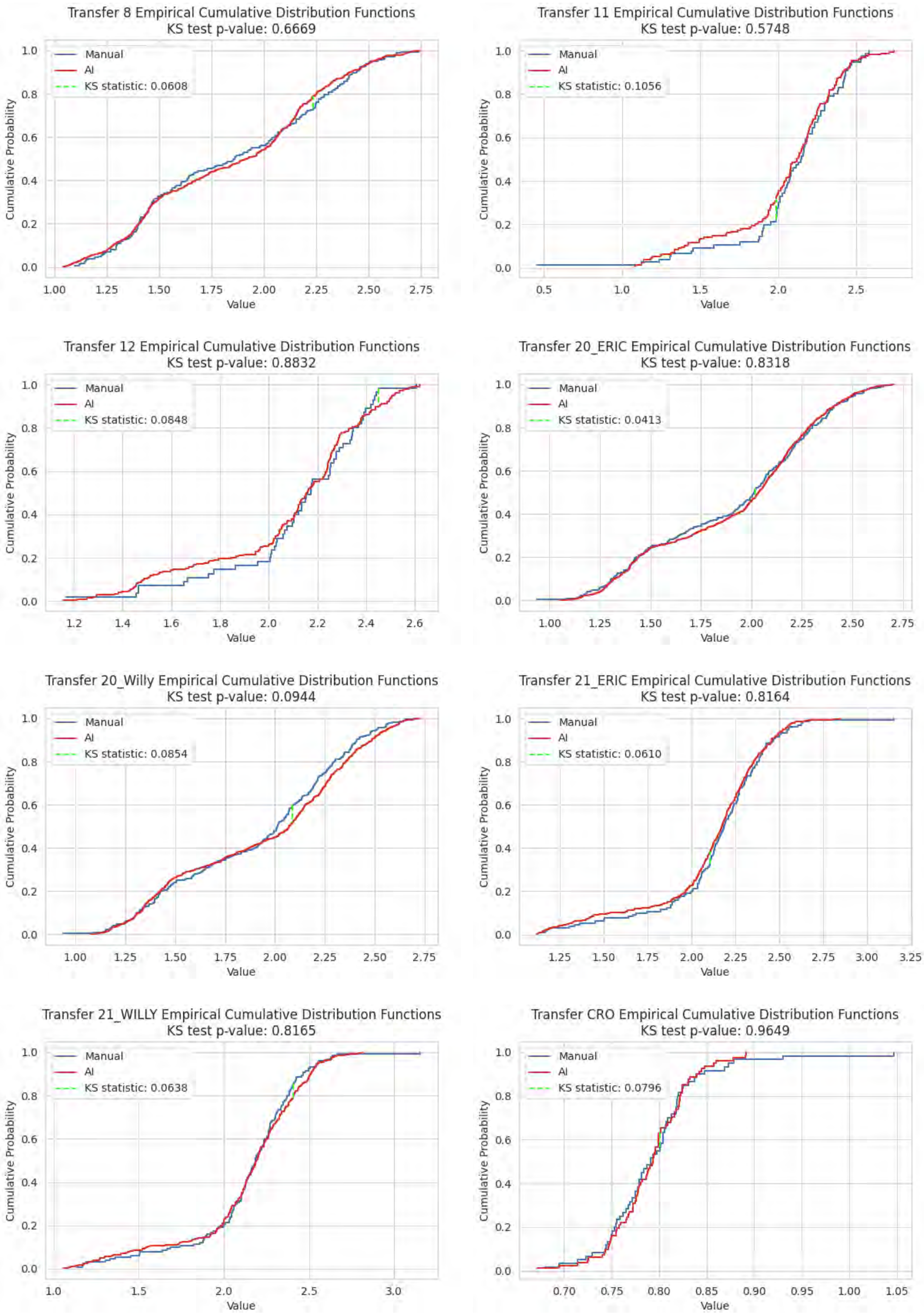


Figure 16: empirical cumulative distribution function (ECDF) (part 1). The ECDF plot shows the cumulative distribution with KS test results

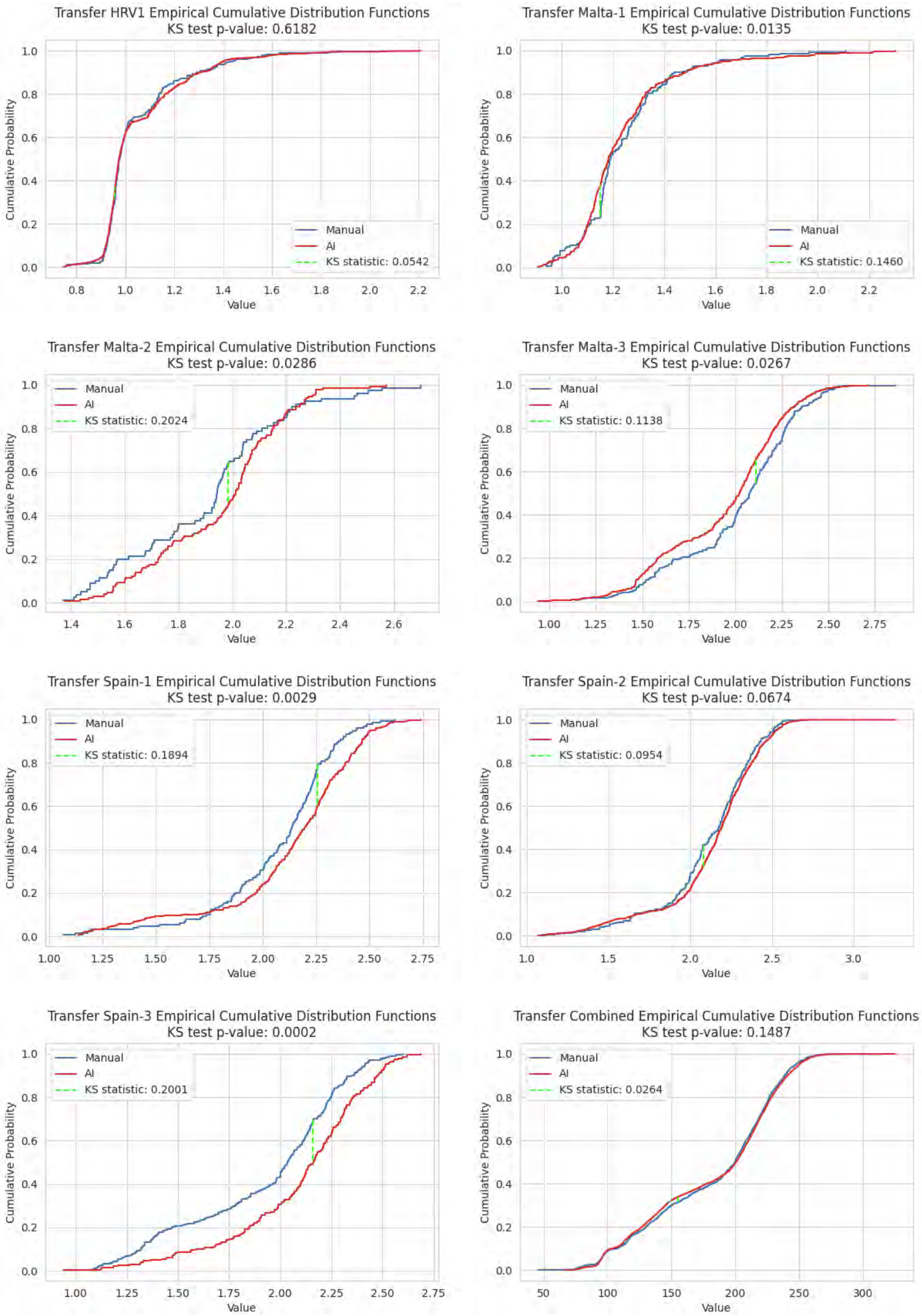


Figure 17: empirical cumulative distribution function (ECDF) (part 2). The ECDF plot shows the cumulative distribution with KS test results

References

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Van Essen, B. C., Awwal, A. A. S. & Asari, V. K. (2019), 'A state-of-the-art survey on deep learning theory and architectures', *Electronics* **8**(3), 292.

Boutros, N., Shortis, M. R. & Harvey, E. S. (2015), 'A comparison of calibration methods and system configurations of underwater stereo-video systems for applications in marine ecology', *Limnology and Oceanography: Methods* **13**(5), 224–236.

Harvey, E., Cappo, M., Shortis, M., Robson, S., Buchanan, J. & Speare, P. (2003), 'The accuracy and precision of underwater measurements of length and maximum body depth of southern bluefin tuna (thunnus maccoyii) with a stereo–video camera system', *Fisheries Research* **63**(3), 315–326.

Harvey, E., Fletcher, D. & Shortis, M. (2002), 'Estimation of reef fish length by divers and by stereo-video', *Fisheries Research* **57**(3), 255–265.

Harvey, E., Shortis, M., Stadler, M. & Cappo, M. (2002), 'A comparison of the accuracy and precision of measurements from single and stereo-video systems', *Marine Technology Society Journal* **36**(2), 38–49.

Shortis, M. (2015), 'Calibration techniques for accurate measurements by underwater camera systems', *Sensors* **15**(12), 30810–30826.