FINAL REPORT

Benchmarking for health and productivity in aquaculture.

Abalone industry

Tracey Bradley, Thomas Teoh, Nick Savva

May 2023

FRDC Project No **2018/180**

ISBN 978-1-76136-282-8

Benchmarking for health and productivity in aquaculture. Abalone industry

2018-180, 2023

Ownership of Intellectual property rights

Creative Commons licence

Disclaimer

Researcher Contact Details

Name: Dr Tracey Bradley
Address: Agriculture Victoria,
475 Mickleham Rd,
Attwood, VIC 3049
Phone: 0429 429 498
Email: Tracey.bradley@agriculture.vic.gov.au

FRDC Contact Details

Address: 25 Geils Court
Deakin ACT 2600
Phone: 02 6122 2100
Email: frdc@frdc.com.au
Web: www.frdc.com.au

In submitting this report, the researcher has agreed to FRDC publishing this material in its edited form.

Contents

Tables

Figures

Acknowledgments

The authors of this report wish to acknowledge that this project was supported by funding from the Fisheries Research and Development Corporation.

The authors would also like to thank Dr Ben Madin from AusVet Services for his assistance in data preparation and ingestion. Staff from Agriculture Victoria provided assistance in the connection between the data storage system and internal PowerBI function – in particular Dean Bellingham and Tony Monteith.

We would like to acknowledge all the Australian farmers who contributed the materials in this report. This project was for them, included their data and will hopefully be of benefit into the future. They are listed below:

Mat Skye – 888 (Western Australia)

Luke Thorpe, Tim Rudge and Louise Vosloo - Yumbah Corporation (SA, Vic and Tas)

Hamish Ebery and Mark Gervis - Southern Ocean Mariculture (Vic)

Carl Hitchings – Jade Tiger Abalone (Vic and Tas)

Nick Savva – previously of AbTas (Tas)

Joel Gilby – JST (Tas) – now known as “Three Friends”

Abbreviations

AAGA – Australian Abalone Grower’s Association.

DJPR – Department of Jobs, Precincts and Regions

FCR – Feed Conversion Ratio

PBI – Power BI – free Microsoft software for displaying data.

Executive Summary

Benchmarking is a form of evaluation undertaken by comparing a measure with a standard. With its widespread adoption across many industries, benchmarking was identified as an important area for development in aquatic industries through the national strategic plan for aquatic animal health (AQUAPLAN 2014 – 2019). Despite this industry recognition and internal identification of its importance, there has been no national benchmark conducted across any Australian aquaculture industry. Prior to the commencement of this project, benchmarking is undertaken and considered as important within companies. The absence of national benchmarking could be related to issues around maintaining data confidentiality and a lack of coordination in the numerous smaller aquatic species sectors. With information on key health and production variables, individual farms will be able to make good decisions around early detection and prevention of disease, improving overall individual and industry farm productivity.

The abalone aquaculture industry in Australia comprises 11 pump-ashore farms; all farms were enlisted to participate in this project. The aims of the project were twofold: to develop and standardise individual farm data collection and review systems in place used by farmers to record health and production parameters. To enable this, the project team set out to provide a secure data storage system with the capacity to allow industry benchmarking against industry agreed - upon standards.

Initial interactions with farms revealed a diaspora of data collection and storage systems used across the industry. Furthermore, there were variations in how key parameters were being measured, which needed to be accounted for. The steps required in this benchmarking project were to: understand what data was collected; process and map the data to a schema suitable for storage in a secure database; ingest and clean data as required; validate data then publish on the individual farm dashboard with access online through reports created in Power BI (PBI), securely hosted in the Agriculture Victoria Microsoft tenancy. Much of the work in this project related to understanding, checking and correcting data quality issues. Overall, five of the eleven farms elected to participate in the individualised, interactive farm dashboard while the remaining farms were provided with access to the overall industry benchmark for the variables of interest.

For the five farms opting to participate in the creation of the individual dashboard, an accurate geospatial map was constructed mapping every production tank using a geographic information system (QGIS, ArcGIS). The main variables benchmarked were associated with tank – level mortality, stocking rates and weight for age measures. Many of these were displayed on the interactive map as a choropleth or "heat map" which represented higher measures for the variable of interest as darker tanks on the map. The interactive nature of the map allowed for drilling down in the data related to both space and time. Further enhancements have been made at the request of the farmers, for example a colour coded map displaying cohorts and stocking rates. Following a farmer meeting held in early July 2022, the addition of information on feed conversion efficiency, water temperature and modelled production versus actual was requested. Currently this development is paused but could form part of a second tranche of variables assuming the data is suitably robust.

The creation of this benchmark is the first industry initiative of its kind in Australia and potentially internationally. A major part of the work of the project was gaining the trust of farms to provide their data to the secure environment. Another large piece of work was ensuring appropriate data quality, rectifying software issues and continual improvement for the participating farms in terms of providing robust and valid data. The farmers have free access to the secure database to visually interrogate their own data and that of the whole industry (as aggregated data). The project is designed to have an enduring legacy; providing data is received from the farms, the benchmarking will continue.

Introduction

The Australian abalone aquaculture industry is an important part of the seafood export landscape. Historically, abalone volume has come predominantly from wild caught abalone. However, this is changing with the contribution of aquaculture abalone product having grown from negligible in 2000-01 to a projected 38% of production by the end of 2026-27 (ABARES, 2022). The overall volume and value of Australian aquaculture was 1429 tonne and $54 313 789 respectively up to the FYE 2021 (Nick Savva, pers comm). Despite the deleterious impact of COVID-19 on abalone consumption internationally, the outlook for the abalone aquaculture industry is positive with plans for expansion in coming years.

To support expansion of the abalone and other aquaculture industries, industry -wide benchmarking has been proposed. Benchmarking can be defined as: "to evaluate (something) by comparison with a standard" (Oxford Language). More specifically, comparing organisational performance against that of their competitors, accepted standards or modelled estimations (Skipper-Horton, 2013). Benchmarking is useful for highlighting both the strengths and weaknesses in a business, and when implemented correctly should contribute to addressing any management areas where changes can improve performance.

The national strategic plan for aquatic animal health (AQUAPLAN 2014 – 2019) recognised the importance of benchmarking for health and production under Activity 3.3: "Undertake aquatic animal health benchmarking for specific aquaculture sectors" (Department of Agriculture, 2014). AQUAPLAN, as a five-year strategic plan is developed through consultation with a number of stakeholders: the Commonwealth government, jurisdictional governments and industry. This project aimed to assist farms in two ways: with an intuitive, farm-specific data display system and by conducting benchmarking across the abalone aquaculture sector. The benchmark provides feedback to individuals on their farm performance compared with the industry mean and other related variables to demonstrate where an improvement in production and health parameters is achievable. This project has strong links with other AQUAPLAN objectives such as Objective 1: "Improving Regional and Enterprise-level Biosecurity". There is clear evidence (particularly in terrestrial industries) of the link between animal health and productivity.

Industry benchmarking for animal health and production characteristics is commonly practiced in many intensive agricultural industries. The dairy industry, for example, routinely has access to many metrics associated with milk production. The national Dairy Farm Monitor Project undertakes a detailed annual survey of farmers using a standardised chart of accounts measuring a large range of variables such as kilograms of milk solids sold per cow or hectare; feed produced per hectare and number of cows milked per staff member (Dairy Australia, 2019). This project has been operational for decades (in the current and previous iterations) and provides feedback to farmers on the economic performance of the whole farm in terms of return on assets and equity. Following deregulation of the dairy industry in 2001, there was a strong push for increased efficiency on farms. Despite a low milk price by international standards, the dairy industry is the 3rd largest agricultural industry nationally worth $4 billion per year with 50% of product exported (Australia is the 3rd largest milk exporter in the world). Benchmarking has played a large part in ensuring that the "total factor productivity", which measures how efficiently and intensely inputs are utilised, is much higher in dairying (1.6% annual increase) compared to other traditional industries such as beef (0.5%). Improvements can occur when the practices undertaken by industry leaders are adopted across the whole sector for the benefit of all.

There is limited international literature related to benchmarking in aquatic industries, most likely due to company concerns about data sharing and maintaining a commercial edge over competitors. In reviewing rainbow trout production on Canadian farms, Skipper – Horton (2013) commented that in the sector there was a substantial variability in performance parameters that would benefit from benchmarking with an improvement and standardisation of data collection methods. Where modelled growth targets were not achieved, farms had to prematurely harvest fish to provide a timely product at the cost to the individual farm. This indicated a need to refine the currently accepted growth models with commercial data. Skipper-Horton (2013) noted that there was a diversity in management strategies and environmental conditions across operations and whilst the producers tracked performance of their own fish, there was little collaboration or sharing of data across farms. This resulted in a lack of industry standards for performance traits they wished to improve and no understanding of their own performance relative to these standards.

Some aquaculture sectors in Australia (particularly smaller industries) do not currently have a coordinated system in place for the routine collection and analysis of health and production data. This deficiency was illustrated in the Murray cod industry FRDC Project 2010-036 (Bradley et al, 2014) where it was found that even basic health and production parameters such as mortality rates were not always recorded. The statement made in the Executive Summary for the Murray cod project is illustrative: "It would appear that there is a strong need for transparent benchmarking of the costs and returns experienced by members of the industry that would enable the whole industry to move towards more profitable production using a more collaborative approach". Initial discussions with a range of industries at the commencement of this project (including abalone, barramundi and Murray cod), indicated that for many players an intuitive and secure system for recording farm productivity and health data was desirable and a first step towards farm benchmarking for health and productivity.

With information on key health and production variables, individual farms can make good decisions around early detection and prevention of disease, therefor improving overall farm productivity. On a larger scale, the ability of farms within an industry to benchmark their individual farm data against an industry standard provides useful information enabling the individual producer to confidentially compare personal farm data with the whole industry and then take remedial action where improvements should be considered (ie trouble shoot areas where the farm is falling behind). Other benefits for the individual producer include the ability to provide data quickly and efficiently to management, for example the Board of Directors, government and others if required; provide supportive evidence in times of stress such as during a disease outbreak, loss of trade etc and to support other initiatives such as compartmentalisation.

At an industry level, collecting a range of health and production variables can be used to support industry quality, certification and marketing. The involvement of industry members in a national benchmarking program could potentially assist with ensuring trade access to markets where disease freedom and biosecurity programs are adopted and integrated into an overall program.

Within the Australian abalone sector there has been limited sharing of data across enterprises. A project was contracted by the Australian Abalone Grower's Association (AAGA) in 2010. This was a repeat of a similar survey from 2007. The project was titled the "2010 AAGA Business Benchmark" and included responses from 6 farms which at the time represented 57% of the industry production. The survey focussed on cost of production and productivity. Not all farms responded to all the questions in the brief survey which was reported back to industry through a series of de-identified histograms. There were concerns related to this project around the ability of farmers to decipher farm identity and hence highly confidential data. It was a non-negotiable requirement of this current project to ensure the data was completely unidentifiable.

The benchmarking project was conducted across all pump-ashore abalone aquaculture farms in Australia (herein referred to as "the abalone industry"). Originally the project was designed to enlist 3 industries, however the abalone industry was considered the ideal candidate for benchmarking given the similarity of management across Australia. The Australian pump ashore industry is currently comprised of 11 farms: one in Western Australia, 2 in South Australia, 4 in Victoria and 4 in Tasmania (Figure 1). Farms under the same company structure are represented by the same colour.

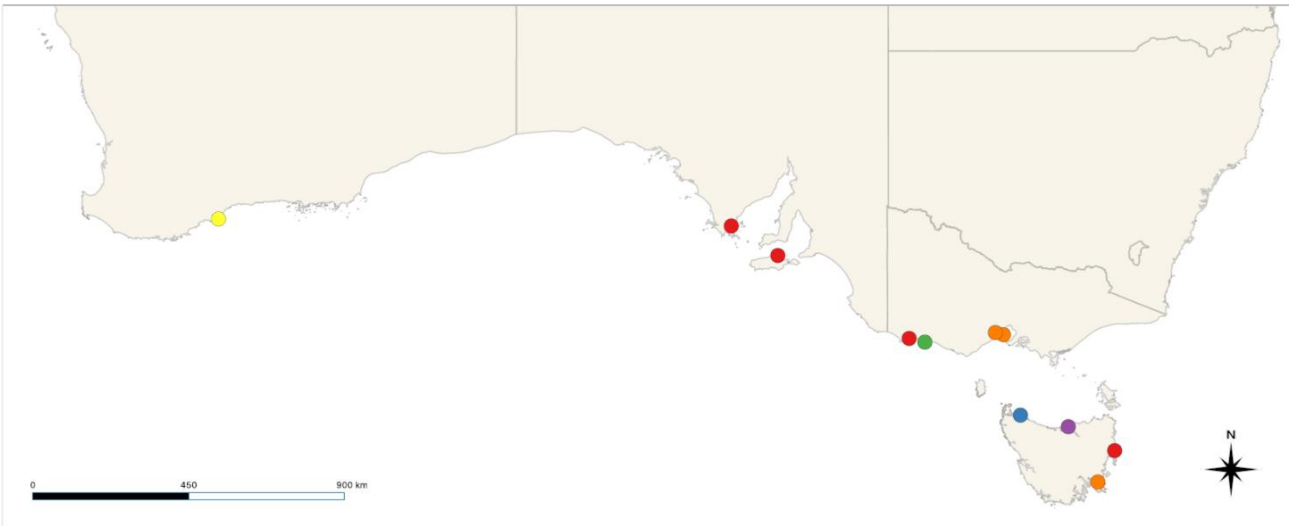


Figure 1. Location of Australian land-based abalone farms.

The level of data collection on farms varied: from minimal collection, to farms recording information but not using it to review farm performance, to larger companies where benchmarking is utilised at a company level to monitor performance and take corrective action where required. The benchmarked variables include mortality rates, growth rates, feed conversion efficiency and other health and production statistics. The lack of coordinated data collection impedes some individual farms in assessing and comparing their disease risk/status and overall performance both within farm and with the industry in general. Furthermore, if data collection and benchmarking was routinely undertaken in the Australian aquaculture industries there would be opportunities for producers to learn from other industry members how best to increase productivity and reduce biosecurity risks on individual farms and across the industry as a whole.

Objectives

In this project our objectives were to:

1 Develop a farm data collection and review system for farmers to record health and production parameters
2 Provide a secure data storage system with the capacity to allow industry benchmarking against industry approved standards - for example the industry median.

Method

1. Recruitment and data security of individual farms:

The initial recruitment and enrolment of the farms in the project was an extended process over many months. Recruitment activities included a presentation at the AAGA Annual General Meeting in 2019 and individual meetings with farm managers and staff (too numerous to document). The occurrence of the COVID-19 pandemic during this time did not assist with engaging the industry. Physical farm visits, where trust was to be built, could not occur due to travel restrictions. Farms were also understandably very concerned about the impact of the pandemic on their sales and how to manage large volumes of mature stock that were ready to harvest but without a market.

A major concern for farmers was around confidentiality. As such, a legally crafted agreement was created and signed by all farmers (Appendix 1). There were many discussions held with a range of stakeholders providing assurances of confidentiality. The desire for a high level of data security is completely understandable and has provided interesting challenges with respect to feeding back on progress in this project to the funding body and participants. Discussions with colleagues embarking on similar activities highlighted that it took 2 years to gain the trust of Chilean salmon farms to provide data related to a project on ISA vaccination (Ausvet Services pers comm).

There were also internal departmental project management and IT requirements that needed to be managed prior to progressing the work which are summarised in the results.

The project agreement incorporated a "Go – No Go" point as the first milestone where a minimum of 5 industry members with appropriate quality data were required for the project to proceed. By enlisting all 11 Australian farms, this requirement was easily met.

Early in the project, it was clear that some farms were more interested in having the capacity to visualise their data through a mapped format. Five farms volunteered to be part of the more extensive dashboard "Hal" which provided farms with a range of information related to an electronic map (tank level) created for the farm.

2. Improve on-farm data collection:

The Australian abalone aquaculture industry is comprised of 11 individual farms (Figure 1). Seven of the farms belong to 2 companies operating across 3 states. The remaining 4 farms are individually owned and operated.

To ensure rigour across all the variables being benchmarked, individual discussions were required with each farm to explore how data was being collected, handled and reported. Responses were generally provided at the company level for the corporate farms and individual holding level for the independent farms. Where there were variations within the company in data management and husbandry, these were noted.

The farmers were surveyed about a range of data collection practices on the farms. A series of questions was developed relating to what the farmers themselves considered important parameters to measure; this process was undertaken in an iterative manner with careful consideration of how they would be quantified according to the farm data. The project did not investigate data related to the nursery stage of growth during the first 6 months of production as there was little data collection relating to this phase of the lifecycle. The results from the discussions around farm data collection are presented within the text in the results section.

The on-farm data collection process can be broken down into 2 main steps:

1. Physical data collection – this includes activities that occur on farm and the systems in place to ensure they are conducted accurately and consistently. For example, collection of numbers of mortalities per tank or measuring the weight or size of a sample of animals to estimate an individual average weight and the overall biomass of a tank. Farmers were consulted about this aspect of data collection, with knowledge captured in a spreadsheet and used to inform subsequent data analyses.

2. Data recording systems – there are a range of electronic data capture systems used by the farms (some being supported by paper-based records on the farm). The systems vary from bespoke, custom developed individual processes to proprietary commercial off-the-shelf software as outlined previously. Samples of all data recording systems were reviewed to determine suitability for data ingestion and to establish the robustness of the overall data recording systems in the way they tracked husbandry metrics. This process involved direct communications with farms, management or the AbTrack software developer to collect, review and negotiate a suitable methodology for ongoing data provision.

Once data was collected in its myriad individual respective forms from the farms, a process of consolidation and standardisation of measurements was undertaken in a complex ETL (Extract/Transform/Load) process to take information from farm input files and store in the Omnisyan database ("Aquamark"). Generally, the steps undertaken for each farm included:

1. A systematic review of the data collection system outputs to identify what form production data took.

2. A process of mapping fields and values to a schema suitable for storage in the database

3. Ingestion of data and identification of non-conformant data (names in a date field), which necessitated either data cleaning at the source or construction of a routine to cope with errant values. This was codified for automation in instances where the form of the data coming in was sufficiently well structured and consistent.

4. Assessment of the ingested data to identify mis-mappings or other issues with the ingestion process

5. Validation of data by individual review of summary data and visualisation for sense checking

Bringing in data from semi-manual systems (such as spreadsheets) involved a significant amount of time and effort invested to handle unpredictable data formats and presentation.

3. Process mapping:

The various steps in production and data collection points for each farm were documented to assist in assessing data comparability and for the development of an appropriate database. The creation of the database structure, mapping the source data, combining data and reporting the data using PBI are outlined in the results.

Given the interconnection between the data collected and process mapping for the farms, the information provided in the results section can be considered evidence of the effort invested in improving the rigor of on-farm data systems.

4. Farmer provision of information:

The initial steps required a frank appraisal of the data farms were already collecting. Samples of the variously formatted spreadsheets were collected alongside detailed discussions about how data was captured and interpreted.

Initially data provision was through the project Principal Investigator; a participating farm would send records or data extracts to the Principal Investigator's email address, who would then manually forward it to the database administrator who would enter the data into the system. For ongoing submissions, an automatic submission system was developed to facilitate continued integration of data into the database; this took the form of an email address which was setup to receive data files as attachments. Once received, a workflow would be triggered which would automatically deploy a parsing routine to read in the data (according to the farm submitting the data) and feed it into the database. Any data files that were unsuccessfully processed would be handled manually by a database operator.

Farmers were provided with the option at all stages to leave the program and be provided with their data. They were given the opportunity to specify whether their data was deleted or archived according to their requirements.

5. Accessing online

The desired aim of the project was to develop an interactive system where farmers could access their data online and be provided with feedback directly.

All farmers that had shown an interest in the “Hal” individual farm dashboard were provided with access online through reports created in Microsoft Power BI (PBI), securely hosted in the Agriculture Victoria Microsoft tenancy. Individual access for stakeholders is provisioned by the sharing of reports to individual email addresses (managed by the Principal Investigator) to only allow for access to a farm’s reports for each individual; this access is further secured by the requirement for multi-factor authentication when first registering on the site. Further security is available in the reporting data model in the form of role level security implemented against each farm.

It should be noted that although all farm data is stored in the cloud (AusVet AWS), the data displayed was only as current as the data provided and ingested.

Results *

1. Farm recruitment and data security

All eleven farms (or representatives for the corporate entities) signed the legally drafted confidentiality agreement whereby the Department agreed to handle all data in a highly secure manner and not allow dissemination to other parties of individual farm-level records (Appendix 1). This applied to the individual farm-level records. The industry level records were made available to all farms as part of the national benchmarking component in an aggregated format. The data aggregation ensured that data was sufficiently obfuscated so no farm had the ability to directly access or discern other farm's data.

The departmental IT security team developed 12 steps to be addressed to ensure security of industry data and further that state government requirements were being met (Table 1).

Task number	Task
1	Solution diagram/information flow
2	Provide db design schema for data storage
3	Quantify and address risk to business of farmer data loss.
4	Determine if personal information stored etc through Cloud
5	Ensure appropriate encryption in place
6	Data destruction upon project completion
7	Host provides vulnerability scanning
8	Confidentiality of authentication data and implement MFA
9	Access to data configured securely for farmers
10	Data sharing agreement between DJPR and AusVet
11	Resources available in DJPR for ongoing maintenance
12	Deliverables made available and kept securely

Table 1. Twelve steps for data security as required by Agriculture Victoria

Data was managed by an external provider (Ausvet Services) who have been contracted by Agriculture Victoria on an ongoing basis to maintain and secure information systems. Confidentiality agreements encompass the contracted relationship that Agriculture Victoria has with Ausvet Services around data held by the Department and that of third parties. Collected data was and continues to be secured in a managed cloud ecosystem backed by Amazon Web Services infrastructure. Access to the data is strictly controlled using user credentials through a front-end portal, and scoped database credentials for whitelisted IP addresses for the database directly.

2. Improving farm data collection and description of practices

The Benchmarking Project aimed to improve data collection practices whilst allowing farms to continue with their existing data recording systems. Many of the farms have been relying on existing systems for decades and there was limited appetite for making changes to these systems. Further, a commitment of the project was to not require the farms to "double enter" data into a system forced upon them. Furthermore, as outlined below, there was wide variation within this small industry in how various key parameters were measured. With these constraints, the development of a suite of harmonised variables within an industry analogous to the creation of a "Chart of Accounts" where all variables measured are consistent across the industry was challenging.

* Data in this section applies to a number of farms and is not representative of any single enterprise.

The data needs of farms are relatively simple: they wish to know how many abalone are dying in a tank or cohort, the biomass (weight) of abalone in a tank (including stocking rates), individual average weights (or sizes) in a tank and when a tank needs to be harvested. The data recording system needs to be able to accurately record movements of stock around farms (which occurs frequently) when stock are graded, culled and harvested (either total or partially of larger stock in the cohort).

In total, across all farms in the project from 1/6/2021 to 30/5/2022 over 2 million records were ingested into the Jock database.

Six of the farms utilise a proprietary software known as "AbTrak". This software was created over 20 years ago and presents a simple click-and-enter windows interface backed by a "clarion topspeed" database. From the developer of AbTrak: "The aim of AbTrak is to provide the best estimate of each tank's biomass at the end of each month" (Tom Carswell pers comm). The program AbTrak records some variables (eg mortality rates) but actually models many of the variables of interest such as the growth, biomass of a tank and average individual abalone weight in a tank. Most farms that use AbTrak will record corrections to the modelled data when there has been an activity undertaken that provides "real-life" data.

Anecdotally the farms believe the modelling used by AbTrak is relatively accurate (within 10%). As can be seen from the 2 examples below (Figure 2) comparing actual and modelled weight and % growth, there appears to be a larger variation which does not support this assertion, although it may be that once averaged across all tank data this approximates to 10%. The adjustment to stocking rates, feed etc relative to the output from AbTrak tends to occur on a more adhoc basis depending on the farm. Experience over decades appears to have made the farm managers less reliant on the predictions of AbTrak. Although over 50% of the farms by number (and much more by volume of production) are using AbTrak, it appears that part of the industry would prefer a more contemporary data solution that is cloud based with real-time reporting capabilities. At the time of writing this solution had not been resolved.

Biomass kg	Actual kg	Actual Wt g	Diff kg real and modelled
360	420	60	17%
366	420	54	15%
440	400	-40	-9%
398	420	22	6%
387	400	13	3%
391	400	9	2%
378	400	22	6%
385	400	15	4%
375	400	25	7%
388	400	12	3%

growth	AB Trak Expect
	% growth
14.00%	6
8.49%	6
15.65%	6
12.85%	6
15.34%	12
	12
	12
22.97%	12
12.57%	12
27.70%	12
8.30%	12
-6.06%	12
3.07%	12

Figure 2. Location of Australian land-based abalone farms. Examples illustrating the difference between modelled data in AbTrack and farm measurements

Two farms have bespoke spreadsheets that have evolved over time. The spreadsheets have the capacity for modelling growth. The data in these spreadsheets is a much more robust reflection of farm activities as it is usually based on more frequently recorded and collected data.

Only one company records data directly into a cloud-based system where daily data is visible to management as it is entered. This farm elected not to provide individual records to the project but was very willing to assist with providing aggregate data.

Although data entry may occur on a daily basis on many farms, the ability to analyse data is generally retrospective when reports are generated on a monthly or quarterly basis. For most farms, this has been traditionally how data has been analysed although many farms see the benefit of receiving contemporaneous feedback on data.

The range in data collection methods is of interest. Some farms collect data (such as tank level daily mortality) on paper as staff move around the farm, and transfer to an electronic format sometime later (the time period varying between farms). Staff on other farms utilise tablets recording data directly into a tablet which connects to a central data repository enabling farm management to access all data in real time. An outcome of this project has been the upskilling of some farms using the AbTrack software to move from paper – based to electronic recording of farm data. This has reduced the need for double handling of data – with the inherent risk for human error. Further, one farm has increased its data upload period from 3 months to one month. One farm also moved from monthly recording of mortalities to daily.

Mortality data:

All farms collect data on mortalities occurring at a tank (or group of tanks) level, usually around 3 times per week. The frequency of mortality data collection varies with the season (usually more frequent in summer) and if there are health issues on the farm (more frequent where there are perceived problems). Further, although some farms recorded mortalities at a tank level 3 times per week, this data may be recorded only monthly in the system.

Mortality data seemed to be the most logical variable to commence benchmarking, however how this fundamental variable was recorded amongst farms varied. Most farms record mortalities at an individual tank level with one farm recording aggregate mortalities across a bank of 16 (or more) tanks. It was difficult to calculate the percentage mortality for some farms where the denominator (ie number of abalone in a tank) was not known for periods of time.

Issues with the AbTrak software in recording % mortalities were identified when the data was initially interrogated. This error was not recognised by the farms and was rectified. The fact that there was an error in the software for some time would indicate that farms were not using this information generated from AbTrak; this is a significant missed opportunity to better understand and improve their operations. Towards the end of the project, there were errors appearing in the submitted data as per Figure 3 which made the ingestion process extremely difficult. Where farms are submitting over 20 000 records per month this is understandable (but unfortunate).

Date	Sp.	Y. Cl.	Age	Biomas	Nom W	Nom Q	Trans Ir	Trans C	Harvest	Growth	Morts	Mort kg	Mort %
25/05/2022	GL	19	30							0	245	22.03	###.###
25/05/2022	GL	19	30							0.04	245	22.03	167.808
25/05/2022	GL	19	30							0.05	245	22.03	135.359
25/05/2022	GL	19	30	20	88.4	222				-21.52	245	21.65	52.463
25/05/2022	GL	19	30	27	89.9	298				-21.88	245	22.03	45.12
25/05/2022	GL	21	6	0	1.3	56				-0.06	44	0.06	44
25/05/2022	GL	19	30	32	89.9	360				-21.87	245	22.03	40.496
25/05/2022	GL	19	30	40	88.4	452				-21.46	245	21.65	35.151
25/05/2022	GL	19	30	57	88.4	649				-21.4	245	21.65	27.405

Figure 3. AbTrak data demonstrating errors in recording Mort %

Most farms have a designated trigger level for when mortalities should be reported to management – this may be clearly marked as a whole number in the recording system. Delays in data input could conceivably cause delays in the reporting of significant mortalities but this seems unlikely.

Over the life of the project, some farms moved from paper-based recording of mortalities with the inherent risk of transcription errors to the use of tablets with AbTrak software installed. We believe that by developing a dialogue between farms such improvements were a direct outcome of the project.

The variables measured for the 5 farms utilising "Hal" and mapped back to tanks were: number of mortalities by date, numbers by date and weight, total kg by date and by date and weight category, weight of mortalities by stocking density and weight category (Figure 4). All farms enrolled in the “Hal” dashboard had an accurate, electronic representation of their tanks in situ. The electronic map is totally interactive such that a date of interest can be specified, individual tanks or groups of tanks on the farm can be selected to provide a history of the variable of interest or a weight class can be selected. This functionality is difficult to describe statically and cannot be demonstrated in this report given the confidential nature of data, but a partial representation is shown in Figure 5.

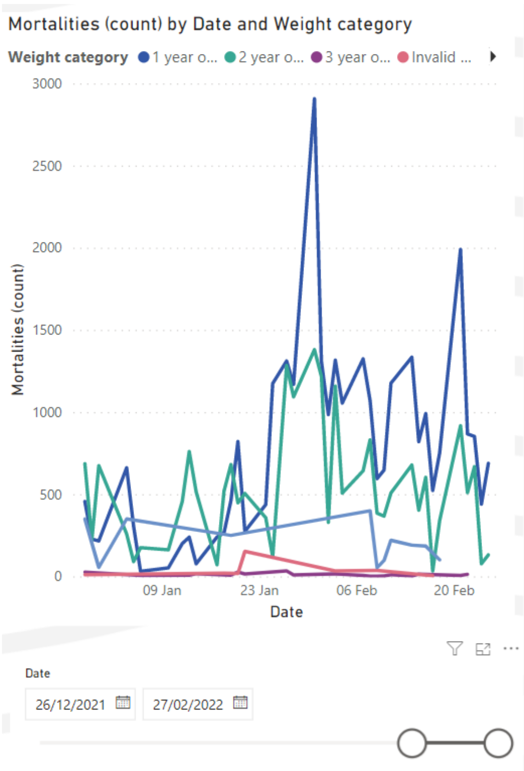


Figure 4. Individual mortality count by date and weight category

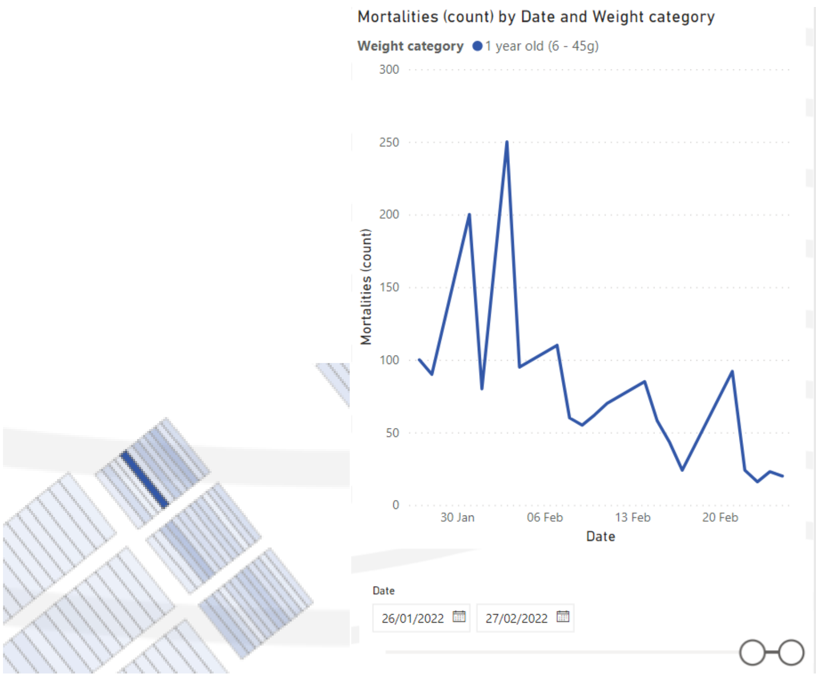


Figure 5. Partial view of dashboard demonstrating function of selecting individual tank for mortality count over a specified period.

The benchmarked parameters across the industry were total mortalities and percent mortalities by date across all farms and total numbers and percent mortalities for cohorts across all farms (Figures 6 and 7).

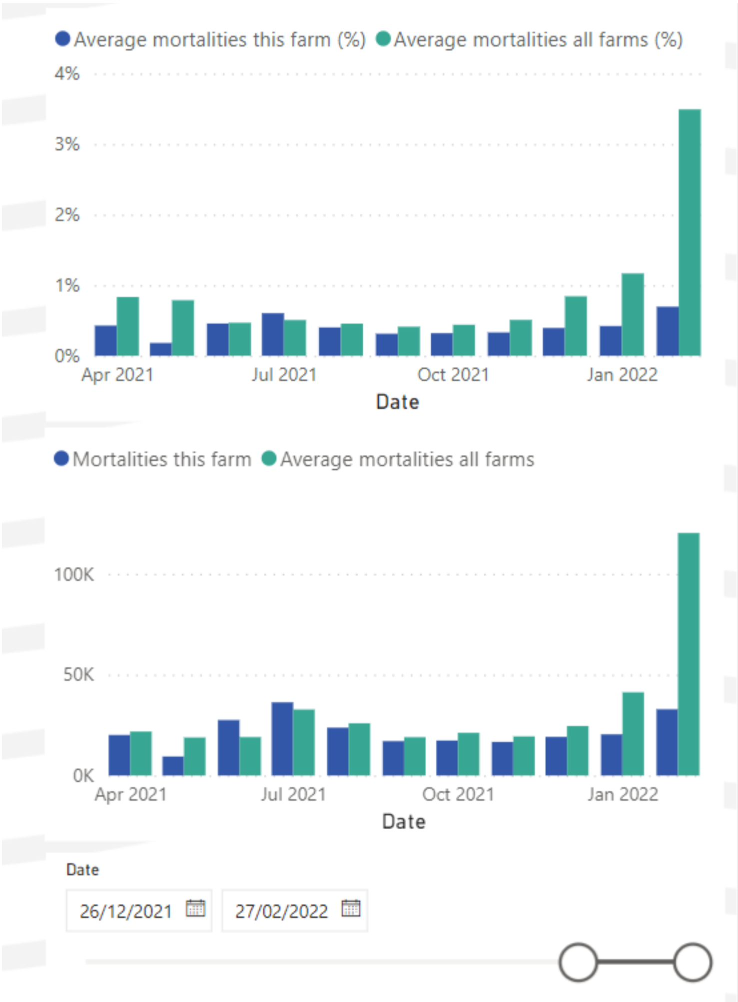


Figure 6. Benchmarked comparison across farms for mortalities this farm (blue) vs all farms (green)

Months since spawning	2017	2018	2019	2020
5				3267 (0.52%)
6				3586 (0.31%)
7				12672 (1.08%)
8				18005 (1.88%)
9				7285 (0.77%)
10				5993 (0.51%)
11				13614 (0.55%)
12				13988 (0.62%)
13				10692 (0.44%)
14				13148 (0.53%)
15				42964 (1.44%)
16				5513 (0.39%)
17			4705 (0.2%)	2497 (0.19%)
18			1082 (0.05%)	3573 (0.28%)
19			1524 (0.08%)	3992 (0.31%)
20			2425 (0.14%)	
21			2467 (0.15%)	
22			3010 (0.2%)	
23			2790 (0.2%)	
24			4916 (0.33%)	
25			6655 (0.43%)	
26			5826 (0.3%)	
27			11550 (0.6%)	
28			1796 (0.19%)	
29		3494 (0.35%)	666 (0.08%)	
30		1467 (0.14%)	1326 (0.13%)	
31		2212 (0.18%)	1751 (0.13%)	
32		2945 (0.29%)	3388 (0.26%)	
33		2974 (0.32%)	2729 (0.21%)	
34		2196 (0.27%)	13036 (1.03%)	
35		2212 (0.31%)		
36		2360 (0.41%)		
37		4363 (1.14%)		

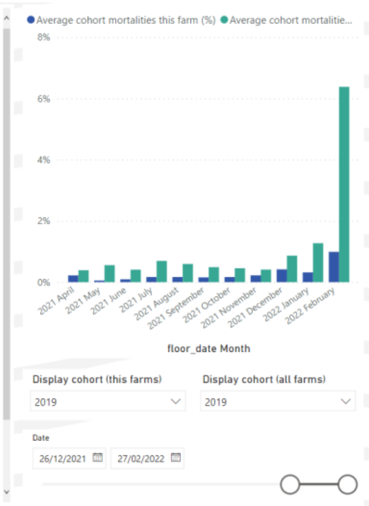


Figure 7. Benchmarked comparison of mortalities for a cohort.

Stocking rate data:

The number of kilograms of abalone per unit of tank area is a commonly used industry metric for stocking rate. Some farms also collect data on % coverage of a tank i.e. how much of the tank area is utilised, this assessment is made by eye. In Australian land-based abalone aquaculture farms pump water ashore and utilise concrete or plastic tanks for holding stock. Most farms have rectangular concrete tanks in the grow out area under shade cloth with varying levels of block out (to 100%) (Figure 8). Three of the eleven farms have outdoor tanks with hides (two farms with round tanks and one with rectangular tanks). The hides are generally constructed of concrete and it can be difficult to calculate the precise surface area of these structures. (Figures 9 and 10). There should be innovative changes in quantifying stocking rates and biomass on farms given a current FRDC project utilising photography and artificial intelligence.

Figure 8. Concrete slab tanks in partially shaded shed.

Figure 9. Round concrete tanks outdoors with hides.

Figure 10. Plastic lined tanks with weaners on concrete hides

Stocking rates were displayed as a choropleth or "heat map" (density) over time for all the "Hal" farms. This attribute was further delineated by cohort year. This enables farms to manage cohorts of stock related to their density so that they can easily visualise where stock may need to be graded or harvested (Figure 11). Colour coding helps management direct staff to specific areas of the farm where the information is visible on a tablet (or other device). Average stocking density for each month across the farm was compared I.e. individual farm versus all farms (Figure 12).

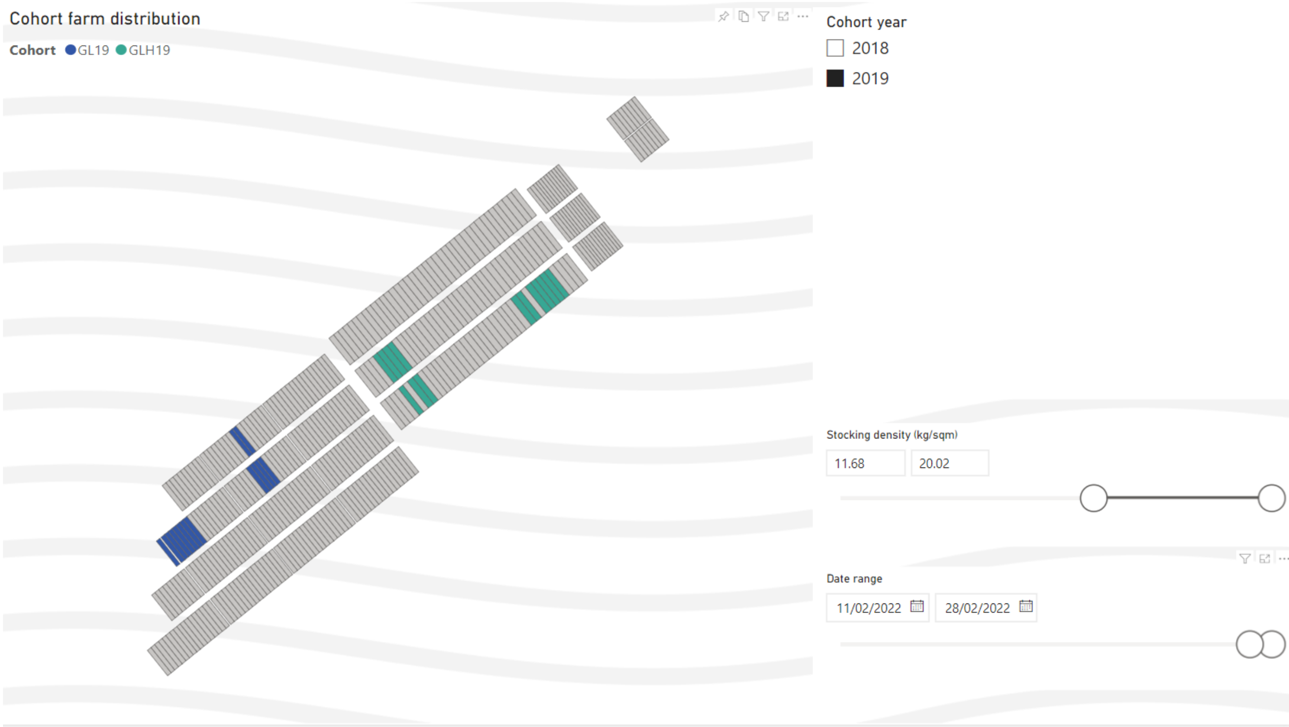


Figure 11. Cohort density map

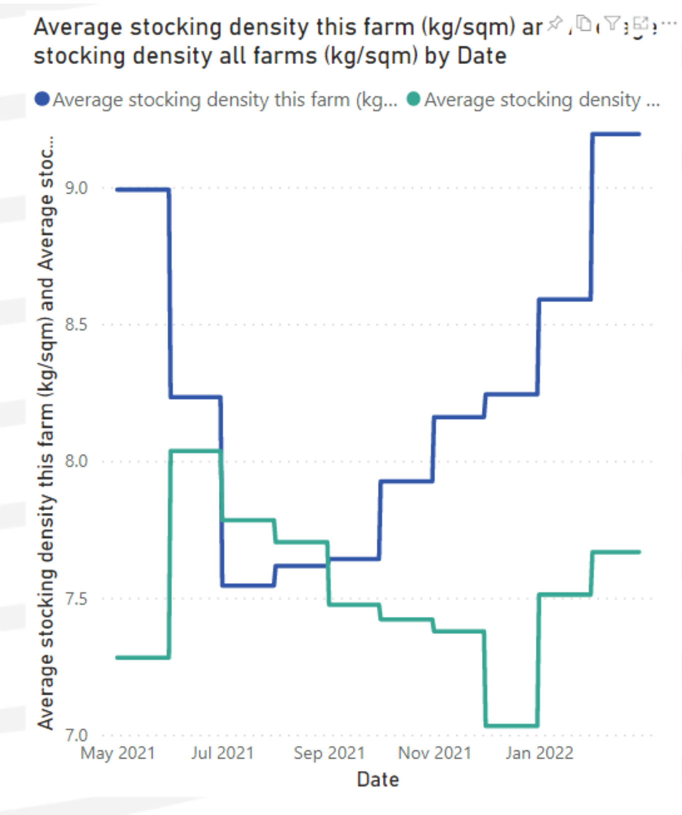


Figure 12. Comparison of average stocking density this farm versus all farms.

Data was also provided as the count of mortalities relative to age of stock and density over a specific period (Figure 13). This enables farms to interrogate the mortality patterns for different age groups relative to stocking density. It can be seen from this graphic that there are a large number of mortalities in the category "invalid weight" where data may need to be reconciled.

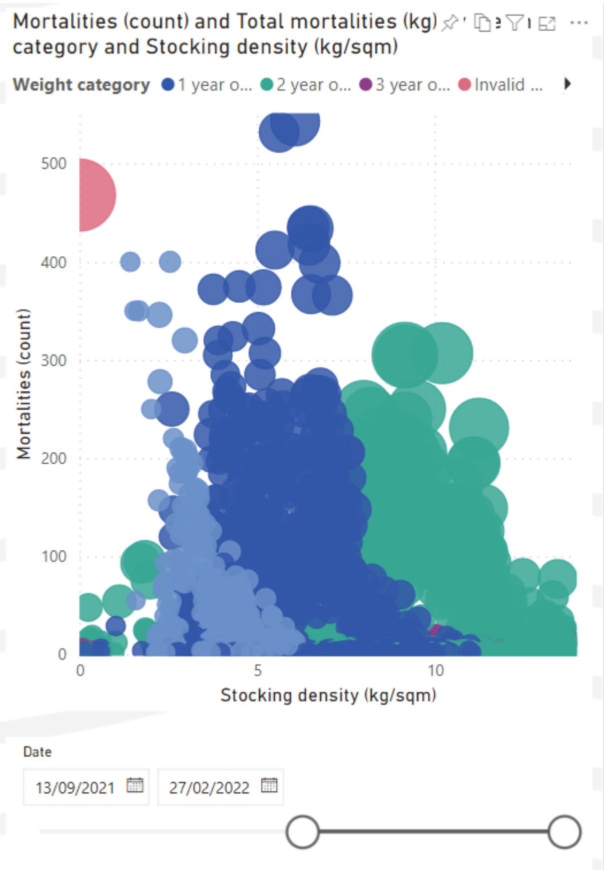


Figure 13. Mortality count relative to age and density of stock

Production data:

As previously indicated, there were major issues establishing a comparable measure for productivity. How production data was arranged and measured varied amongst the farms.

There are 3 main areas which contributed to a lack of clear, comparable data across all farms.

Terminology issues:

A cohort is defined by the farmers in many ways: when abalone are spawned; when animals are set in the nursery; or when they are moved into the grow out area (up to 8 months later). This resulted in the "2018 cohort" being either spawned in 2018 or moved to grow out in 2018 (which in effect means the latter cohort should be labelled "2017 cohort"). Abalone may spawn between the start of October and start of February but the date may be set at November 1. Some farms undertake a small amount of "out of season" spawning in April, a practice that seems to be increasing. Given this variability in terminology, the most logical way to have consistency across the farms was to measure production in terms of months since spawning (as this was usually reliably recorded).

Measurement issues:

Accurately calculating the average weight of individual abalone in a tank, the number of abalone in a tank and the biomass (total kg of stock) of a tank can be challenging.

All farms measure a subsample of abalone in a tank to approximate the weight (or length) of the individuals and the biomass of the whole tank. Where length is measured, a proven constant is used to convert lengths to weights (and vice versa). Taking these measurements across a large number of tanks is very time consuming. How this is activity is performed varies across farms and can therefore add to the variability of the results. Between 30 and 100 abalone were reported as being sampled and measured, the number varying with the size of the stock. In selecting the sample, farms generally tend to randomise and may, for example, select from 3 lines across the tank, a specific quartile of the tank etc. However, stock selection may also be sampled opportunistically.

The regularity of stock sampling varied considerably across the industry: monthly, quarterly, before Board meetings, when a tank is moved, when stock were taken from a tank, or not at all until harvest. For some farms there is no specific program for conducting measurements and there is a heavy reliance on experience. Farms weigh the tanks out at harvest to get total kg of biomass held in the tank and will often estimate average weight and number of abalone at this time. However, in cases where only part of a tank is harvested, it is difficult to know the total biomass of a tank.

Grading occurs either mechanically or manually at harvest but may also occur on farms at different stages in the life cycle. Farms may grade and cull stock at various stages of the growth cycle: e.g. at 6 months out of the nursery, at 1 year or at harvest with no culling occurring at this stage (e.g. stock movement from nursery to weaning to grow out). The whole tank is weighed at harvest.

Data recording issues:

Six of the eleven farms use the proprietary software AbTrak. This software is editable and models the production variables (weight of individuals, biomass, stocking rates and growth). Farms using this software generally make changes to the modelled figures which are reflected in the following month's data. This activity relies on farms measuring and correcting the modelled variables which may not occur across all tanks in an ordered manner.

Farms using bespoke spreadsheets do not encounter the issues outlined above. The data entered is not derived and these farms may have a more regular measurement regime. Modelling future growth occurs through formulae (often complicated) that have been developed over time, and often shared between farms.

There is currently only one company where data is available online at the time of measurement. For all other farms, data is collected and reported retrospectively (usually at the end of the month).

Production measures

The farmers use different measures to calculate production and efficiency including % growth per month, days to reach a certain weight, staff per kg of abalone produced etc (Figure 14).

S	**late grade**		Tank Area cm2	745000
20-Nov-18	12R~15SS			
1-Nov-17				
4700	**LENGTH 0**	68	**Growth, g/day**	0.11
5014	**LENGTH 1**	69	**GAIN, kg**	14
454	**um/day:**	49.90	**% Area**	**34**
50.1	L, mm (actual)	69.3	**HARVEST No.**	0
40.3	L, mm (model)	63.8	**HARVEST kg.**	0

Figure 14. One example of how production data is recorded.

The AbTrack software program models growth and does not reflect actual growth occurring in each tank. The reconciliation of the true biomass and growth in a tank may not occur until harvest so it may take a complete life cycle to have an accurate reading of the biomass in a tank and the growth rate. The frequency and level of accuracy of estimates varies between companies and farms. This issue is extremely complex and is currently being resolved. The current piloted measure of growth is average weight (g) per animal by months since spawning.

Other data and future variables:

Feed:

Farms were interested in benchmarking feed use across the industry. The problem with benchmarking this major cost was the nature of measurement and recording. Many of the farms have sufficient experience to feed “by eye” and not record feed provided at a tank or shed level. These farms will then calculate the feed conversion ratio (FCR) at the end of the financial year by dividing the total feed purchased by the biomass gained. This approach suits the purposes of these farms. Some farms record feed usage on spreadsheets and use these levels as a guide to staff. Other farms weigh feed rations out to tanks every day and adjust to optimise growing conditions. Given this variation in approach, pragmatically, the annual FCR can be calculated and benchmarked across farms if farms provide the data. This was an intended benchmark to add to the final report however data was received by 3 farms only.

Water temperature (and other water quality measures):

All farms have the intention to collect water quality data however in some cases the equipment was currently non-functional. Water temperature is the first parameter of interest (see Figure 15). The location and number of water temperature loggers varies between farms: some had one, some had many, sometimes they were only in the nursery or at the intake pipe where in other cases they were in every bank of tanks. Data was collected from every half hour to once per day. Some farms collect data on oxygen as mg/L or percent saturation but not all farms collect this data. Some farms collect data on other parameters such as salinity.

Despite the importance of water temperature data, due to the variability in equipment, frequency of collection and location of data loggers it was believed to be too difficult to robustly benchmark this variable. Again, given farmer feedback, it is hoped that in the future this variable could be included.

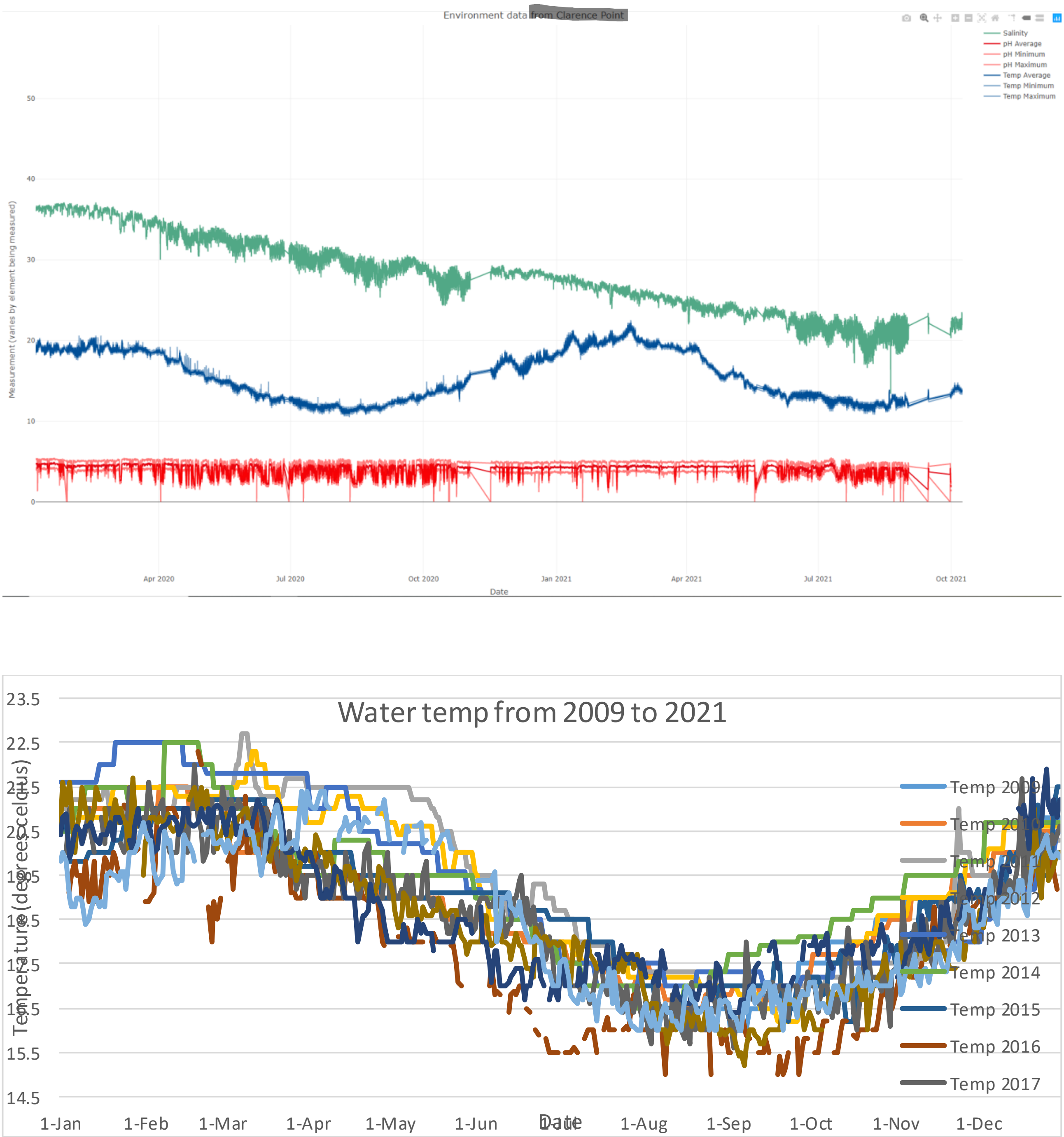


Figure 15. Water quality data from two of the farms.

FTE/unit of production:

Some farms were very interested in benchmarking labour costs per unit of production. Theoretically this should be possible however on smaller farms, there is often a blurring between operational and administrative roles. In the future, if valid data can be collated, this variable will be benchmarked, most likely on an annual basis. In this current project, the farms have agreed that this variable is too difficult to include in the benchmark.

3. Process mapping

Process mapping assists in ensuring the project team is collecting the data that the industry is most interested in above the default set of variables.

Step 1: Database structure:

A system architecture was designed to ensure that data was able to be handled and stored appropriately to allow for reporting desired outputs. There are 3 major stakeholders in this process: the farms, the Department and AusVet (external contracting organising managing the cloud-based data repository).

The initial design of the database was a hierarchical schema (Figure 16) accommodating data from company to site to tanks to records, with cohorts and movements adjacent. This has further grown organically as the project continued and more farm data formats needed to be accommodated to be a modified version of this, but the core structure remains.

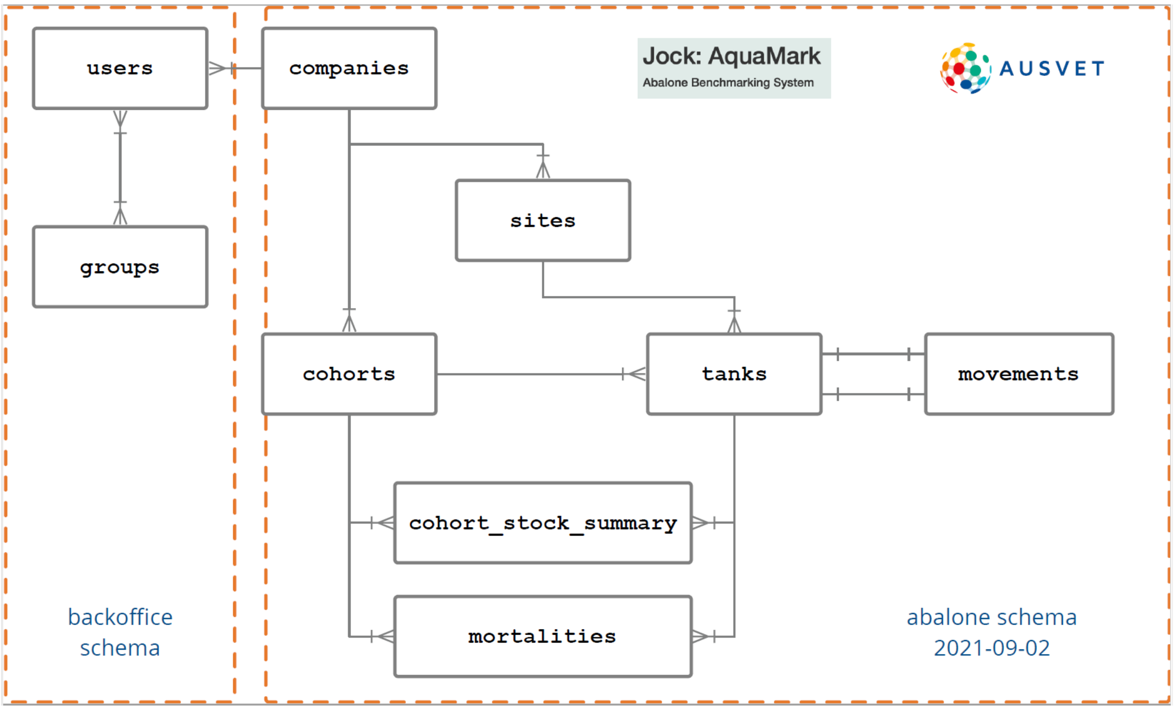


Figure 16. data base schema number 1.

The data was then ingested into a Power BI data model, which was established in the AgVic Microsoft tenancy. This connection was secured using specific database-level credentials only used for this purpose, and the use of a white-listed IP address configuration using an on-premises data gateway in the AgVic Bendigo datacentre (Figure 17).

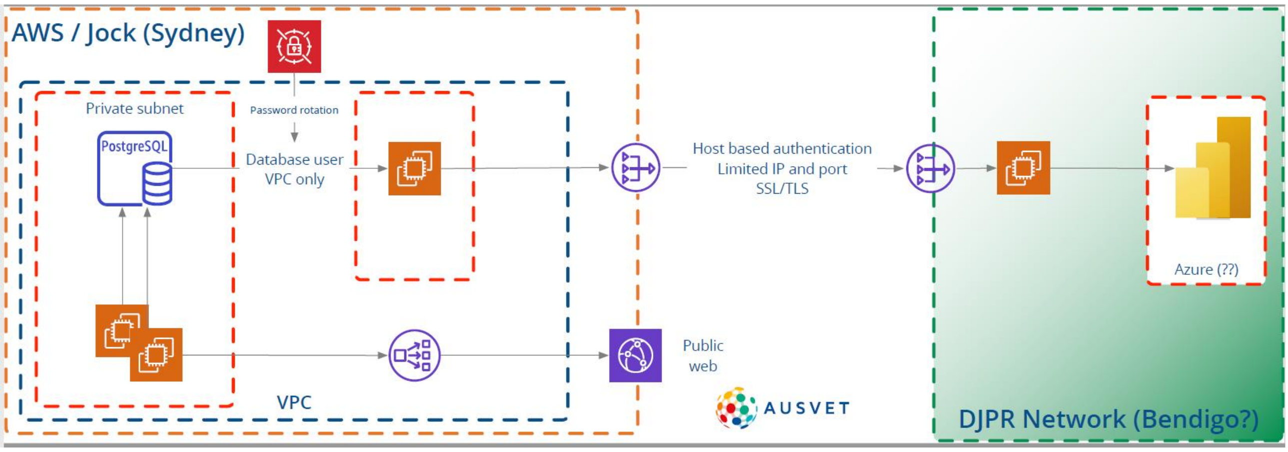


Figure 17. Database schema number 2 - demonstrating interactions

As part of this process, a manual process of data mapping needed to be undertaken at the point of ingestion (Figure 18); while this was laborious, it yielded unexpected findings such as data collection issues which were fed back directly to the participating organisations.

Step 2 (and 3): Mapping the source data

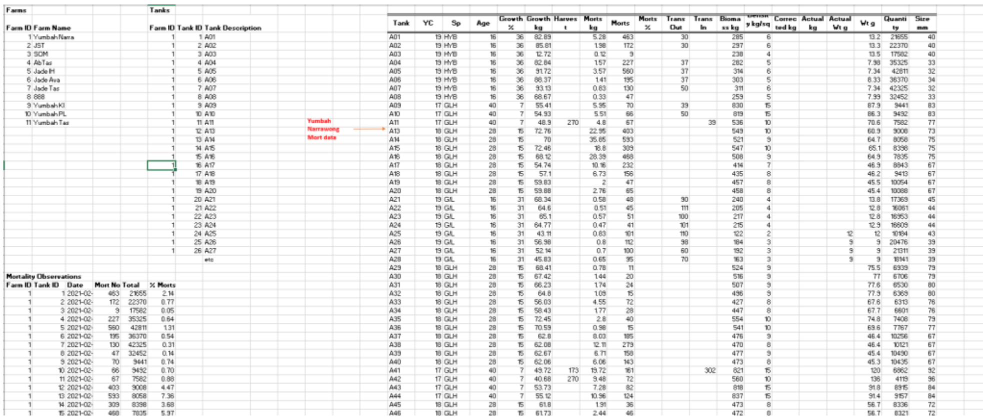


Figure 18. Example of how source data is mapped

Combining data from different tables (Steps 2 and 3)

Record level data is spread across the following tables as the ETL process handles data from different recording systems:

- `cohort_stock_summary`
- `mortalities`

This information is related to each other in the following way, and can be brought together to a summary table:

cohort_stock_summarytable	mortalitiestable	summarytable
id	id	id
record_date	mortdate	record_date
tankid	tankid	tankid
cohortid	cohortid	cohortid
growth_perc	(not available)	growth_perc
growth_kg	(not available)	growth_kg
harvest	(not available)	(not available)
morts_kg	(not available)	morts_kg
morts_num	mortnum	morts_num
morts_perc	mortnum/atrisk	
transfers_out	(not available)	(not available)
transfers_in	(not available)	(not available)
biomass_kg	(not available)	biomass_kg
density_kg_sqm	(not available)	density_kg_sqm
corrected_kg	(not available)	(not available)
actual_kg	(not available)	(not available)
actual_weight_g	(not available)	(not available)
weight_g	(not available)	weight_g
quantity	atrisk	quantity
size_mm	(not available)	size_mm
ratio	(not available)	(not available)
comments	(not available)	(not available)
createdon	(not available)	(not available)
createdby	(not available)	(not available)
modifiedon	modifiedon	(not available)
modifiedby	(not available)	(not available)
del	del	(not available)
morts_diff	(not available)	(not available)

Table 2. Tables in the ETL process

There is aggregate data located in:

- `tank_summaries`

This was not included in the record level data to ensure that double-ups in accounting didn't occur. With the methodology of calculating percentage mortalities in place this should be acceptable.

SQL query as a table

To bring the above information into the one table, the following script was used:

```
-- Define the dataset to be working with
WITH a AS
(    SELECT *
-- Cast variable as a numeric with two decimal places
      -- Otherwise it rounds it off and we lose detail
        ,CAST
        (
            mortnum AS NUMERIC(9,2)
        ) AS mortnum_num
-- Change the atrisk population to 1 where there is an atrisk of 0
-- To ensure that mortalities don't undergo a divide by 0 error
        ,CASE
            WHEN atrisk=0 THEN 1
            ELSE atrisk
        END AS atrisk_nozero
FROM "abalone"."mortalities"
)
-- Bring together the tables using UNION
SELECT
    id,
    record_date,
    tankid,
    morts_kg,
```

```
    morts_num,
    morts_perc,
    quantity,
    growth_perc,
    growth_kg,
    biomass_kg,
    density_kg_sqm,
    weight_g,
    size_mm
FROM
    "abalone"."cohort_stock_summary"
UNION
SELECT
    id,
    mortdate AS record_date,
    tankid,
    NULL AS morts_kg,
    mortnum AS morts_num,
    mortnum_num/atrisk_nozero * 100 AS morts_perc,
    atrisk AS quantity,
    NULL AS growth_perc,
    NULL AS growth_kg,
    NULL AS biomass_kg,
    NULL AS density_kg_sqm,
    NULL AS weight_g,
    NULL AS size_mm
FROM
    a
```

Additional queries were used to aggregate the data to summary tables, for direct comparison across farms.

```
WITH unioned_table AS (
    SELECT
        id,
        record_date,
```

```
        tankid,
        cohortid,
        morts_kg,
        morts_num,
        morts_perc,
        quantity,
        growth_perc,
        growth_kg,
        biomass_kg,
        density_kg_sqm,
        weight_g,
        size_mm
    FROM
        abalone.cohort_stock_summary
    UNION
    SELECT
        id,
        mortdate AS record_date,
        tankid,
        cohortid,
        NULL AS morts_kg,
        mortnum AS morts_num,
        NULL AS morts_perc,
        atrisk AS quantity,
        NULL AS growth_perc,
        NULL AS growth_kg,
        NULL AS biomass_kg,
        NULL AS density_kg_sqm,
        NULL AS weight_g,
        NULL AS size_mm
    FROM
        abalone.mortalities
),
joined_table AS (
    SELECT
        unioned_table.*,
        tanks.name as tank_name,
        cohorts.cohort,
        cohorts.spawn_date,
        to_char(record_date, 'YYYY-MM-01') AS record_month,
        sites.name AS site_name,
        companies.name AS company_name
    FROM
        unioned_table
    LEFT JOIN
    (
        SELECT
            id,
            name,
            siteid
        FROM
            abalone.tanks
    ) tanks
    ON
        unioned_table.tankid = tanks.id
    LEFT JOIN
    (
```

```
        SELECT
            id,
            cohort,
            spawn_date
        FROM
            abalone.cohorts
    ) cohorts
    ON
        unioned_table.cohortid = cohorts.id
    LEFT JOIN
    (
        SELECT
            id,
            name,
            companyid
        FROM
            abalone.sites
    ) sites
    ON
        tanks.siteid = sites.id
    LEFT JOIN
    (
        SELECT
            id,
            name
        FROM
            abalone.companies
    ) companies
    ON sites.companyid = companies.id
    WHERE companies.name != 'Jade Tiger Abalone'
),
grouped_by_tank AS (
    SELECT
        record_month,
        company_name,
        site_name,
        cohort,
        tank_name,
        spawn_date,
        -- Across the month, the total number of mortalities per
        -- tank is the sum
        sum(morts_num) AS morts_num,
        -- The morts_perc approach involved determining the average
        -- morts_perc throughout the month. I.e.
        -- Sum the morts_perc readings then divide by the number of
        -- days these reading pertain to.
        -- Unsure how to determine the number of days
        -- 1. Can't assume 30 days
        -- 2. Can't take the datediff of the beginning and end date
        -- in the month because there may be dates before and after
        -- that were not recorded.
        -- 3. Can't assume that only days that were observed are
        -- a complete record. There will be gaps in days
        -- periodically
        -- For now assume that it is across 30 days
        sum(morts_perc) / 30 AS morts_perc_percentagemethod_needswork,
        -- Grouping by tank, need to estimate average tank pop
```

```
        avg(quantity) AS quantity,
        -- The maximum weight in grams for that month for that cohort in
        -- the tank
        -- I.e. the tank reached X g in the period
        max(weight_g) AS weight_g
    FROM
        joined_table
    GROUP BY
        record_month,
        company_name,
        site_name,
        cohort,
        tank_name,
        spawn_date
),
grouped_by_site AS (
    SELECT
        record_month,
        company_name,
        site_name,
        cohort,
        spawn_date,
        -- Across the month, the total number of mortalities per
        -- site is the sum
        sum(morts_num) AS morts_num,
        -- Continuing on with the previous working for the avg
        -- percentage method, the average of the day-to-day
        -- tank level morts_perc is the morts_perc at the site
        -- level
        avg(morts_perc_percentagemethod_needswork) AS
morts_perc_percentagemethod_needswork,
        -- Across the month, the site population is the sum of
        -- of the averaged day-level tank populations.
        sum(quantity) AS quantity,
        -- The average weight across all tanks within the cohort
        -- and site
        avg(weight_g) as weight_g
    FROM
        grouped_by_tank
    GROUP BY
        record_month,
        company_name,
        site_name,
        cohort,
        spawn_date
)
SELECT *,
    -- Convert to date to allow for better usage later on
    CAST(record_month AS date) AS record_date,
    -- Using the site-level mortalities and population we
    -- can approximate a morts_perc
    morts_num / quantity AS morts_perc_populationmethod
FROM grouped_by_site
UNION
-- Bring in Jade Tiger summarised data
SELECT
    to_char(record_date, 'YYYY-MM-01') AS record_month,
```

```
    companies.name AS company_name,
    sites.name AS site_name,
    spawning AS cohort,
    spawndate AS spawn_date,
    NULL as morts_num,
    cohort_monthly_mort_perc AS morts_perc_percentagemethod_needswork,
    NULL as quantity,
    avg_weight_g AS weight_g,
    record_date AS record_date,
    cohort_monthly_mort_perc AS morts_perc_populationmethod
FROM
    jade_tiger_cohorts records
LEFT JOIN (
    SELECT *
    FROM abalone.sites
) sites
ON records.siteid = sites.id
LEFT JOIN (
    SELECT *
    FROM abalone.companies
) companies
ON sites.companyid = companies.id
```

Other such tables to provide dimension to the data or add additional facts were constructed in a similar fashion to mould and aggregate data to the correct unit of interest as required.

Step 4: Data modelling using Power BI (reporting the data)

A reporting data model was constructed using Microsoft PowerQuery and Microsoft Power BI. These tools were selected to provide a relatively easy interface for development and maintenance, alongside enterprise grade performance and security features.

The model is constructed independently, and individual farm reports will be able to be produced linking to that data.

```
aquamark-abalone-model

  ├ aquamark-abalone-yumbah

  ├ aquamark-abalone-jst

  ├ aquamark-abalone-abtas

  └ etc...
```

Database connections have been setup to allow a secured connection into the Omnisyan (Jock) platform to pull out data. PostgreSQL ODBC driver connections configured at the system ODBC level are used to manage connection strings. Parameters within these strings are relevant for the user and type of access, and allow for the configuration of data gateways to be used to access data. This approach is compliant with the security considerations that are in place, including whitelisting IPs to reduce the potential for unauthorised access.

This pattern of working as been replicated across developers and Agriculture Victoria IT to facilitate seamless migration from development environments to production (Figure 19).

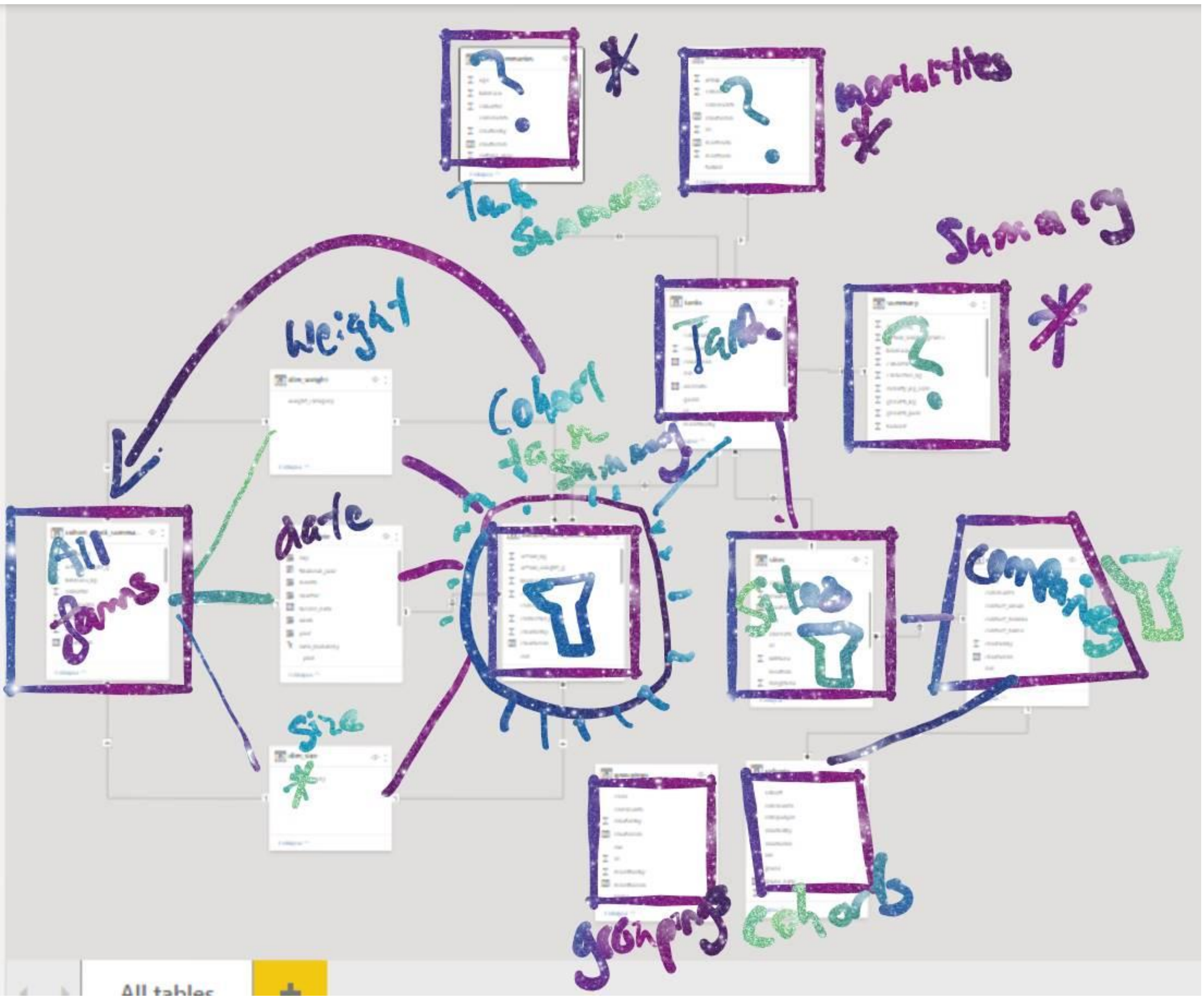


Figure 19. Overall schema

The summary table, imported and renamed to cohort_stock_summary is considered the fact table. This table has been duplicated and joined in different ways to enable us to provide reporting of a farm population vs the whole population.

Additionally, this table has been furnished with additional classifications for the weights of the animals listed on a record-by-record basis, based off definitions from the farmers or estimated from the data.

Sizes

Category	Classification
<35	Less than 35 mm
35 to <55	35 to less than 55 mm
55 to <65	55 to less than 65 mm
65 to <80	65 to less than 80 mm
80 to <90	80 to less than 90 mm
Invalid size	Null values or values out of range (and likely

	erroneous)

Weights

Category	Classification
0 to <6	0 to less than 6 g, Weaners
6 to <45	6 to less than 45 g, 1 year old
45 to <130	45 to less than 130 g, 2 year old
130 to <200	130 to less than 200 g, 3 year old
200 to <300	200 to less than 300 g, 3+ year old
Invalid weight	Null values or values out of range (and likely erroneous)

Table 3. Sizes and weights

The M query code used to calculate this is provided below.

```
   #"Added Conditional Column" = Table.AddColumn(cohort_stock_summary_Table,
"size_category", each if [size_mm] = null then "Invalid size" else if [size_mm] < 35
then "<35" else if [size_mm] < 55 then "35 - <55" else if [size_mm] < 65 then "55 -
<65" else if [size_mm] < 80 then "65 - <80" else if [size_mm] < 90 then "80 - <90" else
if [size_mm] >= 90 then ">90" else "Invalid size"),

#"Added Conditional Column1" = Table.AddColumn(#"Added Conditional Column",
"weight_category", each if [weight_g] = null then "Invalid weight" else if [weight_g] <
0 then "Invalid weight" else if [weight_g] < 6 then "Weaners (0 - 6g)" else if
[weight_g] < 45 then "1 year old (6 - 45g)" else if [weight_g] < 130 then "2 year old
(45 - 130g)" else if [weight_g] < 200 then "3 year old (130 - 200g)" else if [weight_g]
< 300 then "3+ year old (200+ g)" else "Invalid weight"),
```

Calculated dimension tables

Dimension tables are synthetically created to round out the data modelling for reporting. This is done based off data present in the tables themselves and are linked to the fact tables. The dimension tables produced include:

- `dim_date` representing the date dimension
- `dim_size` representing the size dimension
- `dim_weight` representing the weight dimension

The date dimension dim_date

All unique record_dates are extracted from the cohort_stock_summary table using the following calculation:

```
CALCULATETABLE(

DATESBETWEEN(

cohort_stock_summary[record_date],        MIN(cohort_stock_summary[record_date]),
MAX(cohort_stock_summary[record_date])

)

)
```

This fills in any missing dates between the very first and very last dates in the data

From here, calculations can be made about what day, week, month, quarter, year and financial year a date falls within. These can then be used to visualise data.

The size and weight dimensions dim_size and dim_weight

Each of these tables takes the unique categories for size and weight respectively and pulls it into a separate table. For now, just the category label is present but an expansion of this can be planned to accommodate varying definitions from individual farms.

Relationships

Generally the model entity relationships reflect those defined above in the source database. Typically one-way relationships have been setup between the tables, where higher level tables filter lower level tables (i.e. filtering for a weight category will filter the records that are associated with that weight category as well, but filtering for a record will not filter out weight categories visually). Two-way relationships are present where information needs to be kept in lock-step - companies to sites to records are one such example. The other place a two-way relationship is defined is between the core cohort_stock_summary table and the date dimension. This ensured that a farm was able to access information that related to the time period they were operating within; Information outside this period will not be shown to the farm as it is not relevant to them.

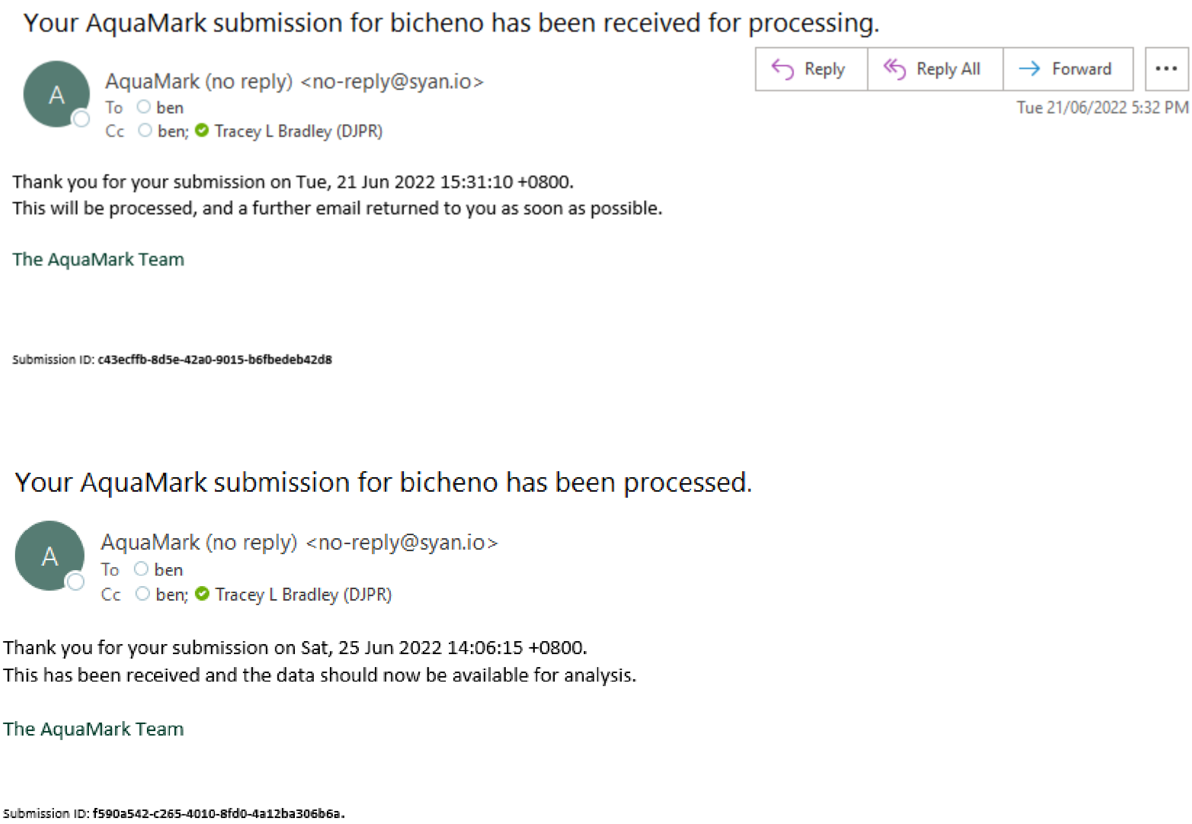
The Power BI output

The first step in creating the individual dashboard was to create an electronic map of the farm located accurately in space as per the GPS coordinates of the farm. This was an extremely time-consuming manual process but allowed each farm to interact with their data related to a group of tanks or an individual tank.

4. Farmer provision of information

Data was sought from the farms as spreadsheets via email. A system where the files can be sent directly and securely to the to the Ausvet cloud service has been developed as per below. This triggers an email acknowledging receipt and then when processed another email (Figure 20).

```
"888.bremer@syan.io",
"som.som@syan.io",
"yumbah.bicheno@syan.io",
"yumbah.ki@syan.io",
"yumbah.narrawong@syan.io",
"yumbah.pl@syan.io
```



Your AquaMark submission for bicheno has been received for processing.

AquaMark (no reply) <no-reply@syan.io>
To ben
Cc ben; Tracey L Bradley (DJPR)

Reply | Reply All | Forward | ...

Tue 21/06/2022 5:32 PM

Thank you for your submission on Tue, 21 Jun 2022 15:31:10 +0800.
This will be processed, and a further email returned to you as soon as possible.

The AquaMark Team

Submission ID: c43ecffb-8d5e-42a0-9015-b6fbedeb42d8

Your AquaMark submission for bicheno has been processed.

AquaMark (no reply) <no-reply@syan.io>
To ben
Cc ben; Tracey L Bradley (DJPR)

Thank you for your submission on Sat, 25 Jun 2022 14:06:15 +0800.
This has been received and the data should now be available for analysis.

The AquaMark Team

Submission ID: f590a542-c265-4010-8fd0-4a12ba306b6a.

Figure 20. Emails acknowledging receipt and processing of data

5. Accessing online

Farms have been provided with free, online access to their Power BI reports through a secure cloud-based Power BI workspace residing within the AgVic tenancy. Stakeholders are onboarded onto the system using their company email addresses – this is tied to their ongoing access to that email address which means the organisations have control over onboarding and offboarding of their staff. Furthermore, each user upon onboarding is required by security policy settings to set a multi-factor authentication method to further lock down their account.

Each user has been provided with access to see their shared reports and interact with them, but not work with the underlying data, download the report file, or make changes. Each of these reports are filtered row-wise to provide farm-level data for their own farm in the report, and summary statistics across periods of time for other farms as a comparison. As the new data is submitted and processed, the Power BI report is automatically updated and the farmer is able to visualise the new data.

The intuitive layout of the report ensures it can easily be shown to management, Boards, government authorities or any other relevant stakeholder.

Discussion

Discussion of results relative to the objectives:

1 Develop a farm data collection and review system for farmers to record health and production parameters

2 Provide a secure data storage system with the capacity to allow industry benchmarking against industry approved standards - for example the industry median.

The abalone Benchmarking Project is the first industry-wide benchmarking activity to be undertaken in Australian aquaculture. Larger companies across a range of species benchmark for health and production at a company level but the data is not shared externally across the industry. All 11 pump ashore abalone aquaculture facilities nationally enrolled in the project and provided data from (or before) 1/6/2021 to 30/5/2022. The involvement of all farmers in the project is a testament to the collaborative nature of the Australian abalone industry.

The project successfully deployed a secure data storage system through a 3rd party cloud service provider. The system has capacity to hold large volumes of farm health and production data. The data is accessible to the farms through highly secure Power BI reports where confidentiality is maintained using user-access permissions and row-level security. The farms were either enrolled in the more interactive dashboard with a suite of health and production parameters or had access to the overall benchmarked variables. Industry data was aggregated for the benchmarked variables ensuring that there was no possibility of identifying individual farm data.

Despite abalone aquaculture being a small and relatively homogenous industry with a large amount of data generated, many of the farms are not maximising the utility of the data. The creation of the dashboard "Hal" for 5 individual farms was an attempt to rectify this situation. "Hal" was deployed to 5 farms with farmer access to visualise and display the data as they wished. This should assist in interpretation of mortality events, optimal stocking rates and aid with on-farm operations.

Data quality issues were a major impediment to creating a robust benchmark for key parameters across the industry. Standard definitions such as year class of stock varied amongst farms. The manner of calculating various parameters such as growth and stocking rate within farms was surprisingly disparate. Considerable work was required to align these parameters to make them meaningful. There is a huge volume of data being generated by farms, and for those using the proprietary software, much of it is modelled data, not actual data. This software has suited industry for many years but inhibits the ability to benchmark actual, not modelled data across all 11 farms. The variation in attention to detail in data collection was apparent with some farms meticulously collecting data and others not. This further affected the robustness of the national benchmark. There is clear evidence that for some farms, the project accelerated progress towards improvements in data quality. By collecting millions of records across one year of data, interrogating the data and providing feedback to farms, improvements in data collection systems could be instituted where there was an appetite from the farm.

One company processes data direct to the cloud for real time analysis and subsequent action. In a capital-intensive industry where small, timely changes in practices can heavily impact affect farm productivity, this innovation seems the next logical step for the whole industry.

Following a farmer meeting, the addition of several new variables were of interest. These included the following:

1. Feed consumed per production (as FCR) as a one off. They would send the annual figures in (easier given it is EOFY).

2. Stocking Rate graphic change to % mort (instead of totals) – this may be unachievable.

3. Gram/month post spawning compared to individual farm modelling. The nature of the data will need to be assessed.

4. Water temperatures and how that relates to morts and growth.

The enhancement of the current project with this additional data is appropriate for another phase of the project dependent on farmer enthusiasm.

Conclusion

The Benchmarking Project has provided all the Australian abalone farms with the opportunity to utilise a purpose-built health and production monitoring dashboard for their own farm and the ability to compare key parameters across all farms. The project is farmer-driven and provides feedback in a totally secure and confidential environment.

All 11 farms contributed data to the project, with 5 electing to participate in the individualised data dashboard (Hal) with electronic farm tank maps.

Despite the abalone industry being small and intensive, there are clear points of difference in how measurements are made and recorded which complicated creating a robust benchmark for health and productivity parameters across all farms.

Within the larger corporations, individual farm benchmarking currently occurs and provides clear benefits. It is believed that this project has already provided interesting insights for farms around their practices. In a collaborative and open environment, the practices of leader farms should assist all in improving their production.

"A rising tide lifts all boats". (John F Kennedy, 1963)

Implications

The Benchmarking Project has provided all the Australian abalone farms with the opportunity to utilise a purpose-built health and production monitoring dashboard for their own farm and the ability to compare key parameters across all farms. The project is farmer-driven and provides feedback in a totally secure and confidential environment.

During the project issues related to data quality were uncovered and identified and rectified where possible. Technical aspects of the project would have been much simpler if a single data recording system were used across the industry. However, the commitment was to allow farmer to use existing systems and not require double entry of data. To that end, most farms rely on retrospective analysis of data, do not necessarily have secure systems for data storage and in many cases are working off modelled rather than real data. A minority of the industry (one company) have already switched to cloud-based systems, and this will be the way of the future. A simple, purpose built, industry - specific data recording system could be developed, however this would require cooperation across the whole industry and this will not happen in the current competitive climate.

Aside from illustrating data in an intuitive format, the dashboard (and benchmark) will be useful to demonstrate health and production parameters to government, Boards of management and other interested 3rd parties. The benchmark can be used by the industry as a whole to support trade where issues related to health and production may have arisen. Aspects of the dashboard are designed to assist with farm management for staff moving around tanks with a tablet.

The project has a legacy component: provided the farms continue to contribute data, the system will be maintained.

Recommendations

This project has been developed in close communication with the abalone farmers. Changes in management on some farms and larger companies have hampered the ability for the project to continue for some participants.

Provided there are a minimum of 5 interested farms, it is hoped that the benchmarking component will continue and the individual dashboards will be maintained.

Some participants have signalled that they will use the concept from this project in developing their own recording systems.

Further development

In the future, the benchmark could potentially be refined to capture a different suite of variables as an additional project if the farmers are motivated.

Ongoing reporting of descriptive statistics from the benchmark to be used in both internal and international negotiations could continue given Agriculture Victoria oversight of the data, with the approval of industry participants.

Extension and Adoption

The project has been completely created for use by the industry. All relevant industry members have had secure access to the PBI dashboard. The dashboard was demonstrated at the AAGA meeting in August 2022 in Portland. It is hoped that the legacy aspect of the PBI dashboard ensures that farmers use the product into the future.

Where farms have elected to not continue with providing data to the project for their own reasons, the benchmarked dataset will be reduced. Communications with industry about the number of farms contributing data will be maintained.

Project materials developed

The project has developed the Aquamark benchmarking dashboard that is freely available to all industry members and has no intellectual property constraints. The benchmarking dashboard is intended to be used by industry into the future as data continues to be provided to the secure, cloud database. An image from the front page of one of the participating farms is displayed below.

Appendices

References – footnotes/references/cross-references

ABARES, 2022 https://www.awe.gov.au/abares/research-topics/fisheries/fisheries-economics/fisheries-forecasts#abalone Outlook for Australian Fisheries and Aquaculture. Accessed 24/6/2022.

Bradley, TL, McCowan, C., Cohen, S., Green, C., Mansell, P. (2014) *Improved fish health management for integrated inland aquaculture through Better Management Practices (BMPs).* Final Report for FRDC project 2010 – 036. Department of Primary Industries, Melbourne.

Dairy Australia (2019). The dairy farm monitor project, Victoria Annual Report 2018 – 2019. Available at https://agriculture.vic.gov.au/about/agriculture-in-victoria/dairy-farm-monitor-project

Department of Agriculture, 2014, AQUAPLAN 2014-2019: Australia's National Strategic Plan for Aquatic Animal Health, Canberra. https://www.awe.gov.au/sites/default/files/sitecollectiondocuments/animal-plant-health/aquatic/aquaplan-2014-2019.pdf. Accessed 24/6/2022

Skipper – Horton, JO (2013). *Benchmarking growth performance and feed efficiency of commercial rainbow trout farms in Ontario, Canada*. [Master's thesis]. University of Guelph

Appendix 1. Confidentiality Agreement

Data and Information Consent and Confidentiality Form - Abalone Benchmarking Project

Purpose:

The purpose of this project is to provide farms with access to a highly secure system for managing data. The national benchmarking aspect of the project will provide farmers with the ability to assess their farm data against agreed - to aggregated industry parameters. This will be beneficial at an industry and farm level to provide supportive evidence during stressful periods such as disease outbreaks or loss of trade where a rapid response is required. The project will assist with trade access to markets where disease freedom and biosecurity programs form part of the pre-requisite conditions to undertake trade.

How the information will be used:

Identifiable participant data – will only be available to participants that provided that data.

Aggregated and deidentified participant data – to assist with establishing and providing industry benchmarks, participants can see how they compare to the industry mean/median and other measures as required.

Individual participants will not be able to see another individual participant's data.

Participants will have the choice to transfer data and information using a secured and encrypted channel (encrypted HTTP, API or direct database connection) or unsecured channel (email). Once data has been collected by DJPR, it is protected physically and electronically. A web accessible front-end will be made available for participants to interact with their data directly.

Disclosure:

DJPR will protect and manage all data and information provided by participants in accordance with the Privacy and Data Protection Act 2014 (Victoria), the Privacy Act 1988 (Cth) and other internal DJPR ICT and privacy policies that apply to the collection, management and use of any data and infomation.

Consent and Acknowledgement:

By signing this document you provide consent to any data or information you provide being collected, managed and used by DJPR for and in connection with the Purpose.

You warrant that to the best of your knowledge, all data and information you provide is accurate, complete and error free.

DJPR will take all reasonable precautions to protect the security and integrity of your data and information. To the extent permitted by law, DJPR is not liable for any loss, liability or expense arising out of or in connection with the collection, management and use of any data and information.

Name of authorised signatory of organisation ____

Organisation____

Signature____

Date_______________

FRDC FINAL REPORT CHECKLIST

The final report checklist can now be filled in when submitting your final report deliverable in FishNet.