Ocean Engineering 139 (2017) 265–274

Contents lists available at ScienceDirect

Ocean Engineering

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

Statistical analysis of vertical accelerations of planing craft: Common pitfalls and how to avoid them

Douglas VanDerwerken, Carolyn Judge*

United States Naval Academy, Annapolis, United States

ARTICLE INFO

Keywords:
High-speed craft
Statistics
Uncertainty propagation
Exponential distribution
Planing hulls

ABSTRACT

A careful review of the literature on the analysis of vertical accelerations of high-speed watercraft reveals that a number of statistical misconceptions are pervasive. These include: (1) ignoring vertical thresholding in analysis, (2) misunderstanding the relationships between the Exponential and other distributions, (3) failing to report uncertainty, and (4) failing to discuss alternative methods for estimation. This article brings these issues to the fore and shows how popular methods, when applied to simulated data, can give inaccurate results. We set forth proper methods for analysis and illustrate their use on a data set from a full-scale planing craft.

1. Introduction

This paper deals with the issues surrounding the statistical methods used to analyze acceleration peaks for planing hulls operating in irregular seas. The acceleration response for planing craft is nonlinear with respect to the wave height, unlike for displacement craft where a Rayleigh distribution can be used to analyze acceleration peaks in irregular waves. There is great interest in determining the best probability distribution for describing these vertical acceleration peak values. In fact, Koelbel (1995) identified vertical acceleration as the "single most pressing problem" facing planing hull designers when determining structural design criteria. Vertical acceleration is a critical parameter when determining structural design standards, classification requirements, habitability, and personnel readiness and safety. If the probability distribution were known, the designer could determine the average of the $1/N^{th}$ highest peaks (or any other feature of the distribution) for input into the structural design procedure. Identifying a probability density function for the vertical accelerations allows for "simple and direct transitions between criteria" (Schleicher, 2008). This paper delves into the methods researchers have been using when analyzing the statistical characterization of high-speed planing craft vertical acceleration data in irregular seas. The process for identifying these acceleration peaks involves decisions on a host of complicated problems including data acquisition, filtering, and peak identification. This paper does not delve into the issues surrounding those decisions, but instead considers the purely statistical questions relating to the set of presumably random draws from the distribution of vertical acceleration peaks. A careful review of the available literature has revealed a number of statistical misconceptions. We identify four common pitfalls in the statistical analysis of vertical acceleration data and present improved methods. The impacts of the statistical misconceptions are demonstrated using simulated data and statistically consistent methods presented are applied to a data set from a full-scale planing craft.

The existence of these misconceptions does not necessarily invalidate the conclusions in the literature. In most cases, the focus was on uncovering the physics of wave slamming and the statistical analysis played only a relatively minor role. Moreover, the history of improved hull design validates much of the findings that resulted from these simple, or even flawed, statistical techniques. For instance, McCue (2012) noted that the historical reliance on the Exponential distribution, even when it may have been suboptimal, has built in a certain conservatism in design that has served the naval architecture community well and should not be abandoned lightly. However, our review of the literature reveals that there are some statistical concepts that appear to be incorrectly applied. Other researchers should be made aware of these problems and the potential solutions.

The first widely recognized analysis on vertical accelerations for high-speed planing craft was done by Fridsma (1969), Fridsma (1971). He performed a systematic investigation of planing hull motions and vertical accelerations using a series of prismatic planing boat models with varying deadrise angles and length-beam ratios, tested with various loadings, trims, speed-length ratios, and sea states. He showed that the craft vertical acceleration data are random and highly nonlinear in relation to wave height. Analyzing all positive peaks, he found that the acceleration peak data tended to follow an Exponential distribution. The results from Fridsma's investigations were incorporated into estimates for expected vertical accelerations (Savitsky and

* Corresponding author.
E-mail address: judge@usna.edu (C. Judge).

http://dx.doi.org/10.1016/j.oceaneng.2017.05.004
Received 17 October 2016; Received in revised form 29 March 2017; Accepted 6 May 2017
0029-8018/ Published by Elsevier Ltd.

Brown, 1976; Allen and Jones, 1978; Hoggard and Jones, 1980), particularly through calculating the average of the highest $1/N^{th}$ (or peak tail average) vertical acceleration peaks under the Exponential assumption. These methods of vertical acceleration prediction were eventually incorporated into the structural design rules of the classification societies (ABS, 2016; DNV, 2013).

As data became more easily acquired and computer resources made it possible to automate analysis, researchers began investigating automated methods of peak identification and looking to recreate Fridsma's results. Different peak identification methods affected reported results and, even within the same peak identification method, statistical results (primarily the average of the highest $1/N^{th}$ result) were found to be sensitive to user-defined vertical threshold levels (Zseleczky and McKee, 1989; Allen et al., 2008; Soletic, 2010; Grimsley, 2010; Grimsley et al., 2010; Riley et al., 2010; Zseleczky, 2012; Somayajula and Falzarano, 2014; Razola et al., 2016). As more vertical acceleration data for high-speed planing craft operating in rough seas became available, the Exponential distribution was found to be a poor fit and, in many instances, other probability distributions were found to fit these data better (Brown and Klosinski, 1980; Zarnick and Turner, 1981; Blount et al., 2006; Schleicher, 2008; Soletic, 2010; Grimsley, 2010; Grimsley et al., 2010; Taunton et al., 2011; McCue, 2012; Bowles and Soja, 2014; Begovic et al., 2016; Razola et al., 2016). Several of these efforts are examined in more depth to provide examples of the current approaches to vertical acceleration peak statistical analysis. Grimsley (2010) and Grimsley et al. (2010) examined existing multiple full-scale and model-scale acceleration data sets to determine if the Exponential distribution was the best distribution. They compared the fits for four distributions: Exponential, Rayleigh, Lognormal and Gumbel. The authors also investigated the effect that peak identification method has on the results of the distribution fitting. They found the Exponential distribution was the most sensitive to the vertical threshold values. Using a minimum distance criterion, they determined the Exponential distribution to be the worst fit to the data. Soletic (2010) tested scale models of a systematic series of planing hull geometries in rough water to study their seakeeping characteristics. He found that the acceleration data did not match the Exponential distribution when using the peak identification method chosen. The tail averages based on the Exponential distribution were significantly lower than those measured directly from the experimental data. Taunton et al. (2011) also conducted an experimental investigation of high-speed hard-chine planing hulls in irregular seas and found that the Gamma distribution fit the acceleration data better than the Exponential distribution. McCue (2012) created an automated algorithm for categorizing wave slam events according to impact types as defined in Riley et al. (2010). The author fit various distributions to the combined data and to the data separated by slam type. McCue used minimum distance to compare the fits of different probability distributions (Exponential, Rayleigh, Lognormal, Gumbel, Weibull, Fréchet, and Generalized Extreme Value or GEV). She found the Exponential distribution to be a poor fit for vertical acceleration peak data for data from a full-scale high-speed planing boat operating in rough seas, generally finding the GEV to do the best job of capturing the tail behaviors with the Weibull distribution being the second best fit. Somayajula and Falzarano (2014) considered similar data to McCue (2012) and investigated the fit of the Generalized Pareto distribution (which includes the Exponential distribution as a special case) using a peaks-over-threshold analysis for motions of a planing craft in random seas. The Generalized Pareto distribution was fit using maximum likelihood. The authors' focus was on predicting the tail or extreme values. Begovic et al. (2016) pursued an experimental program for providing high-speed seakeeping data for realistic service and weather conditions. The authors evaluated several probability distributions for wave, heave, and pitch motions, and considered best-fit distributions for vertical accelerations (Exponential, Gamma, and Weibull). In terms of a likelihood estimation factor, the Exponential distribution was found to be the worst fit, while the Gamma and Weibull distributions appeared to be better choices. Razola et al. (2016) presented a simulated set of impact acceleration data for a high-speed craft operating in irregular seas. By simulating the responses, rather than conducting physical experiments, many more impacts could be evaluated. The authors considered issues involved in determining the statistical characteristics of the impact acceleration process including the slamming time scales, selection of appropriate sampling rates and filtering levels, identification of peak acceleration impact values, statistical distributions and convergence, and the relationship between the peak tail averages. Generalized Pareto, Gumbel, Fréchet, and Weibull distributions were considered as fits to the vertical acceleration peak data. The authors evaluated the goodness of fit using graphical methods (quantile-quantile plots) and a minimum distance criterion. They found that the selected vertical threshold had a strong influence on the resulting peak tail averages, that variation of the horizontal threshold time window had very little influence on the peak tail averages and extreme values, and that the Exponential distribution was not a good fit to the acceleration peak data. The authors conducted an evaluation of the uncertainty in the extreme values through bootstrapping.

Unfortunately, some of the results of these and other wave slam analyses rest on shaky statistical ground. In our review of the literature, we have found four common pitfalls in the statistical analysis of vertical acceleration data that we feel need to be addressed. These include:

1. Vertical thresholding is not accounted for in analysis (and hence tail averages are not well-defined).
2. Implications of Exponential distribution being a special case of other distributions are ignored or misunderstood.
3. Uncertainty is ignored.
4. Alternative methods for estimation are not discussed.

Issues related to data collection procedures and expected distribution features are not addressed in this paper. Most of the statistical analysis conducted in wave slam analyses has focused on attempts to identify the best probability distributions for high-speed craft operating in irregular seas, which would be expected to depend on hull geometry, model scale, instrumentation, and data treatment (Savitsky, 2016). These issues are not relevant to the concerns listed above, as the present paper is focused on the method of analysis for the given peak acceleration values. This paper will proceed as follows. Section 2 gives a thorough account of each of the above pitfalls, demonstrates how they lead to inaccurate inference, and gives improved methods for these situations. Section 3 shows how these methods can be applied using full-scale planing craft acceleration data. We close with a candid discussion of the results.

2. Common Pitfalls

2.1. Vertical thresholding is not accounted for in analysis (and hence tail averages are not well-defined)

In Fridsma (1971), the vertical threshold was zero and data for analysis "include[d] all positive peak accelerations, including those that are wave-induced as well as impact spikes...." Fridsma fit the Exponential distribution to these data, which is entirely appropriate because the Exponential probability density function (pdf) is nonzero on $[0, \infty)$. Since that time, several researchers have found that other vertical acceleration peak data sets are poorly fit by the Exponential distribution, suggesting the Exponential distribution not be used universally. Zseleczky and McKee (1989), Grimsley (2010), and others have raised concerns about the sensitivity of tail averages to the choice of vertical threshold (i.e. the value below which peaks are not chosen). What seems to have gone unnoticed is that this "sensitivity" results not from the choice of threshold per se but from an inconsistent definition