FINAL PROJECT REPORT EVALUATION FORM

Project Number: SARF089	Completion Date: 31 Jan 2016
Project Title: A statistical analysis of sea-lice medicine use and benthic monitoring at Scottish marine salmon farms (2002 – 2014).	
1. In your view have the scientific objectives been achieved. If not, does this need to be addressed by SARF? Incorrect statistical models have been used so the objectives have not been achieved. The authors do not realize that the most important information from a study is extracted from interactions among fixed effects variables. They did not consider any interactions in their modelling. Their models also did not account for the correlations among the responses measured on a farm over time. Their use of site is not appropriate.	
2. Comment on the overall results of the project, including their significance for SARF. The overall results are in question since the models used to base the conclusions on incorrect.	
3. Is there a need for further work? If so, explain. The data need to be reanalysed using appropriate models, mainly models the include interactions of fixed effects, appropriate random effects, and that accounts for the correlations among the repeated measures made on a farm.	
Overall marking	1 - outstanding results 2 - results significantly above expectation 3 - satisfactory results 4 - results below expectation 5 - poor results
Signature	Print Name REF02 Date 31/01/2016

Please indicate whether you wish to receive payment (Yes/No) YES

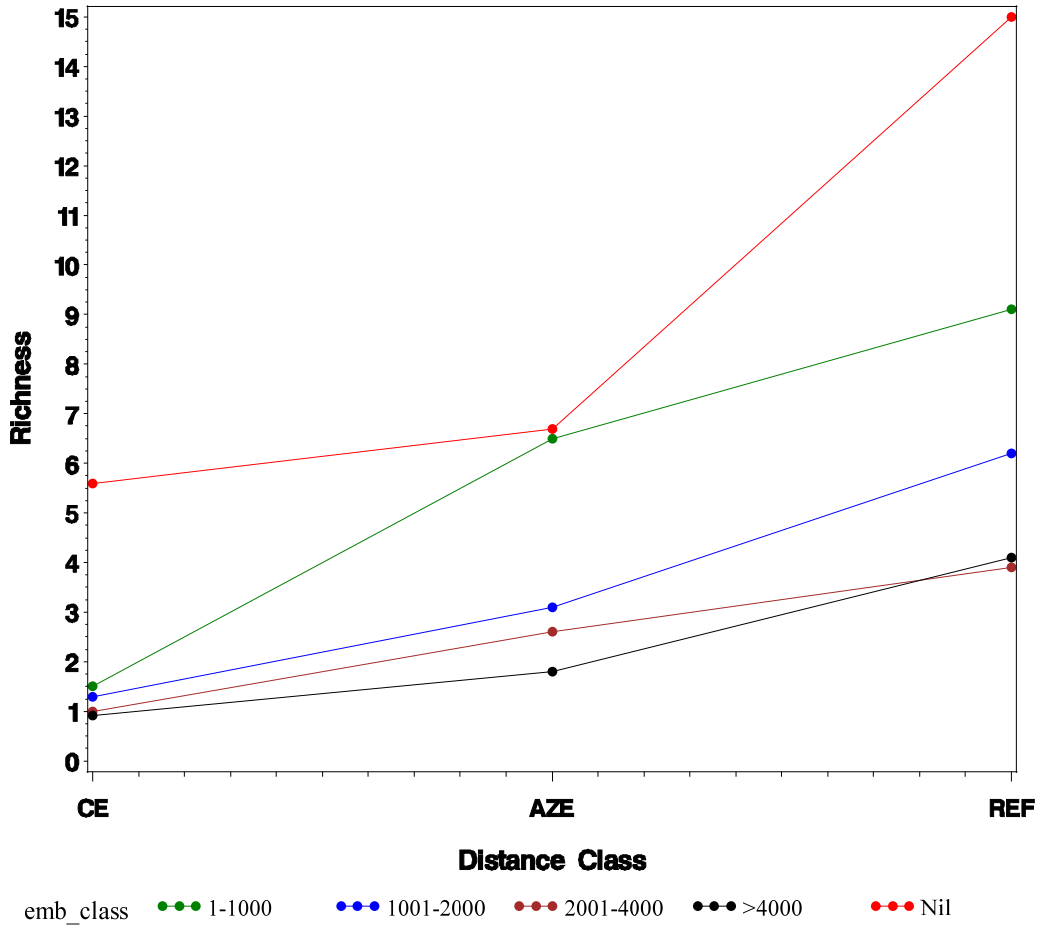
Additional Comments from REF02:
This is an observational study so any conclusions cannot be accessed as cause and effect. Cause and effect can only be assessed via properly designed experiment that is controlled during the execution.

This is a very good study with lots of structure and nicely filtered data. The study involves the collection of data from farms at various sampling times and at three distances from the farm cage. The sampler would take a grab sample of a given size at the three distances at a specific sampling time. Over the study time three different grab sample sizes were used to collect data. The objective is to look at the effect of EMB on crustacean abundance and richness, as well as other response variables. The farms has characteristics measured such as depth of water at cage site, speed of current, type of sediment under the cage, etc. It is assumed that amount of EMB, site characteristics (water depth, speed of current, characteristics of the sediment, size of grab sample) and distance classes have an influence on richness and abundance.

A model can be constructed to describe the richness or abundance that takes into account the sites and their characteristics, distance classes and the amount of EMB applied during a production cycle (or other time period). There were 1259 samples from 99 sites. It is assumed the sites represent a population of sites so inferences can be extended to the population of sites. So sites are to be used as a random effect in the model. The fixed effects in the model consists of the distance classes, then amount of EMB, the site characteristics and possible interactions among the above.

The following are the results from Table 7 of the report which looks at crustacean richness as a function of distance class and EMB class. The following figure displays the richness vs. distance class for each of the EMB classes.

Table 7 Richness by Distance Class for each EMB Class



The lines for EMB class of 2001-4000 and >4000 are almost identical, so it looks like the richness is not affected with higher amounts of EMB. The other lines are not parallel, indicating the slopes of the lines depends on the amount of EMB, i.e., there is an interaction between the levels of EMB and the distance from the cage (distance class).

The next two figures display the richness vs amount of EMB (measured by center of EMB class) and vs the square root of the center of the EMB class for each of the distance classes.

Table 7 Richness by EMB Class for each Distance Class – – Center

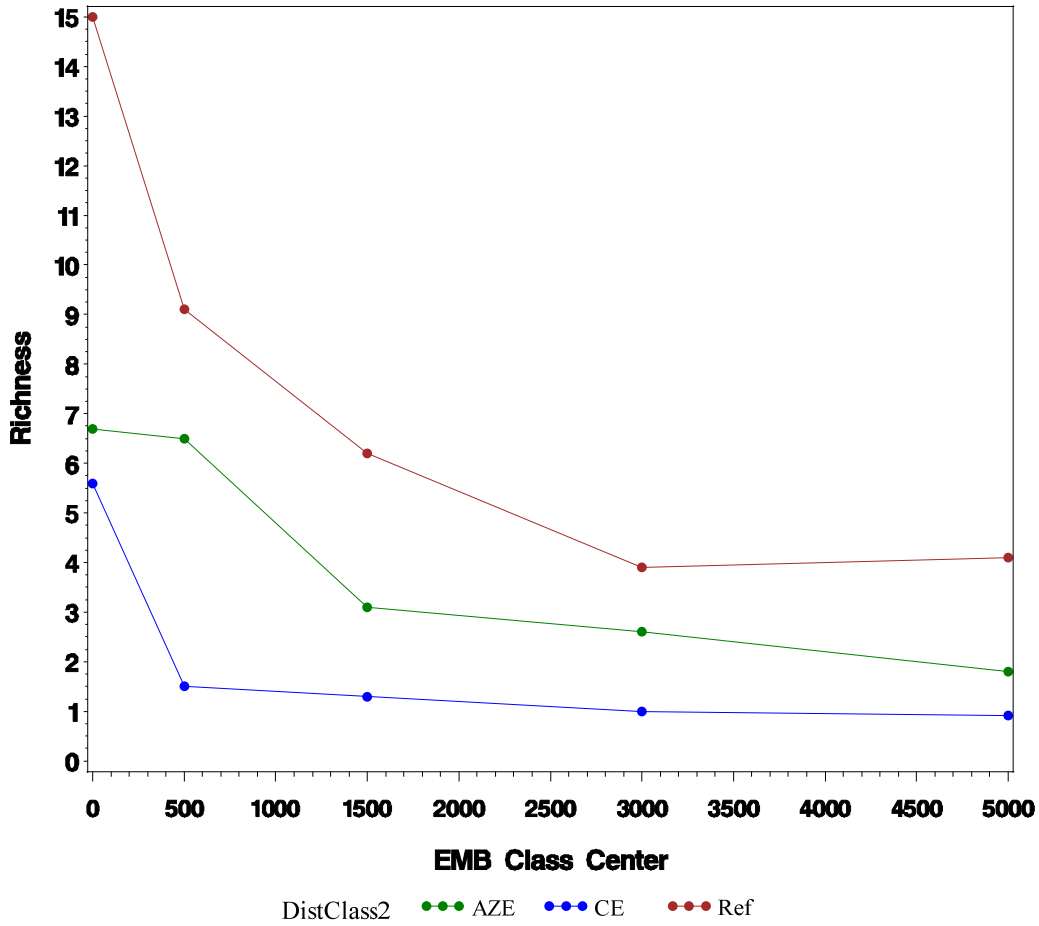
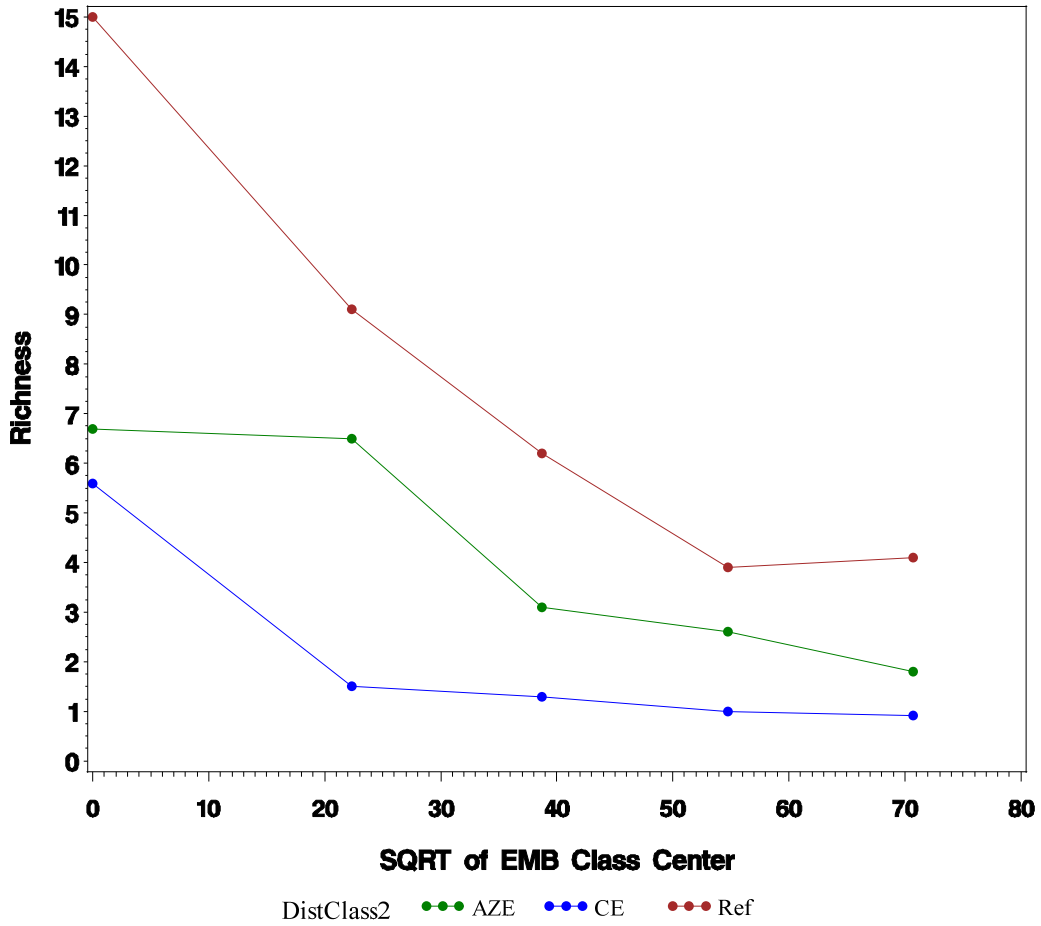


Table 7 Richness by EMB Class for each Distance Class – – SQRT of Center



Either graph shows the slopes of the lines are not equal across the distance classes, again showing there is an interaction among the distances classes and the amount of EMB.

Back to the construction of the model. The above tables indicate there is a strong possibility there is an interaction among the levels of distance class and the amount of EMB. There are not information in the table, but there are possible interactions among the levels of distance and the characteristics of the sites and interactions among the amount of EMB and the sites of the characteristics. There could be other interactions (like 3 way and higher), but there is a great possibility of 2 way interactions among the fixed effect variables. The starting model would have a structure with sites as random effects and fixed effects of distance classes, amount of EMB, site characteristics and all 2 way interactions among these fixed effects.

In response to Comment 31 the authors said (paraphrased) "the possible interactions are enormous, stating there are 9 fixed effects and the number of possible interactions is 2^9-10. And there are 2^{11} -12 possible interactions among the random effects. So it is to enormous to attack the interaction problem". First there are no possibility of interactions among the random effects as they are not cross classified. Second including up to 9 way interactions is not reasonable for describing a biological process, but the inclusion of 2 way and possibly 3 way interaction are very plausible. Instead the authors choose not to

look at any interactions at all. The above graphs show that there are high possibilities of interactions among some of the effects and they **MUST** be investigated, but they were not!

Using sites as a random effect is appropriate that allows for a wide inference across a population of similar sites. But when the sites are grouped as in table 3, called Local_A. When sites are grouped then the groups become fixed effects and sites nested within groups are the random effect. But they then made continued to construct just two groups of sites denoted by, OrkShet and Others and then considered them to be a random effect with 2 levels (with consent of SEPA). They did this to reduce the number of levels of site to help with their modeling. Any mixed software worth its salt can easily handel hundreds or even thousands of levels for a random effect. They should have used site as a random effect. If using OrkShet and Other as two groups of site (call them site_type), then two levels of site_type should be a fixed effect and sites nested within Site_Type should be the random effect. There is no indication that the authors used any process to decide the sites should be put into two groups. They have site characteristics such as depth, current speed, sediment sizes, etc that could be used as information to construct clusters using some type of cluster analysis.

The authors just looked at main effects models on which they based their conclusions. But by graphing the data in tables 6 and 7 (graphs above) you see the slopes are different across the amount of EMB (using EMB classes and square root of center of class) among the three distance classes. Their model need to account for these unequal slopes.

The EMB Class Nil the richness (Table 7 and First Figure) went from 5.6 (CE), 6.7 (AZE) to 15.0 (REF), a 2.7 fold change from CE to REF. So we expect around a 2.7 fold reduction from REF to CE without any EMB being applied. The changes in richness at rates of EMB (>0) need to be compared to the information at Nil, where no EMB has been applied. So the change in richness from 15 to 4.1 at Ref, a 3.7 fold reduction across the levels of EMB classes is not too different from the change of 2.7 fold across the distance classes for the Nil EMB class. So there is a reduction happening at the Nil EMB class that is not being accounted for in the overall discussion. So without EMB there is a 2.7 reduction across the distance classes and only a 3.7 reduction across the EMB use classes at REF. So at most 1 of the 3.7 fold reduction can be accounted by the use of EMB. Using interactions of distance class and amount of EMB would help extract proper information.

Each site was measured over time and total EMB is an accumulation over sampling times, so these measurements are correlated and a repeated measures analysis that accounts for this correlation within a site needs to be fit to the data. Without accounting for this correlation, the results are questionable.

When did they quit using TBZ?

What is the area of the farms compared to the area of possible farms? This is needed so we can make a judgment of the effect of EMB on richness and abundance. If the percent of area of farms compared to possible area of farms is 0.01%, then EMB would have a maximum effect on 0.01% of the richness and abundance. So what is the % area?

Did any of the farms use EMB for a few production cycles and then stop using EMB? If so one could study the recovery using this data.

On page 2, how does the amount Fiori (2012) used, 18 ng/L compare to the amounts used by the farmers? We cannot draw conclusions about EMB use on farms if we do not know the amount of EMB use.

In the extraction of residues, 3 samples were obtained and the Maximum EMB concentration was used as the data point. The median would be a much more representative value or if you do analysis on the maximum you need to do a quantile regression.

Their models do allow for unequal variances at the distance classes, which is great.

In Tables 9 and 10, EMB is next to the smallest effect (a little larger than depth) (Table 9), so there are many other effects in the model that are much more important than EMB on crustacean abundance. In Table 10, EMB is third from the bottom. Current, and distance classes are much more important. In table 12, EMB is next to the bottom of importance. In Table 13, EMB is least important. By including interactions in the modeling I am sure the conclusions would change.

Finally, it makes since to convert the grab sample measurements to a common unit. There are three sizes of grab samples, so the data need to be converted to a common size. The authors used grab sample size as a term in the fixed effects in the model, where it makes more since to convert the data from the grab samples to a common area/volume. By not converting to a common grab sample size is like comparing apples to oranges.

The modeling process could also benefit from using random coefficient models between sites and the fixed effects.