



UNIVERSIDADE FEDERAL DA BAHIA

INSTITUTO DE BIOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E BIOMONITORAMENTO
MESTRADO PROFISSIONAL EM ECOLOGIA APLICADA À GESTÃO AMBIENTAL

MARINA PESSOA FELZEMBURGH MOTTA

Utilização da análise de poder em estudos de avaliação de impactos ambientais

Salvador, Bahia

2016

MARINA PESSOA FELZEMBURGH MOTTA

Utilização da análise de poder em estudos de avaliação de impactos ambientais

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ecologia e Biomonitoramento como requisito parcial para a obtenção do título de mestre em Ecologia Aplicada à Gestão Ambiental, pela Universidade Federal da Bahia.

Orientador: Prof. Dr. Gilson Correia de Carvalho

Salvador, Bahia

2016

Sistema de Bibliotecas da UFBA

Motta, Marina Pessoa Felzemburgh

Utilização da análise de poder em estudos de avaliação de impactos ambientais / Marina Pessoa Felzemburgh Motta, Gilson Correia de Carvalho. -- Salvador, 2016.

57 f.: il.

Orientador: Gilson Correia de Carvalho.

Dissertação (Mestrado - Ecologia Aplicada à Gestão Ambiental) -- Universidade Federal da Bahia, Instituto de Biologia, 2016.

1. Análise de Poder. 2. Poder Estatístico. 3. Impacto Ambiental. 4. Monitoramento Ambiental. I. de Carvalho, Gilson Correia. II. Universidade Federal da Bahia. Instituto de Biologia. III. Título.

CDD -

CDU -

Folha de aprovação

MARINA PESSOA FELZEMBURGH MOTTA

**Utilização da análise de poder em estudos de avaliação de
impactos ambientais**

Programa de Pós-Graduação em Ecologia e Biomonitoramento

Universidade Federal da Bahia

Membros da banca examinadora

Membros titulares:

Prof. Dr. Gilson Correia de Carvalho (Orientador)

(Departamento de Biointeração, Instituto de Ciências da Saúde, Universidade Federal da Bahia)

Prof^a. Dra Camila Magalhães Pigozzo

(Centro Universitário Jorge Amado)

MSc. Gabriel Barros Gonçalves de Souza

(Centro de Ciências da Saúde, Instituto de Biologia, Universidade Federal do Rio de Janeiro)

APRESENTAÇÃO

O uso indiscriminado dos recursos naturais pelo homem tem causado enorme degradação ambiental nos mais diversos ecossistemas do mundo. Com isso, tem crescido a necessidade e a importância de estudos de avaliação ambiental, prévios ou posteriores à presença de variados estressores. Técnicas estatísticas visam auxiliar nas análises de impacto ambiental e são largamente utilizadas na literatura, sob uma gama diversificada de testes estatísticos, desde mais simples até análises bastante complexas. Contudo, serão esses testes suficientes para garantir confiabilidade aos dados, evitando ou minimizando danos ambientais?

No âmbito dos testes de hipóteses, comumente utilizados em estudos de avaliação ambiental, hipóteses investigadas nunca podem ser provadas, podem apenas ser rejeitadas com as ferramentas de inferência estatística (STEIDL, HAYES E SCHAUBER, 1997). Por isso, trabalha-se com uma hipótese nula, a qual geralmente é do tipo “não existe alteração” e que, sendo rejeitada, fortalece a hipótese de trabalho. Neste tipo de investigação, contudo, a cada resultado obtido, dois tipos de erros podem estar sendo cometidos, o de não rejeitar uma hipótese nula falsa ou de rejeitar uma hipótese nula verdadeira.

Para evitar esse segundo tipo de erro, o poder estatístico da análise deve ser avaliado com o intuito de estimar a confiabilidade nas conclusões obtidas com os testes de hipóteses, reduzindo os riscos ao meio ambiente e auxiliando na tomada de decisões por gestores ambientais. O presente trabalho teve por objetivo fazer uma revisão da utilização da análise de poder em estudos de avaliação de impactos antrópicos que utilizaram dados biológicos nas últimas quatro décadas e exemplificar maneiras de aplicação de análises de poder em trabalhos técnicos, com o intuito de auxiliar na popularização deste tipo de abordagem.

Esta dissertação está dividida em dois capítulos, sendo o capítulo I constituído de uma revisão do uso da análise de poder em estudos de avaliação de impactos ambientais e o capítulo II composto por um manual com um passo a passo e exemplos de aplicação de análises de poder.

Palavras-chave: Poder estatístico. Impacto ambiental. Monitoramento.

SUMÁRIO

APRESENTAÇÃO	1
LISTA DE FIGURAS	3
LISTA DE TABELAS	4
CAPÍTULO I	5
RESUMO	6
ABSTRACT	6
1. INTRODUÇÃO	7
2. OBJETIVOS	11
2.1. Objetivos específicos	11
3. MATERIAIS E MÉTODOS	11
4. RESULTADOS	13
5. DISCUSSÃO	18
6. CONSIDERAÇÕES FINAIS	21
7. REFERÊNCIAS BIBLIOGRÁFICAS	23
CAPÍTULO II	29
RESUMO	30
ABSTRACT	30
1. INTRODUÇÃO	31
2. OBJETIVOS	33
2.1. Objetivos específicos	33
3. MATERIAIS E MÉTODOS	34
4. RESULTADOS	36
4.1. Análise de poder <i>a priori</i>	36
4.2. Análise de poder <i>a posteriori</i>	39
5. DISCUSSÃO	40
6. CONSIDERAÇÕES FINAIS	51
7. REFERÊNCIAS BIBLIOGRÁFICAS	52

LISTA DE FIGURAS

Capítulo 1

Figura 1: Variação histórica do número de estudos ambientais utilizando análises de poder publicados até setembro de 2016.	14
Figura 2: Tipo de análise de poder utilizado nos artigos analisados no presente estudo.	15
Figura 3: Níveis de poder considerados adequados nos estudos ambientais analisados no presente estudo.....	15
Figura 4: Testes estatísticos empregados nos estudos ambientais analisados no presente estudo.....	16
Figura 5: Softwares utilizados para realização das análises de poder nos estudos aqui analisados.	16
Figura 6: Fontes de alteração antrópica sobre os ecossistemas e indicadores biológicos nos estudos avaliados.	17
Figura 7: Indicadores biológicos utilizados nos estudos avaliados. Alguns estudos utilizaram mais de um desses indicadores.	18

Capítulo 2

Figura 1: Passo a passo para realização de uma análise de poder a priori no software GPower 3.1.....	37
Figura 2: Outra forma de seleção do tipo de teste, através da aba “Tests”.....	38
Figura 3: Saída dos resultados.....	39
Figura 4: Variação do passo a passo para realização de uma análise de poder post hoc no software GPower 3.1.	39
Figura 5: Plotando um gráfico poder x tamanho amostral.....	42
Figura 6: Gráficos poder x tamanho amostral para 3 tamanhos de efeito diferentes, com $\alpha = 0,05$ (esquerda) e $\alpha = 0,10$ (direita) (situação hipotética 1).	43
Figura 7: Quadros de comparação do poder do teste t (a esquerda) e do teste Wilcoxon Mann-Whitney (a direita) para N, tamanho de efeito e α equivalentes.	43
Figura 8: Gráficos poder x tamanho amostral para 3 tamanhos de efeito diferentes, com $\alpha = 0,05$ (esquerda) e $\alpha = 0,10$ (direita) (situação hipotética 2).	46
Figura 9: Tamanhos amostrais em função do poder para diferentes níveis de significância.....	47

Figura 10: Calculando o tamanho de efeito observado (aba da direita) e o poder do teste (aba da esquerda) no GPower 3.1.	50
---	----

LISTA DE TABELAS

Capítulo 1

Tabela 1: Possíveis resultados para testes de hipótese.	8
--	---

Capítulo 2

Tabela 1: Valores de tamanho de efeito categorizados para diferentes testes estatísticos.	35
--	----

Tabela 2: Tamanhos amostrais obtidos pela combinação dos outros parâmetros da análise de poder a partir de tamanhos de efeito pré-definidos (situação hipotética 1).	41
---	----

Tabela 3: Dados simulados, tamanhos de efeito calculados e N amostrais obtidos para situação hipotética 1.	44
---	----

Tabela 4: Tamanhos amostrais obtidos pela combinação dos outros parâmetros da análise de poder (situação hipotética 2).....	45
---	----

Tabela 5: Valores aleatórios obtidos como representação das amostras coletadas na situação hipotética 3 para os 4 pontos amostrais (norte, sul, leste, oeste), com respectivas médias e desvios-padrão.	49
--	----

CAPÍTULO I

Análise de poder em estudos de avaliação de impactos ambientais: revisão sistemática

Salvador, Bahia

2016

ANÁLISE DE PODER EM ESTUDOS DE AVALIAÇÃO DE IMPACTOS AMBIENTAIS: REVISÃO SISTEMÁTICA

Marina Pessoa Felzemburgh Motta, Gilson Correia de Carvalho

RESUMO

Testes de hipótese são comumente utilizados em estudos de monitoramento de impactos ambientais. Visam aumentar a confiabilidade das análises estatísticas utilizadas como ferramentas adicionais para detectar a ocorrência de impactos antrópicos ao meio ambiente e, conseqüentemente, evitá-los. Tais testes objetivam a identificação de possíveis impactos, porém, convencionalmente focam em minimizar a detecção errônea de um impacto evitando, assim, custos adicionais ao estudo e, conseqüentemente, ao empreendedor. Contudo, seu uso indevido pode negligenciar impactos e, assim, embasar danos ambientais silenciosos. A análise de poder, objeto desta revisão, é uma ferramenta ainda pouco utilizada, cujo objetivo é analisar a capacidade de evitar tais problemas com base no delineamento amostral e no teste estatístico escolhidos. Foram analisados 34 estudos de monitoramento ambiental que realizaram algum tipo de análise de poder. As ciências aquáticas são as que mais tem incorporado a análise de poder em seus estudos. Os peixes e os invertebrados aquáticos foram os grupos mais utilizados nessas pesquisas. O tipo de análise de poder mais comum foi a post hoc, utilizada em 28% dos estudos. O software mais utilizado foi o R, seguido pelo GPower, SYSTAT, PRIMER e SAS e os testes mais empregados foram a ANOVA e o Teste T. Observou-se que a análise de poder ainda permanece uma temática incipiente nas ciências ambientais, devendo ser mais disseminada de modo a obter conclusões mais consistentes e melhor embasar a tomada de decisões por gestores ambientais.

ABSTRACT

Hypothesis testing are commonly used in environmental impact monitoring studies. They aim to increase the reliability of statistics analysis that have been used as additional tools to detect the occurrence of anthropic impacts to the environment, therefore preventing it. Conventionally those studies aim to identify possible environmental impacts, however traditionally focus on minimize impact misdetection,

avoiding additional costs to the studies and consequently to the entrepreneur. However, its misuse might neglects some natural impacts leading to environmental damage. Power analysis, object of this review, is a still underutilized tool whose purpose is analyze the ability to avoid such problems based on the sample design and statistical test chosen. There were analyzed 34 environmental monitoring studies that perform any type of power analysis. The aquatic science is the one the has incorporated more amounts of analysis of power in their studies. Fish and aquatic invertebrates were the most used groups in these researches. The type of Power analysis mostly applied was a post hoc, found in 28% of the studies. The software mostly used was R, followed by GPower, SYSTAT, PRIMER and SAS and the tests mostly employed were ANOVA and T Test. It was observed that the analysis of power still remains an incipient issue in the environmental sciences, and should be more disseminated in order to obtain more consistent conclusions and better support decision making by environmental managers.

1. INTRODUÇÃO

Em estudos ambientais muitas vezes é difícil diferenciar uma alteração devido à variabilidade natural de uma alteração por razões antrópicas, especialmente se esta última é pequena comparada à natural (OSENBERG et al., 1994). Neste sentido, hipóteses são testadas a fim de conferir certo grau de confiabilidade às análises através do nível de significância do teste (α), convencionalmente estabelecido no valor crítico (α_c) de 0,05 (MAPSTONE, 1995), ou seja, o teste deve ter 95% de confiança.

Existe, contudo, um grande problema em utilizar um nível de α pré-definido para toda e qualquer análise estatística: o critério de significância frequentemente é utilizado sem considerar a qualidade do estudo e dos dados e ignora informações críticas sobre o poder estatístico, podendo resultar na não rejeição de hipóteses nulas apesar da existência de efeitos reais e relevantes ou na rejeição de hipóteses nulas quando os efeitos existentes são triviais (MUDGE et al., 2012). Significância estatística não é sinônimo de relevância biológica e, portanto, efeitos biológicos triviais podem ser estatisticamente significativos e ter atenção desnecessária

enquanto efeitos biologicamente importantes podem ser perdidos por não serem estatisticamente significativos devido a um insuficiente poder estatístico (HANSON, 2011).

Poder estatístico é a probabilidade de corretamente rejeitar uma falsa hipótese nula em um determinado teste estatístico (MORRISEY, 1993). No contexto de um monitoramento, a probabilidade de detectar um impacto onde não existe, ou seja, gerar um alarme-falso, é denominada de erro tipo I e é calculada através do nível de significância α . Por outro lado, a probabilidade de errar ao não detectar um impacto que esteja ocorrendo constitui um erro tipo II (β), o qual dá uma falsa sensação de segurança, pois a conclusão pode ser de que não há impacto quando na verdade existe um, e assim nenhuma medida é tomada quando deveria ser (FAIRWEATHER, 1991).

Tabela 1: Possíveis resultados para testes de hipótese.

	Manter H_0	Rejeitar H_0
H_0 verdadeira	Decisão correta ($1 - \alpha$)	Erro Tipo I (α)
H_0 falsa	Erro Tipo II (β)	Decisão correta (poder = $1 - \beta$)

FONTE: Adaptada de Gotelli e Ellison (2011)

A análise de poder é uma importante ferramenta para a interpretação de dados e resultados experimentais: ela é capaz de quantificar a magnitude e o tipo de mudança provável de ser detectada e assim torna-se uma ferramenta útil para o monitoramento ambiental por dar uma medida de confiabilidade e credibilidade às tendências observadas (ZANGRANDI; CATSIKI; AVDIS, 2005).

A análise de poder não adiciona um entendimento da natureza em si, mas eles permite que algum controle de qualidade seja exercido ao longo das inferências derivadas de estudos ambientais (STEINBECK; SCHIEL; FOSTER, 2005).

Em estudos ambientais, a hipótese nula geralmente é entendida como “não há evidência de mudança nas variáveis medidas em resposta ao distúrbio”, ou seja, caso o teste aceite a hipótese nula como verdadeira, assume-se que não há impacto (MORRISEY, 1993). Contudo, ainda segundo este autor, este resultado pode ser

apenas fruto de um teste com poder insuficiente para anular tal hipótese nula e, portanto, fraco ao afirmar que de fato não existe um impacto.

Um alarme-falso pode levar os gestores a tomar ações onde elas não são necessárias, com consequências para as indústrias que constituem as principais fontes de impacto, podendo resultar na interrupção de um trabalho, no fechamento de uma fábrica ou na imposição de cotas de pesca. Por outro lado, um erro tipo II pode representar um aval para a continuidade do impacto silencioso causado, o que pode causar séria degradação ambiental podendo inclusive ser agravada após longos períodos de exposição não detectada (MAXWELL; JENNINGS, 2005; MAPSTONE, 1995). Portanto, falhar em detectar um impacto pode carregar um custo que usualmente é maior do que o custo de erroneamente detectar um impacto (FAIRWEATHER, 1991).

Segundo Mapstone (1995), em questões de impactos ambientais, um resultado “significativo” é motivo de preocupação, ação, litígio, discussão – não importa quão pequeno ou inconsequente seja o impacto; “não significância”, por sua vez, é talvez controverso, frequentemente leva a complacência, mas raramente precipita uma ação, preocupação ou discussão, mesmo que grandes e importantes impactos tenham sido perdidos.

Apesar da sua grande relevância para estudos ambientais, especialmente os de longo prazo, Green (1989) constatou que a análise de poder era bastante subutilizada na ecologia, tanto em estudos científicos quanto em análises técnicas.

Ao contrário do α , que é convencionalmente definido, o valor de β pode ser apenas indiretamente controlado, sendo proporcionalmente relacionado ao valor de significância do teste (α), ao tamanho da amostra e ao tamanho do efeito detectado (FAIRWEATHER, 1991), sendo que este último não é facilmente controlado. Além disso, não é claro o valor definido *a priori* como adequado para o β (e consequentemente para o poder) (FRANCO; QUINTINO; ELLIOTT, 2015).

Uma aproximação ambientalmente conservativa é usar um baixo nível β . Isto pode ser atingido ao custo de um elevado risco de erro tipo I, ou ao custo de um aumento do esforço amostral (HANSON, 2011). Alguns autores defendem que deve haver uma ponderação entre os riscos dos erros tipo I e tipo II, através de uma

equivalência do α e do β (TOFT; SHEA, 1983; PETERMAN, 1990). Mapstone (1995) sugeriu que, na ausência de informações suficientes sobre os custos dos erros, sejam atribuídos valores iguais ao α e β , ao invés da tradicional estratégia de usar um α constante e deixar o β aleatório, equilibrando os riscos de falsas conclusões entre o potencial poluidor e o ecossistema em risco.

Análises de poder podem ser muito úteis como uma ferramenta *a priori* para delinear regimes amostrais e são importantes para cálculos *a posteriori* após análises estatísticas não significativas (GLASBY, 1997). Segundo Peterman (1990) e Steidl et al. (1997), a análise de poder é idealmente utilizada no planejamento de pesquisas, para garantir que os recursos, geralmente escassos, sejam utilizados tão efetivamente quanto possível para atingir os objetivos do estudo. Já a utilização da análise de poder *post hoc* ou *a posteriori* tem ganhado espaço entre estudos cujos resultados foram não significativos, de modo a distinguir se de fato nenhuma alteração foi detectada por inexistência do fato ou por cometimento de erro tipo II (STEIDL et al., 1997). Há uma outra vertente de utilização da análise de poder, que é para determinar a magnitude de tendências antes de se tornarem significativas, possibilitando a tomada de decisões que evitem que populações ameaçadas cheguem a níveis inaceitáveis de risco de extinção (MAXWELL; JENNINGS, 2005) e pode ser utilizada ainda para demonstrar quais testes estatísticos são mais sensíveis e poderosos (FAIRWEATHER, 1991)

Segundo Cohen (1992), estimar o poder da análise para testes paramétricos é relativamente simples, através dos valores de α , tamanho de efeito esperado e variância, que pode ser estimada através de estudos prévios. Testes paramétricos precisam satisfazer uma série de pressupostos para serem aplicados: i. amostras independentes e igualmente distribuídas; ii. variâncias entre grupos homogêneas; iii. resíduos com distribuição normal; iv. correta classificação das amostras; e v. efeitos principais aditivos entre si (GOTELLI & ELLISON, 2011). O problema, contudo, é que em estudos ecológicos frequentemente os dados não cumprem todas as premissas necessárias e por isso é preciso utilizar análises não paramétricas, para as quais a informação disponível é ainda relativamente pequena (PAIS et al., 2014).

Thomas e Krebs (1997) fizeram uma extensa revisão dos softwares utilizados para análises de poder à época e demonstraram, já na década de 90, a existência de

uma ampla gama de softwares capazes de fazer análises de poder tanto *a priori* quanto *a posteriori*, concluindo não haver razões para continuar a ignorar o poder de testes estatísticos em estudos ecológicos. Atualmente novos softwares surgiram, porém a análise de poder permanece subutilizada.

2. OBJETIVOS

Tendo em vista a grande importância deste tema, o presente trabalho objetivou fazer uma revisão das formas de utilização da análise de poder em trabalhos de avaliação ambiental, de modo a verificar quais os métodos e softwares mais utilizados e quais áreas da ecologia tem incorporado essa temática.

2.1. Objetivos específicos

- Quantificar quantos estudos técnicos na área ambiental tem utilizado a análise de poder como ferramenta desde a sua concepção;
- Apontar como a utilização da análise de poder na área ambiental tem crescido ao longo do tempo;
- Avaliar quais os tipos de análise de poder tem sido aplicadas nos estudos ambientais;
- Analisar quais programas tem sido utilizados para a realização de análises de poder em estudos ambientais;
- Verificar quais áreas das ciências ambientais tem incorporado a análise de poder em seus estudos;
- Listar os tipos de impacto ambiental que estão sendo avaliados através de análises de poder;

3. MATERIAIS E MÉTODOS

Para realizar a revisão bibliográfica acerca da utilização da análise do poder estatístico foi feito um levantamento em publicações periódicas de revistas indexadas que abordaram situações de monitoramento/diagnóstico de distúrbios

antrópicos utilizando indicadores biológicos. Este levantamento foi feito nas bases de dados *Web of Science*, *Science Direct* e *Scopus*, através da ferramenta “pesquisa avançada” das mesmas. As palavras-chave utilizadas foram “avaliação de impacto”, “monitoramento” e “avaliação de impacto ambiental” associadas com as palavras-chave “análise de poder” e “tamanho de efeito”. Os três primeiros termos foram combinados com os dois últimos para abranger todas as possibilidades de combinação entre eles. Todos os termos foram pesquisados em inglês e entre aspas para direcionar as buscas para o termo exato pesquisado. Não foi inicialmente delimitado um período para o levantamento bibliográfico de modo a obter resultados desde o princípio da utilização da temática proposta.

Foram utilizados caracteres especiais na composição das fórmulas de pesquisa para auxiliar a obtenção do resultado desejado, filtrando possíveis desvios da temática central. Em alguns termos foram utilizados asteriscos (*) no lugar de uma ou mais letras para evitar a perda de trabalhos que tenham porventura escrito o termo pesquisado de forma diferente ou errada (e.g. Power analysis x Power analyses). As palavras-chave utilizadas nas pesquisas foram: "impact a*e*ment", "monitoring", "Environment* Impact a*e*ment", "Power Analys*s" e "Effect Size"), adicionando-se o termo de busca específico para cada um das bases. Foram utilizados também os termos “OR” e “AND” para fazer a associação de palavras-chave desejadas. Sendo assim, as equações de busca empregadas foram:

- Para a base de dados Web of Science: ts=("impact a*e*ment" OR "monitoring" OR "Environment* Impact a*e*ment") AND ("Power Analys*s" OR "Effect Size")
- Para a base de dados Science Direct: TITLE-ABS-KEY("impact a*e*ment" OR "monitoring" OR "Environment* Impact a*e*ment") AND ("Power Analys*s" OR "Effect Size")
- Para a base de dados Scopus: TAK ("impact a*e*ment" OR "monitoring" OR "Environment* Impact a*e*ment") AND ("Power Analys*s" OR "Effect Size")

Desta forma, a intenção foi obter uma amostra representativa de artigos publicados que utilizaram a análise de poder ou análise do tamanho de efeito em estudos ambientais.

Do resultado inicial obtido foi feita uma triagem a partir da leitura dos resumos das publicações encontradas. Essa triagem buscou a presença de algum tipo de análise de poder vinculada a uma interferência antrópica em determinado indicador biológico. Quando não era observado um impacto (positivo ou negativo) o trabalho era desconsiderado da seleção. Além disso, foram incluídos na seleção apenas aqueles que apresentavam também a presença de aspecto temporal do estudo, com ocorrência de pelo menos duas campanhas de amostragem.

Foi elaborado um banco de dados no software Microsoft Excel, de modo a enumerar os artigos selecionados com algumas informações gerais de cada um, a saber: (I) O tipo de estudo, se diagnóstico ou de monitoramento; (II) Qual a duração do estudo; (III) Qual a análise realizada; (IV) Qual o software utilizado; (V) Se a análise de poder foi realizada antes ou após a coleta dos dados; (VI) Qual o indicativo de poder considerado adequado; (VII) Qual o indicador biológico utilizado; (VIII) Qual foi o foco da análise: tamanho de efeito, tamanho amostral ou o poder do teste; (IX) Qual a pressão antrópica sobre o ecossistema.

De posse deste banco de dados, procedeu-se a uma avaliação de como a análise de poder tem sido realizada nas últimas décadas, quais as metodologias/programas mais utilizados, quais as maiores dificuldades encontradas e como tem sido os resultados obtidos. De modo a ilustrar os resultados, foi realizada uma análise descritiva das informações obtidas utilizando gráficos de frequência gerados através do software Microsoft Excel 2007.

4. RESULTADOS

A partir do levantamento bibliográfico realizado, foi obtido inicialmente um total de 1103 artigos, dos quais após a triagem obteve-se 34 artigos considerados relevantes para a análise. A fase de levantamento bibliográfico foi concluída em setembro de 2016, portanto só contempla parte do referido ano. Os artigos selecionados foram publicados entre 1993 e 2016 e são apresentados no Apêndice A. É um número considerado muito pequeno, dada a importância do tema e o tempo transcorrido desde as primeiras publicações sobre o assunto, refletindo a grande subutilização desta análise em estudos ecológicos.

Cabe ressaltar que um número muito grande de estudos ambientais não chega a ser publicado, principalmente se não foram realizadas através de iniciativa privada e/ou não apresentaram resultados estatisticamente significativos de acordo com a estatística convencional. Tais estudos não foram incluídos na análise, pois grande parte deles é oriunda de trabalhos técnicos de consultoria ambiental, o que envolve questões de propriedade e sigilo dos dados. Desta forma, a amostra obtida pode refletir um universo pequeno dentro dos estudos ambientais.

Os primeiros artigos sobre a análise de poder aplicada à área ambiental foram publicados na década de 80, tendo ganhado força a partir da década de 90, como pode ser observado na Figura 1. Os primeiros artigos publicados, contudo, foram mais conceituais, sem uma análise de poder diretamente aplicada e, por isso, não foram incluídos na contabilização dos artigos aqui analisados. Alguns exemplos disso são Toft e Shea (1983), Green (1989), Peterman (1990), Fairweather (1991) e Cohen (1992).

A maioria dos trabalhos analisados (32) focou em estudos de monitoramento, com pelo menos um ano de duração. Apenas dois trataram de estudos de diagnóstico, prévios ao licenciamento para implantação ou expansão de empreendimentos.

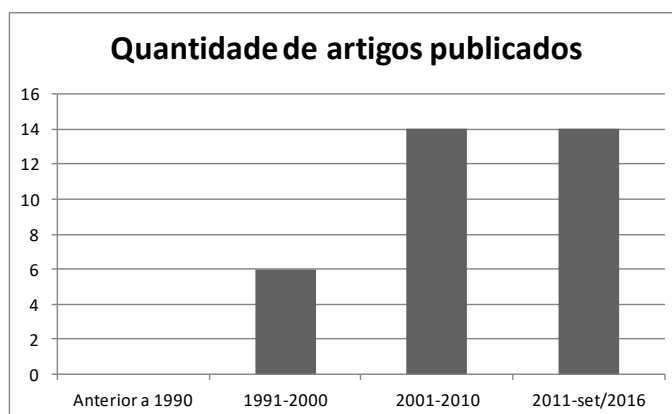


Figura 1: Variação histórica do número de estudos ambientais utilizando análises de poder publicados até setembro de 2016.

Em termos do tipo de análise de poder, o resultado foi relativamente equilibrado, como pode ser observado na Figura 2. Cerca de 23,5% dos estudos focaram a análise na avaliação do poder propriamente dito, 20,6% no tamanho amostral e apenas 11,8% focaram no tamanho de efeito exclusivamente. Já 35% dos artigos

avaliaram dois ou mais fatores (ora tamanho amostral e tamanho de efeito, ora tamanho amostral e poder, ora poder e tamanho de efeito).

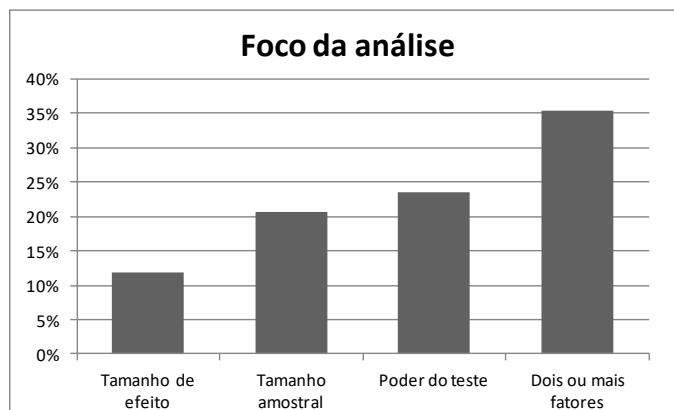


Figura 2: Tipo de análise de poder utilizado nos artigos analisados no presente estudo.

Dos estudos avaliados, 50% utilizaram o nível de 80% de poder como sendo o valor adequado para os testes estatísticos, 15% utilizaram o nível de 90% ou mais e 35% não declararam o valor de poder considerado adequado para um teste estatístico (Figura 3).

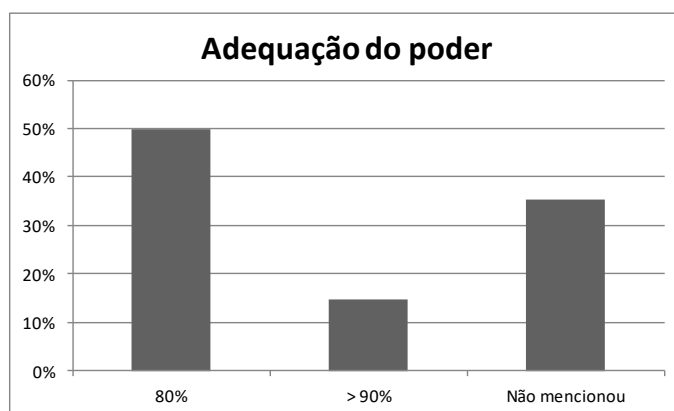


Figura 3: Níveis de poder considerados adequados nos estudos ambientais analisados no presente estudo.

Os testes estatísticos mais utilizados foram a ANOVA e o teste t (em 29,4 e 23,5% dos estudos respectivamente), muitas vezes ajustados dentro de um monitoramento BACI (Before-After Control-Impact), seguidos por análises por simulação (17,6%) e pelos testes de regressão linear (11,7%) (Figura 4). Cinco estudos (14,7%) utilizaram-se da análise de poder para avaliar testes não paramétricos – Wilcoxon Mann-Whitney e Kruskal-Wallis (VAN WYNSBERGE et al., 2013; GEWURTZ et al., 2012; FANG et al., 2009; DIKOU et al, 2009; FRANCO et al, 2015). Como alguns estudos utilizaram mais de um tipo de teste, não foi possível elaborar o gráfico em

percentuais, pois o valor total excedia os 100%. Desta forma, a Figura 4 foi apresentada com os dados brutos obtidos.

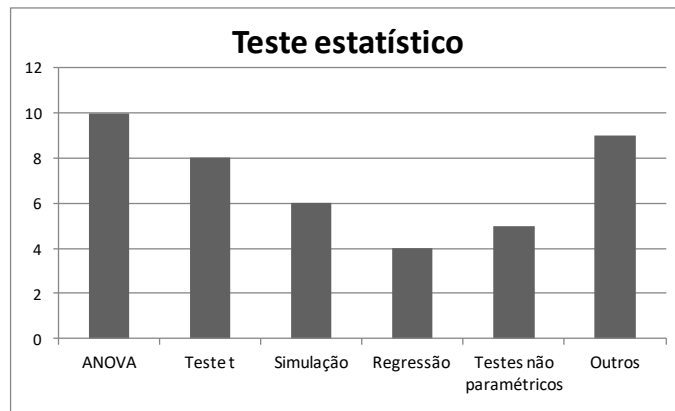


Figura 4: Testes estatísticos empregados nos estudos ambientais analisados no presente estudo.

Dentre os programas utilizados para executar as análises de poder, diversos foram mencionados, sendo o software livre R o mais citado. Além deste, foram utilizados os seguintes programas: SAS, GPower, Systat, Primer, Monitor, Trends, Excel, SPSS e Statistica. Doze estudos realizaram cálculos manualmente ou não mencionaram o software utilizado para realizar as análises e por isso foram agrupados na categoria “outros” no gráfico abaixo (Figura 5: 5).

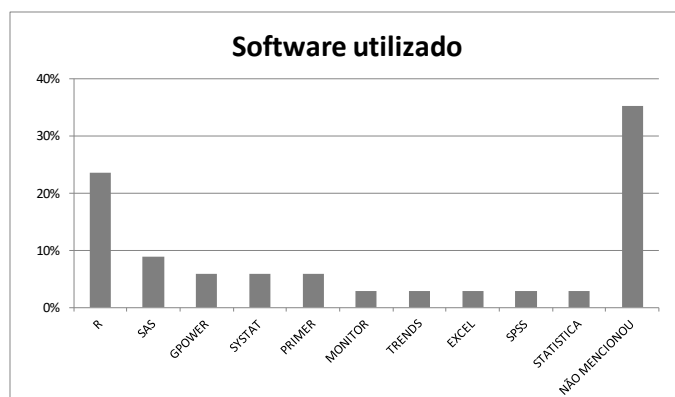


Figura 5: Softwares utilizados para realização das análises de poder nos estudos aqui analisados.

As alterações ambientais de origem antrópica que basearam os estudos aqui analisados foram bastante diversas (Figura 6) **Erro! Fonte de referência não encontrada.** s mais frequentes, contudo, foram poluição, a geração de energia (categoria que agrupou a implantação e operação de usinas eólicas, nucleares e elétricas), a exploração de recursos variados (madeira, pesca, turismo, ovos de crocodilo e óleo)

e o estabelecimento de reservas ecológicas, que representam impactos positivos sobre o ecossistema.

Segundo a norma ISO 14.001:2004 impacto ambiental é “qualquer modificação do meio ambiente, adversa ou benéfica, que resulte, no todo ou em parte, das atividades, produtos ou serviços de uma organização”. Sánchez (2013) sintetizou o conceito de avaliação de impacto ambiental com a necessidade de comparação entre duas situações hipotéticas futuras, uma com a presença do empreendimento ou medida proposta e outra sem. Sendo assim, o estabelecimento de uma reserva ecológica pode ser definido como impacto ambiental, por alterar a qualidade de um habitat em relação ao que se esperaria se nenhuma medida fosse tomada.

As outras fontes de impacto estudadas nos trabalhos aqui analisados, porém em menor quantidade, foram a perda de habitats, emissários submarinos, tráfego de embarcações, dragagem, ameaças ao ecossistema (que inclui a perda de habitats e urbanização), a restauração de rios, tráfego de embarcações e dragagem.

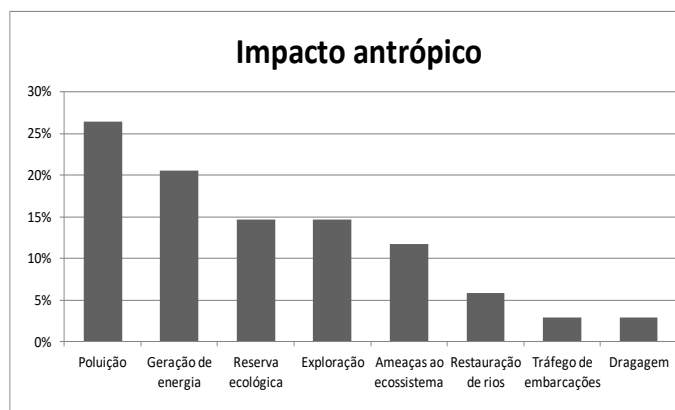


Figura 6: Fontes de alteração antrópica sobre os ecossistemas e indicadores biológicos nos estudos avaliados.

Quanto aos indicadores biológicos utilizados, os grupos predominantemente utilizados os invertebrados aquáticos (38,2%), especialmente as comunidades bentônicas, e os peixes (35,3%). Alguns estudos avaliaram mais de um indicador (invertebrados aquáticos e peixes), por isso foram incluídos separadamente no gráfico demonstrativo. Outros grupos que também foram utilizados mais de uma vez na análise foram as aves (14,7%), as algas (5,9%) e os mexilhões (5,9%). Foram observados alguns estudos em espécies terrestres, tanto animais quanto vegetais,

porém, por terem sido citados apenas uma vez cada, foram agrupados na categoria “outros”, representando 11,8% do total (Figura 7 **Erro! Fonte de referência não encontrada.**). Como alguns estudos utilizaram mais de um tipo de indicador biológico, não foi possível elaborar o gráfico em percentuais, pois o valor total excedia os 100%. Desta forma, a Figura 6 foi apresentada com os dados brutos obtidos.

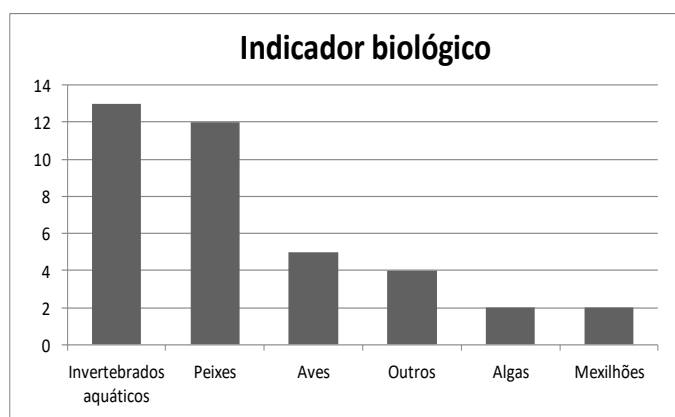


Figura 7: Indicadores biológicos utilizados nos estudos avaliados. Alguns estudos utilizaram mais de um desses indicadores.

5. DISCUSSÃO

A determinação do tamanho amostral ideal em um estudo ambiental é de extrema importância, principalmente por razões econômicas: um estudo subdimensionado pode ser um desperdício de dinheiro por não ser capaz de produzir resultados úteis, enquanto que um estudo sobredimensionado utiliza mais recursos do que o necessário. Lenth (2001) e Thomas & Krebs (1997) defendem que a análise de poder é mais útil quando utilizada no planejamento dos estudos, auxiliando na decisão sobre o tamanho amostral e o nível de significância a serem utilizados. Contudo, poucos trabalhos tem se preocupado com essa questão, como foi evidenciado no presente estudo.

Alguns autores defendem a grande importância da determinação do tamanho de efeito no início do planejamento de um delineamento amostral. Steidl et al. (1997) define o “tamanho de efeito” como sendo o “efeito observado” (diferença entre médias das populações tratamento e controle, $|\mu_T - \mu_C|$) ponderado pela variância através do desvio padrão ($|\mu_T - \mu_C|/\sigma$). Para Ortiz (2002), estimar a magnitude da

perturbação (ou tamanho de efeito) é essencial para aumentar a robustez do teste estatístico bem como para otimizar os custos do programa de amostragem. Rotenberry e Wiens (1985) defendem que a maneira como os resultados da análise de poder são apresentados seja alterada, de modo a focar a atenção no tamanho de efeito e não no poder propriamente dito. Dos artigos analisados, apenas 11,8% avaliaram o tamanho de efeito como foco central das análises.

Apenas sete trabalhos analisaram o tamanho amostral ideal através de análise de poder. Outros 12 trabalhos avaliaram dois ou mais tipos de análise de poder, sendo que a maioria destes (92%, ou 12 estudos) tinha como um dos seus focos centrais a análise do tamanho amostral. Contudo, a maioria dos estudos realizou as análises de poder somente após a obtenção dos resultados, ou seja, fizeram análise de poder *a posteriori*.

Destes trabalhos que realizaram análise de poder *a posteriori*, avaliou-se ora o tamanho do efeito encontrado ora a adequabilidade do tamanho amostral. A importância deste tipo de análise reside no fato que possibilita fazer uma avaliação dos resultados da análise de modo a melhorar o delineamento amostral no caso da continuidade do monitoramento ou para servir de embasamento para trabalhos futuros na mesma linha de pesquisa. Esse tipo de análise é útil após a finalização de estudos, no caso da obtenção de resultados não significativos (THOMAS; KREBS, 1997)

No caso de um resultado não significativo, a análise de poder estatístico *post hoc* serve para determinar a probabilidade de erroneamente ter aceitado a hipótese nula de não diferença (cometido um erro tipo II), adicionando confiança aos resultados negativos obtidos. Além disso, análises de poder *a posteriori* podem também ser usadas para mostrar quando uma amostragem ineficiente pode ter ocorrido, auxiliando assim os gestores ambientais a contabilizar melhor seus esforços amostrais (LOUGHEED; BREAUULT; LANK, 1999).

Foi observada ainda a utilização da análise de poder para avaliar tendências populacionais. Neste caso, alguns trabalhos focaram em avaliar o poder para detectar uma tendência de declínio ou de crescimento de determinado tamanho pré-definido (X% em Y anos), tal como em Zangrandi et al. (2005). Já Cantor et al.

(2012) utilizou a análise de poder para avaliar quantas amostras seriam necessárias para detectar uma tendência de X% em Y anos com poder pré-definido. Além dessas duas possibilidades de uso, Maxwell e Jennings (2005) fez uma análise adicional: avaliou qual o tamanho da tendência que poderia ser detectada em determinado período e com determinado poder.

Muitos autores utilizam o valor de 0,80 como o valor adequado para o poder (STEIDL et al., 1997; LANG; BÜHLER, 2012, MILLER; MITCHELL, 2014), no que chamou-se de “convenção cinco-oitenta – 5-80” ($\alpha = 5\%$, poder = 80%) (DI STEFANO, 2003). Este autor frisa a necessidade de se equilibrar os custos relativos dos erros tipo I e tipo II numa análise, especialmente no campo ambiental, quando um erro tipo II costuma ser mais oneroso que um erro tipo I.

No caso de não se conhecer a relação dos custos dos erros tipo I/tipo II, Mudge et al. (2012) e Hinds (1984) defendem que o risco dos dois tipos de erros sejam iguais ($\alpha = \beta$), equilibrando os riscos entre a indústria e o meio ambiente. Cohen (1988) apud Di Stefano (2003) sugere o uso da convenção 5-80 apenas quando o pesquisador não tem uma base para definir o poder adequado, uma vez que o ideal é definir o poder desejado através da ponderação entre os custos dos dois tipos de erros. Dos estudos avaliados, 50% utilizaram como adequado o nível de poder de 80%, enquadrando-se na tradição acima referida.

Os testes estatísticos mais empregados nos estudos avaliados foram a ANOVA e o Teste t de Student, seguidos por análises por simulação de dados e pela Regressão Linear. Isso reflete a maior utilização da análise de poder em testes paramétricos, corroborando Cohen (1992), que abordou a maior simplicidade das análises paramétricas como ponto favorável à utilização da análise de poder.

Diversos softwares foram utilizados para realização das análises de poder, como demonstrado anteriormente. Alguns específicos para este tipo de análise (como o software GPower) e outros mais gerais para tratamento de dados estatísticos (R, Primer, Statistica, SAS e Systat). Alguns ainda direcionados para análises de tendências temporais, como o Monitor e o TRENDS. O mais utilizado foi o R.

Os impactos antrópicos mais avaliados nos estudos aqui analisados foram a poluição e a geração de energia, e os indicadores biológicos mais empregados nas

análises foram os peixes e invertebrados aquáticos. Isso demonstra que as ciências aquáticas são as que mais tem incorporado a análise de poder estatístico dentro da avaliação dos seus estudos.

Análises estatísticas ainda representam um campo meio obscuro entre as ciências ambientais. Algumas das dificuldades encontradas na interrelação entre as ciências naturais e a estatística enumeradas por Mueller (1992) e que representam obstáculos ainda nos dias de hoje são:

- “- A obtenção de um conjunto adequado de dados ambientais é dispendiosa e consome muito tempo;*
- A maioria dos dados ambientais são multivariados, não se prestando a análises clássicas. Muitos são autocorrelacionados no espaço e no tempo e, frequentemente, são vários os determinantes de um fenômeno ambiental. Suas interações tendem a ser complexas, dificultando o desenvolvimento de modelos eficazes;*
- O ambiente natural é frequentemente incontrolável e guiado por mecanismos ainda não inteiramente identificados;*
- A comunicação entre os estatísticos, os cientistas e entendidos na área do meio ambiente ainda é incipiente.”*

Alia-se a isso ainda o fato que a formação dos pesquisadores e profissionais das áreas ambientais carece de um maior aprofundamento nas áreas exatas e estatísticas, de modo a tornar a interdisciplinaridade entre essas áreas uma prática natural. Essa má formação dos cientistas dificulta a implementação de análises estatísticas para além do senso comum, podendo ser um dos fatores dificultadores da disseminação da análise de poder em estudos ambientais.

6. CONSIDERAÇÕES FINAIS

A despeito da grande importância da aplicação da análise de poder em testes estatísticos, o que se observa é que ela ainda é muito incipiente na área ambiental, sendo bastante subutilizada. A consequência é que dentro de um universo tão amplo de pesquisa e trabalhos técnicos na área ambiental, pouquíssimos estudos se

preocuparam em avaliar se os dados são de fato capazes de inferir sobre a não existência de um impacto com um adequado grau de segurança.

Além disso, grande parte dos estudos que utilizaram a análise de poder o fizeram posteriormente à análise dos dados, de modo que não foi feito um planejamento adequado, só sendo possível concluir sobre a qualidade dos dados e dos estudos realizados. A análise *post hoc* possibilita apenas inferir se o teste realizado foi suficientemente poderoso ou não, mas não é capaz de alterar o resultado.

Os testes mais empregados aliados à análise de poder foram o Teste t e a Análise de Variância (ANOVA) e, de modo geral, a ecologia aquática é a que mais tem incorporado a análise de poder, sendo os peixes e invertebrados aquáticos os grupos mais utilizados. Dentre os softwares mais utilizados estão o R, seguido pelo GPower, SYSTAT, PRIMER e SAS.

Percebeu-se a necessidade de difundir ainda mais a aplicação da análise de poder na fase de planejamento dos estudos, obtendo delineamentos mais adequados como forma de obter conclusões mais consistentes evitando gastos desnecessários e assim, melhor embasando a tomada de decisões pelos gestores ambientais.

7. REFERÊNCIAS BIBLIOGRÁFICAS

Blomberg, S. P. (2014). *Power Analysis Using R*, 13.

Cantor, M., Wedekin, L. L., Daura-Jorge, F. G., Rossi-Santos, M. R., & Simões-Lopes, P. C. (2012). Assessing population parameters and trends of Guiana dolphins (*Sotalia guianensis*): An eight-year mark-recapture study. *Marine Mammal Science*, 28(1), 63–83. <http://doi.org/10.1111/j.1748-7692.2010.00456.x>

Cohen, J. (1992). Statistical Power Analysis. *Current Directions in Psychological Science (Wiley-Blackwell)*, 1, 98–101. <http://doi.org/10.1111/1467-8721.ep10768783>

Di Stefano, J. (2003). Functional Ecology\rHow much power is enough? Against the development of an arbitrary convention for statistical power calculations. *Functional Ecology*, 17, 707–709.

Dikou, A., Ackerman, C., Banks, C., Dempsey, A., Fox, M., Gins, M., ... Thomas, L. (2009). Ecological assessment to detect imminent change in coral reefs of Admiral Cockburn Land and Sea National Park, Turks and Caicos Islands. *Marine Ecology*, 30(4), 425–436. <http://doi.org/10.1111/j.1439-0485.2009.00291.x>

Fairweather, P. G. (1991). Statistical power and design requirements for environmental monitoring. *Australian Journal of Marine and Freshwater Research*, 42(5), 555–567. <http://doi.org/10.1071/MF9910555>

Fang, J. K. H., Wu, R. S. S., Yip, C. K. M., & Shin, P. K. S. (2009). Power analysis for biomarkers in mussels for use in coastal pollution monitoring. *Marine Pollution Bulletin*, 58(8), 1152–1158. <http://doi.org/10.1016/j.marpolbul.2009.04.003>

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.

Franco, A., Quintino, V., & Elliott, M. (2015). Benthic monitoring and sampling design and effort to detect spatial changes: A case study using data from offshore wind farm sites. *Ecological Indicators*, 57, 298–304. <http://doi.org/10.1016/j.ecolind.2015.04.040>

Gewurtz, S. B., De Silva, A. O., Backus, S. M., McGoldrick, D. J., Keir, M. J., Small,

- J., ... Muir, D. C. G. (2012). Perfluoroalkyl contaminants in Lake Ontario Lake Trout: Detailed examination of current status and long-term trends. *Environmental Science and Technology*, 46(11), 5842–5850. <http://doi.org/10.1021/es3006095>
- Glasby, T. M. (1997). Analysing data from post-impact studies using asymmetrical analyses of variance: A case study of epibiota on marinas. *Austral Ecology*, 22(4), 448–459. <http://doi.org/10.1111/j.1442-9993.1997.tb00696.x>
- Gotelli, N., & Ellison, A. M. (2011). *Principios de Estatística em Ecologia*. (Artmed, Ed.). Porto Alegre.
- Green, R. H. (1989). Power analysis and practical strategies for environmental monitoring. *Environmental Research*, 50(1), 195–205. [http://doi.org/http://dx.doi.org/10.1016/S0013-9351\(89\)80058-1](http://doi.org/http://dx.doi.org/10.1016/S0013-9351(89)80058-1)
- Hanson, N. (2011). Using biological data from field studies with multiple reference sites as a basis for environmental management: The risks for false positives and false negatives. *Journal of Environmental Management*, 92(3), 610–619. <http://doi.org/10.1016/j.jenvman.2010.09.026>
- Hinds, W. T. (1984). Towards Monitoring of Long-term Trends in Terrestrial Ecosystems. *Environmental Conservation*, 11(01), 11–18. <http://doi.org/doi:10.1017/S0376892900013448>
- Lang, A., & Bühler, C. (2012). Estimation of required sampling effort for monitoring the possible effects of transgenic crops on butterflies: Lessons from long-term monitoring schemes in Switzerland. *Ecological Indicators*, 13(1), 29–36. <http://doi.org/10.1016/j.ecolind.2011.05.004>
- Lenth, R. V. (2001). Some practical guidelines for effective sample size determination. *The American Statistician*, 55(3), 187–193. <http://doi.org/10.1198/000313001317098149>
- Lougheed, Breault, A. & D. B., & Lank. (1999). Estimating Statistical Power to Evaluate Ongoing Waterfowl Population Monitoring Author (s): Lynn W . Lougheed , André Breault and David B . Lank Source : The Journal of Wildlife Management , Vol . 63 , No . 4 (Oct . , 1999) , pp . 1359-1369 Published b. *The Journal of Wildlife Management*, 63(4), 1359–1369.

Mapstone, B. D. (1995). Scalable Decision Rules for Environmental Impact Studies : Effect Size , Type I , and Type II Errors Author (s): Bruce D . Mapstone Published by : Wiley Stable URL : <http://www.jstor.org/stable/1942031> Accessed : 18-05-2016 00 : 44 UTC Your use of the J. *Ecological Applications*, 5(2), 401–410.

Maxwell, D., & Jennings, S. (2005). Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. *Journal of Applied Ecology*, 42(1), 25–37. <http://doi.org/10.1111/j.1365-2664.2005.01000.x>

Miller, K. M., & Mitchell, B. R. (2014). A new tool for power analysis of fixed plot data : Using simulations and mixed effects models to evaluate forest metrics, 5(September), 1–23.

Morrisey, D. J. (1993). Environmental impact assessment - A review of its aims and recent developments. *Marine Pollution Bulletin*, 26(10), 540–545.

Mudge, J. F., Barrett, T. J., Munkittrick, K. R., & Houlahan, J. E. (2012). Negative consequences of using $\alpha = 0.05$ for Environmental monitoring decisions: A case study from a decade of Canada's environmental effects monitoring program. *Environmental Science and Technology*, 46(17), 9249–9255. <http://doi.org/10.1021/es301320n>

Mueller, C. C. (1992). Situação atual da produção de informações sistemáticas sobre o meio ambiente. *Revista Ciência Da Informação*, 21, 14–22.

Ortiz, M. (2002). Optimum Sample Size to Detect Perturbation Effects : The Importance of Statistical Power Analysis ± A Critique, 23(1), 1–9.

Osenberg, C. W., Schmitt, R. J., Holbrook, S. J., Abu-Saba, K. E., & Flegal, R. (1994). Detection of environmental impacts: Natural variability, effect size, and power analysis. *Ecol. Appl.*, 4(1), 16–30.

Pais, M. P., Henriques, S., Costa, M. J., & Cabral, H. N. (2014). Topographic complexity and the power to detect structural and functional changes in temperate reef fish assemblages: The need for habitat-independent sample sizes. *Ecological Indicators*, 45, 18–27. <http://doi.org/10.1016/j.ecolind.2014.03.018>

Peterman, R. (1990). The Importance of Reporting Statistical Power : The Forest

Decline and Acidic Deposition Example Author (s): Randall M . Peterman Published by : Wiley Stable URL : <http://www.jstor.org/stable/1937612> REFERENCES Linked references are available on JSTOR f. *Ecology*, 71(5), 2024–2027.

Rotenberry, J. T., & Wiens, J. A. (1985). Statistical Power Analysis and Community-Wide Patterns, *125*(1), 164–168.

Sánchez, L. E. (2013). *Avaliação de Impacto Ambiental: conceitos e métodos. Oficina de Textos* (Vol. 2 Ed).

Steidl, R. J., Hayes, J. P., & Schaubert, E. (1997). Statistical power analysis in wildlife. *Journal of Wildlife Management*, 61(2), 270–279.

Steinbeck, J. R., Schiel, D. R., & Foster, M. S. (2005). Detecting long-term change in complex communities: A case study from the rocky intertidal zone. *Ecological Applications*, 15(5), 1813–1832. Retrieved from //000232322600026

Thomas, L., & Krebs, C. (1997). A review of statistical power analysis software. *Bulletin of the Ecological Society of America*, 78(2), 128–139. Retrieved from <http://scholar.google.com/scholar?q=intitle:A+REVIEW+OF+STATISTICAL+POWER+ANALYSIS+SOFTWARE#0>

Toft, C. A., & Shea, P. J. (1983). Detecting community-wide patterns: Estimating power strengthens statistical inference. *The American Naturalist*, 122(5), 618–625. <http://doi.org/10.1086/521238>

Van Wynsberge, S., Gilbert, A., Guillemot, N., Payri, C., & Andréfouët, S. (2013). Alert thresholds for monitoring environmental variables: A new approach applied to seagrass beds diversity in New Caledonia. *Marine Pollution Bulletin*, 77(1-2), 300–307. <http://doi.org/10.1016/j.marpolbul.2013.09.035>

Zangrandi, M., Catsiki, V.-A., & Avdis, O. L. V. (2005). The sensitivity of mussel-watch programmes to detect temporal trends in the bio-availability of metals. A case study of the Saronicos Gulf of Greece. *Journal of Environmental Monitoring : JEM*, 7(8), 798–803. <http://doi.org/10.1039/b419326a>

APÊNDICE

ID	Primeiro Autor	Título	Revista	Ano
1	Schroeter SC	Detecting the ecological effects of environmental impacts - a case-study of kelp forest invertebrates	Ecological applications	1993
2	Hyland J	Environmental impact of offshore oil development on the outer continental shelf and slope off Point Arguello, California	Marine Environmental Research	1994
3	Otway NM	Assessing the impacts of deepwater sewage outfalls on spatially- and temporally-variable marine communities	Marine environmental research	1996
4	Crome FHJ	A novel Bayesian approach to assessing impacts of rain forest logging	Ecological Applications	1996
5	Camphuysen KCJ	Beached Bird Surveys Indicate Decline in Chronic Oil Pollution in the North Sea	Marine Pollution Bulletin	1998
6	Sandim L	The statistical power of selected indicator metrics using macroinvertebrates for assessing acidification and eutrophication of running waters	Hydrobiologia	2000
7	Smith WP	Marbled murrelet surveys: site and annual variation, sampling effort, and statistical power	Wildlife Society Bulletin	2001
8	Stirrat SC	Monitoring <i>Crocodylus porosus</i> populations in the Northern Territory of Australia: a retrospective power analysis	Wildlife Research	2001
9	Carey JM	The variability of estimates of variance, and its effect on power analysis in monitoring design	Environmental Monitoring and Assessment	2002
10	Newmark WD	Power to detect trends in ecological indicators in the East Usambara Mountains, Tanzania	African Journal of Ecology	2003
11	Willis TJ	Protection of exploited fish in temperate regions: high density and biomass of snapper <i>Pagrus auratus</i> (Sparidae) in northern New Zealand marine reserves	Journal of applied ecology	2003
12	Maxwell D	Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish	Journal of Applied Ecology	2005
13	Steinbeck JR	Detecting long-term change in complex communities: A case study from the rocky intertidal zone	Ecological applications	2005
14	Zangrandi M	The sensitivity of mussel-watch programmes to detect temporal trends in the bio-availability of metals. A case study of the Saronicos Gulf of Greece	J. Environ. Monit.	2005
15	Lincoln-Smith MP	Using impact assessment methods to determine the effects of a marine reserve on abundances and sizes of valuable tropical invertebrates	Canadian Journal of Fisheries and Aquatic Sciences	2007
16	de Loma TL	A framework for assessing impacts of marine protected areas in Moorea (French Polynesia)	Pacific Science	2008
17	Baldigo BP	Detecting the response of fish assemblages to stream restoration: Effects of different sampling designs	North American Journal of Fisheries Management	2008
18	Dikou A	Ecological assessment to detect imminent change in coral reefs of Admiral Cockburn Land and Sea National Park, Turks and Caicos Islands	Marine Ecology - An Evolutionary Approach	2009
19	Fang JKH	Power analysis for biomarkers in mussels for use in coastal pollution monitoring	Marine Pollution Bulletin	2009
20	McMullin VA	Spatial Variability of Reproductive and Size Characteristics of the Northern Mummichog (<i>Fundulus heteroclitus macrolepidotus</i>) Collected Near Municipal Wastewater Discharges	Water Qual. Res. J. Can.	2010
21	Styan CA	Is monitoring for mass spawning events in coral assemblages in north Western Australia likely to detect spawning?	Marine Pollution Bulletin	2012
22	Pande A	The Kapiti Marine Reserve (New Zealand): spatial and temporal comparisons of multi-species responses after 8 years of protection	New Zealand Journal of Marine Research	2012
23	Cantor M	Assessing population parameters and trends of Guiana dolphins (<i>Sotalia guianensis</i>): An eight-year mark-recapture study	Marine Mammal Science	2012
24	MacLean IMD	Evaluating the statistical power of detecting changes in the abundance of seabirds at sea	The International Journal of Avian Research	2012
25	Gewurtz SB	Perfluoroalkyl contaminants in Lake Ontario Lake Trout: Detailed examination of current status and long-term trends	Environmental Science & Technology	2012
26	Van Wynsberg A	Alert thresholds for monitoring environmental variables: A new approach applied to seagrass beds diversity in New Caledonia	Marine Pollution Bulletin	2013
27	Maclean IMD	Evaluating the statistical power of detecting changes in the abundance of seabirds at sea	IBIS	2013
28	Hale R	Assessing the likely responses by fishes to stream bank rehabilitation in a large, urban estuary	Austral Ecology	2013
29	Franco A	Benthic monitoring and sampling design and effort to detect spatial changes: A case study using data from offshore wind farm sites	Ecological indicators	2015
30	Vanermen N	Assessing seabird displacement at offshore wind farms: power ranges of a monitoring and data handling protocol	Hydrobiologia	2015
31	Hitt NP	Threshold-dependent sample sizes for selenium assessment with stream fish tissue	Integr. Environ. Assess. Manag.	2015
32	Vaudor L	How sampling influences the statistical power to detect changes in abundance: an application to river restoration	Freshwater biology	2015
33	Dolan, TE	Statistical power to detect change in a mangrove shoreline fish community adjacent to a nuclear power plant	Environmental Monitoring and Assessment	2016
34	Buxton RT	Monitoring burrowing petrel populations: A sampling scheme for the management of an island keystone species	Journal of wildlife management	2016

CAPÍTULO II

Guia para análise de poder aplicada a estudos ambientais

Salvador, Bahia

2016

GUIA PARA ANÁLISE DE PODER APLICADA A ESTUDOS AMBIENTAIS

Marina Pessoa Felzemburgh Motta, Gilson Correia de Carvalho

RESUMO

Com o intuito de incentivar o uso da análise de poder de testes estatísticos aplicados a estudos ambientais na área técnica, o presente trabalho elaborou um protocolo simplificado com um passo a passo das análises mais comuns em estudos ambientais. Foi utilizado o software GPower, por este ter sido considerado o software gratuito de mais fácil utilização e, portanto, mais propício a incorporação por técnicos e pesquisadores da área ambiental. As análises demonstradas no protocolo foram o Teste T e a ANOVA, ambas com exemplos com dados simulados através do Microsoft Excel. Foram apresentados exemplos de análises a priori e post hoc. Pôde-se observar o baixo poder que esses testes possuem ao utilizar o nível de significância convencional de 0,05, sendo necessário um tamanho amostral bastante elevado para atingir resultados satisfatórios em termos de minimização dos riscos de erros. Verificou-se ainda a importância em equilibrar as probabilidades de erros tipo I e tipo II, equalizando os riscos entre o empreendedor e o meio ambiente.

ABSTRACT

In order to encourage the usage of analysis of power of statistical tests applied to environmental studies in technical area, this paper produced a simplified protocol with a walkthrough of the most common analyzes in environmental studies. GPower software has been used as we considered it the easiest one within all free softwares available, therefore more conducive to incorporation by technicians and researchers in the environmental area. The analyzes shown in the protocol were the T-test and ANOVA, both examples with simulated data using Microsoft Excel. It was also shown examples of a priori and post hoc analysis. It was observed low power of these tests when using a level of significancy of 0.05 has been necessary a fairly large sample size to achieve satisfactory results in terms of minimizing the risk of error. It was also noted the importance of balancing the chances of errors type I and type II, this way equaling the risks between the entrepreneur and the environment.

1. INTRODUÇÃO

A partir do capítulo 1 desta dissertação, verificou-se quais foram os softwares e as análises mais utilizados nos estudos ambientais. Com base nisso, foi elaborado um protocolo de como realizar análise de poder, com exemplos a partir de situações hipotéticas, de modo a incentivar a incorporação da análise de poder nas rotinas de pesquisas ambientais.

O software mais utilizado para análises de poder nos trabalhos analisados foi o R. O R é um software livre e gratuito utilizado para análises estatísticas e gráficas. Possui algumas possibilidades de análise de poder e, por ser um software livre, dá ao usuário a liberdade de criar scripts para realização de análises não disponíveis no programa. Software livre não é sinônimo de software gratuito, na medida em que por “gratuito” entende-se a não necessidade de pagamento pelo uso, enquanto “livre” refere-se à liberdade de utilização do software sendo permitido ao usuário não só utilizar, mas também copiar, distribuir, modificar e estudar o software (Infowester, 2016).

Segundo Blomberg (2014), o software R possui em sua estrutura básica um pacote chamado **stats**, o qual possui funções para realização de análises de poder para testes t, ANOVA e testes de proporções. Contudo, sua saída é bem simples, uma vez que ele apenas calcula o 4º parâmetro relativo à análise de poder a partir dos outros 3 fornecidos pelo usuário (tamanho amostral, tamanho de efeito, nível de significância e poder). Para outros tipos de análise mais rebuscadas é necessária a instalação de pacotes adicionais, sendo que os indicados pelo autor supracitado são o **pwr** e o **powerSurvEpi**, os quais acrescentam diversas funções ao R.

Outro software livre e gratuito disponível para realização de análises de poder utilizado em alguns dos trabalhos aqui analisados foi o G*Power. O G*Power é um software de análise de dados gratuito, de interface bastante simples e de fácil utilização, específico para análises de poder. Além das tradicionais análises *a priori*, utilizadas para determinar o tamanho amostral ideal com determinado poder pré-definido, e *post hoc*, empregada para determinar o poder do teste realizado, o G*Power 3 traz a possibilidade de três outros tipos de análise de poder: análise de compromisso, análise de sensibilidade e análise de critério.

Segundo Faul et al (2007), a análise de compromisso é utilizada quando o N a ser aplicado é muito pequeno para atingir os níveis de α e β desejados ou quando o N é tão grande que qualquer pequeno efeito negligenciável seria suficiente para rejeitar H_0 . Esta análise serve para equilibrar os riscos de erro tipo I e tipo II sugerindo o maior tamanho amostral possível dentro da disponibilidade de tempo e recursos, o que pode representar uma análise de custo-benefício. Ainda segundo este autor, a análise de sensibilidade busca identificar que tamanho de efeito o estudo é capaz de detectar com determinado poder pré-estabelecido, ou seja, identifica qual o menor tamanho de efeito para o qual o teste aplicado é suficientemente sensível e, por último, a análise de critério funciona como uma alternativa à análise *post hoc*, útil quando o controle do β é mais importante do que o controle do α , sendo assim, ela calcula o α como uma função do poder, do tamanho amostral e do tamanho de efeito.

O G*Power em sua terceira versão possui uma boa interface gráfica e saída simples dos resultados. Ele plota automaticamente os gráficos de distribuição central e não-central com o critério de decisão e as probabilidades de erros associadas. Tem também a opção de plotar quaisquer dos quatro parâmetros centrais da análise de poder (α , poder, N e tamanho de efeito) em função de algum outro.

Thomas & Krebs (1997) fizeram uma revisão dos softwares disponíveis para análises de poder, na qual foram analisados 29 softwares em termos do escopo, facilidade de uso e facilidade de aprendizagem. Dos softwares gratuitos analisados por estes autores, o GPower foi o mais bem avaliado.

Outros softwares utilizados nas análises de poder dos trabalhos analisados foram SAS, SYSTAT, PRIMER, STATISTICA. Todos estes, porém, são softwares pagos, apresentando restrições à disseminação da análise de poder e por este motivo não foram analisados no presente trabalho. Os softwares MONITOR e TRENDS também foram utilizados em alguns dos artigos analisados, mas devido ao seu escopo limitado, uma vez que ambos são focados apenas em análises de tendências temporais, os mesmos não foram aqui avaliados.

Apesar do grande número de softwares disponíveis para análise de poder, Thomas & Krebs (1997) discutiram a ainda rara utilização desta análise na literatura

ecológica científica. Quase 20 anos após essa percepção, o poder estatístico ainda permanece negligenciado em estudos ambientais na literatura científica e, principalmente, na área técnica. Por este motivo, o presente trabalho optou por focar em um único programa que fosse gratuito, de fácil utilização e de fácil aprendizado pelos mais diversos usuários, de modo a incentivar a incorporação da análise de poder em estudos ambientais.

Sendo assim, os dois softwares considerados mais adequados para popularizar a análise de poder foram o GPower e o R. Contudo, o R possui uma linguagem própria e uma interface não tão habitual a usuários não familiarizados com este tipo de software. Além disso, ele requer a elaboração de scripts específicos para cada análise realizada, o que exige uma maior habilidade do usuário. Como o objetivo do presente estudo é gerar um protocolo simplificado para auxiliar na disseminação e aplicabilidade da análise de poder, a utilização de um software programável para o meio técnico não seria de fácil aceitação e, portanto, o software R não será explorado. O presente trabalho então, irá apresentar um protocolo com exemplos práticos de utilização do GPower com o intuito de incentivar a utilização das análises de poder em estudos ambientais.

2. OBJETIVOS

Elaborar um protocolo simplificado demonstrando como diferentes tipos de análise de poder são realizadas através do software GPower 3.1. A seguir, realizar algumas análises de poder com os testes estatísticos mais empregados em estudos ambientais conforme verificado no capítulo I do presente estudo, seguindo o passo-a-passo apresentado para demonstrar como os resultados são dados pelo programa.

2.1. Objetivos específicos

- Apresentar um passo-a-passo da análise *a priori* através do GPower;
- Apresentar um passo-a-passo da análise *a posteriori* através do GPower;
- Simular dados para aplicação em análises hipotéticas através da aplicação de teste t e ANOVA;

- Realizar análises utilizando os dados simulados para cada tipo de teste com análise *a priori* e para uma análise *a posteriori*.

3. MATERIAIS E MÉTODOS

Segundo Faul et al. (2007), a análise de poder no G*Power 3 é realizada em 4 passos:

1. Seleção do tipo de teste, a qual pode ser baseada na abordagem de distribuição ou de design. No primeiro tipo, o usuário seleciona a família do teste (testes t, F, z, X^2 ou exatos) e depois o tipo da análise. Na segunda opção, o usuário pode escolher o teste baseado no delineamento, informando ao programa o tipo de análise (médias, variâncias, proporções, regressão e correlação ou genérica) e depois o tipo de delineamento (número de grupos, se amostras dependentes ou independentes, etc);
2. Escolher o tipo de análise de poder, dentre uma das 5 opções disponíveis (*a priori*, *a posteriori*, sensibilidade, compromisso e de critério);
3. Fornecer os parâmetros necessários ao teste, solicitados pelo programa e que variam de acordo com o tipo da análise;
4. Calcular.

Uma vez seguidos esses passos, o G*Power fornece um output dos resultados (na aba "Protocol of Power Analysis"), no qual constam todos os parâmetros de entrada, o tipo de análise realizado e os resultados obtidos: os graus de liberdade, o t crítico e os demais parâmetros que variam de acordo com a análise realizada. Em algumas delas, o resultado apresenta um parâmetro chamado "poder real", apesar do poder desejado ter sido especificado anteriormente. Isso acontece porque o programa aproxima o N amostral para o menor inteiro possível, e com isso o poder se eleva um pouco de modo a garantir que o resultado não seja inferior ao poder previsto.

O tamanho de efeito a ser fornecido ao programa pode ser pré-definido pelo usuário ou calculado pelo G*Power 3. Cohen (1977¹, 1988²) apud GPower Tutorial classificou o tamanho de efeito em categorias (pequeno, médio e grande) para cada

¹ Cohen, J. (1977). *Statistical Power Analysis for the Behavioural Sciences*. Academic Press, New York.

² Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.), New York: Academic Press, New York.

tipo de análise (teste t, ANOVA, Chi-Quadrado, etc), como pode ser observado na Tabela 1 a seguir, as quais podem ser utilizadas no caso de se optar em pré-definir o tamanho de efeito esperado de se detectar. Por outro lado, caso não se queira definir previamente, pode-se calcular o tamanho de efeito observado após a realização de uma amostragem, uma vez que se tenham dados disponíveis através de um estudo piloto ou de literatura, utilizando-se o link “Determine =>” do programa e fornecendo as médias e desvios padrão de cada um dos grupos.

Tabela 1: Valores de tamanho de efeito categorizados para diferentes testes estatísticos.

	PEQUENO	MÉDIO	GRANDE
Teste t de médias	0,20	0,50	0,80
Teste t de correlações	0,10	0,30	0,50
Teste f (ANOVA)	0,10	0,25	0,40
Qui-Quadrado	0,10	0,30	0,50

FONTE: Adaptada de Cohen (1977) *apud* GPower Tutorial.

Para demonstrar como é realizada uma análise de poder no software GPower 3.1 foram realizadas algumas análises demonstrando o passo-a-passo no software através de printscreens das telas do programa. Foram feitas simulações com diferentes níveis de tamanho de efeito, poder e alfa para avaliar como varia o tamanho amostral necessário em diferentes situações e contextos.

Foram escolhidos dois tipos de testes estatísticos, o teste t e a ANOVA, por terem sido estas as análises mais utilizadas nos artigos analisados no capítulo anterior. Dentro de cada uma das análises foram gerados diferentes cenários no GPower, com diferentes combinações dos parâmetros da análise de poder. Para cada cenário foi gerado um ou mais gráfico(s) correspondente(s).

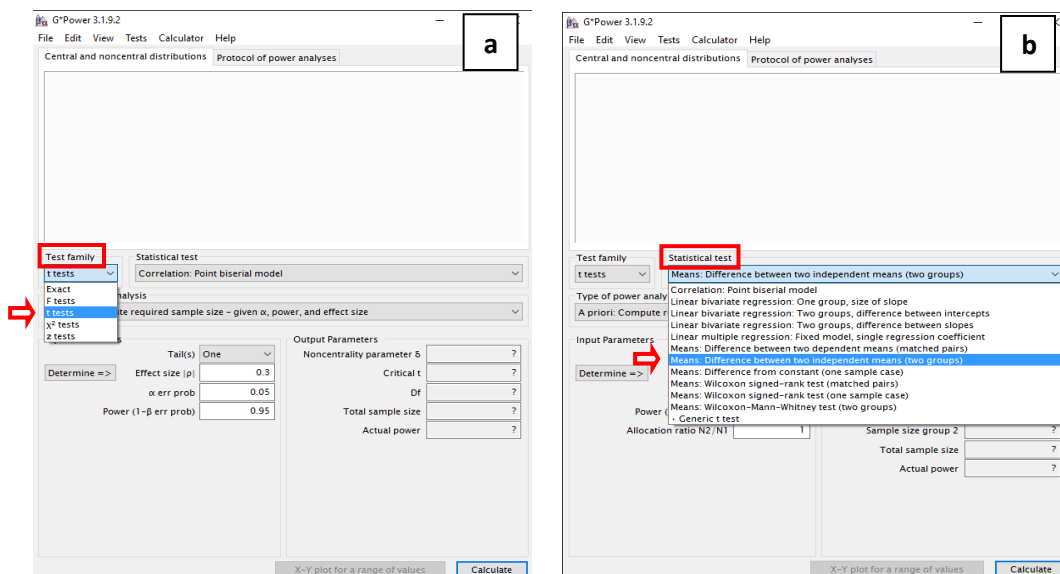
Para simulação dos dados a serem utilizados dentro das situações hipotéticas criadas, foram gerados conjuntos de dados aleatórios através do Microsoft Excel (na opção “análise de dados” → “geração de número aleatório”) de modo a se obter três tamanhos de efeitos distintos (pequeno, médio e grande). Esses dados referem-se à diversidades encontradas numa área de estudo hipotética, onde por exemplo, os códigos C1, C2 e C3 representam as três amostras controle enquanto I1, I2 e I3 representam as três amostras impacto.

4. RESULTADOS

A análise de poder no GPower é bastante simples de ser realizada e o software é fácil de ser manuseado mesmo por usuários sem experiência com o programa. No tópico a seguir é demonstrada a interface do programa ilustrando o passo-a-passo inicial para uma análise de poder.

4.1. Análise de poder *a priori*

Primeiramente é escolhida a família (Figura 1a) e o tipo de teste estatístico (Figura 1b) a ser utilizado. Para este exemplo, foi escolhido um teste t para análise de duas médias independentes. A seguir foi escolhido o tipo da análise de poder, no caso foi escolhida a análise *a priori* (Figura 1c). Por último, é selecionado se o teste é uni ou bicaudal, e são inseridos os parâmetros de entrada no programa (Figura 1d). Para calcular o tamanho amostral necessário em um teste t é necessário informar o tamanho de efeito esperado, o poder e o nível de significância desejados. Caso não se saiba o tamanho de efeito esperado, mas tendo informações das médias das variáveis de análise e de seus desvios padrão através de um estudo piloto ou de literatura pertinente por exemplo, pode-se calcular o tamanho de efeito utilizando a aba “Determine” e inserindo os parâmetros requeridos pelo software (Figura 1e). Colocando-se o mouse sobre o campo do tamanho de efeito, o programa mostra quais os valores convencionados para o tipo de teste escolhido.



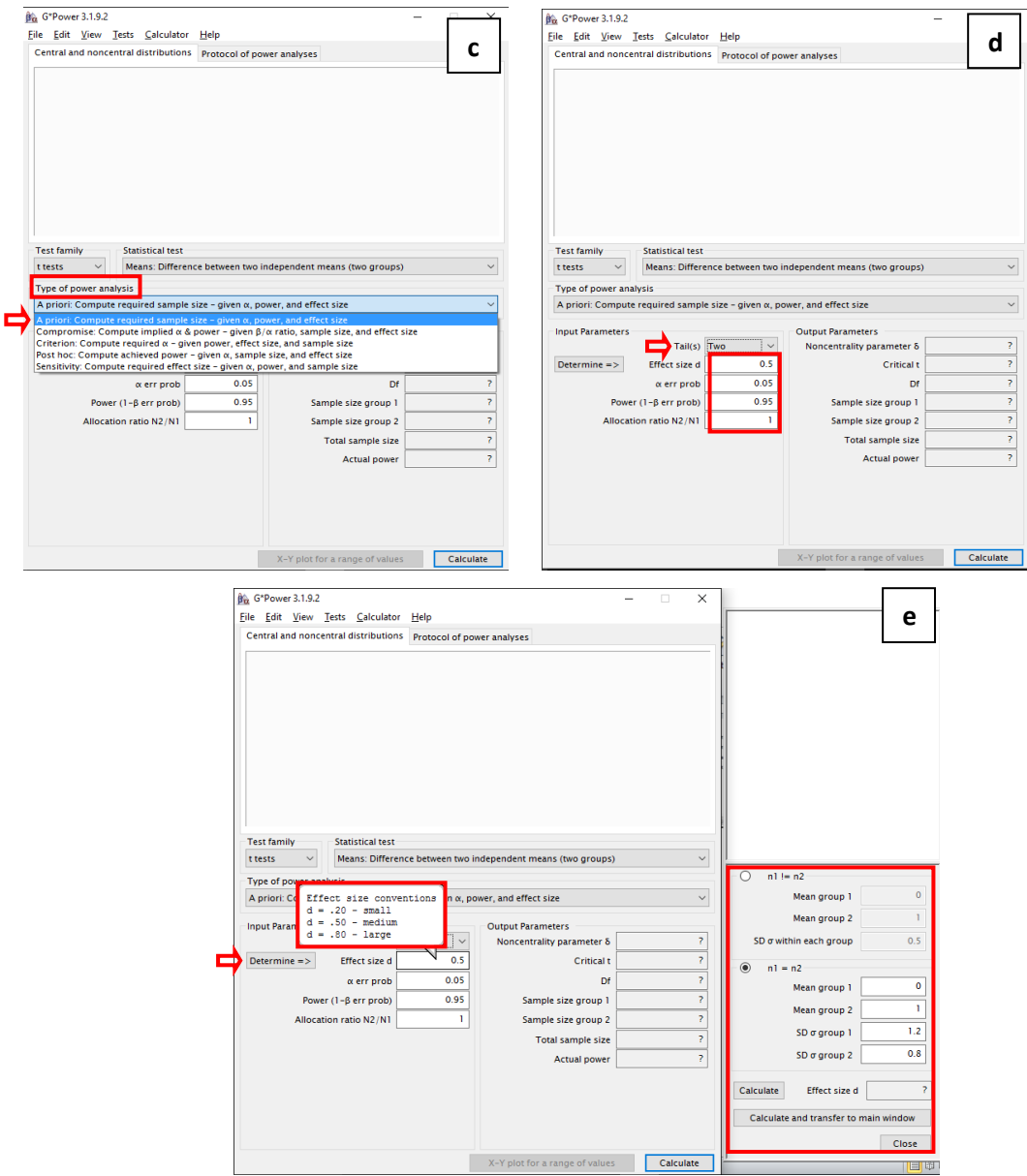


Figura 1: Passo a passo para realizaç o de uma an lise de poder *a priori* no software GPower 3.1.

Outra forma de utilizar o GPower na escolha dos testes   atrav s da aba "Tests" e escolhendo o tipo de an lise desejada. Neste exemplo, foi escolhido o teste de m dias e posteriormente o tipo do teste "muitos grupos e uma vari vel independente – ANOVA 1-fator (ANOVA One-way)". Selecionando desta forma, o pr prio programa se encarrega de escolher a fam lia e o teste apropriado para a an lise (Figura 2).

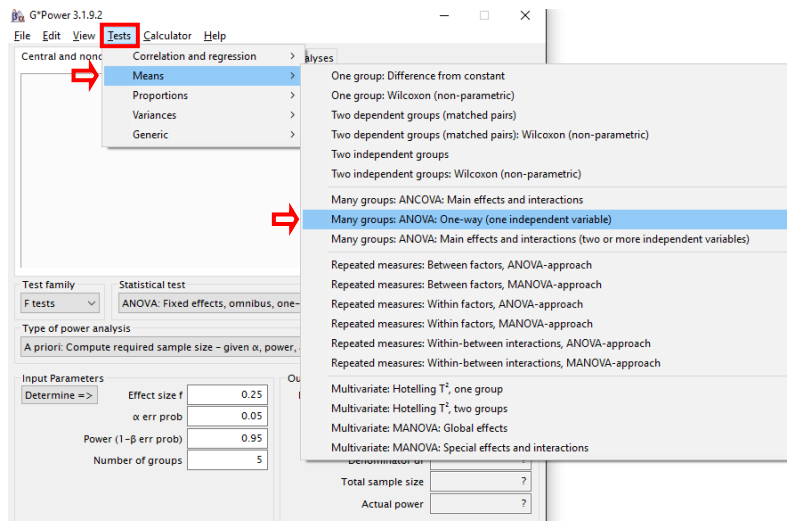


Figura 2: Outra forma de seleção do tipo de teste, através da aba “Tests”.

Tradicionalmente tem se utilizado o valor de 80% como adequado para o poder de testes estatísticos, no que se denomina convenção 5-80, segundo Di Stefano (2003), que ocorre quando se utiliza um α de 0,05 e um poder de 0,8. Contudo, muitos autores defendem a necessidade de se equilibrar o risco dos erros entre a indústria e o meio ambiente através da razão $\alpha/\beta = 1$ (HINDS, 1984; DI STEFANO, 2003; MUDGE et al, 2012). Por este motivo, nas análises aqui apresentadas foram utilizados os valores de poder de 80, 90 e 95% para representar o uso convencional (80%) e valores equilibrados com α de 0,10 e 0,05.

Para demonstrar como são apresentados os resultados da análise de poder no GPower, foi rodada uma análise de poder de teste t com tamanho de efeito médio (0,50), poder de 80% e α de 0,05, conforme a convenção 5-80. A Figura 3 ilustra o output desses resultados.

Primeiramente é apresentado um gráfico com o valor do t crítico e as distribuições central e não-central (na aba “Central and noncentral distributions”). O t crítico define o limite da região de rejeição de H_0 na distribuição central bem como o valor do poder na distribuição não central. Na segunda aba de resultados (“Protocol of power analysis”) são apresentados todos os parâmetros alimentados na análise bem como todos os resultados obtidos, incluindo o t crítico, o número de graus de liberdade o

número de amostras em cada grupo e o poder real obtido (o qual sempre é maior que o poder pré-definido).

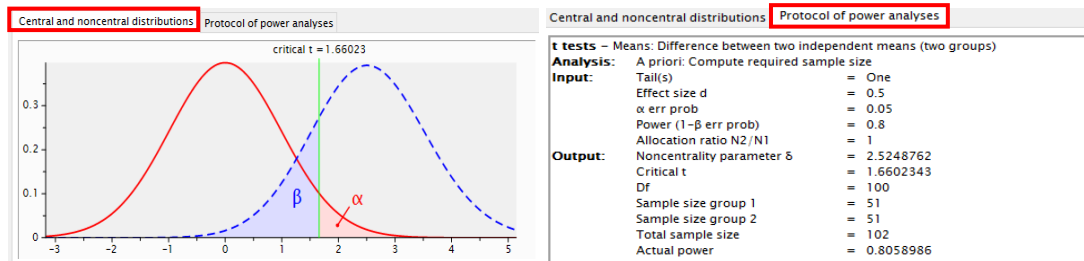


Figura 3: Saída dos resultados.

4.2. Análise de poder *a posteriori*

A análise de poder *a posteriori*, também chamada de *post hoc*, é realizada da mesma forma que a análise *a priori*, apenas selecionando o tipo diferente de análise conforme passo c da Figura 1 apresentada anteriormente. Neste caso, seleciona-se como tipo de análise de poder a análise *post hoc*, como apresentado na Figura 4 a seguir. Variam também os parâmetros de entrada no programa. Para este tipo de análise, é necessário fornecer ao programa o tamanho de efeito, a probabilidade α , o tamanho amostral total e o número de grupos analisados.

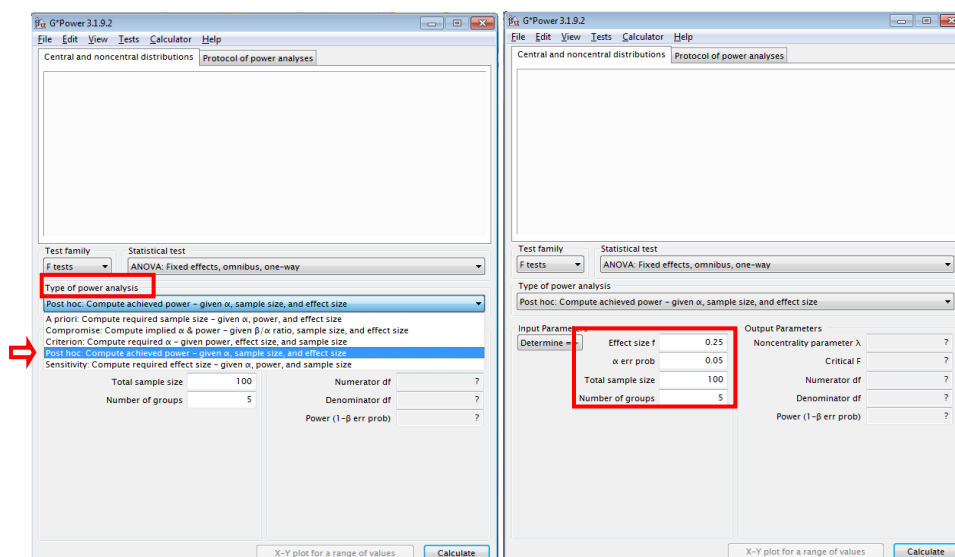


Figura 4: Variação do passo a passo para realização de uma análise de poder *post hoc* no software GPower 3.1.

5. DISCUSSÃO

A discussão a seguir será apresentada na forma de situações hipotéticas, aplicando o passo-a-passo apresentado no item anterior.

Situação hipotética 1:

Supondo-se um estudo sobre o efeito da introdução de uma espécie exótica de coral em um ambiente estuarino onde em alguns dos afluentes o coral invasor já foi observado e encontra-se em expansão e em outros dos seus afluentes não há sinal da presença do coral. Todos os rios encontram-se em condições similares de salinidade, temperatura e demais condições físico-químicas e grau de antropização. Pretende-se investigar se a presença do coral invasor afeta a diversidade de corais nativos na região. Considerando-se três rios com a presença do coral e três sem, deseja-se saber quantas amostras são necessárias para constatar com poder suficientemente alto se a diversidade de corais dos ambientes afetados difere da dos ambientes controle.

Para esta situação deve ser realizado um teste t, para o qual foram feitas simulações com tamanhos de efeito de 0,20, 0,50 e 0,80 (variação de 20, 50 e 80% respectivamente), considerados pequeno, médio e grande para teste t de médias por Cohen (1977) apud GPower Tutorial conforme tabela apresentada anteriormente, poder de 80, 90 e 95% e nível de significância de 0,05 e 0,10. Tais simulações foram feitas com o intuito de verificar de que forma variou o tamanho amostral, numa análise de poder *a priori*. Com os resultados obtidos foi construída uma tabela (Tabela 2) e os gráficos apresentados a seguir.

O N amostral é calculado pelo GPower a partir da combinação dos outros três parâmetros. Sendo assim, o número de amostras de cada grupo (controle x impacto) deve ser metade do N total, o que nem sempre resulta em tamanhos amostrais equivalentes a múltiplos de 3, como estipulado na situação hipotética aqui criada (3 rios por grupo). Neste caso, para obter o N amostral real deve-se aproximar o N amostral por grupo ao menor múltiplo de 3 subsequente ao resultado obtido pelo

programa, o que resultaria em um poder levemente superior ao estipulado pelo usuário.

Tabela 2: Tamanhos amostrais obtidos pela combinação dos outros parâmetros da análise de poder a partir de tamanhos de efeito pré-definidos (situação hipotética 1).

ALFA	TAMANHO DE EFEITO	PODER	N AMOSTRAL TOTAL	N AMOSTRAL POR GRUPO
0,05	0,8	0,8	42	21
0,05	0,8	0,9	56	28
0,05	0,8	0,95	70	35
0,05	0,5	0,8	102	51
0,05	0,5	0,9	140	70
0,05	0,5	0,95	176	88
0,05	0,2	0,8	620	310
0,05	0,2	0,9	858	429
0,05	0,2	0,95	1084	542
0,1	0,8	0,8	30	15
0,1	0,8	0,9	42	21
0,1	0,8	0,95	56	28
0,1	0,5	0,8	74	37
0,1	0,5	0,9	106	53
0,1	0,5	0,95	138	69
0,1	0,2	0,8	452	226
0,1	0,2	0,9	658	329
0,1	0,2	0,95	858	429

Outra funcionalidade do GPower é a geração de gráficos de um parâmetro em função de qualquer outro à sua escolha. Clicando na aba “X-Y plot for a range of values” é aberta uma nova janela onde serão alimentados os parâmetros desejados para o gráfico (marcados em vermelho na aba direita da Figura 5). No caso, como o objeto da análise foi o tamanho amostral, foram gerados gráficos para demonstrar como varia o poder em função do N para diferentes tamanhos de efeito pré-determinados e com o α fixo. Foi selecionada a amplitude do tamanho amostral a ser plotada (de 10 a 1000 amostras), determinada a partir dos resultados obtidos na tabela acima, três plots de tamanho de efeito, de modo a representar efeitos pequeno, médio e grande.

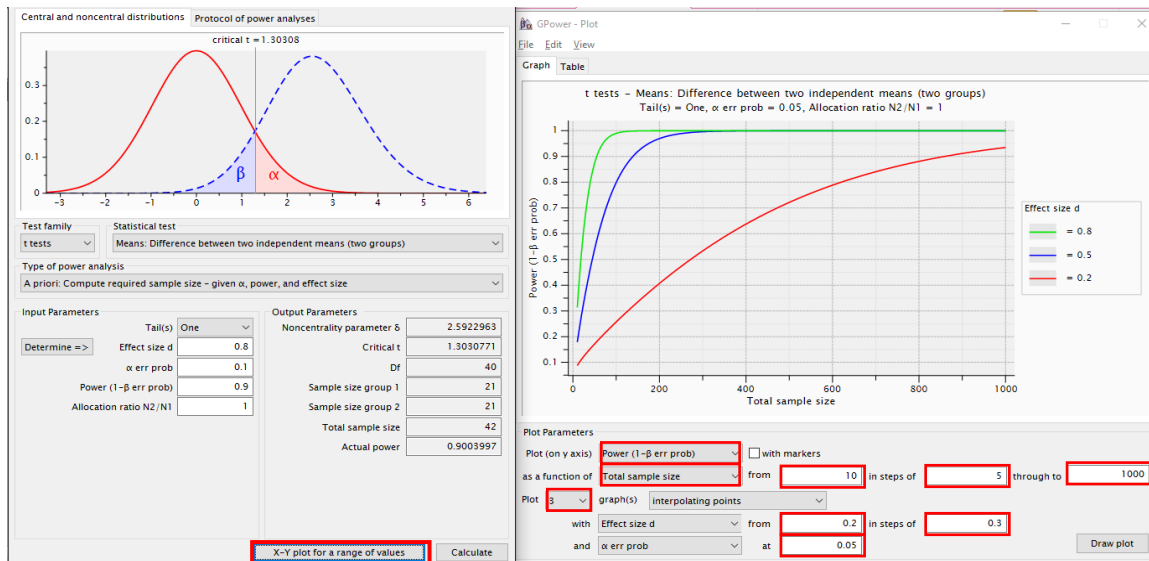


Figura 5: Plotando um gráfico poder x tamanho amostral

Os gráficos a seguir demonstram como variou o poder em função do tamanho amostral para um α de 0,05 e 0,10 para diferentes tamanhos de efeito (Figura 6). Com o primeiro valor de α para um tamanho de efeito grande, com 80% de variação, em qualquer dos níveis de poder analisados seria necessário um n amostral entre 42 e 70 para atingir o resultado esperado. Já para um tamanho de efeito intermediário (50% de variação) são necessárias entre 102 e 176 amostras. Por outro lado, para detectar um tamanho de efeito pequeno (20%), em qualquer dos níveis de poder avaliados, o tamanho amostral necessário é bastante alto, com um total de amostras variando de 620 a 1084. Tais valores são praticamente inviáveis para um estudo ambiental, o que significa que um teste t tem baixo poder pra detectar pequenos efeitos num determinado estudo, uma vez que necessitaria de um número muito grande de amostras. Para α de 0,10, o tamanho amostral diminuiu em relação ao nível de significância anterior, variando entre 30 e 56 para tamanho de efeito de 80%, entre 74 e 138 para tamanho de efeito de 50% e entre 452 e 858 para tamanho de efeito de 20%.

Considerando que os recursos financeiros e de tempo em estudos ambientais quase sempre são bastante limitados, foi possível observar que em qualquer das situações simuladas seria inviável detectar um tamanho de efeito de 0,20, uma vez que em todos os casos avaliados seriam necessárias mais de 200 amostras em cada grupo para alcançar o resultado almejado em um teste t.

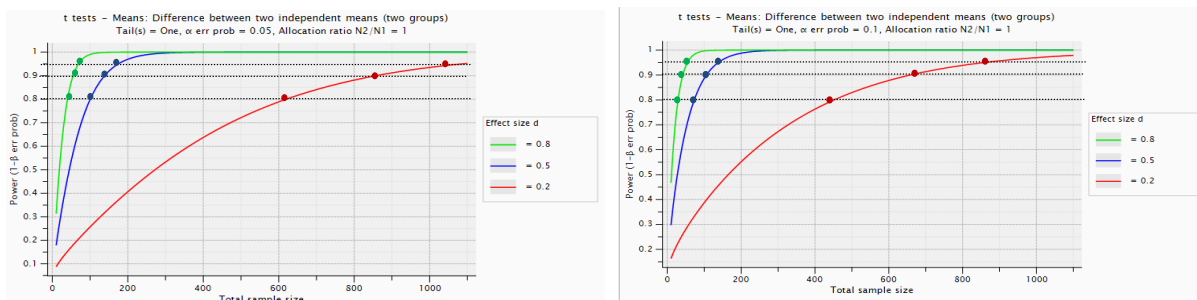


Figura 6: Gráficos poder x tamanho amostral para 3 tamanhos de efeito diferentes, com $\alpha = 0,05$ (esquerda) e $\alpha = 0,10$ (direita) (situação hipotética 1).

Foi feita também uma análise de comparação do poder dos testes paramétricos e seus correspondentes não paramétricos (Figura 7). Assim, gerou-se uma tabela do poder em função do tamanho amostral para um teste t e seu correspondente Wilcoxon Mann-Whitney. Pode-se observar que o teste paramétrico foi levemente mais poderoso que o não paramétrico, atingindo maior poder com um menor número de amostras, mas a diferença entre eles é pequena, em torno de 1 a 2%.

Graph		Table		
t tests - Means: Difference between two independent means (two groups) Tail(s) = One, α err prob = 0.05, Allocation ratio N2/N1 = 1				
#	Total sample size	Effect size d = 0.2 Power (1-β err prob)	Effect size d = 0.5 Power (1-β err prob)	Effect size d = 0.8 Power (1-β err prob)
1	50.0000	0.171645	0.539002	0.873552
2	150.0000	0.335161	0.919687	0.999384
3	250.0000	0.472881	0.989196	0.999998
4	350.0000	0.587974	0.998749	1.000000
5	450.0000	0.681987	0.999869	1.000000
6	550.0000	0.757242	0.999987	1.000000
7	650.0000	0.816470	0.999999	1.000000
8	750.0000	0.862425	1.000000	1.000000
9	850.0000	0.897654	1.000000	1.000000
10	950.0000	0.924378	1.000000	1.000000
11	1050.0000	0.944467	1.000000	1.000000
12	1150.0000	0.959448	1.000000	1.000000

Graph		Table		
t tests - Means: Wilcoxon-Mann-Whitney test (two groups) Tail(s) = One, Parent distribution = Normal, α err prob = 0.05, Allocation ratio N2/N1 = 1				
#	Total sample size	Effect size d = 0.2 Power (1-β err prob)	Effect size d = 0.5 Power (1-β err prob)	Effect size d = 0.8 Power (1-β err prob)
1	50.0000	0.167510	0.522725	0.859450
2	150.0000	0.325001	0.908707	0.999094
3	250.0000	0.458516	0.986337	0.999997
4	350.0000	0.571269	0.998229	1.000000
5	450.0000	0.664520	0.999792	1.000000
6	550.0000	0.740190	0.999977	1.000000
7	650.0000	0.800614	0.999998	1.000000
8	750.0000	0.848212	1.000000	1.000000
9	850.0000	0.885276	1.000000	1.000000
10	950.0000	0.913849	1.000000	1.000000
11	1050.0000	0.935683	1.000000	1.000000
12	1150.0000	0.952239	1.000000	1.000000

Figura 7: Quadros de comparação do poder do teste t (a esquerda) e do teste Wilcoxon Mann-Whitney (a direita) para N, tamanho de efeito e α equivalentes.

Supõe-se que os dados aleatórios gerados (médias e desvios padrão), apresentados na Tabela 3, foram obtidos com base na literatura ou em um estudo piloto e que os dados apresentam distribuição normal. No exemplo 1 o tamanho de efeito “observado” foi de 0,268. Seriam necessárias 121 amostras em cada rio para obter um poder de 95% e 74 amostras por rio para atingir um poder de 80%. No exemplo 2, o tamanho de efeito subiu para 0,598 e, conseqüentemente, o tamanho amostral necessário caiu para 25/rio para poder de 95% e 15/rio para poder de 80%. No exemplo 3 o tamanho de efeito cresceu mais ainda, sendo necessárias apenas 7 amostras por rio para obter um poder de 95% e 5 para um poder de 80%.

Assim como explicado anteriormente, os tamanhos amostrais apresentados na tabela abaixo foram obtidos calculado pelo GPower a partir da combinação dos outros três parâmetros, sendo necessária uma aproximação para se atingir um N compatível com o número de subgrupos utilizados na análise. As análises foram rodadas para testes t bicaudais.

Tabela 3: Dados simulados, tamanhos de efeito calculados e N amostrais obtidos para situação hipotética 1.

AMOSTRA	DIVERSIDADE	DIVERSIDADE MÉDIA	DESVIO PADRÃO	TAMANHO DE EFEITO	N PARA PODER DE 95%	N PARA PODER DE 80%
EXEMPLO 1						
C1	0,29	1,05	0,82	0,268	363 por grupo (121 por rio)	220 por grupo (74 por rio)
C2	1,91					
C3	0,94					
I1	0,00	1,34	1,29			
I2	2,58					
I3	1,43					
EXEMPLO 2						
C1	1,21	1,67	0,49	0,598	74 por grupo (25 por rio)	45 por grupo (15 por rio)
C2	2,19					
C3	1,60					
I1	1,04	1,41	0,37			
I2	1,40					
I3	1,79					
EXEMPLO 3						
C1	2,01	2,62	0,53	1,155	21 por grupo (7 por rio)	13 por grupo (5 por rio)
C2	2,89					
C3	2,97					
I1	1,70	2,03	0,49			
I2	1,80					
I3	2,60					

Situação hipotética 2:

Considerando o impacto gerado pelo lançamento de despejos orgânicos por emissários submarinos sobre as comunidades bentônicas, o alcance das plumas geradas por esses lançamentos deve variar também. Para avaliar como essa variação impacta nas comunidades bentônicas das regiões adjacentes será feito um monitoramento (hipotético) das regiões localizadas nos pontos cardeais num raio de 500m da desembocadura de um emissário. Será feita uma análise de poder com o intuito de detectar quantas amostras deverão ser tomadas em cada ponto (norte, sul,

leste, oeste) para se detectar diferentes tamanhos de efeito e com diferentes valores de poder.

Como o estudo avaliará 4 variáveis categóricas, será realizada uma Análise de Variância (ANOVA). Para avaliar o poder de uma ANOVA foram realizadas simulações a partir de 4 grupos de amostras, considerando que as amostras cumprem as premissas de homogeneidade das variâncias e que os dados são independentes entre si. Foram avaliados tamanhos de efeito pequeno, médio e grande que, segundo a tabela de Cohen (1977) apud GPower Tutorial apresentada anteriormente, correspondem a 10, 25 e 40% para ANOVA. Foram feitas simulações com poder de 80, 90 e 95% e nível de significância de 0,05 e 0,10, com o intuito de verificar de que forma variou o tamanho amostral. Assim como feito anteriormente, foi gerada uma tabela (Tabela 4) e gráficos correspondes (Figura 8) com os tamanhos amostrais necessários para cada uma das combinações dos parâmetros avaliadas.

Tabela 4: Tamanhos amostrais obtidos pela combinação dos outros parâmetros da análise de poder (situação hipotética 2).

ANOVA				
ALFA	TAMANHO DE EFEITO	PODER	N AMOSTRAL POR GRUPO	N AMOSTRAL TOTAL
0,05	0,1	0,8	274	1096
0,05	0,1	0,9	356	1424
0,05	0,1	0,95	431	1724
0,05	0,25	0,8	45	180
0,05	0,25	0,9	58	232
0,05	0,25	0,95	70	280
0,05	0,4	0,8	18	72
0,05	0,4	0,9	24	96
0,05	0,4	0,95	28	112
0,1	0,1	0,8	221	884
0,1	0,1	0,9	296	1184
0,1	0,1	0,95	366	1464
0,1	0,25	0,8	36	144
0,1	0,25	0,9	48	192
0,1	0,25	0,95	60	240
0,1	0,4	0,8	15	60
0,1	0,4	0,9	20	80
0,1	0,4	0,95	24	96

Os gráficos ilustram como varia o poder com o aumento do tamanho amostral para os diferentes tamanhos de efeitos e níveis de significância selecionados (Figura 8). As linhas pretas horizontais representam os níveis de poder de 80 e 95%. Pode-se observar como os níveis de poder tidos como adequados são muito mais rapidamente atingidos com tamanhos de efeito médio a grande do que quando o efeito observado é pequeno.

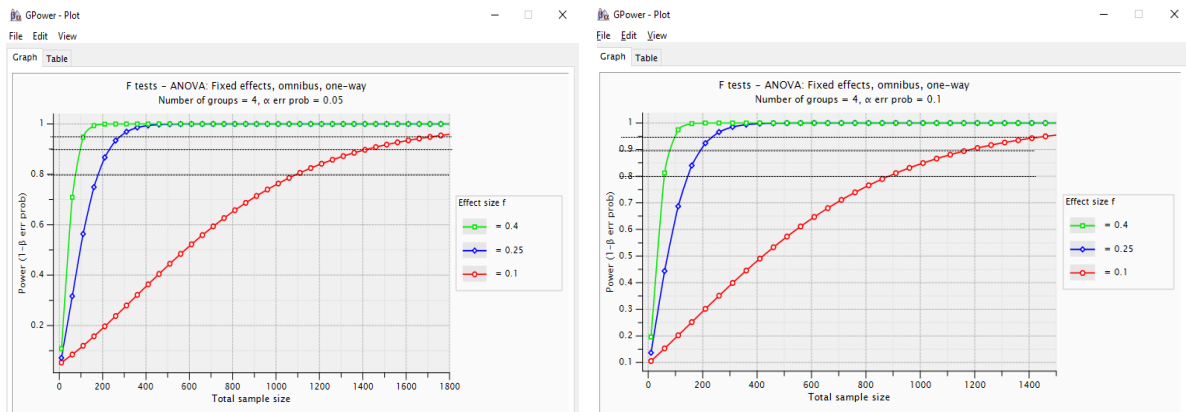


Figura 8: Gráficos poder x tamanho amostral para 3 tamanhos de efeito diferentes, com $\alpha = 0,05$ (esquerda) e $\alpha = 0,10$ (direita) (situação hipotética 2).

No caso da ANOVA, o tamanho amostral obtido é o número total de amostras, ou seja, a soma de todos os grupos definidos no delineamento. Como a situação hipotética aqui analisada refere-se aos quatro pontos cardeais, o número de grupos é 4, então o N necessário em cada estação é o N total/4. Pode-se observar que para se detectar um efeito de 10% com qualquer dos níveis de α pré-definidos é necessário um N superior a 800, o que significa a coleta de mais de 200 amostras por grupo. Na prática, esse é um valor bastante elevado para estudos ambientais, o que significa que muito dificilmente um teste de ANOVA irá detectar um efeito desta pequena magnitude.

Um tamanho de efeito intermediário, em torno de 25% de variação, requer um total amostral entre 144 e 280 amostras, ou seja, entre 36 e 70 amostras por grupo. Já para um tamanho de efeito grande, em torno de 40%, entre 60 e 112 amostras são suficientes (entre 15 e 28 amostras por grupo) variando de acordo com o poder desejado. Esses dois valores já são mais coerentes dentro do âmbito ambiental, podendo ser obtidos a um custo menos inalcançável.

Uma outra forma de exibir os resultados no GPower é através de quadros como os que se seguem (Figura 9). Neles foram avaliados os tamanhos amostrais resultantes de diferentes combinações de poder e tamanho de efeito com um nível de significância pré-definido, assim como foi feito na situação hipotética 1. Esses quadros são gerados automaticamente junto com os gráficos, situando-se na aba “Table”.

Nesses quadros pode-se ter uma visão geral de como o tamanho amostral aumenta consideravelmente com o aumento do poder ou diminuição do α ou do tamanho de efeito. Foram acrescentadas combinações com nível de significância de 0,01 às combinações anteriormente definidas de 0,05 e 0,10. De modo geral, uma diminuição do α de 0,10 para 0,05 (de 90 para 95% de confiança) representou um aumento do N de 17 a 24%, enquanto que uma redução do α de 0,05 para 0,01 significa um aumento correspondente entre 32 e 41% no tamanho amostral. Por exemplo, para o menor tamanho amostral, resultante da combinação de um poder de 0,8 com tamanho de efeito de 0,40, passou de 58 com $\alpha = 0,10$ para 72 com $\alpha = 0,05$ e 102 com $\alpha = 0,01$.

The image shows three tables from the GPower software interface, each representing a different alpha level. Each table has columns for Power (1-β err prob), Total sample size, and three columns for Effect size f (0.1, 0.25, 0.4). The number of groups is fixed at 4.

#	Power (1-β err prob)	Effect size f = 0.1	Effect size f = 0.25	Effect size f = 0.4
1	0.800000	882.909	143.943	58.2469
2	0.850000	1010.77	164.397	66.2260
3	0.900000	1182.77	191.911	76.9627
4	0.950000	1460.40	236.326	94.2995

#	Power (1-β err prob)	Effect size f = 0.1	Effect size f = 0.25	Effect size f = 0.4
1	0.800000	1094.17	178.397	72.1705
2	0.850000	1234.03	200.770	80.9027
3	0.900000	1421.06	230.692	92.5831
4	0.950000	1720.90	278.662	111.312

#	Power (1-β err prob)	Effect size f = 0.1	Effect size f = 0.25	Effect size f = 0.4
1	0.800000	1551.44	253.033	102.376
2	0.850000	1714.40	279.104	112.558
3	0.900000	1930.42	313.666	126.056
4	0.950000	2273.10	368.494	147.469

Figura 9: Tamanhos amostrais em função do poder para diferentes níveis de significância.

É desejável que α e β sejam mantidos nos menores valores possível, porém, como eles são inversamente correlacionados, é difícil minimizar os dois ao mesmo tempo, motivo pelo qual Hanson (2011) defende que eles sejam ajustados considerando a razão entre o risco dos erros tipo I e tipo II, de acordo com a situação e com o nível de aceitação de cada um dos riscos. Para Mudge et al (2012), a definição de um “ α ótimo” (equilibrado entre os custos dos erros tipo I e tipo II) tem o potencial de

reduzir a frequência de erros estatísticos no monitoramento ambiental sem aumentar o número de amostras ou alterar o tamanho de efeito crítico escolhido.

Esse equilíbrio é de fundamental importância num estudo ambiental, para que os estudos não priorizem apenas o risco para o empreendedor ao custo de um risco ambiental mais elevado, ao mesmo tempo que haja um delineamento viável de ser realizado, uma vez que, como foi visto nas simulações aqui demonstradas, para se atingir um poder adequado com um valor de alfa tradicionalmente definido, é necessário uma quantidade de amostras muito grande, muitas vezes inviável em termos de custo e disponibilidade de tempo.

Situação hipotética 3: Análise de poder post hoc

Apesar da análise de poder ser mais útil no planejamento de um estudo, ela também exerce um papel importante após a realização de uma amostragem e subsequente análise dos dados, quando o resultado obtido não é significativo. Tradicionalmente tende-se a negligenciar este tipo de resultado, assumindo-se que não há impacto na situação analisada. Contudo, um resultado não significativo pode ser fruto de um estudo com poder insuficiente para detectar um impacto e, desta forma, importantes impactos sobre o meio ambiente podem acabar sendo desconsiderados.

Como exemplo da utilização deste tipo de análise, denominada de *post hoc*, criou-se com a situação hipotética 3, a partir da situação hipotética 2, para avaliação dos impactos do despejo de esgoto por um emissário submarino sobre a comunidade bentônica da região, porém considerando que a análise já tivesse sido realizada sem um planejamento prévio do poder da análise. A análise foi feita a partir da amostragem da comunidade bentônica em 5 pontos a norte, sul, leste e oeste num raio de 500m da desembocadura do emissário.

Considerando que no local exista uma corrente marítima predominantemente de sentido norte/sul, a pluma de efluentes tende a afetar principalmente as estações do sul e, em menor grau, as do leste. As estações do norte e oeste não sofrem impacto da pluma do emissário. Contudo, a análise dos dados concluiu que não houve impacto algum, uma vez que o p foi superior a 0,07 (valor hipoteticamente estabelecido) e o nível de significância escolhido foi de 95% (*i.e.* $\alpha_{\text{crítico}} = 0,05$).

Foi então aplicada uma análise de poder para verificar se de fato a conclusão do estudo está correta ou se pode ter ocorrido um erro tipo II, quando a existência de um impacto não é percebida por H_0 ter sido considerada verdadeira erroneamente. Foram gerados no Microsoft Excel números aleatórios considerando 5 amostras por ponto, de modo a se obter 4 valores médios com um mesmo desvio padrão para representar os 4 fatores da análise (N, S, L, O). Os valores obtidos são apresentados na Tabela 5.

Tabela 5: Valores aleatórios obtidos como representação das amostras coletadas na situação hipotética 3 para os 4 pontos amostrais (norte, sul, leste, oeste), com respectivas médias e desvios-padrão.

Réplica	N	S	L	O
R1	7	2	3	6
R2	13	8	9	12
R3	14	9	10	13
R4	10	5	6	9
R5	12	7	8	11
Média	11,2	6,2	7,2	10,2
Desvio Padrão	2,86	2,86	2,86	2,86

A análise de poder *post hoc* é bastante simples de ser realizada no GPower 3, uma vez determinado o tamanho de efeito observado na análise, através da quantidade e tamanho dos grupos, médias e desvio padrão. Dessa forma, alimentando o programa para uma ANOVA 1-fator com o número de amostras utilizados (20), quantidade de grupos (4), o nível de significância utilizado (0,05) e o tamanho de efeito observado, obtêm-se o poder da análise e conseqüentemente, a probabilidade do erro tipo II.

A Figura 10 ilustra como foi calculado o tamanho de efeito observado e o poder do teste. Na aba da direita, aberta através do link “Determine”, foi calculado o tamanho de efeito observado a partir dos valores médios, desvio padrão e tamanho dos grupos obtidos no Excel. Com o tamanho de efeito resultante (0,72), retorna-se à aba da esquerda, alimentando o programa com o valor obtido para se calcular o poder do teste que, no caso, foi de 0,66.

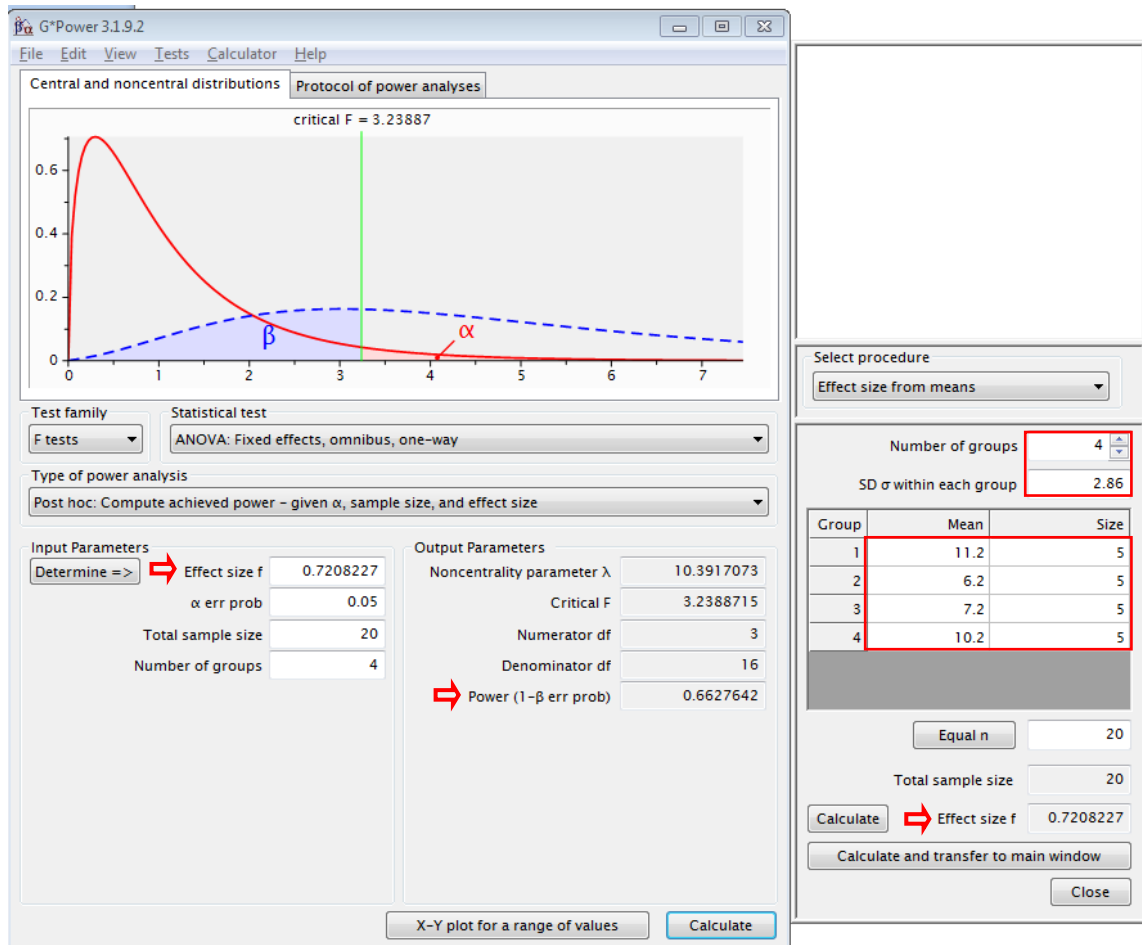


Figura 10: Calculando o tamanho de efeito observado (aba da direita) e o poder do teste (aba da esquerda) no GPower 3.1.

Simulando outros tamanhos de efeito a partir da quantidade de amostras e nível de significância pré-estabelecido, percebeu-se que, para qualquer tamanho de efeito inferior a 0,84 o poder obtido seria menor que 80%. Ou seja, a probabilidade de cometer um erro tipo II (β) e concluir que não houve impacto quando de fato as comunidades ao sul e sudeste do emissário submarino estavam sendo impactadas é superior a 20%, mais de quatro vezes maior do que a probabilidade de cometer um erro tipo I. Ou seja, no exemplo criado a probabilidade de se cometer um erro tipo II é de aproximadamente 33%, valor bastante acima do que o convencionalmente tido como aceitável (20%) ou equilibrado com o α (5%).

6. CONSIDERAÇÕES FINAIS

Através do protocolo aqui elaborado foi demonstrado que a utilização do software GPower permite a realização de análises simplificadas do poder de testes estatísticos devido à facilidade de sua utilização, facilitando a incorporação de análises de poder em estudos ambientais.

Verificou-se a importância da realização de estudos piloto, de forma a estimar previamente algumas informações necessárias ao planejamento adequado de um monitoramento ambiental. Tais informações são importantes para se estimar o tamanho de efeito esperado de se encontrar em cada situação analisada, necessário para o cálculo do tamanho amostral adequado.

Considerando a correlação das variáveis envolvidas em uma análise de poder pôde-se observar que, para ser mais conservativo em relação ao meio ambiente reduzindo ao máximo os impactos ambientais, é necessário adotar práticas de inclusão da análise de poder nos estudos, prévios à sua execução, de modo a estimar o tamanho amostral mais adequado para se minimizar o risco de erros tipo II. Para isso, verificou-se que o nível de significância convencionalmente utilizado em estudos ambientais ($\alpha = 0,05$) pode muitas vezes comprometer os resultados, uma vez que a variável é diretamente proporcional ao poder, ou seja, quanto menor o nível α selecionado, menor será o poder do teste ou, ainda, que para se obter um baixo nível de significância e um alto poder equivalente, seria necessário um tamanho amostral bastante elevado, o que geralmente compromete a viabilidade de um monitoramento ambiental.

A simulação da análise *post hoc* permitiu concluir que, apesar de ter 95% de confiança, isso não se reflete em segurança para o meio ambiente, uma vez que a probabilidade de não cometer um erro tipo II pode ser alta mesmo com um nível de significância reduzido.

Conclui-se com isso sobre a importância de se equilibrar os riscos dos erros entre o ambiente e o empreendedor, definindo previamente o tamanho amostral a ser utilizado com base no tamanho de efeito esperado, no nível de significância e no poder desejável.

7. REFERÊNCIAS BIBLIOGRÁFICAS

BLOMBERG, S. P. (2014). Power Analysis Using R, 13.

DI STEFANO, J. (2003). Functional Ecology\rHow much power is enough? Against the development of an arbitrary convention for statistical power calculations. *Functional Ecology*, 17, 707–709.

FAUL, F., ERDFELDER, E., LANG, A., & BUCHNER, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.

GPOWER TUTORIAL. Disponível em <<http://www.statpower.net/Content/312/Handout/gpower-tutorial.pdf>>. Acessado em 08 de setembro de 2016.

HANSON, N. (2011). Using biological data from field studies with multiple reference sites as a basis for environmental management: The risks for false positives and false negatives. *Journal of Environmental Management*, 92(3), 610–619. <http://doi.org/10.1016/j.jenvman.2010.09.026>

HINDS, W. T. (1984). Towards Monitoring of Long-term Trends in Terrestrial Ecosystems. *Environmental Conservation*, 11(01), 11–18. <http://doi.org/doi:10.1017/S0376892900013448>

MUDGE, J. F., BARRETT, T. J., MUNKITTRICK, K. R., & HOULAHAN, J. E. (2012). Negative consequences of using $\alpha = 0.05$ for Environmental monitoring decisions: A case study from a decade of Canada's environmental effects monitoring program. *Environmental Science and Technology*, 46(17), 9249–9255. <http://doi.org/10.1021/es301320n>

MUELLER, C. C. (1992). Situação atual da produção de informações sistemáticas sobre o meio ambiente. *Revista Ciência Da Informação*, 21, 14–22.

SOFTWARE LIVRE, código aberto e software gratuito: as diferenças. Disponível em <<http://www.infowester.com/freexopen.php>>. Acessado em 12 de setembro 2016.

THOMAS, L., & KREBS, C. (1997). A review of statistical power analysis software.

Bulletin of the Ecological Society of America, 78(2), 128–139. Retrieved from <http://scholar.google.com/scholar?q=intitle:A+REVIEW+OF+STATISTICAL+POWER+ANALYSIS+SOFTWARE#0>