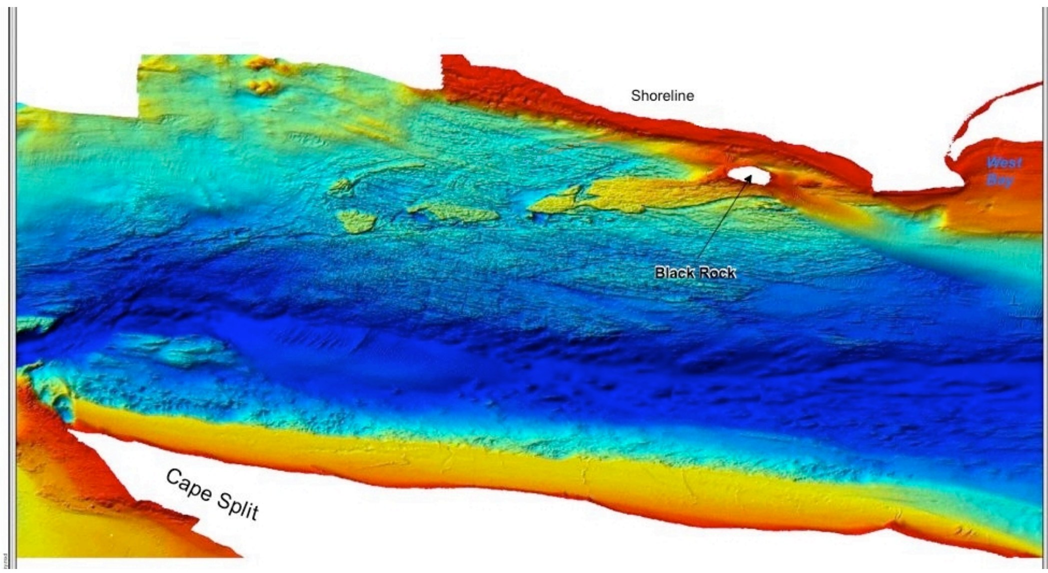
FUNDY TIDAL ENERGY DEMONSTRATION PROJECT

LOBSTER CATCH MONITORING

Summary of Results from Three Surveys with Recommendations for a Revised Survey Design

FINAL REPORT



Prepared for: **FORCE**
53 Prince Street
Hantsport, NS

Prepared by: CEF Consultants Ltd.
5885 Cunard St, Suite 801
Halifax, NS

Date: February, 2011

EXECUTIVE SUMMARY

Commercial fishing for lobster is important in Minas Channel, as well as most of the Bay of Fundy. Lobster fishing is one of the few commercial fishing operations likely to occur near the deployment area and an environmental effects monitoring (EEM) program has been implemented to assess potential changes in fishing success as result of construction and operation of the tidal energy program. The program consists of setting commercial lobster traps within test and control areas. Three surveys, two in the fall of 2009 and one in the spring of 2010, have been conducted to date. This EEM program is planned to continue as turbines are installed to provide input to the environmental management of the tidal energy program.

This report summarizes the key results from the 2009 and 2010 surveys and demonstrates the ability of the EEM program to identify significant effects. It includes an independent statistical review of results from the three surveys. The statistical review is incorporated into the comments on the significance of results observed by the monitoring, as well as recommendations for design improvements.

The primary difference between the two fall surveys conducted in 2009 was that the first survey took place prior to the commercial lobster fishery and the second survey took place while the fishery was active. Comparison of catch rates in the two fall surveys indicated that the catch of market-sized lobster was significantly lower ($P<0.0001$) during the commercial fishery (Fall Survey 2) than before (Fall Survey 1), documenting an effect from the commercial fishery.

The relationship of lobster size is an important variable affecting distributions of lobster, with smaller lobster generally found closer to shore in shallower water. Large lobster, greater than 110 mm carapace length, were less abundant but widely distributed with no statistical correlation with depth. A related finding was that the test area appears to be moderately fished commercially in comparison to a higher level of exploitation around Black Rock. The deeper ridges and banks, particularly in the Western Control Area, do not appear to be significantly fished.

A turbine was installed on November 12, 2009, part way through Fall Survey 2. Statistical comparison of the mean catch from traps 200-300 m from the turbine deployment site were significantly lower ($p=0.0002$) after turbine deployment. This comparison should be not taken to indicate an effect from turbine deployment because there were a number of differences in conditions, including water depth variations from traps being shifted by currents and effects of the commercial fishery, which could explain the differences in catch rates. However, the comparison does indicate the level at which differences can be detected in the EEM program conducted and that the program warrants continuation.

The key recommendation involves combining the two design elements into a single Before – After Control-Impact (BACI) design that will improve the efficiency of the survey, while providing a similar ability to detect potential impacts, both nearfield and farfield. The recommendations also suggest that the existing samples can be used for comparison with

future results. However, it would be advisable, since new sampling stations need to be randomly selected, that pre-deployment data be collected using the new design. A fall survey in 2011, prior to the start of the commercial fishing season, is the top priority, followed by a spring survey to better document movements in and out of the area.

Contents

List of Tables

List of Figures

List of Photos

1 BACKGROUND

1.1 Program Rationale

A pilot tidal energy program is being established in Minas Channel west of Black Rock near Parrsboro, Nova Scotia. Development of the program is being coordinated by the Fundy Ocean Research Centre for Energy (FORCE), an independent body comprising private industry, government and academia. FORCE's Board of Directors consists of technology developers testing with FORCE, one representative from the Province of Nova Scotia, and an independent academic from a relevant discipline appointed by the Province of Nova Scotia.

Different designs of tidal power generators are to be installed within a 1.6 km^2 crown lease area[1] at mid-tide depths of approximately 60 m, with power and communication lines running to shore north of Black Rock. Nova Scotia Power Inc. (NSPI) installed the first prototype tidal generator on November 12, 2009, and additional units are planned for installation by other companies within the same lease area in 2012.

Commercial fishing for lobster is important in this area, as well as most of the Bay of Fundy. Lobster fishing is one of the few commercial fishing operations likely to occur near the deployment area and an environmental effects monitoring (EEM) program has been implemented to assess potential changes in fishing success as result of construction and operation of the tidal energy program. The program consists of setting commercial lobster traps within test and control areas. Three surveys, two in the fall of 2009 and one in the spring of 2010, have been conducted to date. This EEM program is planned to continue as turbines are installed to provide input to the environmental management of the tidal energy program.

1.2 EMAC Review and Input

An Environmental Monitoring Advisory Committee (EMAC) has been established to provide independent expert scientific and traditional ecological knowledge advice to the Fundy Ocean Research Centre for Energy (FORCE) on the adequacy of the environmental effects monitoring programs related to the Fundy Tidal Energy Demonstration Project. EMAC is tasked with:

- providing a forum for review of environmental monitoring program results;
- reviewing EEM programs and making recommendations for improvement; and
- communicating advice to FORCE on EEM programs.

---

[1] The total crown lease area includes the test demonstration area of 1.6 km^2 plus a corridor area of 0.47 km^2 for cables to shore, for a total of 2.07 km^2.

One of the aims of EMAC is to encourage the EEM programs to be as cost-effective as possible while meeting their intended purpose. Direction and advice will be based on "adaptive management" principles, defined as: "an iterative process of planning and implementing an action, monitoring, evaluating and making adjustments as needed. Dr. Robert Miller, a member of EMAC, provided input to the statistical review appended to this report, but he remains independent of the main findings of this report, avoiding any conflict of interest.

This report is intended to provide EMAC with a summary of survey results and recommendations for improvement in survey design and implementation. This report summarizes previous reports focusing on key lessons learned. It also includes a review of the survey design with suggestions about how to make to make it more efficient, but also what kind of environmental change is likely detected by the program.

1.3 This Report

This report summarizes the key results from the 2009 and 2010 surveys and demonstrates the ability of the EEM program to identify significant effects. It includes an independent statistical review of results from the three surveys. The statistical review is incorporated into the comments on the significance of results observed by the monitoring, as well as recommendations for design improvements. A copy of the independent review is appended to this report. Readers are referred to previous data reports for more detailed portrayal of survey results. These surveys were carried out before and after deployment of the NSPI turbine in November of 2009, and thus provide some initial indication of potential effects from deployment and operation. When information is available prior to the potential impact, the design is often referred to as a Before – After Control-Impact (BACI) design (Smith 2002).

It is important to note that while approaches used in other similar EEM programs were utilized in the initial design of these surveys, the environment, especially the very strong bottom tidal currents, is unique and no similar study had been previously conducted in these conditions. Initial survey information collected has provided critical information necessary to evaluate the program design and to improve its efficiency through adaptive management.

2 PROGRAM DESIGN

2.1 Study Location

The crown lease development area and three initial proposed deployment sites are shown in Figure 2-1. The oval dark blue area on Figure 2-1 to the right of the crown lease deployment area contains Black Rock, a major feature of the coastal area. The town of Parrsboro is located approximately 15 km to the east of the study area. A fourth deployment berth has since been established.

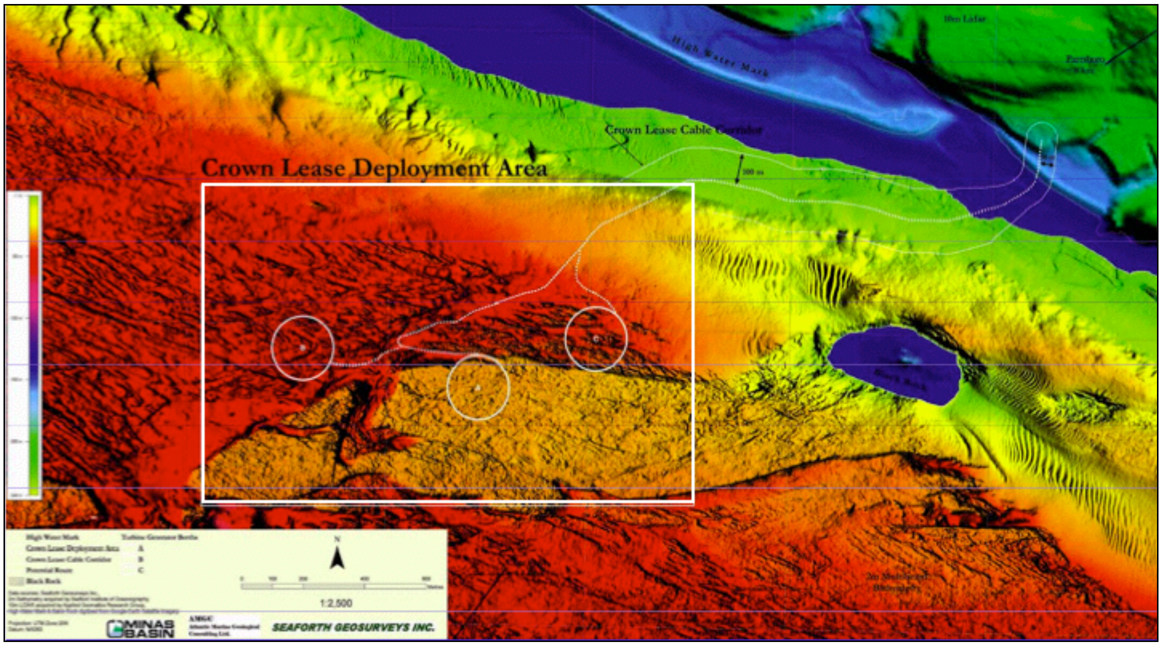


Figure 2-1: Crown Lease Development Area

In addition, a reference site for biological and physical monitoring has been established at 45° 21' 53" N, 64° 27' 32" W. Water depth at the site is 58 metres at mid tide.

2.2 Control and Test Comparisons

The design of the EEM program for lobster was based on measuring catchability within test and control areas. The crown lease deployment area was selected as the test area. Test Area refers to the area within which samples are considered affected by the treatment being monitored, i.e,. tidal power generator deployment and operation. Control areas were selected to the east and west of the test area. These area were selected because together they contained a range of lobster habitats similar to the test area, and they were assumed to be sufficiently far from the treatment that they would be unaffected by it.

The primary evaluation of effect was to be an Analysis of Variance comparing catchability within test and control area. Catchability was to be determined by deployment and retrieval of standard, baited, commercial lobster traps, similar to those used by local fishers. Test and control areas are illustrated in Figure 2-2.

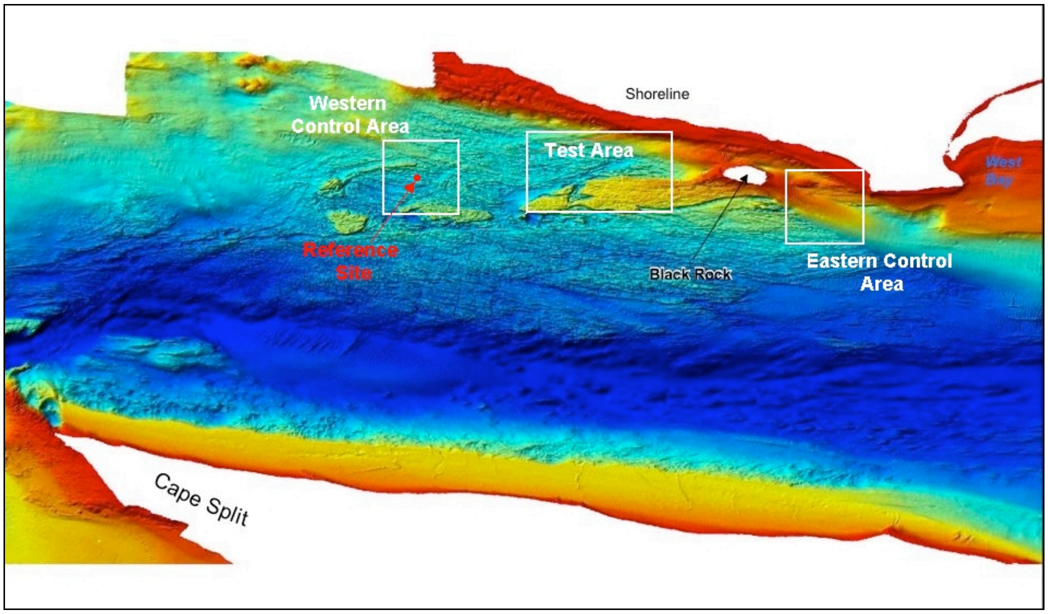


Figure 2-2: Location of Test and Control Areas

All stations were selected randomly and no stratification was used. Initially 25 stations were selected in the test area and 10 stations in each of the two control areas.

2.3 Nearfield Comparisons

NSPI requested specific nearfield monitoring around their turbine before and after deployment. A series of sampling stations were established in a circular grid at 200 m and 500 m from the proposed deployment site. Four stations were established at each distance in north, east, west and south directions from the deployment location. These stations were monitored along with the test and control stations.

2.4 Survey Equipment

Sixty standard commercial lobster traps and associated gear, including buoys, were purchased for use in this study. Traps were weighted with approximately 100 kg of concrete poured into the bottom of each trap. Traps were baited with shad and herring soaked in brine and escape vents were blocked to retain all sizes of lobster. All fishing was carried out at or near slack tide. Eight traps were equipped with thermometers to record bottom temperature.

Traps were set over slack and rising or falling tide whereas recovery of traps could only be done during slack high or low tide because it was only then that buoys were visible. During the first survey, two days were required to set all the traps. In subsequent surveys the number of traps was reduced allowing them to be set in one day.

All fishing was carried out by the fishing vessel Cape Rose (CFV 3089) from the Parrsboro wharf (Photo 2-1). Fishing was conducted under DFO Scientific Licence #324435 in 2009 and #324436 in 2010.

Photo 2-1: Fishing Vessel Cape Rose at the Passboro Wharf at Low Tide

3 RESULTS FALL 2009 AND SPRING 2010 SURVEYS

3.1 Comparison of Surveys

The conduct of the surveys changed over time, with a general reduction in the number of traps in an effort to provide more efficiency and consistency in sampling. The sampling effort and numbers of lobster caught for the three surveys are provided in Table 3-1.

Table 3-1: Number of Traps and Lobster Caught by Surveys, in 2009 and 2010

Survey	Date	Number of Traps Set	Number of Traps Recovered	Number of Trap Sets	Number of Lobster Caught
Fall One, 2009	September 25 - October 3	51	48	132	1387
Fall Two, 2009	November 5 - November 18	48	41	126	1135
Spring, 2010	May 10 - June 4	28	23	192	755

The primary difference between the two fall surveys conducted in 2009 was that the first survey took place prior to the commercial lobster fishery and the second survey took place while the fishery was active. In addition, during the second survey the number of stations was reduced and some traps fished in pairs separated by 60 m of line; buoys were also doubled with the second buoy attached as a trailer buoy to shorten the time required to find the traps. Between the first and second fall 2009 surveys, the number of stations was reduced from 33 to 18 in the test area, and from 20 to 9 in the control areas. As Table 3-1 indicates, these changes only amounted in a small change in the number of traps fished or the number of sets conducted because in many cases, two traps were fished at a single station.

The changes between the two fall surveys improved the efficiency of sampling to some degree, but the paired traps were not as successful as hoped. The movement of traps in areas of particularly high currents was still excessive, in one case being over one kilometer from set to retrieval. This magnitude of trap movement affects the interpretation of the data as well as affecting the time necessary to find and retrieve the traps. In addition, retrieval of the paired traps frequently resulted in safety concerns if the ropes became tangled or other problems arose during hauling. As a result of experience in the second fall survey, further modifications were made to the design of the spring survey.

In the spring of 2010 the number of stations and the number of traps fished in pairs were further reduced. This resulted in only 6 control stations being used and 14 test stations. The number of test stations was more than double the number of control stations because of the need to maintain both nearfield and farfield sampling within the test area. Note, however, as Table 3-1 indicates, that the number of trap sets actually increased because the sampling was more efficient. Generally speaking the objective was to enable recovery of all traps in a single day of fishing and this was accomplished most, but not all of the time.

3.2 Location of Stations

The locations of stations in the fall and spring surveys are illustrated in Figures 3-1 and 3-2 for the test area. Figure 3-2 also provides an indication of the movement of traps between set and retrieval. Note that the movement of traps in the spring survey was lower than in previous surveys largely because previous movement was used as a primary factor in selecting stations for use in the spring survey.

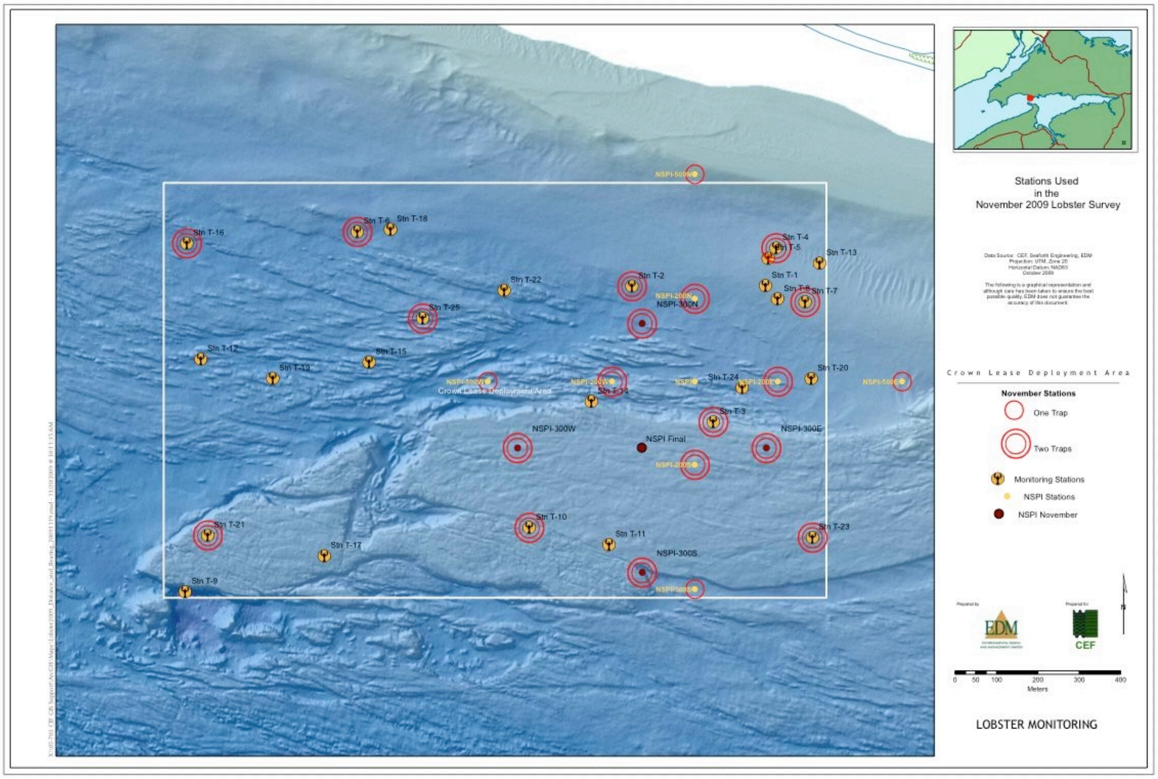


Figure 3-1: Sampling Stations within the Test Area in the Fall 2009 Surveys

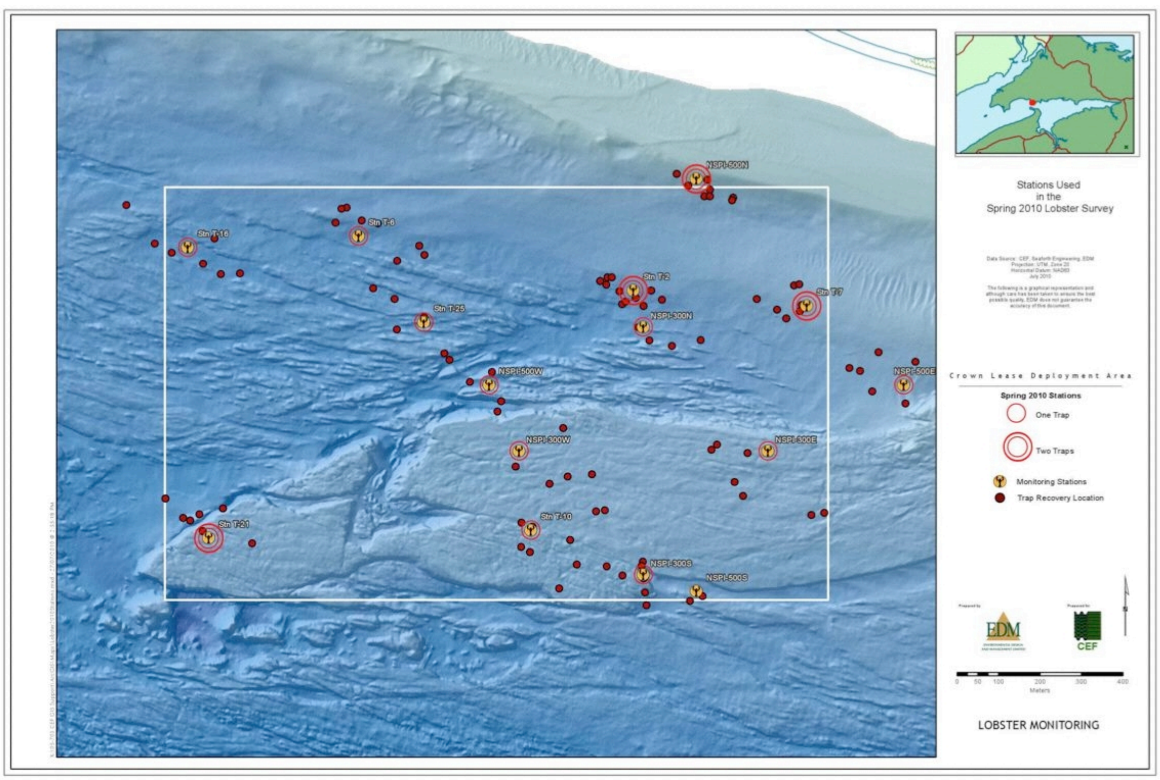


Figure 3-2: Sampling Stations within the Test Area in the Spring 2010 Survey

3.3 Differences between Spring and Fall Survey

Catch rates differed between the three surveys, with the largest difference between the spring and the two fall surveys (Table 3-2). Local fishers have said[2] lobster move out of the Upper Bay of Fundy in the fall and return in the early Spring. Fishing in the fall of 2009 was not carried on long enough to document a decline in catch rates because weather began to deteriorate in November. The survey in the spring of 2010 showed similar distributions of lobster by depth, but did not clearly identify increasing catch rates that would suggest an inmigration of lobster with time. Nonetheless, catch rates in the spring of 2010 were almost half of those in the fall of 2009.

Table 3-2: Catch Rates (Number of Lobster/Trap) in the Three Surveys

Survey	Catch Rate - Low Tide	Catch Rate - High Tide	Average Catch Rate
Fall One 2009	10.4	10.7	10.5
Fall Two 2009	11.2	7.1	9.0
Spring 2010	3.9	4.3	3.9

Lobster catch was divided into four size classes based on carapace length: < 66 mm; 66 to 82 mm; >82 to 110 mm; and > 110 mm. Lobster were considered market size above 82.5 mm carapace length. Statistical comparisons were done examining the relationship between catch rates, size class, water depth, and season using a generalized linear model (GLM). The GLM generalizes linear regression by allowing the linear model to be related to the response variable via a link function and by allowing the magnitude of the variance of each measurement to be a function of its predicted value. In this case, a negative binomial model linking size class and catch rates was used. All interactions between depth and season were statistically significant ($p<0.05$). In addition, the relationship between size class, depth and catch rates remained similar regardless of season.

3.4 Major Variables Affecting Catch

The three most important factors affecting catch rates relevant to effects monitoring are:

- soak time;
- carapace length; and
- water depth.

The distribution of lobster by sex and whether females were berried or not was examined but did not appear to be variable spatially.

---

[2] A project meeting was held with lobster fishermen early in design of the lobster monitoring program. In addition, local fishermen helped in implementation of the program and thus numerous conversations have taken place over the last two years.

3.4.1 Soak Time

Soak time refers to the length of time between setting and retrieval of a trap. Since the trap is baited and bait is usually consumed over a few days, soak time is a factor in trap efficiency. Soak time was considered to be an important potential factor in survey design, but the logistics involved in sampling were unknown. Trap movement from its set position, search time available, weather conditions, and number of traps fished all affected how long it takes between setting and retrieving a trap. Soak time was found to be a significant ($p<0.05$) variable in catch rates, with little interaction with other variables of interest (Bayley, 2010). The statistical review recommended that soak time be included in the analysis of potential project effects even though it would reduce the degrees of freedom in comparisons.

3.4.2 Carapace Length and Depth

Carapace length (CL) is the typical measurement for lobster size. Lobster size was found to be an important variable affecting distributions of lobster, with smaller lobster generally found closer to shore in shallower water. The distribution of lobster by size can affect the potential for impact on a particular size group because of the location of turbine deployment, as well as affecting the interpretation of results of catch rates. Figure 3-3 illustrates the different relationships between size class and water depth for the fall and spring surveys. Large lobster, greater than 110 mm CL, were less abundant but widely distributed with no statistical correlation with depth. The relationships are similar for each size class but the density is lower in the spring for all size groups.

Figures 3-4 and 3-5 illustrate the generalized distribution of lobster by size class from trap results of the fall and spring surveys based on the generalized linear models illustrated in Figure 3-3.

If there is movement in and out of Minas Channel with season, this movement does not appear to substantially affect the size distribution. This is an important finding of the surveys to date. Further surveys to investigate inter-annual variation would be helpful.

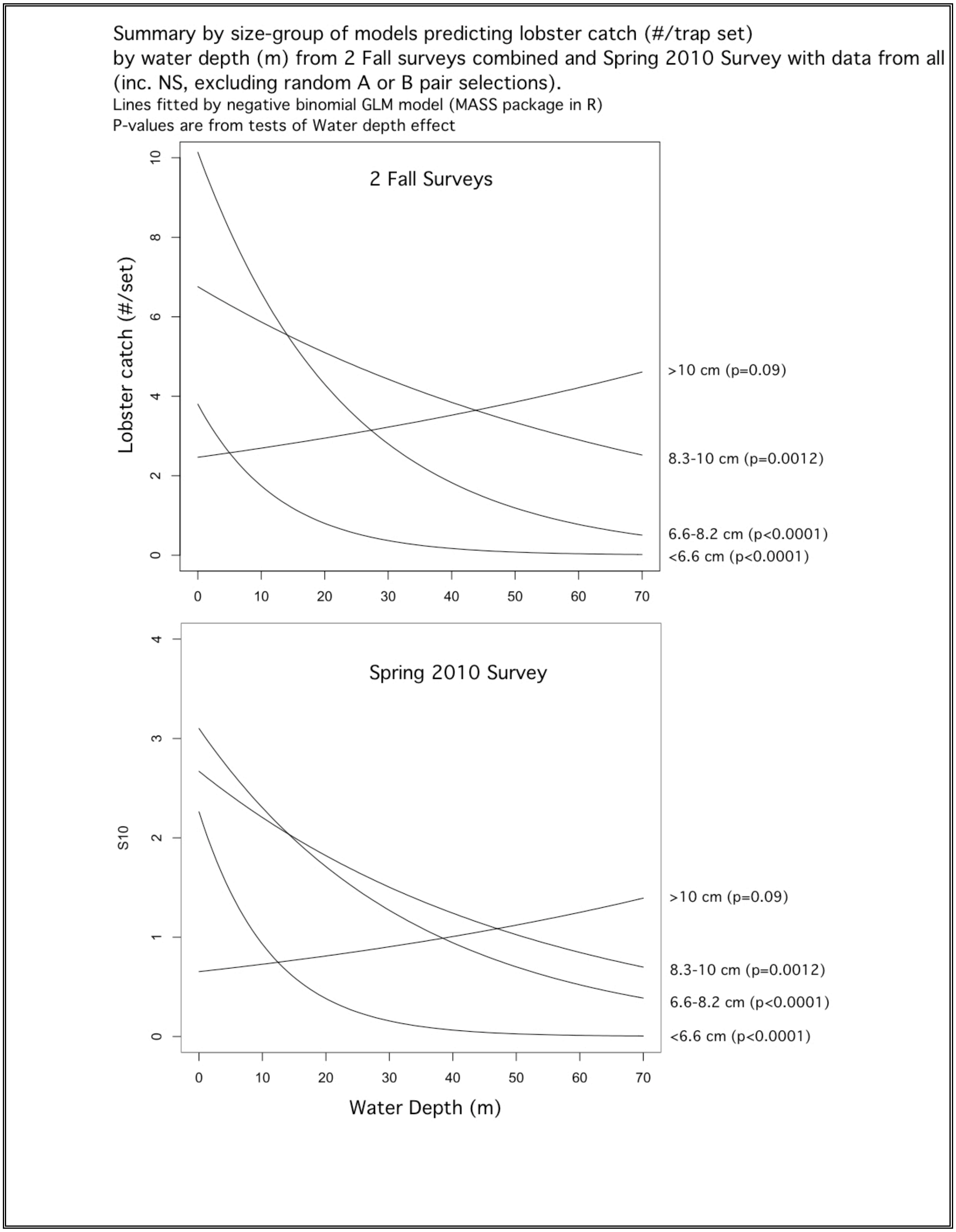


Figure 3-3: Relationships between Size Class and Water Depth by Survey

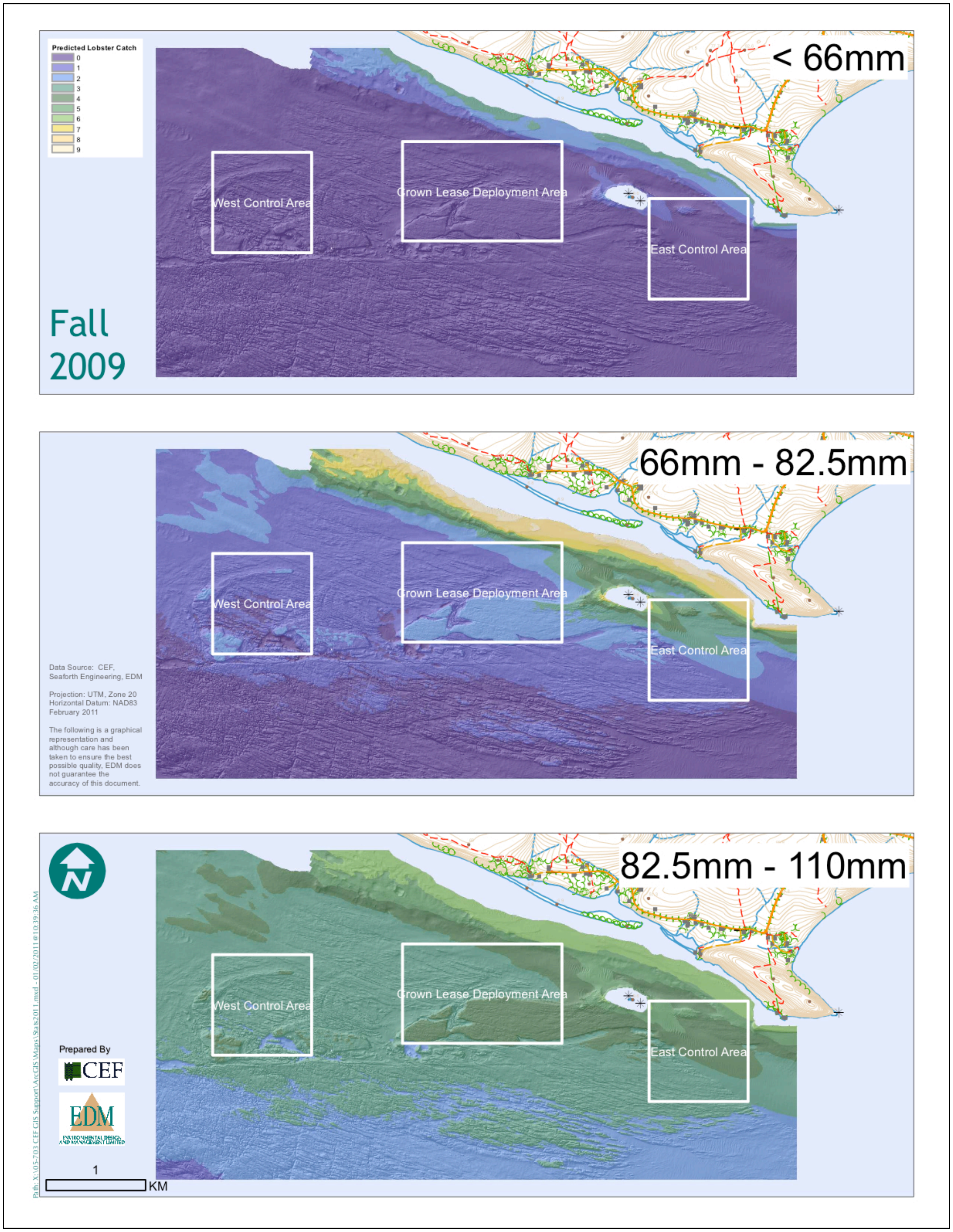


Figure 3-4: Generalized Density of Lobster Based on Fall 2009 Survey Results

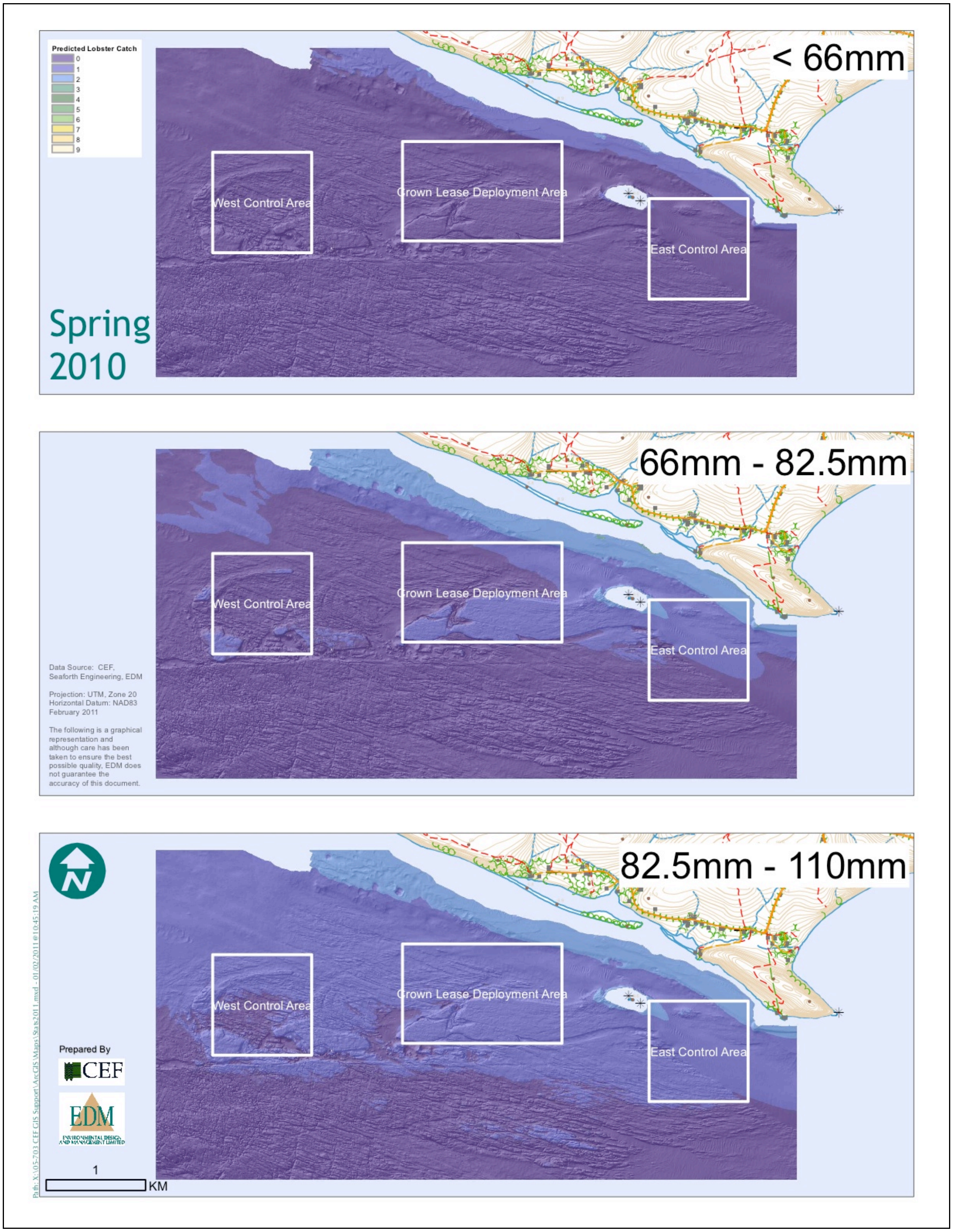


Figure 3-5: Generalized Density of Lobster Based on Spring 2010 Survey Results

3.5 Trap Movement

Information on trap movement was obtained because the recovery location of the trap was recorded by a portable GPS and the trap was always reset at the initial sample site. Movement of traps may provide useful information to current modelers as well as berth holders because it can provide indication of direction and strength of bottom currents.

Trap movement will continue to be an issue, but it appears to affect survey efficiency more than analytical confidence. The movement of traps between set and retrieval is largely a random process – traps do not move in a consistent manner from one day to the next.

In the first survey, traps were moved by high tidal currents an average of 116 m between hauls, with a maximum shift greater than one kilometer. Figure 3-6 representing trap hauls on September 28, 2009 provides an example of the type of movement observed between the setting and hauling of traps.

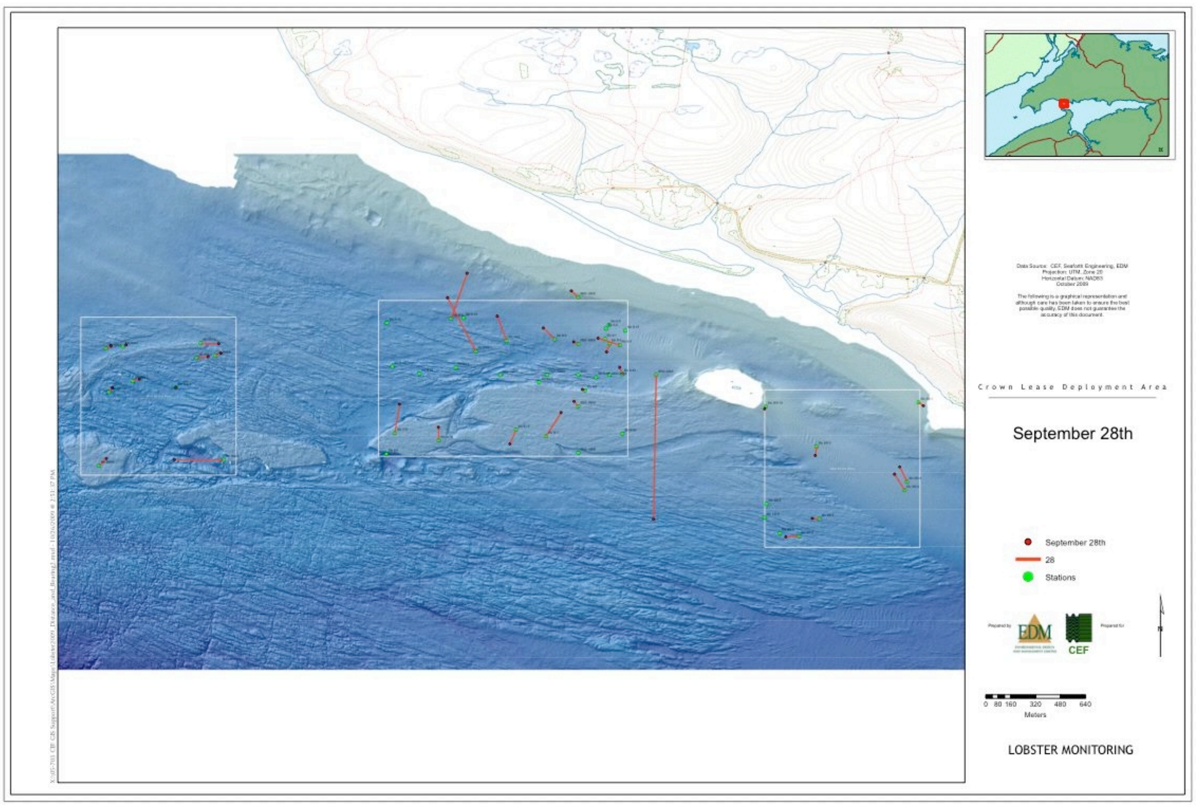


Figure 3-6: Location of Trap Haul Compared to Set Location for September 28, 2009

The pairing of traps in the second fall survey provided better replication in the sets and increased efficiency in the survey, but trap movement was similar with an average shift of 135 m between set and haul locations.

3.6 Effect of the Commercial Fishery

The two fall surveys took place before and during the commercial fishery. Analysis of the data was carried out to determine if the effect of the commercial fishery could be measured, and if so, was it the same in different areas. A generalized linear model using a negative binomial distribution was used to compare the catch rate of market-sized lobster in the Eastern Control Area between the two fall surveys. The comparison indicated that the catch of market-sized lobster was significantly lower ($P<0.0001$) during the commercial fishery (Fall Survey 2) than before (Fall Survey 1), documenting an effect from the commercial fishery in this area.

Similar comparisons were done with the proportion of market-sized lobster in the catch from the Test and Western Control Areas. Whereas the average catch of market-sized lobster dropped by 50% in Eastern Control Area, it only dropped by 29% in the Test Area and was roughly the same in both surveys in the Western Control Area. A comparison of areas indicated these differences were highly significant ($p<0.001$) with only a minor interaction between areas ($p=0.02$).

This comparison supported the general indications of fishing patterns by commercial fishermen in the area. The area around Black Rock (Eastern Control Area) was known to be a major fishing area and that was part of the reason for its selection as a control area. The Western Control Area was more difficult to fish and hence less frequently fished commercially. This comparison supports the value of the two control areas in the monitoring program, as a means to adjust results for the effect of the commercial fishery. It should be noted, however, that fishing patterns may change and continued monitoring will be required.

3.7 Effect of Turbine Deployment

A turbine was installed on November 12, 2009, part way through Fall Survey 2. The NS-samples, those in the grid around the turbine deployment site, provide samples from a before/after situation from close to (2-300 m) and further (about 500 m) from the turbine (Table 3-3).

Table 3-3: Catch Before and After Turbine Deployment from NSPI Stations

Distance from Turbine	200 - 300 m		500 m	
	Number of Samples	Average Catch/Trap	Number of Samples	Average Catch/Trap
Turbine absent	6	2.27	7	2.53
Turbine present	3	1.29	4	2.77

Statistical comparison of the mean catch from the 200-300 m traps was significantly lower ($p=0.0002$) after turbine deployment. In addition, average catch in three survey areas were not significantly different over the same time period (Bayley 2010).

This comparison should be not taken to indicate an effect from turbine deployment because there were a number of differences in conditions, including water depth variations from traps being shifted by currents, that could explain the differences in catch rates. In addtion, as noted in the detailed report on the two fall surveys, the largest drop in numbers was in market-sized lobster and this decrease occurred at the stations to the east, closest to the most intense commercial fishery (CEF 2010, page 32). However, the comparison does indicate the level at which differences can be detected in the EEM program conducted and that the program warrants continuation.

3.8 Survey Efficiency

Improvements were made in survey efficiency between each survey. Experience has indicated that approximately 15 stations can be sampled routinely in a day. More stations can be sampled at lower amplitude tides because search time is longer and traps remain closer to their set location. At extreme high tides, buoys may remain at the surface for less than 30 minutes at each slack tide, allowing recovery of relatively few traps.

Consistency and balance in the design has been improving, but the following recommendations in design changes will lead to further improvements in efficiency and should provide a stable design for future surveys.

4 PROPOSED DESIGN CHANGES

The statistical review carried out by Dr. Peter Bayler (see appendix) has been important in two key areas:

- program design, and
- analytical methods.

4.1 Combining Nearfield and Farfield Designs

The largest change proposed in survey design involves the integration of the nearfield and farfield designs. Integration of the two designs has a number of advantages including:

- increased efficiency in sampling, i.e., obtaining the maximum amount of information for the least effort;
- all deployment areas are treated consistently; and
- there is a single approach to analysis using analysis of variance to separate factors.

Figure 4-1 illustrates the proposed design in the test area for allocation of stations where both the nearfield and farfield components are combined. Sample sites are to be selected randomly but stratified within distance zones established around all deployment sites as evenly as possible.

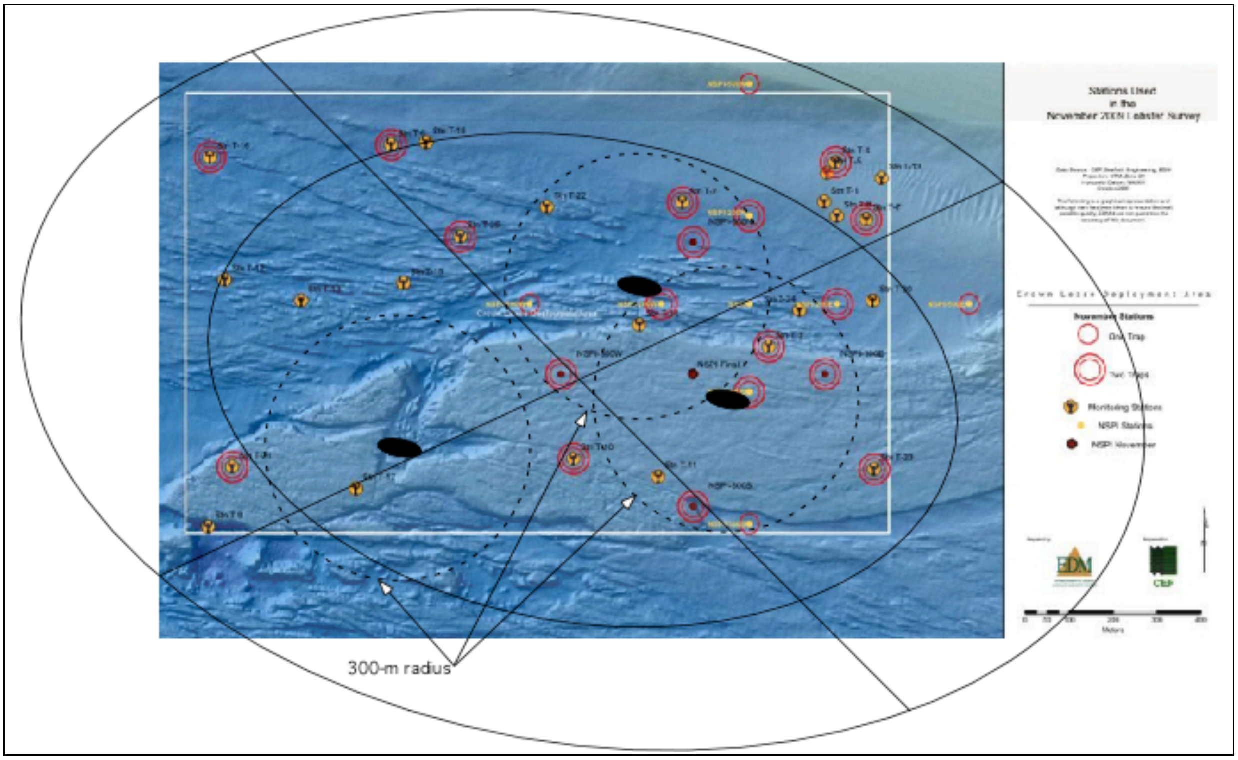


Figure 4-1: Proposed Sampling Design Combining Nearfield and Farfield Components

This proposed change in design requires new stations to be randomly selected. Since depth is known to be a significant variable, some stratification by depth can be introduced into the station selection to ensure that an adequate balance of depths are sampled. It is suggested that stations within the Test Area be selected first, and then stations within the two Control Areas be selected to provide a similar balance of depths within broad categories, such as shallow, intermediate and deep.

4.2 Statistical Comparisons

Surveys conducted to date provide information on variability of catch within the different areas, at different seasons, with and without the commercial fishery, and at different depths. Review of the data allows an estimate of the number of samples needed to detect a change in lobster catchability at a particular level, e.g., a pre-set change in percent of catch. In this case a change in catch of 2 lobster per trap was considered appropriate. After accounting for soak time, approximately 30 samples per group in a simple analysis of variance would be needed to detect a change of 2 lobster/trap-set from a mean of 6 /trap, if Type 1 error is 0.1 and Power is 0.75 (Bayley 2010).

This would translate into a survey with 24 stations divided among eight quadrats for the Test area, and 24 stations divided between the two control areas. Each of these replicated 3 times would provide 72 samples, roughly equivalent to the suggested degree of prediction above.

Analytical methods were also reviewed. Although tests have been done with the standard log(count+1) transformation to produce approximately normal distributions for linear models, future definitive tests, including simple BACI, would be best conducted using generalized linear models with a negative binomial distribution.

5 CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The three surveys, two in the fall of 2009 and one in the spring of 2010, documented the size distribution of lobster and its seasonal change within the proposed tidal energy development area. The study also found a reduction in market size lobster in areas fished more heavily by the commercial lobster fishery, hence demonstrating the kind of effect that can be detected. A related finding was that the test area appears to be moderately fished commercially in comparison to a higher level of exploitation around Black Rock. The deeper ridges and banks, particularly in the Western Control Area, do not appear to be significantly fished.

The abundance of lobster was found to be different in the spring than the fall, although the size distribution in relation to depth and proximity to shore remained similar for all size classes. Lobster smaller than market size (<82.5 mm) were found in shallower water and closer to shore than market lobster. Larger lobster (>100 mm) were less sensitive to depth and they were found widely distributed regardless of depth or proximity to shore.

A significant reduction in catch was observed 200 to 300 m from the NSPI turbine deployment site. This should not be considered to necessarily indicate an effect from the turbine, but that the study was able to detect changes relevant to an EEM program.

5.2 Recommendations

The key recommendation involves combining the two design elements into a single BACI design that will improve the efficiency of the survey, while providing a similar ability to detect potential impacts, both nearfield and farfield. The recommendations also suggest that the existing samples can be used for comparison with future results. However, it would be advisable, since new sampling stations need to be randomly selected, that pre-deployment data be collected using the new design.

The three surveys conducted to date have established important pre-treatment conditions of the spatial distribution of lobster by size, with some indication of seasonal variations. What remains needed is a measure of inter-annual variability prior to deployment of turbines or associated equipment. It will be important to conduct surveys in 2011 to adequately establish baseline conditions for the revised stations, as well as provide a measure of variability between years. A fall survey, prior to the start of the commercial fishing season, is the top priority, followed by a spring survey to better document movements in and out of the area.

6 REFERENCES

Bayley, P. 2010. Comments on the Lobster Monitoring Component of the Fundy Tidal Project. Courtesy Faculty, Dept. Fisheries & Wildlife, Oregon State University, Corvallis, OR 97331. December 23, 2010: 20p.

CEF Consultants Ltd. 2010. Fundy Tidal Energy Demonstration Project, Lobster Catch Monitoring – Analysis of Results from Two Fall Surveys: September 25-October 3 and November 5-18, 2009. Halifax, NS, March 8, 2010: 36p.

Smith, E.P. 2002. BACI Design. *In* El-Shaarawi, A.H. and Piegorsch, W.W. (eds.). Encyclopedia of Environmetrics. Volume 1. John Wiley & Sons, Ltd, Chichester, USA: 141-148.

Appendix A – Comments on the lobster monitoring component of the Fundy Tidal project

Peter B. Bayley revised December 23, 2010
Courtesy Faculty, Dept. Fisheries & Wildlife,
Oregon State University, Corvallis, OR 97331

Executive Summary

A. A simple BACI analysis based on limited samples at 2-300-m from the turbine indicated a significantly greater drop in mean catch following turbine installation compared to the control at ≈500 m. This result, and other considerations prompt the following questions:
1. Does the catch rate at ≈300 m continue at the observed depressed rate relative to ≈500 m during the extended presence of the turbine?
2. Will the catch rate change due to the disturbance of removal of the turbine?
3. If changed, did it recover after removal of the turbine?
4. Are there larger scale consequences of turbine presence?
A small scale, short term design restricted to within 500 m of the turbine addresses questions 1 and 2. A balanced design of 72 samples per survey accounts for distance and directional effects, with allowance for loss of samples and comparability with existing data. Questions 3 and 4 are addressed by larger scale seasonal survey that includes a continuation of the control and treatment strata of Design A sites, plus a similar number of random samples in Areas E and W. These designs are appropriate and with sufficient power for simple BACI tests.

B. Most existing samples are random or can be regarded as being approximately random, and can be used for future comparisons given comparable variances. An exception is the A-B paired samples that are strongly correlated, and produce a heterogeneous variance when combined with single samples. Paired samples should be discontinued, but random selections of one sample from each existing pair can be pooled with single samples for future analyses. The main weakness with existing data is the lack of balance among strata and replications at sites.

C. The soak-time (number of tides between setting and lifting traps) is an important component of fishing effort and significantly reduces bias and the mean square error, and thereby the efficiency of statistical tests. Further efforts to reduce soak-time variation are not recommended. The emphasis should be on maximizing balance by recovering as many traps as possible, even

if there is delay.

D. After accounting for soak time, approximately 30 samples per group in a simple ANOVA (see Glossary) would be needed to detect a change of 2 lobster/trap-set from a mean of 6 /trap, if Type 1 error =0.1 and Power =0.75.

E. Although tests have been done with the standard log(count+1) transformation to produce approximately normal distributions for linear models, future definitive tests, including simple BACI, would be best conducted using generalized linear models with a negative binomial distribution.

Response variables

Counts of lobster per trap-set comprise positive integers that are not normally distributed. In addition the variance is not independent of the mean. The solution is either (a) to adopt an approximate normalizing transformation or (b) apply a generalized linear model with a response appropriate to count data, such as the negative binomial. I recommend (b) as the best approach to definitive analyses in the future. However, for the purpose of illustrating issues with the data collected so far using the simplest linear analyses (ANOVAs, etc.) and for approximate estimates of samples needed, I here adopt (a) by using a standard log-transformation of Ln(count+1) as the response (y). If not stated otherwise, 'count' is the total number of lobster caught per trap-set.

Sample unit considerations

Because of trap movements, recovery rates, and limited recovery time at slack periods due to tidal currents in Survey 1, a portion of stations in Surveys 2 and 3 were sampled with pairs of traps (A & B) connected by a 60-m rope. These samples indicated a high Pearson correlation between A & B of 0.712 (df=96 [=98 A,B pairs], P= 2.2e-16. Lag 1 autocorrelation of all A,B samples was 0.54 compared to 0.37 for other samples, when the order of samples was the original collected order. Finally, the overall variance of A,B response data was 0.701. The remaining single-sample data pooled from the same Surveys (2 & 3) had a variance of 0.949. The variance ratio test indicates that these estimates are significantly different (p = 0.015, 2-tailed).

Therefore there is an issue of a mixture of variances which could compromise the interpretation of analyses [For linear statistical analyses, samples should at least be independent and there should be a single common error variance]. To illustrate the extent of this, I compared ANOVAs (analyses of variance) among surveys (2,3) and Areas (T, NS, EC, WC) between two sets of data. The first set comprised only A and B samples, the second of all other samples plus one random selection from each A,B pair. Therefore, the second set comprises relatively independent samples because they did not include A and B together.

Each survey had one or more samples from predetermined sampling stations at different dates. Therefore data were blocked by station(location) nested within each Area (see Area/Loc in model below).
Here are 2-way ANOVAs for each Set from Surveys 2 and 3:

Analysis of Variance Table for 1st Set (A & B samples only)
Response: LnCnt (=Ln("total lobster count" + 1)
model: lm(LnCnt~survf*Area+ Area/Loc) (Loc = Location = Station)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
survf	1	13.246	13.25	36.061	1.14e-08 ***
Area	3	15.357	5.119	13.935	3.67e-08 ***
survf:Area	3	16.563	5.521	15.030	1.02e-08 ***
Area:Loc	19	29.428	1.549	4.2164	1.49e-07 ***
Residuals	169	62.079	0.367		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
survf:Area = interaction between Survey(surv) and Area
Area:Loc = interaction representing nesting of Loc(station) in Area.

Analysis of Variance Table for 2nd Set (single samples)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
survf	1	36.809	36.809	86.534	2.85e-15 ***
Area	3	11.510	3.837	9.020	2.36e-05 ***
survf:Area	2	1.159	0.580	1.363	0.26
Area:Loc	13	21.973	1.690	3.974	3.14e-05 ***
Residuals	102	43.389	0.425		

[All analyses were done using R: (R Development Core Team (2009).]
While the difference between the survf:Area interaction significances is striking, the important issue is the difference between the Residual (error) mean square which was 0.425 for the independent samples (2nd Set) and 0.367 for the A,B paired samples (1st set). This is not primarily a spatial issue because station(Loc) was nested in Area and in both cases their contribution to sum of squares was considerable and highly significant. The residual error is biased downwards with the A,B pairs because of correlated pairs taken at the same time. Use of these data could result in false positives (i.e., thinking there is a significant difference where none exists) when interpreting tests. One could combine A & B into a single sample, but the degrees of freedom would be

halved and one could still not mix the data with samples from single traps.

I recommend that the A,B pair sampling be discontinued. Apparently there have been problems retrieving the paired traps and there is still trap movement and difficulties with retrieval. For the purpose of using existing data in future analyses, randomly selecting one sample from each A,B pair and pooling those samples with the single samples is recommended. I draw from this independent set (352 samples from the 3 Surveys) in the example analyses below (my particular random selection has been saved and is available for others to check on my results).

Soak Time

Soak Time is the number of tides between setting and retrieving traps. As CEF (LobsterFall Survey R#1BD60D.pdf) noted, as soak time increased, catch tended to increase until some maximum was passed when it started to decrease due to attrition of the catch. While the defined fishing effort is the trap-set, the amount of time it was set comprises an additional 'nuisance variable' (see Glossary). Because it appeared to be non-linear, it does not appear wise to alter the response variable by redefining catch-per-effort as catch per trap-set x tides. Rather, I first explored the data by introducing effort as the nuisance variables, soaktime and soaktime-squared, to simulate a unimodal (parabolic) effect on the standard response (Ln(count+1)).

Again, we look at ANOVAs that are likely to be used in future analyses in which Surveys and Areas are blocked, and station data blocks nested within Areas. This time the covariates soaktime and soaktime-squared are added (available for Surveys 2 and 3), making it an analysis of covariance (and I excluded NS-samples that are reported separately below):

lm(LnCnt~survf*Area+ Area/Loc +survf:soaktime + Area:soaktime + soaktime + I(soaktime^2))

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
survf	1	19.621	19.621	64.407	7.72e-13 ***
Area	2	21.470	10.735	35.238	9.13e-13 ***
soaktime	1	9.406	9.406	30.875	1.69e-07 ***
I(soaktime^2)	1	2.924	2.924	9.598	0.00243 **
survf:Area	2	3.148	1.574	5.167	0.00704 **
Area:Loc	17	25.385	1.493	4.902	7.01e-08 ***
survf:soaktime	1	1.348	1.348	4.426	0.0375 *
Area:soaktime	2	0.321	0.161	0.527	0.59
Residuals	120	36.557	0.3046		

Both soaktime effects were significant. There is a barely significant survf:soaktime interaction. A simplified version omitting these:

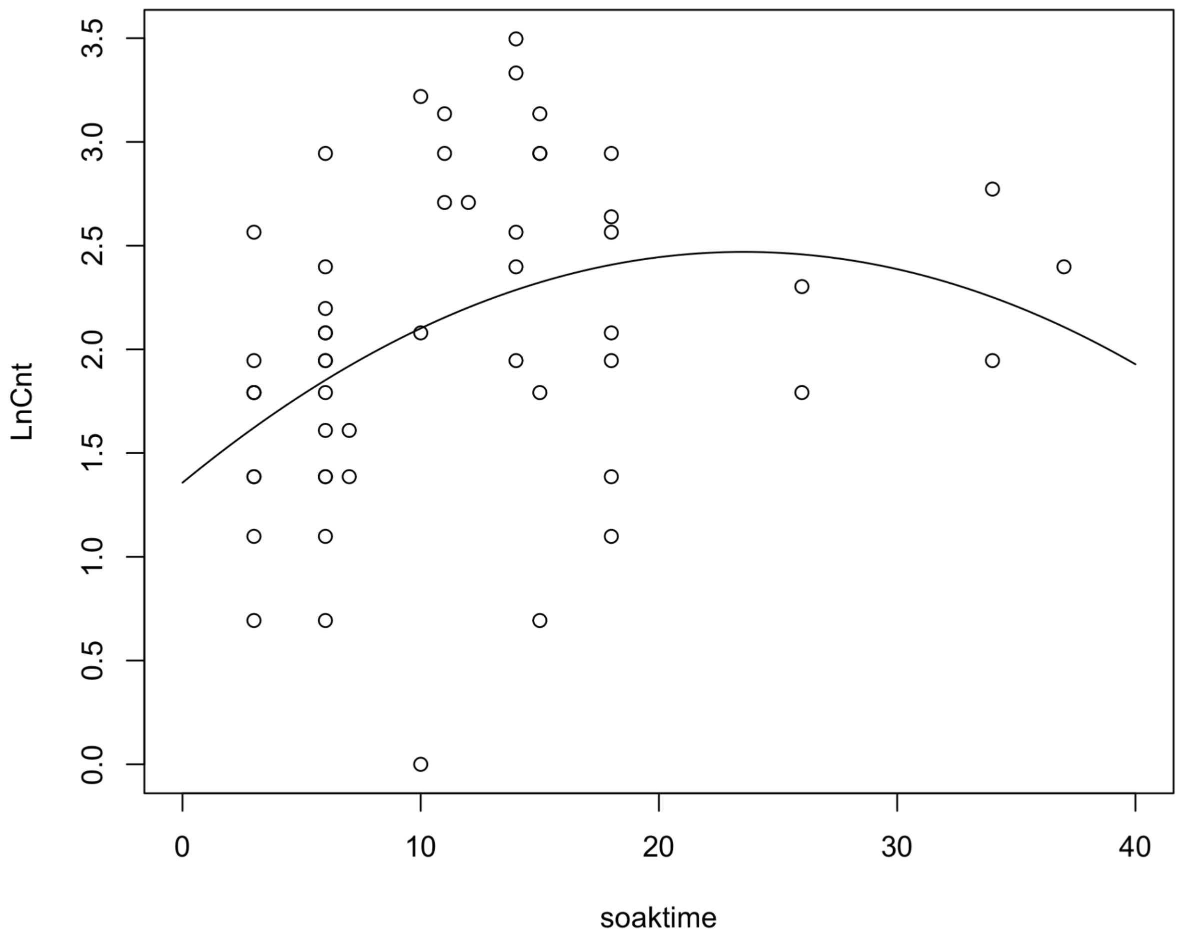
lm(LnCnt~survf*Area+ Area/Loc + soaktime + I(soaktime^2))

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
survf	1	19.621	19.621	63.133	1.05e-12 ***
Area	2	21.470	10.735	34.542	1.24e-12 ***
soaktime	1	9.406	9.406	30.264	2.09e-07 ***
I(soaktime^2)	1	2.924	2.924	9.408	0.00266 **
survf:Area	2	3.148	1.574	5.065	0.00770 **
Area:Loc	17	25.385	1.4932	4.805	9.44e-08 ***
Residuals	123	38.227	0.3108		

Soaktime and its square were highly significant, with signs (positive and negative, respectively) indicating a parabolic response. Apart from potential biases such as weather effects on soak time, the residual error (mean square) is reduced by 19% from 0.3822 without soaktime variables (analysis not shown) to 0.3108 shown above, reducing numbers of samples needed for a given power (see below).

A similar analysis was done with the ‘non-random’ NS samples only, in which there were no interactions with soaktime, and significant soaktime and soaktime^2 coefficients. In this case (analyses not shown) the residual error (mean square) was reduced by 21% from 0.4654 without soaktime variables to 0.3686.

An example is plotted below from Survey 2 Area data with the predicted parabola:



In future analyses soaktime and its square should be considered and included if their coefficients are significant. In the foregoing analyses it was worth the cost of 2 DF to account for soak time and its square. In trial analyses of some subsets I found that the coefficient for Soaktime^2 was not significant (but for soaktime was always significant and positive). Therefore, in future analyses tests for both variables may indicate inclusion of only soaktime at the cost of only 1 DF. Further reducing variation, or degrees of freedom lost, with this variable may be possible by avoiding long soak times, but in any case it is not recommended because spurious interactions may result when combining old data, such as from Survey 2, with new. The emphasis should be on maximizing balance by recovering as many traps as possible, even if there are some longer soak times due to delay.

Water Depth

Depth of sample is strongly related to Area, which is a major stratum in the design. Therefore, analyses that include Area and depth as main effects will be confounded. Mean depths for all samples are shown with corresponding mean Ln(Catch+1) (=LnCnt) by Area (including NS) and Survey are shown below:

	Survey 1		Survey 2		Survey 3	
Area/NS	Depth (n)	LnCnt	Depth (n)	LnCnt	Depth (n)	LnCnt
EC	-30.5(23)	2.82	-21.9(26)	2.19	-24.0(50)	1.87
NS	-36.1(24)	2.20	-34.6(29)	2.15	-31.1(61)	1.30
T	-37.9(57)	2.32	-39.6(56)	1.96	-40.0(71)	0.92
WC	-46.6(28)	1.85	-45.8(15)	1.75	-46.1(10)	0.54

(n = # of samples; Depths in Survey 2 were interpolated as means from samples of corresponding Stations in Survey 1 or 3)

A trend of decreasing mean catch with increasing depth is evident (and of similar significance as soaktime) in all Surveys, despite the occurrence of commercial fishing in Area EC during Survey 2 and varying soak times.

This relationship does not affect BACI-type relationships as long as control Areas along with a consistent sampling design are maintained for before/after comparisons. Because sampling can vary due to trap loss and tweaks in the design, consistency of mean depth in Areas should be checked, so that statistically blocking Areas accounts for depth differences.

Number of samples needed

The power of a test is the ability to detect a difference. It is the complement (1-beta) of the Type II error (beta), which is the probability of failing to detect a difference when one exists. Logically a Type I error (alpha) needs to be posited also, which is the probability of erroneously detecting a difference when none exists. In an ANOVA, the ability to detect a difference between means of groups (power = 1-beta), given alpha, depends on the residual error variance (= variance within groups) and the number of samples, n. Therefore, among the five entities that are interrelated any one can be predicted (albeit with some computing difficulty) from the other four.

Software was only available to predict n when comparing groups in a balanced 1-way ANOVA. In our case, I have presumed a more stringent error tolerance (alpha=0.05, beta = 0.20) and a less-stringent one (alpha=0.10, beta = 0.25). As a guide, an ANOVA is constructed in which a ‘control’ group (before turbine operation) provides a mean of 6 lobster/trap-set (≈equal to the mean from

Area T in Survey 2), while a 'treatment' group (during subsequent turbine operation) registers a mean loss of 2 lobster/trap-set. This 33% drop in raw figures translates to a 17% drop in terms of the transformed response, Ln(count+1). Several approximate predictions are shown here in which the residual mean squares are presumed from the foregoing, more complex ANOVAs:

Using Residual Mean Square of 0.3108 (from previous analysis including soaktime):
For alpha = 0.05 and power = 0.8 (beta = 0.2), n = 44 for each group
For alpha = 0.10 and power = 0.75 (beta = 0.25), n = 30 for each group

or, using Residual Mean Square of 0.3822 (ignoring soaktime):
For alpha = 0.05 and power = 0.8 (beta = 0.2), n = 54 for each group
For alpha = 0.10 and power = 0.75 (beta = 0.25), n = 37 for each group .

As an example in which error variance is about 0.31, if we accept a 1 in 20 chance of a false positive (alpha=0.05) and a 1 in 5 chance of failing to detect a difference (beta=0.2), we would need about 44 samples from a control and 44 from a treatment to detect a mean change of 2 lobsters /trap-set from a control of 6/trap-set.
The effect of a model with soaktime variables decreases the number of samples required by19% (44/54) for alpha=0.05 and beta=0.2.

An example of a turbine-effect test

A turbine was installed on November 12, 2009, part way through Survey 2. There is a before/after situation estimated by NS-samples close to (2-300 m) and further (about 500 m) from the turbine. The means(Ln(count+1)) and sample sizes (in parentheses) were as follows:

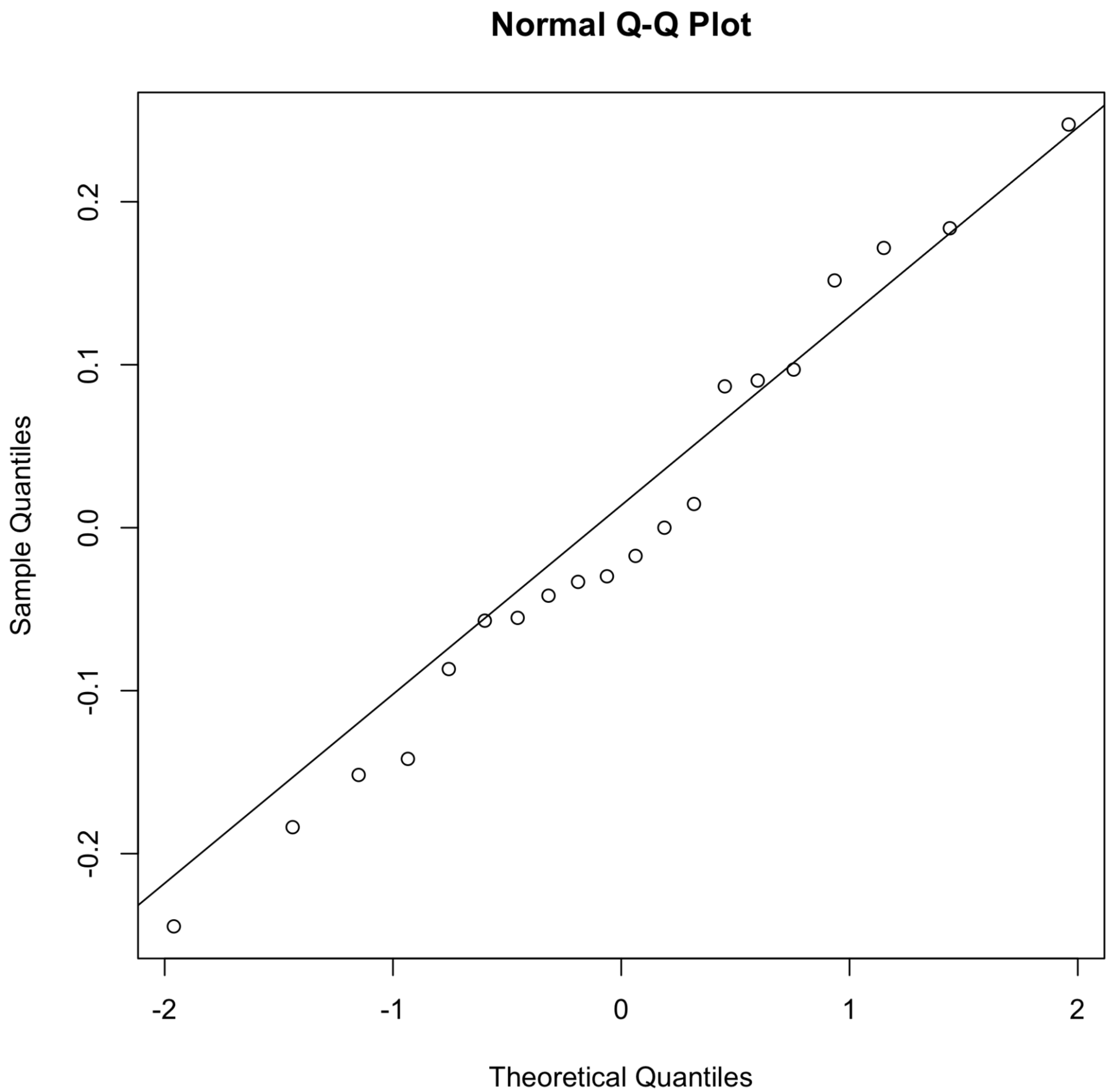
2/300m or 500m (NS23_5):	2/300m	500m
before Turbine (BT):	2.27(6)	2.53(7)
during Turbine (AT):	1.29(3)	2.77(4)

The analysis of covariance is as follows (there were no interactions with soaktime):

lm(LnCnt ~ NS23_5*BT_AT +NS23_5f/Loc + soaktime+ I(soaktime^2))

	Df	SumSq	Mean Sq	F value	Pr(>F)
NS23_5	1	2.2613	2.2613	55.26	7.4e-05 ***
BT_AT	1	0.3972	0.3972	9.71	0.014 *
soaktime	1	1.9774	1.9774	48.32	0.00012 ***
I(soaktime^2)	1	1.4045	1.4045	34.32	0.00038 ***
NS23_5:BT_AT	1	1.7079	1.7079	41.73	0.00020 ***
NS23_5:Loc	6	3.0316	0.5053	12.35	0.00115 **
Residuals	8	0.3274	0.0409		

The important observation here is the interaction (NS23_5:BT_AT, P=0.0002) between the two main effects, before/during turbine (BT_AT) and proximity (NS23_5) to turbine. This interaction indicated that the reduction during turbine presence was greater than any that might have occurred at the 500m distance. The following Normal Q-Q plot of residuals indicates limited departure from normality



The tendency for lower residuals than predicted at the lower left is partly a function of the arbitrary constant, 1, in the Ln(Count+1) transformation, because the deviation is reduced when 0.1 is used.

Given that this result was surprising to some, I have explored four avenues that may support or contradict its validity:

1. Recall that with each paired A/B sample one was randomly selected. With such small samples sizes in this analysis it is possible that the random selection resulted in markedly different estimates. The 2/300-m samples were paired (the 500-m not). The following shows the means for the A/B samples not

selected:

2/300m or 500m (NS23_5):	2/300m	
before Turbine (BT):	2.33(6)	
during Turbine (AT):	0.73(3)	

An analysis of covariance according to the foregoing model again indicated a significant interaction (NS23_5:BT_AT, P=0.0027) between the two main effects, before/during turbine (BT_AT) and proximity (NS23_5) to turbine.

2. Another potential source of bias may be due to the varying distances due to the turbine being placed 200 m SW of the original planned position. The following shows the mean distances of the samples from the actual location (followed by sample size and range):

2/300m or 500m (NS23_5):	2/300m	500m
before Turbine (BT):	260(6, 131-379)	537(7, 362-669)
during Turbine (AT):	223(3, 131-361)	586(4, 362-669)

An ANOVA based on these four means showed that there was no difference in turbine distance between before and after turbine (BT_AT) at either distance group (NS23_5, P=0.87) and no interaction (NS23_5:BT_AT, P=0.52).

3. As mentioned previously (Water Depth), mean water depth might have varied as a result of incomplete replication of all sites:

2/300m or 500m (NS23_5):	2/300m	500m
before Turbine (BT):	-36.9(6)	32.1(7)
during Turbine (AT):	-37.8(3)	27.3(4)

An ANOVA based on these four means showed that there was no difference in water depth between before and after turbine (BT_AT) at either distance group (P=0.58) and no interaction (NS23_5:BT_AT, P=0.50). Even the mean depth between distance groups (NS23_5) was not significant (P=0.10) even though the North 500-m sites were in shallower water.

4. Finally, the three Areas also did not indicate a reduction in mean catch during turbine presence:

Area:	T	EC	WC
before Turbine (BT)	1.92(17)	2.31(10)	1.69(5)
during Turbine (AT)	2.05(11)	2.30(7)	2.03(3)

In conclusion, this analysis indicates a significant (P<0.001) reduction of lobster catch rate at a distance of around 223-m from the turbine. The average magnitude of this drop was 43% in terms of Ln(catch+1) or 70% in terms of the catch (number of lobster per trap-set). It should be stressed that ‘after turbine' (AT) data were recorded within a 6-day period (Nov. 12-17) following turbine installation on Nov. 12. Moreover, it is reported that the turbine was not functioning. Therefore, it is possible that this localized effect was due to the disturbance associated with installation.

This analysis also illustrates how a simple BACI-type analysis (Green (1979); http://www.web-e.stat.vt.edu/vining/smith/B001-_o.pdf) might be applied in the future. It also raises questions leading to one of the proposed designs described below.

Design Considerations

The non-NS Stations were randomly positioned in the 3 Areas T, EC and WC, and varying portions of those locations were resampled and retained in subsequent Surveys. The following shows the numbers of samples (excluding one of each A,B pair) by Survey and Station:

	Station (Loc)								
Surv	EC1	EC10	EC3	EC4	EC5	EC6	EC7	EC8	EC9
1	3	3	3	2	3	2	3	2	2
2	4	4	5	0	0	0	3	0	1
3	8	8	8	0	0	0	8	0	5

	Station (Loc)																					
Surv	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T2	T20	T21	T22	T23	T24	T25	T3	T4	T6	T7	T8
1	3	3	1	3	1	1	3	3	3	3	3	3	3	3	2	3	3	3	1	3	3	3
2	3	0	0	0	0	0	3	0	0	0	2	0	1	0	3	0	2	2	4	3	4	1
3	7	0	0	0	0	0	7	0	0	0	7	0	8	0	0	0	6	0	0	8	8	0

	Station (Loc)									
Surv	WC1	WC10	WC2	WC3	WC4	WC5	WC6	WC7	WC8	WC9
1	3	3	2	3	2	3	3	3	3	3
2	0	0	1	0	0	0	0	3	2	2
3	0	0	0	0	0	0	0	7	0	0

One statistical approach is to pair treatment and control samples by date at fixed locations, allowing differences to be compared by site "before and after" as in a paired BACI scheme (Stewart-Oaten et al. 1986). However, this requires consistent sampling by specific locations and dates, which, as demonstrated by this table, is difficult given the trap losses, trap fishing limitations, and design changes. Conversely a simple BACI approach is easier to balance. However, while it is possible to account for the considerable variance among Stations, as I have done by nesting them in Areas in the forgoing analyses, there is clearly poor balance to the point of producing singularities that restrict the statistical options available. The designs described below attempt to minimize this problem.

Traps in the NS ‘fixed’ Stations moved because of the tides as did their non-NS counterparts, and their positions relative to the turbine changed when the latter was changed. Effectively, they could be regarded as approximately randomly located within larger areal boundaries centered at about 2-300 m and 500 m from the turbine. This permitted the simple BACI outlined in the previous section. In both sampling approaches there is also an element of randomness in the movement of the lobsters themselves. It is noteworthy how similar the results of analyses were between spatially random data from the 3 areas in which many Stations were not revisited, and the relatively fixed NS data (see Soak Time section). In both cases the change in residual mean square was similar with and without soak time variables. The levels of significance were also similar. This similarity, however, cannot be guaranteed, and better designs are recommended below.

Proposed Designs

The proposed changes are with the benefit of hindsight. None of the foregoing design issues could have been avoided without foreseeing the huge effects of tidal forces on trap movement and loss, and on the ability to fish the traps consistently, even when the weather is kind. The following outlines two connected designs at different scales that attempt to maximize balance and the utility of existing data, given trap movement and the risks of trap loss.

To date, we have a suggested, short-term indication of a turbine effect at a local scale, and no indication at a larger scale. Five questions come to mind (with hints of sampling required):

1. Does the catch rate at ≈300 m continue at the observed depressed rate relative to ≈500 m during the extended presence of the turbine? (sample while the turbine is still there);
2. Will the catch rate change due to the disturbance of removal of the turbine? (sample during and immediately after turbine removal);
3. If changed, did it recover after removal of the turbine? (sample at least one season after turbine removal);
4. Are there larger scale consequences of turbine presence? (sample at larger spatial scale consecutive fall and/or spring seasons).
5. How does one approach a multiple turbine scenario?

Relevant to all these questions is the status of the turbine when in place, information which is not available to date. Obviously it is important to know if the turbine is turning and if it is generating. NS Power should devise a method of monitoring the turbine while in situ.

Two related designs, one at a local scale (Design A) to address questions 1 and 2, and a large scale, Design B, to address 3 and 4 are outlined below. Both can incorporate existing data with the constraints outlined in this report. If Design A results are non-significant with sufficient power, funding for Design B may be considered unnecessary.

Design A

To answer questions 1 and 2, a local scale, short-term design should be limited to the turbine area. One approach would be to assign samples to two annuli (strata) around the turbine, at 300-350m (treatment) and 450-500m (control). Sample sites for temporal replication should be assigned randomly in each of four directional substrata (-45-45°; 45-135°; 135-225°; 225-315°) relative to the direction of the turbine (0°) in each stratum (see Figure). These substrata are to account for possible directional effects due to water currents and noise from the turbine. Some traps will undoubtedly drift outside their annulus, hence the 100-m separation between the primary strata. The separation also assures some contrast between strata and comparability with existing NS data.

Regarding numbers of samples, it is important to have back-ups to ensure as balanced a design as possible. Analyses will include nesting of site replications within substrata, and nesting of sites from the latter in the two strata (300-350m, 450-500m). Balance is important at all levels. A suggested design is 12 randomized stations in each stratum, (3 in each substratum -45-45, 45-135°, etc.) each repeated 3 times. Replicating 3 times at each location, and having 3 locations in each substratum would provide good insurance for single losses in locations or site replications, and still retain temporal and spatial replication at the substratum level. If all samples could be completed, the total samples per survey (either 'before' or 'after') would be 72 (=3x3x4x2), meaning 36 for each stratum, which provides good power for the main treatment/control effect. For comparison, the first 3 surveys, excluding half of the A/B pairs, achieved 132, 73, and 147 completed samples respectively. While mean traveling time would be reduced, total time available may be more limiting to answer questions 1, and 2. Whatever variant of this scheme is chosen, it is strongly recommended to not have less than 3 samples in each cell, because if one is not completed there would still the minimum of 2 required to provide random error estimates within cells.

A minimum distance between sites may be desirable to avoid physical interference and spatial correlation (see next Figure). In this case, re-randomizing locations for proposed sites to maintain this minimum would be acceptable. I believe that one spatial randomization procedure is acceptable to define sites for future surveys. The design cannot account for lobster possibly being attracted to the turbine as a protective structure, because of the 300-m safety zone for trap placement.

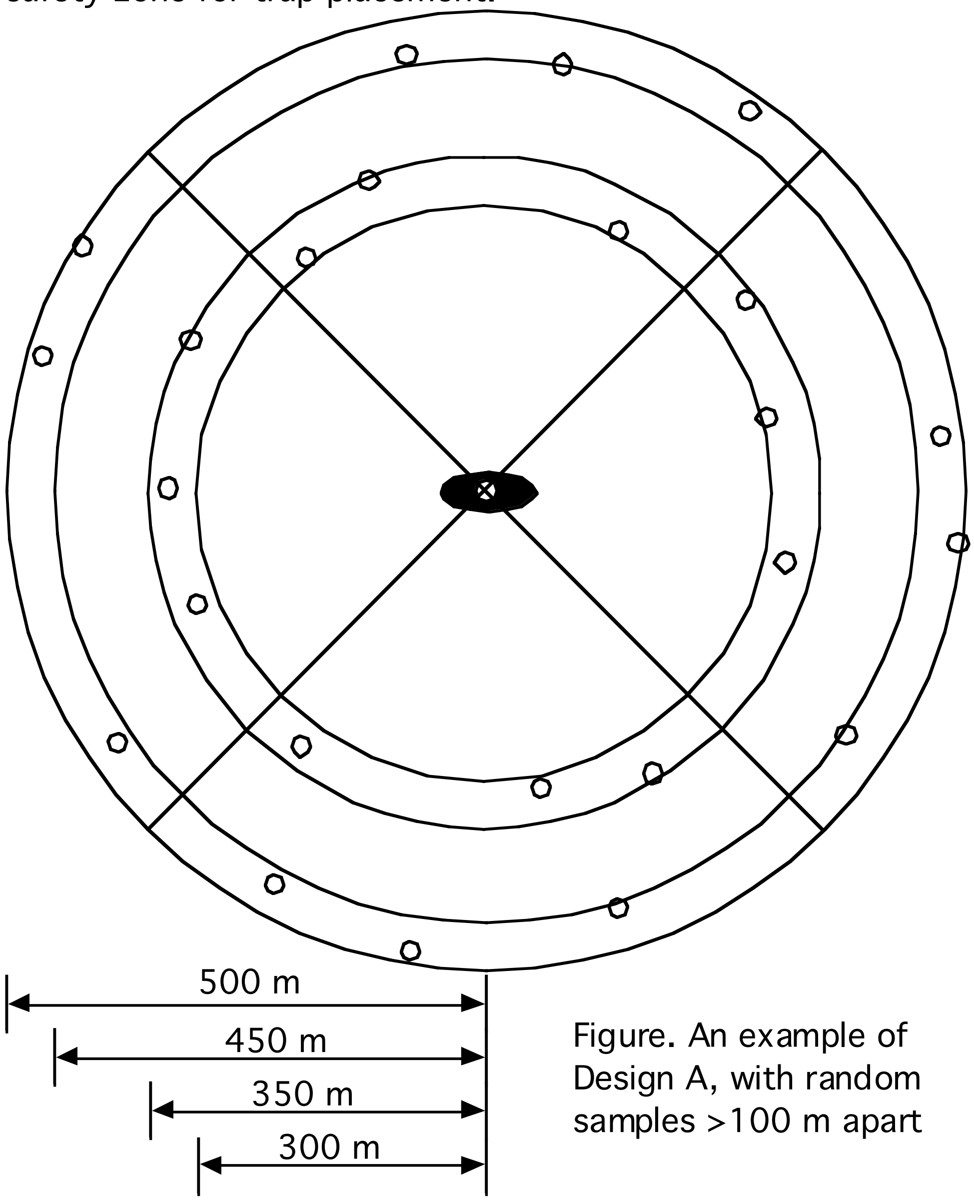
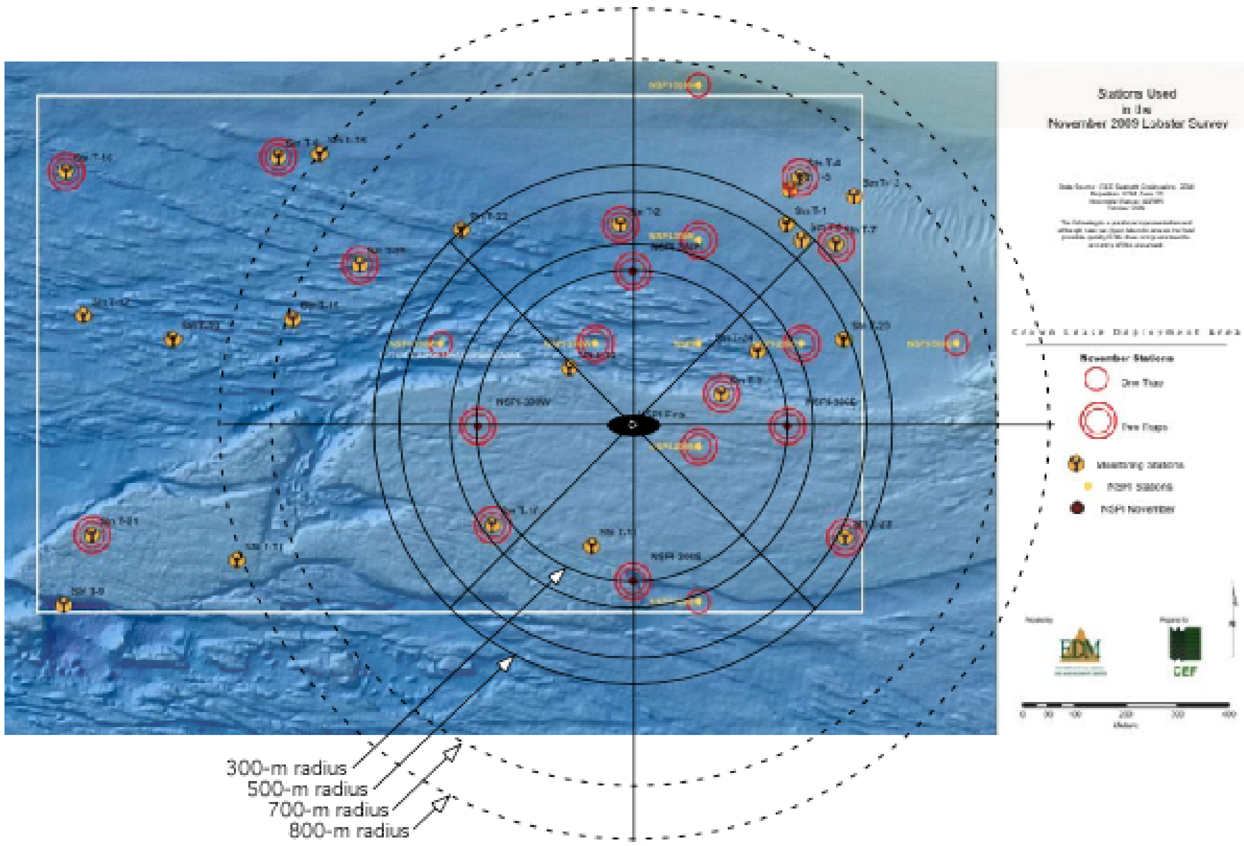


Figure. An example of Design A, with random samples >100 m apart

The following figure shows the placement of this design around the current turbine location in Area T.



I have shown the quadrats oriented as NSEW. It would make more sense to rotate the axis so that mid-points of two opposing quadrats were aligned with the mean tidal current direction. In this way, directional effects of the turbine in action could be tested and compared with the two lateral quadrats.

It has been suggested that wider annuli be added and sampled from and used instead of a continuation of Area T samples. I have drawn annuli of 700-m and 800-m radii merely to facilitate consideration. There may need to be censoring of projected samples in extremely shallow water. However, the data indicate that down to the minimum depth of 5.1–m sampled, there is no indicated change in direction of the increase in catch with decreasing depth as reported above under Water Depth. While tidal current will be slower, it is possible that turbine noise will amplify or be sustained in the water column or via the substrate as depth decreases. The 'before' control data will establish any differences that can be accounted for.

I should also add that for every mean distance chosen, the randomization design could occur along the perimeter of each circle rather than in an annulus. In either case, we are setting up a null hypothesis that there is no catch difference between certain mean distances and directions. Within that hypothesis I have randomized positions as much as possible.

Design B

Questions 3 and 4 and the desirability of continued monitoring will need to be reassessed if results of Design A are non-significant and/or of small effect magnitude. Otherwise, a longer term (question 3) and larger spatial scale (question 4) design needs to be considered. Such a strategy should include a continuation of turbine area sampling as described in Design A, and a continuation of Area E and W samples. Given the desirability of continuation of Design A sites, albeit at a lower intensity, and the inclusion of a significant part of Area T by those samples, I don't believe a continuation of Area T randomized sites is necessary unless it is known that more turbines are planned in that area.

Given the above recommendations for Design A, a similar sampling effort to each annulus in Design A, in terms of number of samples, could be applied to Areas E and W resulting in 36 samples each (total 72) per survey. Therefore, a total sampling effort of 144 samples per survey would include 36 each for Areas E and W. Replicating 3 times for each site, would require 12 randomized sites for Areas E and W. Existing sites can be used or randomly selected from. To summarize, each survey (each fall and/or spring) would dedicate 12 randomized sites (stations) to each of 4 strata: Turbine-300m, Turbine-500m, Area E, and Area W. Each site would be replicated three times. This would provide balance, sufficient power, and insurance for losses (with the option to 'rebalance' by selecting from completed samples for analysis). I cannot comment on the relative merits of fall and spring surveys, but a fall survey should be completed before the start of the fishing season in order to avoid any effect of that factor on experimental catches, especially in Area E.

Several Turbines

Regarding question 5, Design A could only be utilized if the turbines are considerable distances apart. Based on our current limited information "considerable distance" may be 500 to 1000-m. If they are too close to be considered to have independent effects on samples, many randomized locations would need to be censored without necessarily guaranteeing independence. An alternative would be to broaden the question by assessing the joint effect of a cluster of turbines. Instead of distinguishing distance from turbine on a categorical basis (treatment/control), as implied by the strata in Design A, one can take a continuous approach by expressing samples from each site in terms of their distance from one or more turbines (D_i for turbine i). One would not, however, expect D_i itself to best express the drop in intensity of noise or current from a turbine. For example, the decay of, say noise, from a point source over a two-dimensional surface (e.g., the sea bottom), should be approximately proportional to $1/D_i$. A joint effect of n similar turbines would therefore be proportional to $\Sigma(1/D_i)$ (i=1,n), which could be incorporated as a covariate in a simple BACI, in which its interaction with the before/after variable would be tested.

It should be added that considerable cooperation with the NS Power engineers would be required. Installing individual turbines over several seasons would result in a "moving target" and would not leave enough time for the joint impact to be monitored, and would extend the lapse of time between the before and after samples, thereby weakening the assumptions in the BACI analysis. .

Analysis comments
The total lobster catch, either here transformed as Ln(Catch/trap-set+1) or better still untransformed in a negative binomial model, is the most robust response statistically. However, it may be desirable for biological reasons to deal with subsets, such as market-size lobster. A related size grouping of interest would be 70-82 mm under legal size; 82-100 mm legal but immature, and >100 mm mature. Changes in such groups may occur independently of the total catch. Each group could be analyzed separately, but there will be higher CV and lower power. Alternatively, relative changes could be assessed by simultaneously analyzing catch rates of the three size groups using a multinomial response as an extension of the negative binomial, at the cost of 2 DF but with the benefit of using all data. This would test whether proportions of size classes responded differently to the treatment.

Conclusions

In summary, these recommendations are more in line with sampling reality and with feasible statistical analyses. Spatial randomization for each survey among all strata is statistically more appropriate (and was already done in the 3 Areas), yet given the pseudo-random implications ascribed to the variable locations of the current NS data it is not unreasonable to use most of the existing data (after allowing for the A,B separation as recommended above) in future comparisons, if variances are consistent. A further advantage of spatial randomizing within strata and attempting to conduct even sampling among Stations is that the design will remain robust in the face of inevitable trap losses and movement. Simple BACIs that depend on testing for a different change closer to the turbine compared to further afield, as demonstrated in the previous section, would be the most appropriate approach.

References

Green, R.H. (1979) Sampling Design and Statistical Methods for Environmental Biologists, Wiley, Chichester.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.)

Stewart-Oaten, A., Murdoch, W.W. & Parker, K.R. (1986). Environmental impact assessment: pseudoreplication in time? Ecology **67**, 929 – 940.

Appendix: Terms of Reference.
Address the following questions:

1. What is the appropriate statistical design for the base test and control study (BACI)?

2. What is the appropriate statistical design for the NSPI grid stations?

3. What transformations of the data are most appropriate for comparisons?

4. How much of the existing data can be used in future analysis?

5. How has station selection/reduction biased the analysis - is it acceptable?

Glossary

ANOVA = Analysis of variance, linear statistical approach that tests differences in means of groups defined by one or more categorical variables

Anacova = Analysis of covariance, linear statistical approach that tests differences in means of groups defined by one or more categorical variables, including effects of one or more continuous variables.

P or **Pr** = probability (between 0 and 1) that the result (e.g., an estimate of a coefficient) would be different from zero (null hypothesis not disproved) from repeated random sampling.

DF = Degrees of Freedom, which in statistical linear models is the number of samples less the number of fitted coefficients, being a measure of the robustness of the estimated model.

Location(Loc in analyses) is synonymous with Station, meaning a fixed sampling location determined randomly in the design. Of course this 'location' often moves by the time the trap is lifted.

Nuisance variable : is a variable that is not of interest in itself, but influences the results that may be confounded if the nuisance variable(s) were omitted.

Interaction (first order) in statistics is the effect of a change in one explanatory variable on the effect of a second explanatory variable on the response (Y). It is the same when the two explanatory variables are interchanged. It is denoted by ':' between the explanatory variables in the analysis results.