Deep-Sea Research II 110 (2014) 84–91

Contents lists available at ScienceDirect

Deep-Sea Research II

journal homepage: www.elsevier.com/locate/dsr2

# Repeated-measures regression designs and analysis for environmental effects monitoring programs

Michael D. Paine [a,1], Marc A. Skinner [b], Bruce W. Kilgour [c], Elisabeth M. DeBlois [d,*], Ellen Tracy [e]

[a] Paine, Ledge and Associates, 302-1050 Bowron Court, North Vancouver, BC, Canada V7H 2X6
[b] Stantec Consulting Ltd., 102-40 Highfield Park Drive, Dartmouth, NS, Canada B3A 0A3
[c] Kilgour & Associates Ltd, 2285C St. Laurent Boulevard, Unit 16, Ottawa, ON, Canada K1G 4Z6
[d] Elisabeth DeBlois Inc., 456 Marine Drive, Logy Bay, NL, Canada A1K 3E1
[e] Stantec Consulting Ltd., 141 Kelsey Drive, St. John's, NL, Canada A1B 0L2

ARTICLE INFO

Available online 24 October 2014

*Keywords:*
Repeated measures
Regression analysis
Environmental monitoring
Oil and gas fields
Statistical analysis

ABSTRACT

This paper provides a general overview of repeated-measures (RM) regression designs and analysis for marine monitoring programs, in support of sediment chemistry, particle size and benthic macroinvertebrate community analyses provided as part of this series. In RM regression designs, the same $n$ replicates (usually stations in monitoring programs) are re-sampled (i.e., repeatedly measured) at $t>1$ Times (usually years). The stations provide variation in the predictor, or $X$ variables. In the Terra Nova environmental effects monitoring (EEM) program, $n=48$ stations were sampled in each of $t=7$ years from 2000 to 2010. Two distance measures from five drill centres (sources of drilling wastes) were fixed predictor variables. RM regression designs are rarely used in environmental monitoring programs, but are often suitable and would be appropriate if applied to data from many monitoring programs. For the Terra Nova EEM program, carry-over effects, or persistent and usually small-scale variations among stations unrelated to distance, were strong for most sediment quality variables. Whenever natural carry-over effects are strong, RM designs and analysis will usually be more powerful and suitable than alternative approaches to the analysis.



## 1. Introduction

A repeated-measures (RM) regression design based on ordinary least squares (OLS) was used for the sediment quality component (sediment chemistry and particle size; benthic macroinvertebrate communities) of the environmental effects monitoring (EEM) program for the Terra Nova offshore oil development (DeBlois et al., 2014-a; Paine et al., 2014). RM regression designs are not widely used in the environmental sciences, and may be unfamiliar to many readers. RM regression designs are an alternative to the tradition RM analysis of variance (ANOVA) design. As such, our goal is to highlight distinctions between RM ANOVA and RM regression designs and illustrate the utility of the latter in environmental studies, with emphasis on the monitoring effects from offshore marine oil and gas operations.

This paper provides a general review of and guidance on how to interpret spatial–temporal RM regression designs for field monitoring of effects of offshore oil developments and other anthropogenic activities. Selected data from the Terra Nova EEM program are then used to illustrate specific points and some practical limitations of general statistical theory, design, methods and analysis for RM regression.

## 2. RM regression designs

### 2.1. Overview

RM ANOVA designs are widely used in the medical and social sciences. Green (1993), von Ende (1993), Paine (1996) and Quinn and Keough (2002) discuss the use of RM ANOVA in biological, ecological and environmental studies. RM designs involve the re-sampling of the same replicates (Subjects) during each of $t$ Times. Designs that involve the sampling of a different set of replicates during each time period are normally less statistically powerful than RM designs. While it should be evident that the statistical

* Corresponding author.
*E-mail address:* edeblois@nf.sympatico.ca (E.M. DeBlois).
[1] Dr. Michael Paine is now deceased.

http://dx.doi.org/10.1016/j.dsr2.2014.10.017
0967-0645/

approach taken to analyze a dataset will be dictated by the study design established prior to any survey or experiment, this is not always the case. Many examples in the biological literature have documented improper "factors × times" analyses of data that clearly should have been approached with RM-based analyses (Maceina et al., 1995; Paine, 1996).

RM designs are most effective when replicates are rare or expensive, and when there is persistent and large variance among replicates (i.e., carry-over effects) that is unrelated to factors of interest (i.e., the treatment; Paine, 1996; Quinn and Keough, 2002). However, RM designs require that every replicate be sampled in every time period, which is not always possible. Missing any single replicate sample (for whatever reason such as sample damage, sample loss, inability to collect the sample for safety reasons, etc.) will have the result that the sample across all time periods either will not be useable in the RM analysis or the missing values will need to be estimated (Green,1993).

The before–after-control-impact (BACI) family of designs, originating from Green (1979), are unique cases of RM approaches that have been espoused as optimal designs to detect the effect of interventions in environmental situations (Green, 1979; Stewart-Oaten et al., 1986; Underwood, 1989, 1991, 1993, 1994; Kingsford, 1998; Quinn and Keough, 2002). BACI designs are often used to assess the effect of a new point-source discharge(s) on physical, chemical, or biological receptors relative to control sites (Green, 1979). The repeated measurements in BACI designs are the replicate samples (locations) within control and impact treatments, during time periods both before and after the intervention (drilling in the case of the Terra Nova development). Criticisms suggesting BACI designs are pseudoreplictaed (Hurlbert, 1984) are not relevant when multiple impact and/or control sites are sampled more than two times.

RM regression designs are a special case of the general RM design in which sampling and characterization of a *Y* response is carried out not within treatment groups, but over a range of values of an *X* predictor variable. RM regression designs are most effective when *X* predictor variables (e.g., in this paper: distance from a drill centre) are relatively continuously distributed, and relationships between the response and the predictor variables are evident. Draper and Smith (1998) discuss optimization of *X* distributions for various purposes and expected *Y*–*X* relationships.

RM regression designs are rarely used in biological, ecological and environmental studies (but see Green, 1993). RM regression is sometimes used in medical studies, with baseline (pre-treatment) values of the response or *Y* variable (e.g., blood pressure) being used as a predictor (*X* variable) of a future response (see additional examples in Donner, 1984). RM ANCOVA designs, with both a Treatment factor and one or more covariates (*X*), are sometimes used in biological and ecological studies, and frequently used in medical studies (again, often with baseline *Y* values as *X*). Quinn and Keough (2002; pp. 356–357) discuss the use of covariates (*X*) in RM ANCOVA designs. RM regressions as defined in this paper are simplified versions of RM ANCOVA, with no additional treatment factor (such as Study versus Reference). Quinn and Keough (2002) and others provided guidance on the mechanics of RM ANCOVA.

In contrast to RM regression and ANOVA approaches based on OLS, generalized linear mixed models (GLIMM) are based on maximum likelihood estimation (MLE) and therefore not subject the restrictive assumptions of normally distributed data, patterns of variance, or balancing (Quinn and Keough, 2002). Additionally, in longitudinal studies, GLIMM do not require assumptions related to the correlation structure of within-subjects observations as are required in traditional, OLS-based RM approaches. With GLIMM, non-normal data (e.g., poisson, negative binomial, etc.), heterogeneity of variance and/or covariance, or unbalanced data can be incorporated into the linear model. As such, recent statistical literature has proposed GLIMM as a robust alternative to traditional RM approaches (Quinn and Keough, 2002; Wang and Goonewardene, 2004; Misangyi et al., 2006). In the only known primary literature example directly comparing RM regression to GLIMM using common datasets, Misangyi et al. (2006) conclude RM regression is inferior to GLIMM (MLM models by the nomenclature of these authors) under most circumstances. This conclusion aside, these authors' own analysis shows that when designs are balanced and sphericity assumptions are met, RM regression yields the same results and conclusions as GLIMM (Misangyi et al., 2006). Obviously, if one violates the assumptions of any statistical test, unbiased results should not be expected. While we acknowledge GLIMM approaches offer greater analytical flexibility with fewer assumptions, this class of analyses were not widely known when the Terra Nova offshore oil EEM program was designed in the mid-1990s. Further, RM regression is intuitive for the majority of practitioners and in the words of Misangyi et al. (2006), "it is an extension of familiar repeated-measures ANOVA and builds on standard regression analyses."

This paper focuses on a case that involves re-sampling of stations over time at precisely known locations, or distances (*X*s) from a drill centre. Green (2004) discusses the selection of sample times in offshore oil and other marine monitoring programs. In most cases, times are years, with sampling conducted annually or less frequently. Most monitoring programs are restricted to a relatively short time interval, so the number of years ($t$) and the specific years that can be sampled is limited. Investigators cannot a priori select a subset of representative or similar years. Only one or a few baseline (Before) years before some event or intervention can be sampled in most programs. In longer-term programs, simple trend analyses can be conducted in After years (Section 3.2).

### 2.2. Terra Nova sediment quality monitoring design

At the Terra Nova offshore oil development, there are five drill centres where drilling has occurred in the past (active drill centres) (see DeBlois et al., 2014-b). The Northeast (NE), Northwest (NW), Southeast (SE) and Southwest (SW) drill centres are located at the four corners of a Fisheries Exclusion Zone (FEZ) and are referred to in this paper as the FEZ drill centres. The Far East (FE) drill centre is located approximately 5 km east of the centre of the development. Drill cuttings are discharged at the drill centres. Thus, there are five sources (drill centres) of drill cuttings, in contrast to some other sites, where cuttings are discharged directly below a fixed platform and single source.

The regression design used for the sediment quality component of the Terra Nova EEM program involved approximately 50 sample stations located at various distances (*X*) from drill centres (see DeBlois et al., 2014-a). Forty-eight (48) stations were re-sampled in all seven EEM years (2000, 2001, 2002, 2004, 2006, 2008 and 2010). Thus, the design was a RM regression design with few missing year × station values. The regression design allowed sampling over a broad range of distances from each of the five sources ($<1$ to $>20$ km). Re-sampling the same set of stations was convenient for planning and field sampling, since the survey plan (i.e., station locations) did not change over time. Re-randomization, or sampling a different set of stations each year, would have added an unnecessary complicating factor.

Drilling first occurred at: the NE and SW drill centres prior to 2000 sampling; the NW and SE drill centres prior to 2001 sampling; and the FE drill centre prior to 2002 sampling. Thus, a number of Before versus After comparisons were possible. Drilling was not continuous at any drill centre after first drilling (see DeBlois et al., 2014-b, for drilling statistics). Thus, "after" times