ELSEVIER

Journal of Wind Engineering
and Industrial Aerodynamics 95 (2007) 165–182

JOURNAL OF wind engineering AND industrial aerodynamics

www.elsevier.com/locate/jweia

The *r* largest order statistics model for extreme wind speed estimation

Ying An[1], M.D. Pandey*

Department of Civil Engineering, University of Waterloo, Waterloo, Ont., Canada N2L 3G1

Received 14 November 2005; received in revised form 6 April 2006; accepted 22 May 2006
Available online 28 July 2006

Abstract

The paper presents the statistical estimation of extreme wind speed using annually *r* largest order statistics (*r*-LOS) extracted from the time series of wind data. The method is based on a joint generalized extreme value distribution of *r*-LOS derived from the theory of Poisson process. The parameter estimation is based on the method of maximum likelihood. The hourly wind speed data collected at 30 stations in Ontario, Canada, are analyzed in the paper. The results of *r*-LOS method are compared with those obtained from the method of independent storms (MIS) and specifications of the Canadian National Building Code (CNBC-1995). The CNBC estimates are apparently conservative upper bound due to large sampling error associated with annual maxima analysis. Using the *r*-LOS method, the paper shows that the wind pressure data can be suitably modelled by the Gumbel distribution.

Keywords: Wind speed; Extreme value estimation; Generalized extreme value distribution; Order statistics; Annual maxima; Maximum likelihood method; Method of independent storm

1. Introduction

The estimation of design wind speed corresponding to a long return period is generally based on the extreme value theory, which derives the three asymptotic domains of attraction, namely, the Gumbel, Frechet and Weibull distributions [1]. These three distributions can be written in a unified form, referred to as the Generalized Extreme Value distribution (GEV).

*Corresponding author. Tel.: +1 519 888 4567x5858; fax: +1 519 888 4349.
E-mail address: mdpandey@uwaterloo.ca (M.D. Pandey).
[1]Graduate student.

0167-6105/$ - see front matter
doi:10.1016/j.jweia.2006.05.008

Traditionally, a sample of annual maximum wind speed is fitted with the Gumbel distribution using the methods of moments or least squares. However, the statistical extrapolation to estimate wind speed corresponding to 500–1000 year return period is seriously contaminated by sampling and model uncertainty, if data are available for a limited period (20–30 years). This has motivated the development of approaches to enlarge the sample extreme values beyond the annual maxima.

The method of independent storms (MIS), proposed by Cook [2] and refined by Harris [3,4], considers several wind storm maxima, rather than just annual maxima. The extremes of storm maxima are fitted with the Gumbel distribution. The MIS method is limited to the Gumbel distribution, and it discounts the possibility of GEV model representing the data. The Peaks-Over-Threshold (POT) method is another alternative that models the peaks of wind speed time series exceeding a threshold by the Generalized Pareto Distribution (GPD), which is shown to be the domain of attraction of the peaks [5,6]. However, the application of POT is confounded by an erratic variation of a quantile estimate with respect to the threshold used in creating the sample of peaks [7].

The paper presents an alternate extreme value analysis of the Canadian wind speed data that is based on estimation of the joint distribution of annually r largest order statistics (r-LOS) of data. Assuming that r-LOS are generated by an underlying inhomogeneous Poisson process, they can be modelled by a joint GEV distribution [8,9]. The paper shows that the r-LOS method provides a systematic approach to (1) ascertain whether data belong to the Gumbel or the GEV distribution, and (2) estimate the sampling error associated with quantile estimates. Although the theoretical basis of the r-LOS method is well established, the paper illustrates its versatility in the estimation of extreme wind speed.

The paper is organized as follows. A brief review of extreme value theory and MIS is presented in Section 2. The proposed r-LOS method is described in Section 3. Section 4 presents a detailed analysis of wind data collected at 30 sites in Ontario (Canada) using r-LOS and MIS methods. The wind speed quantile estimates are compared with the design values specified in the Canadian National Building Code (CNBC) 1995. Section 5 summarizes the finding of this paper.

2. Extreme value estimation methods

2.1. Background

2.1.1. General concept

Consider a sample of *iid* random variables, $(X_1, X_2, \ldots, X_n)$, with a cumulative distribution (CDF) $F_X(x)$ and denote the maximum value in the sample as $M_n = \max(X_1, \ldots X_n)$. The CDF of M_n can be obtained as

$$Pr\{M_n < x\} = [F_X(x)]^n. \tag{1}$$

According to extreme value theory, if there exist some constants a_n and b_n, then the distribution of extremes converges to a non-trivial result as

$$Pr\{(M_n - b_n)/a_n \leqslant x\} = [F_X(a_n x + b_n)]^n \to G_X(x) \quad \text{as } n \to \infty. \tag{2}$$

The asymptotic distribution, $G_X(x)$, must converge to one of the three types of distributions, namely, the Gumbel, Frechet and Weibull forms [1]. The three distributions